IN THE UNITED STATES DISTRICT COURT

FOR THE DISTRICT OF DELAWARE

ZYMOGENETICS, INC., a Washington Corporation, Plaintiff, v. BRISTOL-MYERS SQUIBB CO., a Delaware Corporation, and DOES 1 through 100, Defendants.	)	C.A. No. 06-500-SLR **PUBLIC VERSION**

DECLARATION OF HANNAH LEE IN SUPPORT OF PLAINTIFF ZYMOGENETICS, INC.'S MOTION FOR SUMMARY JUDGMENT OF VALIDITY OF PATENTS-IN-SUIT WITH RESPECT TO THE CAPON PATENTS AND SEED PATENT VOLUME 2– EXHIBITS 11-16

Philip A. Rovner (#3215)
POTTER ANDERSON & CORROON LLP
Hercules Plaza
P. O. Box 951
Wilmington, DE 19899
(302) 984-6000
provner@potteranderson.com

Attorneys for Plaintiff
ZymoGenetics, Inc.

OF COUNSEL:

Paul J. Andre
Lisa Kobialka
KING & SPALDING LLP
1000 Bridge Parkway
Suite 1000
Redwood City, CA 94065
(650) 590-0700

Dated: May 16, 2008
Public Version: May 22, 2008

I, HANNAH LEE, declare:

1. I am an attorney with the law firm of King & Spalding LLP, counsel of record for Plaintiff ZymoGenetics, Inc. I have personal knowledge of the facts set forth in this declaration and can testify competently to those facts.

2. Attached hereto as Exhibit 1 is a true and correct copy of pages 25-26, 74, 77-78, 81-83, 109, 178, 183, and 186-187 from the transcript of the deposition of Andrzej Sledziewski ("Dr. Sledziewski"), taken on August 21, 2007.

3. Attached hereto as Exhibit 2 is a true and correct copy of pages from the laboratory notebook of Dr. Sledziewski, bearing bates numbers ZG033202, 33206, 33217, 33220, and 58359.

4. Attached hereto as Exhibit 3 is a true and correct copy of pages 178-179 from the transcript of the deposition of Paul E. Purdue, taken on May 5, 2008.

5. Attached hereto as Exhibit 4 is a true and correct copy of pages from the laboratory notebook of Lillian Anne Bell, bearing bates numbers ZG040368, 40464, 40510-12, 40518-25, and 40527.

6. Attached hereto as Exhibit 5 is a true and correct copy of the Expert Report of Jeffrey V. Ravetch, M.D., Ph.D. Regarding Invalidity of U.S. Patent Nos. 5,843,725 and 6,018,026.

7. Attached hereto as Exhibit 6 is a true and correct copy of the Expert Report of Daniel J. Capon, Ph.D. Regarding Invalidity of U.S. Patent Nos. 5,843,725 and 6,018,026.

8. Attached hereto as Exhibit 7 is a true and correct copy of pages 15 and 56-62 from the transcript of the deposition of Wayne Kindsvogel, taken on August 8, 2007.

9. Attached hereto as Exhibit 8 is a true and copy of pages from Oligonucleotide Request form or primer order sheet, bearing bates numbers ZG047284-85.

10. Attached hereto as Exhibit 9 is a true and correct copy of a timeline regarding relevant dates of priority.

11. Attached hereto as Exhibit 10 is a true and correct copy of U.S. Patent No. 5,843,725.

12. Attached hereto as Exhibit 11 is a true and correct copy of U.S. Patent No. 6,018,026.

13. Attached hereto as Exhibit 12 is a true and correct copy of U.S. Patent No. 6,004,781.

14. Attached hereto as Exhibit 13 is a true and correct copy of U.S. Patent No. 5,336,603.

15. Attached hereto as Exhibit 14 is a true and correct copy of U.S. Patent No. 5,565,335.

16. Attached hereto as Exhibit 15 is a true and correct copy of U.S. Patent No. 6,117,655.

17. Attached hereto as Exhibit 16 is a true and correct copy of U.S. Patent No. 6,710,169.

I declare under penalty of perjury under the laws of the State of California and the United States that each of the above statements is true and correct.

Executed on May 16, 2008 in Redwood Shores, California.

Hannah Lee

IN THE UNITED STATES DISTRICT COURT
FOR THE DISTRICT OF DELAWARE

CERTIFICATE OF SERVICE

I, Philip A. Rovner, hereby certify that on May 22, 2008, the within document was filed with the Clerk of the Court using CM/ECF which will send notification of such filing(s) to the following; that the document was served on the following counsel as indicated; and that the document is available for viewing and downloading from CM/ECF.

BY HAND DELIVERY AND E-MAIL

Frederick L. Cottrell, III, Esq.
Kelly E. Farnan, Esq.
Richard Layton & Finger, P.A.
One Rodney Square
P. O. Box 551
Wilmington, DE 19899
cottrell@rlf.com
farnan@rlf.com

I hereby certify that on May 22, 2008 I have sent by E-mail the foregoing document to the following non-registered participant:

Christopher P. Borello, Esq.
Colleen Tracy, Esq.
Fitzpatrick Cella Harper & Scinto
30 Rockefeller Plaza
New York, NY 10112
cborello@fchs.com
ctracy@fchs.com

/s/ Philip A. Rovner
Philip A. Rovner (#3215)
Potter Anderson & Corroon LLP
Hercules Plaza
P. O. Box 951
Wilmington, DE 19899
(302) 984-6000
provner@potteranderson.com

748348

EXHIBIT 11
Part 1

US006018026A

United States Patent [19]

Siedziewski et al.

[11] Patent Number: 6,018,026

[45] Date of Patent: *Jan. 25, 2000

[54] **BIOLOGICALLY ACTIVE DIMERIZED AND MULTIMERIZED POLYPEPTIDE FUSIONS**

[75] Inventors: **Andrzej Z. Siedziewski; Lillian Anne Bell; Wayne R. Kindsvogel,** all of Seattle, Wash.

[73] Assignee: **ZymoGenetics, Inc.,** Seattle, Wash.

[*] Notice: This patent is subject to a terminal disclaimer.

[21] Appl. No.: **08/980,400**

[22] Filed: **Nov. 26, 1997**

Related U.S. Application Data

[62] Division of application No. 08/477,329, Jun. 7, 1995, Pat. No. 5,750,375, which is a continuation of application No. 08/180,195, Jan. 11, 1994, Pat. No. 5,567,584, which is a continuation of application No. 07/634,510, Dec. 27, 1990, abandoned, which is a continuation-in-part of application No. 07/347,291, May 2, 1989, Pat. No. 5,155,027, which is a continuation-in-part of application No. 07/146,877, Jan. 22, 1988, abandoned.

[51] **Int. Cl.**[7] **C07K 19/00**; C07K 16/46; C07K 14/71

[52] **U.S. Cl.** **530/350**; 530/387.3; 530/388.24; 530/399; 530/402

[58] **Field of Search** 530/350, 387.3, 530/388.24, 399, 402; 435/69.7, 172.3

[56] **References Cited**

U.S. PATENT DOCUMENTS

4,474,893	10/1984	Reading	436/547
4,599,311	7/1986	Kawasaki	435/71
4,761,371	8/1988	Bell et al.	435/68
4,766,073	8/1988	Murray et al.	435/172.3
4,816,397	3/1989	Boss et al.	435/68
4,816,567	3/1989	Cabilly et al.	530/387
5,015,571	5/1991	Niman et al.	435/7.92
5,098,833	3/1992	Lasky et al.	435/69.1
5,116,964	5/1992	Capon et al.	536/27
5,155,027	10/1992	Siedziewski et al.	435/69.7
5,750,375	5/1998	Siedziewski et al.	435/69.7

FOREIGN PATENT DOCUMENTS

088 994	9/1983	European Pat. Off. .
116 201	8/1984	European Pat. Off. .
120 694	10/1984	European Pat. Off. .
162 319	11/1985	European Pat. Off. .
171 496	2/1986	European Pat. Off. .
173 494	3/1986	European Pat. Off. .
239 400	9/1987	European Pat. Off. .
244 221A1	11/1987	European Pat. Off. .
256 654	7/1988	European Pat. Off. .
315 062	5/1989	European Pat. Off. .
327 369A2	8/1989	European Pat. Off. .
WO 86/01533	3/1986	WIPO .
WO 9008953	6/1990	WIPO .

OTHER PUBLICATIONS

Kawana et al., *Biochemical and Biophysical Research Communications 149* (3) : 960-968, 1987

Yarden et al., *Nature* 323:226-232, 1986.

Hart et al., *Science 240*:1529-1531, 1988.

Escobedo et al., *Science 240*:1532–1534, 1988.

Heldin et al., *Embo J. 7*:1387–1393, 1988.

Radeke et al., *Nature 325*593–597, 1987.

Ebina et al., *Cell 40*:747–758, 1985.

Rubin et al., *J. Immun. 135*:3172–3177, 1985.

Ellis et al., *J. Cell Biol. 150*:14a, 1987.

Smith et al., *Science 238*:1704–1707, 1987.

Livnch et al., *J. Biol. Chem. 261*: 12490–12497, 1986.

Boni-Schnetzler and Pilch, *Proc. Natl. Acad. Sci. USA 84*:7832–7836, 1987.

Claesson-Welsh et al., *Mol. Cell. Biol.* 8:3476–3486, 1988.

Ullrich et al., *Nature 309*:418–425, 1984.

Ullrich et al., *Nature 313*: 756–761, 1985.

Derynck et al., *Nature 316*:701–705, 1985.

Loh et al., *Cell 33*:85–93, 1983.

Honjo et al., *Cell 18*:559–568, 1979.

Alber and Kawasaki, *J. Mol. Appl. Genet. 1*: 419–434, 1982.

Russell et al., *Nature 304*: 652–654, 1983.

Grosschedl and Baltimore, *Cell 41*: 885–897, 1985.

Hieter et al., *Cell* 22: 197–207, 1980.

Treiger et al., *J. Immunol. 136*:4099–4105, 1986.

Roth et al., *J. Cell. Biol. 102*:1271–1283, 1986.

Riedel et al., *Science 236*:197–200, 1987.

Riedel et al., *Nature 324*:68–70, 1986.

Gascoigne et al., *Proc. Natl. Acad. Sci. USA 84*:2936–2940, 1987.

Maddon et al., *Cell 47*: 333–348, 1986.

Maddon et al., *Cell 42*: 93–104, 1985.

Capon et al., *Nature 337*:525–531, 1989.

Wood et al., *Nature 314*:446–449, 1985.

Carlson et al., *Mol. Cell. Biol. 3*:439–447, 1983.

Marx, *Science 229*:455–456, 1985.

Neuberger et al., *Nature 312*: 604–609, 1984.

Oi and Morrison *Biotechniques 4*:214–221, 1986.

Ballon et al., *Bio/Technology 5*:1195–1198, 1987.

Ballou et al., *J. Biol. Chem. 255*:5986–5991, 1980.

(List continued on next page.)

Primary Examiner—Lorraine Spector
Assistant Examiner—Claire M. Kaufman
Attorney, Agent, or Firm—Townsend and Townsend and Crew

[57] **ABSTRACT**

Methods for producing secreted receptor analogs and biologically active peptide dimers are disclosed. The methods for producing secreted receptor analogs and biologically active peptide dimers utilize a DNA sequence encoding a receptor analog or a peptide requiring dimerization for biological activity joined to a dimerizing protein. The receptor analog includes a ligand-binding domain. Polypeptides comprising essentially the extracellular domain of a human PDGF receptor fused to dimerizing proteins, the portion being capable of binding human PDGF or an isoform thereof, are also disclosed. The polypeptides may be used within methods for determining the presence of and for purifying human PDGF or isoforms thereof.

29 Claims, 21 Drawing Sheets

OTHER PUBLICATIONS

Munro and Pelham, *EMBO J.* 3:3087–3093, 1984.
von Heijne, *Nucl Acids Res. 14*: 4683–4690, 1986.
von Heijne, *Eur. J. Biochem. 133*:17–21, 1983.
Beggs, *Nature*: 104–108, 1978.
Clarke et al., *Meth. Enzym. 68*:436–442, 1979.
Struhl et al., *Proc. Natl. Acad. Sci. USA 76*:1035–1039, 1979.
Broach et al., *Gene 8*:121–133, 1979.
Mariuzza and Winter, *J..Biol. Chem. 264*:7310–7316, 1989.
Gronwald et al., *Proc. Natl. Acad. Sci. USA 85*:3435–3439, 1988.
Riedel et al., *Protein Engineering 1*(3):237, 1987.
Traunecker et al., Highly efficient neutralization of HIV with recombinant CD4–immunoglobulin molecules, Nature, 339: 68–70, May 4, 1989.

```
   1  GGCCCCTCAGCCCTGCTGCCCAGCACGAGCCTGTGCTCGCCCTGCCCAACGCAGACAGCCAGACCCAGG   69
  70  GCGGCCCCCTCTGGCGGCTCTGCTCCTCCCGAAGGATGCTTGGGGAGTGAGGCGAAGCTGGGCGCTCCTC  138
 139  TCCCCTACAGCAGCCCCCTTCCTCCATCCCTCTGTTCTCCTGAGCCTTCAGGAGCCTGCACCAGTCCTG  207
 208  CCTGTCCTTCTACTCAGCTGTTACCCACTCTGGGACCAGCAGTCTTTCTGATAACTGGGAGAGGGCAGT  276
 277  AAGGAGGACTTCCTGGAGGGGGTGACTGTCCAGAGCCTGGAACTGTGCCCACACCAGAAGCCATCAGCA  345

 346  GCAAGGACACCATGCGGCTTCCGGGTGCGATGCCAGCTCTGGCCCTCAAAGGCGAGCTGCTGTTGCTGT  414
                M  R  L  P  G  A  M  P  A  L  A  L  K  G  E  L  L  L  L  S    20

 415  CTCTCCTGTTACTTCTGGAACCACAGATCTCTCAGGGCCTGGTCGTCACACCCCCGGGGCCAGAGCTTG  483
        L  L  L  L  E  P  Q  I  S  Q  G  L  V  V  T  P  P  G  P  E  L  V    43

 484  TCCTCAATGTCTCCAGCACCTTCGTTCTGACCTGCTCGGGTTCAGCTCCGGTGGTGTGGGAACGGATGT  552
        L  N  V  S  S  T  F  V  L  T  C  S  G  S  A  P  V  V  W  E  R  M  S    66

 553  CCCAGGAGCCCCCACAGGAAATGGCCAAGGCCCAGGATGGCACCTTCTCCAGCGTGCTCACACTGACCA  621
        Q  E  P  P  Q  E  M  A  K  A  Q  D  G  T  F  S  S  V  L  T  L  T  N    89

 622  ACCTCACTGGGCTAGACACGGGAGAATACTTTTGCACCCACAATGACTCCCGTGGACTGGAGACCGATG  690
        L  T  G  L  D  T  G  E  Y  F  C  T  H  N  D  S  R  G  L  E  T  D  E   112

 691  AGCGGAAACGGCTCTACATCTTTGTGCCAGATCCCACCGTGGGCTTCCTCCCTAATGATGCCGAGGAAC  759
        R  K  R  L  Y  I  F  V  P  D  P  T  V  G  F  L  P  N  D  A  E  E  L   135

 760  TATTCATCTTTCTCACGGAAATAACTGAGATCACCATTCCATGCCGAGTAACAGACCCACAGCTGGTGG  828
        F  I  F  L  T  E  I  T  E  I  T  I  P  C  R  V  T  D  P  Q  L  V  V   158
```

Fig. 1A

```
829   TGACACTGCACGAGAAGAAAGGGGACGTTGCACTGCCTGTCCCCTATGATCACCAACGTGGCTTTTCTG  897
        T  L  H  E  K  K  G  D  V  A  L  P  V  P  Y  D  H  Q  R  G  F  S  G  181

898   GTATCTTTGAGGACAGAAGCTACATCTGCAAAACCACCATTGGGACAGGAGGTGGATTCTGATGCCT   966
        I  F  E  D  R  S  Y  I  C  K  T  T  I  G  D  R  E  V  D  S  D  A  Y  204

967   ACTATGTCTACAGACTCCAGGTGTCATCCATCAACGTCTCTGTGAACGCAGTGCAGACTGTGGTCCGCC 1035
        Y  V  Y  R  L  Q  V  S  S  I  N  V  S  V  N  A  V  Q  T  V  V  R  Q  227

1036  AGGGTGAGAACATCACCCTCATGTGCATTGTGATCGGGAATGAGGTGGTCAACTTCGAGTGGACATACC 1104
        G  E  N  I  T  L  M  C  I  V  I  G  N  E  V  V  N  F  E  W  T  Y  P  250

1105  CCCGCAAAGAAAGTGGGCGGCTGGTGGAGCCGGTGACTGACTTCCTCTTGGATATGCCTTACCACATCC 1173
        R  K  E  S  G  R  L  V  E  P  V  T  D  F  L  L  D  H  P  Y  M  I  R  273

1174  GCTCCATCCTGCACATCCCCAGTGCCGAGTTAGAAGACTCGGGGACCTACACCTGCAATGTGACGGAGA 1242
        S  I  L  H  I  P  S  A  E  L  E  D  S  G  T  Y  T  C  N  V  T  E  S  296

1243  GTGTGAATGACCATCAGGATGAAAAGGCCATCAACATCACCGTGGTTGAGAGCGGCTACGTGCGGCTCC 1311
        V  N  D  H  Q  D  E  K  A  I  N  I  T  V  V  E  S  G  Y  V  R  L  L  319

1312  TGGGAGAGGTGGGCACACTACAATTGCTGAGCTGCATCGGAGCCGGACACTGCAGGTAGTGTTCGAGG 1380
        G  E  V  G  T  L  Q  F  A  E  L  H  R  S  R  T  L  Q  V  V  F  E  A  342

1381  CCTACCCACCGCCCACTGTCCTGTGGTTCAAAGACAACCGCACCCTGGGCGACTCCAGCGCTGGCGAAA 1449
        Y  P  P  P  T  V  L  W  F  K  D  N  R  T  L  G  D  S  S  A  G  E  I  365

1450  TCGCCCTGTCCACGCGCAACGTGTCGGAGACCCGGTATGTGTCAGAGCTGACACTGGTTCGCGTGAAGG 1518
        A  L  S  T  R  N  V  S  E  T  R  Y  Y  V  S  E  L  T  L  V  R  V  K  V  388
```

Fig. 1B

```
1519  TGGCAGAGGCTGGCCACTACACCATGCGGGCCTTCCATGAGGATGCTGAGGTCCAGCTCTCCTTCCAGC 1587
        A  E  A  G  H  Y  Y  T  M  R  R  A  F  H  E  D  A  E  V  Q  L  S  F  Q  L  411

1588  TACAGATCAATGTCCCTGTCCGAGTGCTGGAGCTAAGTGAGAGCCACCCTGACAGTGGGGAACAGACAG 1656
        Q  I  N  V  P  V  R  V  L  E  L  S  E  S  H  P  D  S  G  E  Q  T  V  434

1657  TCCGCTGTCGTGGCCGGGGCATGCCCCAGCCGAACATCATCTGGTCTGCCTGCAGAGACCTCAAAAGGT 1725
        R  C  R  G  G  M  P  Q  P  N  I  I  W  S  A  C  R  D  L  K  R  C  457

1726  GTCCACGTGAGCTGCCGCCCCACGCTGCTGGGGAACAGTTCCGAAGAGGAGAGCCAGCTGGAGACTAACG 1794
        P  R  E  L  P  P  T  L  L  G  N  S  S  E  E  E  S  Q  L  E  T  N  V  480

1795  TGACGTACTGGGAGGAGGAGCAGGAGTTTGAGGTGGTGAGCACACTGCGTCTGCAGCACGTGGATCGGC 1863
        T  Y  W  E  E  E  Q  E  F  E  V  V  S  T  L  R  L  Q  H  V  D  R  P  503

1864  CACTGTCGGTGCTGCACGCTGCGCAACGCTGTGGGCCAGGACACGCAGGAGGTCATCGTGGTGCCAC 1932
        L  S  V  R  C  T  L  R  N  A  V  G  Q  D  T  Q  E  V  I  V  V  P  H  526

1933  ACTCCTTGCCCTTTAAGGTGGTGGATCTCAGCCATCCTGGCCCTGGTGGCTCACCATCATCTCCC 2001
        S  L  P  F  K  V  V  V  I  S  A  I  L  A  L  V  V  L  T  I  I  S  L  549

2002  TTATCATCCTCATCATGCTTTGGCAGAAGAAGCCACGTTACGAGATCCGATGGAAGGTGATTGAGTCTG 2070
        I  I  L  I  M  L  W  Q  K  K  P  R  Y  E  I  R  W  K  V  I  E  S  V  572

2071  TGAGCTCTGACGGCCATGAGTACATCTACGTGGACCCCCATGCAGCTGCCCTATGACTCCACGTGGGAGC 2139
        S  S  D  G  H  E  Y  I  Y  V  D  P  M  Q  L  P  Y  D  S  T  W  E  L  595

2140  TGCCGCGGGACCAGCTTGTGCTGGGACGCACCCTCGGCTCTGGGGCCTTTGGGCAGGTGGTGGAGGCCA 2208
        P  R  D  Q  L  V  L  G  R  T  L  G  S  G  A  F  G  Q  V  V  E  A  T  618
```

Fig. 1C

```
2209  CGGCTCATGGCCTGAGCCATTCTCAGGCCACCGATGAAAGTGGCCGTCAAGATGCTTAAATCCACAGCCC 2277
        A  H  G  L  S  H  S  Q  A  T  M  K  V  A  V  K  M  L  K  S  T  A  R  641

2278  GCAGCAGTGAGAAGCAAGCCCTTATGTCGGAGCTGAAGATCATGAGTCACCTTGGGCCCCACCTGAACG 2346
        S  S  E  K  Q  A  L  M  S  E  L  K  I  M  S  H  L  G  P  H  L  N  V  664

2347  TGGTCAACCTGTTGGGGGCCTGCACCAAGGAGGACCCATCTATATCATCACTGAGTACTGCCGCTACG 2415
        V  N  L  L  G  A  C  T  K  G  G  P  I  Y  I  I  T  E  Y  C  R  Y  G  687

2416  GAGACCTGGTGGACTACCTGCACCGCAACAAACACACCTTCCTGCAGCACCACTCCGACAAGCGCCGCC 2484
        D  L  V  D  Y  L  H  R  N  K  H  T  F  L  Q  H  H  S  D  K  R  R  P  710

2485  CGCCCAGCGCGGAGCTCTACAGCAATGCTCTGCCCGTTGGGCTCCCCCTGCCCAGCCATGTGTCCTTGA 2553
        P  S  A  E  L  Y  S  N  A  L  P  V  G  L  P  L  P  S  H  V  S  L  T  733

2554  CCGGGGAGAGCGACGGTGGCTACATGGACATGAGCAAGGACGAGTCGGTGGACTATGTGCCCATGCTGG 2622
        G  E  S  D  G  G  Y  M  D  M  S  K  D  E  S  V  D  Y  V  P  M  L  D  756

2623  ACATGAAAGGAGACGTCAAATATGCAGACATCGAGTCCTCCAACTACATGGCCCCTTACGATAACTACG 2691
        M  K  G  D  V  K  Y  A  D  I  E  S  S  N  Y  M  A  P  Y  D  N  Y  V  779

2692  TTCCCTCTGCCCCTGAGAGGACCTGCCGAGCAACTTTGATCAACGAGTCTCCAGTGCTAAGCTACATGG 2760
        P  S  A  P  E  R  T  C  R  A  T  L  I  N  E  S  P  V  L  S  Y  M  D  802

2761  ACCTCGTGGGCTTCAGCTACCAGGTGGCCAATGGCATGGAGTTTCTGGCCTCCAAGAACTGCGTCCACA 2829
        L  V  G  F  S  Y  Q  V  A  N  G  M  E  F  L  A  S  K  N  C  V  H  R  825

2830  GAGACCTGGCGGCTAGGAACGTGCTCATCTGTGAAGGCAAGCTGGTCAAGATCTGTGACTTTGGCCTGG 2898
        D  L  A  A  R  N  V  L  I  C  E  G  K  L  V  K  I  C  D  F  G  L  A  848
```

Fig. 1D

```
2899  CTCGAGACATCATGCGGGACTCGAATTACATCTCCAAAGGCAGCACCTTTTGCCTTTAAAGTGGATGG 2967
        R  D  I  M  R  D  S  N  Y  I  S  K  G  S  T  F  L  P  L  K  W  M  A  871

2968  CTCCGGAGAGCATCTTCAACAGCCTCTACACCACCCTGAGCGACGTGTGGTCCTTCGGGATCCTGCTCT 3036
        P  E  S  I  F  N  S  L  Y  T  T  L  S  D  V  W  S  F  G  I  L  L  W  894

3037  GGGAGATCTTCACCTTGGGTGGCACCCCTTACCCAGAGCTGCCCATGAACGAGCAGTTCTACAATGCCA 3105
        E  I  F  T  L  G  G  T  P  Y  P  E  L  P  M  N  E  Q  F  Y  N  A  I  917

3106  TCAAACGGGGTTACCGCATGGCCCAGCCTGCCCATGCCTCCGACGAGATCTATGAGATCATGCAGAAGT 3174
        K  R  G  Y  R  M  A  Q  P  A  H  A  S  D  E  I  Y  E  I  H  Q  K  C  940

3175  GCTGGGAAGAGAAGTTTGAGATTCGGCCCCCCTTCTCCCAGCTGGTGCTGCTTCTCGAGAGACTGTTGG 3243
        W  E  E  K  F  E  I  R  P  P  F  S  Q  L  V  L  L  L  E  R  L  L  G  963

3244  GCGAAGGTTACAAAAAGAAGTACCAGCAGGTGGATGAGGAGTTTCTGAGGAGTGACCACCCAGCCATCC 3312
        E  G  Y  K  K  K  K  Y  Q  Q  V  D  E  E  F  L  R  S  D  H  P  A  I  L  986

3313  TTCGGTCCCAGGCCCGCTTGCCTGGGTTCCATGGCCTCCGATCTCCCCTGGACACCAGCTCCGTCCTCT 3381
        R  S  Q  A  R  L  P  G  F  H  G  L  R  S  P  L  D  T  S  S  V  L  Y  1009

3382  ATACTGCCGTGCAGCCCAATGAGGGTGACAACGACTATATCATCCCCCTGCCTGACCCCAAACCCGAGG 3450
        T  A  V  Q  P  N  E  G  D  N  D  Y  I  I  P  L  P  D  P  K  P  E  V  1032

3451  TTGCTGACGAGGGCCCACTGGAGGGTTCCCCCAGCCTAGCCAGCTCCACCCTGAATGAAGTCAACACCT 3519
        A  D  E  G  P  L  E  G  S  P  S  L  A  S  S  T  L  N  E  V  N  T  S  1055

3520  CCTCAACCATCTCCTGTGACAGCCCCCTGGAGCCCCAGGACGAACCAGAGCCCCAGCTTGAGC 3588
        S  T  I  S  C  D  S  P  L  E  P  Q  D  E  P  E  P  Q  L  E  L  1078
```

Fig. 1E

```
3589  TCCAGGTGGAGCCGGAGCCAGAGCTGGAACAGTTGCCGGATTCGGGGTGCCCTGCGCCTCGGGCGGAAG 3657
        Q  V  E  P  E  P  E  L  E  Q  L  P  D  S  G  C  P  A  P  R  A  E  A 1101

3658  CAGAGGATAGCTTCCTGTAGGGGGCTGGCCCCCTACCCTGCCCTGCCTGAAGCTCCCCCCCTGCCAGCAC 3726
          E  D  S  F  L                                                    1106

3727  CCAGCATCTCCTGGCCTGGCCTGACCGGGCTTCCTGTCAGCCAGGCTGCCCTTATCAGCTGTCCCCTTC 3795

3796  TGGAAGCTTTCTGCTCCTGACGTGTTGTGCCCCCAAACCCTGGGGCTGGCTTAGGAGGCAAGAAAACTGC 3864

3865  AGGGGCCGTGACCAGCCCTCTGCCTCCAGGGAGGCCAACTGACTCTGAGCCAGGGTTCCCCCAGGGAAC 3933

3934  TCAGTTTTCCCATATGTAAGATGGGAAAGTTAGGCTTGATGACCCAGAATCTAGGATTCTCTCCCTGGC 4002

4003  TGACACGGTGGGGAGACCGAATCCCTCCCTGGGAAGATTCTTGGAGTTACTGAGGTGGTAAATTAACAT 4071

4072  TTTTTCTGTTCAGCCAGCTACCCCTCAAGGAATCATAGCTCTCTCCTCGCACTTTTTATCCACCCAGGA 4140

4141  GCTAGGGAAGAGACCCTAGCCTCCCTGGCTGCTGGCTGAGCTAGGGCCTAGCTTGAGCAGTGTTGCCTC 4209

4210  ATCCAGAAGAAAGCCAGTCTCCTCCCTATGATGCCAGTCCCTGCGTTCCCTGGCCCGAGCTGGTCTGGG 4278

4279  GCCATTAGGCAGCCCTAATTAATGCTGGAGGCTGAGCCAAGTACAGGACACCCCCAGCCTGCAGCCCTTG 4347

4348  CCCAGGGCACTTGGAGCACACGCAGCCATAGCAAGTGCCTGTGTCCCTGTCCTTCAGGCCCATCAGTCC 4416

4417  TGGGGCTTTTTCTTTATCACCCTCAGTCTTAATCCATCCACCAGAGTCTAGAAGGCCAGACGGGCCCCG 4485

4486  CATCTGTGATGAGAATGTAAATGTGCCAGTGTGGAGTGGCCACGTGTGTGTGCCAGTATATGGCCCTGG 4554
```

Fig. 1F

```
4555  CTCTGCATTGGACCTGCTATGAGGCTTTGGAGGAATCCCTCACCCTCTCTGGCCTCAGTTTCCCCTTC  4623

4624  AAAAAATGAATAAGTCGGACTTATTAACTCTGAGTGCCTTGCCAGCACTAACATTCTAGAGTATTCCAG  4692

4693  GTGGTGCACATTTGTCCAGATGAAGCAAGGCCTATACCCTAAACTTCATCCTGGGGGTCAGCTGGGCTC  4761

4762  CTGGGAGATTCCAGATCACACATCACTCTGGGGACTCAGGAACCATGCCCCTTCCCCAGGCCCCCAG  4830

4831  CAAGTCTCAAGAACACAGCTGCACAGGCCTTGACTTAGAGTGACAGCCGGTGTCCTGGAAAGCCCCAAG  4899

4900  CAGCTGCCCCAGGGACATGGGAAGACCACGGGACCTCTTTCACTACCCACGATGACCTCCGGGGGTATC  4968

4969  CTGGGCAAAAGGGACAAAGAGGGCAAATGAGATCACCTCCTGCAGCCCACCACTCCAGCACCTGTGCCG  5037

5038  AGGTCTGCGTCGAAGACAGAATGGACAGTGAGGACAGTTATGTCTTGTAAAAGACAAGAAGCTTCAGAT  5106

5107  GGTACCCCAAGAAGGATGTGAGAGGTGGCTGCTTTGGAGTTTGCCCCCTCACCCCACCAGCTGCCCCAT  5175

5176  CCCTGAGGCATGCGCTCCATGGGGGTATGGTTTTGTCACTGCCCAGACCTAGCAGTGACATCTCATTGT  5244

5245  CCCCAGCCCCAGTGGGCATTGGAGGTGCCAGGGGAGTCAGGGTTGTAGCCAAGACGCCCCCGCACGGGGA  5313

5314  GGGTTGGGAAGGGGGGTGCAGGAAGCTCAACCCCCTCTGGGCACCAACCCTGCATTGCAGGTTGGCACCTT  5382

5383  ACTTCCCTGGGATCCCCCAGAGTTGGTCCAAGGAGGGAGAGTGGGTTCTCAATACGGTACCAAAGATATA  5451

5452  ATCACCTAGGTTTACAAATATTTTTTAGGACTCACGTTAACTCACATTTATACAGCAGAAATGCTATTTT  5520

5521  GTATGCTGTTAAGTTTTTCTATCTGTGTACTTTTTTTTTTAAGGGAAAGATTTT                  5572
```

Fig. 1G

Figure 2

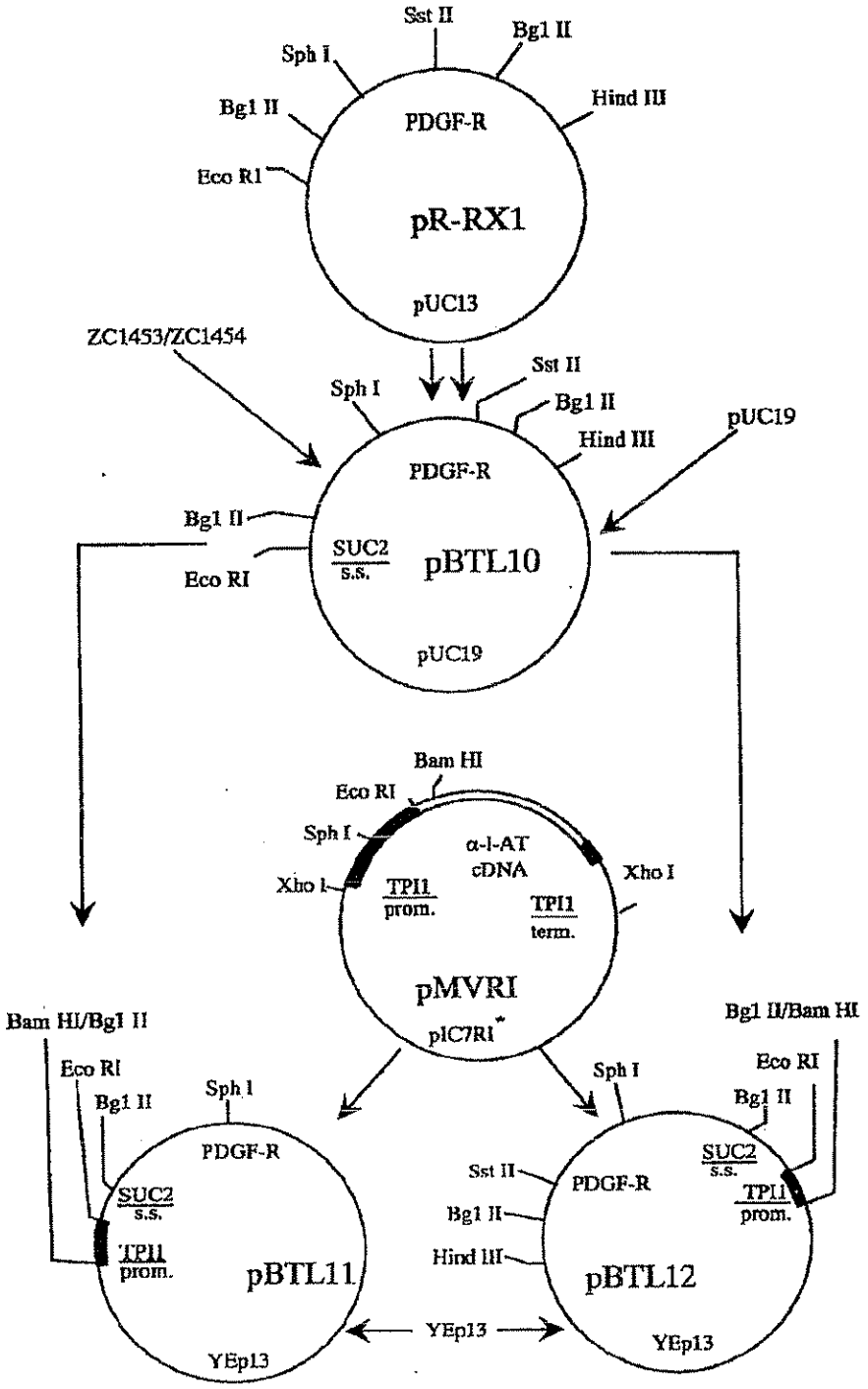

Sst II
Bg1 II
Sph I
Hind III
Bg1 II
PDGF-R
Eco R1
pR-RX1
pUC13
ZC1453/ZC1454
Sst II
Sph I
Bg1 II
pUC19
Hind III
PDGF-R
Bg1 II
SUC2
s.s.
pBTL10
Eco RI
pUC19
Bam HI
Eco RI
Sph I
α-l-AT
cDNA
Xho I
Xho I
TPI1
prom.
TPI1
term.
pMVRI
pIC7RI
Bam HI/Bg1 II
Bg1 II/Bam HI
Eco RI
Sph I
Sph I
Eco RI
Bg1 II
Bg1 II
PDGF-R
SUC2
s.s.
Sst II
PDGF-R
SUC2
s.s.
Bg1 II
TPI1
prom.
Hind III
TPI1
prom.
pBTL11
pBTL12
YEp13
YEp13
YEp13

Figure 3

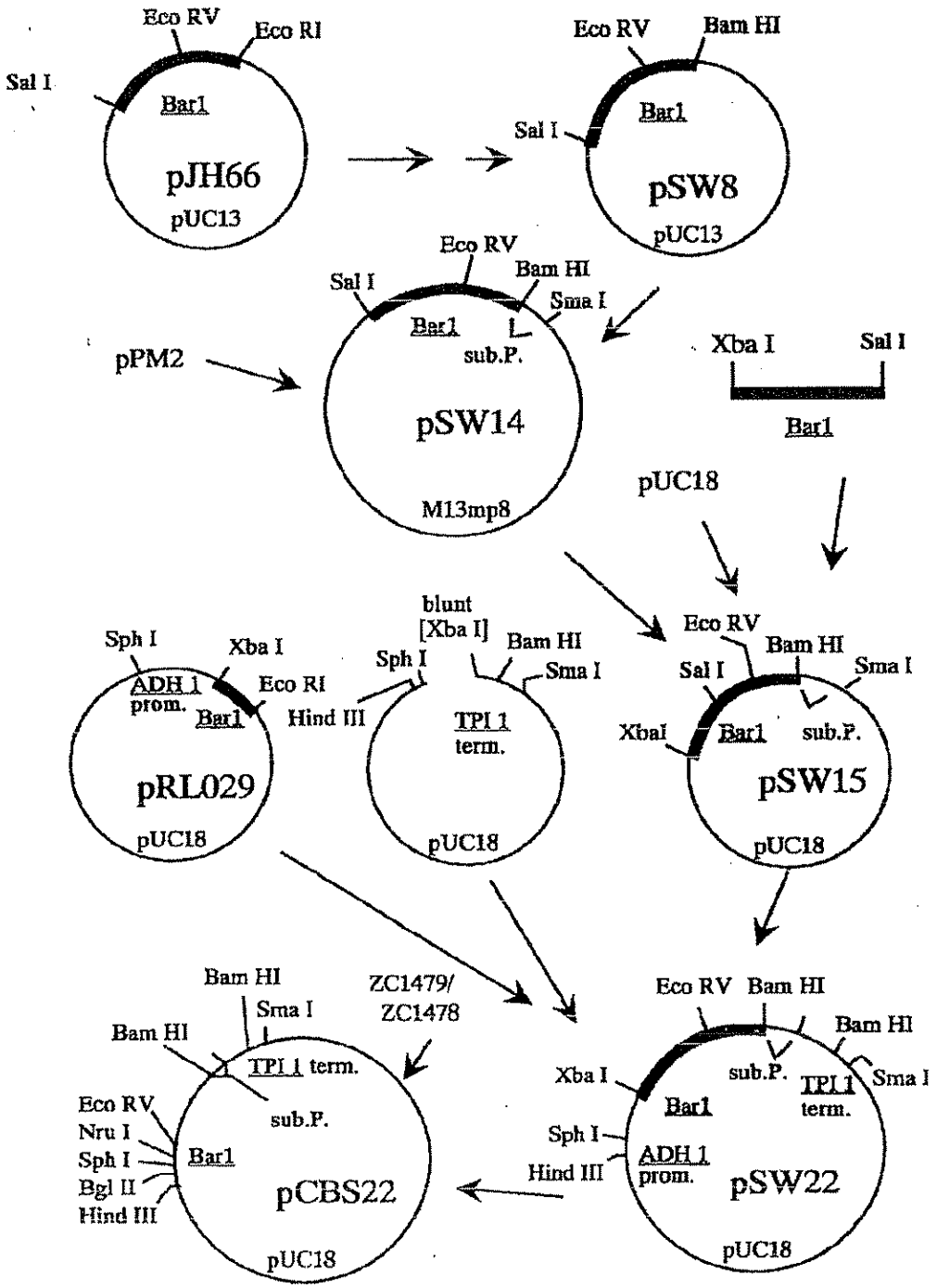

Eco RV
Eco RI
Sal I
Bar1
pJH66
pUC13
Eco RV
Bam HI
Sal I
Bar1
pSW8
pUC13
Eco RV
Bam HI
Sal I
Sma I
Bar1
sub.P.
pPM2
pSW14
M13mp8
Xba I
Sal I
Bar1
pUC18
Sph I
Xba I
ADH 1
prom.
Eco RI
Bar1
pRL029
pUC18
blunt
[Xba I]
Bam HI
Sph I
Sma I
Hind III
TPI 1
term.
pUC18
Eco RV
Bam HI
Sal I
Sma I
XbaI
Bar1
sub.P.
pSW15
pUC18
Bam HI
Sma I
Bam HI
TPI 1 term.
sub.P.
Eco RV
Nru I
Sph I
Bgl II
Hind III
Bar1
pCBS22
pUC18
ZC1479/
ZC1478
Eco RV
Bam HI
Bam HI
Xba I
sub.P.
TPI 1
term.
Sma I
Bar1
Sph I
Hind III
ADH 1
prom.
pSW22
pUC18

Figure 4

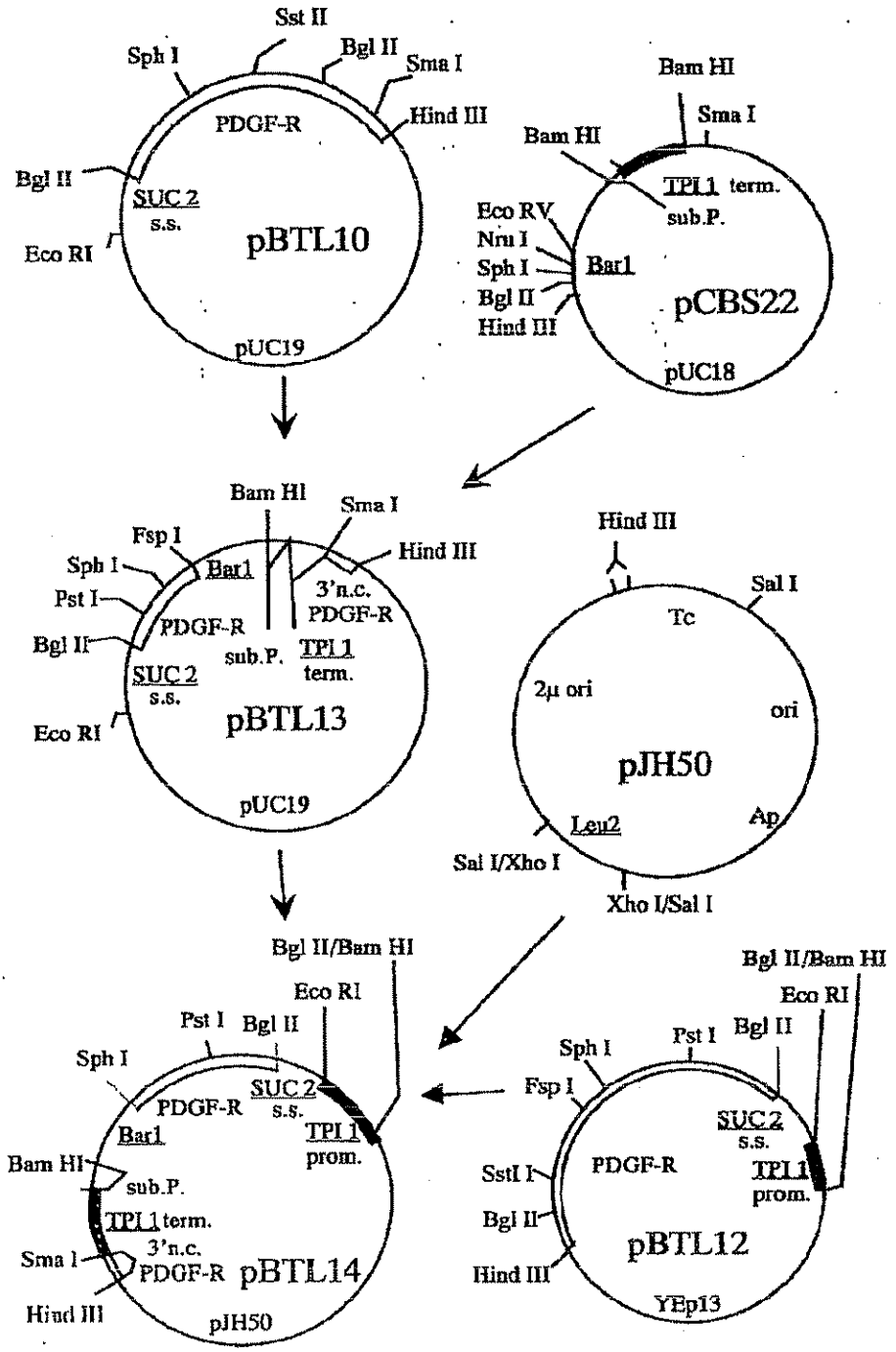

Sst II
Sph I
Bgl II
Sma I
Hind III
PDGF-R
Bgl II
SUC 2
s.s.
Eco RI
pBTL10
pUC19
Bam HI
Sma I
Bam HI
TPI 1 term.
sub.P.
Eco RV
Nru I
Sph I
Bgl II
Hind III
Bar1
pCBS22
pUC18
Bam HI
Sma I
Fsp I
Sph I
Pst I
Bgl II
Bar1
Hind III
3'n.c.
PDGF-R
PDGF-R
sub.P.
TPI 1
term.
SUC 2
s.s.
Eco RI
pBTL13
pUC19
Hind III
Sal I
Tc
2μ ori
ori
pJH50
Leu2
Ap
Sal I/Xho I
Xho I/Sal I
Bgl II/Bam HI
Eco RI
Pst I
Bgl II
Sph I
PDGF-R
SUC 2
s.s.
Bar1
TPI 1
prom.
Bam HI
sub.P.
TPI 1 term.
3'n.c.
Sma I
PDGF-R
pBTL14
Hind III
pJH50
Bgl II/Bam HI
Eco RI
Pst I
Bgl II
Sph I
Fsp I
SUC 2
s.s.
PDGF-R
TPI 1
prom.
SstI I
Bgl II
pBTL12
Hind III
YEp13

Figure 5

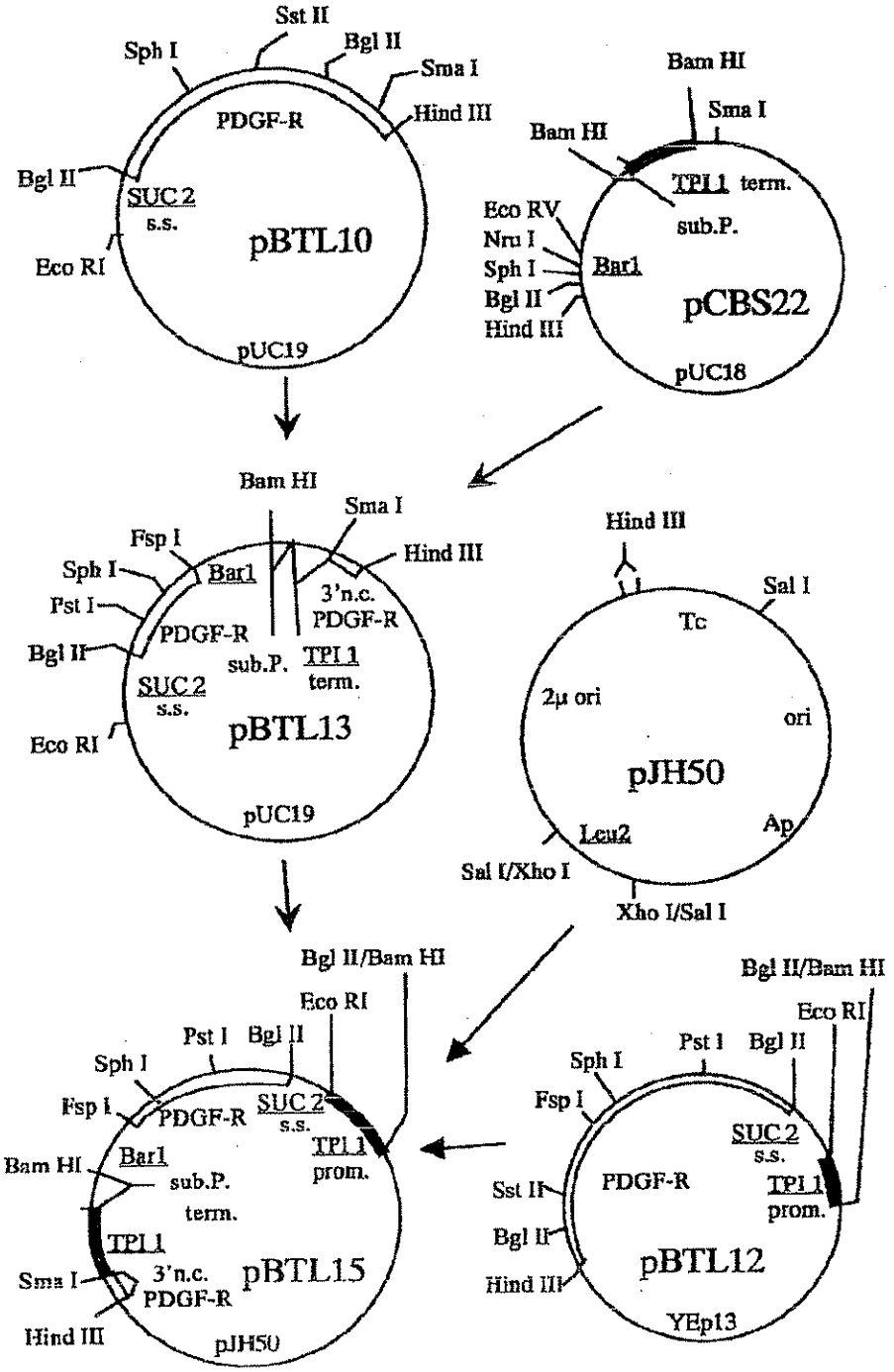

Sst II
Sph I
Bgl II
Sma I
Hind III
PDGF-R
Bgl II
SUC 2
s.s.
pBTL10
Eco RI
pUC19
Bam HI
Sma I
Bam HI
TPI 1 term.
sub.P.
Eco RV
Nru I
Sph I
Bgl II
Hind III
Bar1
pCBS22
pUC18
Bam HI
Sma I
Fsp I
Sph I
Bar1
Hind III
3'n.c.
PDGF-R
Pst I
PDGF-R
Bgl II
sub.P.
TPI 1
term.
SUC 2
s.s.
pBTL13
Eco RI
pUC19
Hind III
Sal I
Tc
2μ ori
ori
pJH50
Leu2
Ap
Sal I/Xho I
Xho I/Sal I
Bgl II/Bam HI
Eco RI
Pst I
Bgl II
Sph I
Fsp I
PDGF-R
SUC 2
s.s.
TPI 1
prom.
Bam HI
Bar1
sub.P.
term.
TPI 1
Sma I
3'n.c.
PDGF-R
pBTL15
Hind III
pJH50
Bgl II/Bam HI
Eco RI
Pst I
Bgl II
Sph I
Fsp I
SUC 2
s.s.
PDGF-R
TPI 1
prom.
Sst II
Bgl II
pBTL12
Hind III
YEp13

Figure 6

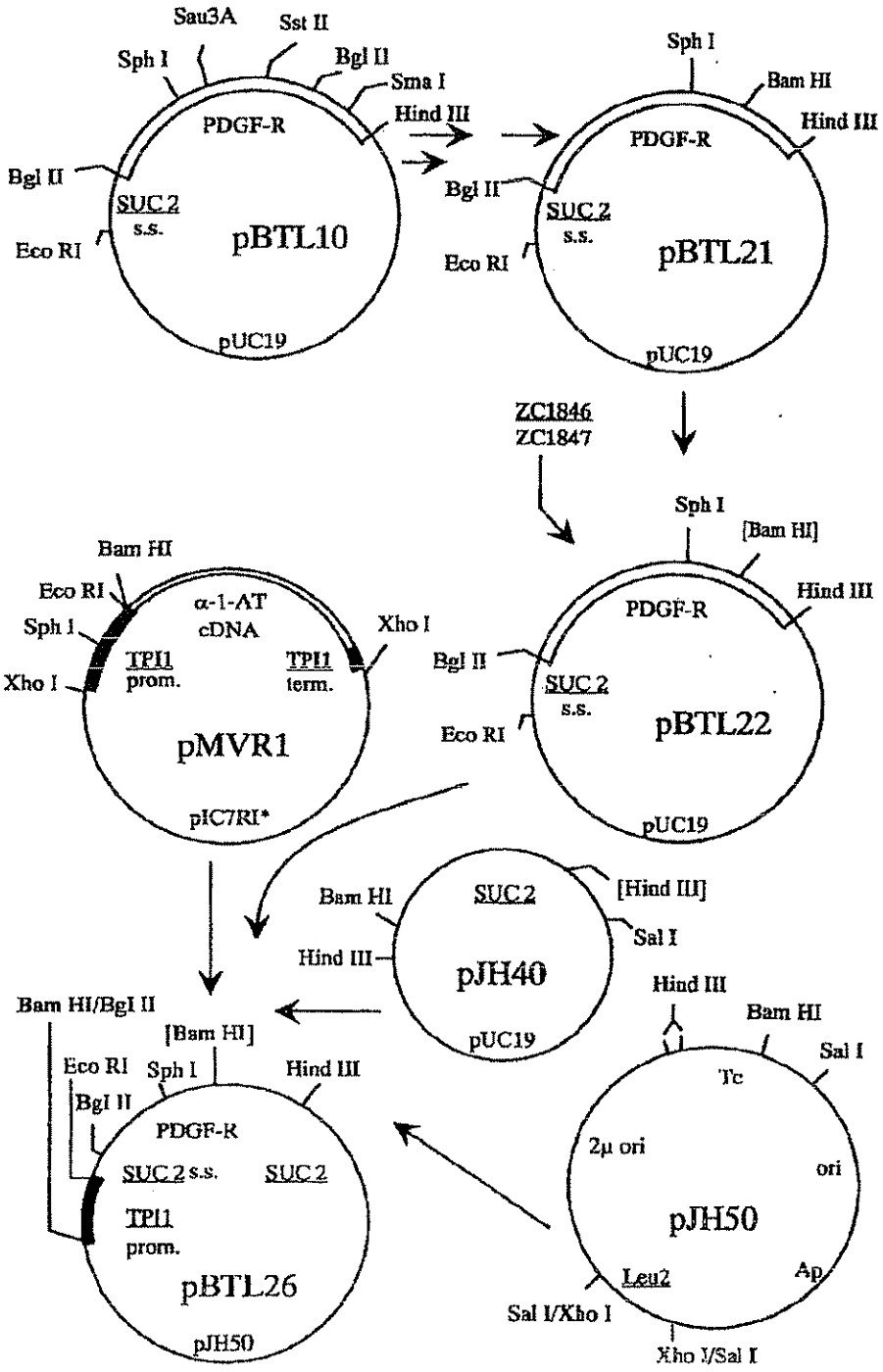

Sau3A
Sst II
Sph I
Bgl II
Sma I
Hind III
PDGF-R
Bgl II
SUC 2
s.s.
Eco RI
pBTL10
pUC19
Sph I
Bam HI
Hind III
PDGF-R
Bgl II
SUC 2
s.s.
Eco RI
pBTL21
pUC19
ZC1846
ZC1847
Sph I
[Bam HI]
Hind III
PDGF-R
Bgl II
SUC 2
s.s.
Eco RI
pBTL22
pUC19
Bam HI
Eco RI
Sph I
Xho I
α-1-AT
cDNA
TPI1
prom.
TPI1
term.
Xho I
pMVR1
pIC7RI*
Bam HI
SUC 2
[Hind III]
Hind III
Sal I
pJH40
pUC19
Bam HI/Bgl II
[Bam HI]
Eco RI
Sph I
Hind III
Bgl II
PDGF-R
SUC 2 s.s.
SUC 2
TPI1
prom.
pBTL26
pJH50
Hind III
Bam HI
Sal I
Tc
2μ ori
ori
pJH50
Leu2
Ap
Sal I/Xho I
Xho I/Sal I

Figure 7

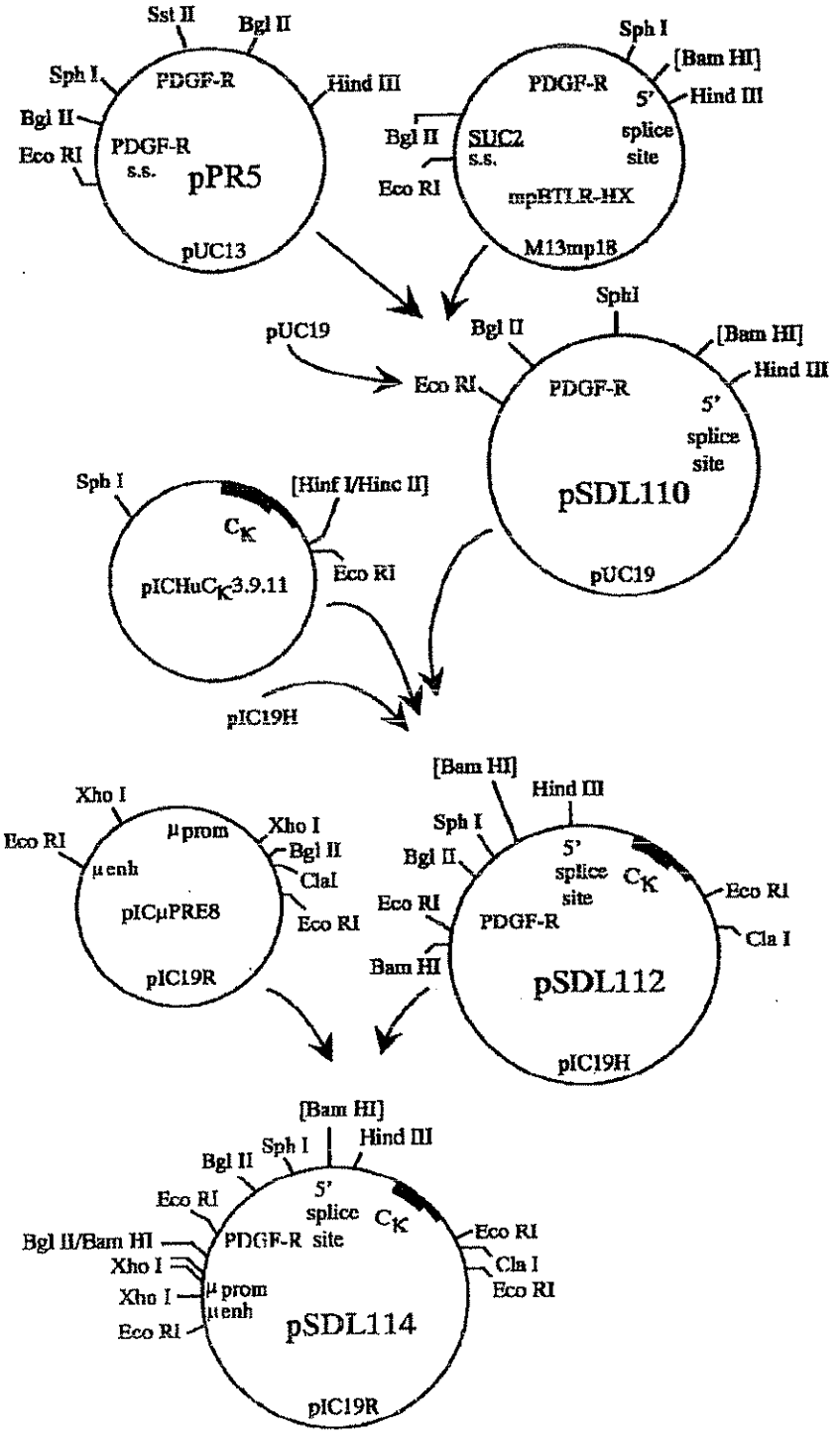

Sst II
Bgl II
Sph I
PDGF-R
Hind III
Bgl II
PDGF-R
Eco RI
s.s.
pPR5
pUC13
Sph I
[Bam HI]
PDGF-R
5'
Hind III
Bgl II
SUC2
splice
s.s.
site
Eco RI
mpRTLR-HX
M13mp18
pUC19
Sphl
Bgl II
[Bam HI]
Eco RI
PDGF-R
Hind III
5'
splice
site
pSDL110
pUC19
Sph I
[Hinf I/Hinc II]
C_K
pICHuC_K3.9.11
Eco RI
pIC19H
Xho I
Eco RI
μprom
Xho I
μenh
Bgl II
ClaI
pICμPRE8
Eco RI
pIC19R
[Bam HI]
Hind III
Sph I
Bgl II
5'
splice
site
C_K
Eco RI
Eco RI
PDGF-R
Cla I
Bam HI
pSDL112
pIC19H
[Bam HI]
Bgl II
Sph I
Hind III
Eco RI
5'
splice
C_K
Eco RI
Bgl II/Bam HI
PDGF-R
site
Cla I
Xho I
Eco RI
Xho I
μprom
μenh
pSDL114
Eco RI
pIC19R

Figure 8

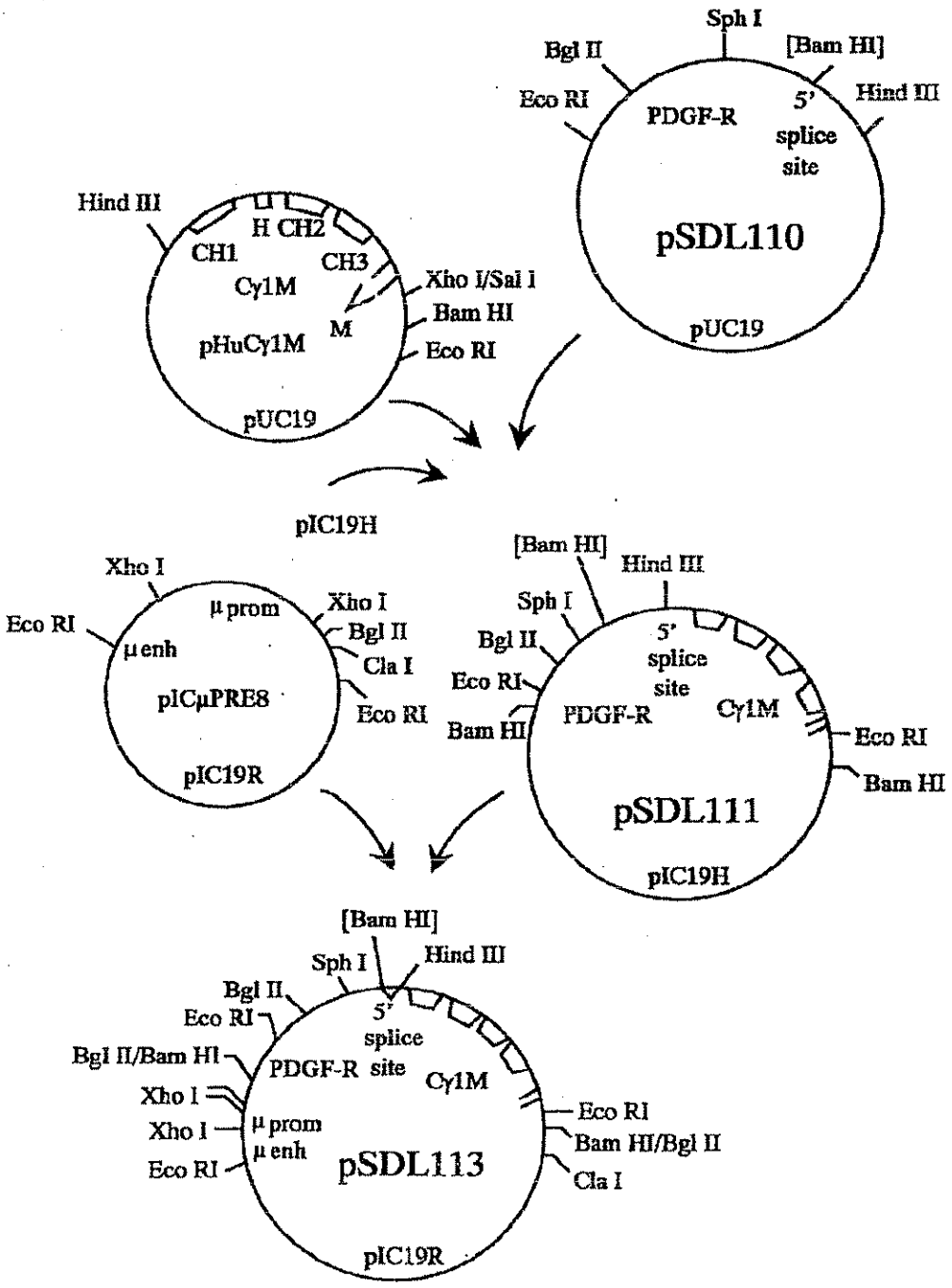

Sph I
Bgl II
[Bam HI]
Eco RI
PDGF-R
5'
splice
site
Hind III
pSDL110
pUC19
Hind III
H CH2
CH1
CH3
Cγ1M
Xho I/Sal I
Bam HI
M
pHuCγ1M
Eco RI
pUC19
pIC19H
Xho I
μ prom
Xho I
Eco RI
Bgl II
μ enh
Cla I
pICμPRE8
Eco RI
pIC19R
[Bam HI]
Hind III
Sph I
Bgl II
5'
splice
site
Eco RI
Bam HI
PDGF-R
Cγ1M
Eco RI
Bam HI
pSDL111
pIC19H
[Bam HI]
Sph I
Hind III
Bgl II
Eco RI
5'
splice
Bgl II/Bam HI
PDGF-R site
Cγ1M
Xho I
Eco RI
Xho I
μ prom
μ enh
pSDL113
Bam HI/Bgl II
Eco RI
Cla I
pIC19R

Figure 9

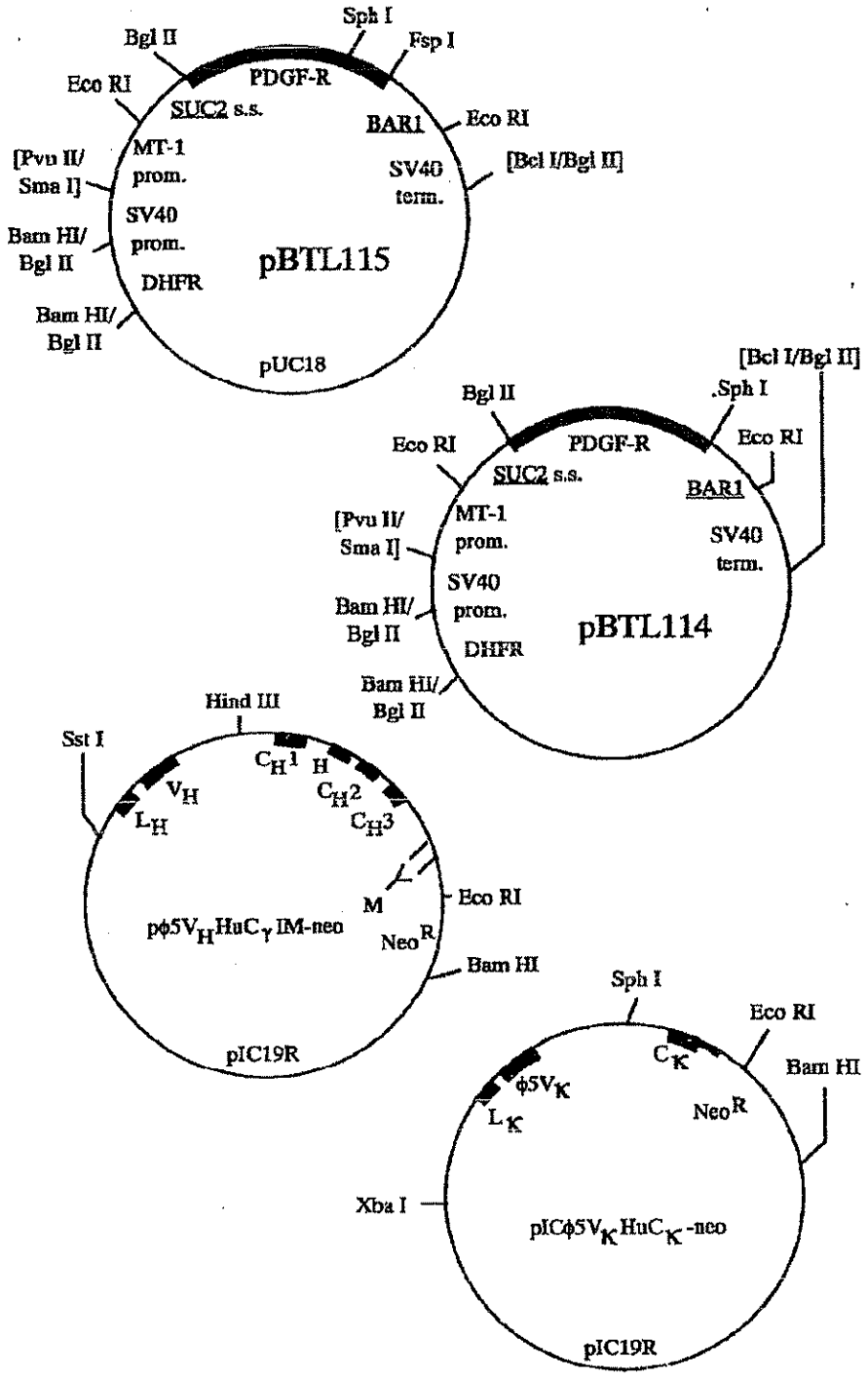
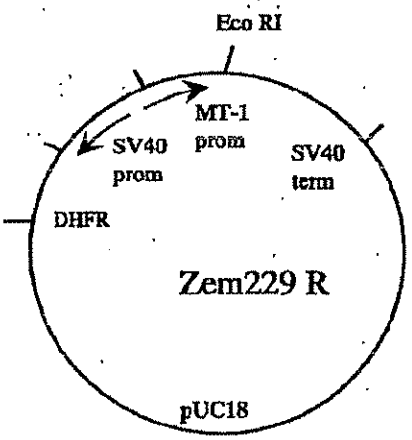
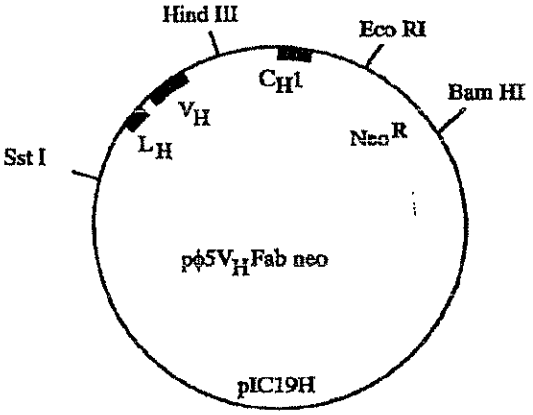

Sph I
Bgl II
Fsp I
Eco RI
PDGF-R
SUC2 s.s.
BAR1
Eco RI
[Pvu II/
Sma I]
MT-1
prom.
SV40
term.
[Bcl I/Bgl II]
SV40
prom.
Bam HI/
Bgl II
pBTL115
DHFR
Bam HI/
Bgl II
pUC18
[Bcl I/Bgl II]
Bgl II
.Sph I
Eco RI
PDGF-R
Eco RI
SUC2 s.s.
BAR1
[Pvu II/
Sma I]
MT-1
prom.
SV40
term.
SV40
prom.
Bam HI/
Bgl II
pBTL114
DHFR
Bam HI/
Bgl II
Hind III
Sst I
C_H1
H
V_H
C_H2
L_H
C_H3
M
Eco RI
$p\phi5V_HHuC_\gamma IM\text{-}neo$
Neo^R
Bam HI
pIC19R
Sph I
Eco RI
C_K
Bam HI
$\phi5V_K$
L_K
Neo^R
Xba I
$pIC\phi5V_KHuC_K\text{-}neo$
pIC19R

Figure 10


Eco RI
MT-1 prom
SV40 prom
SV40 term
DHFR
Zem229 R
pUC18


Hind III
Eco RI
C_H1
Bam HI
V_H
L_H
Neo^R
Sst I
pϕ5V_HFab neo
pIC19H

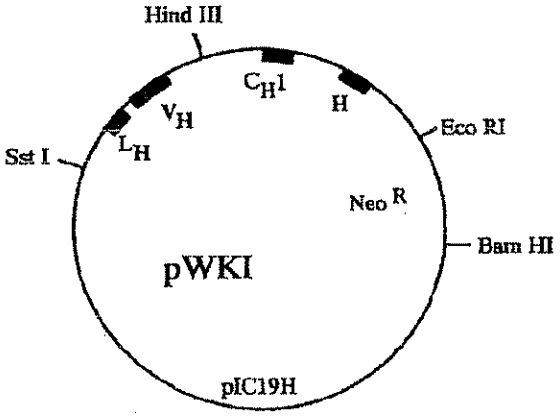

Hind III
C_H1
H
V_H
Eco RI
L_H
Sst I
Neo^R
Bam HI
pWKI
pIC19H

Figure 11A

```
   1  GCCCTGGGGACGGACCGTGGGCGGCGCGCAGCGGCGGGACGCGTTTTGGGGACGTGGTGGCCAGCGCCT
  70  TCCTGCAGACCCACAGGGAAGTACTCCCTTTGACCTCCGGGGAGCTGCGACCAGGTTATACGTTGCTGG

 139  TGGAAAAGTGACAATTCTAGGAAAAGAGCTAAAAGCCGGATCGGTGACCGAAAGTTTCCCAGAGCTATG
                                                                        M
                                                                        1

 208  GGGACTTCCCATCCGGCGTTCCTGGTCTTAGGCTGTCTTCTCACAGGGCTGAGCCTAATCCTCTGCCAG
      G  T  S  H  P  A  F  L  V  L  G  C  L  L  T  G  L  S  L  I  L  C  Q

 277  CTTTCATTACCCTCTATCCTTCCAAATGAAAATGAAAAGGTTGTGCAGCTGAATTCATCCTTTTCTCTG
      L  S  L  P  S  I  L  P  N  E  N  E  K  V  V  Q  L  N  S  S  F  S  L

 346  AGATGCTTTGGGGAGAGTGAAGTGAGCTGGCAGTACCCCATGTCTGAAGAAGAGAGCTCCGATGTGGAA
      R  C  F  G  E  S  E  V  S  W  Q  Y  P  M  S  E  E  E  S  S  D  V  E

 415  ATCAGAAATGAAGAAAACAACAGCGGCCTTTTTGTGACGGTCTTGGAAGTGAGCAGTGCCTCGGCGGCC
      I  R  N  E  E  N  N  S  G  L  F  V  T  V  L  E  V  S  S  A  S  A  A

 484  CACACAGGGTTGTACACTTGCTATTACAACCACACTCAGACAGAAGAGAATGAGCTTGAAGGCAGGCAC
      H  T  G  L  Y  T  C  Y  Y  N  H  T  Q  T  E  E  N  E  L  E  G  R  H

 553  ATTTACATCTATGTGCCAGACCCAGATGTAGCCTTTGTACCTCTAGGAATGACGGATTATTTAGTCATC
      I  Y  I  Y  V  P  D  P  D  V  A  F  V  P  L  G  M  T  D  Y  L  V  I

 622  GTGGAGGATGATGATTCTGCCATTATACCTTGTCGCACAACTGATCCCGAGACTCCTGTAACCTTACAC
      V  E  D  D  D  S  A  I  I  P  C  R  T  T  D  P  E  T  P  V  T  L  H

 691  AACAGTGAGGGGGTGGTACCTGCCTCCTACGACAGCAGACAGGGCTTTAATGGGACCTTCACTGTAGGG
      N  S  E  G  V  V  P  A  S  Y  D  S  R  Q  G  F  N  G  T  F  T  V  G

 760  CCCTATATCTGTGAGGCCACCGTCAAAGGAAAGAAGTTCCAGACCATCCCATTTAATGTTTATGCTTTA
      P  Y  I  C  E  A  T  V  K  G  K  K  F  Q  T  I  P  F  N  V  Y  A  L

 829  AAAGCAACATCAGAGCTGGATCTAGAAATGGAAGCTCTTAAAACCGTGTATAAGTCAGGGGAAACGATT
      K  A  T  S  E  L  D  L  E  M  E  A  L  K  T  V  Y  K  S  G  E  T  I

 898  GTGGTCACCTGTGCTGTTTTTAACAATGAGGTGGTTGACCTTCAATGGACTTACCCTGGAGAAGTGAAA
      V  V  T  C  A  V  F  N  N  E  V  V  D  L  Q  W  T  Y  P  G  E  V  K
```

Figure 11B

```
967  GGCAAAGGCATCACAATACTGGAAGAAATCAAAGTCCCATCCATCAAATTGGTGTACACTTTGACGGTC
     G  K  G  I  T  I  L  E  E  I  K  V  P  S  I  K  L  V  Y  T  L  T  V

1036 CCCGAGGCCACGGTGAAAGACAGTGGAGATTACGAATGTGCTGCCCGCCAGGCTACCAGGGAGGTCAAA
     P  E  A  T  V  K  D  S  G  D  Y  E  C  A  A  R  Q  A  T  R  E  V  K

1105 GAAATGAAGAAAGTCACTATTTCTGTCCATGAGAAAGGTTTCATTGAAATCAAACCCACCTTCAGCCAG
     E  M  K  K  V  T  I  S  V  H  E  K  G  F  I  E  I  K  P  T  F  S  Q

1174 TTGGAAGCTGTCAACCTGCATGAAGTCAAACATTTTGTTGTAGAGGTGCGGGCCTACCCACCTCCCAGG
     L  E  A  V  N  L  H  E  V  K  H  F  V  V  E  V  R  A  Y  P  P  P  R

1243 ATATCCTGGCTGAAAAACAATCTGACTCTGATTGAAAATCTCACTGAGATCACCACTGATGTGGAAAAG
     I  S  W  L  K  N  N  L  T  L  I  E  N  L  T  E  I  T  T  D  V  E  K

1312 ATTCAGGAAATAAGGTATCGAAGCAAATTAAAGCTGATCCGTGCTAAGGAAGAAGACAGTGGCCATTAT
     I  Q  E  I  R  Y  R  S  K  L  K  L  I  R  A  K  E  E  D  S  G  H  Y

1381 ACTATTGTAGCTCAAAATGAAGATGCTGTGAAGAGCTATACTTTTGAACTGTTAACTCAAGTTCCTTCA
     T  I  V  A  Q  N  E  D  A  V  K  S  Y  T  F  E  L  L  T  Q  V  P  S

1450 TCCATTCTGGACTTGGTCGATGATCACCATGGCTCAACTGGGGGACAGACGGTGAGGTGCACAGCTGAA
     S  I  L  D  L  V  D  D  H  H  G  S  T  G  G  Q  T  V  R  C  T  A  E

1519 GGCACGCCGCTTCCTGATATTGAGTGGATGATATGCAAAGATATTAAGAAATGTAATAATGAAACTTCC
     G  T  P  L  P  D  I  E  W  M  I  C  K  D  I  K  K  C  N  N  E  T  S

1588 TGGACTATTTTGGCCAACAATGTCTCAAACATCATCACGGAGATCCACTCCCGAGACAGGAGTACCGTG
     W  T  I  L  A  N  N  V  S  N  I  I  T  E  I  H  S  R  D  R  S  T  V

1657 GAGGGCCGTGTGACTTTCGCCAAAGTGGAGGAGACCATCGCCGTGCGATGCCTGGCTAAGAATCTCCTT
     E  G  R  V  T  F  A  K  V  E  E  T  I  A  V  R  C  L  A  K  N  L  L

1726 GGAGCTGAGAACCGAGAGCTGAAGCTGGTGGCTCCCACCCTGCGTTCTGAACTCACGGTGGCTGCTGCA
     G  A  E  N  R  E  L  K  L  V  A  P  T  L  R  S  E  L  T  V  A  A  A

1795 GTCCTGGTGCTGTTGGTGATTGTGATCATCTCACTTATTGTCCTGGTTGTCATTTGGAAACAGAAACCG
     V  L  V  L  L  V  I  V  I  I  S  L  I  V  L  V  V  I  W  K  Q  K  P
```

Figure 11C

```
1864  AGGTATGAAATTCGCTGGAGGGTCATTGAATCAATCAGCCCGGATGGACATGAATATATTTATGTGGAC
      R  Y  E  I  R  W  R  V  I  E  S  I  S  P  D  G  H  E  Y  I  Y  V  D

1933  CCGATGCAGCTGCCTTATGACTCAAGATGGGAGTTTCCAAGAGATGGACTAGTGCTTGGTCGGGTCTTG
      P  M  Q  L  P  Y  D  S  R  W  E  F  P  R  D  G  L  V  L  G  R  V  L

2002  GGGTCTGGAGCGTTTGGGAAGGTGGTTGAAGGAACAGCCTATGGATTAAGCCGGTCCCAACCTGTCATG
      G  S  G  A  F  G  K  V  V  E  G  T  A  Y  G  L  S  R  S  Q  P  V  M

2071  AAAGTTGCAGTGAAGATGCTAAAACCCACGGCCAGATCCAGTGAAAAACAAGCTCTCATGTCTGAACTG
      K  V  A  V  K  M  L  K  P  T  A  R  S  S  E  K  Q  A  L  M  S  E  L

2140  AAGATAATGACTCACCTGGGGCCACATTTGAACATTGTAAACTTGCTGGGAGCCTGCACCAAGTCAGGC
      K  I  M  T  H  L  G  P  H  L  N  I  V  N  L  L  G  A  C  T  K  S  G

2209  CCCATTTACATCATCACAGAGTATTGCTTCTATGGAGATTTGGTCAACTATTTGCATAAGAATAGGGAT
      P  I  Y  I  I  T  E  Y  C  F  Y  G  D  L  V  N  Y  L  H  K  N  R  D

2278  AGCTTCCTGAGCCACCACCCAGAGAAGCCAAAGAAAGAGCTGGATATCTTTGGATTGAACCCTGCTGAT
      S  F  L  S  H  H  P  E  K  P  K  K  E  L  D  I  F  G  L  N  P  A  D

2347  GAAAGCACACGGAGCTATGTTATTTTATCTTTTGAAAACAATGGTGACTACATGGACATGAAGCAGGCT
      E  S  T  R  S  Y  V  I  L  S  F  E  N  N  G  D  Y  M  D  M  K  Q  A

2416  GATACTACACAGTATGTCCCCATGCTAGAAAGGAAAGAGGTTTCTAAATATTCCGACATCCAGAGATCA
      D  T  T  Q  Y  V  P  M  L  E  R  K  E  V  S  K  Y  S  D  I  Q  R  S

2485  CTCTATGATCGTCCAGCCTCATATAAGAAGAAATCTATGTTAGACTCAGAAGTCAAAAACCTCCTTTCA
      L  Y  D  R  P  A  S  Y  K  K  K  S  M  L  D  S  E  V  K  N  L  L  S

2554  GATGATAACTCAGAAGGCCTTACTTTATTGGATTTGTTGAGCTTCACCTATCAAGTTGCCCGAGGAATG
      D  D  N  S  E  G  L  T  L  L  D  L  L  S  F  T  Y  Q  V  A  R  G  M

2623  GAGTTTTTGGCTTCAAAAAATTGTGTCCACCGTGATCTGGCTGCTCGCAACGTCCTCCTGGCACAAGGA
      E  F  L  A  S  K  N  C  V  H  R  D  L  A  A  R  N  V  L  L  A  Q  G

2692  AAAATTGTGAAGATCTGTGACTTTGGCCTGGCCAGAGACATCATGCATGATTCGAACTATGTGTCGAAA
      K  I  V  K  I  C  D  F  G  L  A  R  D  I  M  H  D  S  N  Y  V  S  K

2761  GGCAGTACCTTTCTGCCCGTGAAGTGGATGGCTCCTGAGAGCATCTTTGACAACCTCTACACCACACTG
      G  S  T  F  L  P  V  K  W  M  A  P  E  S  I  F  D  N  L  Y  T  T  L
```

Figure 11D

```
2830  AGTGATGTCTGGTCTTATGGCATTCTGCTCTGGGAGATCTTTTCCCTTGGTGGCACCCCTTACCCCGGC
      S  D  V  W  S  Y  G  I  L  L  W  E  I  F  S  L  G  G  T  P  Y  P  G

2899  ATGATGGTGGATTCTACTTTCTACAATAAGATCAAGAGTGGGTACCGGATGGCCAAGCCTGACCACGCT
      M  M  V  D  S  T  F  Y  N  K  I  K  S  G  Y  R  M  A  K  P  D  H  A

2968  ACCAGTGAAGTCTACGAGATCATGGTGAAATGCTGGAACAGTGAGCCGGAGAAGAGACCCTCCTTTTAC
      T  S  E  V  Y  E  I  M  V  K  C  W  N  S  E  P  E  K  R  P  S  F  Y

3037  CACCTGAGTGAGATTGTGGAGAATCTGCTGCCTGGACAATATAAAAAGAGTTATGAAAAAATTCACCTG
      H  L  S  E  I  V  E  N  L  L  P  G  Q  Y  K  K  S  Y  E  K  I  H  L

3106  GACTTCCTGAAGAGTGACCATCCTGCTGTGGCACGCATGCGTGTGGACTCAGACAATGCATACATTGGT
      D  F  L  K  S  D  H  P  A  V  A  R  M  R  V  D  S  D  N  A  Y  I  G

3175  GTCACCTACAAAAACGAGGAAGACAAGCTGAAGGACTGGGAGGGTGGTCTGGATGAGCAGAGACTGAGC
      V  T  Y  K  N  E  E  D  K  L  K  D  W  E  G  G  L  D  E  Q  R  L  S

3244  GCTGACAGTGGCTACATCATTCCTCTGCCTGACATTGACCCTGTCCCTGAGGAGGAGGACCTGGGCAAG
      A  D  S  G  Y  I  I  P  L  P  D  I  D  P  V  P  E  E  E  D  L  G  K

3313  AGGAACAGACACAGCTCGCAGACCTCTGAAGAGAGTGCCATTGAGACGGGTTCCAGCAGTTCCACCTTC
      R  N  R  H  S  S  Q  T  S  E  E  S  A  I  E  T  G  S  S  S  S  T  F

3382  ATCAAGAGAGAGGACGAGACCATTGAAGACATCGACATGATGGACGACATCGGCATAGACTCTTCAGAC
      I  K  R  E  D  E  T  I  E  D  I  D  M  M  D  D  I  G  I  D  S  S  D

3451  CTGGTGGAAGACAGCTTCCTGTAACTGGCGGATTCGAGGGGTTCCTTCCACTTCTGGGGCCACCTCTGG
      L  V  E  D  S  F  L
                        1089

3520  ATCCCGTTCAGAAAACCACTTTATTGCAATGCGGAGGTTGAGAGGAGGACTTGGTTGATGTTTAAAGAG
3589  AAGTTCCCAGCCAAGGGCCTCGGGGAGCGTTCTAAATATGAATGAATGGGATATTTTGAAATGAACTTT
3658  GTCAGTGTTGCCTCTTGCAATGCCTCAGTAGCATCTCAGTGGTGTGTGAAGTTTGGAGATAGATGGATA
3727  AGGGAATAATAGGCCACAGAAGGTGAACTTTGTGCTTCAAGGACATTGGTGAGAGTCCAACAGACACAA
3796  TTTATACTGCGACAGAACTTCAGCATTGTAATTATGTAAATAACTCTAACCAAGGCTGTGTTTAGATTG
3865  TATTAACTATCTTCTTTGGACTTCTGAAGAGACCACTCAATCCATCCTGTACTTCCCTCTTGAAACCTG
3934  ATGTAGCTGCTGTTGAACTTTTTAAAGAAGTGCATGAAAAACCATTTTTGAACCTTAAAAGGTACTGGT
4003  ACTATAGCATTTTGCTATCTTTTTTAGTGTTAAAGAGATAAAGAATAATAAG
```

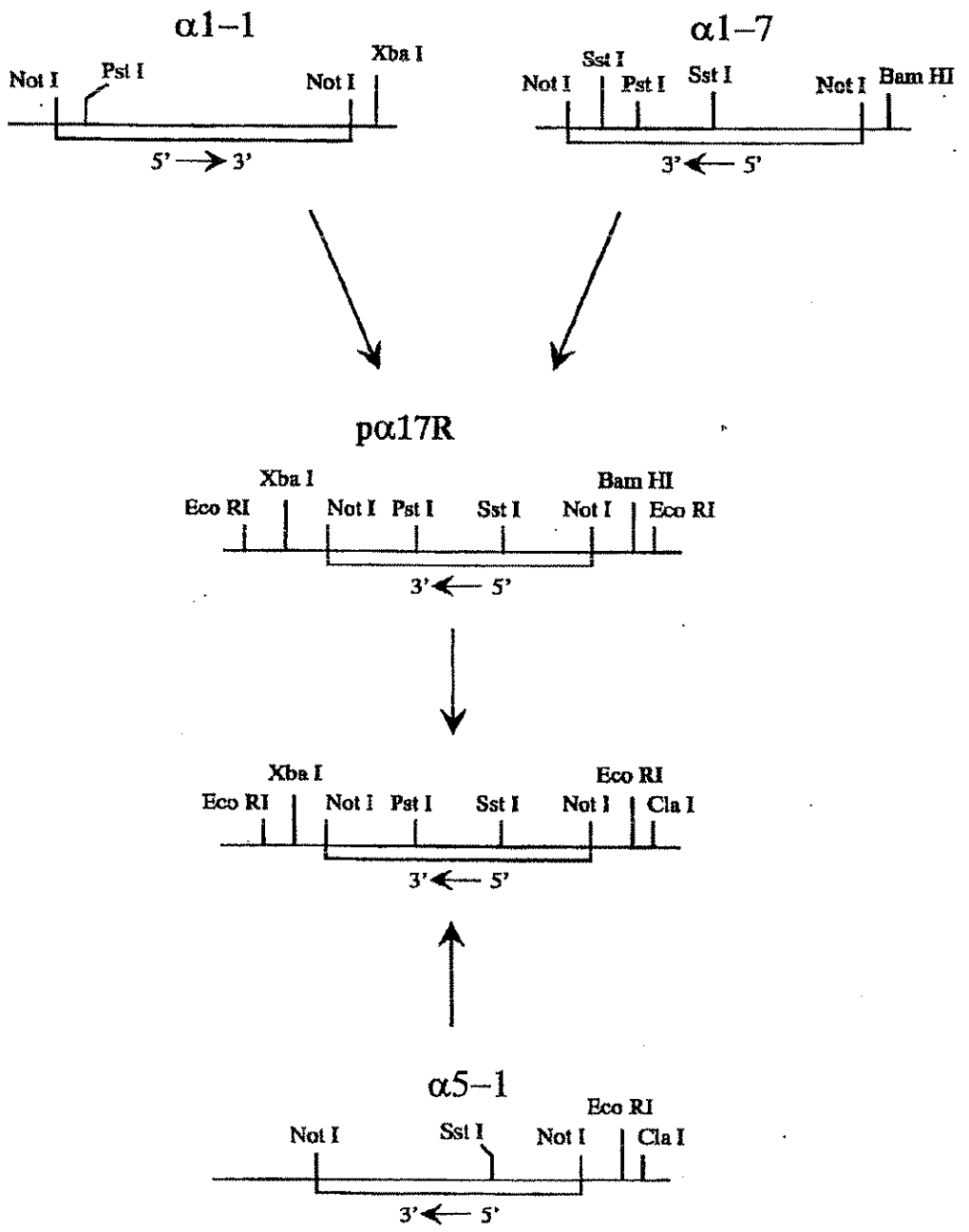

α1–1
Not I
Pst I
Not I
Xba I
5' → 3'
α1–7
Not I
Sst I
Pst I
Sst I
Not I
Bam HI
3' ← 5'
pα17R
Eco RI
Xba I
Not I
Pst I
Sst I
Not I
Bam HI
Eco RI
3' ← 5'
Eco RI
Xba I
Not I
Pst I
Sst I
Not I
Eco RI
Cla I
3' ← 5'
α5–1
Not I
Sst I
Not I
Eco RI
Cla I
3' ← 5'

Figure 12

BIOLOGICALLY ACTIVE DIMERIZED AND MULTIMERIZED POLYPEPTIDE FUSIONS

CROSS REFERENCE TO RELATED APPLICATION

This application is a divisional of U.S. application Ser. No. 8/477,329, filed Jun. 7, 1995, now U.S. Pat. No. 5,750,375 which is a continuation of U.S. application Ser. No. 180,195, filed Jan. 11, 1994,now U.S. Pat. No. 5,567,584, which is a continuation of U.S. application Ser. No. 07/634,510 filed Dec. 27, 1990, now abandoned, which is a continuation-in-part of U.S. application Ser. No. 7/347,291, filed May 2, 1989, now U.S. Pat. No. 5,155,027 which is a continuation-in-part of U.S. application Ser. No/7/146,877, filed Jan. 22, 1988, now abandoned.

TECHNICAL FIELD Technical Field

The present invention is generally directed toward the expression of proteins, and more specifically, toward the expression of growth factor receptor analogs and biologically active dimerized polypeptide fusions.

BACKGROUND OF THE INVENTION

In higher eucaryotic cells, the interaction between receptors and ligands (e.g., hormones) is of central importance in the transmission of and response to a variety of extracellular signals. It is generally accepted that hormones and growth factors elicit their biological functions by binding to specific recognition sites (receptors) in the plasma membranes of their target cells. Upon ligand binding, a receptor undergoes a conformational change, triggering secondary cellular responses that result in the activation or inhibition of intracellular processes. The stimulation or blockade of such an interaction by pharmacological means has important therapeutic implications for a wide variety of illnesses.

Ligands fall into two classes: those that have stimulatory activity, termed agonists; and those that block the effects elicited by the original ligands, termed antagonists. The discovery of agonists that differ in structure and composition from the original ligand may be medically useful. In particular, agonists that are smaller than the original ligand may be especially useful. The bioavailability of these smaller agonists may be greater than that of the original ligand. This may be of particular importance for topical applications and for instances when diffusion of the agonist to its target sites is inhibited by poor circulation. Agonists may also have slightly different spectra of biological activity and/or different potencies, allowing them to be used in very specific situations. Agonists that are smaller and chemically simpler than the native ligand may be produced in greater quantity and at lower cost. The identification of antagonists which specifically block, for example, growth factor receptors has important pharmaceutical applications. Antagonists that block receptors against the action of endogenous, native ligand may be used as therapeutic agents for conditions including atherosclerosis, autocrine tumors, fibroplasia and keloid formation.

The discovery of new ligands that may be used in pharmaceutical applications has centered around designing compounds by chemical modification, complete synthesis, and screening potential ligands by complex and costly screening procedures. The process of designing a new ligand usually begins with the alteration of the structure of the original effector molecule. If the original effector molecule is known to be chemically simple, for example, a catecholamine or prostaglandin, the task is relatively straightforward. However, if the ligand is structurally complex, for example, a peptide hormone or a growth factor, finding a molecule which is functionally equivalent to the original ligand becomes extremely difficult.

Currently, potential ligands are screened using radioligand binding methods (Lefkowitz et al., *Biochem, Biophys. Res. Comm.* 60: 703–709, 1974; Aurbach et al., *Science* 186: 1223–1225, 1974; Atlas et al., *Proc. Natl. Acad. Sci. USA* 71: 4246–4248, 1974). Potential ligands can be directly assayed by binding the radiolabeled compounds to responsive cells, to the membrane fractions of disrupted cells, or to solubilized receptors. Alternatively, potential ligands may be screened by their ability to compete with a known labeled ligand for cell surface receptors.

The success of these procedures depends on the availability of reproducibly high quality preparations of membrane fractions or receptor molecules, as well as the isolation of responsive cell lines. The preparation of membrane fractions and soluble receptor molecules involves extensive manipulations and complex purification steps. The isolation of membrane fractions requires gentle manipulation of the preparation, a procedure which does not lend itself to commercial production. It is very difficult to maintain high biological activity and biochemical purity of receptors when they are purified by classical protein chemistry methods. Receptors, being integral membrane proteins, require cumbersome purification procedures, which include the use of detergents and other solvents that interfere with their biological activity. The use of these membrane preparations in ligand binding assays typically results in low reproducibility due to the variability of the membrane preparations.

As noted above, ligand binding assays require the isolation of responsive cell lines. Often, only a limited subset of cells is responsive to a particular agent, and such cells may be responsive only under certain conditions. In addition, these cells may be difficult to grow in culture or may possess a low number of receptors. Currently available cell types responsive to platelet-derived growth factor (PDGF), for example, contain only a low number (up to 4×10^5; see Bowen-Pope and Ross, *J. Biol. Chem.* 257: 5161–5171, 1982) of receptors per cell, thus requiring large numbers of cells to assay PDGF analogs or antagonists.

Presently, only a few naturally-occurring secreted receptors, for example, the interleukin-2 receptor (IL-2-R) have been identified. Rubin et al. (*J. Immun.* 135: 3172–3177, 1985) have reported the release of large quantities of IL-2-R into the culture medium of activated T-cell lines. Bailon et al. (*BioTechnology* 5: 1195–1198, 1987) have reported the use of a matrix-bound interleukin-2 receptor to purify recombinant interleukin-2.

Three other receptors have been secreted from mammalian cells. The insulin receptor (Ellis et al., *J. Cell Biol.* 150: 14a, 1987), the HIV-1 envelope glyco-protein cellular receptor CD4 (Smith et al., *Science* 238: 1704–1707, 1987), the murine IL-7 receptor (*Cell* 60: 941–951, 1990) and the epidermal growth factor (EGF) receptor (Livneh et al., *J. Biol. Chem.* 261: 12490–12497, 1986) have been secreted from mammalian cells using truncated cDNAs that encode portions of the extracellular domains.

Naturally-occurring, secreted receptors have not been widely identified, and there have been only a limited number of reports of secreted recombinant receptors. Secreted receptors may be used in a variety of assays, which include assays to determine the presence of ligand in biological fluids and assays to screen for potential agonists and antagonists.

Current methods for ligand screening and ligand/receptor binding assays have been limited to those using preparations of whole cells or cell membranes for as a source for receptor molecules. The low reproducibility and high cost of producing such preparations does not lend itself to commercial production. There is therefore a need in the art for a method of producing secreted receptors. There is a further need in the art for an assay system that permits high volume screening of compounds that may act on higher eucaryotic cells via specific surface receptors. This assay system should be rapid, inexpensive and adaptable to high volume screening. The present invention discloses such a method and assay system, and further provides other related advantages.

DISCLOSURE OF INVENTION

Briefly stated, the present invention discloses methods for producing secreted receptor analogs, including receptor analogs and secreted platelet-derived growth factor receptor (PDGF-R) analogs. In addition, the present invention discloses methods for producing secreted biologically active dimerized polypeptide fusions.

Within one aspect of the invention a method for producing a secreted PDGF-R analog is disclosed, comprising (a) introducing into a eukaryotic host cell a DNA construct comprising a transcriptional promoter operatively linked to a secretory signal sequence followed downstream of and in proper reading frame with a DNA sequence encoding at least a portion of the ligand-binding domain of a PDGF-R, the portion including a ligand-binding domain; (b) growing the host cell in an appropriate growth medium under physiological conditions to allow the secretion of a PDGF-R analog encoded by said DNA sequence; and (c) isolating the PDGF-R analog from the host cell.

Within one embodiment of the present invention, a PDGF-R analog comprising the amino acid sequence of FIGS. 1A–1B (Sequence ID Numbers 1 and 2) from isoleucine, number 29, to methionine, number 441, is secreted. Within another embodiment, a PDGF-R analog comprising the amino acid sequence of FIGS. 1A–1B (Sequence ID Numbers 1 and 2) from isoleucine, number 29 to lysine, number 531 is secreted. Within yet another embodiment of the invention, a PDGF-R analog comprising the amino acid sequence of FIGS. 11A–11B (Sequence ID Numbers 35 and 36) from glutamine, number 24 to glutamic acid, number 524 is secreted.

Yet another aspect of the present invention discloses a method for producing a secreted, biologically active dimerized polypeptide fusion. The method generally comprises a) introducing into a eukaryotic host cell a DNA construct comprising a transcriptional promoter operatively linked to a secretory signal sequence followed downstream by and in proper reading frame with a DNA sequence encoding a non-immunoglobulin polypeptide requiring dimerization for biological activity joined to a dimerizing protein, (b) growing the host cell in an appropriate growth medium under physiological conditions to allow the secretion of a dimerized polypeptide fusion encoded by said DNA sequence; and (c) isolating the biologically active dimerized polypeptide fusion from the host cell.

Within one embodiment, the dimerizing protein is yeast invertase. Within another embodiment, the dimerizing protein is at least a portion of an immunoglobulin light chain. Within another embodiment, the dimerizing protein is at least a portion of an immunoglobulin heavy chain.

In another aspect of the invention, a method is disclosed for producing a secreted, biologically active dimerized

polypeptide fusion, comprising (a) introducing into a eukaryotic host cell a first DNA construct comprising a transcriptional promoter operatively linked to a first secretory signal sequence followed downstream by and in proper reading frame with a first DNA sequence encoding a non-immunoglobulin polypeptide requiring dimerization for biological activity joined to an immunoglobulin light chain constant region; (b) introducing into the host cell a second DNA construct comprising a transcriptional promoter operatively linked to a second secretory signal sequence followed downstream by and in proper reading frame with a second DNA sequence encoding an immunoglobulin heavy chain constant region domain selected from the group consisting of C_H1, C_H2, C_H3, and C_HA; (c) growing the host cell in an appropriate growth medium under physiological conditions to allow the secretion of a dimerized polypeptide fusion encoded by said first and second DNA sequences; and (d) isolating the dimerized polypeptide fusion from the host cell. In one embodiment, the second DNA sequence further encodes an immunoglobulin heavy chain hinge region wherein the hinge region is joined to the heavy chain constant region domain. In a preferred embodiment, the second DNA sequence further encodes an immunoglobulin variable region joined upstream of and in proper reading frame with the immunoglobulin heavy chain constant region.

In another aspect of the invention, a method is disclosed for producing a secreted, biologically active dimerized polypeptide fusion, comprising (a) introducing into a eukaryotic host cell a first DNA construct comprising a transcriptional promoter operatively linked to a first secretory signal sequence followed downstream by and in proper reading frame with a first DNA sequence encoding a non-immunoglobulin polypeptide requiring dimerization for biological activity joined to an immunoglobulin heavy chain constant region domain selected from the group consisting of C_H1, C_H2, C_H3, and C_HA; (b) introducing into the host cell a second DNA construct comprising a transcriptional promoter operatively linked to a second secretory signal sequence followed downstream by and in proper reading frame with a second DNA sequence encoding an immunoglobulin light chain constant region; (c) growing the host cell in an appropriate growth medium under physiological conditions to allow the secretion of a dimerized polypeptide fusion encoded by said first and second DNA sequences; and (d) isolating the dimerized polypeptide fusion from the host cell. In one embodiment, the first DNA sequence further encodes an immunoglobulin heavy chain hinge region wherein the hinge region is joined to the heavy chain constant region domain. In a preferred embodiment, the second DNA sequence further encodes an immunoglobulin variable region joined upstream of and in proper reading frame with the immunoglobulin light chain constant region.

In another aspect of the invention, a method is disclosed for producing a secreted, biologically active dimerized polypeptide fusion, comprising (a) introducing into a eukaryotic host cell a DNA construct comprising a transcriptional promoter operatively linked to a secretory signal sequence followed downstream by and in proper reading frame with a DNA sequence encoding a non-immunoglobulin polypeptide requiring dimerization for biological activity joined to an immunoglobulin heavy chain constant region domain selected from the group consisting of C_H1, C_H2, C_H3, and C_HA; (b) growing the host cell in an appropriate growth medium under physiological conditions to allow the secretion of a dimerized polypeptide fusion encoded by said first and second DNA sequences; and (c)

isolating the biologically active dimerized polypeptide fusion from the host cell. In one embodiment, the DNA sequence further encodes an immunoglobulin heavy chain hinge region wherein the hinge region is joined to the heavy chain constant region domain.

In another aspect of the invention, a method is disclosed for producing a secreted, biologically active dimerized polypeptide fusion, comprising (a) introducing into a eukaryotic host cell a first DNA construct comprising a transcriptional promoter operatively linked to a first secretory signal sequence followed downstream by and in proper reading frame with a first DNA sequence encoding a first polypeptide chain of a non-immunoglobulin polypeptide dimer requiring dimerization for biological activity joined to an immunoglobulin heavy chain constant region domain, selected from the group consisting of C_H1, C_H2, C_H3, and C_LA; (b) introducing into the host cell a second DNA construct comprising a transcriptional promoter operatively linked to a second secretory signal sequence followed downstream by and in proper reading frame with a second DNA sequence encoding a second polypeptide chain of the non-immunoglobulin polypeptide dimer joined to an immunoglobulin light chain constant region domain; (c) growing the host cell in an appropriate growth medium under physiological conditions to allow the secretion of a dimerized polypeptide fusion encoded by said first and second DNA sequences wherein said dimerized polypeptide fusion exhibits biological activity characteristic of said non-immunoglobulin polypeptide dimer; and (d) isolating the dimerized polypeptide fusion from the host cell. In one embodiment the first DNA sequence further encodes an immunoglobulin heavy chain hinge region domain wherein the hinge region is joined to the immunoglobulin heavy chain constant region domain.

Within one embodiment of the present invention, a biologically active dimerized polypeptide fusion comprising the amino acid sequence of FIGS. 1A–1B (Sequence ID Numbers 1 and 2) from isoleucine, number 29, to methionine, number 441, is secreted. Within another embodiment, a biologically active dimerized polypeptide fusion comprising the amino acid sequence of FIGS. 1A–1B (Sequence ID Numbers 1 and 2) from isoleucine, number 29 to lysine, number 531 is secreted. Within another embodiment of the invention, a biologically active dimerized polypeptide fusion comprising the amino acid sequence of FIGS. 11A–11B (Sequence ID Numbers 35 and 36) from glutamine, number 24 to glutamic acid, number 524 is secreted. Within yet another embodiment of the invention, a biologically active dimerized polypeptide fusion comprising the amino acid sequence of FIGS. 1A–1B (Sequence ID Numbers 1 and 2) from isoleucine, number 29 to lysine, number 531 dimerized to the amino acid sequence of FIGS. 11A–11B (sequence ID Numbers 35 and 36) from glutamine, number 24 to glutamic acid, number 524 is secreted.

In yet another aspect of the invention, a method is disclosed for producing a secreted receptor analog, comprising (a) introducing into a eukaryotic host cell a DNA construct comprising a transcriptional promoter operatively linked to at least one secretory signal sequence followed downstream by and in proper reading frame with a DNA sequence encoding a ligand-binding domain of a receptor requiring dimerization for biological activity joined to a dimerizing protein; (b) growing the host cell in an appropriate growth medium under physiological conditions to allow the secretion of a receptor analog encoded by said DNA sequence, and (c) isolating the receptor analog from the host cell.

In yet another aspect of the invention, a method is disclosed for producing a secreted receptor analog, comprising (a) introducing into a eukaryotic host cell a first DNA construct comprising a transcriptional promoter operatively linked to a first secretory signal sequence followed downstream by and in proper reading frame with a first DNA sequence encoding a ligand-binding domain of a receptor requiring dimerization for biological activity joined to an immunoglobulin light chain constant region; (b) introducing into the host cell a second DNA construct comprising a transcriptional promoter operatively linked to a second secretory signal sequence followed downstream by and in proper reading frame with a second DNA sequence encoding an immunoglobulin heavy chain constant region domain, selected from the group consisting of C_H1, C_H2, C_H3, and C_LA; (c) growing the host cell in an appropriate growth medium under physiological conditions to allow the secretion of a receptor analog encoded by said first and second DNA sequences; and (d) isolating the receptor analog from the host cell. In one embodiment, the second DNA sequence further encodes an immunoglobulin heavy chain hinge region wherein the hinge region is joined to the heavy chain constant region domain. In a preferred embodiment, the second DNA sequence further encodes an immunoglobulin variable region joined upstream of and in proper reading frame with the immunoglobulin heavy chain constant region.

In another aspect of the invention, a method is disclosed for producing a secreted receptor analog, comprising (a) introducing into a eukaryotic host cell a DNA construct comprising a transcriptional promoter operatively linked to a secretory signal sequence followed downstream by and in proper reading frame with a DNA sequence encoding a ligand-binding domain of a receptor requiring dimerization for biological activity joined to an immunoglobulin heavy chain constant region domain, selected from the group C_H1, C_H2, C_H3, and C_LA; (b) growing the host cell in an appropriate growth medium under physiological conditions to allow the secretion of the receptor analog; and (c) isolating the receptor analog from the host cell. In one embodiment, the DNA sequence further encodes an immunoglobulin heavy chain hinge region wherein the hinge region is joined to the heavy chain constant region domain.

In another aspect of the invention, a method is disclosed for producing a secreted receptor analog, comprising (a) introducing into a eukaryotic host cell a first DNA construct comprising a transcriptional promoter operatively linked to a first secretory signal sequence followed downstream of and in proper reading frame with a first DNA sequence encoding a ligand-binding domain of a receptor requiring dimerization for biological activity joined to an immunoglobulin heavy chain constant region domain, selected from the group C_H1, C_H2, C_H3, and C_LA; (b) introducing into the host cell a second DNA construct comprising a transcriptional promoter operatively linked to a second secretory signal sequence followed downstream by and in proper reading frame with a second DNA sequence encoding an immunoglobulin light chain constant region; (c) growing the host cell in an appropriate growth medium under physiological conditions to allow the secretion of a receptor analog encoded by said first and second DNA sequences; and (d) isolating the receptor analog from the host cell. In one embodiment, the first DNA sequence further encodes an immunoglobulin heavy chain hinge region wherein the hinge region is joined to the heavy chain constant region domain. In a preferred embodiment, the second DNA sequence further encodes an immunoglobulin variable

region joined upstream of and in proper reading frame with the immunoglobulin light chain constant region.

In another aspect of the invention, a method is disclosed for producing a secreted receptor analog, comprising (a) introducing into a eukaryotic host cell a first DNA construct comprising a transcriptional promoter operatively linked to a first secretory signal sequence followed downstream in proper reading frame by a first DNA sequence encoding a first polypeptide chain of a ligand-binding domain of a receptor requiring dimerization for biological activity joined to an immunoglobulin heavy chain constant region domain, selected from the group C_H1, C_H2, C_H3, and C_H4; (b) introducing into the host cell a second DNA construct comprising a transcriptional promoter operatively linked to a second secretory signal sequence followed downstream by and in proper reading frame with a second DNA sequence encoding a second polypeptide chain of the ligand-binding domain of said receptor joined to an immunoglobulin light chain constant region domain; (c) growing the host cell in an appropriate growth medium under physiological conditions to allow the secretion of a receptor analog encoded by said first and second DNA sequences; and (d) isolating the receptor analog from the host cell. In one embodiment the first DNA sequence further encodes an immunoglobulin heavy chain hinge region domain wherein the hinge region is joined to the immunoglobulin heavy chain constant region domain.

Host cells for use in the present invention include cultured mammalian cells and fungal cells. In a preferred embodiment strains of the yeast *Saccharomyces cerevisiae* are used as host cells. Within another preferred embodiment cultured rodent myeloma cells are used as host cells.

Within one embodiment of the present invention, a receptor analog is a PDGF-R analog comprising the amino acid sequence of FIG. 1 (Sequence ID Numbers 1 and 2) from isoleucine, number 29, to methionine, number 441. Within another embodiment a PDGF-R analog comprises the amino acid sequence of FIG. 1 (Sequence ID Numbers 1 and 2) from isoleucine, number 29, to lysine, number 531. Within another embodiment of the invention, a PDGF-R analog comprises the amino acid sequence of FIG. 11 (Sequence ID Numbers 35 and 36) from glutamine, number 24 to glutamic acid, number 524 is secreted. Within yet another embodiment of the invention, a PDGF-R analog comprises the amino acid sequence of FIG. 1 (Sequence ID Numbers 1 and 2) from isoleucine, number 29 to lysine, number 531 and the amino acid sequence of FIG. 11 (Sequence ID Numbers 35 and 36) from glutamine, number 24 to glutamic acid, number 524 is secreted.

PDGF-R analogs produced by the above-disclosed methods may be used, for instance, within a method for determining the presence of human PDGF or an isoform thereof in a biological sample.

A method for determining the presence of human PDGF or an isoform thereof in a biological sample is disclosed and comprises (a) incubating a polypeptide comprising a PDGF receptor analog fused to a dimerizing protein with a biological sample suspected of containing PDGF or an isoform thereof under conditions that allow the formation of receptor/ligand complexes; and (b) detecting the presence of receptor/ligand complexes, and therefrom determining the presence of PDGF or an isoform thereof. Suitable biological samples in this regard include blood, urine, plasma, serum, platelet and other cell lysates, platelet releasates, cell suspensions, cell-conditioned culture media, and chemically or physically separated portions thereof.

These and other aspects of the present invention will become evident upon reference to the following detailed description and attached drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

FIGS. 1A–1D (Sequence ID Numbers 1 and 2) illustrate the nucleotide sequence of a representative PDGF β-receptor cDNA and the derived amino acid sequence of the primary translation product and corresponds to Sequence ID Number 1). Numbers above the lines refer to the nucleotide sequence; numbers below the lines refer to the amino acid sequence.

FIG. 2 illustrates the construction of pBTL10, pBTL11 and pBTL12.

FIG. 3 illustrates the construction of pCBS22.

FIG. 4 illustrates the construction of pBTL13 and pBTL14.

FIG. 5 illustrates the construction of pBTL15.

FIG. 6 illustrates the construction of pBTL22 and pBTL26.

FIG. 7 illustrates the construction of pSDL114. Symbols used are S.S., signal sequence, C_κ, immunoglobulin light chain constant region sequence; μ prom, μ promoter; μ enh, μ enhancer.

FIG. 8 illustrates the construction of pSDLB113. Symbols used are S.S., signal sequence; C_H1 , C_H2, C_H3, immunoglobulin heavy chain constant region domain sequences; H, immunoglobulin heavy chain hinge region sequence; M, immunoglobulin membrane anchor sequences; c_γ1M, immunoglobulin heavy chain constant region and membrane anchor sequences.

FIG. 9 illustrates the constructions pBTL115, pBTL114, $p\phi5V_HHuC_\gamma1M$-neo, $plC\phi5V_\kappa HuC_\kappa$-neo. Symbols used are set forth in FIGS. 7 and 8, and also include L_H, mouse immunoglobulin heavy chain signal sequence; V_H, mouse immunoglobulin heavy chain variable region sequence; E, mouse immunoglobulin heavy chain enhancer sequence; L_κ, mouse immunoglobulin light chain signal sequence; $\phi5V_\kappa$, mouse immunoglobulin light chain variable region sequence; Neo^R, *E. coli* neomycin resistance gene.

FIG. 10 illustrates the constructions Zem229R, $p\phi5V_HFab$-neo and pWKI. Symbols used are set forth in FIG. 9.

FIGS. 11A–11D illustrate the sequence of a representative PDGF α-receptor CDNA and the deduced amino acid sequence (using standard one-letter codes) encoded by the CDNA and corresponds to Sequence ID Numbers 35 and 36. Numbers at the ends of the lines refer to nucleotide positions. Numbers below the sequence refer to amino acid positions.

FIG. 12 illustrates the assembly of a cDNA molecule encoding a PDGF α-receptor. Complementary DNA sequences are shown as lines. Only those portions of the vectors adjacent to the cDNA inserts are shown.

DETAILED DESCRIPTION OF THE INVENTION

Prior to setting forth the invention, it may be helpful to an understanding thereof to set forth definitions of certain terms to be used hereinafter.

DNA Construct

A DNA molecule, or a clone of such a molecule, either single- or double-stranded that has been modified through

human intervention to contain segments of DNA combined and juxtaposed in a manner that as a whole would not otherwise exist in nature.

DNA constructs contain the information necessary to direct the expression and/or secretion of DNA sequences encoding polypeptides of interest. DNA constructs will generally include promoters, enhancers and transcription terminators. DNA constructs containing the information necessary to direct the secretion of a polypeptide will also contain at least one secretory signal sequence.

Secretory Signal Sequence

A DNA sequence encoding a secretory peptide. A secretory peptide is an amino acid sequence that acts to direct the secretion of a mature polypeptide or protein from a cell. Secretory peptides are characterized by a core of hydrophobic amino acids and are typically (but not exclusively) found at the amino termini of newly sythesized proteins. Very often the secretory peptide is cleaved from the mature protein during secretion. Such secretory peptides contain processing sites that allow cleavage of the signal peptides from the mature proteins as it passes through the secretory pathway. Processing sites may be encoded within the signal peptide or may be added to the signal peptide by, for example, in vitro mutagenesis. Certain secretory peptides may be used in concert to direct the secretion of polypeptides and proteins. One such secretory peptide that may be used in combination with other secretory peptides is the third domain of the yeast Barrier protein.

Receptor Analog

A non-immunoglobulin polypeptide comprising a portion of a receptor which is capable of binding ligand and/or is recognized by anti-receptor antibodies. The amino acid sequence of the receptor analog may contain additions, substitutions or deletions as compared to the native receptor sequence. A receptor analog may be, for example, the ligand-binding domain of a receptor joined to another protein. Platelet-derived growth factor receptor (PDGF-R) analogs may, for example, comprise a portion of a PDGF receptor capable of binding anti-PDGF receptor antibodies, PDGF, PDGF isoforms, PDGF analogs, or PDGF antagonists.

Dimerizing Protein

A polypeptide chain having affinity for a second polypeptide chain, such that the two chains associate under physiological conditions to form a dimer. The second polypeptide chain may be the same or a different chain.

Biological Activity

A function or set of activities performed by a molecule in a biological context (i.e., in an organism or an in vitro facsimile thereof). biological activities may include the induction of extracellular matrix secretion from responsive cell lines, the induction of hormone secretion, the induction of chemotaxis, the induction of mitogenesis, the induction of differentiation, or the inhibition of cell division of responsive cells. A recombinant protein or peptide is considered to be biologically active if it exhibits one or more biological activities of its native counterpart.

Ligand

A molecule capable of being bound by the ligand-binding domain of a receptor or by a receptor analog. The molecule may be chemically synthesized or may occur in nature. Ligands may be grouped into agonists and antagonists. Agonists are those molecules whose binding to a receptor induces the response pathway within a cell. Antagonists are those molecules whose binding to a receptor blocks the response pathway within a cell.

Joined

Two or more DNA coding sequences are said to be joined when, as a result of in-frame fusions between the DNA coding sequences or as a result of the removal of intervening sequences by normal cellular processing, the DNA coding sequences are translated into a polypeptide fusion.

As noted above, the present invention provides methods for producing biologically active dimerized polypeptide fusions and secreted receptor analogs, which include, for example, PDGF receptor analogs. Secreted receptor analogs may be used to screen for new compounds that act as agonists or antagonists when interacting with cells containing membrane-bound receptors. In addition, the methods of the present invention provide dimerized non-immunoglobulin polypeptide fusions of therapeutic value that are biologically active only as diners. Moreover, the present invention provides methods of producing polypeptide dimers that are biologically active only as non-covalently associated diners. Secreted, biologically active dimers that may be produced using the resent invention include nerve growth factor, colony stimulating factor-1, factor XIII, and transforming growth factor β.

As used herein, the ligand-binding domain of a receptor is that portion of the receptor that is involved with binding the natural ligand. While not wishing to be bound by theory, the binding of a natural ligand to a receptor is believed to induce a conformational change which elicits a response to the change within the response pathway of the cell. For membrane-bound receptors, the ligand-binding domain is generally believed to comprise the extracellular domain for the receptor. The structure of receptors may be predicted from the primary translation products using the hydrophobicity plot function of, for example, P/C Gene or Intelligenetics Suite (Intelligenetics, Mt. View, Calif.) or may be predicted according to the methods described, for example, by Kyte and Doolittle, *J. Mol. Biol.* 157:105–132, 1982). The ligand-binding domain of the PDGF β-receptor, for example, has been predicted to include amino acids 29–531 of the published sequence (Gronwald et al., ibid.) The ligand-binding domain of the PDGF α-receptor has been predicted to include amino acids 25–500 of the published α-receptor sequence (Matsui et al., ibid.). As used herein, the ligand-binding domain of the PDGF β-receptor includes amino acids 29–441 of the sequence of FIGS. 1A–1D (Sequence ID Number 1) and C-terminal extensions up to and including amino acid 531 (FIG. 1B). The ligand-binding domain of the PDGF α-receptor is understood to include amino acids 24–524 of FIGS. 1A & 1B (Sequence ID Numbers 35 and 36).

Receptor analogs that may be used in the present invention include the ligand-binding domains of the epidermal growth factor receptor (EGF-R) and the insulin receptor. As used herein, a ligand-binding domain is that portion of the receptor that is involved in binding ligand and is generally a portion or essentially all of the extracellular domain that extends from the plasma membrane into the extracellular space. The ligand-binding domain of the EGF-R, for example, resides in the extracellular domain. EGF-R diners have been found to exhibit higher ligand-binding affinity

than EGF-R monomers (Boni-Schnetzler and Pilch, *Proc. Natl. Acad. Sci. USA* §A:7832–7836, 1987). The insulin receptor (Ullrich et al., *Nature* 313:756–761, 1985) requires dimerization for biological activity.

Another example of a receptor that may be secreted from a host cell is a platelet-derived growth factor receptor (PDGF-R). Two classes of PDGF-Rs, which recognized different isoforms of PDGF, have been identified. (PDGF is a disulfide-bonded, two-chain molecule, which is made up of an A chain and a B chain. These chains may be combined as AB heterodimers, AA homodimers or BB homodimers. These dimeric molecules are referred to herein as "isoforms".) The β-receptor (PDGFβ-R), which recognizes only the BB isoform of PDGF (PDGF-BB), has been described (Claesson-Welsh et al., *Mol. Cell, Biol.* 8:3476–3486, 1988; Gronwald et al., *Proc. Natl. Acad. Sci. USA* 85:3435–3439, 1988). The α-receptor (PDGFα-R), which recognizes all three PDGF isoforms (PDGF-AA, PDGF-AB and PDGF-BB), has been described by Matsui et al. (*Science* 243:800–804, 1989) and Kelly and Murray (pending commonly assigned U.S. patent application Ser. No. 07/355,018, which is incorporated herein by reference). The primary translation products of these receptors indicate that each includes an extracellular domain implicated in the ligand-binding process, a transmembrane domain, and a cytoplasmic domain containing a tyrosine kinase activity.

The present invention provides a standardized assay system, not previously available in the art, for determining the presence of PDGF, PDGF isoforms, PDGF agonists or PDGF antagonists using a secreted PDGF receptor analogs. Such an assay system will typically involve combining the secreted PDGF receptor analog with a biological sample under physiological conditions which permit the formation of receptor-ligand complexes, followed by detecting the presence of the receptor-ligand complexes. The term physiological conditions is meant to include those conditions found within the host organism and include, for example, the conditions of osmolarity, salinity and pH. Detection may be achieved through the use of a label attached to the PDGF receptor analog or through the use of a labeled antibody which is reactive with the receptor analog or the ligand. A wide variety of labels may be utilized, such as radionuclides, fluorophores, enzymes and luminescers. Receptor-ligand complexes may also be detected visually, i.e., in immunoprecipitation assays which do not require the use of a label. This assay system provides secreted PDGF receptor analogs that may be utilized in a variety of screening assays for, for example, screening for analogs of PDGF. The present invention also provides a methods for measuring the level of PDGF and PDGF isoforms in biological fluids.

As noted above, the present invention provides methods for producing dimerized polypeptide fusions that require dimerization for biological activity or enhancement of biological activity. Polypeptides requiring dimerization for biological activity include, in addition to certain receptors, nerve growth factor, colony-stimulating factor-1 (CSF-1), transforming growth factor β (TGF-β), PDGF and factor XIII. Nerve growth factor is a non-covalently linked dimer (Harper et al., *J. Biol. Chem.* 257: 8541–8548, 1982). CSF-1, which specifically stimulates the proliferation and differentiation of cells of mononuclear phagocytic lineage, is a disulfide-bonded homodimer (Rettenmier et al., *Mol. Cell. Biol.* 7: 2378–2387, 1987). TGF-β is biologically active as a disulfide-bonded dimer (Assoian et al., *J. Biol. Chem.* 258: 7155–7160, 1983). Factor XIII is a plasma protein that exists as a two chain homodin er n its activated form (Ichinose et al., *Biochem.* 25: 6900–6906, 1986). PDGF, as noted above, is a disulfide-bonded, two chain molecule (Murray et al., U.S. Pat. No. 4,766,073).

The present invention provides methods by which receptor analogs, including receptor analogs and PDGF-R analogs, requiring dimerization for activity may be secreted from host cells. The methods described herein are particularly advantageous in that they allow the production of large quantities of purified receptors. The receptors may be used in assays for the screening of potential ligands, in assays for binding studies, as imaging agents, and as agonists and antagonists within therapeutic agents.

A DNA sequence encoding a human PDGF receptor may be isolated as a cDNA using techniques known in the art (see, for example, Okayama and Berg, *Mol. Cell. Biol.* 2:161–170, 1982; *Mol. Cell. Biol.* 3: 280–289, 1983) from a library of human genomic or cDNA sequences. Such libraries may be prepared by standard procedures, such as those disclosed by Gubler and Hoffman (*Gene* 25: 263–269, 1983). It is preferred that the molecule is a CDNA molecule because CDNA lack introns and are therefore more suited to manipulation and expression in transfected or transformed cells. Sources of mRNA for use in the preparation of a CDNA library include the MG-63 human osteosarcoma cell line (available from ATCC under accession number CRL 1427), diploid human dermal fibroblasts and human embryo fibroblast and brain cells (Matsui et al., ibid.). A CDNA encoding a PDGFβ-R has been cloned from a diploid human dermal fibroblast cDNA library using oligonucleotide probes complementary to sequences of the mouse PDGF receptor (Gronwald et al., ibid.). A PDGFα-R cDNA has been isolated by Matsui et al. (ibid.) from human embryo fibroblast and brain cells. Alternatively, a cDNA encoding a PDGFα-R may be isolated from a library prepared from MG-63 human osteosarcoma cells using a cDNA probe containing sequences encoding the transmembrane and cytoplasmic domains of the PDGFβ-R (described by Kelly and Murray, ibid.). Partial CDNA clones (fragments) can be extended by re-screening of the library with the cloned CDNA fragment until the full sequence is obtained. In one embodiment, a ligand-binding domain of a PDGF receptor is encoded by the sequence of FIGS. 1A & 1B (Sequence ID Number 1) from amino acid 29 through amino acid 441. In another embodiment, a ligand-binding domain of a PDGF receptor is encoded by the sequence of FIGS. 1A & 1B (Sequence ID Number 1) from amino acid 29 through amino acid 531. In yet another embodiment, a ligand-binding domain of a PDGF receptor is encoded by the sequence of FIGS. 11A & 11B (Sequence ID Numbers 35 and 36) from amino acid 24 through amino acid 524. One skilled in the art may envision the use of a smaller DNA sequence encoding the ligand-binding domain of a PDGF receptor containing at least 400 amino acids of the extracellular domain.

DNA sequences encoding EGF-R (Ullrich et al., *Nature* 3: 418–425, 1984), the insulin receptor (Ullrich et al., *Nature* 313: 756–761, 1985), nerve growth factor (Ullrich et al. *Nature* 303: 821–825, 1983), colony stimulating factor-1 (Rettenmier et al., ibid.), transforming growth factor β (Derynck et al., *Nature* 3: 701–705, 1985), PDGF (Murray et al., ibid.), and factor XIII (Ichinose et al., ibid.) may also be used within the present invention.

To direct polypeptides requiring dimerization for biological activity or receptor analogs into the secretory pathway of the host cell, at least one secretory signal sequence is used in conjunction with the DNA sequence of interest. Preferred secretory signals include the alpha factor signal sequence (pre-pro sequence) (Kurjan and Herskowitz, *Cell* 30: 933–943, 1982; Kurjan et al., U.S. Pat. No. 4,546,082;

Brake, EP 116,201, 1983), the PHOS signal sequence (Beck et al., WO 86/00637), the BAR1 secretory signal sequence (MacKay et al., U.S. Pat. No. 4,613,572; MacKay, WO 87/002670), immunoglobulin V_H signal sequences (Loh et al., *Cell* 33: 85–93, 1983; Watson *Nuc, Acids, Res.* 12: 5145–5164, 1984) and immunoglobulin V_κ signal sequences (Watson, ibid.). Particularly preferred signal sequences are the SUC2 signal sequence (Carlson et al., *Mol. Cell. Biol.* 3: 439–447, 1983) and PDGF receptor signal sequences. Alternatively, secretory signal sequences may be synthesized according to the rules established, for example, by von Heinje (*Eur. J. Biochem.* 133: 17–21, 1983; *J. Mol. Biol,* 184: 99–105, 1985; *Nuc. Acids. Res.* 14: 4683–3690, 1986).

Secretory signal sequences may be used singly or may be combined. For example, a first secretory signal sequence may be used singly or combined with a sequence encoding the third domain of Barrier (described in co-pending commonly assigned U.S. patent application Ser. No. 07/104,316, which is incorporated by reference herein in its entirety). The third domain of Barrier may be positioned in proper reading frame 3' of the DNA sequence of interest or 5' to the DNA sequence and in proper reading frame with both the secretory signal sequence and the DNA sequence of interest.

In one embodiment of the present invention, a sequence encoding a dimerizing protein is joined to a sequence encoding a polypeptide chain of a polypeptide dimer or a receptor analog, and this fused sequence is joined in proper reading frame to a secretory signal sequence. As shown herein, the present invention utilizes such an arrangement to drive the association of the polypeptide or receptor analog to form a biologically active molecule upon secretion. Suitable dimerizing proteins include the *S. cerevisiae* repressible acid phosphatase (Mizunaga et al., *J. Biochem.* (Tokyo) 103: 321–326, 1988), the *S. cerevisiae* type 1 killer preprotoxin (Sturley et al., *EMBO J.* 5: 3381–3390, 1986), the *S. calsbergensis* alpha galactosidase melibiase (Sumner-Smith et al., *Gene* 36: 333–340, 1985), the *S. cerevisiae* invertase (Carlson et al., *Mol. Cell. Biol.* 3: 39–447, 1983), the *Neurospora crassa* ornithine decarboxylase (Digangi et al., *J. Biol. Chem.* 6: 7889–7893, 1987), immunoglobulin heavy chain hinge regions (Takahashi et al., *Cell* 29: 671–679, 1982), and other dimerizing immunoglobulin sequences. In a preferred embodiment, *S. cerevisiae* invertase is used to drive the association of polypeptides into dimers. Portions of dimerizing proteins, such as those mentioned above, may be used as dimerizing proteins where those portions are capable of associating as a dimer in a covalent or noncovalent manner. Such portions may be determined by, for example, altering a sequence encoding a dimerizing protein through In vitro mutagenesis to delete portions of the coding sequence. These deletion mutants may be expressed in the appropriate host to determine which portions retain the capability of associating as dimers. Portions of immunoglobulin gene sequences may be used to drive the association of non-immunoglobulin polypeptides. These portions correspond to discrete domains of immunoglobulins. Immunoglobulins comprise variable and constant regions, which in turn comprise discrete domains that show similarity in their three-dimensional conformations. These discrete domains correspond to immunoglobulin heavy chain constant region domain exons, immunoglobulin heavy chain variable region domain exons, immunoglobulin light chain varable region domain exons and immunoglobulin light chain constant region domain exons in immunoglobulin genes (Hood et al., in *Immunology,* The Benjamin/Cummings Publishing Company, Inc., Menlo Park, Calif.; Honjo et al., *Cell* 18: 559–568, 1979; Takahashi et al., *Cell* 29: 671–679, 1982; and Honjo, *Ann. Rev. Immun.* 1:499–528, 1983)). Particularly preferred portions of immunoglobulin heavy chains include Fab and Fab' fragments. (An Fab fragment is a portion of an immunoglobulin heavy chain that includes a heavy chain variable region domain and a heavy chain constant region omain. An Fab' fragment is a portion of an mmunoglobulin heavy chain that includes a heavy chain variable region domain, a heavy chain constant region domain and a heavy chain hinge region domain.)

It is preferred to use an immunoglobulin light chain constant region in association with at least one immunoglobulin heavy chain constant region domain. In a another embodiment, an immunoglobulin light chain constant region is associated with at least one immunoglobulin heavy chain constant region domain joined to an immunoglobulin hinge region. In one set of embodiments, an immunoglobulin light chain constant region joined in frame with a polypeptide chain of a non-immunoglobulin polypeptide dimer or receptor analog and is associated with at least one heavy chain constant region. In a preferred set of embodiments a variable region is joined upstream of and in proper reading frame with at least one immunoglobulin heavy chain constant region. In another set of embodiments, an immunoglobulin heavy chain is joined in frame with a polypeptide chain of a non-immunoglobulin polypeptide dimer or receptor analog and is associated with an immunoglobulin light chain constant region. In yet another set of embodiments, a polypeptide chain of a non-immunoglobulin polypeptide dimer or receptor analog is joined to at least one immunoglobulin heavy chain constant region which is joined to an immunoglobulin hinge region and is associated with an immunoglobulin light chain constant region. In a preferred set of embodiments an immunoglobulin variable region is joined upstream of and in proper reading frame with the immunoglobulin light chain constant region.

Immunoglobulin heavy chain constant region domains include C_H1, C_H2, C_H3, and C_HA of any class of immunoglobulin heavy chain including γ, α, ε, μ, and δ classes (Honjo, ibid., 1983) A particularly preferred immunoglobulin heavy chain constant region domain is human C_H1. Immunoglobulin variable regions include V_H, V_κ, or V_λ.

DNA sequences encoding immunoglobulins may be cloned from a variety of genomic or CDNA libraries known in the art The techniques for isolating such DNA sequences using probe-based methods are conventional techniques and are well known to those skilled in the art. Probes for isolating such DNA sequences may be based on published DNA sequences (see, for example, Hieter et al., *Cell* 2: 197–207, 1980). Alternatively, the polymerase chain reaction (PCR) method disclosed by Mullis et al (U.S. Pat. No. 4,683,195) and Mullis (U.S. Pat. No. 4,683,202), incorporated herein by reference may be used. The choice of library and selection of probes for the isolation of such DNA sequences is within the level of ordinary skill in the art.

Host cells for use in practicing the present invention include eukaryotic cells capable of being transformed or transfected with exogenous DNA and grown in culture, such as cultured mammalian and fungal cells. Fungal cells, including species of yeast (e.g., Saccharomyces spp., Schizosaccharomyces spp.), or filamentous fungi (e.g., Aspergillus spp., Neurospora spp.) may be used as host cells within the present invention. Strains of the yeast *Saccharomyces cerevisiae* are particularly preferred.

Expression units for use in the present invention will generally comprise the following elements, operably linked in a 5' to 3' orientation: a transcriptional promoter, a secre-

tory signal sequence a DNA sequence encoding nonimmunoglobulin polypeptide requiring dimerization for biological activity joined to a dimerizing protein and a transcriptional terminator. The selection of suitable promoters, signal sequences and terminators will be determined by the selected host cell and will be evident to one skilled in the art and are discussed more specifically below.

Suitable yeast vectors for use in the present invention include YRp7 (Struhl et al., *Proc. Natl. Acad. Sci. USA* 7: 1035–1039, 1978), YEp13 (Broach et al., *Gene* 8: 121–133, 1979), pJDB249 and pJDB219 (Beggs, *Nature* 275:104–108, 1978) and derivatives thereof. Such vectors will generally include a selectable marker, which may be one of any number of genes that exhibit a dominant phenotype for which a phenotypic assay exists to enable transformants to be selected. Preferred selectable markers are those that complement host cell auxotrophy, provide antibiotic resistance or enable a cell to utilize specific carbon sources, and include LEU2 (Broach et al. ibid.), URA3 (Botstein et al., *Gene p:* 17, 1979), HIS3 (Struhl et al., ibid.) or POT1 (Kawasaki and Bell, EP 171,142). Other suitable selectable markers include the CAT gene, which confers chloramphenicol resistance on yeast cells.

Preferred promoters for use in yeast include promoters from yeast glycolytic genes (Hitzeman et al., *J. Biol. Chem.* 255: 12073–12080, 1980; Alber and Kawasaki, *J. Mol. A Genet.* 1: 419–434, 1982; Kawasaki, U.S. Pat. No. 4,599, 311) or alcohol dehydrogenase genes (Young et al., in *Genetic Engineering of Microorganisms for Chemicals*, Hollaender et al., (eds.), p. 355, Plenum, New York, 1982; Ammerer, *Meth. Enzymol.* 101: 192–201, 1983). In this regard, particularly preferred promoters are the TPI1 promoter (Kawasaki, U.S. Pat. No. 4,599,311, 1986) and the ADH2–4^c promoter (Russell et al., *Nature* 304: 652–654, 1983 and Irani and Kilgore, described in pending, commonly assigned U.S. patent application Ser. No. 07/183,130, which is incorporated herein by reference). The expression units may also include a transcriptional terminator. A preferred transcriptional terminator is the TPI1 terminator (Alber and Kawasaki, ibid.).

In addition to yeast, proteins of the present invention can be expressed in filamentous fungi, for example, strains of the fungi Aspercillus (McKnight and Upshall, described in commonly assigned U.S. Pat. No. 4,935,349, which is incorporated herein by reference). Examples of useful promoters include those derived from *Aspergillus nidulans* glycolytic genes, such as the ADH3 promoter (McKnight et al., *EMBO J.* 4: 2093–2099, 1985) and the tpiA promoter. An example of a suitable terminator is the ADH3 terminator (McKnight et al., ibid.). The expression units utilizing such components are cloned into vectors that are capable of insertion into the chromosomal DNA of Aspergillus.

Techniques for transforming fungi are well known in the literature, and have been described, for instance, by Beggs (ibid.), Hinnen et al. (*Proc. Natl. Acad. Sci. USA* 75: 1929–1933, 1978), Yelton et al., (*Proc. Natl. Acad. Sci. USA* 81: 1740–1747, 1984), and Russell (*Nature* 301: 167–169, 1983). The genotype of the host cell will generally contain a genetic defect that is complemented by the selectable marker present on the expression vector. Choice of a particular host and selectable marker is well within the level of ordinary skill in the art.

In a preferred embodiment, a *Saccharomyces cerevisiae* host cell that contains a genetic deficiency in a gene required for asparagine-linked glycosylation of glycoproteins is used. Preferably, the *S. cerevisiae* host cell contains a gene a deficiency in the MNN9 gene (described in pending, commonly assigned U.S. patent application Ser. Nos. 116,095 and 189,547 which are incorporated by reference herein in their entirety). Most preferably, the *S. cerevisiae* host cell contains a disruption of the MNN9 gene. *S. cerevisiae* host cells having such defects may be prepared using standard techniques of mutation and selection. Ballou et al. (*J. Biol. Chem.* 255: 5986–5991, 1980) have described the isolation of mannoprotein biosynthesis mutants that are defective in genes which affect asparagine-linked glycosylation. Briefly, mutagenized *S. cerevisiae* cells were screened using fluoresceinated antibodies directed against the outer mannose chains present on wild-type yeast. Mutant cells that did not bind antibody were further characterized and were found to be defective in the addition of asparagine-linked oligosaccharide moieties. To optimize production of the heterologous proteins, it is preferred that the host strain carries a mutation, such as the S. cerevisiae pep4 mutation (Jones, *Genetics* 85: 23–33, 1977), which results in reduced proteolytic activity.

In addition to fungal cells, cultured mammalian cells may be used as host cells within the present invention. Preferred cell lines are rodent myeloma cell lines, which include p3X63Ag8 (ATCC TIB 9), OF (ATCC CRL 1646), NS-1 (ATCC TIB 18) and 210-RCY-Ag1 (Galfre et al., *Nature* 277: 131, 1979). A particularly preferred rodent myeloma cell line is SP2/0-Ag14 (ATCC CRL 1581). In addition, a number of other cell lines may be used within the present invention, including COS-1 (ATCC CRL 1650), BHK, p363.Ag.8.653 (ATCC CRL 1580) Rat Hep I (ATCC CRL 1600), Rat Hep II (ATCC CRL 1548), TCMK (ATCC CCL 139), Human lung (ATCC CCL 75.1), Human hepatoma (ATCC HTB-52), Hep G2 (ATCC HB 8065), Mouse liver (ATCC CC 29.1), 293 (ATCC CRL 1573; Graham et al., *J. Gen. Virol.* 36: 59–72, 1977) and DUKX cells (Urlaub and Chasin, *Proc. Patl. Acad. Sci USA* 77: 4216–4220, 1980) A preferred BHK cell line is the tk^{ts}13 BHK cell line (Waechter and Baserga, *Proc. Natl. Acad. Sci USA* 79: 1106–1110, 1982). A preferred BHK cell line is the tk-ts13 BHK cell line (Waechter and Baserga, *Proc. Natl. Acad. Sci. USA* 79: 1106–1110, 1982). A tk$^-$ BHK cell line is available from the American Type Culture Collection, Rockville, Md., under accession number CRL 1632. A particularly preferred tk$^-$ BHK cell line is BHK 570 which is available from the American Type Culture Collection under accession number 10314.

Mammalian expression vectors for use in carrying out the present invention will include a promoter capable of directing the transcription of a cloned gene or cDNA. referred promoters include viral promoters and cellular promoters. Preferred viral promoters include the major late promoter from adenovirus 2 (Kaufman and Sharp, *Mol. Cell. Biol.* 2: 1304–13199, 1982) and the SV40 promoter (Subramani et al., *Mol. Cell. Biol.* 1: 854–864, 1981). Preferred cellular promoters include the mouse metallothionein 1 promoter (Palmiter et al., *Science* 222: 809–814, 1983) and a mouse V_k promoter (Grant et al., *Nuc. Acids Res.* 11: 5496, 1987). A particularly preferred promoter is a mouse V_H promoter (Loh et al., ibid.). Such expression vectors may also contain a set of RNA splice sites located downstream from the promoter and upstream from the DNA sequence encoding the peptide or protein of interest. Preferred RNA splice sites may be obtained from adenovirus and/or immunoglobulin genes. Also contained in the expression vectors is a polyadenylation signal located downstream of the coding sequence of interest. Polyadenylation signals include the early or late polyadenylation signals from SV40 (Kaufman

and Sharp, ibid.), the polyadenylation signal from the adenovirus 5 E1B region and the human growth hormone gene terminator (DeNoto et al., *Nuc. Acids Res.* 2: 3719–3730, 1981). A particularly preferred polyadenylation signal is the V_H gene terminator (Loh et al., ibid.). The expression vectors may include a noncoding viral leader sequence, such as the adenovirus 2 tripartite leader, located between the promoter and the RNA splice sites. Preferred vectors may also include enhancer sequences, such as the SV40 enhancer and the mouse μ enhancer (Gillies, *Cell* 33: 717–728, 1983). Expression vectors may also include sequences encoding the adenovirus VA RNAs.

Cloned DNA sequences may be introduced into cultured mammalian cells by, for example, calcium phosphate-mediated transfection (Wigler et al., *Cell* 14: 725, 1978; Corsaro and Pearson, *Somatic Cell Genetics* 7: 603, 1981; Graham and Van der Eb, *Virology* 52: 456, 1973.) Other techniques for introducing cloned DNA sequences into mammalian cells, such as electroporation (Neumann et al., *EMBO J.* 1: 841–845, 1982), may also be used. In order to identify cells that have integrated the cloned DNA, a selectable marker is generally introduced into the cells along with the gene or cDNA of interest. Preferred selectable markers for use in cultured mammalian cells include genes that confer resistance to drugs, such as neomycin, hygromycin, and methotrexate. The selectable marker may be an amplifiable selectable marker. A preferred amplifiable selectable marker is the DHFR gene. A particularly preferred amplifiable marker is the DHFR" CDNA (Simonsen and Levinson, *Proc. Natl. Adac. Sci. USA* 80: 2495–2499, 1983). Selectable markers are reviewed by Thilly (*Mammalian Cell Technology*, Butterworth Publishers, Stoneham, Mass.) and the choice of selectable markers is well within the level of ordinary skill in the art.

Selectable markers may be introduced into the cell on a separate plasmid at the same time as the gene of interest, or they may be introduced on the same plasmid. If on the same plasmid, the selectable marker and the gene of interest may be under the control of different promoters or the same promoter, the latter arrangement producing a dicistronic message. Constructs of this type are known in the art (for example, Levinson and Simonsen, U.S. Pat. No. 4,713,339). It may also be advantageous to add additional DNA, known as "carrier DNA" to the mixture which is introduced into the cells.

Transfected mammalian cells are allowed to grow for a period of time, typically 1–2 days, to begin expressing the DNA sequence(s) of interest. Drug selection is then applied to select for growth of cells that are expressing the selectable marker in a stable fashion. For cells that have been transfected with an amplifiable selectable marker the drug concentration may be increased in a stepwise manner to select for increased copy number of the cloned sequences, thereby increasing expression levels.

Host cells containing DNA constructs of the resent invention are grown in an appropriate growth medium. As used herein, the term "appropriate growth medium" means a medium containing nutrients required for the growth of cells. Nutrients required for cell growth may include a carbon source, a nitrogen source, essential amino acids, vitamins, minerals and growth factors. The growth medium will generally select for cells containing the DNA construct by, for example, drug selection or deficiency in an essential nutrient which are complemented by the selectable marker on the DNA construct or co-transfected with the DNA construct. Yeast cells, for example, are preferably grown in a chemically defined medium, comprising a non-amino acid nitrogen source, inorganic salts, vitamins and essential amino acid supplements. The pH of the medium is preferably maintained at a pH greater than 2 and less than 8, preferably at pH 6.5. Methods for maintaining a stable pH include buffering and constant pH control, preferably through the addition of sodium hydroxide. Preferred buffering agents include succinic acid and Bis-Tris (Sigma Chemical Co., St. Louis, Mo.). Yeast cells having a defect in a gene required for asparagine-linked glycosylation are preferably grown in a medium containing an osmotic stabilizer. A preferred osmotic stabilizer is sorbitol supplemented into the medium at a concentration between 0.1 M and 1.5 M., preferably at 0.5 M or 1.0 M. Cultured mammalian cells are generally grown in commercially available serum-containing or serum-free media. Selection of a medium appropriate for the particular cell line used is within the level of ordinary skill in the art.

The culture medium from appropriately grown transformed or transfected host cells is separated from the cell material, and the presence of dimerized polypeptide fusions or secreted receptor analogs is demonstrated. A preferred method of detecting receptor analogs, for example, is by the binding of the receptors or portions of receptors to a receptor-specific antibody. An anti-receptor antibody may be a monoclonal or polyclonal antibody raised against the receptor in question, for example, an anti-PDGF receptor monoclonal antibody may be used to assay for the presence of PDGF receptor analogs. Another antibody, which may be used for detecting substance P-tagged peptides and proteins, is a commercially available rat anti-substance P monoclonal antibody which may be utilized to visualize peptides or proteins that are fused to the C-terminal amino acids of substance P. Ligand binding assays may also be used to detect the presence of receptor analogs. In the case of PDGF receptor analogs, it is preferable to use fetal foreskin fibroblasts, which express PDGF receptors, to compete against the PDGF receptor analogs of the present invention for labeled PDGF and PDGF isoforms.

Assays for detection of secreted, biologically active peptide dimers and receptor analogs may include Western transfer, protein blot or colony filter. A Western transfer filter may be prepared using the method described by Towbin et al. (*Proc. Natl. Acad. Sci. USA* 76: 4350–4354, 1979). Briefly, samples are electrophoresed in a sodium dodecylsulfate polyacrylamide gel. The proteins in the gel are electrophoretically transferred to nitrocellulose paper. Protein blot filters may be prepared by filtering supernatant samples or concentrates through nitrocellulose filters using, for example, a Minifold (Schleicher & Schuell, Keene, N.H.). Colony filters may be prepared by growing colonies on a nitrocellulose filter that has been laid across an appropriate growth medium. In this method, a solid medium is preferred. The cells are allowed to grow on the filters for at least 12 hours. The cells are removed from the filters by washing with an appropriate buffer that does not remove the proteins bound to the filters. A preferred buffer comprises 25 mM Tris-base, 19 mM glycine, pH 8.3, 20% methanol.

The dimerized polypeptide fusions and receptor analogs present on the Western transfer, protein blot or colony filters may be visualized by specific antibody binding using methods known in the art. For example, Towbin et al. (ibid.) describe the visualization of proteins immobilized on nitrocellulose filters with a specific antibody followed by a labeled second antibody, directed against the first antibody. Kits and reagents required for visualization are commercially available, for example, from Vector Laboratories, (Burlingame, Calif.), and Sigma Chemical Company (St. Louis, Mo.).

Secreted, biologically active dimerized polypeptide fusions and receptor analogs may be isolated from the medium of host cells grown under conditions that allow the secretion of the biologically active dimerized polypeptide fusions and receptor analogs. The cell material is removed from the culture medium, and the biologically active dimerized polypeptide fusions and receptor analogs are isolated using isolation techniques known in the art. Suitable isolation techniques include precipitation and fractionation by a variety of chromatographic methods, including gel filtration, ion exchange chromatography and immunoaffinity chromatography. A particularly preferred purification method is immunoaffinity chromatography using an antibody directed against the receptor analog or dimerized polypeptide fusion. The antibody is preferably immobilized or attached to a solid support or substrate. A particularly preferred substrate is CNBr-activated Sepharose (Pharmacia LKB Technologies, Inc., Piscataway, N.J.). By this method, the medium is combined with the antibody/substrate under conditions that will allow binding to occur. The complex may be washed to remove unbound material, and the receptor analog or peptide dimer is released or eluted through the use of conditions unfavorable to complex formation. Particularly useful methods of elution include changes in pH, wherein the immobilized antibody has a high affinity for the ligand at a first pH and a reduced affinity at a second (higher or lower) pH; changes in concentration of certain chaotropic agents; or through the use of detergents.

The secreted PDGF receptor analogs of the present invention can be used within a variety of assays for detecting the presence of and/or screening for native PDGF, PDGF isoforms or PDGF-like molecules. These assays will typically involve combining PDGF receptor analogs, which may be bound to a solid substrate such as polymeric microtiter plate wells, with a biological sample under conditions that permit the formation of receptor/ligand complexes. Screening assays for the detection of PDGF, PDGF isoforms or PDGF-like molecules will typically involve combining soluble PDGF receptor analogs with a biological sample and incubating the mixture with a PDGF isoform or mixture of PDGF isoforms bound to a solid substrate such as polymeric microtiter plates, under conditions that permit the formation of receptor/ligand complexes. Detection may be achieved through the use of a label attached to the receptor or through the use of a labeled antibody which is reactive with the receptor. Alternatively, the labeled antibody may be reactive with the ligand. A wide variety of labels may be utilized, such as radionuclides, fluorophores, enzymes and luminescers. Complexes may also be detected visually, i.e., in immunoprecipitation assays, which do not require the use of a label.

Secreted PDGF receptor analogs of the present invention may also be labeled with a radioisotope or other imaging agent and used for in vivo diagnostic purposes. Preferred radioisotope imaging agents include iodine-125 and technetium-99, with technetium-99 being particularly preferred. Methods for producing protein-isotope conjugates are well known in the art, and are described by, for example, Leikeman et al. (U.S. Pat. No. 4,652,440), Parker et al. (WO 87/05030) and Wilber et al. (EP 203,764). Alternatively, the secreted receptor analogs may be bound to spin label enhancers and used for magnetic resonance (MR) imaging. Suitable spin label enhancers include stable, sterically hindered, free radical compounds such as nitroxides. Methods for labeling ligands for MR imaging are disclosed by, for example, Goffman et al. (U.S. Pat. No. 1,656,026). For administration, the labeled receptor analogs are combined with a pharmaceutically acceptable carrier or diluent, such as sterile saline or sterile water. Administration is preferably by bolus injection, preferably intravenously. These imaging agents are particularly useful in identifying the locations of atherosclerotic plaques, PDGF-producing tumors, and receptor-bound PDGF.

The secreted PDGF receptor analogs of the present invention may also be utilized within diagnostic kits. Briefly, the subject receptor analogs are preferably provided in a lyophilized form or immobilized onto the walls of a suitable container, either alone or in conjunction with antibodies capable of binding to native PDGF or selected PDGF isoform(s) or specific ligands. The antibodies, which may be conjugated to a label or unconjugated, are generally included in the kits with suitable buffers, such as phosphate, stabilizers, inert proteins or the like. Generally, these materials are present in less than about 5% weight based upon the amount of active receptor analog, and are usually present in an amount of at least about 0.001% weight. It may also be desirable to include an inert excipient to dilute the active ingredients. Such an excipient may be present from about 1% to 99% weight of the total composition. In addition, the kits will typically include other standard reagents, instructions and, depending upon the nature of the label involved, reactants that are required to produce a detectable product. Where an antibody capable of binding to the receptor or receptor/ligand complex is employed, this antibody will usually be provided in a separate vial. The antibody is typically conjugated to a label and formulated in an analogous manner with the formulations briefly described above. The diagnostic kits, including the containers, may be produced and packaged using conventional kit manufacturing procedures.

As noted above, the secreted PDGF receptor analogs of the present invention may be utilized within methods for purifying PDGF from a variety of samples. Within a preferred method, the secreted PDGF receptor analogs are immobilized or attached to a substrate or support material, such as polymeric tubes, beads, polysaccharide particulates, polysaccharide-containing materials, polyacrylamide or other water insoluble polymeric materials. Methods for immobilization are well known in the art (Mosbach et al., U.S. Pat. No. 4,415,665; Clarke et al., *Meth. Enzymology* : 436-442, 1979). A common method of immobilization is CNBr activation. Activated substrates are commercially available from a number of suppliers, including Pharmacia (Piscataway, N.J.), Pierce chemical Co. (Rockford, Ill.) and Bio-Rad Laboratories (Richmond, Calif.). A preferred substrate is CNBr-activated Sepharose (Pharmacia, Piscataway, N.J.). Generally, the substrate/receptor analog complex will be in the form of a column. The sample is then combined with the immobilized receptor analog under conditions that allow binding to occur. The substrate with immobilized receptor analog is first equilibrated with a buffer solution of a composition in which the receptor analog has been previously found to bind its ligand. The sample, in the solution used for equilibration, is then applied to the substrate/receptor analog complex. Where the complex is in the form of a column, it is preferred that the sample be passed over the column two or more times to permit full binding of ligand to receptor analog. The complex is then washed with the same solution to elute unbound material. In addition, a second wash solution may be used to minimize nonspecific binding. The bound material may then be released or eluted through the use of conditions unfavorable to complex formation. Particularly useful methods include changes in pH, wherein the immobilized receptor has a high affinity for

PDGF at a first pH and reduced affinity at a second (higher or lower) pH; changes in concentration of certain chaotropic agents; or through the use of detergents.

The secreted PDGF receptor analogs fused to dimerizing proteins of the present invention may be used in pharmaceutical compositions for topical or intravenous application. The secreted PDGF receptor analogs of the present invention may be useful in the treatment of atherosclerosis by, for example, binding endogenous PDGF to prevent smooth muscle cell proliferation. The PDGF receptor analogs fused to dimerizing proteins are used in combination with a physiologically acceptable carrier or diluent. Preferred carriers and diluents include saline and sterile water. Pharmaceutical compositions may also contain stabilizers and adjuvants. The resulting aqueous solutions may be packaged for use or filtered under aseptic conditions and lyophilized, the lyophilized preparation being combined with a sterile aqueous solution prior to administration.

The following examples are offered by way of illustration and not by way of limitation.

EXAMPLES

Enzymes, including restriction enzymes, DNA polymerase I (Klenow fragment), T4 DNA polymerase, T4 DNA ligase and T4 polynucleotide kinase, were obtained from New England Biolabs (Beverly, Mass.), GIBCO-BRL (Gaithersburg, Md.) and Boerhinger-Mannheim Biochemicals (Indianapolis, Ind.) and were used as directed by the manufacturer or as described in Maniatis et al. (*Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, NY, 1982) and Sambrook et al. (*Molecular Cloning: A Laboratory Manual*/Second Edition, Cold Spring Harbor Laboratory, NY, 1989).

Example 1

Cloning PDGF Receptor cDNAs

A. Cloning the PDGF β-Receptor

A cDNA encoding the PDGF β-receptor was cloned as follows. Complementary DNA (cDNA) libraries were prepared from poly(A) RNA from diploid human dermal fibroblast cells, prepared by explant from a normal adult, essentially as described by Hagen et al. (*Proc. Natl. Acad. Sci. USA* 83: 2412–2416, 1986). Briefly, the poly(A) RNA was primed with oligo d(T) and cloned into λgt10 using Eco RI linkers. The random primed library was screened for the presence of human PDGF receptor cDNA's using three oligonucleotide probes complementary to sequences of the mouse PDGF receptor (Yarden et al., *Nature* 323: 226–232, 1986). Approximately one million phage from the random primed human fibroblast cell library were screened using oligonucleotides ZC904, ZC905 and ZC906 (Table 1; Sequence ID Numbers 5, 6 and 7, respectively). Eight positive clones were identified and plaque purified. Two clones, designated RP41 and RP51, were selected for further analysis by restriction enzyme mapping and DNA sequence analysis. RP51 was found to contain 356 bp of 5'-noncoding sequence, the ATG translation initiation codon and 738 bp of the amino terminal coding sequence. RP41 was found to overlap clone RP51 and contained 2649 bp encoding amino acids 43-925 of the β-receptor protein.

TABLE 1

Oligonucleotide Sequences

ZC871 (Sequence ID Number 3)
5' CTC TCT TCC TCA GGT AAA TGA GTG CCA GGG CCG GCA AGC CCC CGC TCC 3'

ZC872 (Sequence ID Number 4)
5' CCG GGG AGC GGC GGC TTG CCG GCC CTG GCA CTC ATT TAC CTG AGG AAG AGA CAG CT 3'

ZC904 (Sequence ID Number 5)
5' CAT GGG CAC GTA ATC TAT AGA TTC ATC CTT GCT CAT ATC CAT GTA 3'

ZC905 (Sequence ID Number 6)
5' TCT TGC CAG GGC ACC TGG GAC ATC TGT TCC CAC ATC ACC GG 3'

ZC906 (Sequence ID Number 7)
5' AAG CTG TCC TCT GCT TCA GCC AGA GGT CCT GGG CAG CC 3'

ZC1380 (Sequence ID Number 8)
5' CAT GGT GGA ATT CCT GCT GAT 3'

ZC1447 (Sequence ID Number 9)
5' TG GTT GTG CAG AGC TGA GGA AGA GAT GGA 3'

ZC1453 (Sequence ID Number 10)
5' AAT TCA TTA TGT TGT TGC AAG CCT TCT TGT TCC TGC TAG CTG GTT TCG CTG TTA A 3'

ZC1454 (Sequence ID Number 11)
5' GAT CTT AAC AGC GAA ACC AGC TAG CAG GAA CAA GAA GGC TTG CAA CAA CAT AAT G 3'

ZC1478 (Sequence ID Number 12)
5' ATC GCG AGC ATG CAG ATC TGA 3'

ZC1479 (Sequence ID Number 13)
5' AGC TTC AGA TCT GCA TGC TGC CGA T 3'

ZC1776 (Sequence ID Number 14)
5' AGC TGA GCG CAA ATG TTG TGT CGA GTG CCC ACC GTG CCC AGC TTA GAA TTC T 3'

ZC1777 (Sequence ID Number 15)
5' CTA GAG AAT TCT AAG CTG GGC ACG GTG GGC ACT CGA CAC AAC ATT TGC GCT C 3'

ZC1846 (Sequence ID Number 16)
5' GAT CGG CCA CTG TCG GTG CGC TGC ACG CTG CGC AAC GCT GTG GGC CAG GAC ACG CAG GAG GTC ATC GTG GTG CCA CAC TCC TTG CCC TTT AAG CA 3'

ZC1847 (Sequence ID Number 17)
5' AGC TTG CTT AAA GGG CAA GGA GTG TGG CAC CAC GAT GAC CTC CTG CGT GTC CTG GCC CAC AGC GTT GCG CAG CGT GCA GCG CAC CGA CAG TGG CC 3'

ZC1886 (Sequence ID Number 18)
5' CCA GTG CCA AGC TTG TCT AGA CTT ACC TTT AAA GGG CAA GGA G 3'

ZC1892 (Sequence ID Number 19)
5' AGC TTG AGC GT 3'

ZC1893 (Sequence ID Number 20)
5' CTA GAC GCT CA 3'

ZC1894 (Sequence ID Number 21)
5' AGC TTC CAG TTC TTC GCC GTC ATG TCA GTT CTT CGC CGT CAT GTG AT 3'

ZC1895 (Sequence ID Number 22)
5' CTA GAT CAC ATG ACG GCG AAG AAC TGA CAT GAG GCC GAA GAA CTG GA 3'

ZC1 1 (Sequence ID Number 23)

TABLE 1-continued

Oligonucleotide Sequences

5' AAT TCG GAT CCA CCA TGG GCA CCA GCC ACC CGG CGT TCC TGG TGT TAG GCT GCC TGC TGA CCG GCC 3'

ZC2182 (Sequence ID Number 24)
5' TGA GCC TGA TCC TGT GCC AAC TGA GCC TGC CAT CGA TCC TGC CAA ACG AGA ACG AGA AGG TTG TGC AGC TA 3'

ZC2183 (Sequence ID Number 25)
5' AAT TTA GCT GCA CAA CCT TCT CGT TCT CGT TTG GCA GGA TGG ATG GCA GGC TCA GTT GGC ACA GGA TCA 3'

ZC2184 (Sequence ID Number 26)
5' GGC TCA GGC CGG TCA GCA GGC AGC CTA ACA CCA GGA ACG CCG GGT GGC TGG TGC CCA TGG TGG ATC CG 3'

ZC2311 (Sequence ID Number 27)
5' TGA TCA CCA TGG CTC AAC TG 3'

ZC2351 (Sequence ID Number 28)
5' CGA ATT CCA C 3'

ZC2352 (Sequence ID Number 29)
5' CAT GGT GGA ATT CGA GCT 3'

ZC2392 (Sequence ID Number 30)
5' ACC TAA GCT TGT CTA GAC TTA CCT TCA CAA CGC AGG GTG GG 3'

The 3'-end of the CDNA was not isolated in the first cloning and was subsequently isolated by screening 6×10^5 phage of the oligo d(T)-primed CDNA library with a 630 bp Sst I-Eco RI fragment derived from the 3'-end of clone RP41. One isolate, designated TO91, was further analyzed by restriction enzyme mapping and DNA sequencing. This clone was found to comprise the 3'-end of the receptor coding region and 1986 bp of 3' untranslated sequence.

Clones RP51, RP41 and TO91 were ligated together to construct a full-length CDNA encoding the entire PDGF β-receptor. RP41 was digested with Ace I and Bam HI to isolate the 2.12 kb fragment. RP51 was digested with Eco RI and Ace I to isolate the 982 bp fragment. The 2.12 kb RP41 fragment and the 982 bp RP51 fragment were joined in a three-part ligation with pUC13, which had been linearized by digestion with Eco RI and Bam HI. The resultant plasmid was designated 51/41. Plasmid 51/41 was digested with Eco RI and Bam HI to isolate the 3 kb fragment comprising the partial PDGF receptor cDNA. OT91 was digested with Bam HI and Xba I to isolate the 1.4 kb fragment containing the 3' portion of the PDGF receptor cDNA. The Eco RI-Bam HI 51/41 fragment, the Bam HI-Xba I TO91 fragment and the Eco RI-Xba I digested pUC13 were joined in a three-part ligation. The resultant plasmid was designated pR-RX1 (FIG. 2).

B. Cloning the PDGF-α Receptor

A cDNA encoding to PDGF α-receptor was cloned as follows. RNA was prepared by the method of Chirgwin et al. (*Biochemistry* 18: 5294, 1979) and twice purified on oligo d(T) cellulose to yield poly(A)+RNA. Complementary DNA was prepared in lgt10 phage using a kit purchased from Invitrogen (San Diego, CA). The resulting λ phage DNA was packaged with a coat particle mixture from Stratagene Cloning Systems (La Jolla, Calif.), infected into *E. coli* strain C600 Hfl⁻ and titered.

Approximately 1.4×10^6 phage recombinants were plated to produce plaques for screening. Nitrocellulose filter lifts were prepared according to standard techniques and were hybridized to a radiolabeled PDGF β-receptor cDNA fragment (Gronwald et al., ibid.) comprising the 1.9 kb Hsp I-Hind III fragment that encodes the transmembrane and cytoplasmic domains of the PDGF β-receptor cDNA. Hybridization was performed for 36 hours at 42° C. in a mixture containing 40% formamide, 5×SSCP (SSC containing 25 mM phosphate buffer, pH 6.5), 200 μg/ml denatured salmon sperm DNA, 3×Denhardt's, and 10% dextran sulfate. Following hybridization, the filters were washed extensively at room temperature in 2×SSC, then for 15 minutes at 47–48° C. Following an exposure to X-ray film, the filters were treated to increasingly stringent wash conditions followed by film recording until a final wash with 0.1×SSC at 65° C. was reached. Film analysis showed that a "family" of plaques that hybridized at lower wash stringency but not at the highest stringency. This "family" was selected for further analysis.

Two λ phage clones from the "family" obtained from the initial screening were subcloned into the Not I site of the pUCtype plasmid vector pBluescript SK⁺ (obtained from Stratagene Cloning Systems, La Jolla, Calif.) and were analyzed by restriction mapping and sequence analysis. Restriction enzyme analysis of a phage clone, designated α1-1, revealed a restriction fragment pattern dissimilar from that of the PDGF β-receptor CDNA with the exception of a common Bgl II-Bgl II band of approximately 160 bp. The PDGF β-receptor cDNA contains two similarly spaced Bgl II sites within the region coding for the second tyrosine kinase domain.

Restriction analysis of a second plasmid subclone (designated α1-7) revealed an overlap of the 5' approximately 1.2 kb of clone α1-1, and an additional approximately 2.2 kb of sequence extending in the 5' direction. Sequence analysis revealed that the 3' end of this clone encodes the second tyrosine kinase domain, which contains regions of near sequence identity to the corresponding regions in the PDGF β-receptor. The 5' end of clone α1-7 contained non-receptor sequences. Two additional β-receptor clones were obtained by probing with α1-1 sequences. Clone α1-1 was digested with Not I and Spe I, and a 230 bp fragment was recovered. Clone α1-1 was also digested with Bam HI and Not I, and a 550 bp fragment was recovered. A clone that hybridized to the 230 bp probe was designated α5-1. This clone contained the 5'-most coding sequence for the PDGF α-receptor. Another clone, designated α6-3, hybridized to the 550 bp probe and was found to contain 3' coding and non-coding sequences, including the poly(A) tail.

Clone α1-1 was radiolabeled (^{32}p) and used to probe a northern blot (Thomas, *Methods Enzymol*, 100: 225–265, 1983) of the MG-63 poly(A)+RNA used to prepare the cDNA library. A single band of approximately 6.6 kb was observed. RNA prepared from receptor-positive cell lines including the human fibroblast SK4, WI-38 and 7573 cell lines; the mouse fibroblast line DI 3T3; the U2-OS human osteosarcoma cell line and baboon aortic smooth muscle cells, and RNA prepared from receptor-negative lines including A431 (an epithelial cell line) and VA 13 (SV40-transformed WI-38 cells) were probed by northern format with the α1-1 cDNA. In all cases, the amount of the 6.6 kb band detected in these RNA correlated well with the relative levels of α-receptor detected on the respective cell surfaces. The 6.6 kb RNA was not detected in RNA prepared from any tested cell line of hematopoietic origin, in agreement with a lack of PDGF α-receptor protein detected on these cell types.

Clones α1-1 and and α1-7 were joined at a unique Pst I site in the region encoding the transmembrane portion of the receptor. Clone α1-1 was digested with Xba I and Pst I and

the receptor sequence fragment was recovered. Clone α1-7 was digested with Pst I and Bam HI and the receptor fragment was recovered. The two fragments were ligated with Xba I+Bam HI-digested pIC19R (Marsh et al. *Gene* 32: 481–486, 1984) to construct plasmid pα17R (FIG. 12).

The remainder of the 5==-most α-receptor sequence was obtained from clone α5-1 as an Sst I-Cla I fragment. This fragment was joined to the Eco RI-Sst I receptor fragment of pα17R and cloned into Eco RI+Cla I-digested pBluescript SK+plasmid to construct plasmid pα17B (FIG. 12). FIGS. 11A–11D (Sequence ID Numbers 35 and 36) shows the nucleotide sequence and deduced amino acid sequence of the cDNA present in pα17B.

Example 2

Construction of a SUC2 Signal Sequence-PDGF β-Receptor Fusion

To direct the PDGF β-receptor into the yeast secretory pathway, the PDGF β-receptor cDNA was joined to a sequence encoding the *Saccharomyces cerevisiae* SUC2 signal sequence. Oligonucleotides ZC1453 and ZC1454 (Sequence ID Numbers 10 and 11; Table 1) were designed to form an adapter encoding the SUC2 secretory signal flanked by a 5' Eco RI adhesive end and a 3' Bgl II adhesive end. ZC1453 and ZC1454 were annealed under conditions described by Maniatis et al. (ibid.). Plasmid pR-RX1 was digested with Bgl II and Sst II to isolate the 1.7 kb fragment comprising the PDGF β-receptor coding sequence from amino acids 28 to 596. Plasmid pR-RX1 was also cut with Sst II and Hind III to isolate the 1.7 kb fragment comprising the coding sequence from amino acids 597 through the translation termination codon and 124 bp of 3' untranslated DNA. The two 1.7 kb pR-RX1 fragments and the ZC1453/ZC1454 adapter were joined with pUC19, which had been linearized by digestion with Eco RI and Hind III. The resultant plasmid, comprising the SUC2 signal sequence fused in-frame with the PDGF β-receptor cDNA, was designated pBTL10 (FIG. 2).

Example 3

Construction of pCBS22

The BAR1 gene product, Barrier, is an exported protein that has been shown to have three domains. When used in conjunction with a first signal sequence, the third domain of Barrier protein has been shown to aid in the secretion of proteins into the medium (MacKay et al., U.S. patent application Ser. No. 104,316).

The portion of the BAR1 gene encoding the third domain of Barrier was joined to a sequence encoding the C-terminal portion of substance P (subP; Munro and Pelham, *EMBO J.* 3: 3087–3093, 1984). The presence of the substance P amino acids on the terminus of the fusion protein allowed the protein to be detected using commercially available anti-substance P antibodies. Plasmid pZV9 (deposited as a transformant in *E. coli* strain RR1, ATCC accession no. 53283), comprising the entire BAR2 coding region and its associated flanking regions, was cut with Sal I and Bam HI to isolate the 1.3 kb BAR1 fragment. This fragment was subcloned into pUC13, which had been cut with Sal I and Bam HI, to generate the plasmid designated pZV17. Plasmid pZV17 was digested with Eco RI to remove the 3'-most 0.5 kb of the BAR1 coding region. The vector-BAR1 fragment was religated to create the plasmid designated pJH66 (FIG. 3). Plasmid pJH66 was linearized with Eco RI and blunt-ended with DNA polymerase I (Klenow fragment). Kinased Bam HI linkers (5'CCG GAT CCG G 3') were added and excess linkers were removed by digestion with Bam HI before religation. The resultant plasmid was designated pSW8 (FIG. 3).

Plasmid pSW81, comprising the TPI1 promoter, the BAR1 coding region fused to the coding region of the C-terminal portion of substance P (Munro and Pelham, *EMBO J.* 3: 3087–3093, 1984) and the TPI1 terminator, was derived from pSW8. Plasmid pSW8 was cut with Sal I and Bam HI to isolate the 824 bp fragment encoding amino acids 252 through 526 of BAR1. Plasmid pPM2, containing the synthetic oligonucleotide sequence encoding the dimer form of the C-terminal portion of substance P (subP) in M13mp8, was obtained from Hugh Pelham (MRC Laboratory of Molecular Biology, Cambridge, England). Plasmid pPM2 was linearized by digestion with Bam HI and Sal I and ligated with the 824 bp BAR1 fragment from pSW8. The resultant plasmid, pSW14, was digested with Sal I and Sma I to isolate the 871 bp BAR1-substance P fragment. Plasmid pSW16, comprising a fragment of BAR1 encoding amino acids 1 through 250, was cut with Xba I and Sal I to isolate the 767 bp BAR1 fragment. This fragment was ligated with the 871 bp BAR1-substance P fragment in a three-part ligation with pUC18 cut with Xba I and Sma I. The resultant plasmid, designated pSW15, was digested with Xba I and Sma I to isolate the 1.64 kb BAR1-substance P fragment. The ADH1 promoter was obtained from pRI029. Plasmid pRI029, comprising the ADH1 promoter and the BAR1 5' region encoding amino acids 1 to 33 in pUC18, -was digested with Sph I and Xba I to isolate the 0.42 kb ADH1 promoter fragment. The TPI1 terminator (Albet and Kawasaki, ibid.) was provided as a linearized fragment containing the TPI1 terminator and pUC18 with a Klenow-filled Xba I end and an Sph I end. This fragment was ligated with the 0.42 kb ADH1 promoter fragment and the 1.64 kb BAR1-substance P fragment in a three-part ligation to produce plasmid pSW22.

The ADH1 promoter and the coding region of BAR1, from the translation initiation ATG through the Eco RV site present in pSW22, were removed by digestion with Hind III and Eco RV. The 3.9 kb vector fragment, comprising the 401 bp between the Eco RV and the Eco RI sites of the BAR1 gene fused to subp and the TPI1 terminator, was isolated by gel electrophoresis. Oligonucleotide ZC1478 (Sequence ID Number 12; Table 1) was kinased and annealed with oligonucleotide ZC1479 (Sequence ID Number 13; Table 1) using conditions described by Maniatis et al. (ibid.). The annealed oligonucleotides formed an adapter comprising a Hind III adhesive end and a polylinker encoding Bgl II, Sph I, Nru I and Eco RV restriction sites. The ZC1479/ZC1478 adapter was ligated with the gel-purified pSW22 fragment. The resultant plasmid was designated pCBS22 (FIG. 3).

Example 4

Construction of pBTL13

In order to enhance the secretion of the β-receptor and to facilitate the identification of the secreted protein, a sequence encoding the third domain of BAR1 fused to the C-terminal amino acids of substance P was fused in frame with the 5' 1240 bp of the PDGF β-receptor. Plasmid pBTL10 (Example 2) was digested with Sph I and Sst I to isolate the 4 kb fragment comprising the SUC2 signal sequence, a portion of the PDGF β-receptor cDNA and the pUC19 vector sequences. Plasmid pCBS22 was digested with Sph I and Sst I to isolate the 1.2 kb fragment compris-

ing the BAR1-subp fusion and the TPI1 terminator. These two fragments were ligated, and the resultant plasmid was designated pBTL13 (FIG. 4).

Example 5

Construction of an Expression Vector Encoding the Entire PDGF β-Receptor

The entire PDGF β-receptor was directed into the secretory pathway by fusing a SUC2 signal sequence to the 5' end of the PDGF β-receptor coding sequence. This fusion was placed behind the TPIL promoter and inserted into the vector YEp13 for expression in yeast.

The TPI1 promoter was obtained from plasmid pTPIC10 (Alber and Kawasaki, *J. Mol. Apnl. Genet.* 1: 410–434, 1982), and plasmid PFATPOT (Kawasaki and Bell, EP 171,142; ATCC 20699). Plasmid pTPIC10 was cut at the unique Kpn I site, the TPI1 coding region was removed with Bal-31 exonuclease, and an Eco RI linker (sequence: GGA ATT CC) was added to the 3' end of the promoter. Digestion with Bgl II and Eco RI yielded a TPI1 promoter fragment having Bgl II and Eco RI sticky ends. This fragment was then joined to plasmid YRp7' (Stinchcomb et al., *Nature* 282: 39–43, 1979) that had been cut with Bgl II and Eco RI (partial). The resulting plasmid, TE32, was cleaved with Eco RI (partial) and Bam III to remove a portion of the tetracycline resistance gene. The linearized plasmid was then recircularized by the addition of an Eco RI-Bam HI linker to produce plasmid TEA32. Plasmid TEA32 was digested with Bgl II and Eco RI, and the 900 bp partial TPI1 promoter fragment was gel-purified. Plasmid pIC19H (Marsh et al., *Gene* 32:481–486, 1984) was cut with Bgl II and Eco RI and the vector fragment was gel purified. The TPI1 promoter fragment was then ligated to the linearized pIC19H and the mixture was used to transform *E. coli* RR1. Plasmid DNA was prepared and screened for the presence of a ~900 bp Bgl II-Eco RI fragment. A correct plasmid was selected and designated PICTPIP.

The TPI1 promoter was then subcloned to place convenient restriction sites at its ends. Plasmid pIC7 (Marsh et al., ibid.) was digested with Eco RI, the fragment ends were blunted with DNA polymerase I (Klenow fragment), and the linear DNA was recircularized using T4 DNA ligase. The resulting plasmid was used to transform *E. coli* RR1. Plasmid DNA was prepared from the transformants and was screened for the loss of the Eco RI site. A plasmid having the correct restriction pattern was designated pIC7RI*. Plasmid pIC7RI* was digested with Hind III and Nar I, and the 2500 bp fragment was gel-purified. The partial TPI1 promoter fragment (ca. 900 bp) was removed from pICTPIP using Nar I and Sph I and was gel-purified. The remainder of the TPI1 promoter was obtained from plasmid pFATPOT by digesting the plasmid with Sph I and Hind III, and a 1750 bp fragment, which included a portion of the TPI1 promoter fragment from PICTPIP, and the fragment from pFATPOT were then combined in a triple ligation to produce pMVR1 (FIG. 2)

The TPI1 promoter was then joined to the SUC2-PDGF β-receptor fusion. Plasmid PBTL10 (Example 2) was digested with Eco RI and Hind III to isolate the 3.4 kb fragment comprising the SUC2 signal sequence and the entire PDGF β-receptor coding region. Plasmid pMVR1 was digested with Bgl II and Eco RI to isolate the 0.9 kb TPI1 promoter fragment. The TPI1 promoter fragment and the fragment derived from pBTL10 were joined with YEp13, which had been linearized by digestion with Bam III and Hind III, in a three-part ligation. The resultant plasmid was designated pBTL12 (FIG. 2).

Example 6

Construction of an Expression Vector Encoding the 5' Extracellular Portion of the PDGF β-Receptor

The extracellular portion of the PDGF β-receptor was directed into the secretory pathway by fusing the coding sequence to the SUC2 signal sequence. This fusion was placed in an expression vector behind the TPI1 promoter. Plasmid pBTL10 (Example 2) was digested with Eco RI and Sph I to isolate the approximately 1.3 kb fragment comprising the SUC2 signal sequence and the PDGF β-receptor extracellular domain coding sequence. Plasmid pMVR1 (Example 5) was digested with Bgl II and Eco RI to isolate the 0.9 kb TPI1 promoter fragment. The TPI1 promoter fragment was joined with the fragment derived from pBTL10 and YEp13, which had been linearized by digestion with Bam HI and Sph I, in a three-part ligation. The resultant plasmid was designated pBTL11 (FIG. 2).

Example 7

Construction of Yeast Expression Vectors pBTL14 and pBTL15, and The Expression of PDGF β-Receptor-BAR1-subP Fusions

A. Construction of pBTL14

The SUC2-PDGFβ-R fusion was joined with the third domain of BAR1 to enhance the secretion of the receptor, and the expression unit was cloned into a derivative of YEp13 termed pJH50. YEp13 was modified to destroy the Sal I site near the LEU2 gene. This was achieved by partially digesting YEp13 with Sal I followed by a complete digestion with Xho I. The 2.0 kb Xho I-Sal I fragment comprising the LEU2 gene and the 8.0 kb linear YEp13 vector fragment were isolated and ligated together. The ligation mixture was transformed into *E. coli* strain RR1. DNA was prepared from the transformants and was analyzed by digestion with Sal I and Xho I. A clone was isolated which showed a single Sal I site and an inactive Xho I site indicating that the LEU2 fragment had inserted in the opposite orientation relative to the parent plasmid YEp13. The plasmid was designated pJH50.

Referring to FIG. 4, plasmid pBTL12 (Example 5) was digested with Sal I and Pst I to isolate the 2.15 kb fragment comprising 270 bp of YEp13 vector sequence, the TPI1 promoter, the SUC2 signal sequence, and 927 bp of PDGF β-receptor cDNA. Plasmid pBTL13 (Example 4) was digested with Pst I and Hind III to isolate the 1.48 kb fragment comprising 313 bp of PDGF β-receptor CDNA, the BAR1-subP fusion and the TPI1 terminator. The fragments derived from pBTL12 and pBTL13 were joined with pJH50, which had been linearized by digestion with Hind III and Sal I, in a three-part ligation. The resultant plasmid was designated pBTL14.

B. Construction of pBTL15

Referring to FIG. 5, a yeast expression vector was constructed comprising the TPI1 promoter, the SUC2 signal sequence, 1.45 kb of PDGF β-receptor CDNA sequence fused to the BAR1-subP fusion and the TPI1 terminator. Plasmid pBTL12 (Example 5) was digested with Sal I and Esp I to isolate the 2.7 kb fragment comprising the TPI1 promoter, the SUC2 signal sequence, the PDGFβ-R coding sequence, and 270 bp of YEp13 vector sequence. Plasmid pBTL13 (Example 4) was digested with Nru I and Hind III to isolate the 1.4 kb fragment comprising the BAR1-subP fusion, the TPI1 terminator and 209 bp of 3' PDGF β-receptor cDNA sequence. The fragments derived from

EXHIBIT 11
Part 2

pBTL12 and pBTL13 were joined in a three-part ligation with pJH50, which had been linearized by digestion with Hind III and Sal I. The resultant plasmid was designated pBTL15.

C. Expression of PDGFβ-R-subP Fusions from pBTL14 and pBTL15

Yeast expression vectors pBTL14 and pBTL15 were transformed into Saccharomyces cerevisiae strains ZY100 (MATa leu2-3,112 ade2-101 suc2-A9 ca12 pep4::TPI1prom-CAT) and ZY400 (MATa leu2-3.112 ade2-101 suc2-A9 gal2 pep4::TPI1prom-CAT Δmnn9::URA3). Transformations were carried out using the method essentially described by Beggs (ibid.). Transformants were selected for their ability to grow on –LEUDS (Table 2).

TABLE 2

Media Recipes'

-LeuThrTrp Amino Acid Mixture

4 g adenine
3 g L-arginine
5 g L-aspartic acid
2 g L-histidine free base
6 g L-isoleucine
4 g L-lysine-mono hydrochloride
2 g L-methionine
6 g L-phenylalanine
5 g L-serine
5 g L-tyrosine
4 g uracil
6 g L-valine

Mix all the ingredients and grind with a mortar and pestle until the mixture is finely ground.

–LEUDS

20 g glucose

6.7 g Yeast Nitrogen Base without amino acids (DIFCO Laboratories Detroit, Mich.)

0.6 g –LeuThrTrp Amino Acid Mixture

182.2 g sorbitol

18 g Agar

Mix all the ingredients in distilled water. Add distilled water to a final volume of 1 liter. Autoclave 15 minutes. After autoclaving add 150 mg L-threonine and 40 mg L-tryptophan. Pour plates and allow to solidify.

–LEUDS+sodium succinate. pH 6.5

20 g Yeast Nitrogen Base without amino acids

0.6 g –LeuTrpThr Amino Acid Mixture

182.2 g sorbitol

11.8 g succinic acid

Mix all ingredients in distilled water to a final volume of 1 liter. Adjust the pH of the solution to pH 6.5. Autoclave 15 minutes. After autoclaving add 150 mg L-threonine and 40 mg L-tryptophan.

Fermentation Medium

7 g/l yeast nitrogen base without amino acids or ammonium sulfate (DIFCO Laboratories)

0.6 g/l ammonium sulfate

0.5 M sorbitol

0.39 g/l adenine sulfate

0.01% polypropylene glycol

Mix all ingredients in distilled water. Autoclave 15 minutes. Add 80 ml 50% glucose for each liter of medium.

Super Synthetic –LEUD pH 6.5 [liquid or solid medium]

6.7 g Yeast Nitrogen Base without amino acids or ammonium sulfate (DIFCO)

6 g ammonium sulfate

160 g adenine

0.6 g –LeuThrTrp Amino Acid Mixture

20 g glucose

11.8g succinic acid

Mix all ingredients and add distilled water to a final volume of 800 ml. Adjust the pH of the solution to pH 6.4. Autoclave 15 minutes. For solid medium, add 18 g agar before autoclaving, autoclave and pour plates.

Super Synthetic –LEUDS, pH 6,4 (liquid or solid medium)

Use the same recipe as Super Synthetic –LEUD, pH 6.4, but add 182.2 g sorbitol before autoclaving.

YEPD

20 g glucose

20 g Bacto Peptone (DIFCO Laboratories)

10 g Bacto Yeast Extract (DIFCO Labloratories)

18 g agar

4 ml adenine 1%

8 ml 1% L-leucine

Mix all ingredients in distilled water, and bring to a final volume of 1 liter. Autoclave 25 minutes and pour plates.

The transformants were assayed for binding to an anti-PDGF β-receptor monoclonal antibody (PR7212) or an anti-substance P antibody by protein blot assay. ZY100 [pBTL14] and ZY100[pBTL15] transformants were grown ovemight at 30° C. in 5 ml Super Synthetic -LEUD, pH 6.4 (Table 2). ZY400[pBTL14] and ZY400[pBTL15] transformants were grown overnight at 30° C. in 5 ml Super Synthetic-LEUDS, pH 6.4 (Table 2). The cultures were pelleted by centrifugation and the supernatants were assayed for the presence of secreted PDGF β-receptor anlogs by protein blot assay using methods described in Example 18. Results of assays using PR7212 are shown in Table 3.

TABLE 3

Results of a protein blot probed with PR7212

Transformant:	
ZY100[pBTL14]	+
ZY400[pBTL14]	++
ZY100[pBTL15]	+
ZY400[pBTL15]	+

Example 8

Construction of a SUC2-PDGFβ-R Fusion Comprising the Complete PDGFβ-R Extracellular Domain

A. Construction of pBTL22

The PDGFβ-R coding sequence present in pBTL10 was modified to delete the coding region 3' to the extracellular PDGFβ-R domain. As shown in FIG. 6, plasmid pBTL10 was digested with Sph I and Bam HI and with Sph I and Sst II to isolate the 4.77 kb fragment and the 466 bp fragment, respectively. The 466 bp fragment was then digested with Sau 3A to isolate the 0.17 kb fragment. The 0.17 kb fragment and the 4.77 kb were joined by ligation. The resultant plasmid was designated pBTL21.

Plasmid pBTL21, containing a Bam HI site that was regenerated by the ligation of the Bam HI and Sau 3A sites, was digested with Hind III and Bam HI to isolate the 4.2 kb fragment. Synthetic oligonucleotides ZC1846 (Sequence ID Number 16; Table 1) and ZC1847 (Sequence ID Number 17; Table 1) were designed to form an adapter encoding the

PDGFβ-R from the Sau 3A site after bp 1856 (FIG. 1B; (Sequence ID Number 1)) to the end of the extracellular domain at 1958 bp (FIG. 1B; Sequence ID Number 1), having a 5' Bam HI adhesive end that destroys the Bam HI site and a 3' Hind III adhesive end. oligonucleotides ZC1846 and ZC1847 were annealed under conditions described by Maniatis et. al. (ibid.). The 4.2 pBTL21 fragment and the ZC1846/ZC1847 adapter were joined by lrgation. The resultant plasmid, designated pBTL22, comprises the SUC2 signal sequence fused in proper reading frame to the extracellular domain of PDGFβ-R in the vector pUC19 (FIG. 6).

B. Construction of pBTL28

An in-frame translation stop codon was inserted immediately after the coding region of the PDGFβ-R in pBTL22 using oligonucleotides ZC1892 (Sequence ID Number 19; Table 1) and ZC1893 (Sequence ID Number 20; Table 1). These oligonucleotides were designed to form an adapter encoding a stop codon in-frame with the PDGFβ-R coding sequence from pBTL22 flanked by a 5' Hind III adhesive end and a 3' Xba I adhesive end. Plasmid pBTL22 was digested with Eco RI and Hind III to isolate the 1.6 kb SUC2-PDGFβ-R R fragment. Plasmid pMVR1 was digested with Eco RI and Xba I to isolate the 3.68 kb fragment comprising the TPI1 promoter, pIC7RI* vector sequences and the TPI1 terminator. Oligonucleotides ZC1892 and ZC1893 were annealed to form a Hind III-Xba I adapter. The 1.6 kb SUC2-PDGFβ-R fragment, the 3.86 kb pMVR1 fragment and the ZC1892/ZC1893 adapter were joined in a three-part ligation. The resultant plasmid was designated pBTL27.

The expression unit present in pBTL27 was inserted into the yeast expression vector pJH50 by first digesting pJH50 with Bam HI and Sal I to isolate the 10.3 kb vector fragment. Plasmid pBTL27 was digested with Bgl II and Eco RI and with Xho I and Eco RI to isolate the 0.9 kb TPI1 promoter fragment and the 1.65 kb fragment, respectively. The 10.3 kb pJH50 vector fragment, the 0.9 kb TPI1 promoter fragment and 1.65 kb fragment were joined in a three-part ligation. The resultant plasmid was designated pBTL28.

C. Construction of Plasmid pBTL30

The PDGFβ-R coding sequence present in plasmid pBTL22 was modified to encode the twelve C-terminal amino acids of substance P and an in-frame stop codon. Plasmid pBTL22 was digested with Eco RI and Hind III to isolate the 1.6 kb SUC2-PDGFβ-R fragment. Plasmid pMVR1 was digested with Eco RI and Xba I to isolate the 3.68 kb fragment comprising the TPI1 promoter, pIC7R1* and the TPI1 terminator. Synthetic oligonucleotides ZC1894 (Sequence ID Number 21; Table 1) and ZC1895 (Sequence ID Number 22; Table 1) were annealed to form an adapter containing the codons for the twelve C-terminal amino acids of substance P followed by an in-frame stop codon and flanked on the 5' end with a Hind III adhesive end and on the 3' end with an Xba I adhesive end. The ZC1894/ZC1895 adapter, the 1.6 kb SUC2-PDGFβ-R fragment and the pMVR1 fragment were joined in a three-part ligation. The resultant plasmid, designated pBTL29, was digested with Eco RI and Xho I to isolate the 1.69 kb SUC2-PDGFβ-R-subP-TPI1 terminator fragment. Plasmid pBTL27 was digested with Bgl II and Eco RI to isolate the 0.9 kb TPI1 promoter fragment. Plasmid pJH50 was digested with Bam HI and Sal I to isolate the 10.3 kb vector fragment. The 1.69 kb pBTL29 fragment, the 0.9 kb TPI1 promoter fragment and the 10.3 kb vector fragment were joined in a three-part ligation. The resulting plasmid was designated pBTL30.

Example 9

Construction and Expression of a SUC2-PDGFβ-R-IgG Hinge

Expression Vector

An expression unit comprising the TPI1 promoter, the SUC2 signal sequence, the PDGFβ-R extracellular domain, an immunoglobulin hinge region and the TPI1 terminator was constructed. Plasmid pBTL22 was digested with Eco RI and Hind III to isolate the 1.56 kb fragment. Plasmid pMVR1 was digested with Eco RI and Xba I to isolate the 3.7 kb fragment, comprising the TPI1 promoter, pIC7RI* vector sequences and the TPI1 terminator. Oligonucleotides ZC1776 (Sequence ID Number 14; Table 1) and ZC1777 (Sequence ID Number 15; Table 1) were designed to form, when annealed, an adapter encoding an immunoglobulin hinge region with a 5' Hind III adhesive end and a 3'Xba I adhesive end. Oligonucleotides ZC1776 and ZC1777 were annealed under conditions described by Maniatis et al. (ibid.). The 1.56 kb pBTL22 fragment, the 3.7 kb fragment and the ZC1776/ZC1777 adapter were joined in a three-part ligation, resulting in plasmid pBTL24.

The expression unit of pBTL24, comprising the TPI1 promoter, SUC2 signal sequence, PDGFβ-R extracellular domain sequence, hinge region sequence, and TPI1 terminator, was inserted into pJH50. Plasmid pBTL24 was digested with Xho I and Hind III to isolate the 2.4 kb expression unit. Plasmid pJH50 was digested with Hind III and Sal I to isolate the 9.95 kb fragment. The 2.4 kb pBTL24 fragment and 9.95 kb pJH50 vector fragment were joined by ligation. The resultant plasmid was designated pBTL25.

Plasmid pBTL25 was transformed into *Saccharomyces cerevisiae* strain ZY400 using the method essentially described by Beggs (ibid.). Transformants were selected for their ability to grow on –LEUDS (Table 2). The transformants were tested for their ability to bind the anti-PDGFβ-R monoclonal antibody PR7212 using the colony assay method described in Example 18. Plasmid pBTL25 transformants were patched onto nitrocellulose filters that had been wetted and supported by YEPD solid medium. Antibody PR7212 was found to bind to the PDGFβ-R-IgG hinge fusion secreted by ZY400[pBTL25] transformants.

Example 10

Construction and Expression of a SUC2 Signal Sequence-PDGFβ-R Extracellular Domain-SUC2 Fusion

As shown in FIG. 6, an expression unit comprising the TPI1 promoter, SUC2 signal sequence, PDGFβ-R extracellular domain sequence, and SUC2 coding sequence was constructed as follows. Plasmid pBTL22 was digested with Eco RI and Hind III to isolate the 1.6 kb SUC2-PDGFβ-R fragment. Plasmid pMVR1 was digested with Bgl II and Eco RI to isolate the 0.9 kb TPI1 promoter fragment. The SUC2 coding region was obtained from pJH40. Plasmid pJH40 was constructed by inserting the 2.0 kb Hind III-Hind III SUC2 fragment from pRB58 (Carlson et al., *Cell* 28:145–154, 1982) into the Hind III site of pUC19 followed by the destruction of the Hind III site 3' to the coding region. Plasmid pJH40 was digested with Hind III and Sal I to isolate the 2.0 kb SUC2 coding sequence. Plasmid pJH50 was digested with Sal I and Bam HI to isolate the 10.3 kb vector fragment. The 0.9 kb Bgl II-Eco RI TPI1 promoter fragment, the 1.6 kb Eco RI-Hind III SUC2-PDGFβ-R, the 2.0 kb Hind III-Sal I SUC2 fragment and the 10.3 kb Bam HI-Sal I vector fragment were joined in a four-part ligation. The resultant plasmid was designated pBTL26 (FIG. 6).

Plasmid pBTL26 was transformed into *Saccharomyces cerevisiae* strain ZY400 using the method essentially described by Beggs (ibid.). Transformants were selected for their ability to grow on -LEUDS (Table 2). ZY400 transformants (ZY400[pBTL26]) were assayed by protein blot (Example 18), colony blot (Example 18) and competition assay.

Protein blot assays were carried out on ZY400[pBTL26] and ZY400[pJH50] (control) transformants that had been grown in flasks. Two hundred-fifty microliters of a 5 ml overnight cultures of ZY400[pBTL26] and ZY400 [pJH50] in −LEUDS+sodium succinate, pH 6.5 (Table 2) were inoculated into 50 ml of −LEUDS+sodium succinate, pH 6.5. The cultures,were incubated for 35 hours in an airbath shaker at 30° C. The culture supernatants were harvested by centrifugation. The culture supernatants were assayed as described in Example 18 and were found to bind PR7212 antibody.

Colony assays were carried out on ZY400[pBTL26] transformants. ZY400[pBTL26] transformants were patched onto wetted nitrocellulose filters that were supported on a YEPD plate. The colony assay carried out as described in Example 8.A. showed that ZY400[pBTL26] antibodies bound PR7212 antibodies.

Competition binding assays were carried out on ZY400 [pBTL26] and ZY400[pJH50] transformants. The transformants were grown in two liters of fermentation medium (Table 2) in a New Brunswick Bioflo2 fermentor (New Brunswick, Philadelphia, Pa.) with continuous pH control at pH 6.4. The cultures were adjusted to pH 7.5 immediately prior to harvesting. Culture supernatants were concentrated in an Amicon concentrator (Amicon, San Francisco, Calif.) using an Amicon 10^4 mw spiral filter cartridge. The concentrated supernatants were further concentrated using Amicon Centriprep 10's. Fifteen milliliters of the concentrated supernatant samples were added to the Centripreps, and the Centripreps were spun in a Beckman GRP centrifuge (Beckman Instruments Inc., Carlsbad, Calif.) at setting 5 for a total of 60 minutes. The concentrates were removed from the Centripreps and were assayed in the competition assay.

The competition binding assay measured the amount of ^{125}I-PDGF left to bind to fetal foreskin fibroblast cells after preincubation with the concentrate containing the PDGFβ-R-SUC2 fusion protein. PDGF-AA and PDGF-AB were iodinated using the Iodopead method (Pierce Chemical). PDGF-$BB_{p,p'}$ was iodinated and purified as described in Example 18.F. The concentrate was serially diluted in binding medium (Table 4). The dilutions were mixed with 0.5 ng of iodinated PDGF-AA, PDGF-$BB_{p,p'}$ or PDGF-AB, and the mixtures were incubated for two hours at room temperature. Three hundred micrograms of unlabeled PDGF-BB was added to each sample mixture. The sample mixtures were added to 24-well plates containing confluent fetal foreskin fibroblast cells (AG1523, available from the Human Genetic Mutant *Cell Repository, Camden, N.J.). The cells were incubated with the mixture for four hours at* 4° C. The supernatants were aspirated from the wells, and the wells were rinsed three times with phosphate buffered saline that was held at 4° C. (PBS; Sigma, St. Louis, Mo.). Five hundred microliters of PBS+1% NP-40 was added to each well, and the plates were shaken on a platform shaker for five minutes. The cells were harvested and the amount of iodinated PDGF was determined. The results of the competition binding assay showed that the PDGFβ-R-SUC2 fusion protein was able to competitively bind all three isoforms of PDGF.

The PDGFβ-R produced from ZY400 [pBTL26] transformants was tested for cross reactivity to fibroblast growth factor (FGF) and transforming growth factorβ (TGF-β) using the competition assay essentially described above. Supernatant concentrates from ZY400[pBTL26] and ZY400 [JH50] (control) transformants were serially diluted in binding medium (Table 4). The dilutions were mixed with 7.9 ng of iodinated FGF or 14 ng of iodinated TGF-β, and the mixtures were incubated for two hours at room temperature. Fourteen micrograms of unlabeled FGF was added to each mixture containing labeled FGF, and 7 μg of unlabeled TGF-β was added to each mixture containing labeled TGF-β. The sample mixtures were added to 24-well plates containing confluent human dermal fibroblast cells. (Human dermal fibroblast cells express both FGF receptors and TGFβ receptors.) The cells were incubated with the mixtures for four hours at 4° C. Five hundred microliters of PBS+1% NP-40 was added to each well, and the plates were shaken on a platform shaker for five minutes. The cells were harvested and the amount of iodinated FGF or TGF-β bound to the cells was determined. The results of these assays showed that the PDGFβ-R-SUC2 fusion protein did not cross react with FGF or TGF-β.

TABLE 4

Reagent Recipes

Binding Medium

500 ml	Ham's F-12 medium
12 ml	1M HEPES, pH 7.4
5 ml	100× PSN (Penicillin/Streptomycin/Neomycin, (Gibco)
1 g	rabbit serum albumin

Western Transfer Buffer

25 mM	Tris, pH 8.3
19 mM	glycine, pH 8.3
20%	methanol

Western Buffer A

50 ml	1 M Tris, pH 7.4
20 ml	0.25 mM EDTA, pH 7.0
5 ml	10% NP-40
37.5 ml	4 M NaCl
2.5 g	gelatin

The Tris, EDTA, NP-40 and NaCl were diluted to a final volume of one liter with distilled water. The gelatin was added to 300 ml of this solution and the solution was heated in a microwave until the gelatin was in solution. The gelatin solution was added back to the remainder of the first solution and stirred at 4° C. until cool. The buffer was stored at 4° C.

Western Buffer B

50 ml	1 M Tris, pH 7.4
20 ml	0.25 M EDTA, pH 7.0
5 ml	10% NP-40
58.4 g	NaCl
2.5 g	gelatin
4 g	N-lauroyl sarcosine

The Tris, EDTA, NP-40, and the NaCl were mixed and diluted to a final volume of one liter. The gelatin was added to 300 ml of this solution and heated in a microwave until the gelatin was in solution. The gelatin solution was added back to the original solution and the N-lauroyl sarcosine was added. The final mixture was stirred at 4° C. until the solids were completely dissolved. This buffer was stored at 4° C.

2× Loading Buffer

36 ml	0.5 M Tris-HCL pH 6.8
16 ml	glycerol
16 ml	20% SDS
4 ml	0.5% Bromphenol Blue in 0.5 M Tris-HCL pH 6.8

Mix all ingredients. Immediately before use, add 100 μl β-mercaptoethanol to each 900 μl dye mix

Example 11

Construction and Expression of PDGF Receptor Analogs From CHO cells

A. Construction of *pBTL114* and *pB T 115*

The portions of the PDGF β-receptor extracellular domain present in pBTL14 and pBTL15 were placed in a mammalian expression vector. Plasmids pBTL14 and pBTL15 were digested with Eco RI to isolate the 1695 bp and 1905 bp SUC2 signal-PDGFβ-R-BAR1 fragments. The 1695 bp fragment and the 1905 bp fragment were each ligated to Zem229R that had been linearized by digestion with Eco RI.

The vector Zem229R was constructed as shown in FIG. 10 from Zem229. Plasmid Zem229 is a pUC18-based expression vector containing a unique Bam HI site for insertion of cloned DNA between the mouse metallothionein-1 promoter and SV40 transcription terminator and an expression unit containing the SV40 early promoter, mouse dihydrofolate reductase gene, and SV40 transcription terminator. Zem229 was modified to delete the Eco RI sites flanking the Bam HI cloning site and to replace the Bamr HI site with a single Eco RI cloning site. The plasmid was partially digested with Eco RI, treated with DNA polymerase I (Klenow fragment) and dNTPs, and religated. Digestion of the plasmid with Bam HI followed by ligaion of the linearized plasmid with a Bam HI-Eco I adapter resulted in a unique Eco RI cloning site. The resultant plasmid was designated Zem229R.

The ligation mixtures were transformed into *E. coli* strain RR1. Plasmid DNA was prepared and the plasmids were subjected to restriction enzyme analysis. A plasmid having the 1695 bp pBTL14 fragment inserted into Zem229R in the correct orientation was designated pBTL114 (FIG. 9). A plasmid having the 1905 bp pBTL15 fragment inserted into Zem229R in the correct orientation was designated pBTL115 (FIG. 9).

B. Expression of Secreted PDGF β-Receptor Analogs in *tk*⁻ ts13 BHK Cells

Plasmids pBTL114 and pBTL115 were each transfected into tk⁻ts13 cells using calcium phosphate precipitation (essentially as described by Gruham and van der Eb, *J. Gen. Virol.* 36: 59–72, 1977). The transfected cells were grown in Dulbecco's modified Eagle's medium (DMEM) containing 10% fetal calf serum, 1×PSN antibiotic mix (Gibco 600–5640), 2.0 mM L-glutamine. The cells were selected in 250 nM methotrexate (MTX) for 14 days, and the resulting colonies were screened by the immunofilter assay (McCracken and Brown, *Biotechniques*, 82–87, March/April 1984). Plates were rinsed with PBS or No Serum medium (DMEM plus 1×PSN antibiotic mix). Teflon® mesh (Spectrum Medical Industries, Los Angeles, Calif.) was then placed over the cells. Nitrocellulose filters were wetted with PBS or No Serum medium, as appropriate, and placed over the mesh. After six hours incubation at 37° C., filters were removed and placed in Wester buffer A (Table 4) overnight at room temperature. The filters were developed using the antibody PR7212 and the procedure described in Example 8. The filters showed that conditioned media from pBTL114-transfected and pBTL115-transfected BHK cells bound the PR7212 antibody indicating the presence of biologically active secreted PDGFβ-R.

Example 12

Expression of PDGF β-Receptor Analogs in Cultured Mouse Myeloma Cells

A. Construction of pIC*μ*PRE8

The immunoglobulin *μ* heavy chain promoter and enhancer were subcloned into pIC19H to provide a unique Hind III site 3' to the promoter. Plasmid pμ (Grosschedl and Baltimore, *Cell* 41: 885–897, 1985) was digested with Sal I and Eco RI to isolate the 3.1 kb fragment comprising the *μ* promoter. Plasmid pIC19H was linearized by digestion with Eco RI and Xho I. The p promoter fragment and the linearized pIC19H vector fragment were joined by ligation. The resultant plasmid, designated pIC*μ*3, was digested with Ava II to isolate the 700 bp *μ* promoter fragment. The 700 bp fragment was blunt-ended by treatment with DNA polymerase I (Klenow fragment) and deoxynucleotide triphosphates. Plasmid pIC19H was linearized by digestion with Xho I, and the adhesive ends were filled in by treatment with DNA polymerase I (Klenow fragement) and deoxynucleotide triphosphates. The blunt-ended Ava II fragment was ligated with the blunt-ended, linearized pIC19H, and the ligation mixture was transformed into *E. coli* JM83. Plasmid DNA was prepared from the transformants and was analyzed by restriction digest. A plasmid with a Bgl II site 5' to the promoter was designated pIC*μ*PR1(–). Plasmid pIC*μ*PR1 (–) was digested with Hind III and Bgl II to isolate the 700 bp *μ* promoter fragment. Plasmid pIC19R was linearized by digestion with Hind III and Bam III. The 700 bp promoter fragment was joined with the linearized pIC19R by ligation. The resultant plasmid, designated pIC*μ*PR7, comprised the *μ* promoter with an unique Sma I site 5' to the promoter and a unique Hind III site 3' to the promoter.

The immunoglobulin heavy chain *μ* enhancer (Gillies et al., *Cell* 33: 717–728, 1983) was inserted into the unique Sma I site to generate plasmid pIC*μ*PRE8. Plasmid pJ4 (obtained from F. Blattner, Univ. Wisconsin, Madison, Wis.), comprising the 1.5 kb Hind III-Eco RI λ enhancer fragment in the vector pAT153 (Amersham, Arlington Heights, Ill.), was digested with Hind III and Eco RI to isolate the 1.5 kb enhancer fragment. The adhesive ends of the enhancer fragment were filled in by treatment with T4 DNA polymerase and deoxynucleotide triphosphates. The blunt-ended fragment and pIC*μ*PR7, which had been linearized by digestion with Sma I, were joined by ligation. The ligation mixture was transformed into *E. coli* RR1. Plasmid DNA was prepared from the transformants, and the plasmids were analyzed by restriction digests. A plasmid comprising the *μ* enhancer and the λ promoter was designated pIC*μ*PRE8 (FIG. 7).

B. Construction of pSDL114

The DNA sequence encoding the extracellular domain of the PDGF β-receptor was joined with the DNA sequence encoding the human immunoglobulin light chain constant region. The PDGF β-receptor extracellular domain was obtained from mpBTL22, which comprised the Eco RI-Hind III fragment from pBTL22 (Example 8.A.) cloned into Eco RI-Hind III cut M13mp18. Single stranded DNA was prepared from a mpBTL22 phage clone, and the DNA was subjected to in vitro mutagenesis using the oligonucleotide ZC1886 (Table 1) and the method described by Kunkel (*Proc. Natl. Acad. Sci. USA* 82: 488–492, 1985). A phage clone comprising the mutagenized PDGFβ-R with a donor splice site (5° splice site) at the 3' end of the PDGFβ-R extracellular domain was designated pBTLR-HX (FIG. 7).

The native PDGFβ-R signal sequence was obtained from pPR5. Plasmid PPR5, comprising 738 bp of 5' coding sequence with an Eco RI site immediately 5' to the translation initiation codon, was constructed by in vitro mutagenesis of the PDGFβ-R cDNA fragment from RP51 (Example 1). Replicative form DNA of RP51 was digested with Eco RI to isolate the 1.09 kb PDGFβ-R fragment. The PDGFβ-R fragment was cloned into the Eco RI site of M13mp18. Single stranded template DNA was prepared from a phage clone containing the PDGFβ-R fragment in the proper orientation. The template DNA was subjected to in vitro mutagenesis using oligonucleotide ZC1380 (Sequence ID Number 8; Table 1) and the method described by Zoller and

Smith (*Meth. Enzymol.* 100: 468–500, 1983). The mutagenesis resulted in the placement of an Eco RI site immediately 5' to the translation initiation codon. Mutagenized phage clones were analyzed by dideoxy sequence analysis. A phage clone containing the ZC1380 mutation was selected, and replicative form (RF) DNA was prepared from the phage clone. The RF DNA was digested with Eco RI and Acc I to isolate the 0.63 kb fragment. Plasmid pR-RXI (Example 1) was digested with Acc I and Eco RI to isolate the 3.7 kb fragment. The 0.63 kb fragment and the 3.7 kb fragment were joined by ligation resulting in plasmid pPR5 (FIG. 7).

As shown in FIG. 7, the PDGFβ-R signal peptide and part of the extracellular domain were obtained from plasmid pPR5 as a 1.4 kb Eco RI-Sph I fragment. Replicative form DNA from phage clone pBTLR-HX was digested with Sph I and Hind III to isolate the approximately 0.25 kb PDGFβ-R fragment. Plasmid pUC19 was linearized by digestion with Eco RI and Hind III. The 1.4 kb Eco RI-Sph I PDGFβ-R fragment, the 0.25 kb Sph I-Hind III fragment from pBTLR-HX and the Eco RI-Hind III cut pUC19 were joined in a three-part ligation. The resultant plasmid, pSDL110, was digested with Eco RI and Hind III to isolate the 1.65 kb PDGFβ-R fragment.

Plasmid pICHuCr3.9.11 was used as the source of the human immunoglobulin light chain gene (FIG. 7). The human immunoglobulin light chain gene was isolated from a human genomic library using an oligonucleotide probe (5' TGT GAC ACT CTC CTG GGA GTT A 3'; Sequence ID Number 32), which was based on a published human kappa C gene sequence (Hieter et al., *Cell* 22: 197–207, 1980). The human light chain (kappa) constant region was subcloned as a 1.1 kb Sph I-Hinf I genomic fragment of the human kappa gene, which has been treated with DNA polymerase DNA I (Klenow fragment) to fill in the Hinf I adhesive end, into Sph I-Hinc II cut pUC19. The 1.1 kb human kappa constant region was susbsequently isolated as a 1.1 kb Sph I-Bam HI fragment that was subdloned into Sph I-Bgl II cut pIC19R (Marsh et al., ibid.). The resultant plasmid was designated pICHuCX3.9.11. Plasmid pICHuC$_{78}$ 3.9.11 was digested with Hind III and Eco RI to isolate the 1.1 kb kappa constant region gene. Plasmid pIC19H was linearized by digestion with Eco RI. The 1.65 kb PDGFβ-R fragment, the 1.1 kb human kappa constant region fragment and the linearized pIC19H were joined in a three part ligation. The resultant plasmid, pSDL112, was digested with Bam HI and Cla I to isolate the 2.75 kb fragment. Plasmid ppPRE8 was linearized with Bgl II and Cla I. The 2.75 kb fragment and the linearized ppPRE8 were joined by ligation. The resultant plasmid was designated pSDL114 (FIG. 7).

Plasmid pSDL114 was linearized by digestion with Cla I and was cotransfected with Pvu I-digested p416 into SP2/0-Ag14 (ATCC CRL 1581) by electroporation using the method essentially described by Neumann et al. (*EMBO J.* 1: 841–845, 1982). (Plasmid p416 comprises the Adenovirus 5 ori, SV40 enhancer, Adenovirus 2 major late promoter, Adenovirus 2 tripartite leader, 5' and 3' splice sites, the DHFR' cDNA, the SV40 polyadenylation signal and 1 ML-1 (Lusky and Botchan, *Nature* 2: 79–81, 1981) vector sequences.) Transfectants were selected in growth medium containing methotrexate.

Media from drug resistant clones were tested for the presence of secreted PDGF β-receptor analogs by enzyme-linked immunosorbant assay (ELISA). Ninety-six well assay plates were prepared by incubating 100 μl of 1 μg/ml polyclonal goat anti-human kappa chain (Cappel Laboratories, McLvern, Pa.) diluted in phosphate buffered saline (PBS; Sigma) overnight at 4° C. Excess antibody was removed by three washes with 0.5% Tween 20 in PBS. One hundred microliters of spent media was added to each well, and the well were incubated for one hour at 4° C. Unbound proteins were removed by eight washes with 0.5% Tween 20 -in PBS. One hundred microliters of peroxidase-conjugated goat anti-human kappa antibody (diluted 1:1000 in a solution containing 5% chicken serum (GIBCO-BRL)+0.5% Tween 20 in PBS) was added to each well and the wells were incubated for one hour at 4° C. One hundred microliters of chromophore (100 ul ABTS (2,2'-Azinobis(3-ethylbenz-thiazoline sulfonic acid) diammonium salt; Sigma)+1 μl 30% H_2O_2+12.5 ml citrate/phosphate buffer (9.04 g/l citric acid, 10.16 g/l Na_2HPO_4)) was added to each well, and the wells were incubated to thirty minutes at room temperature. The samples were measured at 405 nm. The results of the assay showed that the PDGFβ-R analog secreted by the transfectants contained an immunoglobulin light chain.

Spent media from drug resistant clones was also tested for the presence of secreted PDGF β-receptor analogs by immunoprecipitation. Approximately one million drug resistant transfectants were metabolically labeled by growth in DMEM medium lacking cysteine+2% calf serum for 18 hours at 37° C. in the presence of 50 μCi ^{35}S-cysteine. Media was harvested from the labeled cells and 250 μl of the spent media was assayed by immunoprecipitation with the anti-PDGF β-receptor antibody PR7212 to detect the prsence of metabolically labeled PDGF β-receptor analogs. PR7212, diluted in PBS, was added to the media to a final concentration of 2.5 μg per 250 μl spent media. Five microliters of rabbit anti-mouse Ig diluted in PBS was added to the PR7212/media mixtures. The immunocomplexes were precipitated by the addition of 50 μl 10% fixed Staph A (weight/volume in PBS). The immunocomplexes were analyzed on 8% SDS-polyacrylamide gels followed by autoradiography overnight at −70° C. The results of the immunoprecipitation showed that the PDGF β-receptor analog secreted by the transfectants was bound by the anti-PDGF β-receptor antibody. The combined results of the ELISA and immunoprecipitation assays showed that the PDGF β-receptor analog secreted by the transfectants contained both the PDGF β-receptor ligand-binding domain and the human light chain constant region.

C. Cotransfection of pSDL114 with an Immunoglobulin Heavy Chain

Plasmid pSDL114 was cotransfected with pφ5V$_H$huC$_\gamma$1M-neo, which encodes a neomycin resistance gene expression unit and a complete mouse/human chimeric immunoglobulin heavy chain gene expression unit.

Plasmid pφ5V$_H$huC$_\gamma$1M-neo was constructed as follows. The mouse immunoglobulin heavy chain gene was isolated from a lambda genomic DNA library constructed from the murine hybridoma cell line NR-ML-05 (Seralini et al., *Eur. J. Nucl. Med.* 14: 232, 1988) using an oligonucleotide probe designed to span the V$_H$/D/J$_H$ junction (5'GCA TAG TAG TTA CCA TAT CCT CTT GCA CAG 3'; Sequence ID Number 33). The human immunoglobulin gamma-1 C gene was isolated from a human genomic library using a cloned human gamma-4 constant region gene (Ellison et al., *DNA* 1: 11–18, 1981). The mouse immunoglobulin variable region was isolated as a 5.3 kb Sst I-Hind III fragment from the original phage clone and the human gamma-1 C gene was obtained from the original phage clone as a 6.0 kb Hind III-Xho I fragment. The chimeric gamma-1 C gene was created by joining the V$_H$ and C$_H$ fragments via the common Hind III site and incorporating them with the *E. coli* neomycin resistance gene expression unit into pIC19H to yield pφ5V$_H$huC$_\gamma$1M-neo.

Plasmid pSDL114 was linearized by digestion with Cla I and was co-transfected into SP2/0-Ag14 cells with Asp 718 linearized pφ5V$_\mu$huC$_\gamma$1M-neo. The transfectants were selected in growth medium containing methotrexate and neomycin. Media from drug-resistant clones were tested for their ability to bind PDGF in a competition binding assay.

The competition binding assay measured the amount of ^{125}-PDGF left to bind to human dermal fibroblast cells after preincubation with the spent media from pSDL114-pφ5V$_\mu$huC$_\gamma$1M-neo transfected cells. The media were serially diluted in binding medium (Table 4). The dilutions were mixed with 0.5 ng of iodinated PDGF-BB or iodinated PDGF-AA, and the mixtures were incubated for two hours at room temperature. Three hundred micrograms of unlabeled PDGF-BB or unlabeled PDGF-AA was added to one tube from each series. The sample mixtures were added to 24 well plates containing confluent human dermal fibroblast cells. The cells were incubated with the mixture for four hours at 4° C. The supernatants were aspirated from the wells, and the wells were rinsed three times with phosphate buffered saline that was held a 4° C. (PBS; Sigma, St. Louis, Mo.). Five hundred microliters of PBS+1% NP-40 was added to each well, and the plates were shaken on a platform shaker for five minutes. The cells were harvested and the amount of iodinated PDGF was determined. The results of the competition binding assay showed that the protein produced from pSDL114-pφ5V$_\mu$huC$_\gamma$1M- neo transfected cells was able to competetively bind PDGF-BB but did not bind PDGF-AA.

The PDGF β-receptor analog produced from a pSDL114-pφ5V$_\mu$huC$_\gamma$1M-neo transfectant was assayed to determine if the receptor analog was able to bind PDGF-BB with high affinity. Eight and one half milliliters of spent media containing the PDGFβ-R analogs from a pSDL114-pφ5V$_\mu$huC$_\gamma$1M-neo transfectant was added to 425 μl of Sepharose CI-4B-Protein A beads (Sigma, St. Louis, Mo.), and the mixture was incubated for 10 minutes at 4° C. The beads were pelleted by centrifugation and washed with binding medium (Table 4). Following the wash the beads were resuspended in 8.5 ml of binding media, and 0.25 ml aliquots were dispensed to 1.5 ml tubes. Binding reactions were prepared by adding iodinated PDGF-BB$_{Tyr}$ (Example 18.F.) diluted in DMEM+10% fetal calf serum to the identical aliquots of receptor-bound beads to final PDGF-BB$_{Tyr}$ concentrations of between 4.12 pM and 264 pM. Nonspecific binding was determined by adding a 100 fold excess of unlabeled BB to an identical set of binding reactions. Mixtures were incubated overnight at 4° C.

The beads were pelleted by centrifugation, and unbound PDGF-BB was removed with three washes in PBS. The beads were resuspended in 100 μl of PBS and were counted. Results of the assay showed that the PDGFβ-R analog was able to bind PDGF-BB with high affinity.

D. Construction of pSDL113.

As shown in FIG. 8, the DNA sequence encoding the extracellular domain of the PDGF β-receptor was joined with the DNA sequence encoding a human immunoglobulin heavy chain constant region joined to a hinge sequence. Plasmid pSDL110 was digested with Eco RI and Hind III to isolate the 1.65 kb PDGFβ-R fragment. Plasmid pIC1Hu$_\gamma$-1M was used as the source of the heavy chain constant region and hinge region. Plasmid pIC1Hu$_\gamma$-1M comprises the approximately 6 kb Hind III-Xho I fragment of a human gamma-1 C gene subcloned into pUC19 that had been linearized by digestion with Hind III and Sal I. Plasmid pIC1Hu$_\gamma$-1M was digested with Hind III and Eco RI to isolate the 6 kb fragment encoding the human heavy chain constant region. Plasmid pIC19H was linearized by digestion with Eco RI. The 1.65 kb PDGFβ-R fragment, the 6 kb heavy chain constant region fragment and the linearized pIC19H were joined in a three part ligation. The resultant plasmid, pSDL111, was digested with Bam HI to isolate the 7.7 kb fragment. Plasmid pμPRE8 was linearized with Bgl II and was treated with calf intestinal phosphatase to prevent self-ligation. The 7.7 kb fragment and the linearized pμPRE8 were joined by ligation. A plasmid containing the insert in the proper orientation was designated pSDL113 (FIG. 8).

Plasmid pSDL113 is linearized by digestion with Cla I and is cotransfected with Pvu I-digested p416 into SP2/0-Ag14 by electroporation. Transfectants are selected in growth medium containing methotrexate.

Media from drug resistant clones are tested for the presence of secreted PDGFβ-R analogs by immunoprecipitation using the method described in Example 12.B.

E. Cotransfection of pSDL113 with an Immunoglobulin Light Chain Gene

Plasmid pSDL113 is linearized by digestion with Cla I and was cotransfected with pICφ5V$_\kappa$HuC$_\kappa$-Neo, which encodes a neomycin resistance gene and a mouse/human chimeric immunoglobulin light chain gene. The mouse immunoglobulin light chain gene was isolated from a lambda genomic DNA library constructed from the murine hybridoma cell line NR-ML-05 (Woodhouse et al., ibid.) using an oligonucleotide probe designed to span the V$_\kappa$/J$_\kappa$ junction (5' ACC GAA CGT GAG AGG AGT GCT ATA A 3'; Sequence ID Number 34). The human immunoglobulin light chain constant region gene was isolated as described in Example 12.B. The mouse NR-ML-05 immunoglobulin light chain variable region gene was subcloned from the original mouse genomic phage clone into pIC19R as a 3 kb Xba I-Hinc II fragment. The human kappa C gene was subcloned from the original human genomic phage clone into pUC19 as a 2.0 kb Hind III-Eco RI fragment. The chimeric kappa gene was created by joining the NR-ML-05 light chain variable region gene and human light chain constant region gene via the common Sph I site and incorporating them with the *E. coli* neomycin resistance gene into pIC19H to yield pICφ5V$_\kappa$HuC$_\kappa$-Neo (FIG. 9).

The linearized pSDL113 and pICφ5V$_\kappa$HuC$_\kappa$-Neo are transfected into SP2/0-Ag14 cells by electroporation. The transfectants are selected in growth medium containing methotrexate and neomycin.

F. Cotransfection of pSDL113 and pSDL114

A clone of SP2/0-Ag14 stably transfected with pSDL114 and p416 was co-transfected with Cla I-digested pSDL113 and Bam HI-digested pICneo by electroporation. (Plasmid pICneo comprises the SV40 promoter operatively linked to the *E. coli* neomycin resistance gene and pIC19H vector sequences.) Transfected cells were selected in growth medium containing methotrexate and G418. Media from drug-resistanct clones were tested for their ability to bind PDGF-BB or PDGF-AA in a competition binding assay as described in Example 12.C. The results of the assay showed that the transfectants secreted a PDGF β-receptor analog which was capable of competitively binding PDGF-BB but did not detectably bind to PDGF-AA.

G. Cotransfection of pSDL114 with Fab

A clone of SP2/0-AG14 stably transfected with pSDL114 and p416 was transfected with the Fab region of the human gamma-4 gene (γ_4) in plasmid pφ5V$_\mu$Fab-neo.

Plasmid pφ5V$_\mu$Fab-neo was constructed by first digesting plasmid p24BR1 (Ellison et al., *DNA* 1: 11, 1988) was digested with Xma I and Eco RI to isolate the 0.2 kb

fragment comprising the immunoglobulin 3' untranslated region. Synthetic oligonucleotides ZC871 (Sequence ID Number 3; Table 1) and ZC872 (Sequence ID Number 4; Table 1) were kinased and annealed using essentially the methods described by Maniatis et al. (ibid.). The annealed oligonucleotides ZC871/ZC872 formed an Sst I-Xma I adapter. The ZC871/ZC872 adapter, the 0.2 kb p24BRH fragment and Sst I-Eco RI linearized pUC19 were joined in a three-part ligation to form plasmid $p\gamma_A3'$. Plasmid $p\gamma_A3'$ was linearized by digestion with Bam HI and Hind III. Plasmid p24BRH was cut with Hind III and Bgl II to isolate the 0.85 kb fragment comprising the C_H1 region. The $p\gamma_A3'$ fragment and the Hind III-Bgl II p24BRH fragment were joined by ligation to form plasmid $p\gamma_AFab$. Plasmid $p\gamma_AFab$ was digested with Hind III and Eco RI to isolate the 1.2 kb fragment comprising γ_AFab. Plasmid pICneo, comprising the SV40 promoter operatively linked to the *E. coli* neomycin resistance gene and pIC19H vector sequences, was linearized by digestion with Sst I and Eco RI. Plasmid $p\phi5V_H$, comprising the mouse immunoglobulin heavy chain gene variable region and pUC18 vector sequences, was digested with Sst I and Hind III to isolate the 5.3 kb V_H fragment. The linearized pICneo was joined with the 5.3 kb Sst I-Hind III fragment and the 1.2 kb Hind III-Eco RI fragment in a three-part ligation. The resultant plasmid was designated $p\phi5V_HFab$-neo (FIG. 10).

A pSDL114/p416-transfected SP2/0-AG14 clone was transfected with Sca I-linearized $p\phi5V_HFab$-neo. Transfected cells were selected in growth medium containing methotrexate and G418. Media from drug-resistant clones were tested for their ability to bind PDGF in a competition binding assay as described in Example 12.C. The results of the assay showed that the PDGF β-receptor analog secreted from the transfectants was capable of competitively binding PDGF-BB.

H. Cotransfection of pSDL114 with Fab'

A stably transfected SP2/0-AG14 isolate containing pSDL114 and p416 was transfected with plasmid pWK1, which contained the Fab' portion of an immunoglobulin heavy chain gene. Plasmid pWK1 was constructed as follows.

The immunoglobulin gamma-1 Fab' sequence, comprising the C_H1 and hinge regions sequences, was derived from the gamma-1 gene clone described in Example 12.C. The gamma-1 gene clone was digested with Hind III and Eco RI to isolate the 3.0 kb fragment, which was subcloned into Hind III-Eco RI linearized M13mp19. Single-stranded template DNA from the resultant phage was subjected to site-directed mutagenesis using oligonucleotide ZC1447 (Sequence ID Number 9; Table 1) and essentially the method of Zoller and Smith (ibid.). A phage clone was identified having a ZC1447 induced deletion resulting in the fusion of the hinge region to a DNA sequence encoding the amino acids Ala-Leu-His-Asn-His-Tyr-Thr-Glu-Lys-Ser-Leu-Ser-Leu-Ser-Pro-Gly-Lys (Sequence ID Number 31) followed in-frame by a stop codon. Replicative form DNA from a positive phage clone was digested with Hind III and Eco RI to isolate the 1.9 kb fragment comprising the C_H1 and hinge regions. Plasmid $p\phi5V_H$ was digested with Sst I and Hind III to isolate the 5.3 kb fragment comprising the mouse immunoglobulin heavy chain gene variable region. Plasmid pICneo was linearized by digestion with Sst I and Eco RI. The linearized pICneo was joined with the 5.3 kb Hind III-Sst I fragment and the 1.9 kb Hind III-Eco RI fragment in a three-part ligation. The resultant plasmid was designated pWK1 (FIG. 10).

An SP2/0-AG14 clone stably transfected with pSDL114 and p416 was transfected with Asp 718-linearized pWK1. Transfected cells were selected by growth in medium containing methotrexate and G418. Media samples from transfected cells were assayed using the competition assay described in Example 12.C. Results from the assays showed that the transfected cells produced a PDGF β-receptor analog capable of competetively binding PDGF-BB.

Example 13

Purification and Characterization of PDGF β-Receptor Analogs from Mammalian cells Co-transfected With pSDL113 and pSDL114

A. Purification of PDGF β-Receptor Analogs

The PDGF β-receptor analog was purified from conditioned culture media from a clone of transfected cells grown in a hollow fiber system. The media was passed over a protein-A sepharose column, and the column was washed sequentially with phosphate buffered saline, pH 7.2 (PBS; Sigma, St. Louis, Mo.) and 0.1 M citrate, pH 5.0. The PDGF β-receptor analog was eluted from the protein-A column with 0.1 M citrate pH 2.5 and immediately neutralized by the addition of Tris-base, pH 7.4. The eluate fractions containing PDGF β-receptor analog, as determined by silver stain, were pooled and chromatographed over an S-200 column (Pharmacia LKB Technologies, Inc., Piscataway, N.J.) equilibrated with PBS. The peak fractions from the S-200 column were pooled and concentrated on a centriprep-10 concentrator (Amicon). Glycerol (10% final volume) was added to the preparation and the sample frozen at −80° C. PDGF β-receptor analogs purified from pSDL114+pSDL113 co-transfected cells were termed "tetrameric PDGF α-receptors".

B. Measurement of The Relative Binding Affinity of Tetrameric PDGF β-Receptor Analog by Soluble Receptor Assay

Purified tetrameric PDGF β-receptor analog was compared to detergent solubilized extracts of human dermal fibroblasts for ^{125}I-labeled PDGF-BB binding activity in a soluble receptor assay essentially as described by Hart et al. (*J. Biol. Chem.* 262: 10780–10785, 1987). Human dermal fibroblast cells were extracted at 20×10^6 cell equivalents per ml in TNEN extraction buffer (20 mM Tris-HCl, pH 8.0, 100 mM NaCl, 1 mM EDTA, 0.5% Nonidet P-40, 1 mM PMSF, 10% glycerol). Two hundred and fifty thousand PDGF β-receptor-subunits per cell was used to calculate the tetrameric PDGF β-receptor analog number per volume of extract. This value has been previously published by Seifert et al. (*J. Biol Chem.* 264: 8771–8778, 1989). The PDGF β-receptor analog number was determined from the protein concentration of the PDGF β-receptor analog assuming an average molecular weight of 140 kDa for each immunoglobulin-PDGF β-receptor monomer, and four monomers per tetramer. Thus, each tetrameric molecule contains four receptor molecules.

Increasing amounts of either detergent solubilized extracts of human dermal fibroblast cells or purified PDGF β-receptor analog were incubated with 1 ng of ^{125}I-labeled PDGF-BB for one hour at 37° C. The sample was then diluted with 1 ml binding media and was added to monolayers of human dermal fibroblast cells grown in 24-well culture dishes. The samples were incubated for two hours at 4° C. The wells were washed to remove unbound, ^{125}I-labeled PDGF-BB. On half of a milliliter of extraction buffer (PBS+1% Nonidet P-40) was added to each well followed by a 5 minute incubation. The extraction mixtures were harvested and counted in a gamma counter.

The results showed that the PDGF β-receptor analog had the same relative binding affinity as solubilized PDGF

β-receptor-subunit from mammalian cells in a solution phase binding assay.

C. Determination of the Binding Affinity of the PDGF β-Receptor Analog in a Solid Phase Format

The apparent dissociation constant K_D(app) of the PDGF β-receptor analog was determined essentially as described by Bowen-Pope and Ross (*Methods in Enzymology* 109: 69–100, 1985), using the concentration of ^{125}I-labeled PDGF-BB giving half-maximal specific ^{125}I-labeled PDGF-BB binding. Saturation binding assays to determine the concentration of ^{125}I-labeled PDGF-BB that gave half-maximal binding to immobilized PDGF β-receptor analog were conducted as follows.

Affinity purified goat anti-human IgG, H- and L-chain (Commercially available from Cappel Labs) was diluted into 0.1 M Na_2HCO_3, pH 9.6 to a concentration of 2 μg/ml. One hundred microliters of the antibody solution was coated onto each well of 96-well microtiter plates for 18 hours at 4° C. The wells were washed once with ELISA C buffer (PBS+0.05% Tween-20) followed by an incubation with 175 μl/well of ELISA B buffer (PBS+1% BSA+0.05% Tween-20) to block the wells. The wells were washed once with ELISA B buffer. One hundred microliters of 12.1 ng/ml or 24.3 ng/ml of tetrameric PDGF β-receptor analog protein diluted in ELISA B was added to each well and the plates were incubated for 2 hours at 37° C. Unbound protein was removed from the wells by two washes with ELISA C. ^{125}I-labeled PDGF-BB_{Tyr} (Example 18.F.) was serially diluted into binding media (25 mM HEPES, pH 7.2, 0.25% rabbit serum albumin diluted in HAMs F-12 medium (GIBCO-BRL)), and 100 μl of the dilutions were added to the wells. The plates were incubated for two hours at room temperature. The unbound ^{125}I-labeled PDGF-BB was removed, and the wells were washed three times with binding media. Following the last wash, 100 μl of 0.1 M citrate, pH 2.5 was added to each well. After five minutes, the citrate buffer was removed, transferred to a tube and counted in a gamma counter. The counts reflect counts of ^{125}I-labeled PDGF-BB_{Tyr} bound by the receptor analog. Nonspecific binding for each concentration of ^{125}I-labeled PDGF-BB_{Tyr} was determined by a parallel assay wherein separate wells coated only with goat anti-human IgG were incubated with the ^{125}I-labeled PDGF-BB concentrations. Nonspecific binding was determined to be 2.8% of the total input counts per well and averaged 6% of the total counts bound.

Saturation binding assay on 12.1 and 24.3 ng/ml of tetrameric PDGF β-receptor analog gave half-maximal binding at 0.8 and 0.82 ng/ml ^{125}I-labeled PDGF-BB_{Tyr}, respectively. By Scatchard analysis (Scatchard, *Ann. NY Acad. Sci* 51: 660–667, 1949) these values were shown to correspond to a K_D(app) of 2.7×10^{-11} which agree with the published values for PDGF receptors on mammalian cells.

Example 14

Solid Phase Ligand Binding Assay Using the PDGF β-Receptor Analog

A. Solid Phase Radioreceptor Competition Binding Assay

In a solid phase radioreceptor competition binding assay (RRA), the wells of 96-well microtiter plates were coated with 100 μl of 2 μg/ml affinity purified goat anti-human IgG (Cappel Labs) diluted in 0.1 M Na_2HCO_3, pH 9.6. After an eighteen hour incubation at 4° C., the wells were washed once with ELISA C. The wells were blocked by incubation for 2 hours at 37° C. with 175 μl/well ELISA B. The wells were washed once with ELISA B then incubated for 2 hours at 37° C. with 50 ng of tetrameric PDGF β-receptor analog diluted in ELISA B. The unbound receptor was removed, and the test wells were incubated with increasing concentrations of serially diluted, unlabeled PDGF-BB (diluted in binding media. Following a two hour incubation at room temperature, the wells were washed three times with binding media. One hundred microliters of 5 ng/ml ^{125}I-labeled PDGF-BB_{Tyr} (Example 18.F.) was added to each well, and the plates were incubated for an additional two hours at room temperature. The wells were washed three times with binding media followed by a 5 minute incubation with 100 μl/well of 0.1 M citrate, pH 2.5. The samples were harvested and counted in a gamma counter.

Radioreceptor assay (RRA) competition binding curves were generated for PDGF β-receptor analog protein plated at 48.6 ng/ml. The sensitivity of the assays is 1 ng/ml of PDGF-BB, with 8 ng/ml giving 50% inhibition in ^{125}I-PDGF-BB binding, and a working range between 1 and 32 ng/ml of PDGF-BB. The values were similar to those obtained using monolayers of SK-5 cells in an RRA.

B. Use of Tetrameric PDGF β-Receptor Analogs As Antagonists for PDGF-Stimulated Mitogenesis.

A tetrameric PDGF β-receptor analog, purified as described in Example 13, was analyzed for the ability to neutralize PDGF-stimulated mitogenesis in mouse 3T3 cells. Increasing amounts of the purified tetrameric PDGF β-receptor analog were mixed with 5 ng of PDGF. The mixtures were then added to cultures of mouse 3T3 cells. The ability of the PDGF to stimulate a mitogenic response, as measured by the incorporation of ^{3}H-thymidine, was determined essentially as described (Raines and Ross, *Methods in Enzymology* 109: 749–773, 1985, which is incorporated by reference herein). The tetrameric PDGF β-receptor analog demonstrated a dose response inhibition of PDGF-BB-stimulated ^{3}H-thymidine incorporation, while having essentially no effect on PDGF-AA- and PDGF-AB-stimulated ^{3}H-thymidine incorporation.

C. Binding of Tetrameric PDGF β-receptor Analog to Immobilized PDGF.

A tetrameric PDGF β-receptor analog, purified as described in Example 13, was analyzed for its ability to bind to immobilized PDGF. PDGF-BB (100 ng/ml) was coated onto wells a 96-well microtiter plate, and the plates were incubated 18 hours at 4° C. followed by one wash with ELISA C buffer. The wells were incubated for 2 hours 37° C. with ELISA B buffer to block the wells. Increasing concentrations of ^{125}I-labeled tetrameric PDGF β-receptor analog, diluted in binding media, was added to the wells for two hours at room temperature. The wells were washed four times with ELISA C buffer to remove unbound receptor analog. One hundred microliters of 1 M H_2SO_4 was added to each well and the plates were incubated for five minutes at room temperature. The solution was then harvested and transferred to tubes to be counted in a gamma counter. Nonspecific binding was determined to be less than 10% of the total counts bound.

A receptor competition binding assay was developed using this assay format. The assay was carried out as described above, and simultaneous to the addition of the ^{125}I-labeled tetrameric PDGF β-receptor analog, increasing amounts of PDGF-AA, AB or BB were added to the PDGF-BB coated wells. Under these conditions, only PDGF-BB was found to significantly block the binding of the labeled PDGF β-receptor analog to the immobilized PDGF-BB.

Example 15

Construction and Expression of PDGFα-R Analogs in Cultured Mouse Myeloma Cells

A. Construction of an optimized PDGFα-R cDNA

The PDGF α-receptor coding region was optimized for expression in mammalian cells as follows. The 5' end of the cDNA was modified to include an optimized Kozak consensus translation initiation sequence (Kozak, *Nuc. Acids Res.* 12: 857–872, 1984) and Eco RI and Bam HI sites just 5' of the initiation methionine codon. Oligonucleotides ZC2181, ZC2182, ZC2183 and ZC2184 (Sequence ID Numbers 23, 24, 25 and 26, respectively; Table 1) were designed to form, when annealed, an adapter having an Eco RI adhesive end, a Bam HI restriction site, a sequence encoding a Kozak consensus sequence 5' to the initation methionine codon, a mammalian codon optimized sequence encoding amino acids 1–42 of FIGS. 11A–11D, and an Eco RI adhesive end that destroys the Eco RI site within the PDGFα-R coding sequence. The adapter also introduced a diagnostic Cla I site 3' to the initiation methionine codon. Oligonucleotides ZC2181, ZC2182, ZC2183 and ZC2184 were kinased, annealed and ligated. Plasmid pα17B was linearized by partial digestion with Eco RI. The linearized pα17B was ligated with the ZC2181/2C2182/ZC2183/ZC2184 oligonucleotide adapter, and the ligation mixture was transformed into *E. coli*. Plasmid DNA prepared from the transformants was analyzed by restriction analysis and a positive clone having the oligonucleotide adapter in the correct orientation was digested with Eco RI and Pst I to isolate the 1.6 kb fragment. This fragment was subcloned into Eco RI+Pst I-linearized M13mp19. The resultant phage clone was designated 792-8. Single-stranded 792-8 DNA was sequenced to confirm the orientation of the adapter.

A fragment encoding the ligand-binding domain of the PDGF α-receptor (PDGFα-R) was then generated as follows. Restriction sites and a splice donor sequence were introduced at the 3' end of the PDGFα-R extracellular domain by PCR amplification of the 792-8 DNA and oligonucleotides ZC2311 and ZC2392 (Sequence ID Numbers 27 and 30, Table 1). Oligonucleotide ZC2311 is a sense primer encoding nucleotides 1470 to 1489 of FIG. 11B. Oligonucleotide ZC2392 is an antisense primer that encodes nucleotides 1759 to 1776 of FIG. 11B followed by a splice donor and Xba I and Hind III restriction sites. The 792-8 DNA was amplified using manufacturer recommended (Perkin Elmer Cetus, Norwalk, Conn.) conditions and the GeneAmp™ DNA amplification reagent kit (Perkin Elmer Cetus), and blunt-ended 329 bp fragment was isolated. The blunt-end fragment was digested with Nco I and Hind III and ligated with Sma I-digested pUC18. A plasmid having an insert with the Nco I site distal to the Hind III site present in the pUC18 polylinker was designated pUC18 Sma-PCR Nco HIII #13. The Hind III site present in the insert was not regenerated upon ligation with the linearized pUC18. Plasmid pUC18 Sma-PCR Nco I HI #13 was digested with Nco I and Hind III to isolate the 355 bp PDGFα-R containing fragment encoding PDGFαR. oligonucleotides ZC2351 and ZC2352 (Table 1; Sequence ID Numbers 28 and 29) were kinased and annealed to form an Sst I-Nco I adapter encoding an internal Eco RI site and a Kozak consensus translation initiation site. The 355 bp Nco I-Hind III fragment, the ZC2351/ZC2352 adapter and a 1273 bp Nco I fragment comprising the extracellular domain of of PDGF α-R derived from 792-8 were ligated with Hind III+SstI-digested pUC18, and transformed into *E. coli*. Plasmid DNA was isolated from the transformants and analyzed by restriction analysis. None of the isolates contained the 1273 bp Nco I fragment. A plasmid containing the Nco I-Hind III fragment and the ZC2351/ZC2352 adapter was designated pUC18 Hin Sst Δ Nco #46. Plasmid pUC18 Hin Sst Δ Nco #46 was linearized by digestion and joined by ligation with the 1273 bp Nco I fragment comprisising the extracellular domain of the PDGFα-R from clone α18 R-19. The ligations were transformed into *E. coli*, and plasmid DNA was isolated from the transformants. Analysis of the plasmid DNA showed that only clones with the Nco I fragment in the wrong orientation were isolated. A clone having the Nco I fragment in the wrong orientation was digested with Nco I, religated and transformed into *E. coli*. Plasmid DNA was isolated from the transformants and was analyzed by restriction analysis. A plasmid having the Nco I insert in the correct orientation was digested to completion with Hind III and partially digested with Sst I to isolate the 1.6 kb fragment comprising the extracellular domain of the PDGFα-R preceded by a consensus initiation sequence (Kozak, ibid.) and followed by a splice donor site.

B. Construction of pPAB7

The DNA sequence encoding the extracellular domain of the PDGFα-R was joined to the immunoglobulin μ enhancer-promoter and to a DNA sequence encoding an immunoglobulin light chain constant region. The immunoglobulin μ enhancer-promoter was obtained from plasmid pJH1 which was derived from plasmid PICμPRE8 (Example 12.A.) by digestion with Eco RI and Sst I to isolate the 2.2 kb fragment comprising the immunoglobulin enhancer and heavy chain variable region promoter. The 2.2 kb Sst I-Eco RI fragment was ligated with Sst I+Eco RI-linearized pUC19. The resulting plasmid, designated pJH1, contained the immunoglobulin enhancer and heavy chain variable region promoter immediately 5' to the pUC19 linker sequences. Plasmid pJH1 was linearized by digestion with Sst I and Hind III and joined with the 1.6 kb partial Sst I-Hind III fragment containing the PDGFα-R extracellular domain sequences. The resulting plasmid having the immunoglobulin μ enhancer-promoter joined to the PDGFα-R extracellular domain was designated pPAB6. Plasmid pSDL112 was digested with Hind III to isolate the 1.2 kb fragment encoding the immunoglobulin light chain constant region (Cκ). The 1.2 kb Hind III fragment was ligated with Hind III-linearized pPAB6. A plasmid having the C_κ sequence in the correct orientation was desginated pPAB7.

C. Construction of pPAB9

The partial Sst I-Hind III fragment encoding the extracellular domain of the PDGFα-R was joined to the immunoglobulin heavy chain constant region. For convenience, the internal Xba I site in plasmid pJH1 was removed by digestion with Xba I, blunt-ending with T4 DNA polymerase, and religation. A plasmid which did not contain the internal Xba I site, but retained the Xba I site in the polylinker was desginated 11.28.3.6. Plasmid 11.28.3.6 was linearized by digestion with Sst I and Xba I. Plasmid pPAB6 was digested to completion with Hind III and partially digested with Sst I to isolate the 1.6 kb Sst I-Hind III fragment containing the PDGFα-R extracellular domain. Plasmid pφ5V$_\mu$huC$_\gamma$1M-neo (Example 12.C.) was digested with Hind III and Xba I to isolate the 6.0 kb fragment encoding the immunoglobulin heavy chain constant region (huC$_\gamma$1M). The Sst I-Hind III-linearized 11.28.3.6, the 1.6 kb Sst I-Hind III PDGFα-R fragment and the 6.0 kb Hind III-Xba I huC$_\gamma$1M fragment were ligated to form plasmid pPAB9.

D. Expression of pPAB9 in Mammalian Cells

Bgl II-linearized pPAB7 and Pvu I-linearized pPAB9 were cotransfected with Pvu I-linearized pH16 into SP2/0-Ag14 cells by electroporation. Transfected cells were initially selected in growth medium containing 50 nM methotrexate and were subsequently amplified in a growth medium containing 100 μM methotrexate. Media from drug

resistant clones were tested for the presence of secreted PDGF α-receptor analogs by enzyme-linked immunosorbant assay (ELISA). Ninety-six well assay plates were prepared by incubating 100 μl of 1 μg/ml monoclonal antibody 292.1.8 which is specific for the PDGF α-receptor diluted in phosphate buffered saline (PBS; Sigma] overnight at 4° C. Excess antibody was removed by three washes with 0.5% Tween 20 in PBS. One hundred microliters of spent media was added to each well, and the plates were incubated for one hour at 4° C. Unbound proteins were removed by eight washes with 0.5% Tween 20 in PBS. One hundred microliters of peroxidase-conjugated goat anti-human IgG heavy chain antibody (diluted 1:1000 in a solution containing 5% chicken serum (GIBCO-BRL)+0.5% Tween 20 in PBS) was added to each well, and the plates were incubated for one hour at 4° C. One hundred microliters of chromophore (100 μl ABTS [2,2'-Azinobis(3-ethylbenzthiazoline sulfonic acid] diammonium salt; Sigma]+1 μl 30% H_2O_2+12.5 ml citrate/phosphate buffer [9.04 g/l citric acid, 10.16 g/l Na_2HPO_4]) was added to each well, and the wells were incubated for 30 minutes at room temperature. The samples were measured at 405 nm. The results of the assay showed that the PDGF α-receptor analogs secreted by the transfectants contained an immunoglobulin heavy chain.

Analysis of spent media from transfected cells by Northern analysis, Western analysis and by radioimmunoprecipitation showed that the transfectants did not express a PDGF α-receptor analog from the pPAB7 construction. Transfectants were subsequently treated as containing only pPAB9.

Drug resistant clones were also tested for the presence of secreted PDGF α-receptor analogs by immunoprecipitation. For each clone, approximately one million drug resistant transfectants were grown in DMEM lacking cysteine+2% calf serum for 18 hours at 37° C. in the presence of 50 μCi ^{35}S-cysteine. The spent media was harvested from the labeled cells and 250 μl of medium from each clone was assayed for binding to the anti-PDGF α-receptor antibody 292.18. Monoclonal antibody 292.18 diluted in PBS was added to each sample to a final concentration of 2.5 μg per 250 μl spent media. Five microliters of rabbit anti-mouse Ig diluted in PBS was added to each sample, and the immunocomplexes were precipitated by the addition of 50 μl 10% fixed Staph A (weight/volume in PBS). The immunocomplexes were analyzed on 8% SDS-polyacrylamide gels followed by autoradiography overnight at −70° C. The results of the immunoprecipitation showed that the PDGF α-receptor analog secreted by the transfectants was bound by the anti-PDGF α-receptor antibody. The combined results of the ELISA and immunoprecipitation assays showed that the PDGF α-receptor analog secreted by the transfectants contained both the PDGF α-receptor ligand-binding domain and the human heavy chain.

Spent medium from drug-resistant clones were tested for their ability to bind PDGF in a competition binding assay essentially as described in Example 12.C. The results of the assay showed that the transfectants secreted a PDGF α-receptor analog capable of binding PDGF-AA. A clone containing the pPAB9 was designated 3.17.1.57.

E. Co-expression of pPAB7 and pPAB9 in Mammalian Cells

Bgl II-linearized pPAB7 and Bam HI-linearized pICneo were cotransfected into clone 3.17.1.57, and transfected cells were selected in the presence of neomycin. Media from drug resistant cells were assayed for the presence of immunoglobulin heavy chain, immunoglobulin light chain and the PDGF α-receptor ligand-binding domain by ELISA essentially as described above. Briefly, ninety-six well assay plates were prepared by incubating 100 μl of 1 μg/ml goat anti-human IgG Fc antibody (Sigma) or 100 μl of 1 μg/ml 292.18 overnight at 4° C. Excess antibody was removed by three washes with 0.5% Tween 20 in PBS. One hundred microliters of spent media was added to each well of each plate, and the plates were incubated for one hour at 4° C. Unbound proteins were removed by eight washes with 0.5% Tween 20 in PBS. One hundred microliters of peroxidase-conjugated goat anti-human IgG antibody (diluted 1:1000 in a solution containing 5% chicken serum (GIBCO-BRL)+0.5% Tween 20 in PBS) was added to each well of the plate coated with the anti-Fc antibody, and 100 μl of peroxidase-conjugated goat anti human kappa antibody (diluted 1:1000 in a solution containing 5% chicken serum (GIBCO-BRL)+0.5% Tween 20 in PBS) was added to each well of the plate coated with 292.18. The plates were incubated for one hour at 4° C. One hundred microliters of chromophore (100 μl ABTS (2,2'-Azinobis(3-ethylbenz-thiazoline sulfonic acid) diammonium salt; Sigma]+1 μl 30% H_2O_2+12.5 ml citrate/phosphate buffer [9.04 g/l citric acid, 10.16 g/l Na_2HPO_4]) was added to each well of each plate, and the plates were incubated to 30 minutes at room temperature. The samples were measured at 405 nm, the wavelength giving maximal absorbance of the chromogenic substrate, to identify clones having absorbances higher than background indicating the presence of immunoglobulin heavy chain. Clones that gave positive results in both ELISA assays (showing that the clones produced proteins containing heavy chain regions, light chain constant regions and the PDGF α-receptor ligand-binding region) were selected for further characterization.

Drug resistant clones that were positive for both ELISA assays were subsequently tested for the presence of secreted PDGF α-receptor analogs by immunoprecipitation. For each positive clone, approximately one million drug resistant transfectants were grown in DMEM lacking cysteine+2% calf serum for 18 hours at 37° C. in the presence of 50 μCi ^{35}S-cysteine. The spent media was harvested from the labeled cells and 250 μl of medium from each clone was assayed for binding to monoclonal antibody 292.18. Monoclonal antibody 292.18 diluted in PBS was added to each sample to a final concentration of 2.5 μg. Five microliters of rabbit anti-mouse Ig diluted in PBS was added to each sample and the immunocomplexes were precipitated by the addition of 50 μl 10% fixed Staph A (weight/volume in PBS). The immunocomplexes were analyzed on 8% SDS-polyacrylamide gels followed by autoradiography overnight at −70° C. The results of the immunoprecipitation showed that the PDGF α-receptor analog secreted by the transfectants was bound by the anti-PDGF α-receptor antibody. The combined results of the ELISA and immunoprecipitation assays showed that the PDGF α-receptor analog secreted by the transfectants contained the PDGF α-receptor ligand-binding domain, the human heavy chain and the human light chain constant region. A clone that secreted a PDGF α-receptor analog that was positive for both the above-described ELISA assays and the immunoprecipitation assay was designated 5.6.2.1.

Example 16

Purification and characterization of PDGF α-Receptor Analogs

A. Purification of PDGF α-Receptor Analogs From Clone 3.17.1.57

The PDGF α-Receptor analog was purified from the conditioned culture media of clone 3.17.1.57 by cycling cell-conditioned medium over an immunoaffinity column composed of monoclonal antibody 292.18 bound to a CNBr-

activated Sepharose 4B resin, which is specific for the PDGF α-receptor. The column was washed with PBA, then eluted with 0.1 M citrate, pH 3.0. The peak column fractions containing the α-receptor were pooled, neutralized to pH 7.2 by the addition of 2 M Tris, pH 7.4, then passed over a protein-A Sepharose column. This column was washed sequentially with PBS, then with 0.1 M citrate, pH 5.0. The PDGF α-receptor analog was then eluted with 0.1 M citrate, pH 3.0. The peak eluate fractions were pooled, and glycerol was added to a final concentration of lot. The sample was concentrated on a centriprep 10 concentrator (Amicon). The PDGF α-receptor analog purified from clone 3.17.1.57 was termed a "dimeric PDGF α-receptor analog".

B. Purification of PDGF α-Receptor Analogs From Clone 5.6.2.1

The PDGF α-receptor analog was purified from the conditioned culture media of clone 5.6.2.1 by cycling cell-conditioned medium over the immunoaffinity column described above. The column was washed with PBS then eluted with 0.1 M citrate, pH 3.0. The peak column fractions containing the α-receptor were pooled, neutralized to pH 7.2 by the addition of 2 M Tris (what pH 7.4), then passed over a protein-A sepharose column. This column was washed sequentially with PBS then with 0.1 M citrate, pH 5.0. The PDGF α-receptor analog was then eluted with 0.1 M citrate, pH 3.0. The peak eluate fractions were pooled and glycerol was added to a final concentration of 10%. The sample was concentrated on a centriprep 10 concentrator. The PDGF α-receptor analogs purified from clone 5.6.2.1 was termed a "tetrameric PDGF α-receptor analog".

Example 17

A. Use of the PDGF α-receptor Analogs in Ligand Binding Studies

Purified tetrameric PDGF α-receptor analog and purified dimeric PDGF α-receptor analog were compared to monolayers of a control cell line of canine kidney epithelial cells, which do not naturally express the PDGF α-receptor, transfected with the human PDGF α-receptor cDNA for ligand binding activity. The dissociation constant (Kd) of the receptor preparations was determined by saturation binding and subsequent Scatchard analysis.

Ligand binding of the purified PDGF α-receptor analogs was determined using a solid phase binding assay. Affinity-purified goat anti-human IgG was diluted to a concentration of 2 μg/ml in 0.1 M Na_2HCO_3, pH 9.6 and 100 μl/well of the solution was used to coat 96-well microtiter plates for 18 hours at 4° C. Excess antibody was removed from the wells with one wash with ELISA C buffer (PBS, 0.05% Tween-20). The plates were incubated with 175 μl/well of ELISA B buffer (PBS, 1% BSA, 0.05% Tween-20) to block the wells, followed by two washes with ELISA C buffer. One hundred microliters of 50 ng/ml PDGF α-receptor analog (dimeric or tetrameric) diluted in ELISA buffer B was added to each well and the plates were incubated over night at 4° C. Unbound protein was removed from the wells with two washes with ELISA buffer B. ^{125}I-labeled PDGF-AA was serially diluted in binding media (Hams F-12, 25 mM HEPES pH 7.2, 0.25% rabbit serum albumin), and 100 μl of each dilution was added to the wells. The samples were incubated for two hours at room temperature. Unbound ^{125}I-labeled PDGF-AA was removed with three washes with binding media. One hundred microliters of 0.1 M citrate, pH 2.5 was added to each well, and the plates were incubated for five minutes. After the incubation, the citrate buffer was removed and transferred to a tube for counting in a gamma counter. Nonspecific binding for each concentration of ^{125}I-labeled PDGF-AA was determined by a parallel assay wherein separate wells coated only with goat anti-human IgG were incubated with the ^{125}I-labeled PDGF-AA samples.

A saturation binding assay was performed on alpha T-7 cells transfected with the PDGF α-receptor. The cells were grown to confluency in 24-well culture plates. The cells were washed one time with binding media. Iodinated PDGF-AA was serially diluted in binding media. One milliliter of each dilution was added to the wells, and the plates were incubated for 3 hours at 4° C. Unbound ^{125}I-labeled PDGF-AA was removed and the cells were washed three times with binding media. PBS containing 1% Triton X-100 was added to the cells for 5 minutes. The extracts were harvested and counted in a gamma counter. Nonspecific binding was determined at a single concentration of ^{125}-labeled PDGF-AA using a 500-fold excess PDGF-BB.

The dissociation constants determined by Scatchard analysis (ibid.) of the saturation binding assays for the tetrameric PDGF α-receptor analog, dimeric PDGF α-receptor analog and the control cells (Table 5).

Table 5

Dissociation Constants for the Tetrameric PDGF α-Receptor, the Dimeric PDGF α-receptor and control cells Transfected with the PDGF α-receptor

TABLE 5

Dissociation Constants for the Tetrameric PDGF α-Receptor, the Dimeric PDGF α-receptor and control cells Transfected with the PDGF α-receptor	
Receptor	Kd
Tetrameric PDGF α-receptor analog	1.6×10^{-11}
Dimeric PDGF α-receptor analog	8.51×10^{-11}
Control cells [PDGF α-receptor]	3.7×10^{-11}

A solid-phase competition binding assay was established using the tetrameric PDGF α-receptor analog. Ninety six-well microliter plates were coated with goat anti-human IgG (2 μg/ml), the wells blocked with ELISA B buffer, 50 ng/ml of purified tetrameric PDGF α-receptor analog diluted in binding media was added, and the plates were incubated two hours at room temperature. Unbound receptor was removed and the wells were washed with binding media. The plates were incubated for two hours at room temperature with increasing concentrations of either PDGF-AA or PDGF-BB diluted in binding media. The wells were washed, then incubated for two hours at room temperature with 3 ng/ml ^{125}I-labeled PDGF-AA diluted in binding media. Unbound labeled PDGF-AA was removed, the wells were subsequently washed with binding media, and the bound, labeled PDGF-AA was harvested by the addition of 0.1 M citrate, pH 2.5, as described for the saturation binding studies. PDGF-AB, PDGF-AA and PDGF-BB were found to compete for receptor binding with ^{125}I-PDGF-AA.

B. Use of Tetrameric PDGF α-Receptor Analogs As Antagonists for PDGF-Stimulated Mitogenesis.

A dimeric PDGF α-receptor analog, purified as described in Example 16.B., was analyzed for the ability to neutralize PDGF-stimulated mitogenesis in mouse 3T3 cells. Increasing amounts of the purified tetrameric PDGF α-receptor analog were mixed with PDGF-AA, -AB or -BB ranging 0.6 to 5 ng. The mixtures were then added to cultures of confluent mouse 3T3 cells. The ability of the PDGF to stimulate a mitogenic response, as measured by the incorporation of ^{3}H-thymidine, was determined essentially as

described (Raines and Ross, *Methods in Enzymology* 109: 749–773, 1985, which is incorporated by reference herein). The dimeric PDGF α-receptor analog demonstrated a dose response inhibition of PDGF-stimulated ^{3}H-thymidine incorporation for all three isoforms of PDGF.

C. Inverse Ligand-Receptor Radioreceptor Assay

An inverse ligand-receptor radioreceptor assay was designed to screen for the presence of PDGF-BB, PDGF-BB binding proteins, PDGF-BB related molecules, and PDGF-β receptor antagonists in test samples. PDGF-BB (100 ng/ml) was coated onto the walls of 96-well microtiter plates, and the plates were incubated at 4° C. for 16 hours. The wells were washed once with ELISA C buffer and then incubated with ELISA B buffer to block the nonspecific binding sites. To the wells were added 50 μl of either PDGF standard or a test sample and 50 μl of ^{125}I-labeled tetrameric PDGF β-receptor analog. The samples were incubated for one hour at room temperature. The wells were washed once with ELISA C buffer, and 0.1 M citrate, pH 2.5 containing 1% NP-40 was added to each well to disrupt the ligand-receptor analog bond and elute the bound receptor analog. The acid wash was collected and counted in a gamma counter. The presence of PDGF or a molecule which mimics PDGF or otherwise interferes with the binding of the well-bound PDGF-BB with its receptor will cause a decrease in the binding of the radiolabeled tetrameric PDGF β-receptor. Using this assay, PDGF-BB was found to inhibit receptor binding while PDGF-AA and PDGF-AB caused no significant decrease in receptor binding.

Example 18

Assay Methods

A. Preparation of Nitrocellulose Filters for Colony Assay

Colonies of transformants were tested for secretion of PDGF β-receptor analogs by first growing the cells on nitrocellulose filters that had been laid on top of solid growth medium. Nitrocellulose filters (Schleicher & Schuell, Keene, N.H.) were placed on top of solid growth medium and were allowed to be completely wetted. Test colonies were patched onto the wetted filters and were grown at 30° C. for approximately 40 hours. The filters were then removed from the solid medium, and the cells were removed by four successive rinses with Western Transfer Buffer (Table 4). The nitrocellulose filters were soaked in Western Buffer A (Table 4) for one hour at room temperature on a shaking platform with two changes of buffer. Secreted PDGFβ-R analogs were visualized on the filters described below.

B. Preparation of Protein Blot Filters

A nitrocellulose filter was soaked in Western Buffer A (Table 4) without the gelatin and placed in a Minifold (Schleicher & Schuell, Keene, N.H.). Five milliliters of culture supernatant was added without dilution to the Minifold wells, and the liquid was allowed to pass through the nitrocellulose filter by gravity. The nitrocellulose filter was removed from the minifold and was soaked in Western Buffer A (Table 3) for one hour on a shaking platform at room temperature. The buffer was changed three times during the hour incubation.

C. Preparation of Western Blot Filters

The transformants were analyzed by Western blot, essentially as described by Towbin et al. (*Proc. Natl. Acad. Sci. USA* 76: 4350–4354, 1979) and Gordon et al. (U.S. Pat. No. 4,452,901). Culture supernatants from appropriately grown transformants were diluted with three volumes of 95% ethanol. The ethanol mixtures were incubated overnight at −70° C. The precipitates were spun out of solution by centrifugation in an SS-24 rotor at 18,000 rpm for 20 minutes. The supernatants were discarded and the precipitate pellets were resuspended in 200 μl of dH_2O. Two hundred microliters of 2× loading buffer (Table 4) was added to each sample, and the samples were incubated in a boiling water bath for 5 minutes.

The samples were electrophoresed in a 15% sodium dodecylsulfate polyacrylamide gel under non-reducing conditions. The proteins were electrophoretically transferred to nitrocellulose paper using conditions described by Towbin et al. (ibid.). The nitrocellulose filters were then incubated in Western Buffer A (Table 4) for 75 minutes at room temperature on a platform rocker.

D. Processing the Filters for Visualization with Antibody

Filters prepared as described above were screened for proteins recognized by the binding of a PDGF β-receptor specific monoclonal antibody, designated PR7212. The filters were removed from the Western Buffer A (Table 4) and placed in sealed plastic bags containing a 10 ml solution comprising 10 μg/ml PR7212 monoclonal antibody diluted in Western Buffer A. The filters were incubated on a rocking platform overnight at 4° C. or for one hour at room temperature. Excess antibody was removed with three 15-minute washes with Western Buffer A on a shaking platform at room temperature.

Ten microliters biotin-conjugated horse anti-mouse antibody (Vector Laboratories, Burlingame, Calif.) in 20 ml Western Buffer A was added to the filters. The filters were re-incubated for one hour at room temperature on a platform shaker, and unbound conjugated antibody was removed with three fifteen-minute washes with Western Buffer A.

The filters were pre-incubated for one hour at room temperature with a solution comprising 50 μl Vectastain Reagent A (Vector Laboratories) in 10 ml of Western Buffer A that had been allowed to incubate at room temperature for 30 minutes before use. The filters were washed with one quick wash with distilled water followed by three 15-minute washes with Western Buffer B (Table 4) at room temperature. The Western Buffer B washes were followed by one wash with distilled water.

During the preceding wash step, the substrate reagent was prepared. Sixty mg of horseradish peroxidase reagent (Bio-Rad, Richmond, Calif.) was dissolved in 20 ml HPLC grade methanol. Ninety milliliters of distilled water was added to the dissolved peroxidase followed by 2.5 ml 2 M Tris, pH 7.4 and 3.8 ml 4 M NaCl. One hundred microliters of 30% H_2O_2 was added just before use. The washed filters were incubated with 75 ml of substrate and incubated at room temperature for 10 minutes with vigorous shaking. After the 10 minute incubation, the buffer was changed, and the filters were incubated for an additional 10 minutes. The filters were then washed in distilled water for one hour at room temperature. Positives were scored as those samples which exhibited coloration.

E. Processing the Filters For Visualization with an Anti-Substance P Antibody

Filters prepared as described above were probed with an anti-substance P antibody. The filters were removed from the Western Buffer A and rinsed with Western transfer buffer, followed by a 5-minute wash in phosphate buffered saline (PBS, Sigma, St. Louis, Mo.). The filters were incubated with a 10 ml solution containing 0.5 M 1-ethyl-3-3-dimethylamino propyl carbodiimide (Sigma) in 1.0 M NH_4Cl for 40 minutes at room temperature. After incubation, the filters were washed three times, for 5 minutes per wash, in PBS. The filters were blocked with 2% powdered milk diluted in PBS.

The filters were then incubated with a rat anti-substance P monoclonal antibody (Accurate Chemical & Scientific Corp., Westbury, N.Y.). Ten microliters of the antibody was diluted in 10 ml of antibody solution (PBS containing 20% fetal calf serum and 0.5% Tween-20). The filters were incubated at room temperature for 1 hour. Unbound antibody was removed with four 5-minute washes with PBS.

The filters were then incubated with a biotin-conjugated rabbit anti-rat peroxidase antibody (Cappel Laboratories, Melvern, Pa.). The conjugated antibody was diluted 1:1000 in 10 ml of antibody solution for 2 hours at room temperature. Excess conjugated antibody was removed with four 5-minute washes with PBS.

The filters were pre-incubated for 30 minutes at room temperature with a solution containing 50 μl Vectastain Reagent A (Vector Laboratories) and 50 μl Vectastain Reagent B (Vector Laboratories) in 10 ml of antibody solution that had been allowed to incubate for 30 minutes before use. Excess Vectastain reagents were removed by four 5-minute washes with PBS.

During the preceding wash step, the substrate reagent was prepared. Sixty milligrams of horseradish peroxidase reagent (Bio-Rad Laboratories, Richmond, Calif.) was dissolved in 25 ml of HPLC grade methanol. Approximately 100 ml of PBS and 200 μl H_2O_2 were added just before use. The filters were incubated with the substrate reagent for 10 to 20 minutes. The substrate was removed by a vigorous washing distilled water.

F. Iodination of PDGF-BB.

A PDGF-BB mutant molecule having a tyrosine replaing the phenylalanine at position 23 ($PDGF\text{-}BB_{Tyr}$) was iodinated and subsequently purified, using a purification method which produces 125I-labeled PDGF-BB with a higher specific activity than primary-labeled material and which was found to substantially decrease the nonspecific binding component. The $PDGF\text{-}BB_{Tyr}$ was labeled using the Iodobead method (Pierce Chemical). The labeled protein was gel filtered over a C-25 desalting column (Pharmacia LKB Technologies) equilibrated with 10 mM acetic acid, 0.25% gelatin and 100 mM NaCl. The peak fractions were pooled and pH adjusted to 7.2 by the addition of Tris-base. The labeled mixture was chromatographed over an affinity column composed of PDGF β-receptor analog protein coupled to CnBr-activated Sepharose (Pharmacia LKB Technologies, Inc.). The column was washed with phosphate buffered saline and eluted with 0.1 M citrate, pH 2.5 containing 0.25% gelatin. The peak eluate fractions were pooled and assayed by ELISA to determine the PDGF-BB concentration.

Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, it will be evident that certain changes and modifications may be practiced within the scope of the appended claims.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(iii) NUMBER OF SEQUENCES: 36

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4465 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens
(D) DEVELOPMENTAL STAGE: Adult
(F) TISSUE TYPE: Skin
(G) CELL TYPE: fibroblasts

(vii) IMMEDIATE SOURCE:
(B) CLONE: pR-rX1

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 354..3671
(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CGCTCAGCCC TGCTGCCCAG CACGAGCCTG TGCTCGCCCT GCCCAACCCA GACAGCCAGA 60

CCCAGGGCGG CCCCTCTCGC GGCTCTCCTC CTCCCGAAGG ATGCCTGGGG AGTGAGCCGA 120

AGCTGGGCGC TCCTCTCCCC TACAGCAGCC CCCTTCCTCC ATCCCTCTGT TCTCCTGAGC 180

-continued

```
CTTCAGGAGC CTGCACCAGT CCTGCCTGTC CTTCTACTCA GCTGTTACCC ACTCTGGGAC   240

CAGCAGTCTT TCTGATAACT GGGAGAGGGC AGTAAGGAGG ACTTCCTGGA GGGGGTGACT   300

GTCCAGAGCC TGGAACTGTG CCCACACCAG AAGCCATCAG CAGCAAGGAC ACC ATG      356
                                                           Met
                                                            1

CGG CTT CCG GGT GCG ATG CCA GCT CTG GCC CTC AAA GGC GAG CTG CTG      404
Arg Leu Pro Gly Ala Met Pro Ala Leu Ala Leu Lys Gly Glu Leu Leu
          5                  10                  15

TTG CTG TCT CTC CTG TTA CTT CTG GAA CCA CAG ATC TCT CAG GGC CTG      452
Leu Leu Ser Leu Leu Leu Leu Leu Glu Pro Gln Ile Ser Gln Gly Leu
      20                  25                  30

GTC GTC ACA CCC CCG GGG CCA GAG CTT GTC CTC AAT GTC TCC AGC ACC      500
Val Val Thr Pro Pro Gly Pro Glu Leu Val Leu Asn Val Ser Ser Thr
  35                  40                  45

TTC GTT CTG ACC TGC TCG GGT TCA GCT CCG GTG GTG TGG GAA CGG ATG      548
Phe Val Leu Thr Cys Ser Gly Ser Ala Pro Val Val Trp Glu Arg Met
 50                  55                  60                  65

TCC CAG GAG CCC CCA CAG GAA ATG GCC AAG GCC CAG GAT GGC ACC TTC      596
Ser Gln Glu Pro Pro Gln Glu Met Ala Lys Ala Gln Asp Gly Thr Phe
                 70                  75                  80

TCC AGC GTG CTC ACA CTG ACC AAC CTC ACT GGG CTA GAC ACG GGA GAA      644
Ser Ser Val Leu Thr Leu Thr Asn Leu Thr Gly Leu Asp Thr Gly Glu
             85                  90                  95

TAC TTT TGC ACC CAC AAT GAC TCC CGT GGA CTG GAG ACC GAT GAG CGG      692
Tyr Phe Cys Thr His Asn Asp Ser Arg Gly Leu Glu Thr Asp Glu Arg
         100                 105                 110

AAA CGG CTC TAC ATC TTT GTG CCA GAT CCC ACC GTG GGC TTC CTC CCT      740
Lys Arg Leu Tyr Ile Phe Val Pro Asp Pro Thr Val Gly Phe Leu Pro
     115                 120                 125

AAT GAT GCC GAG GAA CTA TTC ATC TTT CTC ACG GAA ATA ACT GAG ATC      788
Asn Asp Ala Glu Glu Leu Phe Ile Phe Leu Thr Glu Ile Thr Glu Ile
 130                 135                 140                 145

ACC ATT CCA TGC CGA GTA ACA GAC CCA CAG CTG GTG GTG ACA CTG CAC      836
Thr Ile Pro Cys Arg Val Thr Asp Pro Gln Leu Val Val Thr Leu His
                 150                 155                 160

GAG AAG AAA GGG GAC GTT GCA CTG CCT GTC CCC TAT GAT CAC CAA CGT      884
Glu Lys Lys Gly Asp Val Ala Leu Pro Val Pro Tyr Asp His Gln Arg
             165                 170                 175

GGC TTT TCT GGT ATC TTT GAG GAC AGA AGC TAC ATC TGC AAA ACC ACC      932
Gly Phe Ser Gly Ile Phe Glu Asp Arg Ser Tyr Ile Cys Lys Thr Thr
         180                 185                 190

ATT GGG GAC AGG GAG GTG GAT TCT GAT GCC TAC TAT GTC TAC AGA CTC      980
Ile Gly Asp Arg Glu Val Asp Ser Asp Ala Tyr Tyr Val Tyr Arg Leu
     195                 200                 205

CAG GTG TCA TCC ATC AAC GTC TCT GTG AAC GCA GTG CAG ACT GTG GTC     1028
Gln Val Ser Ser Ile Asn Val Ser Val Asn Ala Val Gln Thr Val Val
 210                 215                 220                 225

CGC CAG GGT GAG AAC ATC ACC CTC ATG TGC ATT GTG ATC GGG AAT GAG     1076
Arg Gln Gly Glu Asn Ile Thr Leu Met Cys Ile Val Ile Gly Asn Glu
                 230                 235                 240

GTG GTC AAC TTC GAG TGG ACA TAC CCC CGC AAA GAA AGT GGG CGG CTG     1124
Val Val Asn Phe Glu Trp Thr Tyr Pro Arg Lys Glu Ser Gly Arg Leu
             245                 250                 255

GTG GAG CCG GTG ACT GAC TTC CTC TTG GAT ATG CCT TAC CAC ATC CGC     1172
Val Glu Pro Val Thr Asp Phe Leu Leu Asp Met Pro Tyr His Ile Arg
         260                 265                 270

TCC ATC CTG CAC ATC CCC AGT GCC GAG TTA GAA GAC TCG GGG ACC TAC     1220
Ser Ile Leu His Ile Pro Ser Ala Glu Leu Glu Asp Ser Gly Thr Tyr
     275                 280                 285
```

-continued

```
ACC TGC AAT GTG ACG GAG AGT GTG AAT GAC CAT CAG GAT GAA AAG GCC      1268
Thr Cys Asn Val Thr Glu Ser Val Asn Asp His Gln Asp Glu Lys Ala
290                 295                 300                 305

ATC AAC ATC ACC GTG GTT GAG AGC GGC TAC GTG CGG CTC CTG GGA GAG      1316
Ile Asn Ile Thr Val Val Glu Ser Gly Tyr Val Arg Leu Leu Gly Glu
                310                 315                 320

GTG GGC ACA CTA CAA TTT GCT GAG CTG CAT CGG AGC CGG ACA CTG CAG      1364
Val Gly Thr Leu Gln Phe Ala Glu Leu His Arg Ser Arg Thr Leu Gln
            325                 330                 335

GTA GTG TTC GAG GCC TAC CCA CCG CCC ACT GTC CTG TGG TTC AAA GAC      1412
Val Val Phe Glu Ala Tyr Pro Pro Pro Thr Val Leu Trp Phe Lys Asp
        340                 345                 350

AAC CGC ACC CTG GGC GAC TCC AGC GCT GGC GAA ATC GCC CTG TCC ACG      1460
Asn Arg Thr Leu Gly Asp Ser Ser Ala Gly Glu Ile Ala Leu Ser Thr
    355                 360                 365

CGC AAC GTG TCG GAG ACC CGG TAT GTG TCA GAG CTG ACA CTG GTT CGC      1508
Arg Asn Val Ser Glu Thr Arg Tyr Val Ser Glu Leu Thr Leu Val Arg
370                 375                 380                 385

GTG AAG GTG GCA GAG GCT GGC CAC TAC ACC ATG CGG GCC TTC CAT GAG      1556
Val Lys Val Ala Glu Ala Gly His Tyr Thr Met Arg Ala Phe His Glu
                390                 395                 400

GAT GCT GAG GTC CAG CTC TCC TTC CAG CTA CAG ATC AAT GTC CCT GTC      1604
Asp Ala Glu Val Gln Leu Ser Phe Gln Leu Gln Ile Asn Val Pro Val
            405                 410                 415

CGA GTG CTG GAG CTA AGT GAG AGC CAC CCT GAC AGT GGG GAA CAG ACA      1652
Arg Val Leu Glu Leu Ser Glu Ser His Pro Asp Ser Gly Glu Gln Thr
        420                 425                 430

GTC CGC TGT CGT GGC CGG GGC ATG CCC CAG CCG AAC ATC ATC TGG TCT      1700
Val Arg Cys Arg Gly Arg Gly Met Pro Gln Pro Asn Ile Ile Trp Ser
    435                 440                 445

GCC TGC AGA GAC CTC AAA AGG TGT CCA CGT GAG CTG CCG CCC ACG CTG      1748
Ala Cys Arg Asp Leu Lys Arg Cys Pro Arg Glu Leu Pro Pro Thr Leu
450                 455                 460                 465

CTG GGG AAC AGT TCC GAA GAG GAG AGC CAG CTG GAG ACT AAC GTG ACG      1796
Leu Gly Asn Ser Ser Glu Glu Glu Ser Gln Leu Glu Thr Asn Val Thr
                470                 475                 480

TAC TGG GAG GAG GAG CAG GAG TTT GAG GTG GTG AGC ACA CTG CGT CTG      1844
Tyr Trp Glu Glu Glu Gln Glu Phe Glu Val Val Ser Thr Leu Arg Leu
            485                 490                 495

CAG CAC GTG GAT CGG CCA CTG TCG GTG CGC TGC ACG CTG CGC AAC GCT      1892
Gln His Val Asp Arg Pro Leu Ser Val Arg Cys Thr Leu Arg Asn Ala
        500                 505                 510

GTG GGC CAG GAC ACG CAG GAG GTC ATC GTG GTG CCA CAC TCC TTG CCC      1940
Val Gly Gln Asp Thr Gln Glu Val Ile Val Val Pro His Ser Leu Pro
    515                 520                 525

TTT AAG GTG GTG GTG ATC TCA GCC ATC CTG GCC CTG GTG GTG CTC ACC      1988
Phe Lys Val Val Val Ile Ser Ala Ile Leu Ala Leu Val Val Leu Thr
530                 535                 540                 545

ATC ATC TCC CTT ATC ATC CTC ATC ATG CTT TGG CAG AAG AAG CCA CGT      2036
Ile Ile Ser Leu Ile Ile Leu Ile Met Leu Trp Gln Lys Lys Pro Arg
                550                 555                 560

TAC GAG ATC CGA TGG AAG GTG ATT GAG TCT GTG AGC TCT GAC GGC CAT      2084
Tyr Glu Ile Arg Trp Lys Val Ile Glu Ser Val Ser Ser Asp Gly His
            565                 570                 575

GAG TAC ATC TAC GTG GAC CCC ATG CAG CTG CCC TAT GAC TCC ACG TGG      2132
Glu Tyr Ile Tyr Val Asp Pro Met Gln Leu Pro Tyr Asp Ser Thr Trp
        580                 585                 590

GAG CTG CCG CGG GAC CAG CTT GTG CTG GGA CGC ACC CTC GGC TCT GGG      2180
Glu Leu Pro Arg Asp Gln Leu Val Leu Gly Arg Thr Leu Gly Ser Gly
    595                 600                 605
```

-continued

```
GCC TTT GGG CAG GTG GTG GAG GCC ACG GCT CAT GGC CTG AGC CAT TCT     2228
Ala Phe Gly Gln Val Val Glu Ala Thr Ala His Gly Leu Ser His Ser
610                 615                 620                 625

CAG GCC ACG ATG AAA GTG GCC GTC AAG ATG CTT AAA TCC ACA GCC CGC     2276
Gln Ala Thr Met Lys Val Ala Val Lys Met Leu Lys Ser Thr Ala Arg
                630                 635                 640

AGC AGT GAG AAG CAA GCC CTT ATG TCG GAG CTG AAG ATC ATG AGT CAC     2324
Ser Ser Glu Lys Gln Ala Leu Met Ser Glu Leu Lys Ile Met Ser His
            645                 650                 655

CTT GGG CCC CAC CTG AAC GTG GTC AAC CTG TTG GGG GCC TGC ACC AAA     2372
Leu Gly Pro His Leu Asn Val Val Asn Leu Leu Gly Ala Cys Thr Lys
        660                 665                 670

GGA GGA CCC ATC TAT ATC ATC ACT GAG TAC TGC CGC TAC GGA GAC CTG     2420
Gly Gly Pro Ile Tyr Ile Ile Thr Glu Tyr Cys Arg Tyr Gly Asp Leu
    675                 680                 685

GTG GAC TAC CTG CAC CGC AAC AAA CAC ACC TTC CTG CAG CAC CAC TCC     2468
Val Asp Tyr Leu His Arg Asn Lys His Thr Phe Leu Gln His His Ser
690                 695                 700                 705

GAC AAG CGC CGC CCG CCC AGC GCG GAG CTC TAC AGC AAT GCT CTG CCC     2516
Asp Lys Arg Arg Pro Pro Ser Ala Glu Leu Tyr Ser Asn Ala Leu Pro
                710                 715                 720

GTT GGG CTC CCC CTG CCC AGC CAT GTG TCC TTG ACC GGG GAG AGC GAC     2564
Val Gly Leu Pro Leu Pro Ser His Val Ser Leu Thr Gly Glu Ser Asp
            725                 730                 735

GGT GGC TAC ATG GAC ATG AGC AAG GAC GAG TCG GTG GAC TAT GTG CCC     2612
Gly Gly Tyr Met Asp Met Ser Lys Asp Glu Ser Val Asp Tyr Val Pro
        740                 745                 750

ATG CTG GAC ATG AAA GGA GAC GTC AAA TAT GCA GAC ATC GAG TCC TCC     2660
Met Leu Asp Met Lys Gly Asp Val Lys Tyr Ala Asp Ile Glu Ser Ser
    755                 760                 765

AAC TAC ATG GCC CCT TAC GAT AAC TAC GTT CCC TCT GCC CCT GAG AGG     2708
Asn Tyr Met Ala Pro Tyr Asp Asn Tyr Val Pro Ser Ala Pro Glu Arg
770                 775                 780                 785

ACC TGC CGA GCA ACT TTG ATC AAC GAG TCT CCA GTG CTA AGC TAC ATG     2756
Thr Cys Arg Ala Thr Leu Ile Asn Glu Ser Pro Val Leu Ser Tyr Met
                790                 795                 800

GAC CTC GTG GGC TTC AGC TAC CAG GTG GCC AAT GGC ATG GAG TTT CTG     2804
Asp Leu Val Gly Phe Ser Tyr Gln Val Ala Asn Gly Met Glu Phe Leu
            805                 810                 815

GCC TCC AAG AAC TGC GTC CAC AGA GAC CTG GCG GCT AGG AAC GTG CTC     2852
Ala Ser Lys Asn Cys Val His Arg Asp Leu Ala Ala Arg Asn Val Leu
        820                 825                 830

ATC TGT GAA GGC AAG CTG GTC AAG ATC TGT GAC TTT GGC CTG GCT CGA     2900
Ile Cys Glu Gly Lys Leu Val Lys Ile Cys Asp Phe Gly Leu Ala Arg
    835                 840                 845

GAC ATC ATG CGG GAC TCG AAT TAC ATC TCC AAA GGC AGC ACC TTT TTG     2948
Asp Ile Met Arg Asp Ser Asn Tyr Ile Ser Lys Gly Ser Thr Phe Leu
850                 855                 860                 865

CCT TTA AAG TGG ATG GCT CCG GAG AGC ATC TTC AAC AGC CTC TAC ACC     2996
Pro Leu Lys Trp Met Ala Pro Glu Ser Ile Phe Asn Ser Leu Tyr Thr
                870                 875                 880

ACC CTG AGC GAC GTG TGG TCC TTC GGG ATC CTG CTC TGG GAG ATC TTC     3044
Thr Leu Ser Asp Val Trp Ser Phe Gly Ile Leu Leu Trp Glu Ile Phe
            885                 890                 895

ACC TTG GGT GGC ACC CCT TAC CCA GAG CTG CCC ATG AAC GAG CAG TTC     3092
Thr Leu Gly Gly Thr Pro Tyr Pro Glu Leu Pro Met Asn Glu Gln Phe
        900                 905                 910

TAC AAT GCC ATC AAA CGG GGT TAC CGC ATG GCC CAG CCT GCC CAT GCC     3140
Tyr Asn Ala Ile Lys Arg Gly Tyr Arg Met Ala Gln Pro Ala His Ala
    915                 920                 925
```

-continued

```
TCC GAC GAG ATC TAT GAG ATC ATG CAG AAG TGC TGG GAA GAG AAG TTT      3188
Ser Asp Glu Ile Tyr Glu Ile Met Gln Lys Cys Trp Glu Glu Lys Phe
930                 935                 940                 945

GAG ATT CGG CCC CCC TTC TCC CAG CTG GTG CTG CTT CTC GAG AGA CTG      3236
Glu Ile Arg Pro Pro Phe Ser Gln Leu Val Leu Leu Leu Glu Arg Leu
                950                 955                 960

TTG GGC GAA GGT TAC AAA AAG AAG TAC CAG CAG GTG GAT GAG GAG TTT      3284
Leu Gly Glu Gly Tyr Lys Lys Lys Tyr Gln Gln Val Asp Glu Glu Phe
            965                 970                 975

CTG AGG AGT GAC CAC CCA GCC ATC CTT CGC TCC CAG GCC CGC TTG CCT      3332
Leu Arg Ser Asp His Pro Ala Ile Leu Arg Ser Gln Ala Arg Leu Pro
        980                 985                 990

GGG TTC CAT GGC CTC CGA TCT CCC CTG GAC ACC AGC TCC GTC CTC TAT      3380
Gly Phe His Gly Leu Arg Ser Pro Leu Asp Thr Ser Ser Val Leu Tyr
    995                 1000                1005

ACT GCC GTG CAG CCC AAT GAG GGT GAC AAC GAC TAT ATC ATC CCC CTG      3428
Thr Ala Val Gln Pro Asn Glu Gly Asp Asn Asp Tyr Ile Ile Pro Leu
1010                1015                1020                1025

CCT GAC CCC AAA CCC GAG GTT GCT GAC GAG GGC CCA CTG GAG GGT TCC      3476
Pro Asp Pro Lys Pro Glu Val Ala Asp Glu Gly Pro Leu Glu Gly Ser
                1030                1035                1040

CCC AGC CTA GCC AGC TCC ACC CTG AAT GAA GTC AAC ACC TCC TCA ACC      3524
Pro Ser Leu Ala Ser Ser Thr Leu Asn Glu Val Asn Thr Ser Ser Thr
            1045                1050                1055

ATC TCC TGT GAC AGC CCC CTG GAG CCC CAG GAC GAA CCA GAG CCA GAG      3572
Ile Ser Cys Asp Ser Pro Leu Glu Pro Gln Asp Glu Pro Glu Pro Glu
        1060                1065                1070

CCC CAG CTT GAG CTC CAG GTG GAG CCG GAG CCA GAG CTG GAA CAG TTG      3620
Pro Gln Leu Glu Leu Gln Val Glu Pro Glu Pro Glu Leu Glu Gln Leu
    1075                1080                1085

CCG GAT TCG GGG TGC CCT GCG CCT CGG GCC GAA GCA GAG GAT AGC TTC      3668
Pro Asp Ser Gly Cys Pro Ala Pro Arg Ala Glu Ala Glu Asp Ser Phe
1090                1095                1100                1105

CTG TAGGGGGCTG GCCCCTACCC TGCCCTGCCT GAAGCTCCCC CCCTGCCAGC           3721
Leu

ACCCAGCATC TCCTGGCCTG GCCTGACCGG GCTTCCTGTC AGCCAGGGTG CCCTTATCAG   3781

CTGTCCCCTT CTGGAAGCTT TCTGCTCCTG ACGTGTTGTG CCCCAAACCC TGGGGCTGGC   3841

TTAGGAGGCA AGAAAACTGC AGGGGCCGTG ACCAGCCCTC TGCCTCCAGG GAGGCCAACT   3901

GACTCTGAGC CAGGGTTCCC CCAGGGAACT CAGTTTTCCC ATATGTAAGA TGGGAAAGTT   3961

AGGCTTGATG ACCCAGAATC TAGGATTCTC TCCCTGGCTG ACACGGTGGG GAGACCGAAT   4021

CCCTCCCTGG GAAGATTCTT GGAGTTACTG AGGTGGTAAA TTAACATTTT TTCTGTTCAG   4081

CCAGCTACCC CTCAAGGAAT CATAGCTCTC TCCTCGCACT TTTTATCCAC CCAGGAGCTA   4141

GGGAAGAGAC CCTAGCCTCC CTGGCTGCTG GCTGAGCTAG GGCCTAGCTT GAGCAGTGTT   4201

GCCTCATCCA GAAGAAAGCC AGTCTCCTCC CTATGATGCC AGTCCCTGCG TTCCCTGGCC   4261

CGAGCTGGTC TGGGGCCATT AGGCAGCCTA ATTAATGCTG GAGGCTGAGC CAAGTACAGG   4321

ACACCCCCAG CCTGCAGCCC TTGCCCAGGG CACTTGGAGC ACACGCAGCC ATAGCAAGTG   4381

CCTGTGTCCC TGTCCTTCAG GCCCATCAGT CCTGGGGCTT TTTCTTTATC ACCCTCAGTC   4441

TTAATCCATC CACCAGAGTC TAGA                                           4465
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1106 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

—continued

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
Met Arg Leu Pro Gly Ala Met Pro Ala Leu Ala Leu Lys Gly Glu Leu
 1               5                  10                  15

Leu Leu Leu Ser Leu Leu Leu Leu Leu Glu Pro Gln Ile Ser Gln Gly
            20                  25                  30

Leu Val Val Thr Pro Pro Gly Pro Glu Leu Val Leu Asn Val Ser Ser
        35                  40                  45

Thr Phe Val Leu Thr Cys Ser Gly Ser Ala Pro Val Val Trp Glu Arg
    50                  55                  60

Met Ser Gln Glu Pro Pro Gln Glu Met Ala Lys Ala Glu Asp Gly Thr
65                  70                  75                  80

Phe Ser Ser Val Leu Thr Leu Thr Asn Leu Thr Gly Leu Asp Thr Gly
                85                  90                  95

Glu Tyr Phe Cys Thr His Asn Asp Ser Arg Gly Leu Glu Thr Asp Glu
            100                 105                 110

Arg Lys Arg Leu Tyr Ile Phe Val Pro Asp Pro Thr Val Gly Phe Leu
        115                 120                 125

Pro Asn Asp Ala Glu Glu Leu Phe Ile Phe Leu Thr Glu Ile Thr Glu
    130                 135                 140

Ile Thr Ile Pro Cys Arg Val Thr Asp Pro Gln Leu Val Val Thr Leu
145                 150                 155                 160

His Glu Lys Lys Gly Asp Val Ala Leu Pro Val Pro Tyr Asp His Gln
                165                 170                 175

Arg Gly Phe Ser Gly Ile Phe Glu Asp Arg Ser Tyr Ile Cys Lys Thr
            180                 185                 190

Thr Ile Gly Asp Arg Glu Val Asp Ser Asp Ala Tyr Tyr Val Tyr Arg
        195                 200                 205

Leu Gln Val Ser Ser Ile Asn Val Ser Val Asn Ala Val Gln Thr Val
    210                 215                 220

Val Arg Gln Gly Glu Asn Ile Thr Leu Met Cys Ile Val Ile Gly Asn
225                 230                 235                 240

Glu Val Val Asn Phe Glu Trp Thr Tyr Pro Arg Lys Glu Ser Gly Arg
                245                 250                 255

Leu Val Glu Pro Val Thr Asp Phe Leu Leu Asp Met Pro Tyr His Ile
            260                 265                 270

Arg Ser Ile Leu His Ile Pro Ser Ala Glu Leu Glu Asp Ser Gly Thr
        275                 280                 285

Tyr Thr Cys Asn Val Thr Glu Ser Val Asn Asp His Gln Asp Glu Lys
    290                 295                 300

Ala Ile Asn Ile Thr Val Val Glu Ser Gly Tyr Val Arg Leu Leu Gly
305                 310                 315                 320

Glu Val Gly Thr Leu Gln Phe Ala Glu Leu His Arg Ser Arg Thr Leu
                325                 330                 335

Gln Val Val Phe Glu Ala Tyr Pro Pro Pro Thr Val Leu Trp Phe Lys
            340                 345                 350

Asp Asn Arg Thr Leu Gly Asp Ser Ser Ala Gly Glu Ile Ala Leu Ser
        355                 360                 365

Thr Arg Asn Val Ser Glu Thr Arg Tyr Val Ser Glu Leu Thr Leu Val
    370                 375                 380

Arg Val Lys Val Ala Glu Ala Gly His Tyr Thr Met Arg Ala Phe His
385                 390                 395                 400
```

-continued

Glu Asp Ala Glu Val Gln Leu Ser Phe Gln Leu Gln Ile Asn Val Pro
405 410 415

Val Arg Val Leu Glu Leu Ser Glu Ser His Pro Asp Ser Gly Glu Gln
420 425 430

Thr Val Arg Cys Arg Gly Arg Gly Met Pro Gln Pro Asn Ile Ile Trp
435 440 445

Ser Ala Cys Arg Asp Leu Lys Arg Cys Pro Arg Glu Leu Pro Pro Thr
450 455 460

Leu Leu Gly Asn Ser Ser Glu Glu Glu Ser Gln Leu Glu Thr Asn Val
465 470 475 480

Thr Tyr Trp Glu Glu Glu Gln Glu Phe Glu Val Val Ser Thr Leu Arg
485 490 495

Leu Gln His Val Asp Arg Pro Leu Ser Val Arg Cys Thr Leu Arg Asn
500 505 510

Ala Val Gly Gln Asp Thr Gln Glu Val Ile Val Val Pro His Ser Leu
515 520 525

Pro Phe Lys Val Val Val Ile Ser Ala Ile Leu Ala Leu Val Val Leu
530 535 540

Thr Ile Ile Ser Leu Ile Ile Leu Ile Met Leu Trp Gln Lys Lys Pro
545 550 555 560

Arg Tyr Glu Ile Arg Trp Lys Val Ile Glu Ser Val Ser Ser Asp Gly
565 570 575

His Glu Tyr Ile Tyr Val Asp Pro Met Gln Leu Pro Tyr Asp Ser Thr
580 585 590

Trp Glu Leu Pro Arg Asp Gln Leu Val Leu Gly Arg Thr Leu Gly Ser
595 600 605

Gly Ala Phe Gly Gln Val Val Glu Ala Thr Ala His Gly Leu Ser His
610 615 620

Ser Gln Ala Thr Met Lys Val Ala Val Lys Met Leu Lys Ser Thr Ala
625 630 635 640

Arg Ser Ser Glu Lys Gln Ala Leu Met Ser Glu Leu Lys Ile Met Ser
645 650 655

His Leu Gly Pro His Leu Asn Val Val Asn Leu Leu Gly Ala Cys Thr
660 665 670

Lys Gly Gly Pro Ile Tyr Ile Ile Thr Glu Tyr Cys Arg Tyr Gly Asp
675 680 685

Leu Val Asp Tyr Leu His Arg Asn Lys His Thr Phe Leu Gln His His
690 695 700

Ser Asp Lys Arg Arg Pro Pro Ser Ala Glu Leu Tyr Ser Asn Ala Leu
705 710 715 720

Pro Val Gly Leu Pro Leu Pro Ser His Val Ser Leu Thr Gly Glu Ser
725 730 735

Asp Gly Gly Tyr Met Asp Met Ser Lys Asp Glu Ser Val Asp Tyr Val
740 745 750

Pro Met Leu Asp Met Lys Gly Asp Val Lys Tyr Ala Asp Ile Glu Ser
755 760 765

Ser Asn Tyr Met Ala Pro Tyr Asp Asn Tyr Val Pro Ser Ala Pro Glu
770 775 780

Arg Thr Cys Arg Ala Thr Leu Ile Asn Glu Ser Pro Val Leu Ser Tyr
785 790 795 800

Met Asp Leu Val Gly Phe Ser Tyr Gln Val Ala Asn Gly Met Glu Phe
805 810 815

Leu Ala Ser Lys Asn Cys Val His Arg Asp Leu Ala Ala Arg Asn Val
820 825 830

```
Leu Ile Cys Glu Gly Lys Leu Val Lys Ile Cys Asp Phe Gly Leu Ala
        835                 840                 845

Arg Asp Ile Met Arg Asp Ser Asn Tyr Ile Ser Lys Gly Ser Thr Phe
    850                 855                 860

Leu Pro Leu Lys Trp Met Ala Pro Glu Ser Ile Phe Asn Ser Leu Tyr
865                 870                 875                 880

Thr Thr Leu Ser Asp Val Trp Ser Phe Gly Ile Leu Leu Trp Glu Ile
                885                 890                 895

Phe Thr Leu Gly Gly Thr Pro Tyr Pro Glu Leu Pro Met Asn Glu Gln
            900                 905                 910

Phe Tyr Asn Ala Ile Lys Arg Gly Tyr Arg Met Ala Gln Pro Ala His
        915                 920                 925

Ala Ser Asp Glu Ile Tyr Glu Ile Met Gln Lys Cys Trp Glu Glu Lys
    930                 935                 940

Phe Glu Ile Arg Pro Pro Phe Ser Gln Leu Val Leu Leu Leu Glu Arg
945                 950                 955                 960

Leu Leu Gly Glu Gly Tyr Lys Lys Lys Tyr Gln Gln Val Asp Glu Glu
                965                 970                 975

Phe Leu Arg Ser Asp His Pro Ala Ile Leu Arg Ser Gln Ala Arg Leu
            980                 985                 990

Pro Gly Phe His Gly Leu Arg Ser Pro Leu Asp Thr Ser Ser Val Leu
        995                 1000                1005

Tyr Thr Ala Val Gln Pro Asn Glu Gly Asp Asn Asp Tyr Ile Ile Pro
    1010                1015                1020

Leu Pro Asp Pro Lys Pro Glu Val Ala Asp Glu Gly Pro Leu Glu Gly
1025                1030                1035                1040

Ser Pro Ser Leu Ala Ser Ser Thr Leu Asn Glu Val Asn Thr Ser Ser
                1045                1050                1055

Thr Ile Ser Cys Asp Ser Pro Leu Glu Pro Gln Asp Glu Pro Glu Pro
            1060                1065                1070

Glu Pro Gln Leu Glu Leu Gln Val Glu Pro Glu Pro Glu Leu Glu Gln
        1075                1080                1085

Leu Pro Asp Ser Gly Cys Pro Ala Pro Arg Ala Glu Ala Glu Asp Ser
    1090                1095                1100

Phe Leu
1105
```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 57 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(vii) IMMEDIATE SOURCE:
(B) CLONE: ZC871

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

ATTATACGCT CTCTTCCTCA GGTAAATGAG TGCCAGGGCC GGCAAGCCCC CGCTCCA 57

(2) INFORMATION FOR SEQ ID NO:4:

-continued

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 56 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(vii) IMMEDIATE SOURCE:
(B) CLONE: ZC872

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

CCGGGGAGCG GGGGCTTGCC GGCCCTGGCA CTCATTTACC TGAGGAAGAG AGAGCT 56

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 45 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(vii) IMMEDIATE SOURCE:
(B) CLONE: ZC904

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CATGGGCACG TAATCTATAG ATTCATCCTT GCTCATATCC ATGTA 45

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 38 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(vii) IMMEDIATE SOURCE:
(B) CLONE: ZC906

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

AAGCTGTCCT CTGCTTCAGC CAGAGGTCCT GGGCAGCC 38

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 38 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(vii) IMMEDIATE SOURCE:
(B) CLONE: ZC916

-continued

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

AAGCTGTCCT CTGCTTCAGC CAGAGGTCCT GGGCAGCC 38

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(vii) IMMEDIATE SOURCE:
(B) CLONE: ZC1380

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

CATGGTGGAA TTCCTGCTGA T 21

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(vii) IMMEDIATE SOURCE:
(B) CLONE: ZC1447

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

TGGTTGTGCA GAGCTGAGGA AGAGATGGA 29

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 55 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(vii) IMMEDIATE SOURCE:
(B) CLONE: ZC1453

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

AATTCATTAT GTTGTTGCAA GCCTTCTTGT TCCTGCTAGC TGGTTTCGCT GTTAA 55

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 55 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

-continued

(ii) MOLECULE TYPE: Other nucleic acid

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(vii) IMMEDIATE SOURCE:
(B) CLONE: ZC1454

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GATCTTAACA GCGAAACCAG CTAGCAGGAA CAAGAAGGCT TGCAACAACA TAATG 55

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(vii) IMMEDIATE SOURCE:
(B) CLONE: ZC1478

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

ATCGCGAGCA TGCAGATCTG A 21

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(vii) IMMEDIATE SOURCE:
(B) CLONE: ZC1479

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

AGCTTCAGAT CTGCATGCTG CCGAT 25

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 52 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(vii) IMMEDIATE SOURCE:
(B) CLONE: ZC1776

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

AGCTCAGCCC AAATGTTCTG TCGAGTGCCC ACCGTCCCCA GCTTAGAATT CT 52

-continued

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 52 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(vii) IMMEDIATE SOURCE:
(B) CLONE: ZC1777

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

CTAGAGAATT CTAAGCTGGG CACGGTGGGC ACTCGACACA ACATTTGCGC TC 52

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 95 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(vii) IMMEDIATE SOURCE:
(B) CLONE: ZC1846

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

GATCGGCCAC TGTCGGTGCG CTGCACGCTG CGCAACGCTG TGGGCCAGGA CACGCAGGAG 60

GTCATCGTGG TGCCACACTC CTTGCCCTTT AAGCA 95

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 95 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(vii) IMMEDIATE SOURCE:
(B) CLONE: ZC1847

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

AGCTTGCTTA AAGGGCAAGG AGTGTGGCAC CACGATGACC TCCTGCGTGT CCTGGCCCAC 60

AGCGTTGCGC AGCGTGCAGC GCACCGACAG TGGCC 95

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 43 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

-continued

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(vii) IMMEDIATE SOURCE:
(B) CLONE: ZC1886

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

CCAGTGCCAA GCTTGTCTAG ACTTACCTTT AAAGGGCAAG GAG 43

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 11 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(vii) IMMEDIATE SOURCE:
(B) CLONE: ZC1892

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

AGCTTGAGCG T 11

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 11 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(vii) IMMEDIATE SOURCE:
(B) CLONE: ZC1893

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

CTAGACGCTC A 11

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 47 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(vii) IMMEDIATE SOURCE:
(B) CLONE: ZC1894

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

AGCTTGCAGT TCTTCGGCCT CATGTCAGTT CTTCGGCCTC ATGTGAT 47

(2) INFORMATION FOR SEQ ID NO:22:

-continued

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 47 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(vii) IMMEDIATE SOURCE:
(B) CLONE: ZC1895

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

CTAGATCACA TGAGGCCGAA GAACTGACAT GAGGCCGAAG AACTGGA 47

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 66 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(vii) IMMEDIATE SOURCE:
(B) CLONE: ZC2181

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

AATTCGGATC CACCATGGGC ACCAGCCACC CGGCGTTCCT GGTGTTAGGC TGCCTGCTGA 60

CCGGCC 66

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 71 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(vii) IMMEDIATE SOURCE:
(B) CLONE: ZC2182

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

TGAGCCTGAT CCTGTGCCAA CTGAGCCTGC CATCGATCCT GCCAAACGAG AACGAGAAGG 60

TTGTGCAGCT A 71

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 69 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(iii) HYPOTHETICAL: N

-continued

(iv) ANTI-SENSE: N

(vii) IMMEDIATE SOURCE:
(B) CLONE: ZC2183

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

AATTTAGCTG CACAACCTTC TCGTTCTCGT TTGGCAGGAT CGATGGCAGG CTCAGTTGGC 60

ACAGGATCA 69

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 68 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(vii) IMMEDIATE SOURCE:
(B) CLONE: ZC2184

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

GGCTCAGGCC GGTCAGCAGG CAGCCTAACA CCAGGAACGC CGGGTGGCTG GTGCCCATGG 60

TGGATCCG 68

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(vii) IMMEDIATE SOURCE:
(B) CLONE: ZC2311

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

TGATCACCAT GGCTCAACTG 20

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(vii) IMMEDIATE SOURCE:
(B) CLONE: ZC2351

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

CGAATTCCAG 10

-continued

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(vii) IMMEDIATE SOURCE:
(B) CLONE: ZC2352

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

ATTATACGCA TGGTGGAATT CGAGCT 26

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 41 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(vii) IMMEDIATE SOURCE:
(B) CLONE: ZC2392

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

ACGTAAGCTT GTCTAGACTT ACCTTCAGAA CGCAGGGTGG G 41

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: C-terminal

(vii) IMMEDIATE SOURCE:
(B) CLONE: pWK1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

Ala Leu His Asn His Tyr Thr Glu Lys Ser Leu Ser Leu Ser Pro Gly
1 5 10 15

Lys

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

TGTGACACTC TCCTGGGAGT TA 22

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

GCATAGTAGT TACCATATCC TCTTGCACAG 30

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

ACCGAACGTG AGAGGAGTGC TATAA 25

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4054 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(vii) IMMEDIATE SOURCE:
(B) CLONE: p-alpha-17B

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 205..3471
(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

GCCCTGGCGA CGGACCGTGG CCGGCGCGCA GCCGCGGGAC GCGTTTTGGG GACGTGGTGG 60

CCAGCGCCTT CCTGCAGACC CACACGGAAG TACTCCCTTT GACCTCCGGG GAGCTGCGAC 120

CAGTTTATAC GTTGCTGGTC GAAAAGTGAC AATTCTAGGA AAACAGCTAA AAGCCGGATC 180

G.TGACCGAA AGTTTCCCAC AGCT ATG GGG ACT TCC CAT CCG GCG TTC CTG 231
Met Gly Thr Ser His Pro Ala Phe Leu
1 5

-continued

```
GTC TTA GGC TGT CTT CTC ACA GGG CTG AGC CTA ATC CTC TGC CAG CTT      279
Val Leu Gly Cys Leu Leu Thr Gly Leu Ser Leu Ile Leu Cys Gln Leu
 10                  15                  20                  25

TCA TTA CCC TCT ATC CTT CCA AAT GAA AAT GAA AAG GTT GTG CAG CTG      327
Ser Leu Pro Ser Ile Leu Pro Asn Glu Asn Glu Lys Val Val Gln Leu
             30                  35                  40

AAT TCA TCC TTT TCT CTC AGA TGC TTT GGG GAG AGT GAA GTG AGC TGG      375
Asn Ser Ser Phe Ser Leu Arg Cys Phe Gly Glu Ser Glu Val Ser Trp
         45                  50                  55

CAG TAC CCC ATG TCT GAA GAA GAG AGC TCC GAT GTG GAA ATC AGA AAT      423
Gln Tyr Pro Met Ser Glu Glu Glu Ser Ser Asp Val Glu Ile Arg Asn
     60                  65                  70

GAA GAA AAC AAC AGC GGC CTT TTT GTG ACG GTC TTG GAA GTG AGC AGT      471
Glu Glu Asn Asn Ser Gly Leu Phe Val Thr Val Leu Glu Val Ser Ser
 75                  80                  85

GCC TCG GCG GCC CAC ACA GGG TTG TAC ACT TGC TAT TAC AAC CAC ACT      519
Ala Ser Ala Ala His Thr Gly Leu Tyr Thr Cys Tyr Tyr Asn His Thr
 90                  95                 100                 105

CAG ACA GAA GAG AAT GAG CTT GAA GGC AGG CAC ATT TAC ATC TAT GTG      567
Gln Thr Glu Glu Asn Glu Leu Glu Gly Arg His Ile Tyr Ile Tyr Val
            110                 115                 120

CCA GAC CCA GAT GTA GCC TTT GTA CCT CTA GGA ATG ACG GAT TAT TTA      615
Pro Asp Pro Asp Val Ala Phe Val Pro Leu Gly Met Thr Asp Tyr Leu
        125                 130                 135

GTC ATC GTG GAG GAT GAT GAT TCT GCC ATT ATA CCT TGT CGC ACA ACT      663
Val Ile Val Glu Asp Asp Asp Ser Ala Ile Ile Pro Cys Arg Thr Thr
    140                 145                 150

GAT CCC GAG ACT CCT GTA ACC TTA CAC AAC AGT GAG GGG GTG GTA CCT      711
Asp Pro Glu Thr Pro Val Thr Leu His Asn Ser Glu Gly Val Val Pro
155                 160                 165

GCC TCC TAC GAC AGC AGA CAG GGC TTT AAT GGG ACC TTC ACT GTA GGG      759
Ala Ser Tyr Asp Ser Arg Gln Gly Phe Asn Gly Thr Phe Thr Val Gly
170                 175                 180                 185

CCC TAT ATC TGT GAG GCC ACC GTC AAA GGA AAG AAG TTC CAG ACC ATC      807
Pro Tyr Ile Cys Glu Ala Thr Val Lys Gly Lys Lys Phe Gln Thr Ile
            190                 195                 200

CCA TTT AAT GTT TAT GCT TTA AAA GCA ACA TCA GAG CTG GAT CTA GAA      855
Pro Phe Asn Val Tyr Ala Leu Lys Ala Thr Ser Glu Leu Asp Leu Glu
        205                 210                 215

ATG GAA GCT CTT AAA ACC GTG TAT AAG TCA GGG GAA ACG ATT GTG GTC      903
Met Glu Ala Leu Lys Thr Val Tyr Lys Ser Gly Glu Thr Ile Val Val
    220                 225                 230

ACC TGT GCT GTT TTT AAC AAT GAG GTG GTT GAC CTT CAA TGG ACT TAC      951
Thr Cys Ala Val Phe Asn Asn Glu Val Val Asp Leu Gln Trp Thr Tyr
235                 240                 245

CCT GGA GAA GTG AAA GGC AAA GGC ATC ACA ATA CTG GAA GAA ATC AAA      999
Pro Gly Glu Val Lys Gly Lys Gly Ile Thr Ile Leu Glu Glu Ile Lys
250                 255                 260                 265

GTC CCA TCC ATC AAA TTG GTG TAC ACT TTG ACG GTC CCC GAG GCC ACG     1047
Val Pro Ser Ile Lys Leu Val Tyr Thr Leu Thr Val Pro Glu Ala Thr
            270                 275                 280

GTG AAA GAC AGT GGA GAT TAC GAA TGT GCT GCC CGC CAG GCT ACC AGG     1095
Val Lys Asp Ser Gly Asp Tyr Glu Cys Ala Ala Arg Gln Ala Thr Arg
        285                 290                 295

GAG GTC AAA GAA ATG AAG AAA GTC ACT ATT TCT GTC CAT GAG AAA GGT     1143
Glu Val Lys Glu Met Lys Lys Val Thr Ile Ser Val His Glu Lys Gly
    300                 305                 310

TTC ATT GAA ATC AAA CCC ACC TTC AGC CAG TTG GAA GCT GTC AAC CTG     1191
Phe Ile Glu Ile Lys Pro Thr Phe Ser Gln Leu Glu Ala Val Asn Leu
315                 320                 325
```

-continued

```
CAT GAA GTC AAA CAT TTT GTT GTA GAG GTG CGG GCC TAC CCA CCT CCC    1239
His Glu Val Lys His Phe Val Val Glu Val Arg Ala Tyr Pro Pro Pro
330                 335                 340                 345

AGG ATA TCC TGG CTG AAA AAC AAT CTG ACT CTG ATT GAA AAT CTC ACT    1287
Arg Ile Ser Trp Leu Lys Asn Asn Leu Thr Leu Ile Glu Asn Leu Thr
                350                 355                 360

GAG ATC ACC ACT GAT GTG GAA AAG ATT CAG GAA ATA AGG TAT CGA AGC    1335
Glu Ile Thr Thr Asp Val Glu Lys Ile Gln Glu Ile Arg Tyr Arg Ser
            365                 370                 375

AAA TTA AAG CTG ATC CGT GCT AAG GAA GAA GAC AGT GGC CAT TAT ACT    1383
Lys Leu Lys Leu Ile Arg Ala Lys Glu Glu Asp Ser Gly His Tyr Thr
        380                 385                 390

ATT GTA GCT CAA AAT GAA GAT GCT GTG AAG AGC TAT ACT TTT GAA CTG    1431
Ile Val Ala Gln Asn Glu Asp Ala Val Lys Ser Tyr Thr Phe Glu Leu
    395                 400                 405

TTA ACT CAA GTT CCT TCA TCC ATT CTG GAC TTG GTC GAT GAT CAC CAT    1479
Leu Thr Gln Val Pro Ser Ser Ile Leu Asp Leu Val Asp Asp His His
410                 415                 420                 425

GGC TCA ACT GGG GGA CAG ACG GTG AGG TGC ACA GCT GAA GGC ACG CCG    1527
Gly Ser Thr Gly Gly Gln Thr Val Arg Cys Thr Ala Glu Gly Thr Pro
                430                 435                 440

CTT CCT GAT ATT GAG TGG ATG ATA TGC AAA GAT ATT AAG AAA TGT AAT    1575
Leu Pro Asp Ile Glu Trp Met Ile Cys Lys Asp Ile Lys Lys Cys Asn
            445                 450                 455

AAT GAA ACT TCC TGG ACT ATT TTG GCC AAC AAT GTC TCA AAC ATC ATC    1623
Asn Glu Thr Ser Trp Thr Ile Leu Ala Asn Asn Val Ser Asn Ile Ile
        460                 465                 470

ACG GAG ATC CAC TCC CGA GAC AGG AGT ACC GTG GAG GGC CGT GTG ACT    1671
Thr Glu Ile His Ser Arg Asp Arg Ser Thr Val Glu Gly Arg Val Thr
    475                 480                 485

TTC GCC AAA GTG GAG GAG ACC ATC GCC GTG CGA TGC CTG GCT AAG AAT    1719
Phe Ala Lys Val Glu Glu Thr Ile Ala Val Arg Cys Leu Ala Lys Asn
490                 495                 500                 505

CTC CTT GGA GCT GAG AAC CGA GAG CTG AAG CTG GTG GCT CCC ACC CTG    1767
Leu Leu Gly Ala Glu Asn Arg Glu Leu Lys Leu Val Ala Pro Thr Leu
                510                 515                 520

CGT TCT GAA CTC ACG GTG GCT GCT GCA GTC CTG GTG CTG TTG GTG ATT    1815
Arg Ser Glu Leu Thr Val Ala Ala Ala Val Leu Val Leu Leu Val Ile
            525                 530                 535

GTG ATC ATC TCA CTT ATT GTC CTG GTT GTC ATT TGG AAA CAG AAA CCG    1863
Val Ile Ile Ser Leu Ile Val Leu Val Val Ile Trp Lys Gln Lys Pro
        540                 545                 550

AGG TAT GAA ATT CGC TGG AGG GTC ATT GAA TCA ATC AGC CCG GAT GGA    1911
Arg Tyr Glu Ile Arg Trp Arg Val Ile Glu Ser Ile Ser Pro Asp Gly
    555                 560                 565

CAT GAA TAT ATT TAT GTG GAC CCG ATG CAG CTG CCT TAT GAC TCA AGA    1959
His Glu Tyr Ile Tyr Val Asp Pro Met Gln Leu Pro Tyr Asp Ser Arg
570                 575                 580                 585

TGG GAG TTT CCA AGA GAT GGA CTA GTG CTT GGT CGG GTC TTG GGG TCT    2007
Trp Glu Phe Pro Arg Asp Gly Leu Val Leu Gly Arg Val Leu Gly Ser
                590                 595                 600

GGA GCG TTT GGG AAG GTG GTT GAA GGA ACA GCC TAT GGA TTA AGC CGG    2055
Gly Ala Phe Gly Lys Val Val Glu Gly Thr Ala Tyr Gly Leu Ser Arg
            605                 610                 615

TCC CAA CCT GTC ATG AAA GTT GCA GTG AAG ATG CTA AAA CCC ACG GCC    2103
Ser Gln Pro Val Met Lys Val Ala Val Lys Met Leu Lys Pro Thr Ala
        620                 625                 630

AGA TCC AGT GAA AAA CAA GCT CTC ATG TCT GAA CTG AAG ATA ATG ACT    2151
Arg Ser Ser Glu Lys Gln Ala Leu Met Ser Glu Leu Lys Ile Met Thr
    635                 640                 645
```

-continued

```
CAC CTG GGG CCA CAT TTG AAC ATT GTA AAC TTG CTG GGA GCC TGC ACC     2199
His Leu Gly Pro His Leu Asn Ile Val Asn Leu Leu Gly Ala Cys Thr
650                 655                 660                 665

AAG TCA GGC CCC ATT TAC ATC ATC ACA GAG TAT TGC TTC TAT GGA GAT     2247
Lys Ser Gly Pro Ile Tyr Ile Ile Thr Glu Tyr Cys Phe Tyr Gly Asp
                670                 675                 680

TTG GTC AAC TAT TTG CAT AAG AAT AGG GAT AGC TTC CTG AGC CAC CAC     2295
Leu Val Asn Tyr Leu His Lys Asn Arg Asp Ser Phe Leu Ser His His
            685                 690                 695

CCA GAG AAG CCA AAC AAA CAG CTG GAT ATC TTT GGA TTG AAC CCT GCT     2343
Pro Glu Lys Pro Lys Lys Glu Leu Asp Ile Phe Gly Leu Asn Pro Ala
        700                 705                 710

GAT GAA AGC ACA CGG AGC TAT GTT ATT TTA TCT TTT GAA AAC AAT GGT     2391
Asp Glu Ser Thr Arg Ser Tyr Val Ile Leu Ser Phe Glu Asn Asn Gly
    715                 720                 725

GAC TAC ATG GAC ATG AAG CAG GCT GAT ACT ACA CAG TAT GTC CCC ATG     2439
Asp Tyr Met Asp Met Lys Gln Ala Asp Thr Thr Gln Tyr Val Pro Met
730                 735                 740                 745

CTA GAA AGG AAA GAG GTT TCT AAA TAT TCC GAC ATC CAG AGA TCA CTC     2487
Leu Glu Arg Lys Glu Val Ser Lys Tyr Ser Asp Ile Gln Arg Ser Leu
                750                 755                 760

TAT GAT CGT CCA GCC TCA TAT AAG AAG AAA TCT ATG TTA GAC TCA GAA     2535
Tyr Asp Arg Pro Ala Ser Tyr Lys Lys Lys Ser Met Leu Asp Ser Glu
            765                 770                 775

GTC AAA AAC CTC CTT TCA GAT GAT AAC TCA GAA GGC CTT ACT TTA TTG     2583
Val Lys Asn Leu Leu Ser Asp Asp Asn Ser Glu Gly Leu Thr Leu Leu
        780                 785                 790

GAT TTG TTG AGC TTC ACC TAT CAA GTT GCC CGA GGA ATG GAG TTT TTG     2631
Asp Leu Leu Ser Phe Thr Tyr Gln Val Ala Arg Gly Met Glu Phe Leu
    795                 800                 805

GCT TCA AAA AAT TGT GTC CAC CGT GAT CTG GCT GCT CGC AAC GTC CTC     2679
Ala Ser Lys Asn Cys Val His Arg Asp Leu Ala Ala Arg Asn Val Leu
810                 815                 820                 825

CTG GCA CAA GGA AAA ATT GTG AAG ATC TGT GAC TTT GGC CTG GCC AGA     2727
Leu Ala Gln Gly Lys Ile Val Lys Ile Cys Asp Phe Gly Leu Ala Arg
                830                 835                 840

GAC ATC ATG CAT GAT TCG AAC TAT GTG TCG AAA GGC AGT ACC TTT CTG     2775
Asp Ile Met His Asp Ser Asn Tyr Val Ser Lys Gly Ser Thr Phe Leu
            845                 850                 855

CCC GTG AAG TGG ATG GCT CCT GAG AGC ATC TTT GAC AAC CTC TAC ACC     2823
Pro Val Lys Trp Met Ala Pro Glu Ser Ile Phe Asp Asn Leu Tyr Thr
        860                 865                 870

ACA CTG AGT GAT GTC TGG TCT TAT GGC ATT CTG CTC TGG GAG ATC TTT     2871
Thr Leu Ser Asp Val Trp Ser Tyr Gly Ile Leu Leu Trp Glu Ile Phe
    875                 880                 885

TCC CTT GGT GGC ACC CCT TAC CCC GGC ATG ATG GTG GAT TCT ACT TTC     2919
Ser Leu Gly Gly Thr Pro Tyr Pro Gly Met Met Val Asp Ser Thr Phe
890                 895                 900                 905

TAC AAT AAG ATC AAG AGT GGG TAC CGG ATG GCC AAG CCT GAC CAC GCT     2967
Tyr Asn Lys Ile Lys Ser Gly Tyr Arg Met Ala Lys Pro Asp His Ala
                910                 915                 920

ACC AGT GAA GTC TAC GAG ATC ATG GTG AAA TGC TGG AAC AGT GAG CCG     3015
Thr Ser Glu Val Tyr Glu Ile Met Val Lys Cys Trp Asn Ser Glu Pro
            925                 930                 935

GAG AAG AGA CCC TCC TTT TAC CAC CTG AGT GAG ATT GTG GAG AAT CTG     3063
Glu Lys Arg Pro Ser Phe Tyr His Leu Ser Glu Ile Val Glu Asn Leu
        940                 945                 950

CTG CCT GCA CAA TAT AAA AAG AGT TAT GAA AAA ATT CAC CTG GAC TTC     3111
Leu Pro Gly Gln Tyr Lys Lys Ser Tyr Glu Lys Ile His Leu Asp Phe
    955                 960                 965
```

-continued

```
CTG AAG AGT GAC CAT CCT GCT GTG GCA CGC ATG CGT GTG GAC TCA GAC     3159
Leu Lys Ser Asp His Pro Ala Val Ala Arg Met Arg Val Asp Ser Asp
970                 975                 980                 985

AAT GCA TAC ATT GGT GTC ACC TAC AAA AAC GAG GAA GAC AAG CTG AAG     3207
Asn Ala Tyr Ile Gly Val Thr Tyr Lys Asn Glu Glu Asp Lys Leu Lys
                990                 995                1000

GAC TGG GAG GGT GGT CTG GAT GAG CAG AGA CTG AGC GCT GAC AGT GGC     3255
Asp Trp Glu Gly Gly Leu Asp Glu Gln Arg Leu Ser Ala Asp Ser Gly
            1005                1010                1015

TAC ATC ATT CCT CTG CCT GAC ATT GAC CCT GTC CCT GAG GAG GAG GAC     3303
Tyr Ile Ile Pro Leu Pro Asp Ile Asp Pro Val Pro Glu Glu Glu Asp
        1020                1025                1030

CTG GGC AAG AGG AAC AGA CAC AGC TCG CAG ACC TCT GAA GAG AGT GCC     3351
Leu Gly Lys Arg Asn Arg His Ser Ser Gln Thr Ser Glu Glu Ser Ala
    1035                1040                1045

ATT GAG ACG GGT TCC AGC AGT TCC ACC TTC ATC AAG AGA GAG GAC GAG     3399
Ile Glu Thr Gly Ser Ser Ser Ser Thr Phe Ile Lys Arg Glu Asp Glu
1050                1055                1060                1065

ACC ATT GAA GAC ATC GAC ATG ATG GAC GAC ATC GGC ATA GAC TCT TCA     3447
Thr Ile Glu Asp Ile Asp Met Met Asp Asp Ile Gly Ile Asp Ser Ser
                1070                1075                1080

GAC CTG GTG GAA GAC AGC TTC CTG TAACTGGCGG ATTCGAGGGG TTCCTTCCAC    3501
Asp Leu Val Glu Asp Ser Phe Leu
            1085

TTCTGGGGCC ACCTCTGGAT CCCGTTCAGA AAACCACTTT ATTGCAATGC GGAGGTTGAG   3561

AGGAGGACTT GGTTGATGTT TAAAGAGAAG TTCCCAGCCA AGGGCCTCGG GGAGCGTTCT   3621

AAATATGAAT GAATGGGATA TTTTGAAATG AACTTTGTCA GTGTTGCCTC TTGCAATGCC   3681

TCAGTAGCAT CTCAGTGGTG TGTGAAGTTT GGAGATAGAT GGATAAGGGA ATAATAGGCC   3741

ACAGAAGGTG AACTTTGTGC TTCAAGGACA TTGGTGAGAG TCCAACAGAC ACAATTTATA   3801

CTGCGACAGA ACTTCAGCAT TGTAATTATG TAAATAACTC TAACCAAGGC TGTGTTTAGA   3861

TTGTATTAAC TATCTTCTTT GGACTTCTGA AGAGACCACT CAATCCATCC TGTACTTCCC   3921

TCTTGAAACC TGATGTAGCT GCTGTTGAAC TTTTTAAAGA AGTGCATGAA AAACCATTTT   3981

TGAACCTTAA AAGGTACTGG TACTATAGCA TTTTGCTATC TTTTTTAGTG TTAAAGAGAT   4041

AAAGAATAAT AAG                                                      4054
```

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1089 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

```
Met Gly Thr Ser His Pro Ala Phe Leu Val Leu Gly Cys Leu Leu Thr
 1               5                  10                  15

Gly Leu Ser Leu Ile Leu Cys Gln Leu Ser Leu Pro Ser Ile Leu Pro
            20                  25                  30

Asn Glu Asn Glu Lys Val Val Gln Leu Asn Ser Ser Phe Ser Leu Arg
        35                  40                  45

Cys Phe Gly Glu Ser Glu Val Ser Trp Gln Tyr Pro Met Ser Glu Glu
    50                  55                  60

Glu Ser Ser Asp Val Glu Ile Arg Asn Glu Glu Asn Asn Ser Gly Leu
65                  70                  75                  80

Thr Val Thr Val Leu Glu Val Ser Ser Ala Ser Ala Ala His Thr Gly
```

-continued

```
                85                  90                  95

Leu Tyr Thr Cys Tyr Tyr Asn His Thr Gln Thr Glu Glu Asn Glu Leu
            100                 105                 110

Glu Gly Arg His Ile Tyr Ile Tyr Val Pro Asp Pro Asp Val Ala Phe
        115                 120                 125

Val Pro Leu Gly Met Thr Asp Tyr Leu Val Ile Val Glu Asp Asp Asp
    130                 135                 140

Ser Ala Ile Ile Pro Cys Arg Thr Thr Asp Pro Glu Thr Pro Val Thr
145                 150                 155                 160

Leu His Asn Ser Glu Gly Val Val Pro Ala Ser Tyr Asp Ser Arg Gln
                165                 170                 175

Gly Phe Asn Gly Thr Phe Thr Val Gly Pro Tyr Ile Cys Glu Ala Thr
            180                 185                 190

Val Lys Gly Lys Lys Phe Gln Thr Ile Pro Phe Asn Val Tyr Ala Leu
        195                 200                 205

Lys Ala Thr Ser Glu Leu Asp Leu Glu Met Glu Ala Leu Lys Thr Val
    210                 215                 220

Tyr Lys Ser Gly Glu Thr Ile Val Val Thr Cys Ala Val Phe Asn Asn
225                 230                 235                 240

Glu Val Val Asp Leu Gln Trp Thr Tyr Pro Gly Glu Val Lys Gly Lys
                245                 250                 255

Gly Ile Thr Ile Leu Glu Glu Ile Lys Val Pro Ser Ile Lys Leu Val
            260                 265                 270

Tyr Thr Leu Thr Val Pro Glu Ala Thr Val Lys Asp Ser Gly Asp Tyr
        275                 280                 285

Glu Cys Ala Ala Arg Gln Ala Thr Arg Glu Val Lys Glu Met Lys Lys
    290                 295                 300

Val Thr Ile Ser Val His Glu Lys Gly Phe Ile Glu Ile Lys Pro Thr
305                 310                 315                 320

Phe Ser Gln Leu Glu Ala Val Asn Leu His Glu Val Lys His Phe Val
                325                 330                 335

Val Glu Val Arg Ala Tyr Pro Pro Pro Arg Ile Ser Trp Leu Lys Asn
            340                 345                 350

Asn Leu Thr Leu Ile Glu Asn Leu Thr Glu Ile Thr Thr Asp Val Glu
        355                 360                 365

Lys Ile Gln Glu Ile Arg Tyr Arg Ser Lys Leu Lys Leu Ile Arg Ala
    370                 375                 380

Lys Glu Glu Asp Ser Gly His Tyr Thr Ile Val Ala Gln Asn Glu Asp
385                 390                 395                 400

Ala Val Lys Ser Tyr Thr Phe Glu Leu Leu Thr Gln Val Pro Ser Ser
                405                 410                 415

Ile Leu Asp Leu Val Asp Asp His His Gly Ser Thr Gly Gly Gln Thr
            420                 425                 430

Val Arg Cys Thr Ala Glu Gly Thr Pro Leu Pro Asp Ile Glu Trp Met
        435                 440                 445

Ile Cys Lys Asp Ile Lys Lys Cys Asn Asn Glu Thr Ser Trp Thr Ile
    450                 455                 460

Leu Ala Asn Asn Val Ser Asn Ile Ile Thr Glu Ile His Ser Arg Asp
465                 470                 475                 480

Arg Ser Thr Val Glu Gly Arg Val Thr Phe Ala Lys Val Glu Glu Thr
                485                 490                 495

Ile Ala Val Arg Cys Leu Ala Lys Asn Leu Leu Gly Ala Glu Asn Arg
            500                 505                 510
```

-continued

Glu Leu Lys Leu Val Ala Pro Thr Leu Arg Ser Glu Leu Thr Val Ala
515 520 525

Ala Ala Val Leu Val Leu Leu Val Ile Val Ile Ile Ser Leu Ile Val
530 535 540

Leu Val Val Ile Trp Lys Gln Lys Pro Arg Tyr Glu Ile Arg Trp Arg
545 550 555 560

Val Ile Glu Ser Ile Ser Pro Asp Gly His Glu Tyr Ile Tyr Val Asp
565 570 575

Pro Met Gln Leu Pro Tyr Asp Ser Arg Trp Glu Phe Pro Arg Asp Gly
580 585 590

Leu Val Leu Gly Arg Val Leu Gly Ser Gly Ala Phe Gly Lys Val Val
595 600 605

Glu Gly Thr Ala Tyr Gly Leu Ser Arg Ser Gln Pro Val Met Lys Val
610 615 620

Ala Val Lys Met Leu Lys Pro Thr Ala Arg Ser Ser Glu Lys Gln Ala
625 630 635 640

Leu Met Ser Glu Leu Lys Ile Met Thr His Leu Gly Pro His Leu Asn
645 650 655

Ile Val Asn Leu Leu Gly Ala Cys Thr Lys Ser Gly Pro Ile Tyr Ile
660 665 670

Ile Thr Glu Tyr Cys Phe Tyr Gly Asp Leu Val Asn Tyr Leu His Lys
675 680 685

Asn Arg Asp Ser Phe Leu Ser His His Pro Glu Lys Pro Lys Lys Glu
690 695 700

Leu Asp Ile Phe Gly Leu Asn Pro Ala Asp Glu Ser Thr Arg Ser Tyr
705 710 715 720

Val Ile Leu Ser Phe Glu Asn Asn Gly Asp Tyr Met Asp Met Lys Gln
725 730 735

Ala Asp Thr Thr Gln Tyr Val Pro Met Leu Glu Arg Lys Glu Val Ser
740 745 750

Lys Tyr Ser Asp Ile Gln Arg Ser Leu Tyr Asp Arg Pro Ala Ser Tyr
755 760 765

Lys Lys Lys Ser Met Leu Asp Ser Glu Val Lys Asn Leu Leu Ser Asp
770 775 780

Asp Asn Ser Glu Gly Leu Thr Leu Leu Asp Leu Leu Ser Phe Thr Tyr
785 790 795 800

Gln Val Ala Arg Gly Met Glu Phe Leu Ala Ser Lys Asn Cys Val His
805 810 815

Arg Asp Leu Ala Ala Arg Asn Val Leu Leu Ala Gln Gly Lys Ile Val
820 825 830

Lys Ile Cys Asp Phe Gly Leu Ala Arg Asp Ile Met His Asp Ser Asn
835 840 845

Tyr Val Ser Lys Gly Ser Thr Phe Leu Pro Val Lys Trp Met Ala Pro
850 855 860

Glu Ser Ile Phe Asp Asn Leu Tyr Thr Thr Leu Ser Asp Val Trp Ser
865 870 875 880

Tyr Gly Ile Leu Leu Trp Glu Ile Phe Ser Leu Gly Gly Thr Pro Tyr
885 890 895

Pro Gly Met Met Val Asp Ser Thr Phe Tyr Asn Lys Ile Lys Ser Gly
900 905 910

Tyr Arg Met Ala Lys Pro Asp His Ala Thr Ser Glu Val Tyr Glu Ile
915 920 925

Met Val Lys Cys Trp Asn Ser Glu Pro Glu Lys Arg Pro Ser Phe Tyr
930 935 940

-continued

```
His Leu Ser Glu Ile Val Glu Asn Leu Leu Pro Gly Gln Tyr Lys Lys
945                 950                 955                 960

Ser Tyr Glu Lys Ile His Leu Asp Phe Leu Lys Ser Asp His Pro Ala
                965                 970                 975

Val Ala Arg Met Arg Val Asp Ser Asp Asn Ala Tyr Ile Gly Val Thr
            980                 985                 990

Tyr Lys Asn Glu Glu Asp Lys Leu Lys Asp Trp Glu Gly Gly Leu Asp
        995                 1000                1005

Glu Gln Arg Leu Ser Ala Asp Ser Gly Tyr Ile Ile Pro Leu Pro Asp
    1010                1015                1020

Ile Asp Pro Val Pro Glu Glu Glu Asp Leu Gly Lys Arg Asn Arg His
1025                1030                1035                1040

Ser Ser Gln Thr Ser Glu Glu Ser Ala Ile Glu Thr Gly Ser Ser Ser
                1045                1050                1055

Ser Thr Phe Ile Lys Arg Glu Asp Glu Thr Ile Glu Asp Ile Asp Met
            1060                1065                1070

Met Asp Asp Ile Gly Ile Asp Ser Ser Asp Leu Val Glu Asp Ser Phe
        1075                1080                1085

Leu
```

We claim:

1. A biologically active, dimerized polypeptide fusion, comprising:

first and second polypeptide chains, wherein each of said polypeptide chains comprises a non-immunoglobulin polypeptide requiring dimerization for biological activity joined to a dimerizing protein heterologous to said non-immunoglobulin polypeptide.

2. The biologically active, dimerized polypeptide fusion of claim 1 which is a homodimer.

3. The biologically active, dimerized polypeptide fusion of claim 1 wherein the dimerizing protein of one of said polypeptide chains comprises an immunoglobulin heavy chain constant region.

4. The biologically active, dimerized polypeptide fusion of claim 3 wherein the immunoglobulin heavy chain constant region is joined to an immunoglobulin hinge region.

5. The biologically active, dimerized polypeptide fusion of claim 3 wherein the immunoglobulin heavy chain constant region is joined to an immunoglobulin variable region.

6. The biologically active, dimerized polypeptide fusion of claim 5 wherein the immunoglobulin variable region is selected from the group consisting of V_H, Vκ, and Vλ.

7. The biologically active, dimerized polypeptide fusion of claim 1 wherein the dimerizing protein one of said polypeptide chains comprises an immunoglobulin heavy chain constant region domain selected from the group consisting of C_H1, C_H2, C_H3, and C_H4 of a γ, α, ε, μ, or δ class immunoglobulin heavy chain.

8. The biologically active, dimerized polypeptide fusion of claim 1 wherein the dimerizing protein of one of said polypeptide chains comprises an immunoglobulin light chain constant region.

9. A biologically active, multimerized polypeptide fusion, comprising:

a non-immunoglobulin polypeptide requiring multimerization for biological activity joined to an immunoglobulin light chain constant region; and

an immunoglobulin heavy chain constant region domain selected from the group consisting of C_H1, C_H2, C_H3, and C_H4.

10. The biologically active, multimerized polypeptide fusion of claim 9 which is a tetramer comprising four polypeptide fusions each having a non-immunoglobulin polypeptide joined to a multimerizing protein.

11. The biologically active, multimerized polypeptide fusion of claim 9 including a multimerizing protein which comprises an immunoglobulin heavy chain constant region domain selected from the group consisting of C_H1, C_H2, C_H3, and C_H4 of a γ, α, ε, μ, or δ class immunoglobulin heavy chain.

12. A biologically active, multimerized polypeptide fusion comprising: a non-immunoglobulin polypeptide requiring multimerization for biological activity joined to an immunoglobulin heavy chain constant region; and

an immunoglobulin heavy chain.

13. The biologically active, multimerized polypeptide fusion of claim 12 wherein the immunoglobulin heavy chain constant region is joined to an immunoglobulin hinge region.

14. The biologically active, multimerized polypeptide fusion of claim 12 wherein the immunoglobulin heavy chain constant region is joined to an immunoglobulin variable region.

15. The biologically active, multimerized polypeptide fusion of claim 14 wherein the immunoglobulin variable region is selected from the group consisting of V_H, Vκ, and Vλ.

16. A biologically active, heteromultimeric polypeptide fusion, comprising:

a first polypeptide fusion comprising a first non-immunoglobulin polypeptide joined to a first multimerizing protein heterologous to said first non-immunoglobulin polypeptide and a second polypeptide fusion comprising a second non-immunoglobulin polypeptide joined to a second multimerizing protein heterologous to said second non-immunoglobulin polypeptide, wherein said first multimerizing protein and said second multimerizing protein are capable of associating to form a multimer.

17. The biologically active, heteromultimeric polypeptide fusion of claim 16 wherein the first and second non-

immunoglobulin polypeptides each comprise an amino acid sequence selected from the group consisting of (A) the amino acid sequence Sequence ID NO:2, and (B) the amino acid sequence of Sequence ID NO:36.

18. The biologically active, heteromultimeric polypeptide fusion of claim 16 wherein the first multimerizing protein is different from the second multimerizing protein.

19. The biologically active, heteromultimeric polypeptide fusion of claim 18 wherein the first and second non-immunoglobulin polypeptides are the same.

20. The biologically active, heteromultimeric polypeptide fusion of claim 16 wherein the first and second multimerizing proteins each comprise an immunoglobulin heavy chain constant region or an immunoglobulin light chain constant region.

21. The biologically active, heteromultimeric polypeptide fusion of claim 16 which comprises a first polypeptide fusion having a first non-immunoglobulin polypeptide joined to a first immunoglobulin constant region and a second polypeptide fusion having a second non-immunoglobulin polypeptide fused to a second immunoglobulin constant region different from the first immunoglobulin constant region.

22. The biologically active, heteromultimeric polypeptide fusion of claim 21 wherein the first multimerizing protein comprises an immunoglobulin heavy chain constant region and the second multimerizing protein comprises an immunoglobulin light chain constant region.

23. The biologically active, heteromultimeric polypeptide fusion of claim 20 wherein one of said multimerizing proteins comprises an immunoglobulin heavy chain constant region joined to an immunoglobulin hinge region.

24. The biologically active, heteromultimeric polypeptide fusion of claim 20 wherein one of said multimerizing proteins comprises an immunoglobulin heavy chain constant region joined to an immunoglobulin variable region.

25. The biologically active, heteromultimeric polypeptide fusion of claim 24 wherein the immunoglobulin variable region is selected from the group consisting of V_H, Vκ, and Vλ.

26. The biologically active, heteromultimeric polypeptide fusion of claim 16 wherein one of said multimerizing proteins comprises an immunoglobulin heavy chain constant region domain selected from the group consisting of C_H1, C_H2, C_H3, and C_H4 of a γ, α, ε, μ, or δ class immunoglobulin heavy chain.

27. The biologically active, heteromultimeric polypeptide fusion of claim 16 wherein said multimerized polypeptide fusion comprises a first polypeptide fusion comprising a first receptor or receptor domain requiring multimerization for activity joined to a first immunoglobulin constant region and a second polypeptide fusion comprising a second receptor or receptor domain requiring multimerization for activity joined to a second immunoglobulin constant region.

28. The biologically active, heteromultimeric polypeptide fusion of claim 27 wherein one of said receptor domains comprises a ligand binding domain.

29. The biologically active, heteromultimeric polypeptide fusion of claim 16 which is a heterotetramer comprising four polypeptide fusions each having a non-immunoglobulin polypeptide joined to a multimerizing protein.

* * * * *

EXHIBIT 12

US006004781A

United States Patent [19]

Seed

[11] Patent Number: 6,004,781

[45] Date of Patent: Dec. 21, 1999

[54] **NUCLEIC ACID ENCODING IG-CD4 FUSION PROTEINS**

[75] Inventor: **Brian Seed**, Boston, Mass.

[73] Assignee: **The General Hospital Corporation**, Charlestown, Mass.

[21] Appl. No.: **08/191,708**

[22] Filed: **Feb. 4, 1994**

Related U.S. Application Data

[63] Continuation of application No. 08/057,952, Apr. 12, 1993, abandoned, which is a continuation of application No. 07/896,781, Jun. 9, 1992, abandoned, which is a continuation of application No. 07/299,596, Jan. 23, 1989, abandoned, which is a continuation-in-part of application No. 07/147,351, Jan. 22, 1988, abandoned.

[51] **Int. Cl.**[6] **C12N 15/62**

[52] **U.S. Cl.** **435/69.7**; 435/252.3; 435/320.1; 536/23.4

[58] **Field of Search** 435/69.7, 252.3, 435/320.1; 536/23.4

[56] **References Cited**

U.S. PATENT DOCUMENTS

4,675,382	6/1987	Murphy	530/350
4,816,567	3/1989	Cabilly et al.	530/387
5,336,603	8/1994	Capon et al.	435/69.7

FOREIGN PATENT DOCUMENTS

0 120 694	10/1984	European Pat. Off. .
0 125 023	11/1984	European Pat. Off. .
0314317 A1	5/1989	European Pat. Off. .
WO88/01304	2/1988	WIPO .
WO89/01940	3/1989	WIPO .
WO89/02922	4/1989	WIPO .

OTHER PUBLICATIONS

Deen, K.C., et al., "A soluble form of CD4 (T4) protein inhibits AIDS virus infection," Nature 331:82–84 (Jan. 7, 1988).

Fisher, R.A., et al., "HIV infection is blocked in vitro by recombinant soluble CD4," Nature 331:76–78 (Jan. 7, 1988).

Article entitled "Prevention of HIV–1 IIIB Infection in Chimpanzees by CD4 Immunoadhesin" by Ward et al., Nature, 352:434–436 (1991).

Article entitled "Prevention of HIV–2 and SIVsm Infection by Passive Immunization in Cynomolgus Monkeys" by Putkonen et al., Nature 352:436–438 (1991).

Mizukami, T., et al., "Expression and Characterization of Chimeric Proteins Containing Human CD4 Linked to Human Immunoglobulin Heavy Chain Constant Regions," in Morisset, R.A. (ed.) V International Conference on AIDS, Abstract No. M.C.P. 89 (Montreal, Canada 1989), p. 556.

Morrison, et al., "Chimeric human antibody molecules: Mouse antigen–binding domains with human constant region domains," Proc. Natl. Acad. Sci. USA 81:6851–6855, 1984.

McDougal, et al., "Binding of HTLV–III/LAV to T4+ T Cells by a Complex of the 110K Viral Protein and the T4 Molecule," Science 231:382–385, 1986.

Maddon, et al., "The Isolation and Nucleotide Sequence of a cDNA Encoding the T Cell Surface Protein T4: A New Member of the Immunoglobulin Gene Family," Cell 42:93–104, 1985.

Capon, et al., "Designing CD4 Immunoadhesine for AIDS Therapy," Nature 337:525–531 1989.

Gascoigne, N.R.J. et al., "Secretion of a Chimeric T–cell Receptor–immunoglobulin protein," Proc. Natl. Acad. Sci. (USA) 84:2936–2940 (1987).

Lasky, L. A. et al., "Delineation of a Region of the Human Immuno–deficiency Virus Type 1 gp120 Glycoprotein Critical for Interaction with the CD4 Receptor," Cell 50:975–985 (1987).

Traunecker, A. et al., "Soluble CD4 Molecules Neutralize Human Immuno–deficiency Virus Type 1," Nature 331:84–86 (1988).

Kowalski, M. et al., "Functional Regions of the Envelope Glycoprotein of Human Immunodeficiency Virus Type I," Science 237:1351–1355 (1987).

Stricker, et al., "An AIDS–related Cytotoxic Autoantibody Reacts with a Specific Antigen on Stimulated CD4+ T Cells," Nature 327:710–713 (1987).

Smith, D. H. et al., "Blocking of HIV–1 Infectivity by a Soluble, Secreted Form of the CD4 Antigen," Science 238:1704–1707 (1987).

Estess, P. et al., "Analysis of T–Cell Receptor Structure and Function Using Chimeric T–Cell Receptor/Immunoglobulin Molecules," J. Cell. Biochem. (Suppl. 11D):Abstract 331, p. 258 (1987).

Clark, S. et al., "Peptide and nucleotide sequences of rat CD4 (W3/25) antigen: Evidence for derivation from a structure with four immuno–globulin–related domains," Proc. Natl. Acad. Sci. (USA) 84:1649 (1987).

Palker, T. J. et al., "A Conserved Region at the COOH Terminus of Human Immunodeficiency Virus gp120 Envelope Protein Contains an Immunodominant Epitope," Proc. Natl. Acad. Sci. (USA) 84:2479–2483 (1987).

Neuberger, M. S. et al., "Recombinant Antibodies Possessing Novel Effector Functions," Nature 312:604–608 (1984).

(List continued on next page.)

Primary Examiner—John Ulm

Attorney, Agent, or Firm—Greenlee, Winner and Sullivan, PC

[57] **ABSTRACT**

The invention relates to a fusion protein which comprises an immunoglobulin of the IgM, IgG1 or IgG3 immunoglobulin class, wherein the variable region of the light or heavy chain has been replaced with CD4 or fragment thereof which is capable of binding to gp120. The invention also relates to an immunoglobulin-like molecule comprising the fusion protein of the invention together with an immunoglobulin light or heavy chain. The invention also relates to a method of treating HIV or SIV infection comprising administering the fusion proteins or immunoglobulin-like molecules of the invention to an animal. The invention also relates to assays for HIV or SIV comprising contacting a sample suspected of containing HIV or SIV gp120 with the immunoglobulin-like molecule or fusion protein of the invention, and detecting whether a complex is formed.

14 Claims, No Drawings

OTHER PUBLICATIONS

Rusche, J. A. et al., "Antibodies that Inhibit Fusion of Human Immuno-deficiency Virus-Infected Cells Bind a 24-Amino Acid Sequence of the Viral Envelope, gp120," Proc. Natl. Acad. Sci. (USA) 85:3198-3202 (1988).

Boulianne, G. L. et al., "Production of Functional Chimaeric Mouse/Human Antibody," Nature 312:643-646 (1984).

The International Search Report for corresponding International Application No. PCT/US89/00238.

NUCLEIC ACID ENCODING IG-CD4 FUSION PROTEINS

CROSS-REFERENCE TO RELATED APPLICATION

This application is a continuation of application Ser. No. 08/057,952, filed Apr. 12, 1993, now abandoned, which is a continuation application of Ser. No. 07/896,781, filed Jun. 9, 1992, now abandoned, which is a continuation of application Ser. No. 07/299,596, filed Jan. 23, 1989, now abandoned, which is a continuation-in-part of application Ser. No. 07/147,351, filed Jan. 22, 1988, now abandoned.

FIELD OF THE INVENTION

The invention is in the field of recombinant genetics.

BACKGROUND OF THE INVENTION

The human and simian immunodeficiency viruses HIV and SIV are the causative agents of Acquired Immune Deficiency Syndrome (AIDS) and Simian Immunodeficiency Syndrome (SIDS), respectively. See Curren, J. et al., *Science* 329:1359–1357 (1985); Weiss, R. et al., *Nature* 324:572–575 (1986). The HIV virus contains an envelope glycoprotein, gp120 which binds to the CD4 protein present on the surface of helper T lymphocytes, macrophages and other cells. Dalgleish et al. *Nature*, 312:763 (1984). After the gp120 binds to CD4, virus entry is facilitated by an envelope-mediated fusion of the viral target cell membranes.

During the course of infection, the host organism develops antibodies against viral proteins, including the major envelope glycoproteins gp120 and gp41. Despite this humoral immunity, the disease progresses, resulting in a lethal immunosuppression characterized by multiple opportunistic infections, parasitemia, dementia and death. The failure of host anti-viral antibodies to arrest the progression of the disease represents one of the most vexing and alarming aspects of the infection, and augurs poorly for vaccination efforts based upon conventional approaches.

Two factors may play a role in the inefficacy of the humoral response to immunodeficiency viruses. First, like other RNA viruses (and like retroviruses in particular), the immunodeficiency viruses show a high mutation rate which allows antigenic variation to progress at a high rate in response to host immune surveillance. Second, the envelope glycoproteins themselves are heavily glycosylated molecules presenting few epitopes suitable for high affinity antibody binding. The poorly antigenic, "moving" target which the viral envelope presents, allows the host little opportunity for restricting viral infection by specific antibody production.

Cells infected by HIV express the gp120 glycoprotein on their surface. Gp120 mediates fusion events among $CD4^+$ cells via a reaction similar to that by which the virus enters the uninfected cell, leading to the formation of short-lived multinucleated giant cells. Syncytium formation is dependent on a direct interaction of the gp120 envelope glycoprotein with the CD4 protein. Dalgleish et al., supra, Klatzmann, D. et al., *Nature* 312:763 (1984); McDougal, J. S. et al. *Science*, 231:382 (1986); Sodroski, J. et al., *Nature*, 322:470 (1986); Lifson, J. D. et al., *Nature*, 323:725 (1986); Sodroski, J. et al., *Nature*, 321:412 (1986).

The CD4 protein consists of a 370 amino acid extracellular region containing four immunoglobulin-like domains, a membrane spanning domain, and a charged intracellular region of 40 amino acid residues. Maddon, P. et al., *Cell* 42:93 (1985); Clark, S. et al., *Proc. Natl. Acad. Sci.* (USA) 84:1649 (1987).

Evidence that CD4-gp120 binding is responsible for viral infection of cells bearing the CD4 antigen includes the finding that a specific complex is formed between gp120 and CD4. McDougal et al., supra. Other workers have shown that cell lines, which were non-infective for HIV, were converted to infectable cell lines following transfection and expression of the human CD4 cDNA gene. Maddon et al., *Cell* 47:333–348 (1986).

In contrast to the majority of antibody-envelope interactions, the receptor-envelope interaction is characterized by a high affinity ($K_a \approx 10^8$l/mole) immutable association. Moreover, the affinity of the virus for CD4 is at least 3 orders of magnitude higher than the affinity of CD4 for its putative endogenous ligand, the MHC class II antigens. Indeed, to date, a specific physical association between monomeric CD4 and class II antigens has not been demonstrated.

In response to bacterial or other particle infection, the host organism usually produces serum antibodies that bind to specific proteins or carbohydrates on the bacterial or particle surface, coating the bacteria. This antibody coat on the bacterium or other particle stimulates cytolysis by Fc-receptor-bearing lymphoid cells by antibody-dependent cellular toxicity (ADCC). Other serum proteins, collectively called complement (C), bind to antibody-coated targets, and also can coat foreign particles nonspecifically. They cause cell death by lysis, or stimulate ingestion by binding to specific receptors on the macrophage called complement receptors. See Darnell J. et al., in *Molecular Cell Biology*, Scientific American Books, pp. 641 and 1087 (1986).

The most effective complement activating classes of human Ig are IgM and IgG1. The complement system consists of 14 proteins that, acting in order, cause lysis of cells. Nearly all of the C proteins exist in normal serum as inactive precursors. When activated, some become highly specific proteolytic enzymes whose substrate is the next protein in a sequential chain reaction.

The entire C sequence can be triggered by either of two initiation pathways. In one (the classic pathway), Ab—Ag complexes bind and activate C1, C4 and C2 to form a C3-splitting enzyme. In the second pathway, polysaccharides commonly on the surface of many bacteria and fungi bind with trace amounts of a C3 fragment and then with two other proteins (factor B and properdin) to form another C3-splitting enzyme. Once C3 is split by either pathway, the way is open for the remaining sequence of steps which lead to cell lysis. See Davis, B. D., et al., In *Microbiology*, 3rd ed., Harper and Row, Philadelphia, Pa., pp. 452–466 (1980).

A number of workers have disclosed methods for preparing hybrid proteins. For example, Murphy, U.S. Pat. No. 4,675,382 (1987), discloses the use of recombinant DNA techniques to make hybrid protein molecules by forming the desired fused gene coding for a hybrid protein of diptheria toxin and a polypeptide ligand such as a hormone, followed by expression of the fused gene.

Many workers have prepared monoclonal antibodies (Mabs) by recombinant DNA techniques. Monoclonal antibodies are highly specific well-characterized molecules in both primary and tertiary structure. They have been widely used for in vitro immunochemical characterization and quantitation of antigens. Genes for heavy and light chains have been introduced into appropriate hosts and expressed, followed by reaggregation of the individual chains into functional antibody molecules (see, for example, Munro,

Nature 312:597 (1984); Morrison, S. L., *Science* 229:1202 (1985); Oi et al., *Biotechniques* 4:214 (1986); Wood et al., *Nature* 314:446–449 (1985)). Light- and heavy-chain variable regions have been cloned and expressed in foreign hosts wherein they maintained their binding ability (Moore et al., European Patent Application 0088994 (published Sep. 21, 1983)).

Chimeric or hybrid antibodies have also been prepared by recombinant DNA techniques. Oi and Morrison, *Biotechniques* 4:214 (1986) describe a strategy for producing such chimeric antibodies which include a chimeric human IgG anti-leu3 antibody.

Gascoigne, N. R. J., et al., *Proc. Natl. Acad. Sci.* (USA) 84:2936–2940 (1987) disclose the preparation of a chimeric gene construct containing a T-cell receptor α-chain variable (V) domain and the constant (C) region coding sequence of an immunoglobulin γ2a molecule. Cells transfected with the chimeric gene synthesize a protein product that expresses immunoglobulin and T-cell receptor antigenic determinants as well as protein A binding sites. This protein associates with a normal λ chain to form an apparently normal tetrameric (H_2L_2, where H=heavy and L=light) immunoglobulin molecule that is secreted.

Sharon, J., et al., *Nature* 309:54 (1984), disclose construction of a chimeric gene encoding the variable (V) region of a mouse heavy chain specific for the hapten azophenylarsonate and the constant (C) region of a mouse kappa light chain (V_HC_K). This gene was introduced into a mouse myeloma cell line. The chimeric gene was expressed to give a protein which associated with light chains secreted from the myeloma cell line to give an antibody molecule specific for azophenylarsonate.

Morrison, *Science* 229:1202 (1985), discloses that variable light- or variable heavy-chain regions can be attached to a non-Ig sequence to create fusion proteins. This article states that the potential uses for the fusion proteins are three: (1) to attach antibody specifically to enzymes for use in assays; (2) to isolate non-Ig proteins by antigen columns; and (3) to specifically deliver toxic agents.

Recent techniques for the stable introduction of immunoglobulin genes into myeloma cells (Banerji, J., et al., *Cell* 33:729–740 (1983); Potter, H., et al., *Proc. Natl. Acad. Sci.* (USA) 81:7161–7165 (1984)), coupled with detailed structural information, have permitted the use of in vitro DNA methods such as mutagenesis, to generate recombinant antibodies possessing novel properties.

PCT Application WO87/02671 discloses methods for producing genetically engineered antibodies of desired variable region specificity and constant region properties through gene cloning and expression of light and heavy chains. The mRNA from cloned hybridoma B cell lines which produce monoclonal antibodies of desired specificity is isolated for cDNA cloning. The generation of light and heavy chain coding sequences is accomplished by excising the cloned variable regions and ligating them to light or heavy chain module vectors. This gives cDNA sequences which code for immunoglobulin chains. The lack of introns allows these cDNA sequences to be expressed in prokaryotic hosts, such as bacteria, or in lower eukaryotic hosts, such as yeast.

The generation of chimeric antibodies in which the antigen-binding portion of the immunoglobulin is fused to other moieties has been demonstrated. Examples of non-immunoglobulin genes fused to antibodies include *Staphylococcus aureus* nuclease, the mouse oncogene c-myc, and the Klenow fragment of *E. coli* DNA polymerase I (Neuberger, M. S., et al., *Nature* 312:604–612 (1984); Neuberger, M. S., *Trends in Biochemical Science,* 347–349 (1985)). European Patent Application 120,694 discloses the genetic engineering of the variable and constant regions of an immunoglobulin molecule that is expressed in *E. coli* host cells. It is further disclosed that the immunoglobulin molecule may be synthesized by a host cell with another peptide moiety attached to one of the constant domains. Such peptide moieties are described as either cytotoxic or enzymatic. The application and the examples describe the use of a lambda-like chain derived from a monoclonal antibody which binds to 4-hydroxy-3-nitrophenyl (NP) haptens.

European Patent Application 125,023 relates to the use of recombinant DNA techniques to produce immunoglobulin molecules that are chimeric or otherwise modified. One of the uses described for these immunoglobulin molecules is for whole-body diagnosis and treatment by injection of the antibodies directed to specific target tissues. The presence of the disease can be determined by attaching a suitable label to the antibodies, or the diseased tissue can be attacked by carrying a suitable drug with the antibodies. The application describes antibodies engineered to aid the specific delivery of an agent as "altered antibodies."PCT Application WO83/101533 describes chimeric antibodies wherein the variable region of an immunoglobulin molecule is linked to a portion of a second protein which may comprise the active portion of an enzyme.

Boulianne et al., *Nature* 32:643 (1984) constructed an immunoglobulin gene in which the DNA segments that encode mouse variable regions specific for the hapten trinitrophenol (TNP) are joined to segments that encode human mu and kappa regions. These chimeric genes were expressed to give functional TNP-binding chimeric IgM.

Morrison et al., *P.N.A.S.* (USA) 81:6851 (1984), disclose a chimeric molecule utilizing the heavy-chain variable region exons of an anti-phosphoryl choline myeloma protein G, which were joined to the exons of either human kappa light-chain gene. The genes were transfected into mouse myeloma cell lines, generating transformed cells that produced chimeric mouse-human IgG with antigen-binding function.

Despite the progress that has been achieved on determining the mechanism of HIV infection, a need continues to exist for methods of treating HIV infections.

SUMMARY OF THE INVENTION

The invention relates to a gene comprising a DNA sequence which encodes a fusion protein comprising 1) CD4, or a fragment thereof which binds to HIV gp120, and 2) an immunoglobulin light or heavy chain; wherein said CD4 or HIV gp120-binding fragment thereof replaces the variable region of the light or heavy immunoglobulin chain.

The invention also relates to vectors containing the gene of the invention and hosts transformed with the vectors.

The invention also relates to a method of producing a fusion protein comprising CD4, or fragment thereof which binds to HIV gp120, and an immunoglobulin light or heavy chain, wherein the variable region of the immunoglobulin light or heavy chain has been substituted with CD4, or HIV gp120-binding fragment thereof, which comprises:

cultivating in a nutrient medium under protein producing conditions, a host strain transformed with the vector containing the gene of the invention, said vector further comprising expression signals which are recognized by said host strain and direct expression of said fusion protein, and

recovering the fusion protein so produced.

The invention also relates to a fusion protein comprising CD4, or fragment thereof which is capable of binding to HIV gp120, fused at the C-terminus to a second protein which comprises an immunoglobulin light or heavy chain, wherein the variable region of said light or heavy chain is substituted with CD4 or a HIV gp120 binding fragment thereof.

The invention also relates to an immunoglobulin-like molecule comprising the fusion protein of the invention together with an immunoglobulin light or heavy-chain, wherein said immunoglobulin like molecule binds HIV gp120.

The IgG1 fusion proteins and immunoglobulin-like molecules may be useful for both complement-mediated and cell-mediated (ADCC) immunity, while the IgM fusion proteins are useful principally through complement-mediated immunity.

The invention also relates to a complex between the fusion proteins and immunoglobulin-like molecule of the invention and HIV gp120.

The invention also relates to a method for treating HIV or SIV infections comprising administering the fusion protein or immunoglobulin-like molecule of the invention to an animal.

The invention further relates to a method for detecting HIV gp120 in a sample comprising contacting a sample suspected of containing HIV or gp120 with the fusion protein or immunoglobulin-like molecule of the invention, and detecting whether a complex has formed.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

The invention is directed to a protein gene which comprises

1) a DNA sequence which codes for CD4, or fragment thereof which binds to HIV gp120, fused to
2) a DNA sequence which encodes an immunoglobulin heavy chain.

Preferably, the antibody has effector function.

The invention is also directed to a protein gene which comprises

1) a DNA sequence which codes for CD4, or fragment thereof which binds to HIV gp120, fused to
2) a DNA sequence which encodes an immunoglobulin light chain; wherein said sequence which codes for CD4, or HIV gp120-binding fragment thereof, replaces the variable region of the light immunoglobulin chain.

The invention is also directed to the expression of these novel fusion proteins in transformed hosts and the use thereof to treat and diagnose HIV infections. In particular, the invention relates to expressing said genes in mammalian hosts which express complementary light or heavy chain immunoglobulins to give immunoglobulin-like molecules which have antibody effector function and also bind to HIV or SIV gp120.

The term "antibody effector function" as used herein denotes the ability to fix complement or to activate ADCC.

The fusion proteins and immunoglobulin-like molecules may be administered to an animal for the purpose of treating HIV or SIV infections. By the term "HIV infections" is intended the condition of having AIDS, AIDS related complex (ARC) or where an animal harbors the AIDS virus but does not exhibit the clinical symptoms of AIDS or ARC. By the term "SIV infections" is intended the condition of being infected with simian immunodeficiency virus.

By the term "animal" is intended all animals which may derive benefit from the administration of fusion proteins and immunoglobulin-like molecules of the invention. Foremost among such animals are humans; however, the invention is not intended to be so limited.

By the term "fusion protein" is intended a fused protein comprising CD4, or fragment thereof which is capable of binding to gp120, linked at its C-terminus to an immunoglobulin chain wherein a portion of the N-terminus of the immunoglobulin is replaced with CD4. In general, that portion of immunoglobulin which is deleted is the variable region. The fusion proteins of the invention may also comprise immunoglobulins where more than just the variable region has been deleted and replaced with CD4 or HIV gp120 binding fragment thereof. For example, the V_H and CH1 regions of an immunoglobulin chain may be deleted. Preferably, any amount of the N-terminus of the immunoglobulin heavy chain can be deleted as long as the remaining fragment has antibody effector function. The minimum sequence required for binding complement encompasses domains CH2 and CH3. Joining of Fc portions by the hinge region is advantageous for increasing the efficiency of complement binding.

The CD4 portion of the fusion protein may comprise the complete CD4 sequence, the 370 amino acid extracellular region and the membrane spanning domain, or the extracellular region. The fusion protein may comprise fragments of the extracellular region obtained by cutting the DNA sequence which encodes CD4.

The fusion proteins and immunoglobulin-like molecules may be administered to an animal for the purpose of treating HIV or SIV infections. By the term "HIV infections" is intended the condition of having AIDS, AIDS related complex (ARC) or where an animal harbors the AIDS virus but does not exhibit the clinical symptoms of AIDS or ARC. By the term "SIV infections" is intended the condition of being infected with simian immunodeficiency virus.

By the term "animal" is intended all animals which may derive benefit from the administration of fusion proteins and immunoglobulin-like molecules of the invention. Foremost among such animals are humans; however, the invention is not intended to be so limited.

Where the fusion protein comprises an immunoglobulin light chain, it is necessary that no more of the Ig chain be deleted than is necessary to form a stable complex with a heavy chain Ig. In particular, the cysteine residues necessary for disulfide bond formation must be preserved on both the heavy and light chain moieties.

When expressed in a host, e.g., a mammalian cell, the fusion protein may associate with other light or heavy Ig chains secreted by the cell to give a functioning immunoglobulin-like molecule which is capable of binding to gp120. The gp120 may be in solution, expressed on the surface of infected cells, or may be present on the surface of the HIV virus itself. Alternatively, the fusion protein may be expressed in a mammalian cell which does not secrete other light or heavy Ig chains. When expressed under these conditions, the fusion protein may form a homodimer.

Genomic or CDNA sequences may be used in the practice of the invention. Genomic sequences are expressed efficiently in myeloma cells, since they contain native promoter structures.

The constant regions of the antibody cloned and used in the chimeric immunoglobulin-like molecule may be derived from any mammalian source. The constant regions may be complement binding or ADCC active. However, preliminary work (see Examples) indicates that the fusion proteins of the invention may mediate HIV or SIV infected cell death by an ADCC or complement-independent mechanism. The con-

stant regions may be derived from any appropriate isotype, including IgG1, IgG3, or IgM.

The joining of various DNA fragments, is performed in accordance with conventional techniques, employing blunt-ended or staggered-ended termini for ligation, restriction enzyme digestion to provide appropriate termini, filling in of cohesive ends as appropriate, alkali and phosphatase treatment to avoid undesirable joining, and ligation with appropriate ligases. The genetic construct may optionally encode a leader sequence to allow efficient expression of the fusion protein. For example, the leader sequence utilized by Maddon et al., *Cell* 42:93–104 (1985) for the expression of CD4 may be used.

For cDNA, the cDNA may be cloned and the resulting clone screened, for example, by use of a complementary probe or by assay for expressed CD4 using an antibody as disclosed by Dalgleish et al., *Nature* 312:763–766 (1984); Klatzmann et al., *Immunol. Today* 7:291–297 (1986); McDougal et al., *J. Immunol.* 135:3151–3162 (1985); and McDougal, J. et al., *J. Immunol,* 137:2937–2944 (1986).

To express the fusion hybrid protein, transcriptional and translational signals recognized by an appropriate host element are necessary. Eukaryotic hosts which may be used include mammalian cells capable of culture in vitro, particularly leukocytes, more particularly myeloma cells or other transformed or oncogenic lymphocytes, e.g., EBV-transformed cells. Alternatively, non-mammalian cells may be employed, such as bacteria, fungi, e.g., yeast, filamentous fungi, or the like.

Preferred hosts for fusion protein production are mammalian cells, grown in vitro in tissue culture or in vivo in animals. Mammalian cells provide post translational modification to immunoglobulin protein molecules which provide for correct folding and glycosylation of appropriate sites. Mammalian cells which may be useful as hosts include cells of fibroblast origins such as VERO or CHO-K1 or cells of lymphoid origin, such as the hybridoma SP2/0-AG14 or the myeloma P3x63Sgh, and their derivatives. For the purpose of preparing an immunoglobulin-like molecule, a plasmid containing a gene which encodes a heavy chain immunoglobulin, wherein the variable region has been replaced with CD4 or fragment thereof which binds to gp120, may be introduced, for example, into J558L myeloma cells, a mouse plasmacytoma expressing the lambda-1 light chain but which does not express a heavy chain (see Oi et al., *P.N.A.S.* (USA) 80:825–829 (1983)). Other preferred hosts include COS cells, BHK cells and hepatoma cells.

The constructs may be Joined together to form a single DNA segment or may be maintained as separate segments, by themselves or in conjunction with vectors.

Where the fusion protein is not glycosylated, any host may be used to express the protein which is compatible with replicon and control sequences in the expression plasmid. In general, vectors containing replicon and control sequences are derived from species compatible with a host cell are used in connection with the host. The vector ordinarily carries a replicon site, as well as specific genes which are capable of providing phenotypic selection in transformed cells. The expression of the fusion protein can also be placed under control with other regulatory sequences which may be homologous to the organism in its untransformed state. For example, lactose-dependent *E. coli* chromosomal DNA comprises a lactose or lac operon which mediates lactose utilization by elaborating the enzyme beta-galactosidase. The lac control elements may be obtained from bacterial phage lambda plac5, which is inlective for *E. coli*. The lac promoter-operator system can be induced by IPTG.

Other promoters/operator systems or portions thereof can be employed as well. For example, colicin E1, galactose, alkaline phosphatase, tryptophan, xylose, tax, and the like can be used.

For mammalian hosts, several possible vector systems are available for expression. One class of vectors utilize DNA elements which are derived from animal viruses such as bovine papilloma virus, polyoma virus, adenovirus, vaccinia virus, baculovirus, retroviruses (RSV, MMTV or MOMLV), or SV40 virus. Cells which have stably integrated the DNA into their chromosomes may be selected by introducing one or more markers which allow selection of transfected host cells. The marker may provide for prototropy to an auxotrophic host, biocide resistance, e.g., antibiotics, or heavy metals such as copper or the like. The selectable marker gene can be either directly linked to the DNA sequences to be expressed, or introduced into the same cell by cotransformation. Additional elements may also be needed for optimal synthesis of mRNA. These elements may include splice signals, as well as transcriptional promoters, enhancers, and termination signals. The cDNA expression vectors incorporating such elements includes those described by Okayama, H., *Mol. Cel. Biol.,* 1:280 (1983) and others.

Once the vector or DNA sequence containing the constructs has been prepared for expression, the DNA constructs may be introduced to an appropriate host. Various techniques may be employed, such as protoplast fusion, calcium phosphate precipitation, electroporation or other conventional techniques. After the fusion, the cells are grown in media and screened for the appropriate activity. Expression of the gene(s) results in production of the fusion protein. This expressed fusion protein may then be subject to further assembly to form the immunoglobulin-like molecule.

The host cells for immunoglobulin production may be immortalized cells, primarily myeloma or lymphoma cells. These cells may be grown in appropriate nutrient medium in culture flasks or injected into a synergistic host, e.g., mouse or a rat, or immunodeficient host or host site, e.g., nude mouse or hamster pouch. In particular, the cells may be introduced into the abdominal cavity of an animal to allow production of ascites fluid which contains the immunoglobulin-like molecule. Alternatively, the cells may be injected subcutaneously and the chimeric antibody is harvested from the blood of the host. The cells may be used in the same manner as hybridoma cells. See Diamond et al., *N. Eng. J. Med.* 304:1344 (1981), and Kennatt, McKearn and Bechtol (Eds.), *Monoclonal Antibodies: Hybridomas:—A New Dimension in Biologic Analysis,* Plenum, 1980.

The fusion proteins and immunoglobulin-like molecules of the invention may be isolated and purified in accordance with conventional conditions, such as extraction, precipitation, chromatography, affinity chromatography, electrophoresis or the like. For example, the IgG1 fusion proteins may be purified by passing a solution through a column which contains immobilized protein A or protein G which selectively binds the Fc portion of the fusion protein. See, for example, Reis, K. J., et al., *J. Immunol.* 132:3098–3102 (1984); PCT Application, Publication No. WO87/00329. The chimeric antibody may the be eluted by treatment with a chaotropic salt or by elution with aqueous acetic acid (1 M).

Alternatively the fusion proteins may be purified on anti-CD4 antibody columns, or on anti-immunoglobulin antibody columns.

In one embodiment of the invention, cDNA sequences which encode CD4, or a fragment thereof which binds

gp120, may be ligated into an expression plasmid which codes for an antibody wherein the variable region of the gene has been deleted. Methods for the preparation of genes which encode the heavy or light chain constant regions of immunoglobulins are taught, for example, by Robinson, R. et al., PCT Application, Publication No. WO87-02671.

Preferred immunoglobulin-like molecules which contain CD4, or fragments thereof, contain the constant region of an IgM, IgG1 or IgG3 antibody which binds complement at the Fc region.

The fusion protein and immunoglobulin-like molecules of the invention may be used for the treatment of HIV viral infections. The fusion protein complexes to gp120 which is expressed on infected cells. Although the inventor is not bound by a particular theory, it appears that the Fc portion of the hybrid fusion protein may bind with complement, which mediates destruction of the cell. In this manner, infected cells are destroyed so that additional viral particle production is stopped.

For the purpose of treating HIV infections, the fusion protein or immunoglobulin-like molecule of the invention may additionally contain a radiolabel or therapeutic agent which enhances destruction of the HIV particle or HIV-infected cell.

Examples of radioisotopes which can be bound to the fusion protein or immunoglobulin-like molecule of the invention for use in HIV therapy are ^{125}I, ^{131}I, ^{90}Y, ^{67}Cu, ^{217}Bi, ^{211}At, ^{212}Pb, ^{47}Sc, and ^{109}Pd. Optionally, a label such as boron can be used which emits α and β particles upon bombardment with neutron radiation.

For in vivo diagnosis radionucleotides may be bound to the fusion protein or immunoglobulin-like molecule of the invention either directly or by using an intermediary functional group. An intermediary group which is often used to bind radioisotopes, which exist as metallic cations, to antibodies is diethylenetriaminepentaacetic acid (DTPA). Typical examples of metallic cations which are bound in this manner are ^{99m}Tc ^{123}I, ^{111}In, ^{131}I, ^{97}Ru, ^{67}Cu, ^{67}Ga, and ^{68}Ga.

Moreover, the fusion protein and immunoglobulin-like molecule of the invention may be tagged with an NMR imaging agent which include paramagnetic atoms. The use of an NMR imaging agent allows the in vivo diagnosis of the presence of and the extent of HIV infection within a patient using NMR techniques. Elements which are particularly useful in this manner are ^{157}Gd, ^{55}Mn, ^{162}Dy, ^{52}Cr, and ^{56}Fe.

Therapeutic agents may include, for example, bacterial toxins such as diphtheria toxin, or ricin. Methods for producing fusion proteins comprising fragment A of diphtheria toxin are taught in U.S. Pat. No. 4,675,382 (1987). Diphtheria toxin contains two polypeptide chains. The B chain binds the toxin to a receptor on a cell surface. The A chain actually enters the cytoplasm and inhibits protein synthesis by inactivating elongation factor 2, the factor that translocates ribosomes along mRNA concomitant with hydrolysis of ETP. See Darnell, J. et al., in *Molecular Cell Biology*, Scientific American Books, Inc., page 662 (1986). Alternatively, a fusion protein comprising ricin, a toxic lectin, may be prepared.

Introduction of the chimeric molecules by gene therapy may also be contemplated, for example, using retroviruses or other means to introduce the genetic material encoding the fusion proteins into suitable target tissues. In this embodiment, the target tissues having the cloned genes of the invention may then produce the fusion protein in vivo.

The dose ranges for the administration of the fusion protein or immunoglobulin-like molecule of the invention are those which are large enough to produce the desired effect whereby the symptoms of HIV or SIV infection are ameliorated. The dosage should not be so large as to cause adverse side effects, such as unwanted cross-reactions, anaphylactic reactions, and the like. Generally, the dosage will vary with the age, condition, sex and extent of disease in the patient, counterindications, if any, immune tolerance and other such variables, to be adjusted by the individual physician. Dosage can vary from 0.01 mg/kg to 50 mg/kg, preferably 0.1 mg/kg to 1.0 mg/kg, of the immunoglobulin-like molecule in one or more administrations daily, for one or several days. The immunoglobulin-like molecule can be administered parenterally by injection or by gradual perfusion over time. They can be administered intravenously, intraperitoneally, intramuscularly, or subcutaneously.

Preparations for parenteral administration include sterile or aqueous or non-aqueous solutions, suspensions, and emulsions. Examples of non-aqueous solvents are propylene glycol, poly-ethylene glycol, vegetable oils such as olive oil, and injectable organic esters such as ethyl oleate. Aqueous carriers include water, alcoholic/aqueous solutions, emulsions or suspensions, including saline and buffered media. Parenteral vehicles include sodium chloride solution, Ringer's dextrose, dextrose and sodium chloride, lactated Ringer's, or fixed oils. Intravenous vehicles include fluid and nutrient replenishers, electrolyte replenishers, such as those based on Ringer's dextrose, and the like. Preservatives and other additives may also be present, such as, for example, antimicrobials, antioxidants, chelating agents, inert gases and the like. See, generally, *Remington's Pharmaceutical Science*, 16th Ed., Mack Eds., 1980.

The invention also relates to a method for preparing a medicament or pharmaceutical composition comprising the components of the invention, the medicament being used for therapy of HIV or SIV infection in animals.

The detection and quantitation of antigenic substances and biological samples frequently utilized immunoassay techniques. These techniques are based upon the formation of the complex between the antigenic substance, e.g., gp120, being assayed and an antibody or antibodies in which one or the other member of the complex may be detectably labeled. In the present invention, the immunoglobulin-like molecule or fusion protein may be labeled with any conventional label.

Thus, the hybrid fusion protein or immunoglobulin-like molecule of the invention can also be used in assay for HIV or SIV viral infection in a biological sample by contacting a sample, derived from an animal suspected of having an HIV or SIV infection, with the fusion protein or immunoglobulin-like molecule of the invention, and detecting whether a complex with gp120, either alone or on the surface of an HIV-infected cell, has formed.

For example, a biological sample may be treated with nitrocellulose, or other solid support which is capable of immobilizing cells, cell particles or soluble protein. The support may then be washed with suitable buffers followed by treatment with the fusion protein which may be detectably labeled. The solid phase support may then be washed with the buffer a second time to remove unbound fusion protein and the label on the fusion protein detected.

In carrying out the assay of the present invention on a sample containing gp120, the process comprises:

a) contacting a sample suspected containing gp120 with a solid support to effect immobilization of gp120, or cell which expresses gp120 on its surface;

b) contacting said solid support with the detectably labeled immunoglobulin-like molecule or fusion protein of the invention;

c) incubating said detectably labeled immunoglobulin-like molecule with said support for a sufficient amount of time to allow the immunoglobulin-like molecule or fusion protein to bind to the immobilized gp120 or cell which expresses gp120 on its surface;

d) separating the solid phase support from the incubation mixture obtained in step c); and

e) detecting the bound immunoglobulin-like molecule or fusion protein and thereby detecting and quantifying gp120.

Alternatively, labeled immunoglobulin-like molecule (or fusion protein) -gp120 complex in a sample may be separated from a reaction mixture by contacting the complex with an immobilized antibody or protein which is specific for an immunoglobulin or, e.g., protein A, protein G, anti-IgM or anti-IgG antibodies. Such anti-immunoglobulin antibodies may be monoclonal or polyclonal. The solid support may then be washed with suitable buffers to give an immobilized gp120-labeled immunoglobulin-like molecule antibody complex. The label on the fusion protein may then be detected to give a measure of endogenous gp120 and, thereby, the presence of HIV.

This aspect of the invention relates to a method for detecting HIV or SIV viral infection in a sample comprising

(a) contacting a sample suspected of containing gp120 with a fusion protein or immunoglobulin-like molecule comprising CD4, or fragment thereof which binds to gp120, and the Fc portion of an immunoglobulin chain,

(b) detecting whether a complex is formed.

The invention also relates to a method of detecting gp120 in a sample, further comprising

(c) contacting the mixture obtained in step (a) with an Fc binding molecule, such as an antibody, protein A, or protein G, which is immobilized on a solid phase support and is specific for the hybrid fusion protein, to give a gp120 fusion protein-immobilized antibody complex

(d) washing the solid phase support obtained in step (c) to remove unbound fusion protein,

(e) and detecting the label on the hybrid fusion protein.

Of course, the specific concentrations of detectably labeled immunoglobulin-like molecule (or fusion protein) and gp120, the temperature and time of incubation, as well as other assay conditions may be varied, depending on various factors including the concentration of gp120 in the sample, the nature of the sample, and the like. Those skilled in the art will be able to determine operative and optimal assay conditions for each determination by employing routine experimentation.

Other such steps as washing, stirring, shaking, filtering and the like may be added to the assays as is customary or necessary for the particular situation.

One of the ways in which the immunoglobulin-like molecule or fusion protein of the present invention can be detectably labeled is by linking the same to an enzyme. This enzyme, in turn, when later exposed to its substrate, will react with the substrate in such a manner as to produce a chemical moiety which can be detected as, for example, by spectrophotometric, fluorometric or by visual means. Enzymes which can be used to detectably label the immunoglobulin-like molecule or fusion protein of the present invention include, but are not limited to, malate dehydrogenase, staphylococcal nuclease, delta-V-steroid isomerase, yeast alcohol dehydrogenase, alpha-glycerophosphate dehydrogenase, triose phosphate isomerase, horseradish peroxidase, alkaline phosphatase, asparaginase, glucose oxidase, beta-galactosidase, ribonuclease, urease, catalase, glucose-VI-phosphate dehydrogenase, glucoamylase and acetylcholine esterase.

The immunoglobulin-like molecule or fusion protein of the present invention may also be labeled with a radioactive isotope which can be determined by such means as the use of a gamma counter or a scintillation counter or by autoradiography. Isotopes which are particularly useful for the purpose of the present invention are: ^{3}H, ^{125}I, ^{131}I, ^{32}P, ^{35}S, ^{14}C, ^{51}Cr, ^{36}Cl, ^{57}Co, ^{58}Co, ^{59}Fe and ^{75}Se.

It is also possible to label the immunoglobulin-like molecule or fusion protein with a fluorescent compound. When the fluorescently labeled immunoglobulin-like molecule is exposed to light of the proper wave length, its presence can then be detected due to the fluorescence of the dye. Among the most commonly used fluorescent labelling compounds are fluorescein isothiocyanate, rhodamine, phycoerytherin, phycocyanin, allophycocyanin, o-phthaldehyde and fluorescamine.

The immunoglobulin-like molecule or fusion protein of the invention can also be detectably labeled using fluorescence emitting metals such as ^{152}Eu, or others of the lanthanide series. These metals can be attached to the immunoglobulin-like molecule or fusion protein using such metal chelating groups as diethylenetriaminepentaacetic acid (DTPA) or ethylenediaminetetraacetic acid (EDTA).

The immunoglobulin-like molecule or fusion protein of the present invention also can be detectably labeled by coupling it to a chemiluminescent compound. The presence of the chemiluminescent-tagged immunoglobulin-like molecule or fusion protein is then determined by detecting the presence of luminescence that arises during the course of a chemical reaction. Examples of particularly useful chemiluminescent labeling compounds are luminol, isoluminol, theromatic acridinium ester, imidazole, acridinium salt and oxalate ester.

Likewise, a bioluminescent compound may be used to label the immunoglobulin-like molecule or fusion protein of the present invention. Bioluminescence is a type of chemiluminescence found in biological systems in which a catalytic protein increases the efficiency of the chemiluminescent reaction. The presence of a bioluminescent protein is determined by detecting the presence of luminescence. Important bioluminescent compounds for purposes of labeling are luciferin, luciferase and aequorin.

Detection of the immunoglobulin-like molecule or fusion protein may be accomplished by a scintillation counter, for example, if the detectable label is a radioactive gamma emitter, or by a fluorometer, for example, if the label is a fluorescent material. In the case of an enzyme label, the detection can be accomplished by colorimetric methods which employ a substrate for the enzyme. Detection may also be accomplished by visual comparison of the extent of enzymatic reaction of a substrate in comparison with similarly prepared standards.

The assay of the present invention is ideally suited for the preparation of a kit. Such a kit may comprise a carrier means being compartmentalized to receive in close confinement therewith one or more container means such as vials, tubes and the like, each of said container means comprising the separate elements of the immunoassay. For example, there may be a container means containing a solid phase support, and further container means containing the detectably labeled immunoglobulin-like molecule or fusion protein in solution. Further container means may contain standard solutions comprising serial dilutions of analytes such as gp120 or fragments thereof to be detected. The standard

solutions of these analytes may be used to prepare a standard curve with the concentration of gp120 plotted on the abscissa and the detection signal on the ordinate. The results obtained from a sample containing gp120 may be interpolated from such a plot to give the concentration of gp120.

The immunoglobulin-like molecule or fusion protein of the present invention can also be used as a stain for tissue sections. For example, a labeled immunoglobulin-like molecule comprising CD4 or fragment thereof which binds to gp120 may be contacted with a tissue section, e.g., a brain biopsy specimen. This section may then be washed and the label detected.

The following examples are illustrative, but not limiting the method and composition of the present invention. Other suitable modifications and adaptations which are obvious to this skill in the art are within the spirit and scope of this invention.

EXAMPLES

Example 1

Preparation of CD4-Ig cDNA Constructs

The extracellular portion of the CD4 molecule (See Madden, P. J., et al., *Cell* 42:93–104 (1985)) was fused at three locations in a human IgG1 heavy chain constant region gene by means of a synthetic splice donor linker molecule. To exploit the splice donor linker, a BamHI linker having the sequence CGCGGATCCGCG was first inserted at amino acid residue 395 of the CD4 precursor sequence (nucleotide residue 1295). A synthetic splice donor sequence

```
GATCCCGAGGGTGAGTACTA

    GGCTCCCACTCATGATTCGA
```

bounded by BamHI and HindIII complementary ends was created and fused to the HindIII site in the intron preceding the CH1 domain, to the EspI site in the intron preceding the hinge domain, and to the BanI site preceding the CH2 domain of the IgG1 genomic sequence. Assembly of the chimeric genes by ligation at the BamHI site afforded molecules in which either the variable (V) region, the V+CH1 regions, or the V, CH1 and hinge regions were replaced by CD4. In the last case, the chimeric molecule is expected to form a monomer structure, while in the former, a dimeric molecule is expected.

One such genetic construct which contains the DNA sequence which encodes CD4 linked to human IgG1 at the Hind3 site upstream of the CH1 region (fusion protein CD4H-1) is depicted in Table 1. The plasmid containing this genetic construct (pCD4Hγ1) has been deposited in *E. coli* (MC1061/P3) at the American Type Culture Collection (ATCC) under the terms of the Budapest Treaty and given accession number 67611.

A second genetic construct which contains the DNA sequence which encodes CD4 linked to human IgG1 at the Esp site upstream of the hinge region (fusion protein CD4Eγ1) is depicted in Table 2. The plasmid containing this genetic construct (pCD4Eγ1) has been deposited in *E. coli* (MC1061/P3) at the ATCC under the terms of the Budapest Treaty and given accession number 67610.

A third genetic construct which contains the DNA sequence which encodes CD4 linked to human IgM at the Mst2 site upstream of the CH1 region (fusion protein CD4Mg) is depicted in Table 3. The plasmid containing this genetic construct (pCD4Mμ) has been deposited in *E. coli* (MC1061/P3) at the ATCC under the terms of the Budapest Treaty and given accession number 67609.

A fourth genetic construct which contains the DNA sequence which encodes CD4 linked to human IgM at the Pst site upstream of the CH2 region (fusion protein CD4Pμ) is depicted in Table 4. The plasmid containing this genetic construct (pCD4Pμ) has been deposited in *E. coli* (MC1061/P3) at the ATCC under the terms of the Budapest Treaty and given accession number 67608.

A fifth genetic construct which contains the DNA sequence which encodes CD4 linked to human IgG1 at the BanI site downstream from the hinge region (fusion protein CD4Bγ1) is depicted in Table 5.

Two similar constructs were prepared from the human IgM heavy chain constant region by fusion with the introns upstream of the μ CH1 and CH2 domains at an MstII site and a PstI site respectively. The fusions were made by joining the PstI site of the CD4/IgG1 construct fused at the Esp site in IgG1 gene to the MstII and Pst sites in the IgM gene. In the first instance, this was performed by treatment of the Pst end with T4 DNA Polymerase and the MstII end with *E. coli* DNA Polymerase, followed by ligation; and in the second instance, by ligation alone.

Immunoprecipitation of the fusion proteins with a panel of monoclonal antibodies directed against CD4 epitopes showed that all of the epitopes were preserved. A specific high affinity association is demonstrated between the chimeric molecules and HIV envelope proteins expressed on the surface of cells transfected with an attenuated (reverse transcriptase deleted) proviral construct.

TABLE 1

```
                F N                                  S          B
                N S        B     M     H           DHA          S
                U P        B     N     G           RAU          T
                4 B        V     L     A           AE9          X
                H 2        1     1     1           236          1
                                                     /
   GCCTGTTTGAGAAGCAGCGGGCAAGAAAGACGCAAGCCCAGAGGCCCTGCCATTTCTGTG
1  ---------+---------+---------+---------+---------+---------+ 60
   CGGACAAACTCTTCGTCGCCCGTTCTTTCTGCGTTCGGGTCTCCGGGACGGTAAAGACAC

     B    PS              S                                    S
   DBS ADNPA        D   DHNA              M    HM           HNC
   DAP VRLUU        D   RALU              N    AN           PRC
   EN1 AAAM9        E   AEA9              L    EL           AIF
   122 22416        1   2346              1    31           211
    /  / //           /                     /            /
   GGCTCAGGTCCCTACTGGCTCAGGCCCCTGCCTCCCTCGGCAAGGCCACAATGAACCGGG
61 ---------+---------+---------+---------+---------+---------+ 120
   CCGAGTCCAGGGATGACCGAGTCCGGGGACGGAGGGAGCCGTTCCGGTGTTACTTGGCCC
```

TABLE 1-continued

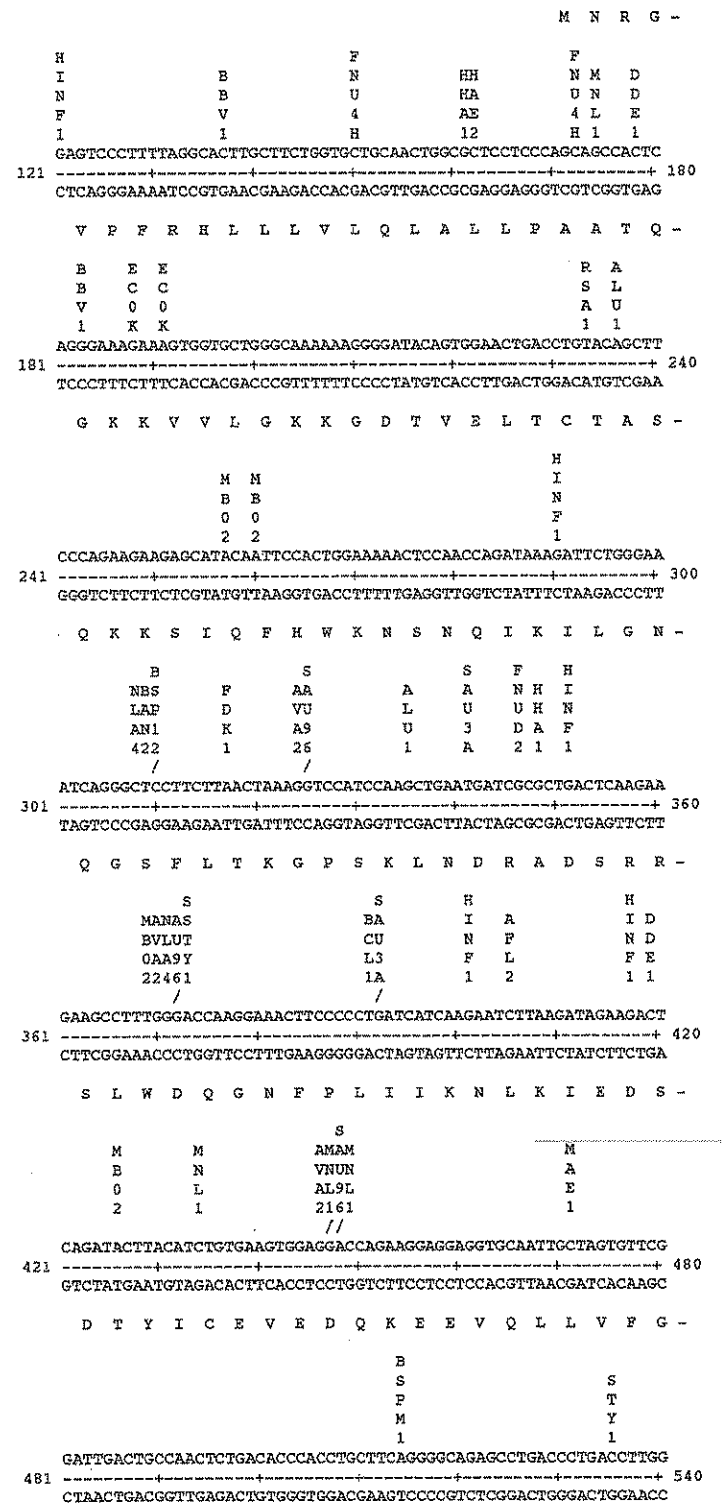


TABLE 1-continued

```
      L  T  A  N  S  D  T  H  L  L  Q  G  Q  S  L  T  L  T  L  E -

          B  BS                          H
         BS  SC          D        M     I   S
         AP  TR          D        N     N   T
         N1  NF          E        L     F   Y
         22  11          1        1     1   1
          /   /
     AGAGCCCCCCTGGTAGTAGCCCCTCAGTGCAATGTAGGAGTCCAAGGGGTAAAAACATAC
541  ---------+---------+---------+---------+---------+---------+ 600
     TCTCGGGGGGACCATCATCGGGGAGTCACGTTACATCCTCAGGTTCCCCATTTTTGTATG

      S  P  P  G  S  S  P  S  V  Q  C  R  S  P  R  G  K  N  I  Q -

                                N       BBH S  B       BS
                       M   MD   ASP   A BSSGSC  S  B N  SC
                       B   ND   LPV   L APTIAR  T  A L  TR
                       0   LE   UBU   U N1NACF  X  N A  NF
                       2   11   122   1 221111  1  1 4  11
                                //      / ///          /
     AGGGGGGGAAGACCCTCTCCGTGTCTCAGCTGGAGCTCCAGGATAGTGGCACCTGGACAT
601  ---------+---------+---------+---------+---------+---------+ 660
     TCCCCCCCTTCTGGGAGAGGCACAGAGTCGACCTCGAGGTCCTATCACCGTGGACCTGTA

      G  G  K  T  L  S  V  S  Q  L  E  L  Q  D  S  G  T  W  T  C -

      N
     NS                            M                   NM  A
     LP                            B                   HA  L
     AH                            0                   EE  U
     31                            2                   11  1
      /
     GCACTGTCTTGCAGAACCAGAAGAAGGTGGAGTTCAAAATAGACATCGTGGTGCTAGCTT
661  ---------+---------+---------+---------+---------+---------+ 720
     CGTGACAGAACGTCTTGGTCTTCTTCCACCTCAAGTTTTATCTGTAGCACCACGATCGAA

      T  V  L  Q  N  Q  K  K  V  E  F  K  I  D  I  V  V  L  A  F -

            HS         M  M
            AT         N  N
            EU         L  L
            31         1  1
             /
     TCCAGAAGGCCTCCAGCATAGTCTATAAGAAAGAGGGGGAACAGGTGGAGTTCTCCTTCC
721  ---------+---------+---------+---------+---------+---------+ 780
     AGGTCTTCCGGAGGTCGTATCAGATATTCTTTCTCCCCCTTGTCCACCTCAAGAGGAAGG

      Q  K  A  S  S  I  V  Y  K  K  E  G  E  Q  V  E  F  S  F  P -

                         A                 A       M
                         L                 L       N
                         U                 U       L
                         1                 1       1
     CACTCGCCTTTACAGTTGAAAAGCTGACGGGCAGTGGCGAGCTGTGGTGGCAGGCGGAGA
781  ---------+---------+---------+---------+---------+---------+ 840
     GTGAGCGGAAATGTCAACTTTTCGACTGCCCGTCACCGCTCGACACCACCGTCCGCCTCT

      L  A  F  T  V  E  K  L  T  G  S  G  E  L  W  W  Q  A  E  R -

                     P  S
                H   M FM A                          M
                P   N LN U                          B
                H   L ML 3                          0
                1   1 11 A                          2
     GGGCTTCCTCCTCCAAGTCTTGGATCACCTTTGACCTGAAGAACAAGGAAGTGTCTGTAA
841  ---------+---------+---------+---------+---------+---------+ 900
     CCCGAAGGAGGAGGTTCAGAACCTAGTGGAAACTGGACTTCTTGTTCCTTCACAGACATT

      A  S  S  S  K  S  W  I  T  F  D  L  K  N  K  E  V  S  V  K -

        B      BS   PS
        SM     SCADNPAD   A                 A H
        TA     TRVRLUUD   L                 L P
        EE     NFAAAM9E   U                 U H
        23     11224161   1                 1 1
         /      / / //
     AACGGGTTACCCAGGACCCTAAGCTCCAGATGGGCAAGAAGCTCCCGCTCCACCTCACCC
901  ---------+---------+---------+---------+---------+---------+ 960
```

TABLE 1-continued

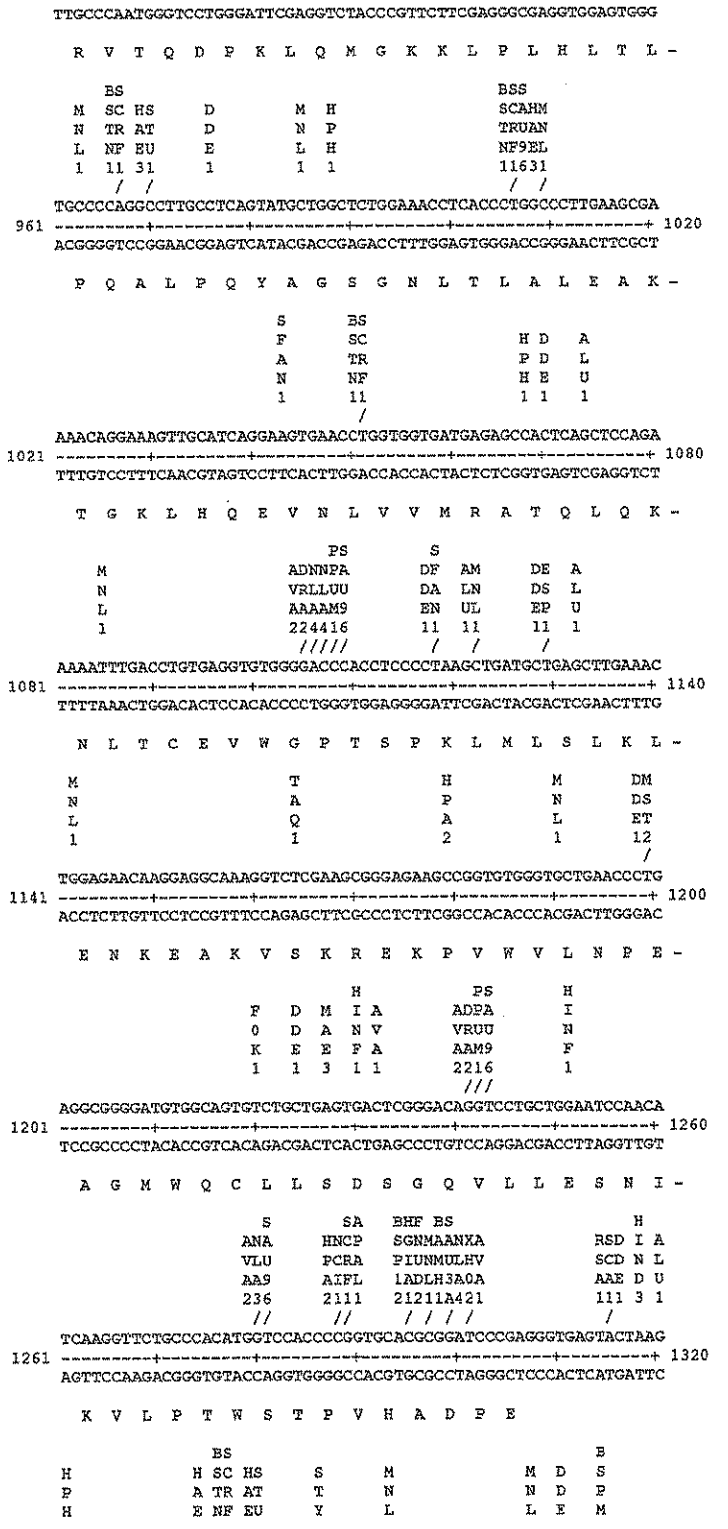


TABLE 1-continued

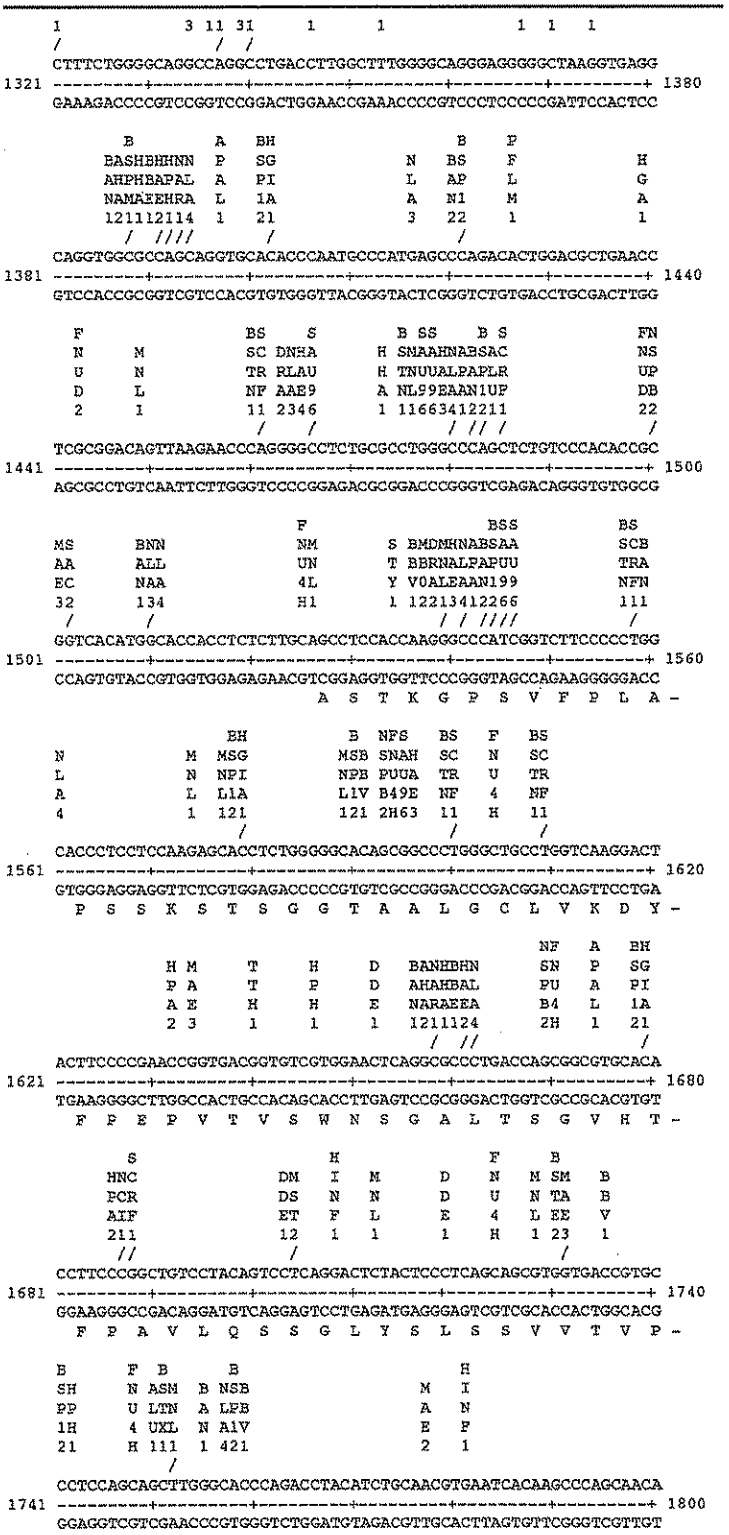


TABLE 1-continued

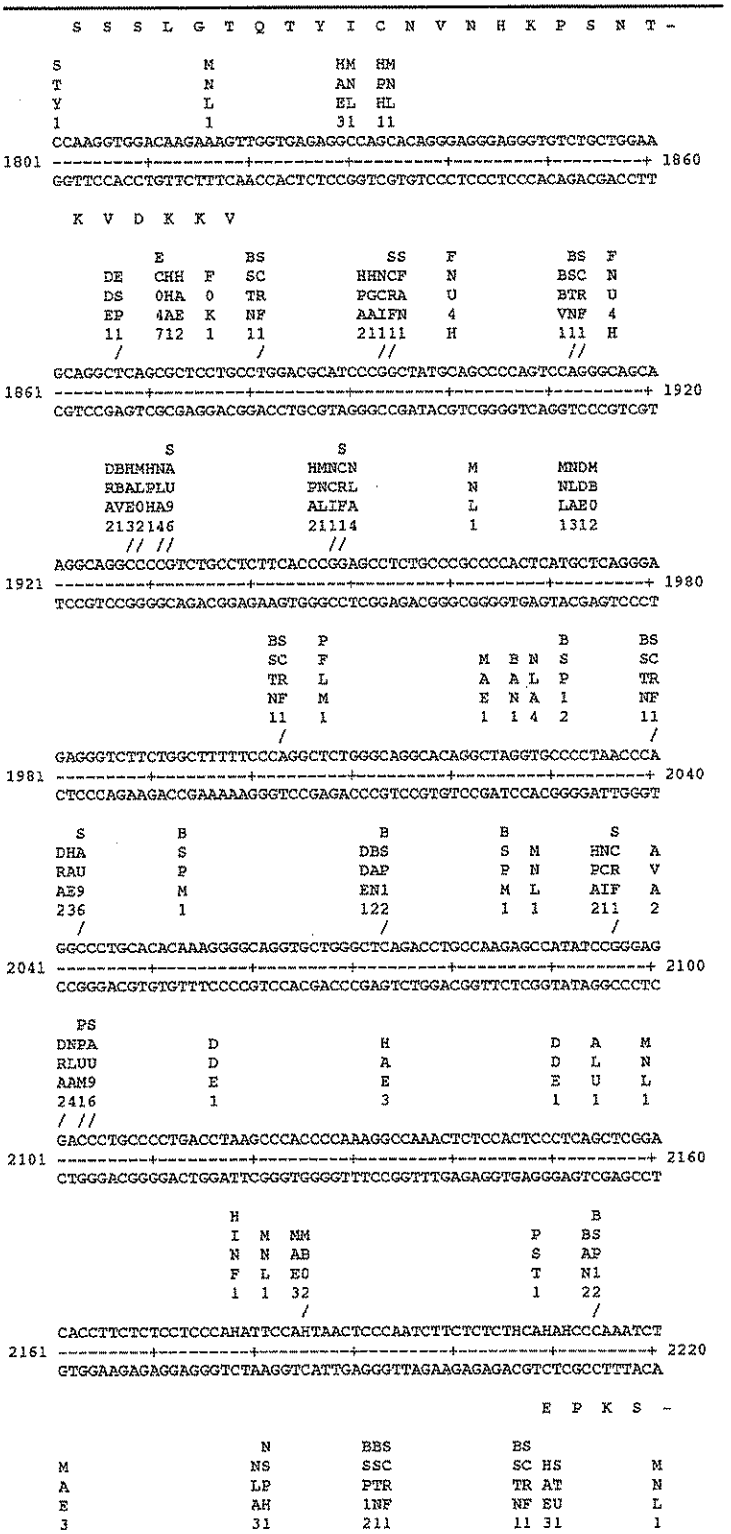


TABLE 1-continued

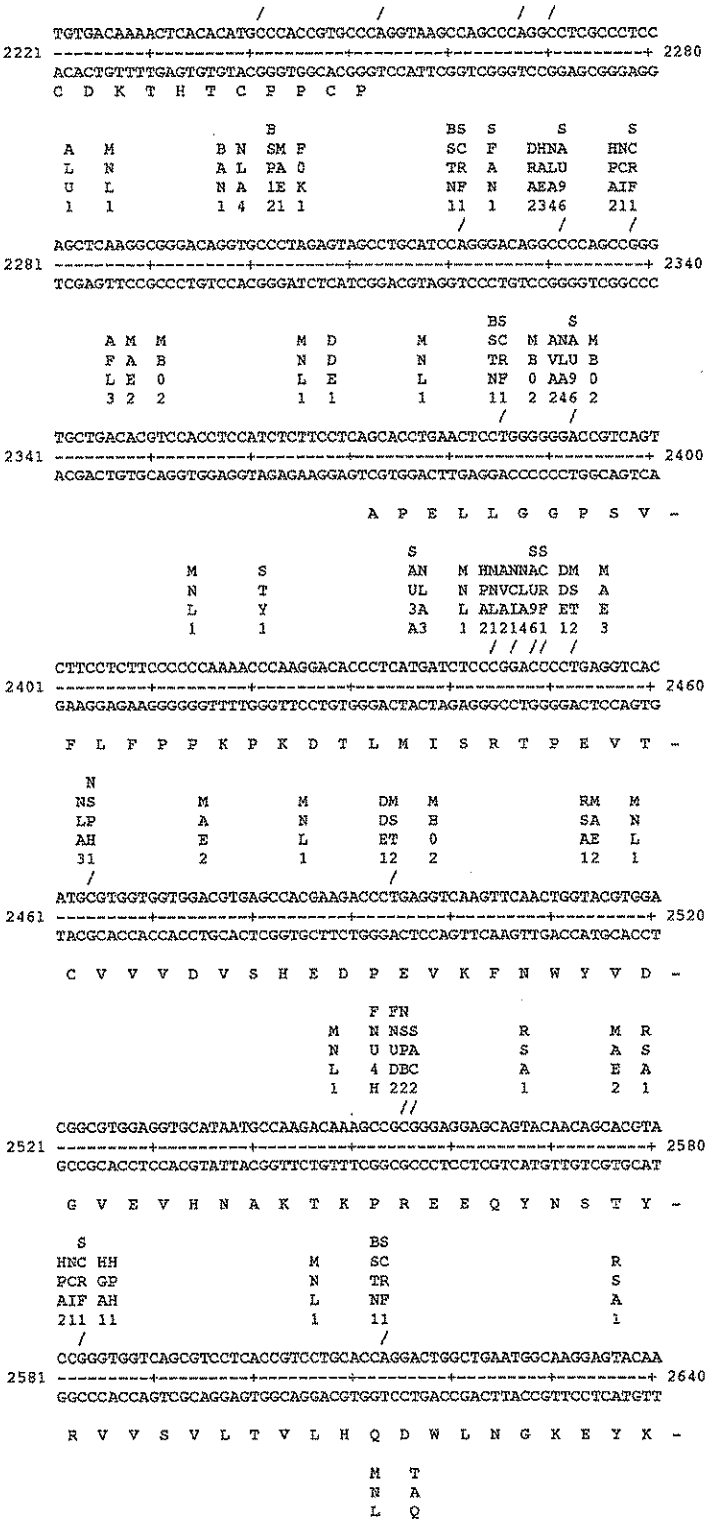


TABLE 1-continued

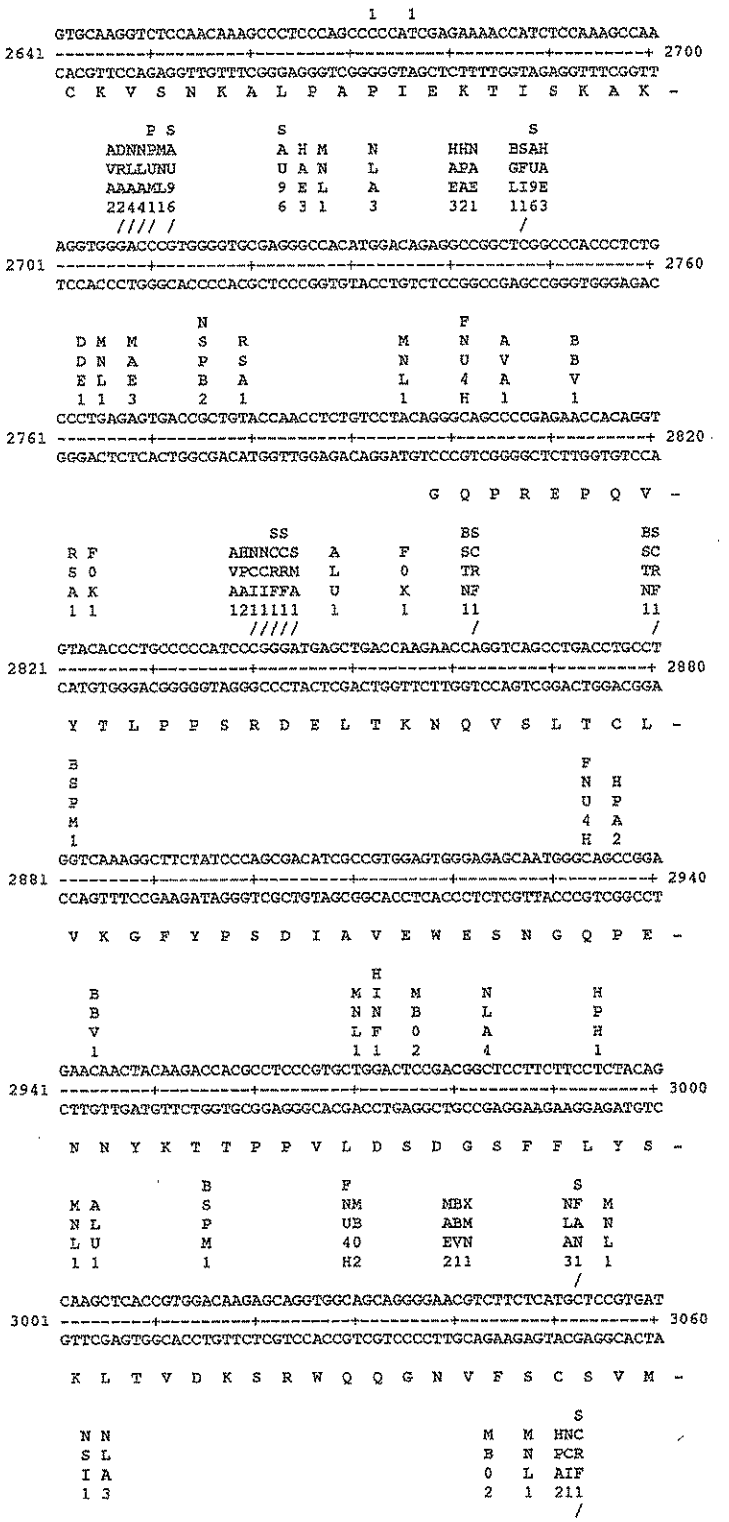


TABLE 1-continued

```
     GCATGAGGCTCTGCACAACCACTACACGCAGAAGAGCCTCTCCCTGTCTCCGGGTAAATG
3061 ---------+---------+---------+---------+---------+---------+ 3120
     CGTACTCCGAGACGTGTTGGTGATGTGCGTCTTCTCGGAGAGGGACAGAGGCCCATTTAC

      H  E  A  L  H  N  H  Y  T  Q  K  S  L  S  L  S  P  G  K  *

            CXHHN
            FMAPA
            RAEAE
            13321
             /
     AGTGCGACGGCCG
3121 ---------+--- 3133
     TCACGCTGCCGGC
```

TABLE 2

```
                    F N                                   S           B
                    N S          B     M      H         DHA          S
                    U P          B     N      G         RAU          T
                    4 B          V     L      A         AE9          X
                    H 2          1     1      1         236          1
                                                          /
     GCCTGTTTGAGAAGCAGCGGGCAAGAAAGACGCAAGCCCAGAGGCCCTGCCATTTCTGTG
   1 ---------+---------+---------+---------+---------+---------+ 60
     CGGACAAACTCTTCGTCGCCCGTTCTTTCTGCGTTCGGGTCTCCGGGACGGTAAAGACAC

        B    PS              S                                    S
      DBS ADNPA        D    DHNA              M    HM           HNC
      DAP VRLUU        D    RALU              N    AN           PCR
      EN1 AAAM9        E    AEA9              L    EL           AIF
      122 22416        1    2346              1    31           211
       /  / //                /                   /             /
     GGCTCAGGTCCCTACTGGCTCAGGCCCCTGCCTCCCTCGGCAAGGCCACAATGAACCGGG
  61 ---------+---------+---------+---------+---------+---------+ 120
     CCGAGTCCAGGGATGACCGAGTCCGGGGACGGAGGGAGCCGTTCCGGTGTTACTTGGCCC

                                                        M  N  R  G -

     H                              F                    F
     I                B             N          HH        N M   D
     N                B             U          HA        U N   D
     F                V             4          AE        4 L   E
     1                1             H          12        H 1   1
     GAGTCCCTTTTAGGCACTTGCTTCTGGTGCTGCAACTGGCGCTCCTCCCAGCAGCCACTC
121 ---------+---------+---------+---------+---------+---------+ 180
     CTCAGGGAAAATCCGTGAACGAAGACCACGACGTTGACCGCGAGGAGGGTCGTCGGTGAG

       V  P  F  R  H  L  L  L  V  L  Q  L  A  L  L  P  A  A  T  Q -

       B     E  E                                       R   A
       B     C  C                                       S   L
       V     D  D                                       A   U
       1     K  K                                       1   1
     AGGGAAAGAAAGTGGTGCTGGGCAAAAAAGGGGATACAGTGGAACTGACCTGTACAGCTT
181 ---------+---------+---------+---------+---------+---------+ 240
     TCCCTTTCTTTCACCACGACCCGTTTTTTCCCCTATGTCACCTTGACTGGACATGTCGAA

       G  K  K  V  V  L  G  K  K  G  D  T  V  E  L  T  C  T  A  S -

                                                        H
                     M  M                               I
                     B  B                               N
                     0  0                               F
                     2  2                               1
     CCCAGAAGAAGAGCATACAATTCCACTGGAAAAACTCCAACCAGATAAAGATTCTGGGAA
241 ---------+---------+---------+---------+---------+---------+ 300
     GGGTCTTCTTCTCGTATGTTAAGGTGACCTTTTTGAGGTTGGTCTATTTCTAAGACCCTT

       Q  K  K  S  I  Q  F  H  W  K  N  S  N  Q  I  K  I  L  G  N -

             B              S                 S    F    H
           NBS       F      AA          A     A    N H  I
           LAP       0      VU          L     U    U H  N
           AN1       K      A9          U     3    D A  F
           422       1      26          1     A    2 1  1
```

TABLE 2-continued

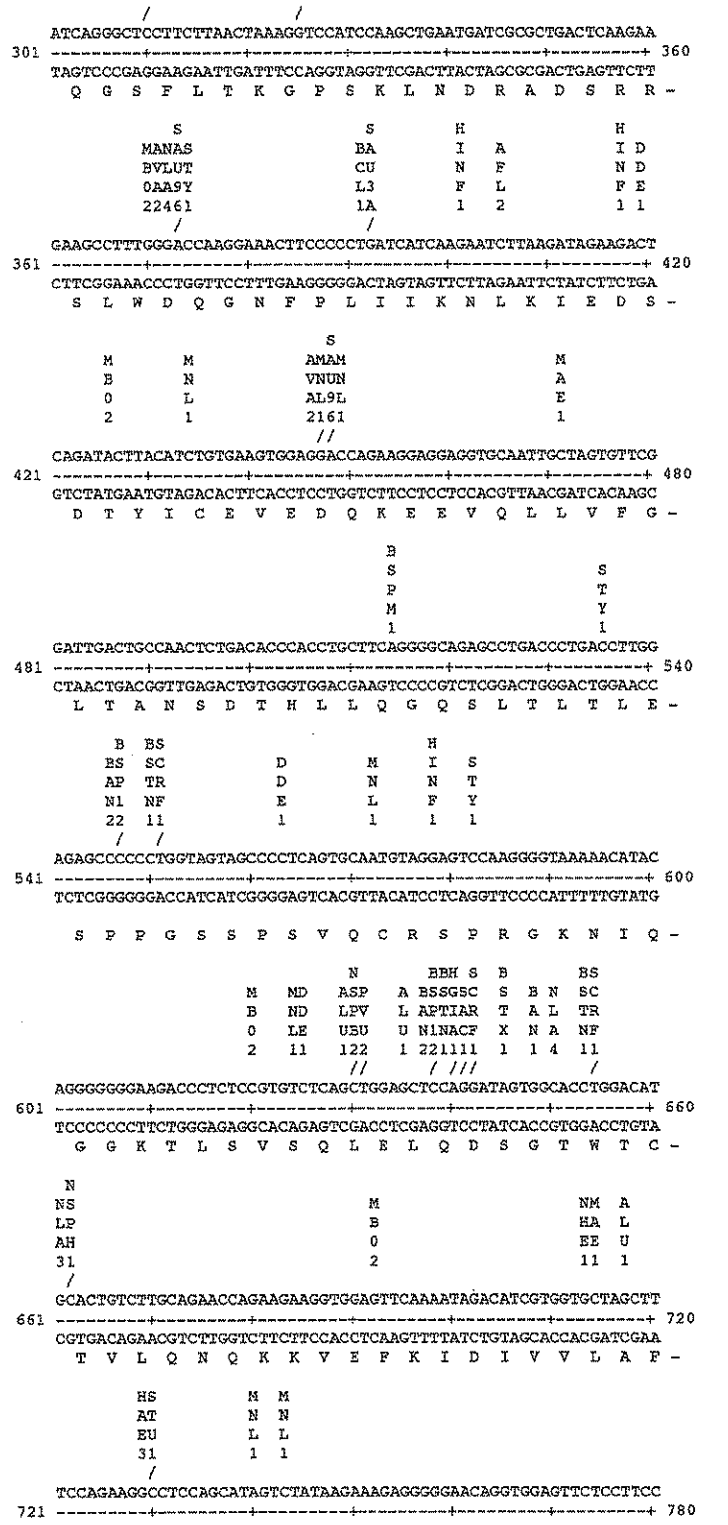


TABLE 2-continued

```
    AGGTCTTCCGGAGGTCGTATCAGATATTCTTTCTCCCCCTTGTCCACCTCAAGAGGAAGG

      Q  K  A  S  S  I  V  Y  K  K  E  G  E  Q  V  E  F  S  F  P -

                           A                 A       M
                           L                 L       N
                           U                 U       L
                           1                 1       1
    CACTCGCCTTTACAGTTGAAAAGCTGACGGGCAGTGGCGAGCTGTGGTGGCAGGCGGAGA
781 ---------+---------+---------+---------+---------+---------+ 840
    GTGAGCGGAAATGTCAACTTTTCGACTGCCCGTCACCGCTCGACACCACCGTCCGCCTCT

      L  A  F  T  V  E  K  L  T  G  S  G  E  L  W  W  Q  A  E  R -

                     P  S
                H    M FM A                           M
                P    N LN U                           B
                H    L ML 3                           0
                1    1 11 A                           2
    GGGCTTCCTCCTCCAAGTCTTGGATCACCTTTGACCTGAAGAACAAGGAAGTGTCTGTAA
841 ---------+---------+---------+---------+---------+---------+ 900
    CCCGAAGGAGGAGGTTCAGAACCTAGTGGAAACTGGACTTCTTGTTCCTTCACAGACATT

      A  S  S  S  K  S  W  I  T  F  D  L  K  N  K  E  V  S  V  K -

        B       BS   PS
        SM      SCADNPAD   A                 A H
        TA      TRVRLUUD   L                 L P
        EE      NFAAAM9E   U                 U H
        23      11224161   1                 1 1
         /       / / //
    AACGGGTTACCCAGGACCCTAAGCTCCAGATGGGCAAGAAGCTCCCGCTCCACCTCACCC
901 ---------+---------+---------+---------+---------+---------+ 960
    TTGCCCAATGGGTCCTGGGATTCGAGGTCTACCCGTTCTTCGAGGGCGAGGTGGAGTGGG
      R  V  T  Q  D  P  K  L  Q  M  G  K  K  L  P  L  H  L  T  L -

         BS                                      BSS
      M  SC HS      D        M  H                SCAHM
      N  TR AT      D        N  P                TRUAN
      L  NF EU      E        L  H                NF9EL
      1  11 31      1        1  1                11631
          / /                                     / /
    TGCCCCAGGCCTTGCCTCAGTATGCTGGCTCTGGAAACCTCACCCTGGCCCTTGAAGCGA
961 ---------+---------+---------+---------+---------+---------+ 1020
    ACGGGGTCCGGAACGGAGTCATACGACCGAGACCTTTGGAGTGGGACCGGGAACTTCGCT
      P  Q  A  L  P  Q  Y  A  G  S  G  N  L  T  L  A  L  E  A  K -

                           S       BS
                           F       SC                H D   A
                           A       TR                P D   L
                           N       NF                H E   U
                           1       11                1 1   1
                                    /
    AAACAGGAAAGTTGCATCAGGAAGTGAACCTGGTGGTGATGAGAGCCACTCAGCTCCAGA
1021 ---------+---------+---------+---------+---------+---------+ 1080
    TTTGTCCTTTCAACGTAGTCCTTCACTTGGACCACCACTACTCTCGGTGAGTCGAGGTCT
      T  G  K  L  H  Q  E  V  N  L  V  V  M  R  A  T  Q  L  Q  K -

                                 PS        S
        M                       ADNNPA       DF  AM      DE  A
        N                       VRLLUU       DA  LN      DS  L
        L                       AAAAM9       EN  UL      EP  U
        1                       224416       11  11      11  1
                                 /////        /   /       /
    AAAATTTGACCTGTGAGGTGTGGGGACCCACCTCCCCTAAGCTGATGCTGAGCTTGAAAC
1081 ---------+---------+---------+---------+---------+---------+ 1140
    TTTTAAACTGGACACTCCACACCCCTGGGTGGAGGGGATTCGACTACGACTCGAACTTTG
      N  L  T  C  E  V  W  G  P  T  S  P  K  L  M  L  S  L  K  L -

       M                       T              H           M       DM
       N                       A              P           N       DS
       L                       Q              A           L       ET
       1                       1              2           1       12
                                                                   /
    TGGAGAACAAGGAGGCAAAGGTCTCGAAGCGGGAGAAGCCGGTGTGGGTGCTGAACCCTG
1141 ---------+---------+---------+---------+---------+---------+ 1200
    ACCTCTTGTTCCTCCGTTTCCAGAGCTTCGCCCTCTTCGGCCACACCCACGACTTGGGAC

      E  N  K  E  A  K  V  S  K  R  E  K  P  V  W  V  L  N  P  E -
```

TABLE 2-continued

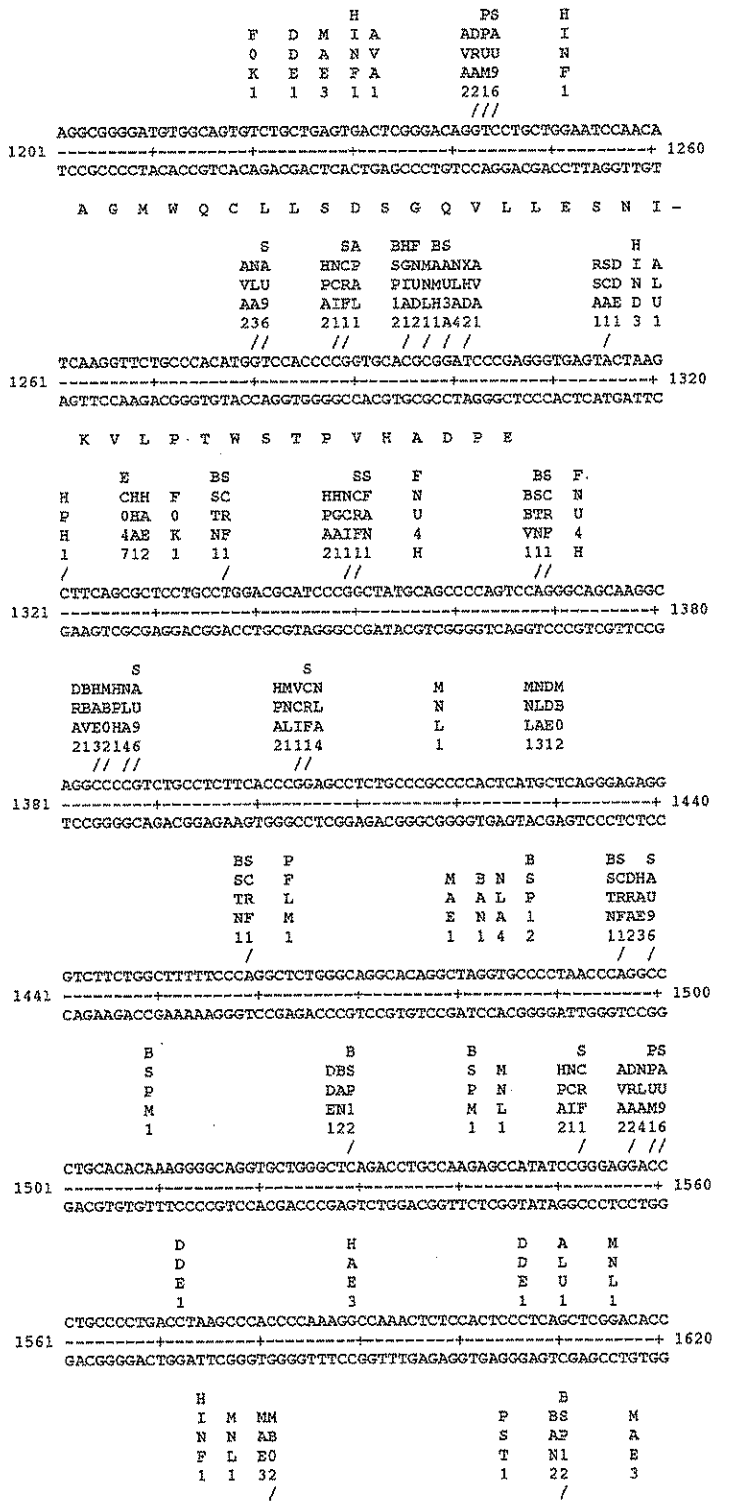


TABLE 2-continued

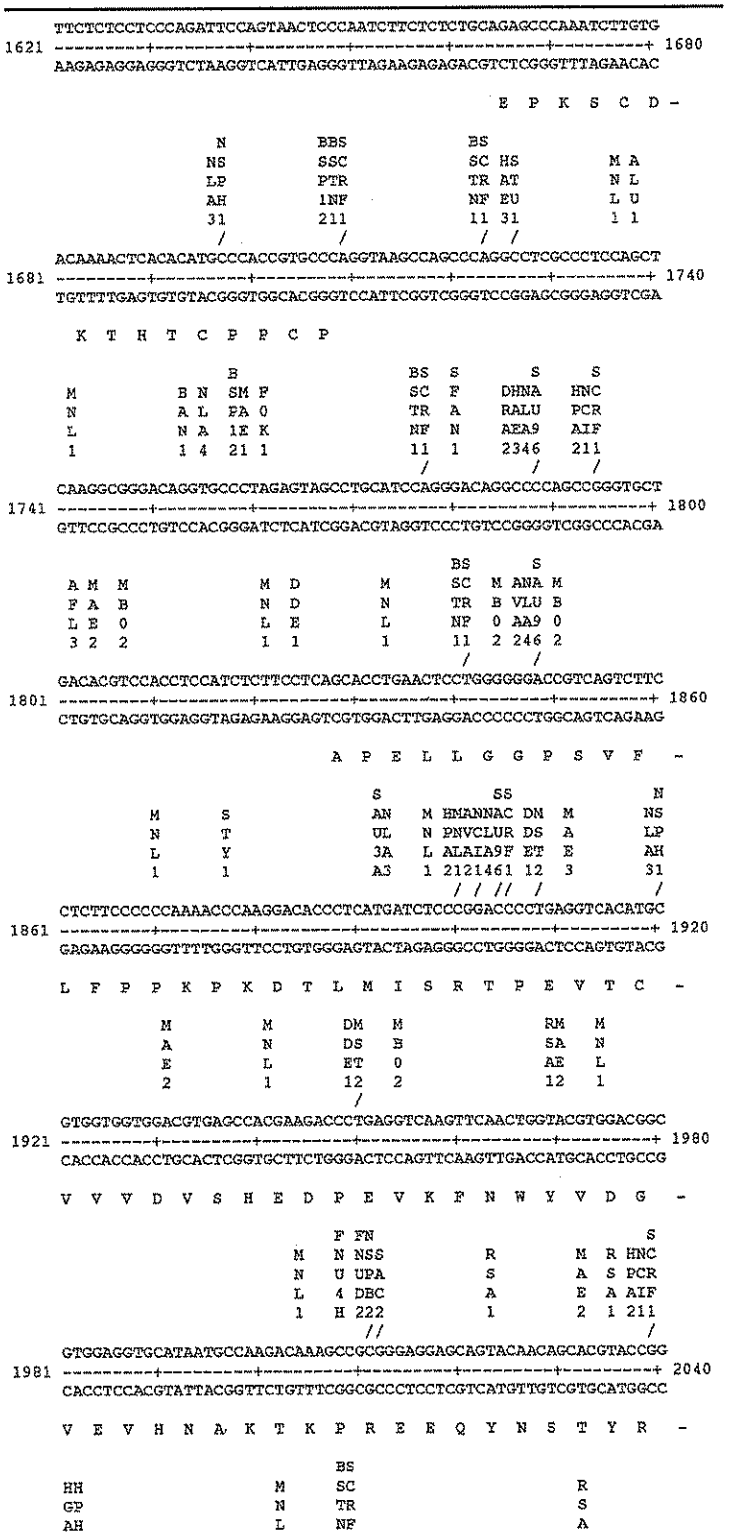


TABLE 2-continued

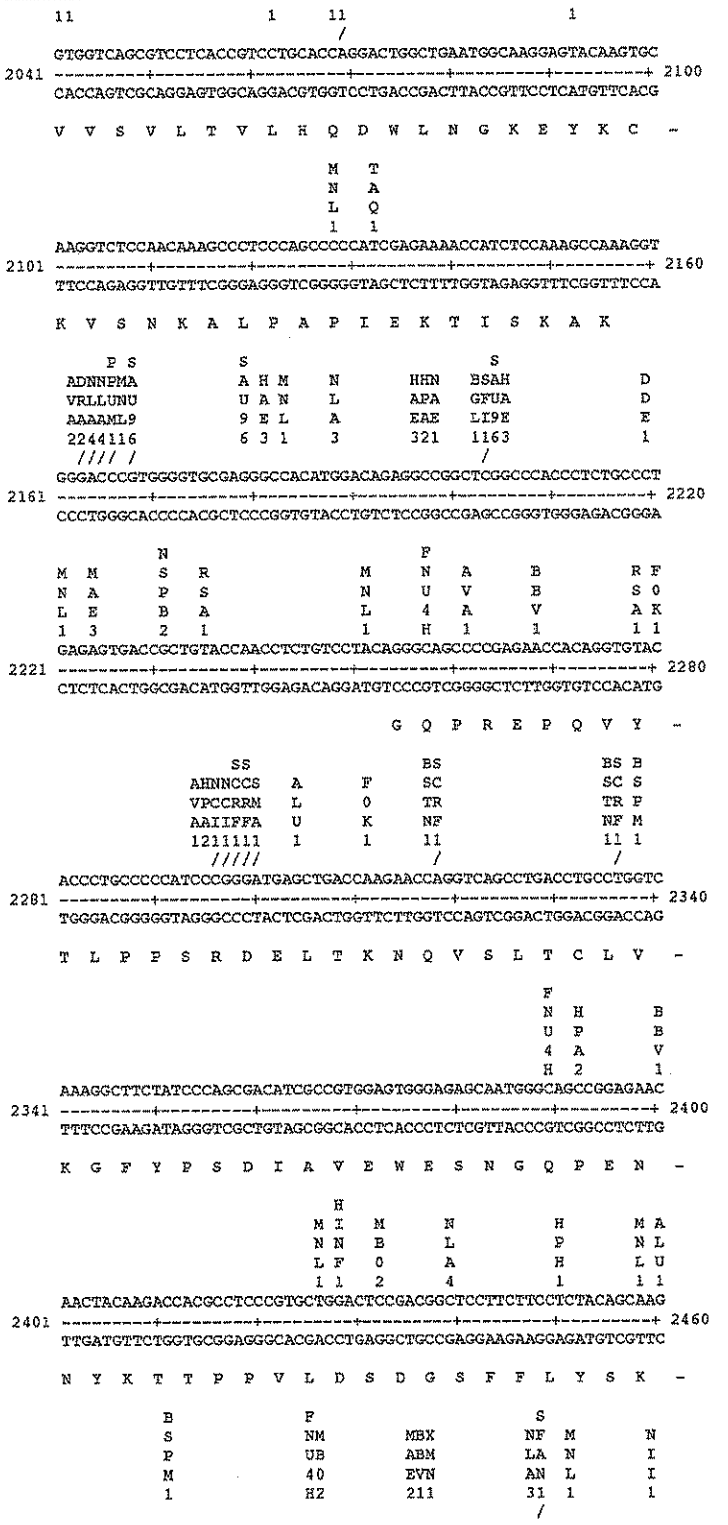


TABLE 2-continued

```
     CTCACCGTGGACAAGAGCAGGTGGCAGCAGGGGAACGTCTTCTCATGCTCCGTGATGCAT
2461 ---------+---------+---------+---------+---------+---------+ 2520
     GAGTGGCACCTGTTCTCGTCCACCGTCGTCCCCTTGCAGAAGAGTACGAGGCACTACGTA

     L  T  V  D  K  S  R  W  Q  Q  G  N  V  F  S  C  S  V  M  H  -

                                                   S
     N                                    M   M  HNC
     L                                    B   N  PCR
     A                                    0   L  AIF
     3                                    2   1  211
                                                   /
     GAGGCTCTGCACAACCACTACACGCAGAAGAGCCTCTCCCTGTCTCCGGGTAAATGAGTG
2521 ---------+---------+---------+---------+---------+---------+ 2580
     CTCCGAGACGTGTTGGTGATGTGCGTCTTCTCGGAGAGGGACAGAGGCCCATTTACTCAC

     E  A  L  H  N  H  Y  T  Q  K  S  L  S  L  S  P  G  K  *

        CXH
        FMA
        RAE
        133
         /
     CGACGGCCG
2581 --------- 2589
     GCTGCCGGC
```

TABLE 3

```
                  F N                          S          B
                  N S        B    M      H    DHA         S
                  U P        B    N      G    RAU         T
                  4 B        V    L      A    AE9         X
                  H 2        1    1      1    236         1
                                                /
     GCCTGTTTGAGAAGCAGCGGGCAAGAAAGACGCAAGCCCAGAGGCCCTGCCATTTCTGTG
   1 ---------+---------+---------+---------+---------+---------+ 60
     CGGACAAACTCTTCGTCGCCCGTTCTTTCTGCGTTCGGGTCTCCGGGACGGTAAAGACAC

        B    PS              S                              S
      DBS ADNPA       D   DHNA              M   HM         HNC
      DAP VRLUU       D   RALU              N   AN         PCR
      EN1 AAAM9       E   AEA9              L   EL         AIF
      122 22416       1   2346              1   31         211
        /  / //              /                  /           /
     GGCTCAGGTCCCTACTGGCTCAGGCCCCTGCCTCCCTCGGCAAGGCCACAATGAACCGGG
  61 ---------+---------+---------+---------+---------+---------+ 120
     CCGAGTCCAGGGATGACCGAGTCCGGGGACGGAGGGAGCCGTTCCGGTGTTACTTGGCCC

                                                      M  N  R  G -

     H                            F                    F
     I              B             N          HH        N M   D
     N              B             U          HA        U N   D
     F              V             4          AE        4 L   E
     1              1             H          12        H 1   1
     GAGTCCCTTTTAGGCACTTGCTTCTGGTGCTGCAACTGGCGCTCCTCCCAGCAGCCACTC
 121 ---------+---------+---------+---------+---------+---------+ 180
     CTCAGGGAAAATCCGTGAACGAAGACCACGACGTTGACCGCGAGGAGGGTCGTCGGTGAG

       V  P  F  R  H  L  L  L  V  L  Q  L  A  L  L  P  A  A  T  Q -

       B    E  E                                        R   A
       B    C  C                                        S   L
       V    0  0                                        A   U
       1    K  K                                        1   1
     AGGGAAAGAAAGTGGTGCTGGGCAAAAAAGGGGATACAGTGGAACTGACCTGTACAGCTT
 181 ---------+---------+---------+---------+---------+---------+ 240
     TCCCTTTCTTTCACCACGACCCGTTTTTTCCCCTATGTCACCTTGACTGGACATGTCGAA

       G  K  K  V  V  L  G  K  K  G  D  T  V  E  L  T  C  T  A  S -

                                                      H
                     M  M                             I
                     B  B                             N
                     0  0                             F
```

TABLE 3-continued

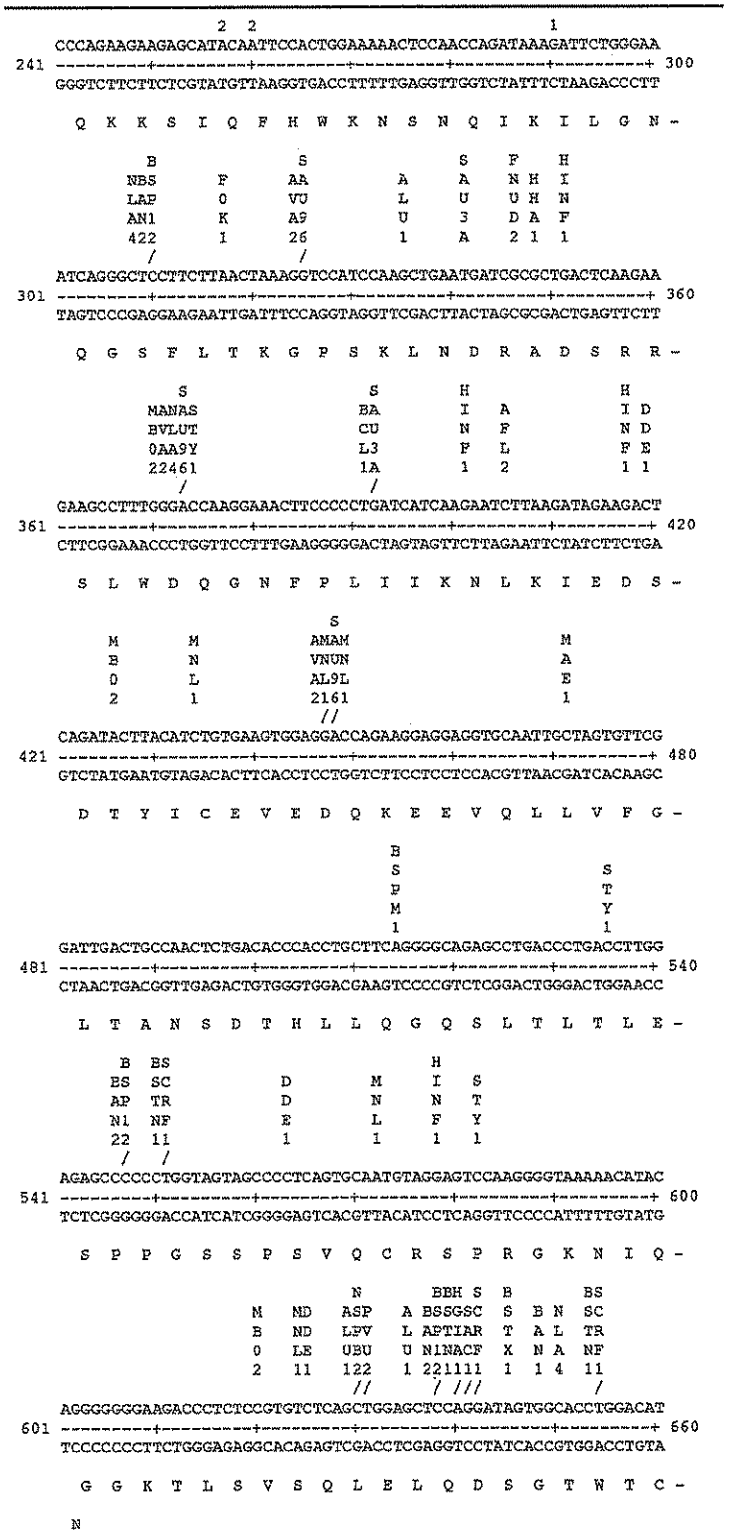


TABLE 3-continued

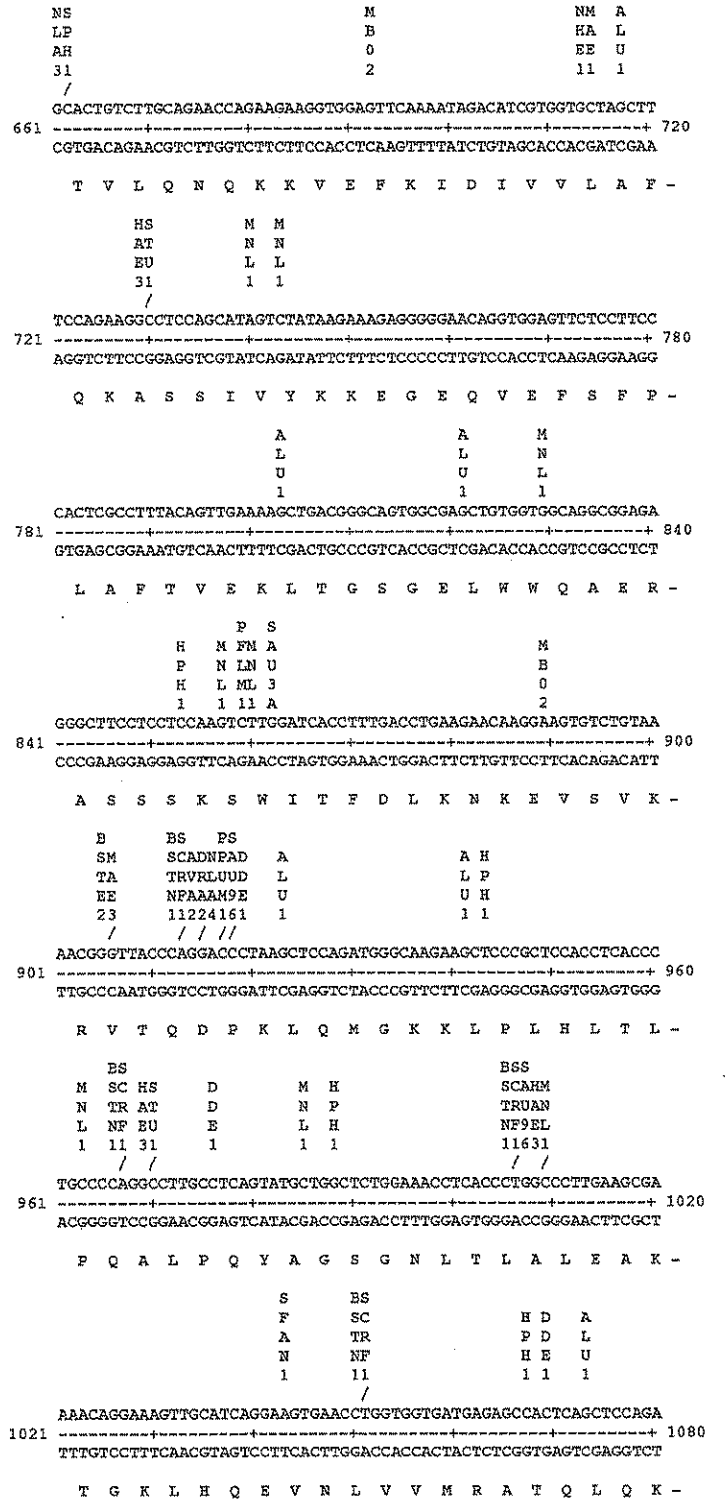


TABLE 3-continued

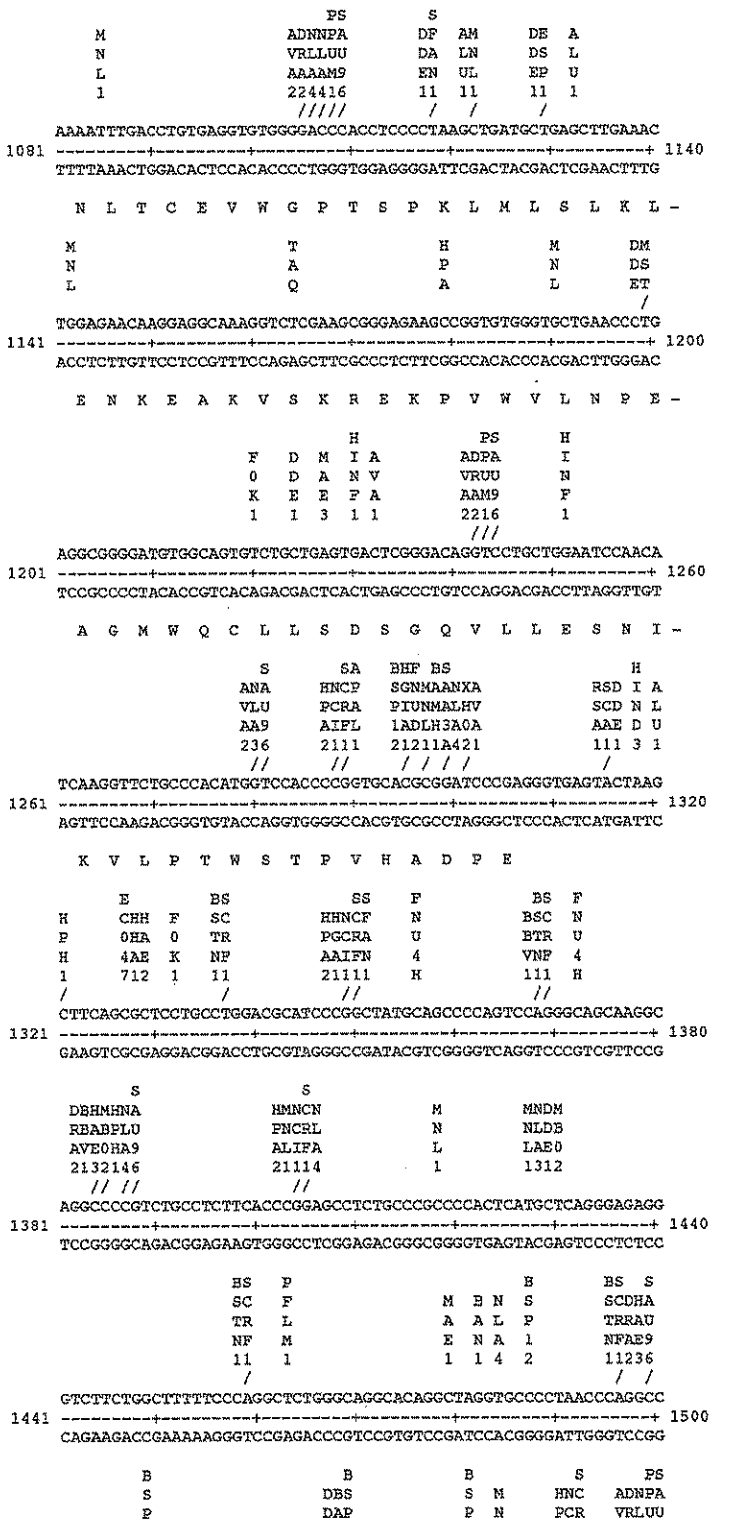


TABLE 3-continued

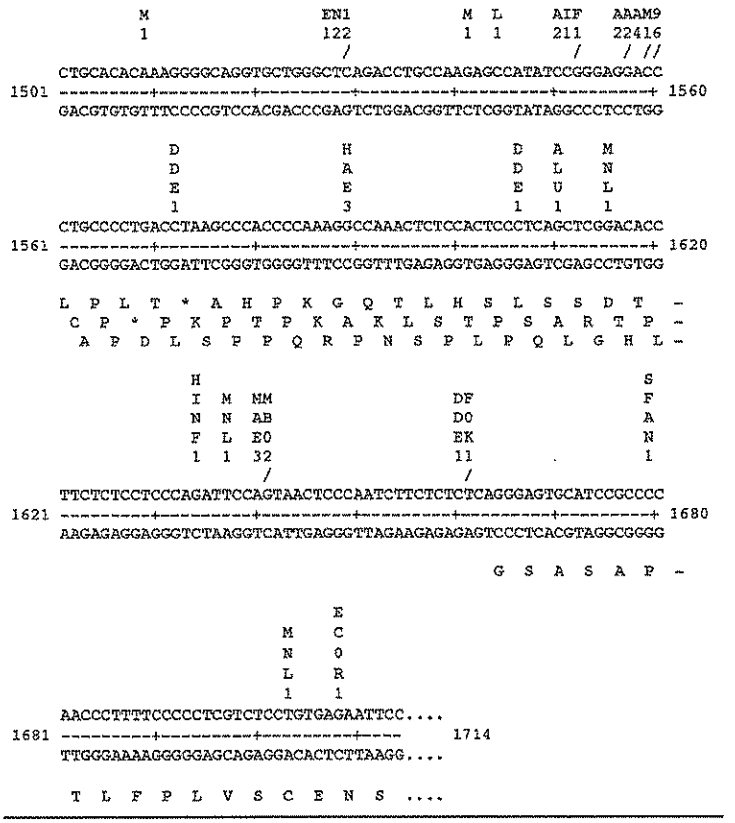


TABLE 4

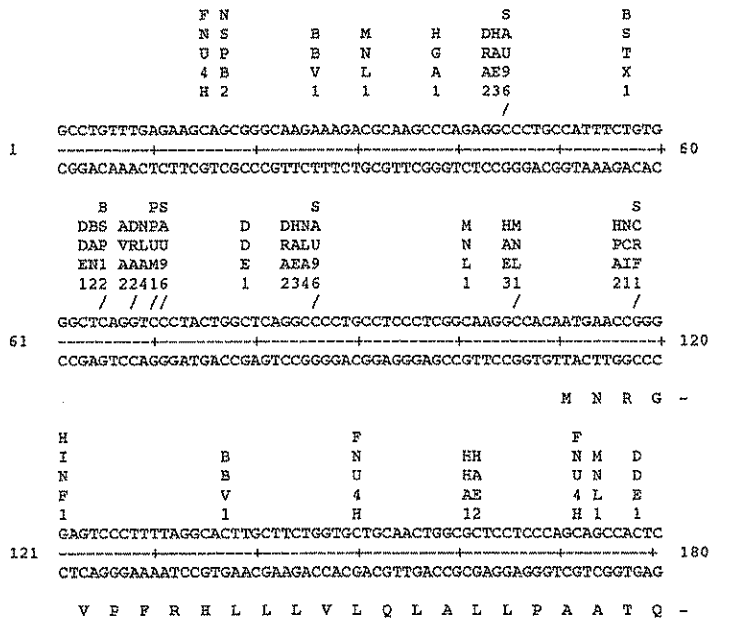


TABLE 4-continued

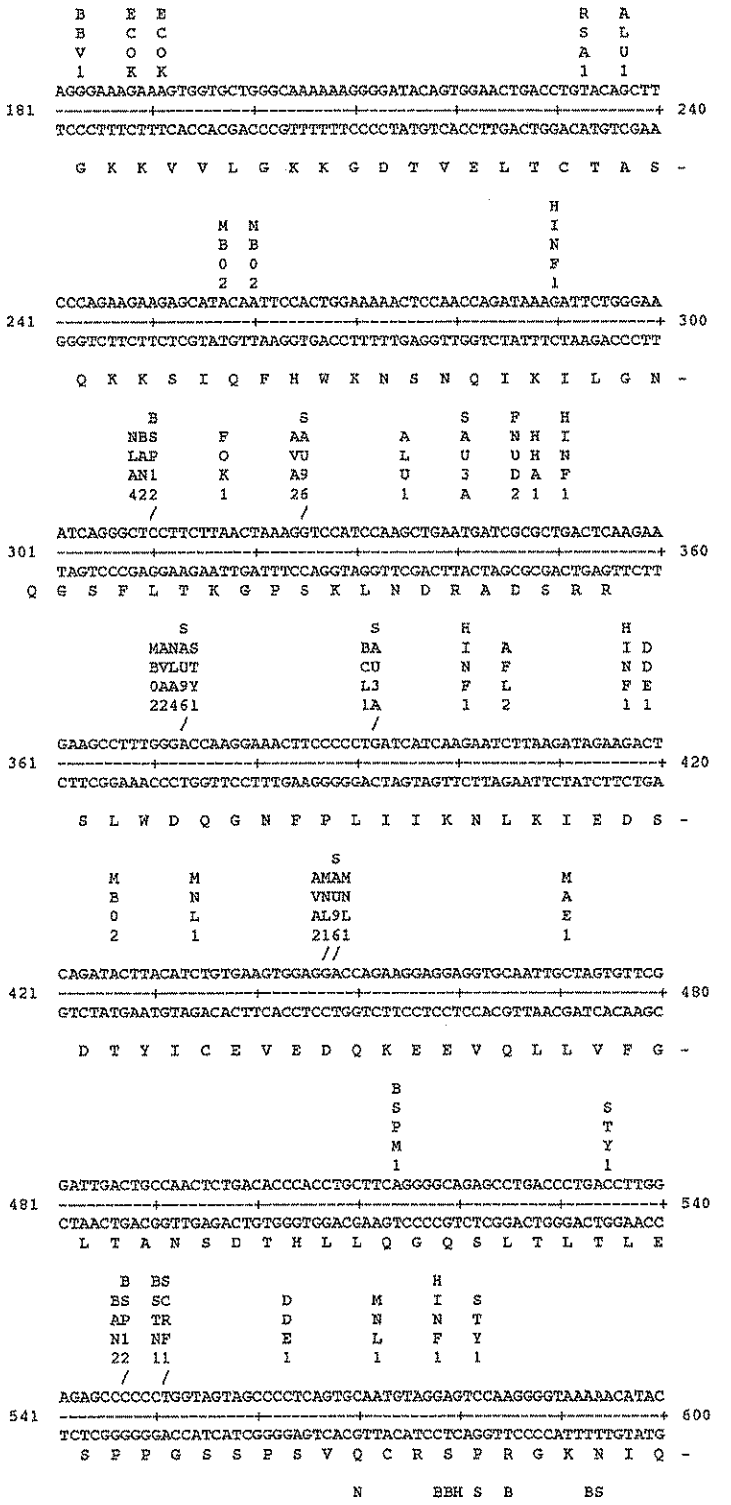


TABLE 4-continued

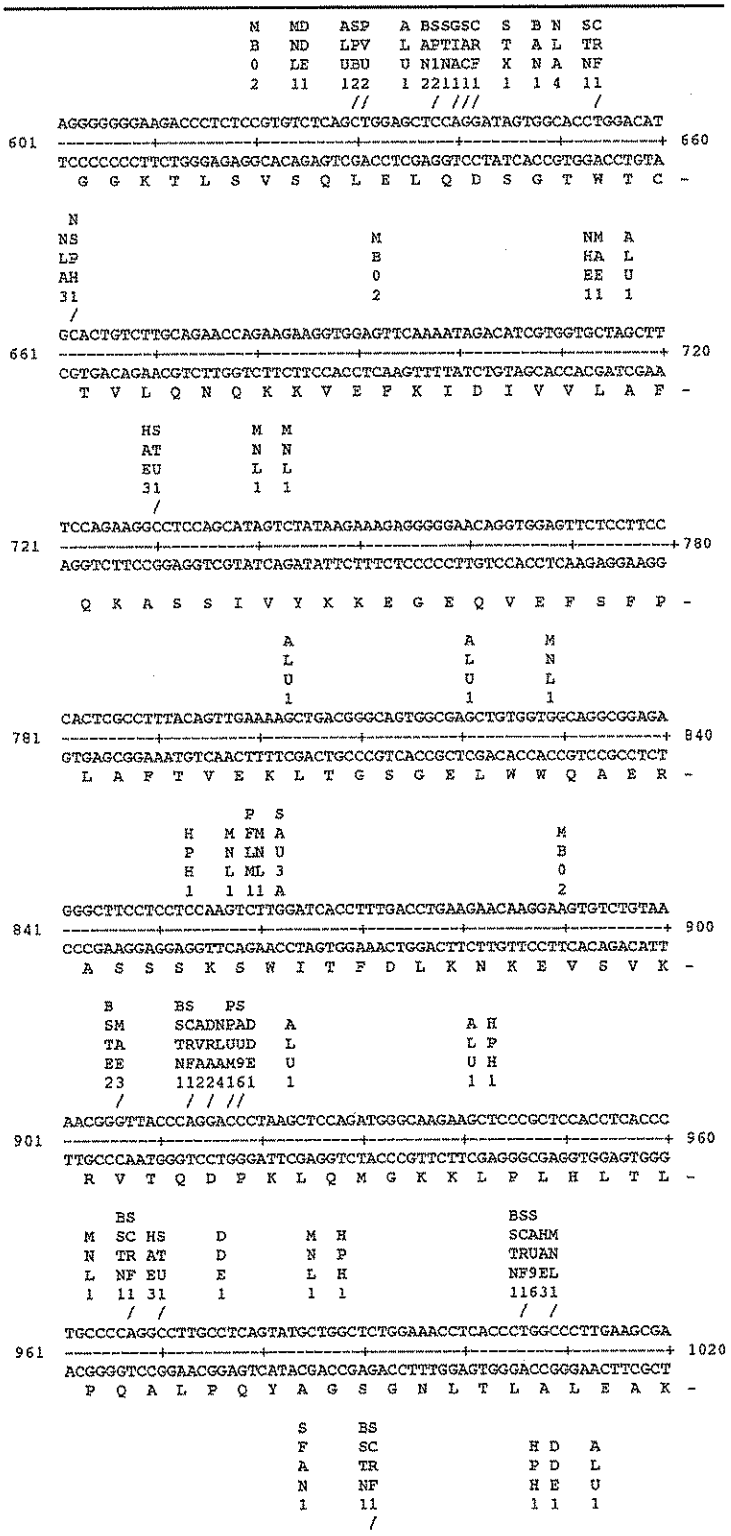


TABLE 4-continued

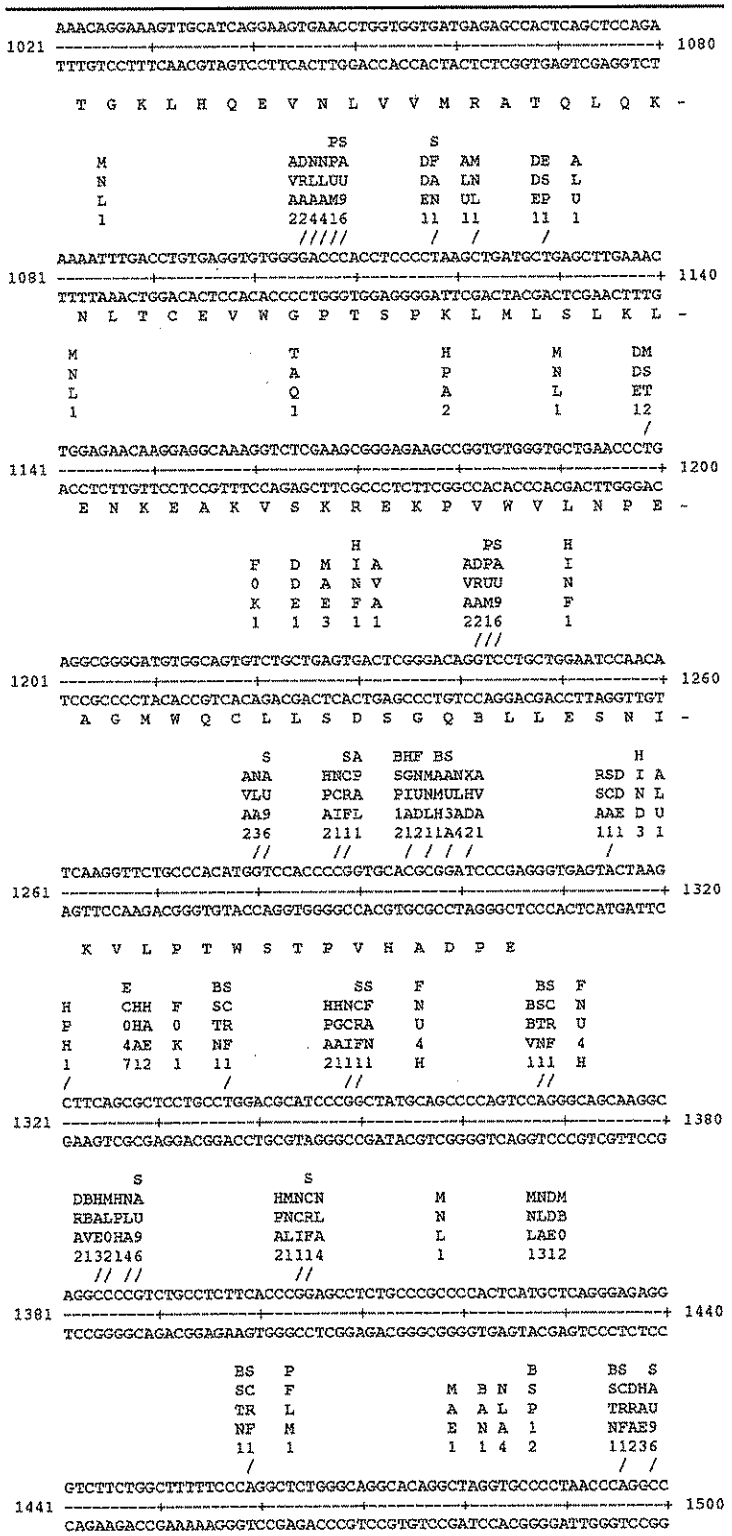


TABLE 4-continued

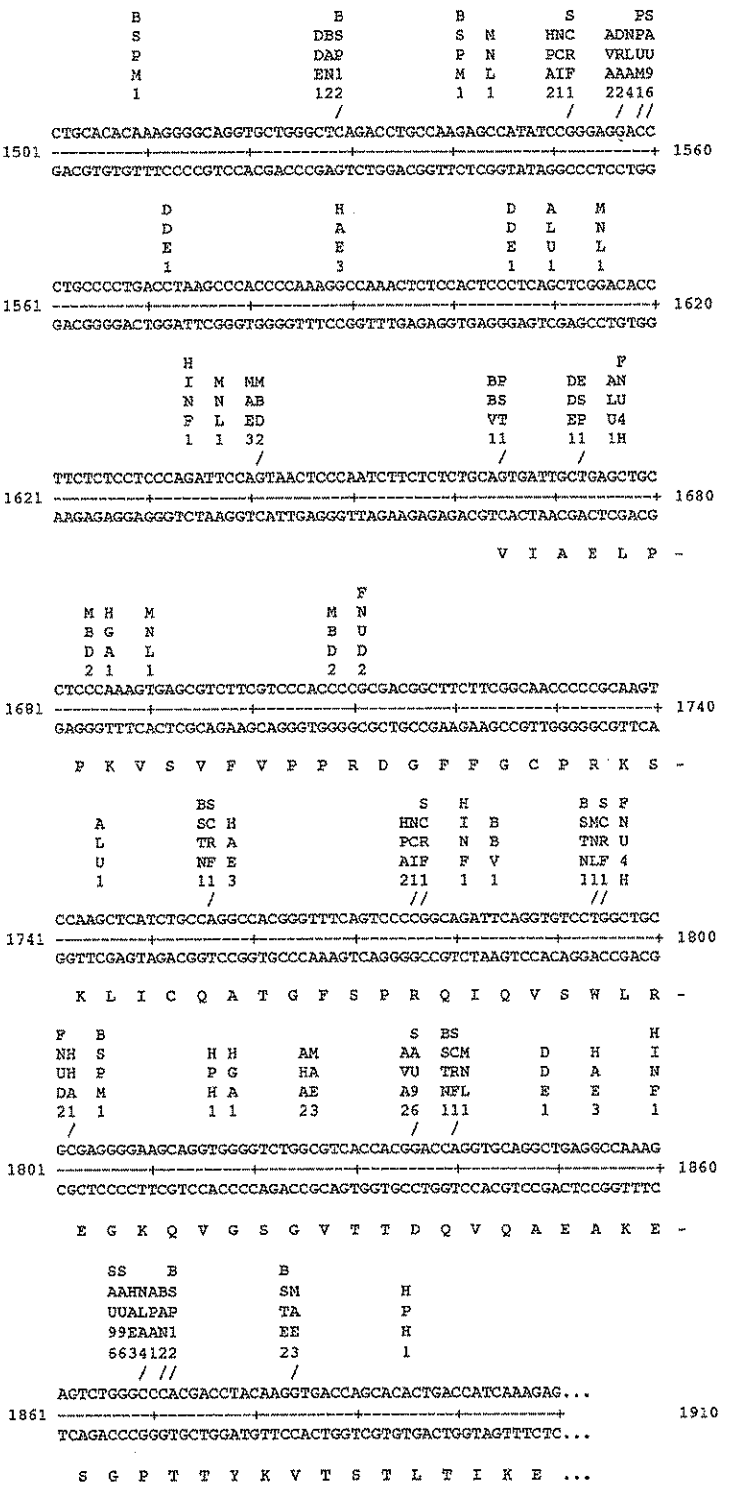
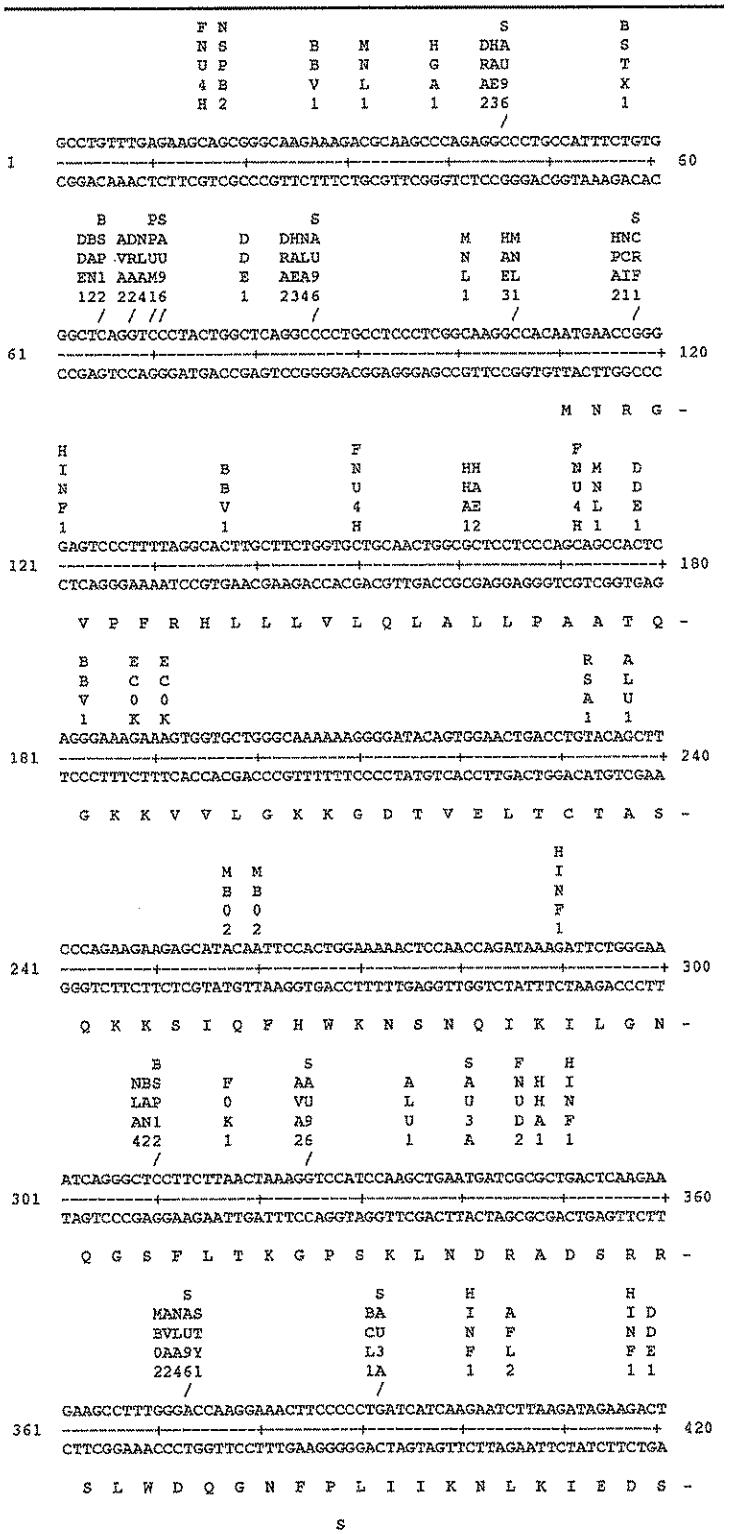


TABLE 5


GCCTGTTTGAGAAGCAGCGGGCAAGAAAGACGCAAGCCCAGAGGCCCTGCCATTTCTGTG
CGGACAAACTCTTCGTCGCCCGTTCTTTCTGCGTTCGGGTCTCCGGGACGGTAAAGACAC
GGCTCAGGTCCCTACTGGCTCAGGCCCCTGCCTCCCTCGGCAAGGCCACAATGAACCGGG
CCGAGTCCAGGGATGACCGAGTCCGGGGACGGAGGGAGCCGTTCCGGTGTTACTTGGCCC
GAGTCCCTTTTAGGCACTTGCTTCTGGTGCTGCAACTGGCGCTCCTCCCAGCAGCCACTC
CTCAGGGAAAATCCGTGAACGAAGACCACGACGTTGACCGCGAGGAGGGTCGTCGGTGAG
AGGGAAAGAAAGTGGTGCTGGGCAAAAAAGGGGATACAGTGGAACTGACCTGTACAGCTT
TCCCTTTCTTTCACCACGACCCGTTTTTTCCCCTATGTCACCTTGACTGGACATGTCGAA
CCCAGAAGAAGAGCATACAATTCCACTGGAAAAACTCCAACCAGATAAAGATTCTGGGAA
GGGTCTTCTTCTCGTATGTTAAGGTGACCTTTTGAGGTTGGTCTATTTCTAAGACCCTT
ATCAGGGCTCCTTCTTAACTAAAGGTCCATCCAAGCTGAATGATCGCGCTGACTCAAGAA
TAGTCCCGAGGAAGAATTGATTTCCAGGTAGGTTCGACTTACTAGCGCGACTGAGTTCTT
GAAGCCTTTGGGACCAAGGAAACTTCCCCCTGATCATCAAGAATCTTAAGATAGAAGACT
CTTCGGAAACCCTGGTTCCTTTGAAGGGGGACTAGTAGTTCTTAGAATTCTATCTTCTGA

S

TABLE 5-continued

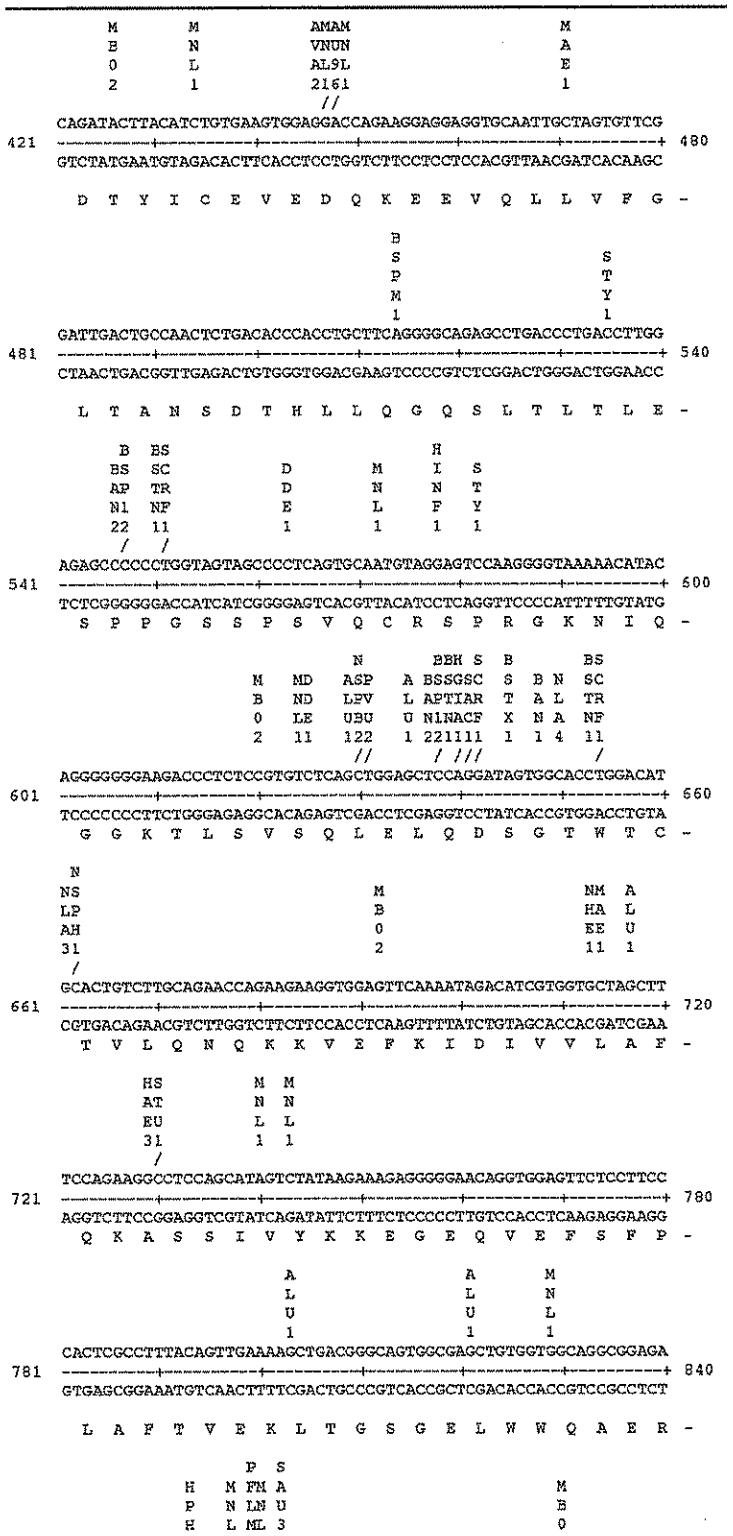


TABLE 5-continued

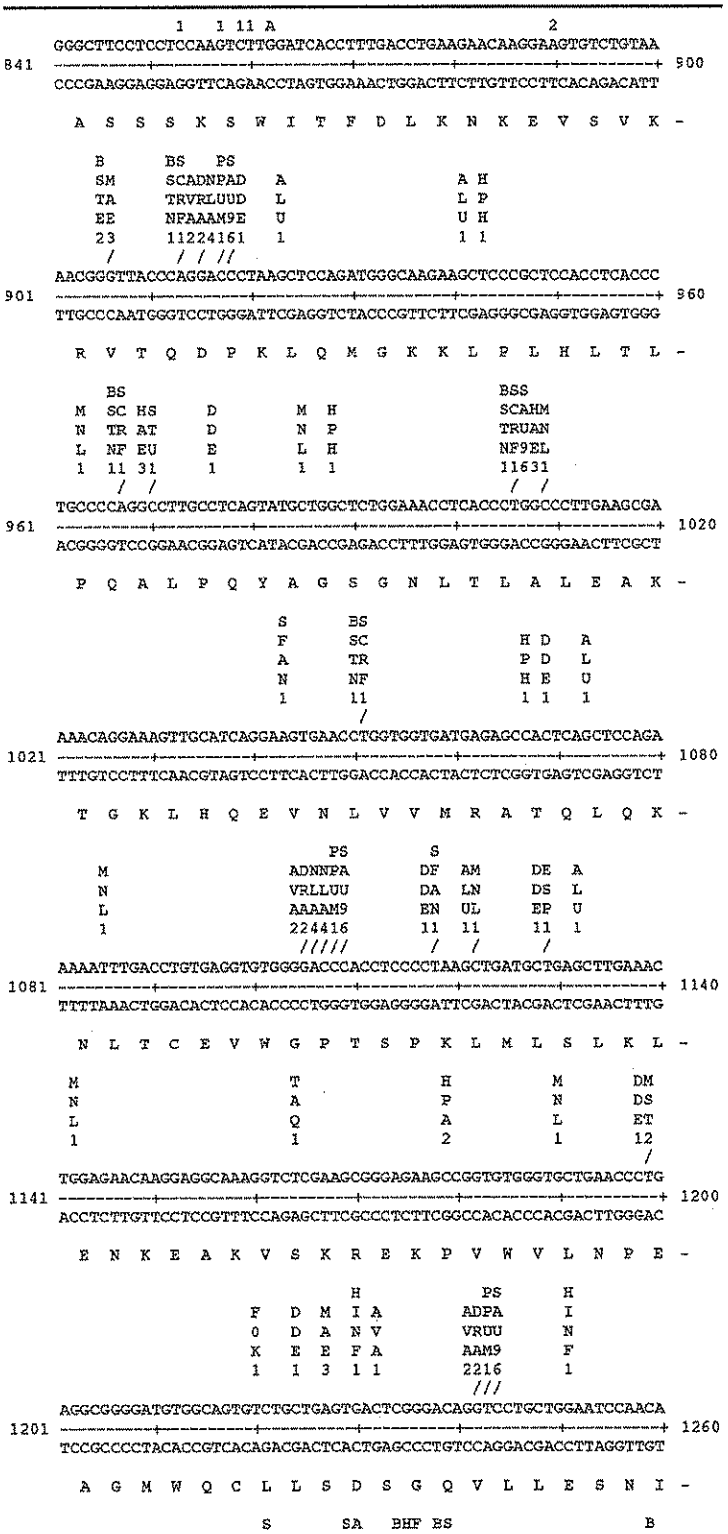


TABLE 5-continued

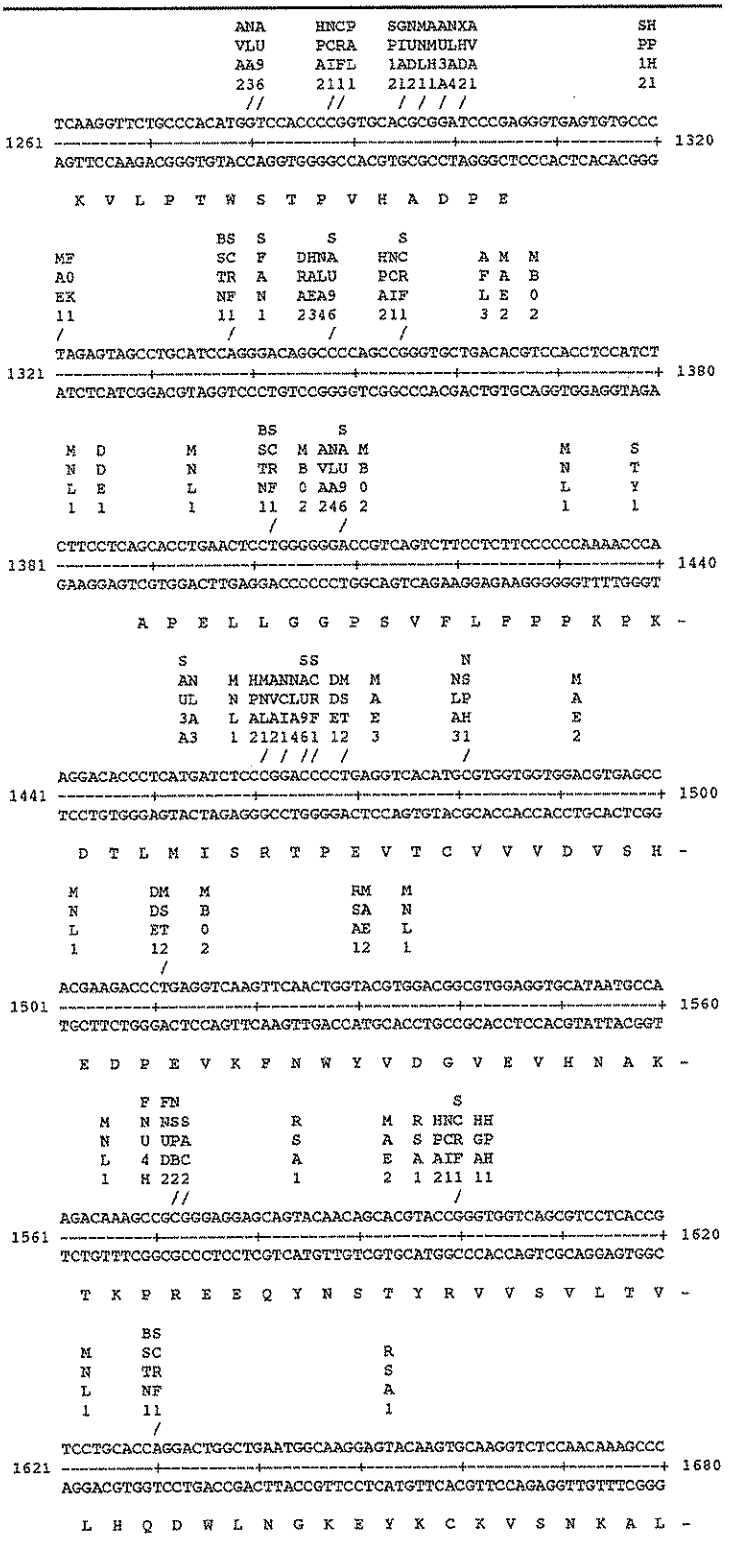


TABLE 5-continued

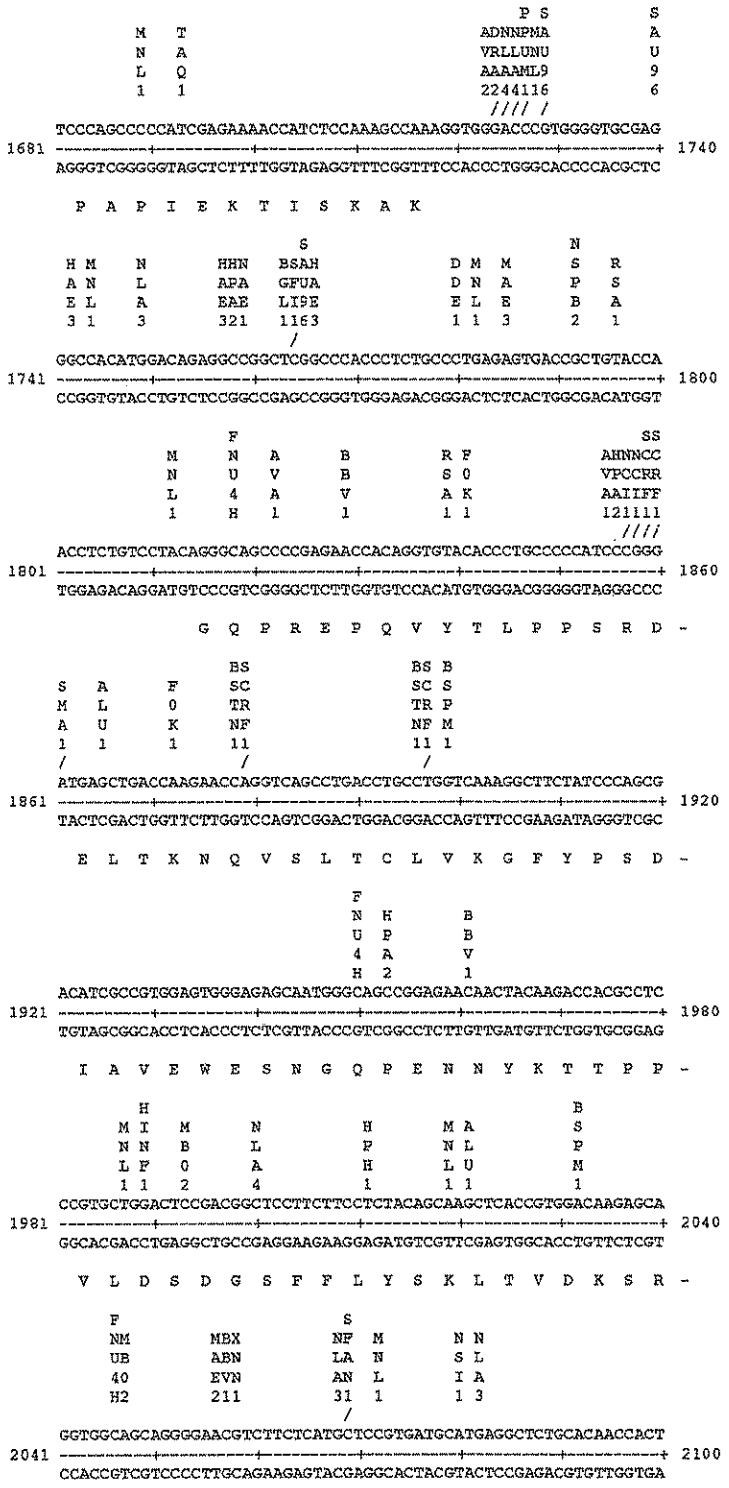


TABLE 5-continued

```
      W  Q  Q  G  N  V  F  S  C  S  V  M  H  E  A  L  H  N  H  Y  -

                                  S
                         M   M  HNC                CXH
                         B   N  PCR                FMA
                         0   L  AIF                RAE
                         2   1  211                133
                                  /                 /
     ACACGCAGAAGAGCCTCTCCCTGTCTCCGGGTAAATGAGTGCGACGGCCG
2101 ----------+----------+----------+----------+----------+          2150
     TGTGCGTCTTCTCGGAGAGGGACAGAGGCCCATTTACTCACGCTGCCGGC

      T  Q  K  S  L  S  L  S  P  G  K  *
```

Example 2

Preparation of the Fusion Proteins from Supernatants of COS Cells

COS cells grown in DME medium supplemented with 10% Calf Serum and gentamicin sulfate at 15 μg/ml were split into DME medium containing 10% NuSerum (Collaborative Research) and gentamicin to give 50% confluence the day before transfection. The next day, CsCl purified plasmid DNA was added to a final concentration of 0.1 to 2.0 μg/ml followed by DEAE Dextran to 400 μg/ml and chloroquine to 100 μM. After 4 hours at 37° C., the medium was aspirated and a 10% solution of dimethyl sulfoxide in phosphate buffered saline was added for 2 minutes, aspirated, and replaced with DME/10% Calf Serum. 8 to 24 hours later, the cells were trypsinized and split 1:2.

For radiolabeling, the medium was aspirated 40 to 48 hours after transfection, the cells washed once with phosphate buffered saline, and DME medium lacking cysteine or methionine was added. 30 minutes later, ^{35}S-labeled cysteine and methionine were added to final concentrations of 30–60 μci and 100–200 μci respectively, and the cells allowed to incorporate label for 8 to 24 more hours. The supernatants were recovered and examined by electrophoresis on 7.5% polyacrylamide gels following denaturation and reduction, or on 5% polyacrylamide following denaturation without reduction. The CD4Bγ1 protein gave the same molecular mass with or without reduction, while the CD4Eγ1 and CD4Hγ1 fusion proteins showed molecular masses without reduction of twice the mass observed with reduction, indicating that they formed dimer structures. The CD4 IgM fusion proteins formed large multimers beyond the resolution of the gel system without reduction, and monomers of the expected molecular mass with reduction.

Unlabeled proteins were prepared by allowing the cells to grow for 5 to 10 days post transfection in DME medium containing 5% NuSerum and gentamicin as above. The supernatants were harvested, centrifuged, and purified by batch adsorption to either protein A trisacryl, protein A agarose, goat anti-human IgG antibody agarose, rabbit anti-human IgM antibody agarose, or monoclonal anti-CD4 antibody agarose. Antibody agarose conjugates were prepared by coupling purified antibodies to cyanogen bromide activated agarose according to the manufacturer's recommendations, and using an antibody concentration of 1 mg/ml. Following batch adsorption by shaking overnight on a rotary table, the beads were harvested by pouring into a sintered glass funnel and washed a few times on the funnel with phosphate buffered saline containing 1% Nonidet P40 detergent. The beads were removed from the funnel and poured into a small disposable plastic column (Quik-Sep QS-Q column, Isolab), washed with at least 20 column volumes of phosphate buffered saline containing 1% Nonidet P40, with 5 volumes of 0.15 M NaCl, 1 mM EDTA (pH 8.0), and eluted by the addition of either 0.1 M acetic acid, 0.1 M acetic acid containing 0.1 M NaCl, or 0.25 M glycine-HCl buffer, pH 2.5.

Example 3

Blockage of Syncytium Formation by the Fusion Proteins

Purified or partially purified fusion proteins were added to HPB-ALL cells infected 12 hours previously with a vaccinia virus recombinant encoding HIV envelope protein. After incubation for 6–8 more hours, the cells were washed with phosphate buffered saline, fixed with formaldehyde, and photographed. All of the full-length CD4 immunoglobulin fusion proteins showed inhibition of syncytium formation at a concentration of 20 μg/ml with the exception of the 4Hγ1 protein, which was tested only at 5 μg/ml and showed partial inhibition of syncytium formation under the same conditions.

Example 4

Chromium Release Cytolysis Assay

The purified fusion proteins were examined for ability to fix complement in a chromium release assay using vaccinia virus infected cells as a model system. Namalwa (B cell) or HPB-ALL (T cell) lines were infected with vaccinia virus encoding HIV envelope protein, and 18 hours later were radiolabeled by incubation in 1 mci/ml sodium ^{51}chromate in phosphate buffered saline for 1 hour at 37°. The labeled cells were centrifuged to remove the unincorporated chromate, and incubated in microtiter wells with serial dilutions of the CD4 immunoglobulin fusion proteins and rabbit complement at a final concentration of 40%. After 1 hour at 37°, the cells were mixed well, centrifuged, and the supernatants counted in a gamma-ray counter. No specific release could be convincingly documented.

Example 5

Binding of the CD4Eγ1 Protein to Fc Receptors

Purified CD4Eγ1 fusion protein was tested for its ability to displace radiolabeled human IgG1 from human Fc receptors expressed on COS cells in culture. The IgG1 was radiolabeled with sodium ^{125}iodide using 1 mci of iodide, 100 μg of IgG1, and two idobeads (Pierce). The labeled protein was separated from unincorporated counts by passage over a Sephadex G25 column equilibrated with phosphate buffered saline containing 0.5 mM EDTA and 5% nonfat milk. Serial dilutions of the CD4Eγ1 fusion protein or unlabeled IgG1 were prepared and mixed with a constant amount of radiolabeled IgG1 tracer. After incubation with COS cells bearing the FcRI and RcRII receptors at 4° C. for at least 45 minutes in a volume of 20 μl, 200 μl of a 3:2

mixture of dibutyl to dioctyl phthalates were added, and the cells separated from the unbound label by centrifugation in a microcentrifuge for 15 to 30 seconds. The tubes were cut with scissors, and the cell pellets counted in a gamma-ray counter. The affinity of the CD4Eγ1 protein for receptors was measured in parallel with the affinity of the authentic IgG1 protein, and was found to be the same, within experimental error.

Example 6

Stable Expression of the Fusion Construct pCD4Eγ1 in Baby Hamster Kidney Cells

Twenty-four hours before transfection, 0.5×10^6 baby hamster kidney cells (BHK; ATCC CCL10) were seeded in a 25 cm^2 culture flask in Dulbecco's modified Eagle's medium (DMEM) containing 10% of fetal calf serum (FCS). The cells were cotransfected with a mixture of the plasmids pCD4Eγ1 (20 μg), pSV2dhfr (5 μg; Lee et al., *Nature* 294:228–232 (1981)) and pRMH140 (5 μg, Hudziak et al., *Cell* 31:137–146 (1982)) according to a modified calcium phosphate transfection technique as described in Zettlmeissl et al. (*Behring Inst. Res. Comm.* 82:26–34 (1988)). 72 h post-transfection, cells were split 1:3 to 1:4 (60 mm culture dishes) and resistant colonies were selected in DMEM medium containing 10% FCS, 400 μg/ml G418 (Geneticin, Gibco) and 1 μM methotrexate (selection medium). The medium was changed twice a week. The resistant colonies (40–100/transfection) appeared 10–15 day post-transfection and were further propagated either as a mixture of clones (i.e., BHK-MK1) or as individually isolated clones. For the determination of the relative expression levels, clone mixtures or individual clones were grown to confluency in T25 culture flasks, washed twice with protein-free DMEM medium, and incubated for 24 h with 5 ml protein-free DMEM medium. These media were collected and subjected to a human IgG specific ELISA in order to determine the relative expression levels of the CD4-IgG1 fusion protein CD4Eγ1. For further analysis an individual clone (BHK-UC3) was chosen due to its high relative expression levels.

Example 7

Detection of the CD4Eγ1 Protein in Culture Supernatants

For ^{35}S methionine labeling of cells, the clone BHK-UC3 and untransfected BHK cells (control) were grown to confluency in T25 culture flasks and subsequently incubated for two hours in HamF12 medium without methionine. Labeling was achieved by incubating 24 h in 2.5 ml of the same medium containing 100 μCi ^{35}S methionine (1070 Ci/mmole, Amersham). For the preparation of cell lysates, the labeled cells were harvested in 1 ml of phosphate buffered saline, pH 7.2 (PBS) and lysed by repetitive freezing and thawing. Cleared lysates (after centrifugation 20000 rpm, 20 min) and culture supernatants were incubated with Protein A-Sepharose (Pharmacia) and the bound material was analyzed on a 10% SDS-Protein A-Sepharose (Pharmacia) and the bound material was analyzed on a 10% SDS-gel according to Laemmli (*Nature* 227:680–685 (1970)), which was subsequently autoradiographed. A specific band of about 80 KDa can be detected only in the supernatant of clone BHK-UC3, which is absent in the lysate of clone BHK-UC3 and in the respective controls.

Example 8

Purification of the Protein CD4Eγ1 from Culture Supernatants

In order to demonstrate that the fusion protein coded by the plasmid pCD4Eγ1 can be obtained in high quantities, the clone BHK-UC3 was grown in 1750 cm^2 roller bottles in selection medium (500 ml). Confluent monolayers were washed twice with protein-free DMEM medium (200 ml) and further incubated for 48 h with protein-free DMEM medium (500 ml). The conditioned culture supernatants (1–2 l) and respective supernatants from untransfected BHK cells were cleared by centrifugation (9000 rpm, 30 min) and microfiltered through a 0.45 μm membrane (Nalgene). After addition of 1% (v/v) of 1.9 M Tris-HCl buffer, pH 8.6, the conditioned medium was absorbed to a Protein A-Sepharose column equilibrated with 50 mM Tris-HCl pH 8.6 buffer containing 150 mM NaCl (4° C.). The loaded column was washed with 10 column volumes of equilibration buffer. Elution of the CD4-IgG1 fusion protein CD4Eγ1 was achieved with 0.1 M sodium citrate buffer, pH 3, followed by immediate neutralization of the column efflux to pH 8 by Tris-base. The peak fractions were pooled, and the pool was analyzed on a Coomassie blue stained SDS-gel resulting in a band of the expected size (80 KDa), and which reacted with a polyclonal anti-human IgG heavy chain antibody and a mouse monoclonal anti-CD4 antibody (BMA040, Behringwerke) in Western Blots. The yields of purified fusion proteins obtained by the given procedure is 5–18 mg/24 h/l culture supernatant. The respective value for a BHK clone mixture (about 80 resistant clones; BHK-MK1) as described above was 2–3 mg/24 h/l.

Example 9

Physical and Biological Characterization of the CD4Eγ1 Fusion Protein

As proven by SDS-electrophoresis on 10–15% gradient gels (Phast-System, Pharmacia) under non-reductive conditions, the CD4Eγ1 fusion protein migrates at the position of a homodimer (about 160 KDa) like a non-reduced mouse monoclonal antibody. This result is supported by analytical equilibrium ultracentrifugation, where the fusion protein behaves as a homogeneous dimeric molecule of about 150 KDa. The absorbance coefficient of the protein was determined as A_{280}=18 cm^2/mg using the quantitative protein determination according to Bradford (*Anal. Biochem.* 72:248–254 (1976)).

The CD4Eγ1-fusion protein shows specific complex formation with a solubilized βgal-gp120 fusion protein (pMB1790; Broker et al., *Behring Inst. Res. Commun.* 82:338–348 (1988)) expressed in *E. coli*. In this protein (110 KDa), a major part of the HIV gp120 protein (Val_{49}-Trp_{645}) is fused to f-galactosidase (amino acids 1–375). In a control experiment a 67-KDa βgal-HIV 3'orf fusion protein (βgal1–375; 3'orf Pro14-Asp123) showed no complex formation. In these experiments, the CD4Eγ1-protein was incubated with the respective fusion protein in molar rations of about 5:1. The complex was isolated by binding to Protein A-Sepharose and the Protein A-Sepharose bound proteins—together with relevant controls—were analyzed on 10–15% gradient SDS-gels (Phast-System, Pharmacia).

The CD4Eγ1 fusion protein binds to the surface of HIV (HIV1/HTLV-IIIB) infected cultured T4-lymphocytes as determined by direct immunofluorescence with fluorescein-isothiocyanate (FITC) labeled CD4Eγ1 protein. It blocks syncytia formation in cultured T4-lymphocytes upon HIV infection (0.25 TCID/cell) at a concentration of 10 μg/ml. Furthermore, HIV-infected cultured T4-lymphocytes (subclone of cell line H9) are selectively killed upon Incubation with CD4Eγ1 in the presence or absence of complement: To a highly (>50%) HIV infected culture of T4-lymphocytes (10^6 cells/ml) 50, 10 or 1 μg/ml CD4Eγ1 fusion protein was added in the presence or absence of guinea pig complement. Cells were observed for specific killing by the fusion protein, which is defined by the

percentage of killed cells after 3 days in relation to viable cells in the culture at the beginning of the experiment corrected by the values for unspecific killing observed in control cultures, lacking the CD4Eγ1 fusion protein (Table 5, Experiment I). Surprisingly, addition of CD4Eγ1 protein to the infected T4 cells in the absence of complement resulted in similar specific killing rates as in the presence of complement (Table 5, Experiment II). This result demonstrates a complement independent cytolytic effect of CD4Eγ1 on HIV infected T-lymphocytes in culture.

TABLE 5

No. Experiment	Assay System	Specific Killing (%)
I	non-infected T4-cells + 50 μg/ml CD4Eγ1 + Compl.	0.7
	infected T4-cells + 50 μg/ml CD4Eγ1 + Compl.	35.1
	infected T4-cells + 10 μg/ml CD4Eγ1 + Compl.	25.1
	infected T4-cells + 1 μg/ml CD4Eγ1 + Compl.	25
II	infected T4-cells + 10 μg/ml CD4Eγ1 + Compl.	49.9
	infected T4-cells + 10 μg/ml CD4Eγ1 + Compl.	69.4

Having now fully described this invention, it will be appreciated by those skilled In the art that the same can be performed with any wide range of equivalent parameters of composition, conditions, and methods of preparing such fusion proteins without departing from the spirit or scope of the invention or any embodiment thereof.

What is claimed is:

1. A nucleic acid molecule encoding a fusion protein comprising 1) a DNA sequence encoding amino acids 1–173 of CD4, and 2) a DNA sequence encoding an immunoglobulin heavy chain, wherein the DNA sequence which encodes at least the variable region of said immunoglobulin chain has been replaced with the DNA sequence which encodes amino acids 1–173 of CD4, whereby a nucleic acid molecule encoding a fusion protein capable of being secreted is formed.

2. The fusion protein gene of claim 1, wherein said immunoglobulin chain is of the class IgM, IgG1 or IgG3.

3. A vector comprising the nucleic acid molecule of claim 1.

4. A host cell transformed with the vector of claim 3.

5. The host cell of claim 4 which expresses an immunoglobulin light chain together with the expression product of said fusion protein gene.

6. A method of producing a fusion protein which is capable of being secreted comprising extracellular CD4 through amino acid 173 of CD4, which fusion protein binds to gp120, and an immunoglobulin heavy chain, wherein at least the variable region of the immunoglobulin chain has been substituted with extracellular CD4 through amino acid 173, which binds to HIV or SIV gp120, which comprises

cultivating in a nutrient medium under protein-producing conditions, a host strain transformed with the vector of claim 3, said vector further comprising expression signals which are recognized by said host strain and direct expression of said fusion protein, and

recovering the fusion protein so produced.

7. The method of claim 6, wherein said host cell produces immunoglobulin light chains and said fusion protein comprises an immunoglobulin heavy chain of the class IgM, IgG1 or IgG3.

8. A nucleic acid molecule encoding a fusion protein comprising 1) a DNA sequence encoding amino acids 1–173 of CD4, and 2) a DNA sequence encoding an immunoglobulin light chain, where in the DNA sequence which encodes the variable region of said immunoglobulin light chain has been replaced with the DNA sequence which encodes amino acids 1–173 of CD4, whereby a nucleic acid molecule encoding a fusion protein capable of being secreted is formed.

9. A vector comprising the nucleic acid molecule of claim 8.

10. A host cell transformed with the vector of claim 9.

11. The host cell of claim 10 which expresses an immunoglobulin heavy chain together with the expression product of said fusion protein gene.

12. The host cell of claim 11, wherein said immunoglobulin heavy chain is of the immunoglobulin class IgM, IgG1 or IgG3.

13. A method of producing a fusion protein which is capable of being secreted comprising extracellular CD4, or fragment thereof which binds to gp120 when fused to an immunoglobulin chain, and an immunoglobulin light chain, wherein the variable region of the immunoglobulin chain has been substituted with extracellular CD4, or said fragment thereof which binds to HIV or SIV gp120, which comprises

cultivating in a nutrient medium under protein-producing conditions, a host cell transformed with the vector of claim 9, said vector further comprising expression signals which are recognized by said host strain and direct expression of said fusion protein, and

recovering the fusion protein so produced.

14. The method of claim 13, wherein said host cell produces immunoglobulin heavy chains of the class IgM, IgG1 or IgG3, together with said fusion protein.

* * * * *

UNITED STATES PATENT AND TRADEMARK OFFICE

CERTIFICATE OF CORRECTION

PATENT NO. : 6,004,781

DATED : December 21, 1999

INVENTOR(S) : Brian SEED

It is certified that error appears in the above-identified patent and that said Letters Patent is hereby corrected as shown below:

In Column 73, line 42, please delete "fusion protein gene" and replace with --nucleic acid molecule--.

Signed and Sealed this

Nineteenth Day of September, 2000

Attest:

Q. TODD DICKINSON

Attesting Officer

Director of Patents and Trademarks

EXHIBIT 13

US005336603A

United States Patent [19]

Capon et al.

[11] Patent Number: 5,336,603

[45] Date of Patent: Aug. 9, 1994

[54] CD4 ADHESON VARIANTS

[75] Inventors: Daniel J. Capon, San Mateo; Timothy J. Gregory, Hillsborough, both of Calif.

[73] Assignee: Genentech, Inc., South San Francisco, Calif.

[21] Appl. No.: 936,190

[22] Filed: Aug. 26, 1992

Related U.S. Application Data

[63] Continuation of Ser. No. 842,777, Feb. 18, 1992, abandoned, which is a continuation of Ser. No. 250,285, Sep. 28, 1988, abandoned, which is a continuation-in-part of Ser. No. 104,329, Oct. 2, 1987, abandoned.

[51] Int. Cl.[5] C07K 13/00; C12N 15/62

[52] U.S. Cl. 435/69.7; 435/257.3; 435/330.1; 536/350; 536/387.3; 536/23.4; 424/134.1

[58] Field of Search 435/69.7, 252.3, 320.1, 435/5; 530/250, 387, 27; 536/23.4

[56] References Cited

U.S. PATENT DOCUMENTS

3,959,080	5/1976	Orth et al.	195/63
4,002,531	1/1977	Royer	195/68
4,055,635	10/1977	Green et al.	424/78
4,301,144	11/1981	Iwashita et al.	424/78
4,412,989	11/1983	Iwashita et al.	424/177
4,444,878	4/1984	Paulus	435/7
4,745,055	5/1988	Schenk et al.	435/7
4,761,371	8/1988	Bell et al.	435/68
4,766,106	8/1988	Katre et al.	514/12
4,816,567	3/1989	Cabilly et al.	530/387
4,879,211	11/1989	Wang et al.	435/5

FOREIGN PATENT DOCUMENTS

0068763	1/1983	European Pat. Off. .
0088695	9/1983	European Pat. Off. .
0120694	10/1984	European Pat. Off. .
0139416	5/1985	European Pat. Off. .
0173494	3/1986	European Pat. Off. .
0296786	12/1986	European Pat. Off. .
0244221	11/1987	European Pat. Off. .
0255694	2/1988	European Pat. Off. .
0256654	2/1988	European Pat. Off. .
0266663	5/1988	European Pat. Off. .
0278776	8/1988	European Pat. Off. .
0325262	1/1989	European Pat. Off. .
0313377	4/1989	European Pat. Off. .
0319815	6/1989	European Pat. Off. .
8503947	9/1985	PCT Int'l Appl. .
8703600	6/1987	PCT Int'l Appl. .
8801304	2/1988	PCT Int'l Appl. .
8803559	5/1988	PCT Int'l Appl. .
8809344	12/1988	PCT Int'l Appl. .
8901940	3/1989	PCT Int'l Appl. .

OTHER PUBLICATIONS

Chaudhary et al., *Proc. natl. Acad. Sci. U.S.A.* 84: 4538–4542 (1987).

Murre et al., *Mol. Cell. Biol.* 6: 1315–1319 (1986).

Oi, V. T., *BioTechniques* 4: 214 (1986).

Peterson, A. S., Ph.D., Thesis, Harvard Univ. Chapter 1 (1988).

McDougal, et al., *Molecular Biology of Homosapiens*, Cold Spring harbor, N.Y. Jnn. 1986.

Springer et al., Proc. Natl. Acad. Sci. U.S.A. 73(7): 2481–2485 (1976).

Seed, Nature 329:840–842 (Oct. 1987).

Goverman et al. J. Cell. Biochem. Suppl. (11 Part D) 1987, AB 334, p. 259.

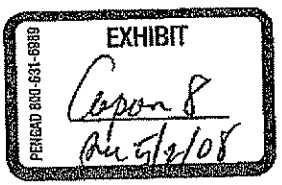
(List continued on next page.)

Primary Examiner—Robert J. Hill, Jr.
Assistant Examiner—John D. Ulm
Attorney, Agent, or Firm—Ginger R. Dreger; Janet E. Hasak

[57] ABSTRACT

Novel derivatives of cell surface proteins which are homologous to the immunoglobulin superfamily (adhesons) are provided. Amino acid sequence variations are introduced into adheson, the most noteworthy of which are those in which the transmembrane and, preferably, cytoplasmic domains are rendered functionally inactive, and in which adheson extracellular domains replace an immunoglobulin variable region. These variants are useful in therapy or diagnostics.

12 Claims, 13 Drawing Sheets



OTHER PUBLICATIONS

Johnson et al., J. Cell. Biochem. Suppl. (11 Part D) 1987 AB 337, p. 260.
Salzawa et al., J. Cell. Biochem. Suppl. (11 Part D) 1987, AB 421, p. 273.
Hashimoto et al., J. Cell. Biochem. Suppl. (11 Part D) 1987 AB 434, p. 278.
Ivars et al., J. Cell. Biochem. Suppl. (11 Part D) 1987, AB 435, p. 278.
Sleckman et al., Nature 328:351 (Jul. 1987).
Gascoigne et al. PNAS U.S.A. 84:2936–2940 (1987).
Anderson et al. Immunology Today 9(7&8):199–203 (1988).
Littman, Ann. Rev. Immunol. 5:561–584 (1987).
Estess et al., J. Cell. Biochem. Suppl. (11 Part D) 1987, AB. 331, p. 258.
Gascoigne et al., J. Cell Biochem. Suppl. (11 Part D) 1987, AB.333, p. 259.
Littman et al., Cell 40:237–246 (1985).
Beauchamp et al., Analytical Biochem. 131:25–33 (1983).
Terhorst et al., Science 209:520–521 (1980).
McDougal et al., Science 231:382–385 (1986).
Maddon et al., Cell 47:333–348 (1986).
Morrison, S. L. Science 229:1202–1207 (1985).
Wills et al., J. Cell Biol. 99:2011–2023 (1984).
Rose et al., Cell 30:753–762 (1982).
Traunecker et al., Nature 331:84–86 (1988).
Traunecker et al., Nature 339:68–70 (1989).
Maddon et al., P.N.A.S. U.S.A. 84:9155–9159 (Dec. 1987).
Morrison et al., P.N.A.S. U.S.A. 81:6851–6855 (1984).
Morrison, S. L. J. Immunol. 123(2):793–800 (1979).
Neuberger et al., Natnre 312:604–608 (1984).
Vitetta et al., Science 238:1098–1104 (1987).
Falkner et al., Nature 298:286–288 (1982).
Köhler, G. P.N.A.S. U.S.A. 77(4):2197–2199 (1980).
Maddon et al., Cell 42(1): 93–104 (1985).
Clark et al., P.N.A.S. U.S.A. 84:1649–1653 (1987).
Smith et al., Science 238:1704–1707 (1987).
Morrison et al., Paul et al., Ed., Ann. Rev. Immunol. 2:239–256 (1984).
Sharon et al., Nature 309:364–367 (1984).
Boulianne et al., Nature 312:643–646 (1984).
Rosenblum Michael G. et al., Cancer Research 45:2421–2424 (Jun. 1985).
Osborn, Laurelee et al., Cell 59:1203–1211 (Dec. 1989).

FIG. 1A

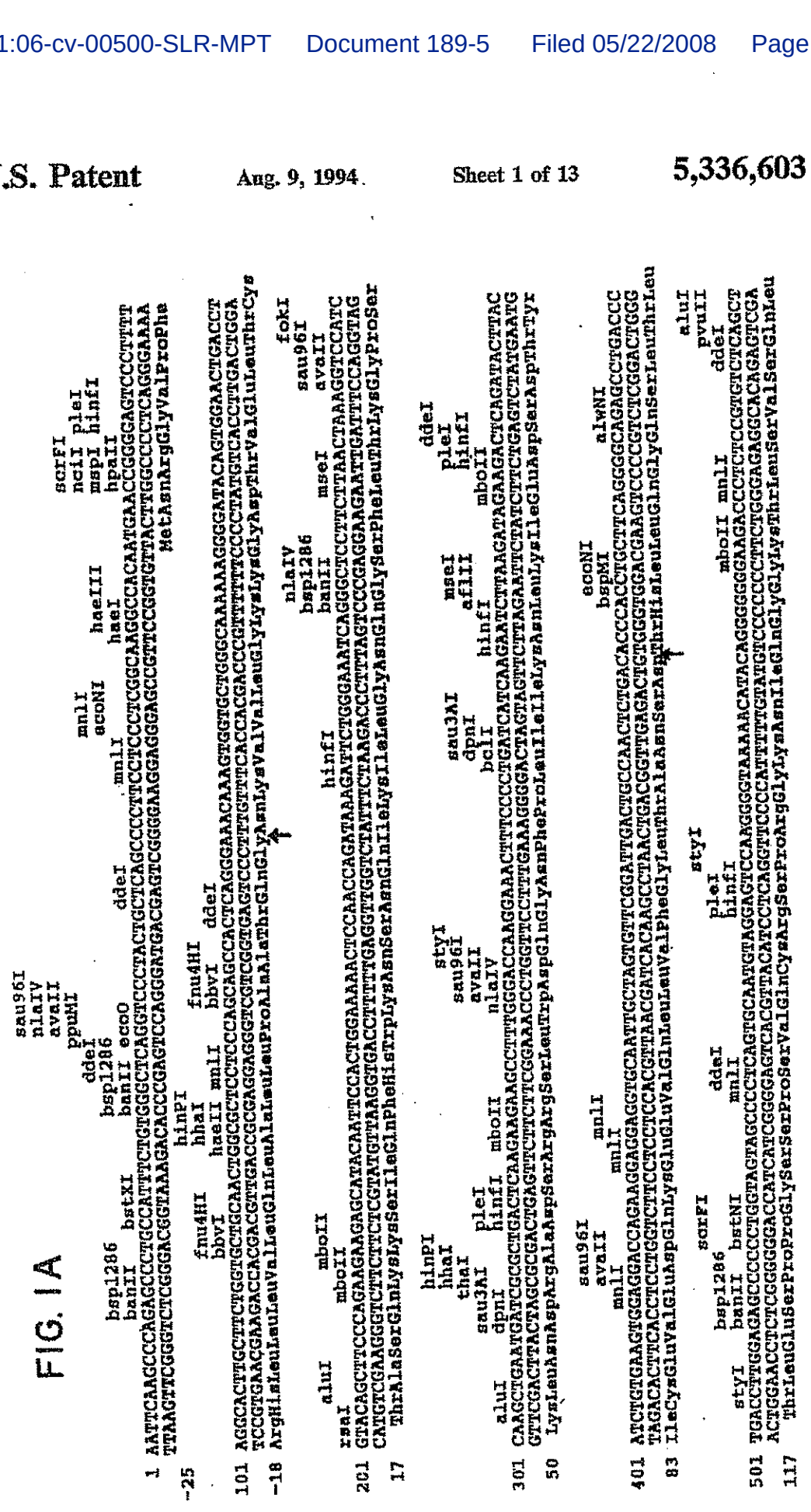
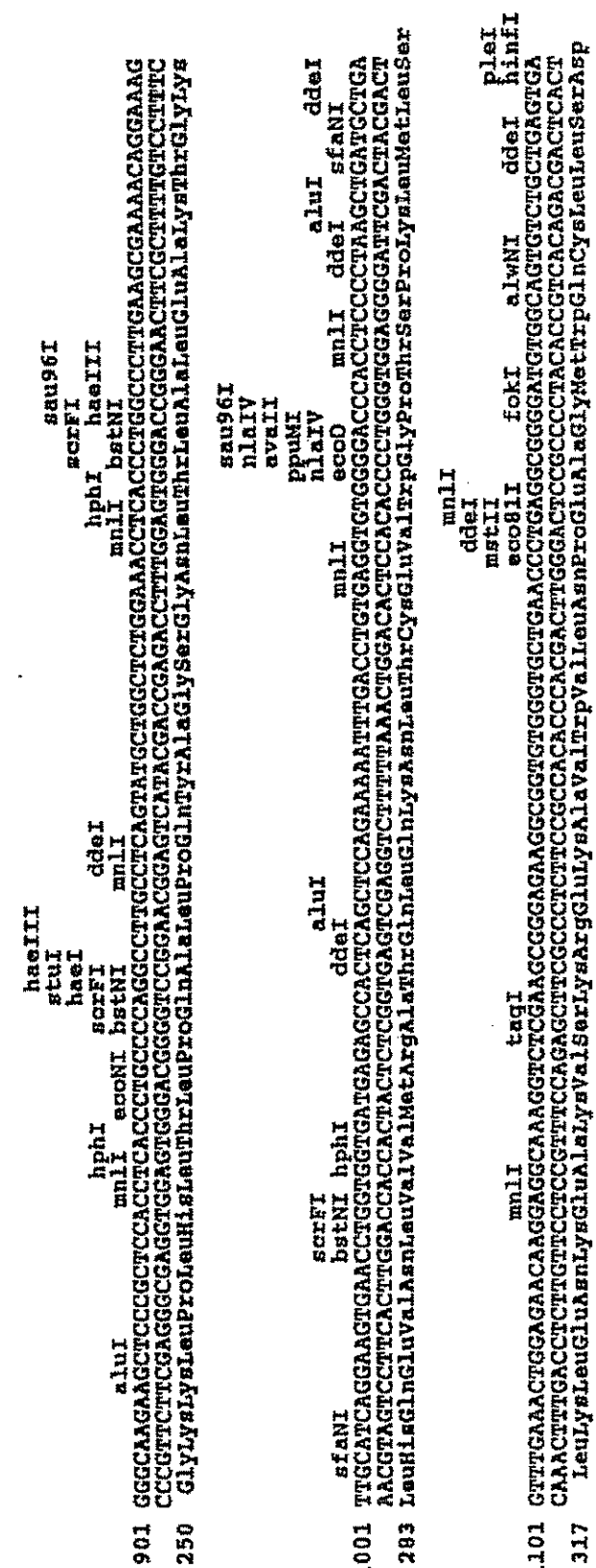

sau96I
nlaIV
ava II
ppuMI
ddeI
bsp1286
banII ecoO
bstXI
hinPI
hhaI
fnu4HI
bbvI
haeII mnlI
ddeI
mnlI
acoNI
haeI
haeIII
scrFI
ncil pleI
mspI hinfI
hpaII
nlaIV
bsp1286
banII
hinfI
mseI
bsp1286
banII
mseI
pleI
hinfI
aflII
mboII
dpnI
hinfI
sau3AI
dpnI
bclI
ddeI
sau96I
ava II
nlaIV
sau3AI
dpnI
ddeI
ecoNI
bspMI
alwNI
sau96I
ava II
mnlI
mnlI
styI
sau96I
ava II
styI
pleI
hinfI
mboII
mnlI
styI
ddeI
mnlI
pleI
hinfI
mboII
mnlI
ddeI
aluI
pvuII
aluI
rsaI
mboII
aluI
mboII
hinPI
hhaI
thaI
sau3AI
dpnI
pleI
hinfI
mboII
aluI
hinPI
hhaI
thaI
sau3AI
dpnI
pleI
hinfI
mboII
nlaIV
aluI
-25
-18
1
17
50
83
117
201
301
401
501
GCTTAAGTTCGGGTCTCGGGACGGTAAAGACACCCGAGTCCAGGGATGACGAGTCGGGGAAGGGAGGCAGCCGTTCCGGTGTTACTTGGCCCCTCAGGGAAAA
CGAATTCAAGCCCAGAGCCCTGCCATTTCTGTGGGCCTCAGGTCCCTACTGCTCAGCCCCTTCCTCCCTCGGCAAGGCCACAATGAACCGGGGAGTCCCTTTT
Met Asn Arg Gly Val Pro Phe
TCCGTGAAÇGAAGACCACGACGGTTGACCGCGAGGAGGGGTCGTCGGGTGAGTCCCTTTGTTTCACCACGACCCGTTTTTTCCCCTATGTCACCTTGACTGGA
AGGCACTTGCTTCTGGTGCTGCAACTGGCGCTCCTCCCAGCAGCCACTCAGGGAAACAAAGTGGTGCTGGGCAAAAAAGGGGATACAGTGGAACTGACCT
Arg His Leu Leu Val Leu Leu Gln Leu Ala Leu Leu Pro Ala Leu Ala Thr Gln Gly Asn Lys Val Val Leu Gly Lys Lys Gly Asp Thr Val Glu Leu Thr Cys
GTACAGCTTCCCCAGAAGAAGAGCATACAATTCCACTGGAAAAACTCCAACCAGATAAGATTCTGGGAAATCAGGGCTCCTTCTTAACTAAAGGTCCATC
CATGTCGAAGGGGTCTTCTTCTCGTATGTTAAGGTTGGGTCTATTTTCTAAGACCCTTTAGTCCCGAGGGAAGAATTGATTTCCAGGTAG
Thr Ala Ser Gln Lys Lys Ser Ile Gln Phe His Trp Lys Asn Ser Arg Asn Gln Ile Lys Ile Leu Gly Asn Gln Gly Ser Phe Leu Thr Lys Gly Pro Ser
GTTCGACTTACTAGCGCGACTGAGTTCTTCTTCGGAAACCCTGGTTCCTTTGAAAGGGGACTAGTAGTTCTTAGAATTCTATCTTCTGAGTCTATGAATG
CAAGCTGAATGATCGCGCTGACTCAAGAAGAAGCCCTTTGGGACCAAGGAAACTTTCCCCTGATCATCAAGAATCTTAAGATAGAAGACTCAGATACTTAC
Lys Leu Asn Asp Arg Ala Asp Ser Arg Arg Ser Leu Trp Asp Gln Gly Asn Phe Pro Leu Ile Leu Lys Asn Leu Lys Ile Glu Asp Ser Asp Thr Tyr
ATCTGTGAAGTGGAGGACCAGAAGGAGGAGGTGCAATTGCTAGTGTTCGGATTGACTGCCAACTCTGACACCCACCTGCTTCAGGGGCAGAGCCCTGACCC
TAGACACTTCACCTCCTGGTCTTCCTCCTCCACGTTAACGATCACAAGCCTAACTGACGGGTTGAGACTGTGGGTGGACGAAGTCCCCGTCTCGGACTGGG
Ile Cys Glu Val Glu Asp Gln Lys Glu Glu Val Gln Leu Leu Val Pro Phe Gly Leu Thr Ala Asn Ser Arg Thr His Ile Leu Leu Gln Leu Gln Gly Leu Ser Leu Thr Leu
TGACCTTGGAGAGCCCCCCCTGGTAGTAGGCCCCCTCAGTGCAATGTAGGAGTCCAAGGGGTAAAAACATACAGGGGGGGAAGACCCTCTCCCGTGTCTCAGCT
ACTGGAACCTCTCGGGGGGGACCATCATCGTTACATCCTCAGGTTCCCCCATTTTTGTATGTCCCCCCCCTTCTGGGAGAGGCACAGAGTCGA
Thr Leu Glu Ser Pro Pro Gly Ser Ser Pro Ser Val Gln Cys Arg Ser Pro Arg Gly Lys Lys Asn Ile Glu Gln Gly Gly Lys Thr Leu Ser Val Ser Gln Leu

FIG. 1B-1



```
          bstXI
       aluI
      sacI bstNI          scrFI                                                                                  mnlI
      hgiAI               bstNI                                                                                 haeIII
      bsp1286         nlaIV                                                                        aluI        stuI
      banII           banI      nlaIII                  mboII                                  nheI            haeI
601  GGAGCTCCAGGATAGTGGCACCTGGACATGCACTGTCTTGCAGAACCAGAAGAAGGTGGAGTTCAAAATAGACATCGTGGTGCTAGCTTTCCAGAAGGCC
     CCTCGAGGTCCTATCACCGTGGACCTGTACGTGACAGAACGTCTTGGTCTTCTTCCACCTCAAGTTTTATCTGTAGCACCACGATCGAAAGGTCTTCCGG
150   GluLeuGlnAspSerGlyThrTrpThrCysThrValLeuGlnAsnGlnLysLysValGluPheLysIleAspIleValValLeuAlaPheGlnLysAla
                                                                                                        ↑
                          mnlI                                                  aluI                aluI
701  TCCAGCATAGTCTATAAGAAAGAGGGGGAACAGGTGGAGTTCTCCTTCCCACTCGCCTTTACAGTTGAAAAGCTGACGGGCAGTGGCGAGCTGTGGTGGC
     AGGTCGTATCAGATATTCTTTCTCCCCCTTGTCCACCTCAAGAGGAAGGGTGAGCGGAAATGTCAACTTTTCGACTGCCCGTCACCGCTCGACACCACCG
183  SerSerIleValTyrLysLysGluGlyGluGlnValGluPheSerPheProLeuAlaPheThrValGluLysLeuThrGlySerGlyGluLeuTrpTrpGln

                                                                                                sau96I
                                                                                                nlaIV
                                                                                                avaII
                                      hphI                                                     ppuMI
                                     sau3AI                                                  scrFI
                       mnlI          dpnI                                                    bstNI      aluI
          mnlI    mnlI  pflMI    alwI            mboII                          bstEII  ecoO  ddeI
801  AGGCGGAGAGGGCTTCCTCCTCCAAGTCTTGGATCACCTTTGACCTGAAGAACAAGGAAGTGTCTGTAAAACGGGTTACCCAGGACCCTAAGCTCCAGAT
     TCCGCCTCTCCCGAAGGAGGAGGTTCAGAACCTAGTGGAAACTGGACTTCTTGTTCCTTCACAGACATTTTGCCCAATGGGTCCTGGGATTCGAGGTCTA
217   AlaGluArgAlaSerSerSerLysSerTrpIleThrPheAspLeuLysAsnLysGluValSerValLysArgValThrGlnAspProLysLeuGlnMet
```



```
                                                haeIII
                                               stuI                                                  sau96I
                                               haeI                                               scrFI
                                 hphI       scrFI       ddeI                                hphI   haeIII
                 aluI          mnlI  ecoNI bstNI      mnlI                                mnlI  bstNI
 901 GGGCAAGAAGCTCCCGCTCCACCTCACCCTGCCCCAGGCCTTGCCTCAGTATGCTGGCTCTGGAAACCTCACCCTGGCCCTTGAAGCGAAAACAGGAAAG
     CCCGTTCTTCGAGGGCGAGGTGGAGTGGGACGGGGTCCGGAACGGAGTCATACGACCGAGACCTTTGGAGTGGGACCGGGAACTTCGCTTTTGTCCTTTC
 250  GlyLysLysLeuProLeuHisLeuThrLeuProGlnAlaLeuProGlnTyrAlaGlySerGlyAsnLeuThrLeuAlaLeuGluAlaLysThrGlyLys

                                                                                                sau96I
                                                                                                nlaIV
                                                                                                avaII
                                                                                               ppuMI
                                scrFI                    aluI                                  nlaIV                aluI    ddeI
       sfaNI                bstNI hphI              ddeI                          mnlI     ecoO    mnlI  ddeI   sfaNI
1001 TTGCATCAGGAAGTGAACCTGGTGGTGATGAGAGCCACTCAGCTCCAGAAAAATTTGACCTGTGAGGTGTGGGGACCCACCTCCCCTAAGCTGATGCTGA
     AACGTAGTCCTTCACTTGGACCACCACTACTCTCGGTGAGTCGAGGTCTTTTTAAACTGGACACTCCACACCCCTGGGTGGAGGGGATTCGACTACGACT
 283 LeuHisGlnGluValAsnLeuValValMetArgAlaThrGlnLeuGlnLysAsnLeuThrCysGluValTrpGlyProThrSerProLysLeuMetLeuSer

                                                                                    mnlI
                                                                                   ddeI
                                                                                  mstII                                  pleI
                                mnlI        taqI                                  eco81I    fokI    alwNI    ddeI  hinfI
1101 GTTTGAAACTGGAGAACAAGGAGGCAAAGGTCTCGAAGCGGGAGAAGGCGGTGTGGGTGCTGAACCCTGAGGCGGGGATGTGGCAGTGTCTGCTGAGTGA
     CAAACTTTGACCTCTTGTTCCTCCGTTTCCAGAGCTTCGCCCTCTTCCGCCACACCCACGACTTGGGACTCCGCCCCTACACCGTCACAGACGACTCACT
 317   LeuLysLeuGluAsnLysGluAlaLysValSerLysArgGluLysAlaValTrpValLeuAsnProGluAlaGlyMetTrpGlnCysLeuLeuSerAsp
```

FIG. 1B-2

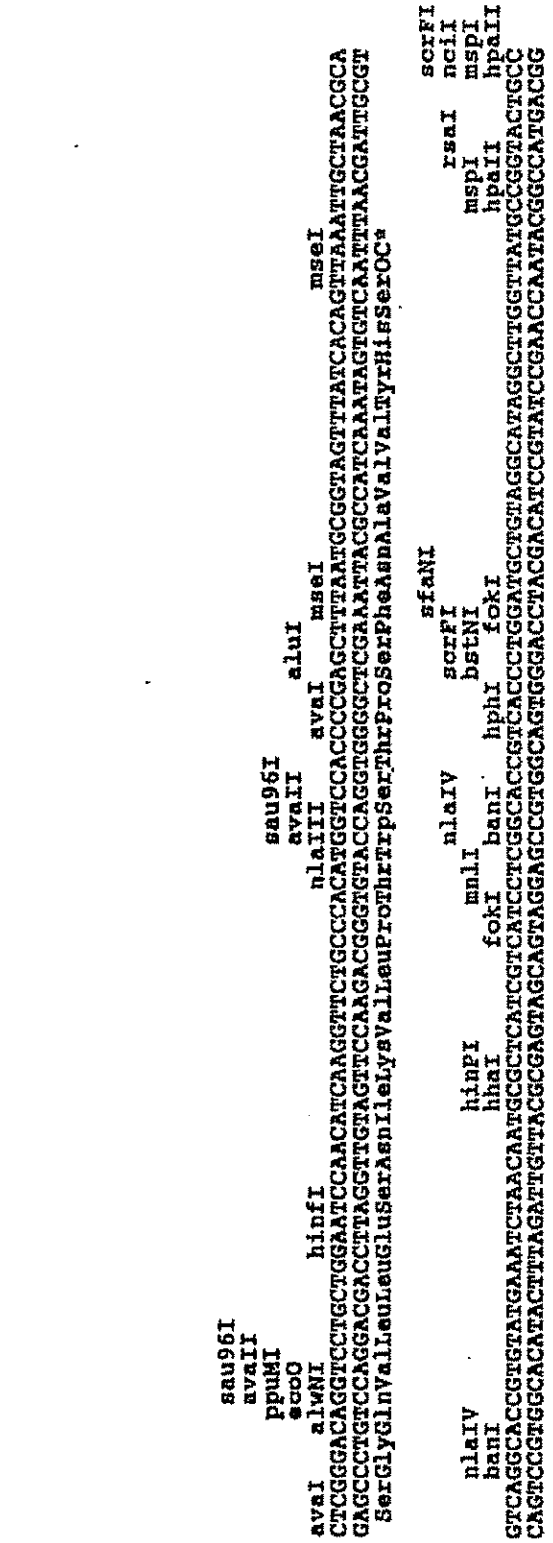

sau96I
avaII
ppuMI
sco0
alwNI
hinfI
nlaIII
sau96I
avaII
aluI
aval
msel
scrFI
ncil
rsaI
mspI
hpaII
sfaNI
nlaIV
scrFI
bstNI
fokI
hinPI
hhaI
mnlI
fokI
banI
hphI
mnlI
banI
haeIII
sau96I
1201 CTCGGGACAGGGTCCTGCTGGAATCCAACATCAAGGTTCTGCCCACATGGTCCACCCCGAGCTTTAATGCGGTAGTTTATCACAGTTAAATTGCTAACGCA
GAGCCCTGTCCCAGGGACGACCTTAGGTTGTAGTTCCAAGACGGGTGTACCAGGTGGGGCTCGAAATTACGCCATCAAATAGTGTCAATTTAACGATTGCGT
350 SerGlyGlnValLeuLeuGluSerAsnIleLysValLeuProThrTrpSerThrProSerPheAsnAlaValLeuTyrHisSerOC*
GTCAGGCACCGTGGTATGAAATCTAACAATGCGCTCATCGTCATCCTCGGCACCGTCACCCTGGATGCTGTAGGCATAGGCTTGGTTATGCCGGTACTGCC
1301
CAGTCCGTGGCACATACTTTAGATTGTTACGCGAGTAGGAGCCGTGGCAGTGGGACCTACGACATCCGTATCCGAACCAATACGGCCATGACGG
1401 GGGCCTCTTGCGGGAT
CCCGGAGAACGCCCTA

FIG. 1C

FIG. 2A

FIG. 2B-1

```
                                                
```

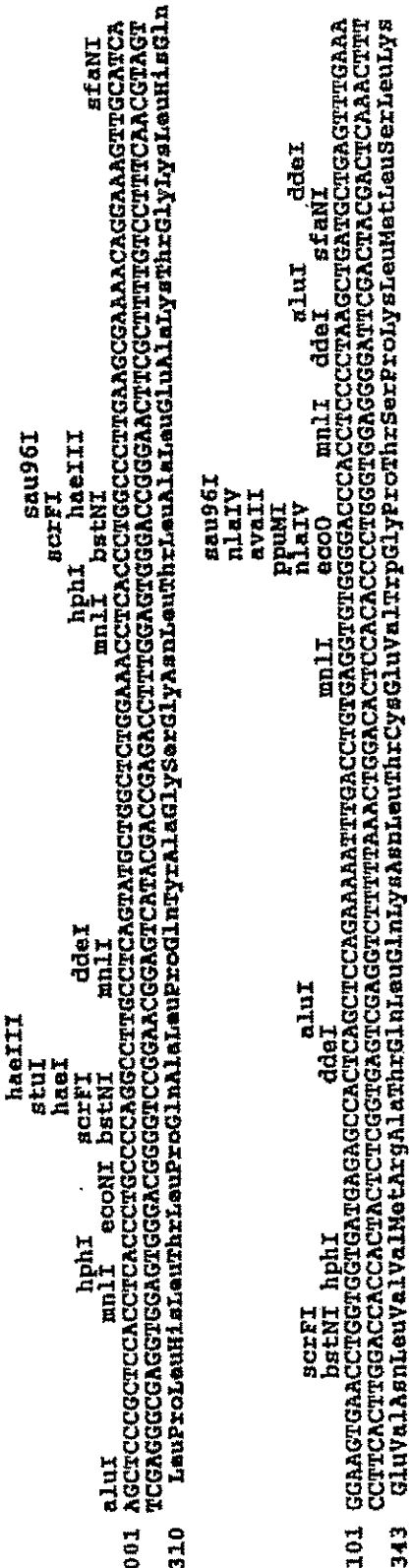
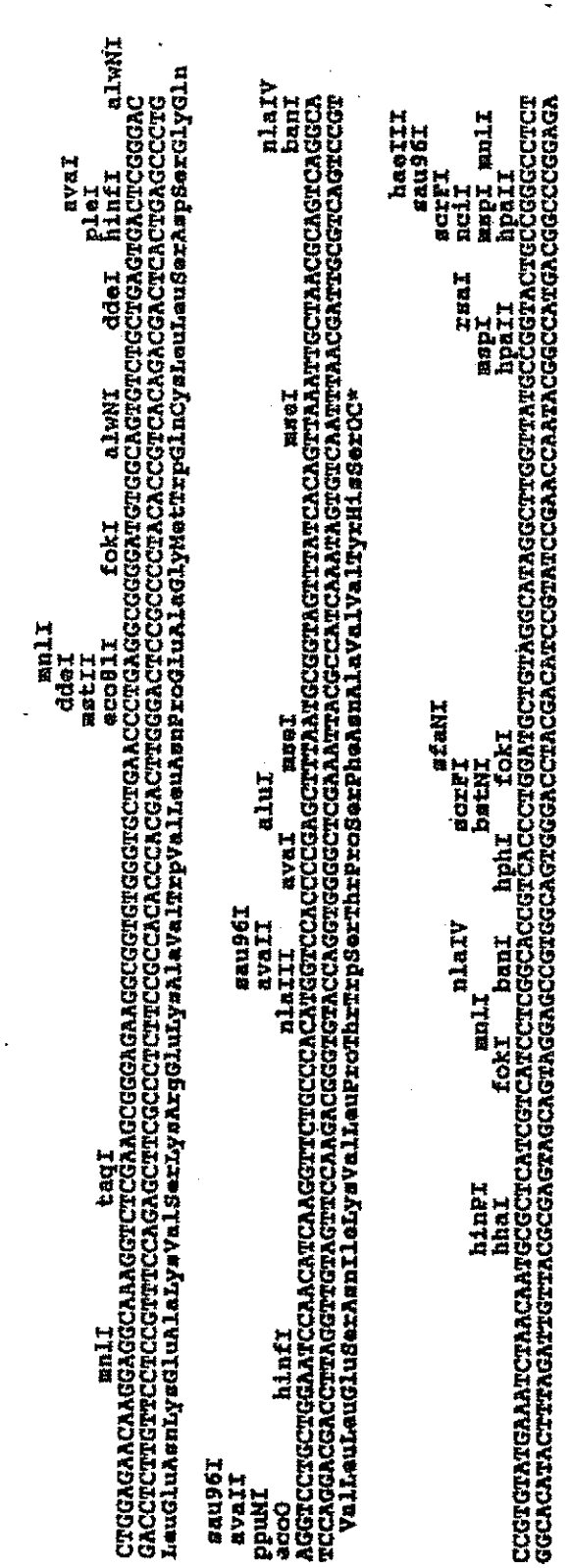

haeIII
stuI
haeI
aluI
hphI
mnlI
ecoNI
bstNI
scrFI
ddeI
mnlI
sau96I
scrFI
hphI
haeIII
mnlI
bstNI
sfaNI
sau96I
nlaIV
avaII
ppuMI
nlaIV
aluI
ddeI
mnlI
ecoO
mnlI
ddeI
sfaNI
aluI
ddeI
scrFI
bstNI
hphI
1001 AGCTCCCGCTCCACCTCACCCTGCCCCCAGGCCTTGCCTCAGTATGCTGGCTCTGGAAACCTCACCCTGGCCCCTTGAAGCGAAAACAGGAAAGTTGCATCA
TCGAGGGCGAGGTGGAGTGGGACGGGGGTCCGGAACGGAGTCATACGACCGAGACCTTTGGAGTGGGACCGGGAACTTCGCTTTTGTCCTTTCAACGTAGT
310 LeuProLeuHisLeuThrLeuProGlnAlaLeuProGlnTyrAlaGlySerGlyAsnLeuThrLeuAlaLeuGluAlaLysThrGlyLysLeuHisGln
343 GluValAsnLeuValValHetArgAlaThrGlnLeuGlnLysAspLeuThrCysGluValTrpGlyProThrSerProLysLeuMetLeuSerLeuLys

FIG. 2B-2



```
                                                                  mnlI
                                                                ddeI                                avaI
                                                               mstII                               pleI
                mnlI        taqI                               eco81I    fokI    alwNI    ddeI  hinfI    alwNI
1201 CTGGAGAACAAGGAGGCAAAGGTCTCGAAGCGGGAGAAGGCGGTGTGGGTGCTGAACCCTGAGGCGGGGATGTGGCAGTGTCTGCTGAGTGACTCGGGAC
     GACCTCTTGTTCCTCCGTTTCCAGAGCTTCGCCCTCTTCCGCCACACCCACGACTTGGGACTCCGCCCCTACACCGTCACAGACGACTCACTGAGCCCTG
 376 LeuGluAsnLysGluAlaLysValSerLysArgGluLysAlaValTrpValLeuAsnProGluAlaGlyMetTrpGlnCysLeuLeuSerAspSerGlyGln

      sau96I
      avaII                                sau96I
     ppuMI                                 avaII      aluI                                                nlaIV
     ecoO       hinfI                    nlaIII    avaI    mseI                    mseI                    banI
1301 AGGTCCTGCTGGAATCCAACATCAAGGTTCTGCCCACATGGTCCACCCCGAGCTTTAATGCGGTAGTTTATCACAGTTAAATTGCTAACGCAGTCAGGCA
     TCCAGGACGACCTTAGGTTGTAGTTCCAAGACGGGTGTACCAGGTGGGGCTCGAAATTACGCCATCAAATAGTGTCAATTTAACGATTGCGTCAGTCCGT
 410   ValLeuLeuGluSerAsnIleLysValLeuProThrTrpSerThrProSerPheAsnAlaValValTyrHisSerOC*

                                                                                                  haeIII
                                                                                                  sau96I
                                                          sfaNI                                  scrFI
                                          nlaIV      scrFI                                rsaI  nciI
                        hinPI        mnlI            bstNI                            mspI     mspI mnlI
                        hhaI       fokI   banI   hphI   fokI                            hpaII    hpaII
1401 CCGTGTATGAAATCTAACAATGCGCTCATCGTCATCCTCGGCACCGTCACCCTGGATGCTGTAGGCATAGGCTTGGTTATGCCGGTACTGCCGGGCCTCT
     GGCACATACTTTAGATTGTTACGCGAGTAGCAGTAGGAGCCGTGGCAGTGGGACCTACGACATCCGTATCCGAACCAATACGGCCATGACGGCCCGGAGA

1501 TGCGGGAT
     ACGCCCTA
```

FIG. 2C

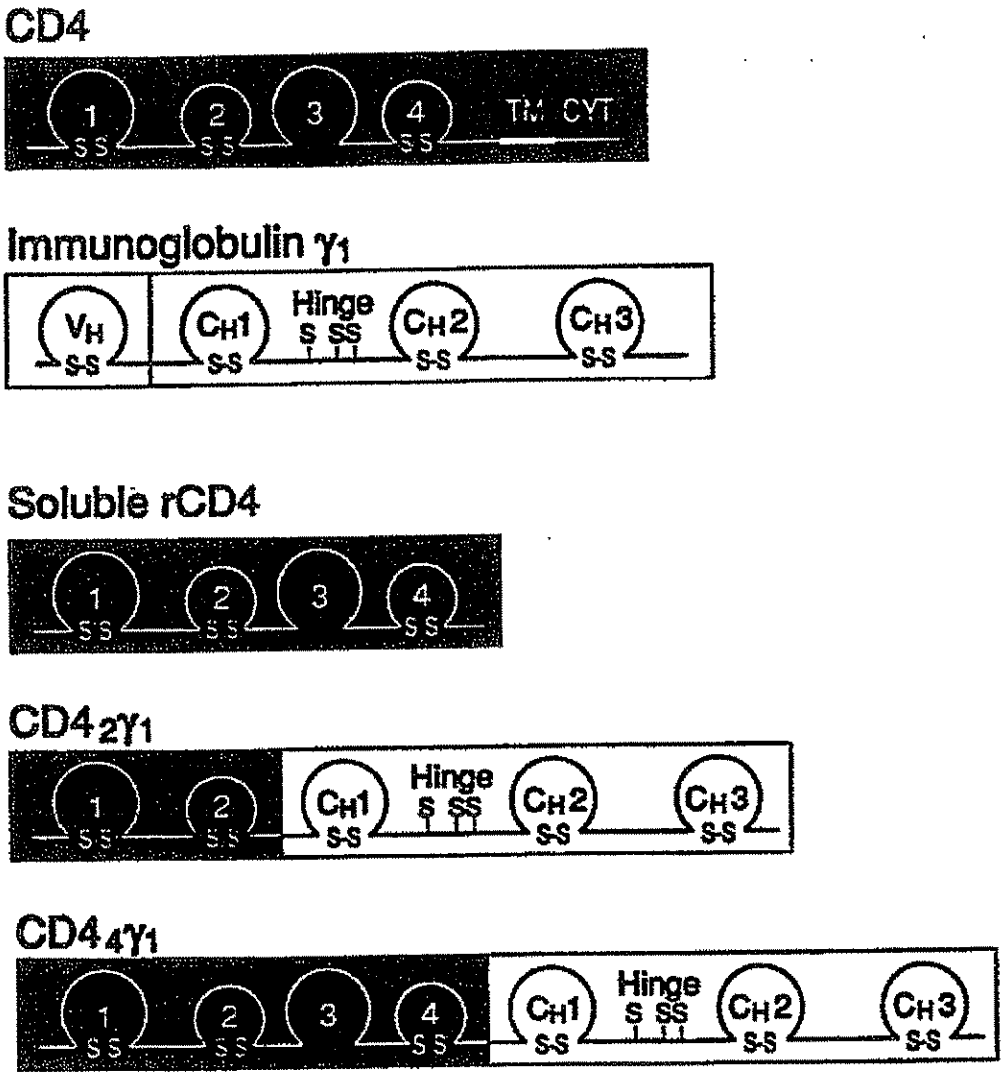

CD4
1
2
3
4
TM
CYT
Immunoglobulin γ1
VH
CH1
Hinge
CH2
CH3
Soluble rCD4
CD4 2γ1
CD4 4γ1

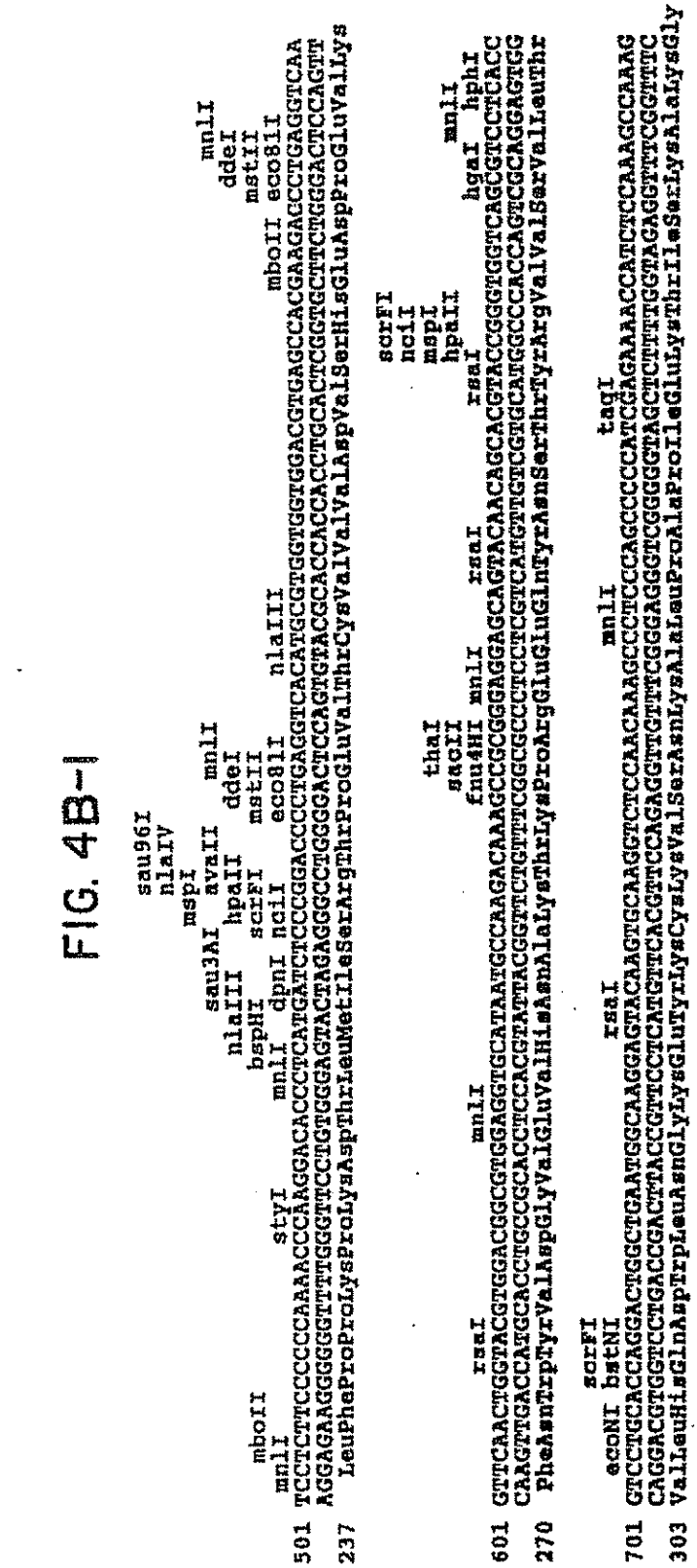
FIG. 3

FIG. 4A

sau96I
nlaIV
eco0
scrFI
bstNI
tagI
sau3AI
dpnI
alwI
hgaI
rsaI
ecoRI
fnu4HI
xmaIII
eaeI
thaI
haeIII
sau96I
nlaIV
bsp1286
hphI
bstEII
apaI
haeIII
eco0
styI
haeIII
mboII
bstNI
scrFI
mnlI
hgiAI
bsp1286
bsp1286
haeIII
mnlI
mnlI
bstNI
mnlI
scrFI
banI
nlaIV
hinPI
nlaIV
narI
haeII
banI
ahaII
ddeI
hhaI
fnu4HI
hgiAI
bsp1286
apaLI
mspI
hpaII
scrFI
ncoI
hphI
mspI
tth111I
hpaII
fnu4HI
ddeI
mstII
mnlI
bbvI
scrFI
bstNI
ecoNI
bstNI
alwNI
bsp1286
banII
nlaIII
bsp1286
bstNI
nlaIV
scrFI
avaII
mboII
styI
1
72
101
103
137
170
201
203
301
401

FIG. 4B

FIG. 4B-1


sau96I
nlaIV
mspI
sau3AI avaII mnlI
nlaIII hpaII ddeI
bspHI scrFI mstII
mboII stvI mnlI dpnI nciI eco81I nlaIII
mboII eco81I
ddeI
mstII
mnlI
scrFI
nciI
mspI
hpaII
thaI
sagII
fnu4HI mnlI rsaI
rsaI hgaI hphI
mnlI
rsaI
mnlI
scrFI
rsaI
mnlI
taqI
ecoNI bstNI
501 TCCTCTTTCCCCCCAAACCCAAGGACACCCTCATGATCTCCCGGACCCCTGAGGTCACATGCGTGGTGGTGGACGTGAGCCACGAAGACCCTGAGGTCAA
237 AGGAGAAAGGGGGGTTTGGGTTCCTGTGGGAGTACTAGAGGGCCTGGGGACTCCAGTGTACGCACCACCTGCACTCGGTGCTTCTGGGACTCCAGTT
LeuPheProProLysProLysAspThrLeuMetIleSerArgThrProGluValThrCysValValValAspValSerHisGluAspProGluValLys
601 GTTCAACTGGTACGTGGACGGCGTGGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAGCAGTACAACAGCACGTACCGGGTGGTCAGCGTCCTCACC
270 PheAsnTrpTyrValAspGlyValGluValHisAsnAlaLysThrLysProArgGluGluGlnTyrAsnSerThrTyrArgValValSerValLeuThr
701 GTCCTGCACCAGGACTGGCTGAATGGCAAGGAGTACAAGTGCAAGGTCTCCAACAAAGCCCTCCCAGCCCCCATCGAGAAAACCATCTCCAAAGCCAAAG
303 ValLeuHisGlnAspTrpLeuAsnGlyLysGluTyrLysCysLysValSerAsnLysAlaLeuProAlaProIleGluLysThrIleSerLysAlaLysGly

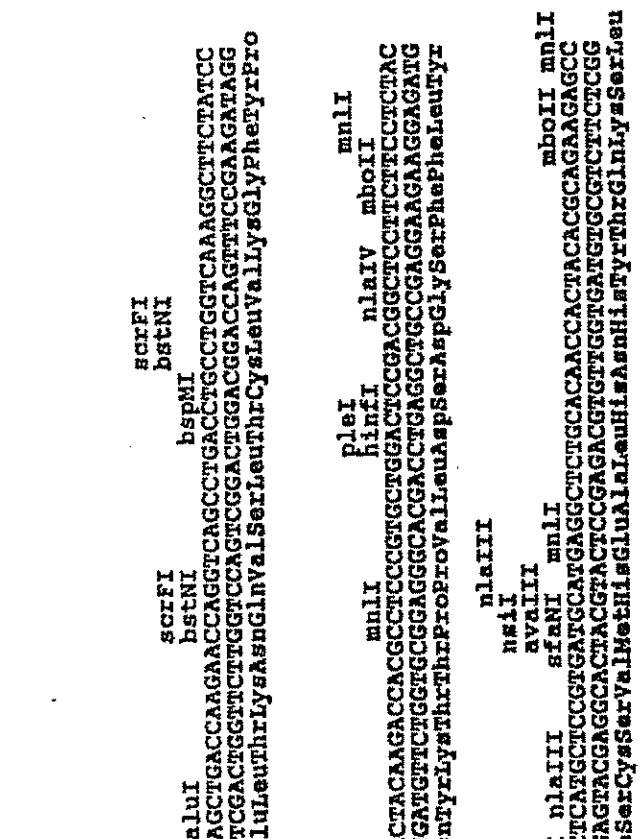

scrFI
ncil
mspI
hpaII
smaI
scrFI
ncil
aval
scrFI
bstNI
fnu4HI
bbvI aval
fokI fokI alul
bspMI
bstNI
rsaI
801 GGCAGCCCCCGAGAACCACAGGTGTACACCCTGCCCCCATCCCGGGATGAGCCTGACCAAGAACCAGGTCAGCCTGACCTGCCTGGTCAAAGGCTTCTATCC
CCGTCGGGGGCTCTTGGTGTCCACATGTGGGACGGGGGTAGGGCCCTACTCGACTGGTTCTTGGTCCAGTCGGACTGGACGGACCAGTTTCCGAAGATAGG
337 GlnProArgGluProGlnValTyrThrLeuProProSerArgAspGluLeuThrLysAsnGlnValSerLeuThrCysLeuValLysGlyPheTyrPro
mspI
hpaII
fnu4HI
bbvI
hinfI
pleI
nlaIV
mboII
mnlI
901 CAGCGACATCGCCGTGGAGTGGGAGAGCAATGGGCAGCCGGAGAACAACTACAAGACCACGCCTCCCGTGCTGGACTCCGACGGCTCCTTCTTCCTCTAC
GTCGCTGTAGCGGCACCTCACCCTCTCGTTACCCGTCGGCCTCTTGTTGATGTTCTGGTGCGGAGGGCACGACCTGAGGCTGCCGAAGGAAGAGGAGATG
370 SerAspIleAlaValGluTrpGluSerAsnGlyGlnProGluAsnAsnTyrLysThrThrProProValLeuAspSerAspGlySerPhePheLeuTyr
nlaIII
nsiI
avaIII
sfaNI mnlI
fnu4HI
bbvI xmnI mboII nlaIII
bspMI
mboII mnlI
1001 AGCAAGCTCACCGTGGACAAGAGCAGGTGGCAGCAGGGGAACGTCTTCTCATGCTCCGTGATGCATGAGGCTCTGCACAACCACTACACGCAGAAGAGCC
TCGTTCGAGTGGCACCTGTTCTCGTCCACCGTCGTCCCCTTGCAGAAGAGTACGAGGCACTACGTACTCCGAGACGTGTTGGTGATGTGCGTCTTCTCGG
403 SerLysLeuThrValAspLysSerArgTrpGlnGlnGlyAsnValPheSerCysSerValMetHisGluAlaLeuHisAsnHisTyrThrGlnLysSerLeu
mspI
hpaII
aluI
hphI
scrFI
ncil
mspI
hpaII
haeIII
xmaIII
eaeI
1101 TCTCCCTGTCTCCGGGTAAATGAGTGCGACGGCCG
AGAGGGACAGAGGCCCATTTACTCACGCTGCCGGC
437 SerLeuSerProGlyLysOp*

FIG. 4B-2

```
                                fnu4HI                                                                  sau96I
                               pstI                                                                     avaII
                            haeIII                                                                     nlaIV
    ecoRI     hphI      fnu4HI bbvI   mboII                                                          mnlI  styI
  1 GAATTCACTCTCACCATCAGCGGCCTGCAGCCTGAAGATTTTGCAACTTATTACTGCCAACAGTATAAGAGTTTGTCGCTCACTTTCGGCGGAGGGACCA
    CTTAAGTGAGAGTGGTAGTCGCCGGACGTCGGACTTCTAAAACGTTGAATAATGACGGTTGTCATATTCTCAAACAGCGAGTGAAAGCCGCCTCCCTGGT
 72       ThrLeuThrIleSerGlyLeuGlnProGluAspPheAlaThrTyrTyrCysGlnGlnTyrLysSerLeuSerLeuThrPheGlyGlyGlyThrLys

          sau3AI          fnu4HI
          dpnI            bbvI         mboII mboII                                  mnlI
101 AGGTGGAGATCAAACGAACTGTGGCTGCACCATCTGTCTTCATCTTCCCGCCATCTGATGAGCAGTTGAAATCTGGAACTGCCTCTGTTGTGTGCCTGCT
    TCCACCTCTAGTTTGCTTGACACCGACGTGGTAGACAGAAGTAGAAGGGCGGTAGACTACTCGTCAACTTTAGACCTTGACGGAGACAACACACGGACGA
104   ValGluIleLysArgThrValAlaAlaProSerValPheIlePheProProSerAspGluGlnLeuLysSerGlyThrAlaSerValValCysLeuLeu
          ← VKJK CK →
                        haeIII
                       haeI                                                  scrFI
    xmnI               mnlI     rsaI                    mnlI                  bstNI
201 GAATAACTTCTATCCCAGAGAGGCCAAAGTACAGTGGAAGGTGGATAACGCCCTCCAATCGGGTAACTCCCAGGAGAGTGTCACAGAGCAGGACAGCAAG
    CTTATTGAAGATAGGGTCTCTCCGGTTTCATGTCACCTTCCACCTATTGCGGGAGGTTAGCCCATTGAGGGTCCTCTCACAGTGTCTCGTCCTGTCGTTC
137  AsnAsnPheTyrProArgGluAlaLysValGlnTrpLysValAspAsnAlaLeuGlnSerGlyAsnSerGlnGluSerValThrGluGlnAspSerLys

                                                                                                      sacI
                                                                                                      hgiAI
                                                                                                      bsp1286
                                                                                                  haeIII
                       fnu4HI                                                                    sau96I aluI
                   ddeI              ddeI                                                       ecoO   banII
                  mnlI bbvI      hgaI                                  accI            hphI    alwNI ddeI
301 GACAGCACCTACAGCCTCAGCAGCACCCTGACGCTGAGCAAAGCAGACTACGAGAAACACAAAGTCTACGCCTGCGAAGTCACCCATCAGGGCCTGAGCT
    CTGTCGTGGATGTCGGAGTCGTCGTGGGACTGCGACTCGTTTCGTCTGATGCTCTTTGTGTTTCAGATGCGGACGCTTCAGTGGGTAGTCCCGGACTCGA
170 AspSerThrTyrSerLeuSerSerThrLeuThrLeuSerLysAlaAspTyrGluLysHisLysValTyrAlaCysGluValThrHisGlnGlyLeuSerSer

                                                                    ddeI
                  aluI                    mnlI     bsp1286 bspMI  mnlI
401 CGCCCGTCACAAAGAGCTTCAACAGGGGAGAGTGTTAGAGGGAGAAGTGCCCCCACCTGCTCCTCAGT
    GCGGGCAGTGTTTCTCGAAGTTGTCCCCTCTCACAATCTCCCTCTTCACGGGGGTGGACGAGGAGTCA
204   ProValThrLysSerPheAsnArgGlyGluCysAM*
```

FIG. 5

CD4 ADHESON VARIANTS

This is a continuation of co-pending application Ser. No. 07/842,777 filed on Feb. 18, 1992, now abandoned, which is a continuation of application Ser. No. 07/250,785 filed on Sep. 28, 1988, now abandoned.

BACKGROUND OF THE INVENTION

This application relates to compositions for antiviral or immunomodulatory therapy.

The primary immunologic abnormality resulting from infection by HIV is the progressive depletion and functional impairment of T lymphocytes expressing the CD4 cell surface glycoprotein (H. Lane et al., Ann. Rev. Immunol. 3:477 [1985]). CD4 is a non-polymorphic glycoprotein with homology to immunoglobulin gene superfamily (P. Maddon et al., Cell 42:93 [1985]). Together with the CD8 surface antigen, CD4 defines two distinct subsets of mature peripheral T cells (E. Reinherz et al., Cell 19:821 [1980]), which are distinguished by their ability to interact with nominal antigen targets in the context of class I and class II major histocompatibility complex (MHC) antigens, respectively (S. Swain, Proc. Natl. Acad. Sci. 78:7101 [1981]; E. Engleman et al., J. Immunol. 127:2124 [1981]; H. Spitz et al., J. Immunol. 129:1563 [1982]; W. Biddison et al., J. Exp. Med. 156:1065 [1982]; and D. Wilde et al., J. Immunol. 131:2178 [1983]). For the most part, CD4 T cells display the helper/inducer T cell phenotype (E. Reinherz, supra), although CD4 T cells characterized as cytotoxic/-suppressor T cells have also been identified (Y. Thomas et al., J. Exp. Med. 154:459 [1981]; S. Meuer et al., Proc. Natl. Acad. Sci. USA 79:4395 [1982]; and A. Krensky et al., Proc. Natl. Acad. Sci. USA 79:2365 [1982]). The loss of CD4 helper/inducer T cell function probably underlies the profound defects in cellular and humoral immunity leading to the opportunistic infections and malignancies characteristic of the acquired immunodeficiency syndrome (AIDS) (H. Lane supra).

Studies of HIV-I infection of fractionated CD4 and CD8 T cells from normal donors and AIDS patients have revealed that depletion of CD4 T cells results from the ability of HIV-I to selectively infect, replicate in, and ultimately destroy this T lymphocyte subset (D. Klatzmann et al., Science 225:59 [1984]). The possibility that CD4 itself is an essential component of the cellular receptor for HIV-I was first indicated by the observation that monoclonal antibodies directed against CD4 block HIV-I infection and syncytia induction (A. Dalgleish et al., Nature [London]312:767 [1984]; J. McDougal et al., J. Inunmol. 135:3151 [1985]). This hypothesis has been confirmed by the demonstration that a molecular complex forms between CD4 and gp120, the major envelope glycoprotein of HIV-I (J. McDougal et al., Science 231:382 [1986]; and the finding that HIV-I tropism can be conferred upon ordinarily non-permissive human cells following the stable expression of a CD4 cDNA (P. Maddon et al., Cell 47:333 [1986]). Furthermore, the neurotropic properties of HIV-I, reflected by a high incidence of central nervous system dysfunction in HIV-I infected individuals (W. Snider et al., Ann. Neurol. 14:403 [1983]), and the ability to detect HIV-I in the brain tissue and cerebrospinal fluid of AIDS patients (G. Shaw et al., Science 227:177 [1985]; L. Epstein, AIDS Res. 1:447 [1985]; S. Koenig, Science 233:1089 [1986]; D. Ho et al., N. Engl. J. Med. 313:1498 [1985]; J. Levy et al., Lancet II:586 [1985]), appears to have its explanation in the expression of CD4 in cells of neuronal, glial and monocyte/macrophage origin (P. Maddon, Cell 47:444 [1986]; I. Funke et al., J. Exp. Med. 165:1230 [1986]; B. Tourvieille et al., Science 234:610 [1986]).

In addition to determining the susceptibility to HIV-I infection, the manifestation of cytopathic effects in the infected host cell appears to involve CD4. Antibody to CD4 was found to inhibit the fusion of uninfected CD4 T cells with HIV-I infected cells in vitro; moreover, the giant multinucleated cells produced by this event die shortly after being formed resulting in the depletion of the population of CD4 cells (J. Lifson et al., Science 232:1123 [1986]). Formation of syncytia also requires gp120 expression, and can be elicited by coculturing CD4-positive cell lines with cell lines expressing the HIV-I env gene in the absence of other viral structural or regulatory proteins (J. Sodroski et al., Nature 322:470 [1986]; J. Lifson et al., Nature 323:725 [1986]). Thus, in mediating both the initial infection by HIV-I as well as eventual cell death, the interaction between gp120 and CD4 constitutes one of several critical entry points in the viral life cycle amenable to therapeutic intervention (H. Mitsuya et al., Nature 325:773 [1987]).

The known sequence of the CD4 precursor predicts a hydrophobic signal peptide, an extracellular region of approximately 370 amino acids, a highly hydrophobic stretch with significant identity to the membrane-spanning domain of the class II MHC beta chain, and a highly charged intracellular sequence of 40 residues (P. Madden, Cell 42:93 [1985]). The extracellular domain of CD4 consists of four contiguous regions each having amino acid and structural similarity to the variable and joining (V-J) domains of immunoglobulin light chains as well as related regions in other members of the immunoglobulin gene superfamily (a subclass of which are defined herein by the coined term "adhesons". These structurally similar regions of CD4 are termed the V_1, V_2, V_3 and V_4 domains (denominated 1–4 in FIG. 3).

A successful strategy in the development of drugs for the treatment of many receptor mediated abnormalities has been the identification of antagonists which block binding of the natural ligand. Since the CD4 adheson ordinarily binds to the recognition sites of the HIV envelope it would appear to be a candidate for therapeutically sequestering these HIV sites, thereby blocking viral infectivity. However, full length CD4 and other adhesons are cell membrane proteins which are anchored in the lipid bilayer of cells. The presence of membrane components will be undesirable from the standpoint of manufacturing and purification. In addition, since adhesons are normally present only on cell surfaces, it would be desirable to produce adhesons in a form which is more stable in the circulation. Additionally, even truncated, soluble CD4 adheson (generally referred to as CD4T) may not be optimally effective as a therapeutic since it possesses a relatively short biological half-life, binds to HIV no better than cell surface CD4, may not cross the placental or other biological barriers and since it merely sequesters the HIV recognition sites without in itself bearing an infected-cell killing or virus killing functionality.

Accordingly, it is an object of this invention to produce soluble, secreted adhesons. It is another object to produce CD4 derivatives. Still another object is to prepare adhesons fused to other polypeptides in order to provide molecules with novel functionalities such as diagnostic reagents for the in vitro assay of adhesons or

their ligands. In particular, it is an objective to prepare molecules for directing toxins or effector molecules (for example the Fc domain of immunoglobulin) to cells bearing receptors for the adhesons, e.g. HIV gp120 in the case of CD4, and for use in facilitating purification of the adhesons. It is a further object to provide stable, highly purified adheson preparations.

SUMMARY

The objects of this invention are accomplished by providing nucleic acid encoding an amino acid sequence variant of an adheson, in particular a variant in which the trans-membrane domain is modified so that it is no longer capable of becoming lodged in the cell membrane. In the case of CD4 such variants are termed soluble CD4.

Variant adhesons are produced by a method comprising (a) transforming a host cell with nucleic acid encoding an amino acid sequence variant of an adheson, (b) culturing the host cell and (c) recovering the variant adheson from the host cell culture media or from lysates of the host cell.

In specific embodiments, the objects of this invention are accomplished by providing an adheson variant selected from the group consisting of (a) an adheson amino acid sequence variant having an inactivated transmembrane domain and (b) a polypeptide comprising an adheson extracellular domain fused to the sequence of a polypeptide which is different from the adheson, this latter, for example, selected from a cytotoxin, an immunogen or a protein with a long plasma half life such as an immunoglobulin constant domain.

In a preferred embodiment a polypeptide comprising a gp120 binding domain of the CD4 adheson is fused at its C-terminus to an immunoglobulin constant domain, or is linked to a cytotoxic polypeptide such as ricin.

The CD4 adheson variants provided herein are purified and formulated in pharmacologically acceptable vehicles for administration to patients.

BRIEF DESCRIPTION OF THE DRAWINGS

FIGS. 1A, 1B-1, 1B-2 and 1C depict the amino acid and nucleotide sequence of a secreted form of the CD4 adheson. The signal processing site is designated with an arrow.

FIGS. 2A, 2B-1, 2B-2 and 2C depict the amino acid and nucleotide sequence of a fusion of the herpes gD leader and N-terminal 27 residues to the putative mature N-terminus of CD4T.

FIG. 3 depicts the structural elements of the native and soluble CD4 adheson, the native human IgG_1 ($\gamma 1$) heavy chain and two exemplary heavy chain-CD4 chimeras.

FIGS. 4A, 4B-1 and 4B-2 are a map of the linkered human IgG_1 ($\gamma 1$) chain fragment employed in the preparation of CD4 fusions. Insert sites are designated $\gamma 1$ and Fc.

FIG. 5 is a map of the human κ light chain fragment useful for CD4 fusions at the arrow flanked by $V_\kappa J_\kappa$(light variable and joining) and C_κ(light constant).

DETAILED DESCRIPTION

Adhesons are cell surface polypeptides having an extracellular domain which is homologous to a member of the immunoglobulin gene superfamily, excluding, however, highly polymorphic members of this superfamily selected from the group of class I and class II major histocompatibility antigens, immunoglobulins and T-cell receptor α, β, γ and δ chains. Examples of adhesons include CD1, CD2, CD4, CD8, CD28, the γ, δ and ϵ chains of CD3, OX-2, Thy-1, the intercellular or neural cell adhesion molecules (I-CAM or N-CAM), lymphocyte function associated antigen-3 (LFA-3), neurocytoplasmic protein (NCP-3), poly-Ig receptor, myelin-associated glycoprotein (FLAG), high affinity IgE receptor, the major glycoprotein of peripheral myelin (Po), platelet derived growth factor receptor, colony stimulating factor-1 receptor, macrophage Fc receptor, Fc gamma receptors and carcinoembryonic antigen. Homologous as defined herein means having the sequence of a member of the immunoglobulin gene superfamily or having a sequence therewithin which has substantially the same as (or a greater degree of) amino acid sequence homology to a known member of the superfamily as the specific examples given above have to the sequence of an immunoglobulin variable or constant domain. Preferred adhesons are CD4, CD8 and high affinity IgE Fc receptor.

This invention is particularly concerned with amino acid sequence variants of adhesons. Amino acid sequence variants of adhesons are prepared with various objectives in mind, including increasing the affinity of the adheson for its binding partner, facilitating the stability, purification and preparation of the adheson, increasing its plasma half life, introducing additional functionalities and lessening the severity or occurrence of side effects during therapeutic use of the adheson. Amino acid sequence variants of adhesons fall into one or a combination of the following classes: insertional, substitutional or deletional variants.

Insertional amino acid sequence variants are those in which one or more amino acid residues extraneous to the adheson are introduced into a predetermined site in the adheson including the C or N termini. Such variants are referred to as fusions of the adheson and a different polypeptide. Such other polypeptides contain sequences other than those which are normally found in the adheson at the inserted position. Several groups of fusions are contemplated herein. Immunologically active adheson fusions comprise an adheson and a polypeptide containing a non-adheson epitope. The non-adheson epitope is any immunologically competent polypeptide, i.e., any polypeptide which is capable of eliciting an immune response in the animal to which the fusion is to be administered or which is capable of being bound by an antibody raised against the non-adheson polypeptide. Typical non-adheson epitopes will be those which are borne by allergens, autoimmune epitopes, or other potent immunogens or antigens recognized by pre-existing antibodies in the fusion recipient, including bacterial polypeptides such as trpLE, beta-galactosidase, viral polypeptides such as herpes gD protein, and the like. Immunogenic fusions are produced by cross-linking in vitro or by recombinant cell culture transformed with DNA encoding an immunogenic polypeptide. It is preferable that the immunogenic fusion be one in which the immunogenic sequence is joined to or inserted into the adheson antigen or fragment thereof by a peptide bond(s). These products therefore consist of a linear polypeptide chain containing adheson epitopes and at least one epitope foreign to the adheson. It will be understood that it is within the scope of this invention to introduce the epitopes anywhere within the adheson molecule or fragment thereof. Such fusions are conveniently made in recombinant host cells or by the use of bifunctional cross-linking agents. The use of a cross-

linking agent to fuse the adheson to the immunogenic polypeptide is not as desirable as a linear fusion because the cross-linked products are not as easily synthesized in structurally homogeneous form.

These immunogenic insertions are particularly useful when formulated into a pharmacologically acceptable carrier and administered to a subject in order to raise antibodies against the adheson, which antibodies in turn are useful in diagnostics or in purification of adheson by immunoaffinity techniques known per se. Alternatively, in the purification of adhesons, binding partners for the fused non-adheson polypeptide, e.g. antibodies, receptors or ligands, are used to adsorb the fusion from impure admixtures, after which the fusion is eluted and, if desired, the adheson is recovered from the fusion, e.g. by enzymatic cleavage.

Other fusions, which may or may not also be immunologically active, include fusions of the adheson sequence with a signal sequence heterologous to the adheson, fusions of transmembrane-modified CD4 adhesons, for example, to polypeptides having enhanced plasma half life (ordinarily >about 20 hours) such as immunoglobulin chains or fragments thereof, and fusions with cytotoxic functionalities. Signal sequence fusions are employed in order to more expeditiously direct the secretion of the adheson. The heterologous signal replaces the native adheson signal, and when the resulting fusion is recognized, i.e. processed and cleaved by the host cell, the adheson is secreted. Signals are selected based on the intended host cell, and may include bacterial yeast, mammalian and vital sequences. The herpes gD glycoprotein signal is suitable for use in mammalian expression systems.

Plasma proteins which have enhanced plasma half-life longer than that of transmembrane modified CD4 include serum albumin, immunoglobulins, apolipoproteins, and transferrin. Preferably, the adheson-plasma protein fusion is not significantly immunogenic in the animal in which it is used and the plasma protein does not cause undesirable side effects in patients by virtue of its normal biological activity.

In a specific embodiment the adheson immunoglobulin-like domain which may be homologous either to the constant or to the variable region domains is conjugated with an immunoglobulin constant region sequence. The resulting products are referred to herein as immunoadhesons. Immunoglobulins and certain variants thereof are known and many have been prepared in recombinant cell culture. For example, see U.S. Pat. No. 4,745,055; EP 256,654; Faulkner et al., Nature 298:286 (1982); EP 120,694; EP 125,023; Morrison, J. Immun. 123:793 (1979); Kohler et al., P.N.A.S. USA 77:2197 (1980); Raso et al., Cancer Res. 41:2073 (1981); Morrison et al., Ann. Rev. Immunol. 2:239 (1984); Morrison, Science 229:1202 (1985); Morrison et al., P.N.A.S. USA 81:6851 (1984); EP 255,694; EP 266,663; and WO 88/03559. Reassorted immunoglobulin chains also are known. See for example U.S. Pat. No. 4,444,878; WO 88/03565; and EP 68,763 and references cited therein.

Ordinarily, the domains of adhesons that are homologous to immunoglobulins and extracellular in their native environment are fused C-terminally to the N-terminus of the constant region of immunoglobulins in place of the variable region(s) thereof, retaining at least functionally active hinge, CH2 and CH3 domains of the constant region of an immunoglobulin heavy chain. This ordinarily is accomplished by constructing the appropriate DNA sequence and expressing it in recombinant cell culture. Immunoglobulins and other polypeptides having enhanced plasma half life are fused to the extracellular or ligand. binding domains of other adhesons in the same fashion.

The boundary domains for the CD4 V-like regions (V1–V4) are, respectively, about 100–109, about 175–184, about 289–298, and about 360–369 (based on the precursor CD4 ,amino acid sequence in which the initiating met is -25; FIG. 1*a*). CD4 sequences containing any of the CD4 V domains are fused to the immunoglobulin sequence. It is preferable that the V1V2 or V1V2V3V4 be fused at their C-termini to the immunoglobulin constant region. The precise site at which the fusion is made is not critical; the boundary domains noted herein are for guidance only and other sites neighboring or within the V regions may be selected in order to optimize the secretion or binding characteristics of the CD4. The optimal site will be determined by routine experimentation. In general, it has been found that the fusions are expressed intracellularly, but a great deal of variation is encountered in the degree of secretion of the fusions from recombinant hosts. For instance, the following table demonstrates the various immunoglobulin fusions that have been obtained by the method of this invention. In all examples of CD4 immunoadhesons, the CD4 signal was used to direct secretion from 293 cells. Lower case m represents murine origin, while the lower case h designates human origin. V and C are abbreviations for immunoglobulin variable and constant domains respectively. The numerical subscripts indicate the number of parenthetical units found in the designated multimer. It will be understood that the chains of the multimers are believed to be disulfide bonded in the same fashion as native immunoglobulins. The CD4 immunoadhesons typically contained either the first N-terminal 366 residues of CD4 ($CD4_4$) or the first 180 N-terminal residues of CD4 ($CD4_2$) linked at their C-terminus to the κ (light) chain or IgG1 heavy chain constant region ($\gamma 1$).

TABLE I

Transfected Gene	Secreted Product
$mV_\kappa C_\kappa$	$mV_\kappa C_\kappa$ and/or $(mV_\kappa C_\kappa)_2$
$mV_{\gamma 1} C_{\gamma 1}$	ND
$mV_\kappa C_\kappa + mV_{\gamma 1} C_{\gamma 1}$	$(mV_\kappa C_\kappa)_2 (mV_{\gamma 1} C_{\gamma 1})_2$ + $mV_\kappa C_\kappa$ and/or $(mV_\kappa C_\kappa)_2$
$hCD4\text{-}mC_\kappa$	$hCD4\text{-}mC_\kappa$ and/or $(hCD4\text{-}mC_\kappa)_2$
$hCD4\text{-}mC_{\gamma 1}$	ND
$hCD4\text{-}mC_\kappa + hCD4\text{-}mC_{\gamma 1}$	$(hCD4\text{-}mC_\kappa)_2 (hCD4\text{-}mC_{\gamma 1})_2$ + $hCD4\text{-}mC_\kappa$ and/or $(hCD4\text{-}mC_\kappa)_2$
$hCD4\text{-}hC_\kappa$	$hCD4\text{-}hC_\kappa$ and/or $(hCD4\text{-}hC_\kappa)_2$
$hCD4\text{-}hC_{\gamma 1}$	$(hCD4\text{-}hC_{\gamma 1})_2$
$hCD4\text{-}hC_\kappa + hCD4\text{-}hC_{\gamma 1}$	$(hCD4\text{-}hC_\kappa)_2 (hCD4\text{-}hC_{\gamma 1})_2$ + $hCD4\text{-}hC_\kappa$ and/or $(hCD4\text{-}hC_\kappa)_2$
$mV_\kappa C_\kappa + hCD4\text{-}hC_{\gamma 1}$	$(mV_\kappa C_\kappa)_2 (hCD4\text{-}hC_{\gamma 1})_2$ + $mV_\kappa C_\kappa$ and/or $(mV_\kappa C_\kappa)_2$

*ND = Not detected

It is interesting to observe from this table that the CD4-human heavy chain immunoadheson was secreted as a dimer whereas the analogous murine construction was not detected (this not excluding the intracellular accumulation of the protein, however). The ability of the hCD4-hCγ1 transformants to produce heavy chain dimer was unexpected since previous work had suggested that immunoglobulin heavy chains are not secreted unless the hosts are cotransformed with nucleic acid encoding both heavy and light chain (Valle et al., Nature 241:338 [1981]). According to this invention,

CD4-IgG immunoadhesion chimeras are readily secreted wherein the CD4 epitope is present in heavy chain dimers, light chain monomers or dimers, and heavy and light chain heterotetramers wherein the CD4 epitope is present fused to one or more light or heavy chains, including heterotetramers wherein up to and including all four variable region analogues are derived from CD4. Where light-heavy chain non-CD4 variable domain is present, a heterofunctional antibody thus is provided.

Various exemplary hetero-and chimeric immunoadhesion antibodies produced in accordance with this invention are schematically diagrammed below. "A" means at least a portion of the extracellular domain of an adhesion containing its ligand binding site; V_L, V_H, C_L and C_H represent light or heavy chain variable or constant domains of an immunoglobulin; n is an integer; and Y designates a covalent cross-linking moiety.

(a) AC_L;

(b) AC_L-AC_L;

(c) AC_H-[AC_H, AC_L-AC_H, AC_L-V_HC_H, V_LC_L-AC_H, or V_LC_L-V_HC_H];

(d) AC_L-AC_H-[AC_H, AC_L-AC_H, AC_L-V_HC_H, V_LC_L-AC_H, or V_LC_L-V_HC_H];

(e) AC_L-V_HC_H-[AC_H, AC_L-AC_H, AC_L-V_HC_H, V_LC_L-AC_H, or V_LC_L-V_HC_H];

(f) V_LC_L-AC_H-[AC_H, AC_L-AC_H, V_LC_L-AC_H, or V_LC_L-V_HC_H]; or

(g) [A-Y]$_n$-[V_LC_L-V_HC_H]$_2$.

The structures shown in this table show only key features, e.g. they do not show joining (J) or other domains of the immunoglobulins, nor are disulfide bonds shown. These are omitted in the interests of brevity. However, where such domains are required for binding activity they shall be construed as being present in the ordinary locations which they occupy in the adheson, immunoadheson or immunoglobulin molecules as the case may be. These examples are representative of divalent antibodies; more complex structures would result by employing immunoglobulin heavy chain sequences from other classes, e.g. IgM. The immunoglobulin V_LV_H antibody combining site also designated as the companion immunoglobulin, preferably is capable of binding to a predetermined antigen.

Suitable companion immunoglobulin combining sites and fusion partners are obtained from IgG-1, -2, -3, or -4 subtypes, IgA, IgE, IgD or IgM, but preferably IgG-1.

A preferred embodiment is a fusion of an N-terminal portion of CD4, which contains the binding site for the gp120 envelope protein of HIV, to the C-terminal F_c portion of an antibody, containing the effector functions of immunoglobulin G_1. There are two preferred embodiments of this sort; in one, the entire heavy chain constant region is fused to a portion of CD4; in another, a sequence beginning in the hinge region just upstream of the papain cleavage site which defines IgG F_c chemically (residue 216, taking the first residue of heavy chain constant region to be 114 [Kobat et al., "Sequences of Proteins of Immunological Interest" 4th Ed., 1987], or analogous sites of other immunoglobulins) is fused to a portion of CD4. These embodiments are described in the examples.

More particularly, those variants in which one or more immunoglobulin-like domains of an adheson are substituted for the variable region of an immunoglobulin chain are believed to exhibit improved in vivo plasma half life. These chimeras are constructed in a fashion similar to chimeric antibodies in which a variable domain from an antibody of one species is substituted for the variable domain of another species. See, for example, EP 0 125 023; Munro, Nature 312: (13 Dec. 1984); Neuberger et al., Nature 312: (13 Dec. 1984); Sharon et al., Nature 309: (24 May 1984); Morrison et al., Proc. Natl. Acad. Sci. USA 81:6851–6855 (1984); Morrison et al., Science 229:1202–1207 (1985); and Boulianne et al., Nature 312:643–646 (13 Dec. 1984). The DNA encoding the adheson immunoglobulin-like domain(s) is cleaved by a restriction enzyme at or proximal to the 3' end of the DNA encoding the immunoglobulin-like domain(s) and at a point at or near the DNA encoding the N-terminal end of the mature adheson polypspride (where use of a different leader is contemplated) or at or proximal to the N-terminal coding region for the adheson (where the native adheson signal is employed). This DNA fragment then is readily inserted into DNA encoding an immunoglobulin light or heavy chain constant region and, if necessary, tailored by deletional mutagenesis. DNA encoding immunoglobulin light or heavy chain constant regions is known or readily available from cDNA libraries or is synthesized. See for example, Adams et al., Biochemistry 19:2711–2719 (1980); Gough et al., Biochemistry 19:2702–2710 (1980); Dolby etal., P.N.A.S. USA, 77:6027–6031 (1980); Rice etal., P.N.A.S. USA 79:7862–7865 (1982); Falkner et. al., Nature 298:286–288 (1982); and Morrison et al., Ann. Rev. Immunol. 2:239–256 (1984).

DNA encoding the immunoglobulin or immunoadheson chimeric chain(s) is transfected into a host cell for expression. If the host cell is producing an immunoglobulin prior to transfection then one need only transfect with the adheson fused to light or to heavy chain to produce a heteroantibody. The aforementioned immunoglobulins having one or more arms bearing the adheson domain and one or more arms bearing companion variable regions result in dual specificity for adheson ligand and for an antigen. These are produced by the above-described recombinant methods or by in vitro procedures. In the latter case, for example, F(ab')$_2$ fragments of the adheson fusion and an immunoglobulin are prepared, the F(ab')2 fragments converted to Fab' fragments by reduction under mild reducing conditions, and then reoxidized in each other's presence under acidic conditions in accord with methods known per se. See also U.S. Pat. No. 4,444,878.

Additionally, procedures are known for producing intact heteroantibodies from immunoglobulins having different specifictries. These procedures are adopted for the in vitro production of heterochimeric antibodies by simply substituting the immunoadheson chains for one of the previously employed immunoglobulins.

In an alternative method for producing a heterofunctional antibody, host cells producing an adheson-immunoglobulin fusion, e.g. transfected myelomas, also are fused with B cells or hybridomas which secrete antibody having the desired companion specificity for an antigen. Heterobifunctional antibody is recovered from the culture medium of such hybridomas, and thus may be produced somewhat more conveniently than by conventional in vitro resorting methods (EP 68,763).

Another group of fusions are those in which an adheson is conjugated with a toxic substance, e.g. a polypeptide such as ricin (including deglycosylated ricin A chain), diptheria toxin A, or a non-peptidyl cytotoxin. Where the toxin is a polypeptide it is convenient to cross-link the polypeptide to the adheson or its trans-

membrane-deleted variant by conventional in vitro protein cross-linking agents (for suitable methods for linking ricin A chain or deglycosylated A chain to CD4 see, for example, Duncan et al., "Analy. Biochem."132:68–73 [1983]; Thorpe et al., "Cancer Res." 47:5924 [1987]; and Ghotie et al., "Cancer Res." 48:2610 [1988]) or by recombinant synthesis as a fusion (see for example, U.S. Pat. No. 4,765,382). Alternatively, where companion antibodies are anti-ricin antibody immunoglobulin variable domains, such immunoglobulin heteroantibodies are employed to deliver ricin to HIV infected cells following the general procedure of Raso et al., Cancer Research, 41:2073 (1981).

Another class of adheson variants are deletional variants. Deletions are characterized by the removal of one or more amino acid residues from a adheson sequence. Typically, the transmembrane and cytoplasmic domains of adhesons ar deleted. In the case of CD4, at least residues 368 to 395 (the transmembrane region), and ordinarily 396–433 as well (the cytoplasmic domain), will be deleted to obtain secreted forms of this adheson. Parenthetically, the amino acid residues follow the numbers given for mature CD4 as noted, for example, in FIGS. 1*a*–1*c*. Thus, CD4T molecules generally will terminate in the vicinity of about residues 366–368, or at any other suitable site N-terminal thereto which preserves the gp120-binding capability of the CD4 variant.

Substitutional variants are those in which at least one residue in the adheson sequence has been removed and a different residue inserted in its place. The native N-terminal residue for mature CD4 is now known to be lysine. Thus, the sequence shown in FIGS. 1*a*–1*c*, with an N-terminal asparagine, is an amino acid sequence variant of native mature CD4. Table 2 below describes substitutions which in general will result in fine modulation of the characteristics of the CD antigen.

TABLE 2

Original Residue	Exemplary Substitutions
Ala	ser
Arg	lys
Asn	gln; his
Asp	glu
Cys	ser; ala
Gln	asn
Glu	asp
Gly	pro
His	asn; gln
Ile	leu; val
Leu	ile; val
Lys	arg; gln; glu
Met	leu; ile
Phe	met; leu; tyr
Ser	thr
Thr	ser
Trp	tyr
Tyr	trp; phe
Val	ile; leu

Substantial changes in function or immunological identity are made by selecting substitutions that are less conservative than those in Table 2, i.e., selecting residues that differ more significantly in their effect on maintaining (a) the structure of the polypeptide backbone in the area of the substitution, for example as a sheet or helical conformation, (b) the charge or hydrophobicity of the molecule at the target site or (c) the bulk of the side chain. The substitutions which in general are expected to produce the greatest changes in adheson properties will be those in which (a) a hydrophilic residue, e.g. seryl or threonyl, is substituted for (or by) a hydrophobic residue, e.g. leucyl, isoleucyl, phenylalanyl, valyl or alanyl; (b) a cysteinyl or prolyl is substituted for (or by) any other residue; (c) a residue having an electropositive side chain, e.g., lysyl, arginyl, or histidyl, is substituted for (or by) an electronegative residue, e.g., glutamyl or aspartyl; or (d) a residue having a bulky side chain, e.g., phenylalanyl, is substituted for (or by) one not having a side chain, e.g., glycyl.

A preferred class of substitutional or deletional variants are those involving the transmembrane region of the adheson. The transmembrane region of the adheson is a highly hydrophobic or lipophilic domain that is the proper size to span the lipid bilayer of the cellular membrane. It is believed to anchor the adheson in the cell membrane.

Deletion or substitution of the transmembrane domain will facilitate recovery and provide a soluble form of the adheson by reducing its cellular or membrane lipid affinity and improving its water solubility. If the transmembrane and cytoplasmic domains are deleted one avoids the introduction of potentially immunogenic epitopes, either by exposure of otherwise intracellular polypeptides that might be recognized by the body as foreign or by insertion of heterologous polypeptides that are potentially immunogenic. A principal advantage of the transmembrane deleted adheson is that it is secreted into the culture medium of recombinant hosts. This variant is water soluble and does not have an appreciable affinity for cell membrane lipids, thus considerably simplifying its recovery from recombinant cell culture.

It will be amply apparent from the foregoing discussion that substitutions, deletions, insertions or any combination thereof are introduced to arrive at a final construct. As a general proposition, all variants will not have a functional transmembrane domain and preferably will not have a functional cytoplasmic sequence. This is generally accomplished by deletion of the relevant domain, although adequate insertional or substitutional mutagens also can be effective for this purpose. For example, the transmembrane domain is substituted by any amino acid sequence, e.g. a random or homopolynucleic sequence of about 5 to 50 sertne, threonine, lysine, arginine, glutamine, aspartic acid and like hydrophilic residues, which altogether exhibit a hydrophilic hydropathy profile, so that it is secreted into the culture medium of recombinant hosts. This variant should also be considered to be an adheson variant.

These variants ordinarily are prepared by site specific mutagenesis of nucleotides in the DNA encoding the adheson, thereby producing DNA encoding the variant, and thereafter expressing the DNA in recombinant cell culture. However, variant adhesons also are prepared by in vitro synthesis. Obviously, variations made in the DNA encoding the variant adhesons must not place the sequence out of reading frame and preferably will not create complementary regions that could produce secondary mRNA structure deleterious to expression (EP 75,444A). The CD4 variants typically exhibit the same gp120 binding activity as does the naturally-occurring prototype, although variants also are selected in order to modify the characteristics of the CD4 adheson as indicated above.

While the site for introducing an amino acid sequence variation is predetermined, the mutation per se need not be predetermined. For example, in order to optimize the performance of a mutation at a given site, random mutagenesis may be conducted at the target codon or region

and the expressed adheson variants screened for the optimal combination of desired activities. Techniques for making substitution mutations at predetermined sites in DNA having a known sequence are well known, for example M13 primer mutagenesis.

Adheson variants that are not capable of binding HIV gp120 are useful nonetheless as immunogens for raising antibodies to the adheson or as immunoassay kit components (labelled, as a competitive reagent for gp120 assay, or unlabelled as a standard for an adheson assay) so long as at least one adheson epitope remains active.

The DNA encoding adhesons is obtained by known procedures. See Williams, Immunol. Today 8:298–303 (1987) and citations therein. In general, prokaryotes are used for cloning of CD4 variant DNA sequences. For example, *E. coli* strain SR101 (for propagating m13 phage, a λ-resistant strain of JM 101; Messing et al., Nucl. Acids. Res. 9(2):309–321 [1981]); and *E. coli* K12 strain 294 (ATCC No. 31446) are particularly useful. Other microbial strains which may be used include *E. coli* B, UM101 and *E. coli* χ1776 (ATCC No. 31537). These examples are illustrative rather than limiting.

DNA encoding the variant adhesons are inserted for expression into vectors containing promoters and control sequences which are derived from species compatible with the intended host cell. The vector ordinarily, but need not, carry a replication site as well as one or more marker sequences which are capable of providing phenotypic selection in transformed cells. For example, *E. coli* is typically transformed using a derivative of pBR322 which is a plasmid derived from an E. coli species (Bolivar, et al., Gene 2: 95 [1977]). pBR322 contains genes for ampicillin and tetracycline resistance and thus provides easy means for identifying transformed cells. The pBR322 plasmid, or other microbial plasmid must also contain or be modified to contain promoters and other control elements commonly used in recombinant DNA constructions.

Promoters suitable for use with prokaryotic hosts illustratively include the β-lactamase and lactose promoter systems (Chang et al., Nature, 275: 615 [1978]; and Goeddel et al., Nature 281: 544 [1979]), alkaline phosphatase, the tryptophan (trp) promoter system (Goeddel, Nucleic Acids Res. 8: 4057 [1980]and EPO Appln. Publ. No. 36,776) and hybrid promoters such as the tac promoter (H. de Boer et al., Proc. Natl. Acad. Sci. USA 80: 21–25 [1983]). However, other functional bacterial promoters are suitable. Their nucleotide sequences are generally known, thereby enabling a skilled worker operably to ligate them to DNA encoding the adheson variant using linkers or adaptors to supply any required restriction sites (Siebenlist et al., Cell 20: 269 [1980]). Promoters for use in bacterial systems also will contain a Shine-Dalgarno (S.Do) sequence operably linked to the DNA encoding the antigen.

In addition to prokaryotes, eukaryotic microbes such as yeast cultures also are useful as cloning or expression hosts. *Saccharomyces cerevisiae,* or common baker's yeast is the most commonly used eukaryotic microorganism, although a number of other strains are commonly available. For expression in Saccharomyces, the plasmid YRp7, for example, (Stinchcomb, et al., Nature 282: 39 [1979]; Kingsman et al, Gene 1: 141 [1979]; Tschemper et al., Gene 10: 157 [1980]) is commonly used. This plasmid already contains the trp1 gene which provides a selection marker for a mutant strain of yeast lacking the ability to grow in tryptophan, for example ATCC no. 44076 or PEP4–1 (Jones, Genetics 85: 12 [1977]). The presence of the trp1 lesion as a characteristic of the yeast host cell genome then provides an effective means of selection by growth in the absence of tryptophan.

Suitable promoting sequences for use with yeast hosts include the promoters for 3-phosphoglycerate kinase (Hitzeman et al., J. Biol. Chem. 255: 2073 [1980]) or other glycolytic enzymes (Hess et al., J. Adv. Enzyme Reg. 7: 149 [1968]; and Holland, Biochemistry 17: 4900 [1978]), such as enolase, glyceraldehyde-3-phosphate dehydrogenase, hexoktnase, pyruvate decarboxylase, phosphofructokinase, glucose-6-phosphate isomerase, 3phosphoglycerate mutase, pyruvate kinase, triosephosphate isomerase, phosphoglucose isomerase, and glucokinase.

Other yeast promoters, which are inducible promoters having the additional advantage of transcription controlled by growth conditions, are the promoter regions for alcohol dehydrogenase 2, isocytochrome C, acid phosphatase, degradative enzymes associated with nitrogen metabolism, metallothionein, glyceraldehyde-3-phosphate dehydrogenase, and enzymes responsible for maltose and galactose utilization. Suitable vectors and promoters for use in yeast expression are further described in R. Hitzeman et al., European Patent Publication No. 73,657A. Yeast enhancers also are advantageously used with yeast promoters.

Promoters for controlling transcription from vectors in mammalian host cells may be obtained from various sources, for example, the genomes of viruses such as: polyoma, Simian Virus 40 (SV40), adenovirus, retroviruses, hepatitis-B virus and most preferably cytomegalovirus, or from heterologous mammalian promoters, e.g. the beta actin promoter. The early and late promoters of the SV40 virus are conveniently obtained as an SV40 restriction fragment which also contains the SV40 viral origin of replication. Fiers et al., Nature, 273: 113 (1978). The immediate early promoter of the human cytomegalovirus is conveniently obtained as a HindIII E restriction fragment. Greenaway, P. J. et al., Gene 18: 355–360 (1982). Of course, promoters from the host cell or related species also are useful herein.

DNA transcription in higher eukaryotes is increased by inserting an enhancer sequence into the vector. Enhancers are cisacting elements of DNA, usually from about 10 to 300bp, that act to increase the transcription initiation capability of a promoter. Enhancers are relatively orientation and position independent having been found 5' (Laimins, L. et al., Proc.Natl.Acad. Sci. 78: 993 [1981]) and 3' (Lusky, M. L., et al., Mol. Cell Bio. 3: 1108 [1983]) to the transcription unit, within an intron (Banerji, J. L. et al., Cell 33: 729 [1983]) as well as within the coding sequence itself (Osborne, T. F., et al., Mol. Cell Bio. 4: 1293 [1984]). Many enhancer sequences are now known from mammalian genes (globin, elastase, albumin, α-fetoprotein and insulin). Typically, however, one will use an enhancer from a eukaryotic cell virus. Examples include the SV40 enhancer on the late side of the replication Lt origin (bp 100–270), the cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers.

Expression vectors used in eukaryotic host cells (yeast, fungi, insect, plant, animal, human or nucleated cells) may also contain sequences necessary for the termination of transcription which may affect mRNA expression. These regions are transcribed as polyadeny-

lated segments in the untranslated portion of the mRNA encoding the adheson.

Expression vector systems generally will contain a selection gene, also termed a selectable marker. Examples of suitable selectable markers for mammalian cells are dihydrofolate reductase (DHFR), thymtdine kinase or neomycin. When such selectable markers are successfully transferred into a mammalian host cell, the transformed mammalian host cell can survive if placed under selective pressure. There are two widely used distinct categories of selective regimes. The first category is based on a cell's metabolism and the use of a mutant cell line which lacks the ability to grow independent of a supplemented medium. Two examples are: CHO DHFR⁻cells and mouse LTK⁻cells. These cells lack the ability to grow without the addition of such nutrients as thymidine or hypoxanthine. Because these cells lack certain genes necessary for a complete nucleotide synthesis pathway, they cannot survive unless the missing nucleotides are provided in a supplemented medium. An alternative to supplementing the medium is to introduce an intact DHFR or TK gene into cells lacking the respective genes, thus altering their growth requirements. Individual cells which were not transformed with the DHFR or TK gene will not be capable of survival in non supplemented media.

The second category is dominant selection which refers to a selection scheme used in any cell type and does not require the use of a mutant cell line. These schemes typically use a drug to arrest growth of a host cell. Those cells which have a novel gene would express a protein conveying drug resistance and would survive the selection. Examples of such dominant selection use the drugs neomycin, Southern P. and Berg, P., J. Molec. Appl. Genet. 1: 327 (1982), mycophenolic acid, Mulligan, R. C. and Berg, P. Science 209: 1422 (1980) or hygromycin, Sugden, B. et al., Mol. Cell. Biol. 5: 410–413 (1985). The three examples given above employ bacterial genes under eukaryotic control to convey resistance to the appropriate drug G418 or neomycin (geneticin), xgpt (mycophenolic acid) or hygromycin, respectively.

"Amplification" refers to the increase or replication of an isolated region within a cell's chromosomal DNA. Amplification is achieved using a selection agent e.g. methotrexate (MTX) which inactivates DHFR. Amplification or the making of successive copies of the DHFR gene results in greater amounts of DHFR being produced in the face of greater amounts of MTX. Amplification pressure is applied notwithstanding the presence of endogenous DHFR, by adding ever greater amounts of MTX to the media. Amplification of a desired gene can be achieved by cotransfecting a mammalian host cell with a plasmid having a DNA encoding a desired protein and the DHFR or amplification gene permitting cointegration. One ensures that the cell requires more DHFR, which requirement is met by replication of the selection gene, by selecting only for cells that can grow in the presence of ever-greater MTX concentration. So long as the gene encoding a desired heterologous protein has cointegrated with the selection gene replication of this gene gives rise to replication of the gene encoding the desired protein. The result is that increased copies of the gene, i.e. an amplified gene, encoding the desired heterologous protein express more of the desired heterologous protein.

Preferred host cells for expressing the CD antigert variants of this invention are mammalian cell lines, examples including: monkey kidney CV1 line transformed by SV40 (COS-7, ATCC CRL 1651); human embryonic kidney line (293, Graham, F. L. et al., J. Gen Virol. 36: 59 [1977]and 293S cells [293 subclones selected for better suspension growth]); baby hamster kidney cells (BHK, ATCC CCL 10); chinese hamster ovary-cells-DHFR (CHO, Urlaub and Chasin, Proc.Natl.Acad. Sci. (USA) 77: 4216, [1980]); mouse sertoli cells (TM4, Mather, J. P., Biol. Reprod. 23: 243–251 [1980]); monkey kidney cells (CV1 ATCC CCL 70); african green monkey kidney cells (VERO-76, ATCC CRL-1587); human cervical carcinoma cells (HELA, ATCC CCL 2); canine kidney cells (MDCK, ATCC CCL 34); buffalo rat liver cells (BRL 3A, ATCC CRL 1442); human lung cells (W138, ATCC CCL 75); human liver cells (Hep G2, HB 8065); mouse mammary tumor (MMT 060562, ATCC CCL51 cells); and TRI cells (Mather, J. P. et al., Annals N.Y. Acad. Sci. 383: 44–68 [1982]).

"Transformation" means introducing DNA into an organism so that the DNA is replicable, either as an extrachromosomal element or by chromosomal integration. One suitable for transformation of the host cells is the method of Graham, F. and van der Eb, A., Virology 52: 456–457 (1973). However, other methods for introducing DNA into cells such as by nuclear injection or by protoplast fusion may also be used. If prokaryotic cells or cells which contain substantial cell walls are used as hosts, the preferred method of transfection is calcium treatment using calcium chloride as described by Cohen, F. N. et al., Proc. Natl. Acad. Sci. (USA), 69:2110 (1972).

Construction of suitable vectors containing the desired coding and control sequences employ standard and manipulative ligation techniques. Isolated plasmids or DNA fragments are cleaved, tailored, and religated in the form desired to form the plasmids required. Suitable procedures are well known for the construction described herein. See, for example, (Manjarls, T. et al., *Molecular Cloning.* 133–134 Cold Spring Harbor, [1982]; "Current Protocols in Molecular Biology", edited by Ausubel et al., [1987], pub. by Greene Publishing Associates & Wiley-Interscience).

Correct plasmid sequences are confirmed by transforming *E. coli* K12 strain 294 (ATCC 31446) with ligation mixtures, successful transformants selected by ampicillin or tetracycline resistance where appropriate, plasmids from the transformants prepared, and then analyzed by restriction enzyme digestion and/or sequenced by the method of Messing et al., Nucleic Acids Res. 9: 309 (1981) or by the method of Maxam et al., Methods in Enzymology 65: 499 (1980).

Host cells are transformed with the expression vectors of this invention. Thereafter they are cultured in appropriate culture media, e.g. containing substances for inducing promoters, selecting transformants or amplifying genes. The culture conditions, such as temperature, pH and the like, are those previously used with the host cell selected for expression, and will be apparent to the ordinarily skilled artisan.

The secreted adheson variants are recovered and purified from the culture supernatants or lysates of recombinant hosts. Typically, the supernatants are concentrated by ultrafiltration, contacted with a ligand affinity or immunoaffinity matrix so as to adsorb the adheson variant, and eluted from the matrix. Optionally, the adheson is purified by ion exchange chromatography.

Surprisingly, purification of soluble CD4 adheson from culture medium was unexpectedly difficult. Notwithstanding that the hydrophobic transmembrane region of the antigen had been deleted, the antigen exhibited a strong tendency to form aggregates that could be readily removed from suspension by centrifugation at 1000×g, and which avidly coat surfaces such as ultrafiltration membranes. This appears to result from the reduction in concentration of albumin or other serum protein (ordinarily present in the crude preparation) to a particular level, below which the truncated antigen no longer remains soluble. This phenomenon appears to be aggravated by exposure of the CD4 adheson to low pH (<about pH 4). As a result, separation procedures (particularly those that employ acid elution, such as immunoaffinity) should be modified so that the eluate is maintained at, or immediately returned to, about neutrality. Further, a surfactant, e.g. a detergent such as Tween 80, should be included with the antigen during the separation procedure. The final purified product will be stabilized with a predetermined protein such as albumin, and/or a detergent.

The purified adheson is formulated into conventional pharmacologically acceptable excipients.

It is administered to patients at a dosage capable of maintaining a concentration of greater than about 100 ng of soluble CD4 adheson/ml plasma. For CD4 adheson variants having different molecular weights, about 2 picomoles of soluble receptor per ml of plasma will be initially evaluated clinically in order to establish a stoichiometric equivalence with native (membrane bound) and soluble receptor. The ordinary dosage of soluble CD4 is 100 μg/kg of patient weight/day.

In order to facilitate understanding of the following examples certain frequently occurring methods and/or terms will be described.

"Plasmids" are designated by a lower case p preceded and/or followed by capital letters and/or numbers. The starting plasmids herein are either commercially available, publicly available on an unrestricted basis, or can be constructed from available plasmids in accord with published procedures. In addition, equivalent plasmids to those described are known in the art and will be apparent to the ordinarily skilled artisan.

"Digestion" of DNA refers to catalytic cleavage of the DNA with a restriction enzyme that acts only at certain sequences in the DNA. The various restriction enzymes used herein are commercially available and their reaction conditions, cofactors and other requirements were used as would be known to the ordinarily skilled artisan. For analytical purposes, typically 1 μg of plasmid or DNA fragment is used with about 2 units of enzyme in about 20 μl of buffer solution. For the purpose of isolating DNA fragments for plasmid construction, typically 5 to 50 μg of DNA are digested with 20 to 250 units of enzyme in a larger volume. Appropriate buffers and substrate amounts for particular restriction enzymes are specified by the manufacturer. Incubation times of about 1 hour at 37° C. are ordinarily used, but may vary in accordance with the supplier's instructions. After digestion the reaction is electrophoresed directly on a polyacrylamide gel to isolate the desired fragment.

"Recovery" or "isolation" of a given fragment of DNA from a restriction digest means separation of the digest on polyacrylamide or agarose gel by electrophoresis, identification of the fragment of interest by comparison of its mobility versus that of marker DNA fragments of known molecular weight, removal of the gel section containing the desired fragment, and separation of the gel from DNA. This procedure is known generally (Lawn, R. et al., Nucleic Acids Res. 9: 6103–6114 [1981], and Goeddel, D. et al., Nucleic Acids Res. 8: 4057 [1980]).

"Dephosphorylation" refers to the removal of the terminal 5' phosphates by treatment with bacterial alkaline phosphatase (BAP). This procedure prevents the two restriction cleaved ends of a DNA fragment from "circularizing" or forming a closed loop that would impede insertion of another DNA fragment at the restriction site. Procedures and reagents for dephosphorylation and other recombinant manipulations are conventional. Reactions using BAP are carried out in 50mM Tris at 68° C. to suppress the activity of any exonucleases which may be present in the enzyme preparations. Reactions were run for 1 hour. Following the reaction the DNA fragment is gel purified.

"Ligation" refers to the process of forming phosphodiester bonds between two double stranded nucleic acid fragments (Maniatis, T. et al., Id. at 146). Unless otherwise provided, ligation may be accomplished using known buffers and conditions with 10 units of T4 DNA ligase ("ligase") per 0.5 μg of approximately equimolar amounts of the DNA fragments to be ligated.

"Filling" or "blunting" refers to the procedures by which the single stranded end in the cohesive terminus of a restriction enzyme-cleaved nucleic acid is converted to a double strand. This eliminates the cohesive terminus and forms a blunt end. This process is a versatile tool for converting a restriction cut end that may be cohesive with the ends created by only one or a few other restriction enzymes into a terminus compatible with any blunt-cutting restriction endonuclease or other filled cohesive terminus. Typically, blunting is accomplished by incubating 2–15 μg of the target DNA in 10mM $MgCl_2$, 1mM dithiothreitol, 50mM NaCl, 10mM Tris (pH 7.5) buffer at about 37° C. in the presence of 8 units of the Klenow fragment of DNA polymerase I and 250 μM of each of the four deoxynucleoside triphosphates. The incubation generally is terminated after 30 min. phenol and chloroform extraction and ethanol precipitation.

The following examples merely illustrate the best mode now contemplated for practicing the invention, but should not be construed to limit the invention. All literature citations herein are expressly incorporated by reference.

EXAMPLE 1

Construction of Vectors for the Expression of Native CD4 and Secreted Derivatives

Section 1

The plasmid used for recombinant synthesis of human CD4 was pSVeCD4DHFR. The plasmid was constructed as follows:

λCD4P1 containing most of the coding sequence of human CD4 (obtained from a human placental cDNA library using oligonucleotide probes based on the published sequence [Maddon et al. 1985]) was digested with EcoRI to produce the cDNA insert. This fragment was recovered by polyacrylamide gel electrophoresis (fragment 1).

pUC18 was digested with EcoRI and the single fragment recovered by polyacrylamide gel electrophoresis (fragment 2). Fragment 1 was ligated to fragment 2 and the ligation mixture transformed into *E. coli* strain 294.

The transformed culture was plated on ampicillin media plates and resistant colonies selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct DNA fragments. This plasmid is referred to as pUCCD4.

pSVeE'DHFR (Muesing et al., Cell 48:691–701 [1987]) was digested with KpnI and BamHI and blunted with *E. coli* DNA polymerase I (Klenow fragment) and the four dNTPs. Fragment 3 containing the pML-Amp^r region, SV40 early promoter, the HIV LTR, and the mouse DHFR gene was recovered by gel electrophoresis, ligated and the ligation mixture transformed into *E. coli* strain 294. The transformed culture was plated on ampicillin media plates and resistant colonies selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the BamHI restriction site and the absence of the KpnI restriction site. This plasmid is referred to as pSVeΔBKDHFR and allows EcoRI-BamHI fragments to be inserted after the SV40 early promoter and transcribed under its control, following transfection into an appropriate cell line.

Synthetic oligonucleotides (adaptors 1–8, below) were made to extend from 76 bp 5' of the initiation codon of CD4 translation to the RsaI restriction site at 121 bp 3' of the initiator, with the sequence AATT at the 5' end of the sense strand to generate an end which could ligate to an EcoRI restriction fragment. These oligonucleotides were ligated and the 204 bp fragment containing the entire sequence recovered by gel electrophoresis (fragment 4).

CD4 adaptor 1: AATTCAAGCCCAGAGCCCTG-
CCATTTCTGTGGGCTCAGGTCCCT

CD4 adaptor 2: pACTGCT-
CAGCCCCTTCCTCCCTCGGCAAGG-
CCACAATGAACCGGGGAGTC

CD4 adaptor 3: pCCTTTTAGG-
CACTTGCTTCTGGTGCT-
GCAACTGGCGCTCCTCCCAGC

CD4 adaptor 4: pAGCCACTCAGG-
GAAACAAAGTGGTGCTGG-
GCAAAAAAGGGGATACAGTGGAACT-
GACCTGT

CD4 adaptor 5: pACAGGTCAGTTCCACT-
GTATCCCCTTTTTTGCCCAGCAC-
CACTTTGTTTCC

CD4 adaptor 6: pCTGAGTGGCTGCTGGGAG-
GAGCGCCAGTTGCAGCACCAGAAG-
CAAGT

CD4 adaptor 7: pGCCTAAAAGG-
GACTCCCCGGTTCATTGTGGCCTTGC-
CGAGGGAGGAAGGG

CD4 adaptor 8: GCTGAGCAGTAGGGACCT-
GAGCCCACAGAAATG-
GCAGGGCTCTGGGCTTG

pUCCD4 was digested with RsaI and SstI and the 401 bp fragment containing part of the CD4 coding sequence recovered by gel electrophoresis (fragment 5). pUC18 was digested with EcoRI and SstI and the fragment comprising the bulk of the plasmid recovered by gel electrophoresis (fragment 6). Fragments 4 and 5 were ligated to fragment 6 and the ligation mixture transformed into *E. coli strain* 294. The transformed culture was plated on ampicillin media plates and resistant colonies selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct fragment. The sequence of the inserted synthetic DNA was checked by excising the 605 bp EcoRI-SstI fragments from several transformants and ligating them to M13mp19 which had been digested with the same enzymes. After transformation into *E. coli* strain JM101, single-stranded DNA was prepared and sequenced. One plasmid which contained the correct sequence was selected, and is referred to as pCD4int.

pCD4int was digested with EcoRI and SstI and fragment 7 containing the 5' end of the CD4 coding region was recovered by gel electrophoresis. pUCCD4 was digested with SstI and BamHI and the 1139 bp fragment containing the remainder of the CD4 coding region (fragment 8) recovered by gel electrophoresis.

pSVeΔBKDHFR was digested with EcoRI and BamHI and fragment 9 comprising the bulk of the plasmid was isolated. Fragments 7, 8 and 9 were ligated and the ligation mixture transformed into *E. coli* strain 294. The transformed culture was plated on ampicillin media plates and the resistant colonies selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct fragment. This plasmid is referred to as pSVeCD4DHFR, and was used to direct synthesis of recombinant intact CD4.

Section 2

A plasmid was constructed to direct the synthesis of a CD4 derivative lacking the putative transmembrane domain and most of the putative cytoplasmic domain (Maddon et al.). This was done with the intention of creating a secreted form of CD4, based on the assumption that these domains anchor the CD4 glycoprotein to the cell membrane, and that their deletion would result in the secretion of the product. This plasmid is referred to as pSVeCD4ΔN1aDHFR and was constructed as follows:

pUCCD4 was digested with SstI and TaqI and the 531 bp fragment (fragment 10) recovered. pUCCD4 was digested with N1aIII and TagI and the 112 bp fragment (fragment 11) recovered. pUCCD4 was digested with HamHi and N1aIII and the 301 bp fragment (fragment 12) recovered. pCD4tnt was digested with SstI and BamHI and fragment 13 comprising the bulk of the plasmid recovered. Fragments 10, 11, and 12 were ligated together with fragment 13 and the ligation mixture transformed into *E. coli* strain 294. The transformed culture was plated on ampicillin media plates and resistant colonies selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct fragment. Plasmid DNA from several transformants was sequenced to ensure that the 195 bp N1aIII fragment had been deleted and that the proper reading frame was restored. The resulting plasmid is referred to as pCD4ΔN1a.

pCD4ΔN1a was digested with EcoRI and BamHI and the 1541 bp fragment containing the sequence of a CD4 derivative lacking the transmembrane and cytoplasmic domains recovered (fragment 14) and ligated to fragment 9 and the ligation mixture transformed into *E. coli* strain 294. The transformed culture was plated on ampicillin media plates and resistant colonies selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct fragment. This plasmid is referred to as pSVeCD4ΔN1aDHFR.

Both pSVeCD4DHFR and pSVeCD4ΔN1aDHFR were transfected into CHO cells by the same method used to establish cell lines stably expressing HIV-I poly-

peptides (Muesing, Smith and Capon, Cell 48:6910701 [1987]). These cells were assayed for production by radioimmunoprecipitation as described below. While no product was detected in initial experiments, subsequent experiments showed that the above described coding segment could indeed direct the synthesis of a soluble CD4 adheson variant both in CHO and 293 cells.

Section 3

A different expression system was initially used for the synthesis and expression of a CD4 variant lacking completely the cytoplasmic and transmembrane domains. This system uses the cytomegalovirus promoter and can be used in cultured cells of human origin. The first plasmid constructed for use in this system contained the entire coding region for CD4 and was intended to function as a control in the following studies. It is referred to as pRKCD4, and was constructed as follows:

pSVeCD4DHFR was digested with EcoRI and BamHI and fragment 15 containing the entire CD4 coding region was isolated. pRK5 (U.S.S.N. 07,472, filed Sep. 11, 1987, now abandoned) was digested with EcoRI and BamHI and fragment 16 comprising the bulk of the plasmid recovered by gel electrophoresis, ligated to fragment 15, and the ligation mixture transformed into *E. coli* strain 294. The transformed culture was plated on ampicillin media plates and resistant colonies selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct fragment. This plasmid is referred to as pRKCD4.

Section 4

The next plasmid constructed was designed to direct the expression of the above-mentioned (Section 3) secreted derivative of CD4. The coding region of CD4 was fused after amino acid residue 368 of mature CD4 to a sequence from pBR322 which codes for 9 more residues before a translation termination codon. This removes the putative CD4 transmembrane and cytoplasmic domains, which are presumed to anchor CD4 to the cell surface. The plasmid is referred to as pRKCD4T (and which produces protein called CD4T), and was constructed as follows:

pSVeCD4DHFR was digested with HpaII, blunted with Klenow fragment and the four dNTPs, and digested with BstEII. The 382 bp fragment (fragment 17) containing part of the CD4 coding sequence was recovered by gel electrophoresis. pSVeCD4DHFR was digested with EcoRI and BstEII and the 874 bp fragment (fragment 18) recovered. pBR322 was digested with HindIII, blunted with Klenow fragment and the four dNTPs, and digested with EcoRI. Fragment 19 comprising the bulk of the plasmid was isolated and ligated to fragments 17 and 18 and the ligation mixture transformed into *E. coli* strain 294. The transformed culture was plated on ampicillin media plates and resistant colonies selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct fragment. This plasmid is referred to as pCD4Tint.

pRK5 was digested with EcoRI and SmaI and fragment 20 comprising the bulk of the plasmid isolated. pCD4Tint was digested with EcoRI and EcoRV and the 1410 bp fragment containing the CD4 coding sequence to the HpaII site at 1176 bp 3' of the initiating codon and the 154 bp HindIII-ECoRV fragment of pBR322 was recovered (fragment 21). Fragments 20 and 21 were ligated and the ligation mixture transformed into *E. coli* strain 294. The transformed culture was plated on ampicillin media plates and resistant colonies selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct fragment. This plasmid is referred to as pRKCD4T.

Section 5*a*

In order to create a secreted form of CD4 which could be purified with an antibody directed to herpes virus type I glycoprotein D, a plasmid was constructed to express a derivative of CD4T in which the region coding for the mature, processed CD4T polypeptide was fused to a sequence coding for the signal peptide and the first 27 residues of the mature type I Herpes Simplex Virus gD glycoprotein. This plasmid is referred to as pRKGDCD4T, and was constructed as follows:

pgDTrunc. DHFR was digested with EcoRI and PvuII and the fragment containing the coding region for the signal peptide and first 27 residues of the mature HSV I gD glycoprotein was isolated (fragment 22). pRKCD4T was digested with EcoRI and BstEII and fragment 23 containing the 3' end of the CD4 coding sequence and the pRK5 region was isolated.

Synthetic oligonucleotides GD (adaptors 1–2, below) containing the coding sequence of CD4 from the codon for the amino terminal residue of mature CD4 to the RsaI site at 121 bp 3' of translation initiation, and containing the sequence CTGCTCGAG at the 5' end of the sense strand were prepared (fragment 24). pRKCD4 was digested with RsaI and BstEII and the 665 bp fragment containing part of the coding region for CD4 was recovered (fragment 25) and ligated to fragment 24. After digestion with BstEII to ensure that only monomeric fragment was present, the 724 bp fragment containing both sequences was recovered by gel electrophoresis (fragment 26).

Fragments 22, 23 and 26 were ligated and the ligation mixture transformed into *E. coli* strain 294. The transformed culture was plated on ampicillin media plates and resistant colonies selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct fragment. The sequence of several transformants was checked to ensure that the synthetic insert was correct and that reading frame was preserved. This plasmid is referred to as pRKGDCD4T.

These pRK5 derived plasmids preferably were transfected into 293S cells for stable expression according to Muesing, et al. Cell 48:691 (1987) with the exception that in addition to the plasmid of interest a plasmid expressing the neomycin resistance gene pRSV neo (Gorman et al. Science 221:553–555 (1985)) was co-transfected. 293 cells also are used satisfactorily as host cells. 2 days after transfection, the cells were passaged into standard medium (1:1 F12/DME supplemented with L-glutamine, penicillin-streptomycin and 10% FBS) with 0.5 mg/ml G418 (Genticin sulfate; Gibco) for selection of stable cell lines, rather than in media containing methotrexate as shown by Muesing et al. Cells were assayed for production of CD4 or CD4 analogs by radioimmunoprecipitation. Binding studies (section 5*c*) used conditioned supernatants from these cells in the 1:1 F12/DME medium. Materials used in

infectivity assays (section 5b) were obtained as described in section 8 below.

gDCD4 adaptor 1: CTGCTCGAGCAGG-GAAACAAAGTGGTGCTGG-GCAAAAAAGGGGATACAGTGGAACTGAC

gDCD4 adaptor 2: pACAGGTCAGTTCGACT-GTATCCCCTTTTTTGCCCAGCAC-CACTTTGTTTCCCTGCTCGA

Section 5b

The following constitutes a study of the neutralization of HIV-1 infectivity by soluble CD4 analogs. A modification of the neutralization procedure of Robert-Guroff et al., Nature 316:72 (1985) was followed. Equal volumes of inhibitor supernatant and virus (60 microliters) were incubated at 4 degrees C. for 1 hour, then the same volume of H9 (Gallo et al., Science 224:500, 1984) at 5×10^6/ml was added and incubation continued for 1 hour at 37 degrees C. Following absorption, 2.5×10^5 cells in 150 microliters were transferred to 2 ml of incubation media. After 4 days at 37 degrees C., the cultures were split 1:2 with fresh media and incubated for an additional 3 days. Cultures were harvested, reverse transcriptase activity was measured (Groopman et al., AIDS Research and Human Retroviruses 3:71, 1987), and immunofluorescence reactivity with HIV-1 positive serum was determined as described (Poiesz et al., Proc. Acad. Nat. Sci. USA 77:7415, 1980). Inhibitor supernatants were obtained from confluent plate cultures of 293S/CDT4, 293S/gDCD4T cells or untransfected 293S cells by replacing the growth medium incubation media and harvesting the supernatants 24 hours later. Inhibitor supernatant replaced part or all of the incubation media during the first three days of culture as indicated in the second column of Table 3. Challenge dose of virus was 100 $TCID_{50}$ (Groopman et al., supra) of HIV-1 strain HTLV-IIIB grown in H9 cells assayed in the same system. Incubation media consisted of RPMI 1640 media containing 2mM L-glutamine, 100 units/ml penicillin, 100 micrograms/ml streptomycin, 2 micrograms/ml polybrene and 20% fetal calf serum (M. A. Bioproducts).

TABLE 3

Inhibitor supernatant	Dilution of Inhibitor supernatant	Indirect immunofluorescence (% positive cells	Reverse transcriptase (cpm/ml × 10^5)
mock-transfected	undil.; 1:4	65.3 65.5	21.8 23.9
mock-transfected	undil.; 1:4	61.2 61.1	18.5 28.1
CD4T	undil.; 1:4	0.4 18.0	0.11 5.94
CD4T	undil.; 1:4	0.8 16.1	0.15 3.72
gDCD4T	undil.; 1:4	0.4 26.8	0.14 9.92
gDCD4T	undil.; 1:4	1.4 36.1	0.23 11.3

Both forms of soluble CD4 virtually abolished the growth of HIV-1, when incubated with virus-infected cells without prior dilution (Table 2). At a dilution of 1:4 the soluble CD4 preparations were only partially effective in inhibiting virus growth, however the level of fluorescent-positive cells and reverse transcriptase was still significantly lower than cultures receiving mock-transfected cell supernatants (Table 2). Since there was no significant difference in virus growth between diluted and undiluted control supernatants, nor did any of the supernatants affect the growth of uninfected H9 cells (data not shown), soluble CD4 proteins present in these supernatants were concluded to be responsible for the neutralization of HIV-1 infection of H9 cells.

Section 5c

To determine the affinity constant for interactions between gp120 and CD4 or CD4 variants, saturation binding analysis was carried out with soluble CD4 (supra) and detergent solubilized intact CD4 (Lasky et al. Cell 50:975 [1987]) employing radioiodinated gp120 labeled with lactoperoxidase. Binding reactions consisted of ^{125}I-gp120 (3 ng to 670 ng, 2.9 nCi/ng) incubated for 1 hour at 0 degrees C. with cell lysates containing intact CD4 (Laskey et al., op cit.) or cell supernatants containing unlabeled CD4T or gDCD4T prepared as described in section 5a. Reactions (0.2ml) had a final composition of 0.5× McDougal Lysts Buffer (McDLB) (1×McDLB contains 0.5% Nontdet NP-40, 0.2% Na deoxycholate, 0.12 M NaCl, 0.02 M Tris-Hc1, pH 8.0) and were performed in duplicate, both in the presence or absence of 50 micrograms of unlabeled purified gp120 (74 fold or greater excess). Following incubation, bound gp120 was quantitated by immunoprecipitation and counted in a gamma counter. For immunoprecipitation, binding reaction solutions were preabsorbed with 5 microliters of normal rabbit serum for one hour at 0° C., and cleared with 40 microliters of Pansorbin (10% w/v, Calbiochem) for 30 minutes at 0 degrees C. Samples were then incubated overnight at 0 degrees C. with 2 microliters of normal serum or 5 microliters (0.25 microgram) of OKT4 monoclonal antibody (Ortho) followed by collection of immune complexes with 10 microliters of Pansorbin. Precipitates were washed twice in iX McDLB and once in water, then eluted by eluting at 100 degrees C. for 2 minutes in sample buffer (0.12 M Tris-HC1 pH 6.8, 4% SDS, 0.7 M mercaptoethanol, 20% glycerol, and 0.1% bromophenol blue). CD4 molecules were bound saturably by gp120, and yielded a simple mass action binding curve. Supernatants from mock-transfected cells gave a level of specifically bound gp120 less than 1% that found for supernatants containing soluble CD4. Scatchard analysis revealed a single class of binding sites on each molecule, with apparent dissociation constants (Kd) of 1.3×10^9 M, 0.83×10^9 M and 0.72×10^{-9} M for intact CD4, CD4T and gDCD4T, respectively. The values obtained for CD4-gp120 binding in solution are comparable to the affinity previously measured for gp120 binding to CD4 on whole cells ($Kd=4.0\times10^{-9}$ M. Lasky, Cell, supra).

Section 6

In order to produce secreted derivatives of CD4 which are free of extraneous amino acid residues, two plasmids were constructed for expression in 293 cells. The plasmids contain CD4 genes which have been truncated without the addition of extra residues, and are referred to as pRKCD4ΔN1a and pRKCD4TP (and which produce proteins called CD4TP and CD4ΔN1a), and were constructed as follows:

Fragment 14 containing the CD4 gene with the 195 bp N1aIII restriction fragment deleted was ligated to fragment 16, which is pRK5 digested with EcoRI and BamHI. The ligation mixture was transformed into *E. coli* strain 294, the transformed culture plated on ampicillin media plates and resistant colonies selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the

correct fragment. The resulting plasmid is referred to as pRKCD4ΔN1a.

Synthetic DNA (5'CGT GAT AGA AGC TTT CTA GAG 3') was made to attach to the HpaII site at 1176bp and which when so attached would terminate translation after amino acid residue 368 of mature CD4 (fragment 27). The other end of this fragment was designed to ligate to BamHI restriction fragments. pUCCD4 was digested with BstEII and HpaII and the 382bp fragment containing part of the CD4 gene was recovered (fragment 28). Fragments 27 and 28 were ligated and then digested with BstEII to reduce dimerized fragments to monomers, and the resulting 401bp fragment was recovered (fragment 29).

pRKCD4 was digested with BstII and BamHI and the fragment comprising the bulk of the plasmid (fragment 30) was isolated and ligated to fragment 29. The ligation mixture was transformed into *E. coli* strain 294, the transformed culture plated on epicillin media plates and resistant colonies selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct fragment. The resulting plasmid is referred to as pRKCD4TP. Both plasmids are transfected into 293 cells to generate stable variant CD4-expressing cell lines as described above.

Section 7

Two plasmids were constructed to direct the expression of secreted CD4 lacking extraneous amino acid residues in CHO cells. These are referred to as pSVeCD4ΔN1aSVDHFR and pSVeCD4TPSVDHFR (and which encode proteins having the primary sequence of CD4ΔN1a and CD4TP), and were constructed as follows:

pE348HBV.E400D22 was digested with PvuI and EcoRI and the fragment containing the SV40 early promoter and part of the β-lactamase gene was recovered (fragment 31). pE348HBV.E400D22 was digested with PvuI and BamHI and the large fragment containing the balance of the β-lactamase gene as well as the SV40 early promoter and the DHFR gene was isolated (fragment 32).

Fragments 31 and 32 were ligated together with fragment 14 and transformed into *E. coli* strain 294. The transformed culture was plated on ampicillin media plates and resistant colonies selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct fragment. The resulting plasmid is referred to as pSVECD4ΔN1aSVDHFR. This plasmid contains the same DNA fragment encoding the soluble CD4 molecule found in the above-mentioned plasmid pSVeCD4ΔN1aDHFR (Section 2).

pRKCD4TP was digested with EcoRI and BamHI and the fragment containing the truncated CD4 coding region was isolated and ligated to fragments 31 and 32. The ligation mixture was transformed into *E. coil* strain 294, the transformed culture plated on ampicillin media plates and resistant colonies selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct fragment. The resulting plasmid is referred to as pSVeCD4TPSVDHFR. Both of these plasmids are transfected into CHO cells and amplified transfectants selected by methotrexate using conventional procedures.

EXAMPLE 2

Fusions of the V region of the CD4 gene, which is homologous to the variable region of immunoglobulin genes (ref Maddon et al. 1985), to the constant (C) region of human immunoglobuitn κ and γ2 chains are constructed as follows:

Synthetic DNA is made to code for the C region of human κ chain (residues 109–214) based on the sequence published by Morin et al., Proc. Natl. Acad. Sci. 82:7025–7029, with the addition at the 5' end of the coding strand of the sequence GGGG, which allows this fragment to be ligated to the BspMI site at the end of the putative V-like region of CD4. At the 3' end of the coding region, a translational stop codon is added as well as a sequence which allows this end to be ligated to BamHI restriction fragments. The synthetic DNA is made in 8 fragments, 4 for each strand, 70–90 bases long. These are then allowed to anneal and ligated prior to isolation on a polyacrylamide gel (fragment 33).

pRKCD4 is digested with EcoRI and BspMI and the 478bp fragment containing the region coding for the putative V-like domain of CD4 is recovered (fragment 34). Fragments 33 and 34 are ligated together with fragment 16 (from the expression vector pRK5). The ligation mixture is transformed into *E. coli* strain 294, the transformed culture plated on ampicillin media plates and resistant colonies selected. Plasmid DNA is prepared from transformants and checked by restriction analysis for the presence of the correct fragment. The resulting plasmid is referred to as pRKCD4Ck.

A plasmid encoding a fusion of the CD4 V-like domain to the human immunoglobulin Cγ2 region is constructed in a similar fashion, and is referred to as pRKCD4Cγ2. Both of these plasmids are transfected into 293 cells, myeloma cells or other competent cells in order to obtain cell lines expressing variant CD4 molecules as described above.

EXAMPLE 3

The gDCD4T secreted by the method of Example 1 was purified from cell culture fluid containing either 10% FBS (fetal bovine serum) or no added FBS. The conditioned cell culture fluid was first concentrated by ultrafiltration then purified by immunoaffinity chromatography. The immunoaffinity column was produced by coupling murine monoclonal antibody 5B6 (whose epitope is on the HSV-1 gD portion of the gDCD4T molecule) to glyceryl coated controlled pore glass by the method of Roy et al., 1984. The concentrated cell culture fluid is applied directly to the column and the contaminating proteins are washed away with neutral pH buffer. The column is then washed with neutral buffer containing tetramethylammonium chloride followed by neutral buffer containing Tween 80. The bound gDCD4T is eluted from the column with buffer at pH3 containing Tween 80 (0.1% w/v) and is neutralized immediately as it is eluted. The eluted neutralized gDCD4T is then concentrated by ultrafiltration and dialyzed/diafiltered to exchange the buffer for a physiological salt solution containing Tween 80 at approximately 0.1% w/v.

If the detergent is not present the gDCD4T forms aggregates as evidenced by the ability of centrifugation at approximately 10,000 Xg for 2 minutes to remove the gDCD4T from the solution. Incubation of gDCD4T at 4° C. in 0.1M sodium acetate, 0.5M NaCl and 0.25M tris at pH 7 together with BSA, Tween 80 or glycerol as

candidate stabilizers showed that, in the absence of a stabilizer the gDCD4T gradually aggregated over the space of 12 days to the point where only about 60–70% of the protein was soluble. However, use of 0.1% w/v Tween 80 or (0.5 mg/ml BSA ensured that about 100% or 80%, respectively, of the gDCD4T remained soluble over this period. Surprisingly glycerol was ineffective as a stabilizer and produced results inferior even to the control-at 8 days about 80% of the gDCD4T was aggregated when stored in the presence of glycerol.

EXAMPLE 4

Plasmids were constructed to direct the expression of proteins containing differing lengths of the amino-terminal, extracellular domain of CD4 fused to the constant region of human immunoglobulin γ1. These plasmids are referred to as pRKCD4$_{2\gamma1}$, pRKCD4$_{e4\gamma1}$, pRKCD4$_{2\gamma1}$, pRKCD4$_{e2\gamma1}$, pRKCD4$_{1\gamma1}$, and pRKCD4$_{e1\gamma1}$.

Plasmid pRKCD4$_{4\gamma1}$ contains the portion of the CD4 gene from the initiation codon to the fusion site after the codon for serine reside 366 of the mature CD4 polypeptide, immediately followed by the sequence coding for the constant region of human immunoglobulin γ1, starting at the codon for serine residue 114 of mature human immunoglobulin γ1 (Kabat et al.).

Plasmid pRKCD4$_{e4\gamma1}$ contains the portion of the CD4 gene from the initiation codon to the fusion site after the codon for lysine residue 360 of the mature CD4 polypeptide, immediately followed by the sequence coding for the constant region of human immunoglobulin γ1, starting at the codon for serine residue 114 of mature human immunoglobulin γ1 (Kabat et al.).

Plasmid pRKCD4$_{2\gamma1}$ contains the portion of the CD4 gene from the initiation codon to the fusion site after the codon for glutamine residue 180 of the mature CD4 polypeptide, immediately followed by the sequence coding for the constant region of human immunoglobulin γ1, starting at the codon for serine residue 114 of mature human immunoglobulin γ1 (Kabat et al.).

Plasmid pRKCD4$_{e2\gamma1}$ contains the portion of the CD4 gene from the initiation codon to the fusion site after the codon for leucine residue 177 of the mature CD4 polypeptide, immediately followed by the sequence coding for the constant region of human immunoglobulin γ1, starting at the codon for serine residue 114 of mature human immunoglobulin γ1 (Kabat et al.).

Plasmid pRKCD4$_{1\gamma1}$ contains the portion of the CD4 gene from the initiation codon to the fusion site after the codon for aspartic acid residue 105 of the mature CD4 polypeptide, immediately followed by the sequence coding for the constant region of human immunoglobulin γ1, starting at the codon for serine residue 114 of mature human immunoglobulin γ1 (Kabat et al.).

Plasmid pRKCD4$_{e1\gamma1}$ contains the portion of the CD4 gene from the initiation codon to the fusion site after the codon for leucine residue 100 of the mature CD4 polypeptide, immediately 25 followed by the sequence coding for the constant region of human immunoglobulin γ1, starting at the codon for serine residue 114 of mature human immunoglobulin γ1 (Kabat et al.).

Construction of these plasmids required the prior construction of plasmid pRKCD4TP/γ1. It was constructed as follows:

A cDNA clone coding for human immunoglobulin γ1 was obtained from a human spleen cDNA library (Clontech Laboratories, Inc.) using oligonucleotides based on the published sequence (Ellison et al., "Nucl. Acids Res." 10:4071–4079 [1982]). and an EcoRI-BagI fragment (the EcoRI site was contributed by a linker; see FIG. 4*a*,*b*) containing part of the variable and all of the constant region was obtained. This fragment was blunted with Klenow fragment, and recovered by gel electrophoresis (Fragment al).

Plasmid pRKCD4TP-kk, encoding a substitutional variant of soluble CD4 (residues 1-368) containing a lysine residue instead of asparagine at position 1 of the mature polypeptide, was constructed from plasmid pRKCD4TP by site-directed mutagenesis. A synthetic oligonucleotide was made as a primer for a mutagenesis reaction to obtain the desired coding sequence. This was synthesized as a 51-mer which contained two silent mutations from the natural sequence in addition to the substitution mutation, and 21 bases on each side of the mutated codons:

5'-CCC TTT TTT GCC CAG CAC CAC CTT CTT GCC CTG-AGT GGC TGC TGG GAG GAG-3'

Plasmid pRKCD4TP was transformed into *E. coli* strain SR101 and the transformed colonies plated on ampicillin media plates. Resistant colonies were selected and grown in the presence of m13K07 helper bacteriophage to yield secreted, encapsidated single-stranded templates of pRKCD4TP. The single-stranded plasmid DNA was isolated and used as the template for mutagenesis reactions with the synthetic oligonucleotides described above as primers. The mutagenesis reaction was transformed *E. coli* SR101 and the transformed culture plated on ampicillin media plates. Transformants were screened by colony hybridization (ref. Grunstein-Hogness) for the presence of the appropriate sequence using the following 16 mer as the probe.

5'-C CAC CTT CTT GCC CTG -3'

The hybridization conditions chosen were sufficiently stringent that the probe only detects the correctly fused product. Colonies identified as positive were selected and plasmid DNA was isolated and transformed into *E. coli* strain SR101. The transformed cultures were plated on ampicillin media plates, and resistant colonies were selected and grown in the presence of m13K07 bacteriophage. Templates were prepared as above and screened by sequencing.

Plasmid pRKCD4TP-kk was digested with XbaI and treated with Klenow Enzyme, and Fragment a2, containing the linearized plasmid was recovered by gel electrophoresis, and ligated with fragment al. The ligation mixture was transformed into *E. coli* strain 294, the transformed culture plated on ampicillin media plates and resistant colonies selected. Plasmid DNA was prepared from the transformants and checked by restriction analysis for the presence of the correct fragment in the correct orientation (i.e., the immunoglobulin coding region in the same orientation as the CD4 coding region, and at the 3' end of the CD4 coding region). This plasmid is referred to as pRKCD4TP/γ1.

Synthetic oligonucleotides were made as primers for deletional mutagenesis reactions to fuse the appropriate coding sequences of IgG1 and CD4 as described above. These were synthesized as 48-mers comprising 24 nucleotides on each side of the desired fusion site (i.e., corresponding to the COOH-terminal 8 residues of the desired CD4 moiety, and the NH_2-terminal 8 residues of the desired immunoglobulin moiety). Plasmid pRKCD4TP/γ1 was transformed into *E. coli* strain SR101 and the transformed cultures plated on ampicil-

lin media plates. Resistant colonies were selected and grown in the presence of m13KO7 helper bacteriophage to yield secreted, encapsidated single-stranded templates of pRKCD4TP/γ1. The single-stranded plasmid DNA was isolated and used as the template for mutagenesis reactions with the synthetic oligonucleotides described above as primers. The mutagenesis reactions were transformed *E. coli* SR101 and the transformed culture plated on ampicillin media plates. Transformants were screened by colony hybridization (ref. Grunstein-Hogness) for the presence of the appropriate fusion site, using 16mers as probes. These 16mers comprise 8 bases on either side of the fusion site, and the hybridization conditions chosen were sufficiently stringent that the probes only detect the correctly fused product. Colonies identified as positive were selected and plasmid DNA was isolated and transformed into *E. coli* strain SR101. The transformed cultures were plated on ampicillin media plates, and resistant colonies were selected and grown in the presence of m13K07 bacteriophage. Templates were prepared as above and screened by sequencing.

The plasmids were transfected into 293 cells using standard procedures and assayed for expression and production as described above.

	Expressed	Secreted
$pRKCD4_{1\gamma1}$	+	–
$pRKCD4_{e1\gamma1}$	+	+
$pRKCD4_{2\gamma1}$	+	+
$pRKCD4_{e4\gamma1}$	+	+
$pRKCD4_{4\gamma1}$	+	+

Plasmids also were constructed to direct the expression of fusion proteins containing differing lengths of the amino-terminal, extracellular domain of CD4 fused to the truncated portion of the constant region of human immunoglobulin γ1, comprising only the hinge region and constant domains CH_2 and CH_3.

Synthetic oligonucleotides were made as primers for mutagenesis reactions to delete the immunoglobulin sequence from Ser114 to Cys215 inclusive (Kabat et al.). These were synthesized as 48-mers comprising 24 nucleotides on each side of the desired fusion site (i.e., corresponding to the COOH-terminal 8 residues of the desired CD4 moiety, and the NH_2-terminal 8 residues of the desired immunoglobulin moiety). Plasmids $pRKCD4_{4\gamma1}$, $pRKCD4_{2\gamma1}$ and $pRKCD4_{1\gamma1}$ were separately transformed into *E. coli* strain SR101 and the transformed culture plated on ampicillin media plates. Resistant colonies were selected and grown in the presence of m13K07 helper bacteriophage to yield secreted, encapsidated single-stranded templates of these plasmids. The single-stranded plasmid DNA was isolated and used as the template for mutagenesis reactions with the synthetic oligonucleotides described above as primers. The mutagenesis reactions were transformed *E. coli* SR101 and the transformed culture plated on ampicillin media plates. Transformants were screened by colony hybridization (Grunstein-Hoghess) for the presence of the appropriate fusion site, using 16mers as probes. These 16mers comprise 8 bases on either side of the fusion site, and the hybridization conditions chosen were sufficiently stringent that the probes only detect the correctly fused product. Colonies identified as positive were selected and plasmid DNA was isolated and transformed into *E. coli* strain SR101. The transformed cultures were plated on ampicillin media plates, and resistant colonies were selected and grown in the presence of m13KO7 bacteriophage. Templates were prepared as above and screened by sequencing.

The plasmid derived from plasmid $pRKCD4_{4\gamma1}$ is referred to as $pRKCD4_{4Fcl}$, that derived from plasmid $pRKCD4_{2\gamma1}$ is referred to as $pRKCD4_{2Fcl}$ and that derived from plasmid $pRKCD4_{1\gamma1}$ is referred to as $pRKCD4_{1Fcl}$.

$pRKCD4_{2Fcl}$, $pRKCD4_{1Fcl}$ and $pRKCD4_{4Fcl}$ are cultured in the same fashion as described above and CH1-deleted CD4 immunoadhesons recovered as described elsewhere herein.

Light Chain Fusions

Plasmids were constructed to direct the expression of proteins containing differing lengths of the amino terminal, extracellular domain of CD4 fused to the constant region of human immunoglobulin κ. These plasmids are referred to as $pRKCD4_{4\kappa}$, and $pRKCD4_{e4\kappa}$.

Plasmid $pRKCD4_{4\kappa}$contains the portion of the CD4 gene from the initiation codon to the fusion site after the codon for serine residue 366 of the mature CD4 polypeptide, immediately followed by the sequence for the constant region of human immunoglobulin κ, starting at the codon for threonine residue 109 of the mature human immunoglobulin κ. (Kabat et al.)

Plasmid $pRKCD4_{e4\kappa}$contains the portion of the CD4 gene from the initiation codon to the fusion site after the codon for lysine residue 360 of the mature CD4 polypeptide, immediately followed by the sequence for the constant region of human immunoglobulin κ, 2O starting at the codon for threonine residue 109 of the mature human immunoglobulin κ. (Kabat et al.)

These plasmids were constructed in a manner analogous to plasmids $pRKCD4_{4\gamma1}$ and $pRKCD4_{e4\gamma1}$ described above, with the following exception:

The human immunoglobulin κ coding sequence (FIG. 5) was obtained from a human spleen cDNA library (Clontech Laboratories, Inc.) using oligonucleotides based on the published sequence (Hieter, P. A. et al., Cell 22:197–207 [1980]) and an EcoRI-BspMI fragment containing part of the variable region and the entire constant region was obtained (see FIG. 5). This fragment was blunted with Klenow fragment and the four dNTPs. This fragment was used instead of fragment a1, and was used to construct plasmid pRKCD4TP/hκ.

Expression in CHO Cells

Plasmids were or are constructed to direct the expression of the immunoadhesons described above in CHO cells. These are referred to as $pSVeCD4_{4\gamma1}$SVDHFR, $pSVeCD4_{2\gamma1}$SVDHFR, $PSVeCD4_{e1\gamma1}$SVDHFR, $pSVeCD4_{4Fcl}$SVDHFR, $pSVeCD4_{2Fcl}$SVDHFR, $psVeCD4_{el\gamma1}$SVDHFR, $pSVeCD4_{4Fcl}$SVDHFR, $pSVeCD4_{2Fcl}$SVDHFR, $pSVeCD4_{1Fcl}$SVDHFR, $pSVeCD4_{4\kappa}$SVDHFR and $pSVeCD4_{2\kappa}$SVDHFR.

Fragment 31 was prepared as described above. Fragment 32*a* was prepared by digesting plasmid pE348HB-V.E400 D22 with BamHI, blunting with Klenow fragment and the four dNTPs, then digesting with PvuI and isolating the large fragment containing the balance of the β-lactamase gene and the SV40 early promoter and the DHFR gene. Plasmids $pRKCD4_{4\gamma1}$, $pRKCD4_{2\gamma1}$, $pRKCD4_{1\gamma1}$, $pRKCD4_{e4\gamma1}$, $pRKCD4_{4\kappa}$and $pRKCD4_{2\kappa}$were separately digested with HindIII, blunted with Klenow fragment and the four clNTPs,

then digested with EcoRI and the fragments encoding the CD4-Ig fusion protein were isolated. The resulting DNA fragments were ligated together with fragments 31 and 32*a* and transformed into *E. coli* strain 294. Colonies were selected and checked for the presence of the correct plasmid as above, then transfected into CHO cells and amplified by methotrexate selection using conventional procedures.

EXAMPLE 5

Culture Purification and formulation of CD4 variants

Plasmids encoding soluble CD4 adhesons such as CD4T, CD4TP, or soluble CD4 immunoadhesons were calcium phosphate transletted into CHO-DP7 (a proinsulin-transformed autotrine host cell derived from CHO; U.S.S.N. 07,472, abandoned) and the transformants grown in selective medium (1:1 HAM F12/DMEM GHT⁻containing 1–10% diafiltered or dialyzed bovine serum). Other suitable host cells are CHO cells or 293S human embryonic kidney cells. The transformants were amplified by methotrexate selection in the same medium but containing 500 nm methotrexate. A subclone capable of secreting CD4TP, CD4tp 500 b, was selected. CD4tp 500 b is cultured in a DMEM/HAM F12 medium at about 37° C. until CD4TP accumulates in the culture, after which the medium is separated from the cells and insoluble matter by centrifuging.

Culture fluid from CD4TP transformants was concentrated and diafiltered to lower the ionic strength. The concentrate was passed through a large volume of Q-Sepharose anion exchange resin (previously equilibrated with 25 mM NaCl pH 8.5) in order to adsorb contaminants from the culture fluid. The isoelectric point of CD4TP is about 9.5, thus making it possible to discriminate between truncated forms of CD4 and most contaminants by alternate adsorption, respectively, on a cation exchange resin such as carboxymethyl or sulfonyl Sepharose, and an anion exchange resin such as quaternary ammonium Sepharose. In addition, since highly electropositive domains are present in the extracellular segment of CD4 any CD4-containing variant is purified in the same fashion as CD4TP. The unadsorbed culture fluid from the anion exchange resin step was then passed through a cation exchange resin (previously equilibrated with 25 mM NaCl at pH 8.5) whereby CD4TP was adsorbed to the resin. The CD4TP was eluted with a NaCl gradient at pH 8.5, this CD4 variant eluting at about 0.2 M NaCl. Ammonium sulfate was added to the eluate to a concentration of 1.7M and the solution passed through a column of hydrophobic interaction chromatography resin (phenyl or butyl Sepharose). The CD4TP was eluted from the hydrophobic interaction column with a gradient of ammonium sulfate, the CD4TP emerging at about 0.7M ammonium sulfate. The eluate was concentrated and buffer exchanged on a G-25 column using phosphate buffered saline containing 0.02% (w/v) Tween 20 or Tween 80. The CD4TP was soluble and stable in this solution, which was sterile filtered and filled into vials as an aqueous formulation. Other polymeric nontonic surfactants are suitably used with the CD4 formulations, including Pluronic block copolymers or polyethylene glycol.

It is also possible to employ immunoaffinity purification of soluble CD4 wherein the CD4 is adsorbed onto an immobilized antibody against CD4. This method suffers from the disadvantage that elution of the soluble CD4 under acidic conditions leads to protein aggregation that is only thoroughly ameliorated at relatively higher levels of surfactant. The foregoing procedure permits the use of much lower quantities of surfactant, about from 0.01 to 0.10% (w/v) surfactant.

The procedure followed for the purification of CD4 fusions with immunoglobulin heavy chain was to concentrate recombinant supernatants by ultrafiltration and thereafter adsorb the fusion onto resin-immobilized Staphylococcal protein A. The fusion was eluted with 0.1M citrate buffer pH 3 with no salt or detergent. This preparation is buffered into Tris buffer at pH 7.5. The immunoglobulin fusions with CD4 V1–V4 optionally are further purified by the procedure described above for unfused CD4 variants. CD4 immunoglobulin fusions with CD4 V1–V2 also may be purified by the procedure above, except that it is not expected that the isoelectric point of this class of molecules will be as alkaline as that of species containing all four V regions of CD4.

EXAMPLE 6

The characteristics of several adheson variants were determined. As shown in table 4 the immunoadhesons $CD4_{4\gamma1}$ and $CD4_{2\gamma1}$ show improved plasma half-life in rabbits, coupled with high-affinity gp120 binding and an affinity for Fcγ receptor (determined with U937 cells) that is comparable to that of bulk human IgG1.

TABLE 4

	gp120 KD (nM)#	$Fc_\gamma R$ KD (nM)+	Plasma Half-Life++ In Rabbits (Hrs.)
CD4T§	2.3 ± 0.4	Not detected	0.25
$CD4_{4\gamma1}$	1.2 ± 0.1	2.83 ± 0.25	6.4
$CD4_{2\gamma1}$	1.4 ± 0.1	3.01 ± 0.68	40.6
human IgG1	ND**	3.52 ± 0.5	21 days*

*determined in humans
+KD was determined by the method of Anderson et al., "J. Immunol." 125:2735–2741 (1980).
#determined by the method of Smith et al., "Science" 238:1704–07 (1987).
§residues 1–368 only
++The adheson variant was injected intravenously into rabbits and samples of blood were collected periodically and assayed for the presence of the adheson variant.
**Not done.

We claim:

1. A heterofunctional immunoadheson comprising a fusion protein in which a polypeptide comprising a human CD4 antigen variable (V) region is fused at its C-terminus to the N-terminus of a polypeptide comprising a constant region of an immunoglobulin chain disulfide bonded to a companion immunoglobulin heavy chain-light chain pair bearing a V_LV_H antibody combining site capable of binding a predetermined antigen.

2. The immunoadheson of claim 1 selected from the group consisting of AC_H-(V_LC_L-V_HC_H), (AC_L-AC_H)-(V_LC_L-V_HC_H), (AC_L-V_HC_H)-(V_LC_L-V_HC_H), and (V_LC_L-AC_H)-(V_LC_L-V_HC_H), wherein A is a human

CD4 antigen V region, V_L and V_H are the variable domains of an immunoglobulin light and heavy chain, respectively, and C_L and C_H are the constant domains of an immunoglobulin light and heavy chain, respectively.

3. The immunoadheson of claim 1 wherein the V_LV_H antibody combining site is from an anti-ricin antibody.

4. The immunoadheson of claim 1 wherein the immunoglobulin sequences are from IgG-1 or IgG-3 subtypes.

5. The immunoadheson of claim 1 wherein the immunoglobulin sequences are from IgA or IgM.

6. The immunoadheson of claim 1 wherein the fusion protein comprises the V-J domains of the human CD4 antigen.

7. The immunoadheson of claim 1 wherein the human CD4 antigen comprises the sequence Asn-Lys-Val-Val-Leu-Gly-Lys-Lys.

8. The immunoadheson of claim 1 wherein the V-region comprises the amino acid sequence extending about from residue 25 to residue 117 of the precursor of human CD4 in which the initiating Met is 1.

9. A composition comprising a heterofunctional immunoadheson comprising a fusion protein in which a polypeptide comprising a human CD4 antigen variable (V) region is fused at its C-terminus to the N-terminus of a polypeptide comprising a constant region of an immunoglobulin chain disulfide bonded to a companion immunoglobulin heavy chain-light chain pair bearing a V_LV_H antibody combining site capable of binding a predetermined antigen.

10. The composition of claim 9 which is sterile and which further comprises a physiologically acceptable carrier.

11. A method of preparing a heterobifunctional immunoadheson comprising transfecting a host cell expressing nucleic acid encoding an immunoglobulin having a variable region directed against a predetermined antigen, with nucleic acid encoding a fusion protein in which a polypeptide comprising a human CD4 antigen variable (V) region is fused at its C-terminus to the N-terminus of a polypeptide comprising a constant region of an immunoglobulin chain, and culturing the resulting transfected host cell under conditions which permit the expression of said heterobifunctional immunoadheson by said transfected host cell.

12. The method of claim 11 wherein the heterofunctional immunoadheson is recovered from the host cell culture.

* * * * *

EXHIBIT 14

US005565335A

United States Patent [19]

Capon et al.

[11] Patent Number: 5,565,335

[45] Date of Patent: Oct. 15, 1996

[54] ADHESION VARIANTS

[75] Inventors: Daniel J. Capon, San Mateo; Timothy J. Gregory, Hillsborough, both of Calif.

[73] Assignee: Genentech, Inc., South San Francisco, Calif.

[21] Appl. No.: 236,311

[22] Filed: May 2, 1994

Related U.S. Application Data

[63] Continuation of Ser. No. 936,190, Aug. 26, 1992, Pat. No. 5,336,603, which is a continuation of Ser. No. 842,777, Feb. 18, 1992, abandoned, which is a continuation of Ser. No. 250,785, Sep. 28, 1988, abandoned, which is a continuation-in-part of Ser. No. 104,329, Oct. 2, 1987, abandoned.

[51] Int. Cl.[6] C07K 14/705; C07K 14/73; C07K 14/00; C12N 15/62

[52] U.S. Cl. 435/64.7; 435/252.3; 435/320.1; 514/2; 530/350; 530/387.1; 530/387.3; 536/23.4

[58] Field of Search 435/69.7, 252.3, 435320.1; 530/350, 387.1, 387.3; 536/23.4; 514/2

[56] References Cited

U.S. PATENT DOCUMENTS

3,959,080	5/1976	Orth et al.	95/63
4,002,531	1/1977	Royer	195/68
4,055,635	10/1977	Green et al.	424/78
4,301,144	11/1981	Iwashita et al.	424/78
4,412,989	11/1983	Iwashita et al.	424/177
4,444,878	4/1984	Paulus	435/7
4,745,055	5/1988	Schenk et al.	435/7
4,761,371	8/1988	Bell et al.	435/68
4,766,106	8/1988	Katre et al.	514/12
4,816,567	3/1989	Cabilly et al.	530/387
4,879,211	11/1989	Wang et al.	435/5

FOREIGN PATENT DOCUMENTS

068763	1/1983	European Pat. Off. .
088695	9/1983	European Pat. Off. .
120694	10/1984	European Pat. Off. .
139416	5/1985	European Pat. Off. .
173494	3/1986	European Pat. Off. .
244221	11/1987	European Pat. Off. .
255694	2/1988	European Pat. Off. .
256654	2/1988	European Pat. Off. .
266663	5/1988	European Pat. Off. .
278776	8/1988	European Pat. Off. .
296786	12/1988	European Pat. Off. .
313377	4/1989	European Pat. Off. .
319815	6/1989	European Pat. Off. .
325262	7/1989	European Pat. Off. .
WO85/03947	9/1985	WIPO .
WO87/03600	6/1987	WIPO .
WO88/01304	2/1988	WIPO .
WO88/03559	5/1988	WIPO .
WO88/09344	12/1988	WIPO .
WO89/01940	3/1989	WIPO .

OTHER PUBLICATIONS

Miller et al., *Science* 244:334–337, 21 Apr. 1984.

Blank et al., *Nature* 337:187–189, Jan. 12, 1989.

Anderson et al., "Regulatory Interactions Between Members of the Immunoglobulin Superfamily", *Immun. Today*, 9(7&8):199–203 (1988).

Beauchamp et al., "A New Procedure for the Synthesis of Polyethylene Glycol–Protein Adducts; Effects on Function, Receptor Recognition, and Clearance of Superoxide Dismutase, Lactoferrin, and α_2–Macroglobulin", *Analytical Biochem.*, 131:25–33, (1983).

Boulianne et al., "Production of Functional Chimaeric Mouse/Human Antibody", 312:643–646, (1984).

Chaudhary et al., "Activity of a Recombinant Fusion Protein Between Transforming Growth Factor Type α and *Pseudomonas* Toxin", *Proc. Natl. Acad. Sci. USA*, 84:4538–4542, (1987).

Clark et al., "Peptide and Nucleotide Sequences of Rat CD4 (W3/25) Antigen: Evidence for Derivation from a Structure with Four Immunoglobulin–Related Domains", *Proc. Natl. Acad. Sci., USA*, 84:1649–1653, (1987).

Estess et al., "Analysis of T–Cell Receptor Structure and Function Using Chimeric T–Cell Receptor/Immunoglobulin Molecules", *J. Cell. Biochem. Suppl.* (11 Part D), Abs. 331, p. 258, (1987).

Falkner et al., "Expression of Mouse Immunoglobulin Genes in Monkey Cells", *Nature*, 298:286–288, (1982).

Gascoigne et al., "Secretion of a Chimeric T-Cell Receptor-Immunoglobulin Protein", *Proc. Nat. Acad. Sci.*, 84:2936–2940 (1987).

Gascoigne et al., "Secretion of Chimeric T Cell Receptor–Immunoglobulin Fusion Proteins", *J. Cell. Biochem. Suppl.*, (11 Part D), Abs. 333, p. 259, (1987).

Goverman et al., "Chimeric T–Cell Receptor Genes as Tools in Analyzing T–Cell/Target–Cell Interactions" *J. Cell. Biochem. Suppl.*, (11 part D), Abs. 334, p. 259, (1987).

Hashimoto et al., "Rearrangement and Expression of T Cell Receptor Genes in Pre–T Cells", *J. Cell. Biochem. Suppl.*, (11 Part D), Abs. 434, p. 278, (1987).

Ivars et al., "Expression of Cloned T Cell Receptor Genes", *J. Cell. Biochem. Suppl.* (11 Part D), Abs. 435, p. 278, (1987).

Johnson et al., "Correlation of T Cell Receptor Structure and Function", *J. cell. Biochem. Suppl.*, (11 Part D), Abs. 337, p. 260, (1987).

Kohler, "Immunoglobulin Chain Loss in Hybridoma Lines", *Proc. Natl. Acad. Sci., USA*, 77(4):2197–2199, (1980).

Littman, "The Structure of the CD4 and CD8 Genes", *Ann. Rev. Immunol.*, 5:561–584, (1987).

Littman et al., "The Isolation and Sequence of the Gene Encoding T8: A Molecule Defining Functional Classes of T Lymphocytes", *Cell*, 40:237–246, (1985).

Maddon et al., "The Isolation and Nucleotide Sequence of a cDNA Encoding the T Cell Surface Protein T4: A New Member of the Immunoglobulin Gene Family", *Cell*, 42(1):93–104, (1985).

Maddon et al., "The T4 Gene Encodes the AIDS Virus Receptor and Is Expressed in the Immune System and the Brain", *Cell*, 47:333–348, (1986).

Maddon et al., "Structure and Expression of the Human and Mouse T4 Genes", *Proc. Natl. Acad. Sci., USA*, 84:9155–9159, (1987).

McDougal et al., "The T4 Glycoprotein Is A Cell–Surface Receptor for the AIDS Virus" *Molecular Biology of Homosapiens, , Cold Spring Harbor Symposia on Quantitative Biology, vol. LI (CSHL, NY), (1986).*

McDougal et al., "Binding of HTLV–III/LAV to T4 T Cells by a Complex of the 110K Viral Protein and T4 Molecule", *Science*, 231:382–385, (1986).

Morrison et al., "Transfer and Expression of Immunoglobulin Genes", *Ann. Rev. Immunol.*, 2:239–256, (1984).

Morrison, S. L., "Transfectomas Provide Novel Chimeric Antibodies", *Science*, 229:1202–1207, (1985).

Morrison et al., "Chimeric Human Antibody Molecules: Mouse Antigen–Binding Domains with Human Constant Region Domains", *Proc. Natl. Acad. Sci., USA*, 81:6851–6855, (1984).

Morrison, "Sequentially Derived Mutants of the Constant Region of the Heavy Chain of Murine Immunoglobulins" *J. Immunol.*, 123(2):793–800, (1979).

Murre et al., "Biochemical and Functional Analyses of a Secreted H–2L$^{d_{Molecule}}$. *Mol. Cell. Biol.*, 6:1315–1319, (1986).

Neuberger et al., "Recombinant Antibodies Possessing Novel Effector Functions" *Nature*, 312:604–608 (1984).

Oi, V. T., "Chimeric Antibodies", *Biotechniques*, 4:214–221, (1986).

Osborn, et al., "Direct Expression Cloning of Vascular Cell Adhesion Molecule 1, a Cytokine–Induced Endothelial Protein That Binds to Lymphocytes", *Cell*, 59:1203–1211, (1989).

Peterson, A. S. PhD. "Genetic Analysis of CD2/LFA–3 and CD4/HIV Interactions", Thesis Harvard University, Chapter 1 (1988).

Rose, et al., "Expression from Cloned cDNA of Cell–Surface Secreted Forms of the Glycoprotein of Vesicular Stomatitis Virus in Eucaryotic Cells", *Cell*, 30:753–762, (1982).

Rosenblum et al., "Modification of Human Leukocyte Interferon Pharmacology with a Monoclonal Antibody", *Cancer Research*, 45:2421–2424, (1985).

Salzawa et al., "The CD4 Molecule is Associated with the T–Cell Receptor", *J. Cell. Biochem. Suppl.*, (11 Part D), Abs. 421, p. 273, (1987).

Seed, "An LFA–3 cDNA Encodes a Phospholipid–Linked Membrane Protein Homologous to its Receptor CD2", Nature 329:840–842, (1987).

Sharon et al., "Expression of a $V_\lambda c_k$ Chimaeric Protein in Mouse Myeloma Cells", *Nature*, 309:364–367, (1984).

Sleckman et al., "Expression and Function of CD4 in A Murine T–Cell Hybridoma", *Nature*, 328:351–353, (1987).

Smith et al., "Blocking of HIV–1 Infectivity by a Soluble, Secreted Form of the CD4 Antigen", *Science*, 238:1704–1707, (1987).

Springer et al., "Detergent–Soluble HLA Antigens Contain A Hydrophilic Region at the COOH–Terminus and a Penultimate Hydrophobic Region", *Proc. Natl. Acad. Sci., USA*, 73(7):2481–2485, (1976).

Terhorst et al., "Biochemical Analysis of Human and T Lymphocyte Differentiation Antigens T4 and T5", *Science*, 209:520–521, (1980).

Traunecker et al. "Soluble CD4 Molecules Neutralize Human Immunodeficiency Virus Type 1", *Nature, 331:84–86 (1988).*

Traunecker et al., "Highly Efficient Neutralization of HIV with Recombinant CD4–Immunoglobulin Molecules", *Nature, 339:68–70, (1989).*

Wills, et al., "Mutations of the Rous Sarcoma Virus env Gene That Affect the Transport and Subcellular Location of the Glycoprotein Products", *J. Cell Bio.*, 99:2011–2023, (1984).

Vitetta et al., "Redesigning Nature's Poisons to Create Anti–Tumor Reagents", *Science*, 238:1098–1104, (1987).

Allaway et al., "Expression and Characterization of CD4–IgG$_2$, a Novel Heterotetramer That Neutralizes Primary HIV Type 1 Isolates" *AIDS Research and Retroviruses* 11(5):533–539 (1995).

Aruffo et al., "CD44 is the principal cell surface receptor for hyaluronate" *Cell* 61:1303–1313 (1990).

Byrn et al., "Biological properties of a CD4 immunoadhesin" *Nature* 344:667–670 (1990).

Capon et al., "Designing CD4 immunoadhesins for AIDS therapy" *Nature* 337:525–531 (1989).

Chamow et al., "A Humanized, Bispecific Immunoashesion–Antibody That Retargets CD3 Effectors to Kill HIV–1–Infected Cells" *Journal of Immunology* 153:4268–4280 (1994).

Clerici et al., "Effect of a Recombinant CD4–IgG on In Vitro T Helper Cell Function: Data From a Phase I/II Study of Patients with AIDS" *The Journal of Infectious Diseases* 168:1012–1016 (1993).

Haak–Frendscho et al., "Human IgE receptor alpha chain–IgG chimera blocks passive cutaneous anaphylaxis reaction in vivo" *J. Immunol.* 151:351 (1993).

Hakimi et al., "The α subunit of the human IgE receptor (FcERI) is sufficient for high affinity IgE binding" *Journal of Biological Chemistry* 265(36):22079–22081 (1990).

Kahn et al., "A Phase 1 Study of Recombinant Human CD4 Immunoglobulin G (rCD4–IgG) In Patients With HIV–Associated Immune Thrombocytopenic Purpura" (W.B. 2156) p. 221 (1991).
Linsley et al., "Binding of the B cell activation antigen B7 to CD28 Costimulates T cell proliferation and interleukin 2 mRNA accumulation" *Journal of Experimental Medicine* 173:721–731 (1991).
Linsley et al., "CTLA–4 is a second receptor for the B cell activation antigen B7" *Journal of Experimental Medicine* 174:561–569 (1991).
Schacker et al., "The Effects of High–Dose Recombinant Soluble CD4 on Human Ommunodeficiency Virus Type 1 Viremia" *The Journal of Infectious Diseases* 169:37–40 (1994).
Shearer et al., "Transport of Recombinant Human C–Immunoglobulin G Across the Human Placenta: Pharmacokinetics and Safety in Six Mother–Infant Pairs in AIDS Clinical Trial Group Protocol 146" *Clinical and Diagnostic Laboratory Immunology* pp. 281–285 (1995).
Stamenkovic et al., "The B lymphocyte adhesion molecule CD22 interacts with leukocyte common antigen CD45RO on T cells and α2–6 sialyltransferase, CD75, on B cells" *Cell* 66:1133–1144.
Ushijima et al., "Synergistic Effect of Recombinant CD4–Immunoglobulin in Combination with Azidothymidine, Dideooxyinosine and 0.5 β–monoclonal Antibody on Human Immunodeficiency Virus infection in vitro" *Letters in Applied Microbiology*, 19:1–5 (1994).
Ward et al., "Prevention of HIV–1 IIIB Infection in Chimpanzees by CD4 Immunoadhesin" *Nature* 352:434–436 (1991).

Primary Examiner—John Ulm
Attorney, Agent, or Firm—Jeffrey S. Kubinec

[57] **ABSTRACT**

Novel derivatives of cell surface proteins which are homologous to the immunoglobulin superfamily (adhesons) are provided. Amino acid sequence variations are introduced into the adheson, the most noteworthy of which are those in which the transmembrane and, preferably, cytoplasmic domains are rendered functionally inactive, and in which adheson extracellular domains replace an immunoglobulin variable region. These variants are useful in therapy or diagnostics, in particular, CD4 variants are therapeutically useful in the treatment of HIV infections.

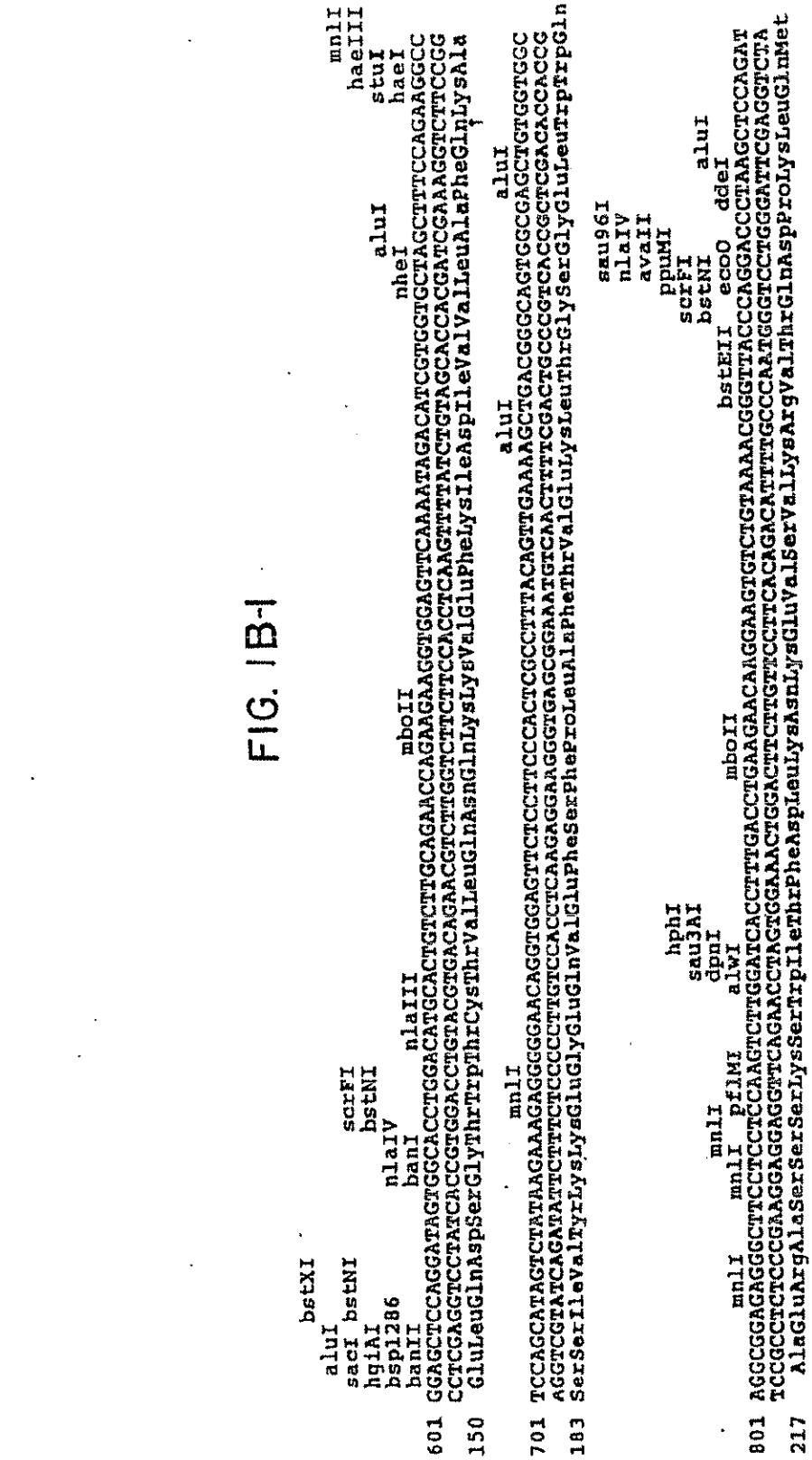
18 Claims, 13 Drawing Sheets

FIG. 1A



```
                                                sau96I
                                                nlaIV
                                                avaII
                                               ppuMI                                                                  scrFI
                                            ddeI                                                  mnlI                  nciI pleI
                 bsp1286                 bsp1286                                                  ecoNI    haeIII       mspI hinfI
                 banII   bstXI           banII ecoO          ddeI          mnlI          haeI                 hpaII
  1 AATTCAAGCCCAGAGCCCTGCCATTTCTGTGGGCTCAGGTCCCTACTGCTCAGCCCCTTCCTCCCTCGGCAAGGCCACAATGAACCGGGGAGTCCCTTTT
    TTAAGTTCGGGTCTCGGGACGGTAAAGACACCCGAGTCCAGGGATGACGAGTCGGGGAAGGAGGGAGCCGTTCCGGTGTTACTTGGCCCCTCAGGGAAAA
-25                                                                                                  MetAsnArgGlyValProPhe
                                  hinPI
                      fnu4HI      hhaI          fnu4HI
                      bbvI       haeII mnlI     bbvI    ddeI
101 AGGCACTTGCTTCTGGTGCTGCAACTGGCGCTCCTCCCAGCAGCCACTCAGGGAAACAAAGTGGTGCTGGGCAAAAAAGGGGATACAGTGGAACTGACCT
    TCCGTGAACGAAGACCACGACGTTGACCGCGAGGAGGGTCGTCGGTGAGTCCCTTTGTTTCACCACGACCCGTTTTTTCCCCTATGTCACCTTGACTGGA
-18 ArgHisLeuLeuLeuValLeuGlnLeuAlaLeuLeuProAlaAlaThrGlnGlyAsnLysValValLeuGlyLysLysGlyAspThrValGluLeuThrCys
                                                                                                  nlaIV                  fokI
                                                                                                 bsp1286              sau96I
        aluI          mboII                                                                 hinfI          banII     mseI     avaII
    rsaI          mboII
201 GTACAGCTTCCCAGAAGAAGAGCATACAATTCCACTGGAAAAACTCCAACCAGATAAAGATTCTGGGAAATCAGGGCTCCTTCTTAACTAAAGGTCCATC
    CATGTCGAAGGGTCTTCTTCTCGTATGTTAAGGTGACCTTTTTGAGGTTGGTCTATTTCTAAGACCCTTTAGTCCCGAGGAAGAATTGATTTCCAGGTAG
 17   ThrAlaSerGlnLysLysSerIleGlnPheHisTrpLysAsnSerAsnGlnIleLysIleLeuGlyAsnGlnGlySerPheLeuThrLysGlyProSer

                 hinPI                                                                                                ddeI
                 hhaI                                  styI                                                          pleI
                thaI                                 sau96I                  sau3AI          mseI                    hinfI
              sau3AI    pleI                         avaII                   dpnI           aflII
      aluI    dpnI      hinfI   mboII               nlaIV                   bclI       hinfI          mboII
301 CAAGCTGAATGATCGCGCTGACTCAAGAAGAAGCCTTTGGGACCAAGGAAACTTTCCCCTGATCATCAAGAATCTTAAGATAGAAGACTCAGATACTTAC
    GTTCGACTTACTAGCGCGACTGAGTTCTTCTTCGGAAACCCTGGTTCCTTTGAAAGGGGACTAGTAGTTCTTAGAATTCTATCTTCTGAGTCTATGAATG
 50  LysLeuAsnAspArgAlaAspSerArgArgSerLeuTrpAspGlnGlyAsnPheProLeuIleIleLysAsnLeuLysIleGluAspSerAspThrTyr

                 sau96I
                 avaII          mnlI                                                                  ecoNI
                mnlI          mnlI                                                                  bspMI          alwNI
401 ATCTGTGAAGTGGAGGACCAGAAGGAGGAGGTGCAATTGCTAGTGTTCGGATTGACTGCCAACTCTGACACCCACCTGCTTCAGGGGCAGAGCCTGACCC
    TAGACACTTCACCTCCTGGTCTTCCTCCTCCACGTTAACGATCACAAGCCTAACTGACGGTTGAGACTGTGGGTGGACGAAGTCCCCGTCTCGGACTGGG
 83 IleCysGluValGluAspGlnLysGluGluValGlnLeuLeuValPheGlyLeuThrAlaAsnSerAspThrHisLeuLeuGlnGlyGlnSerLeuThrLeu
                                                                                                                                aluI
                      scrFI                                  styI                                                                pvuII
              bsp1286                 ddeI                 pleI                                                          mboII mnlI      ddeI
      styI    banII   bstNI          mnlI                 hinfI
501 TGACCTTGGAGAGCCCCCCTGGTAGTAGCCCCTCAGTGCAATGTAGGAGTCCAAGGGGTAAAAACATACAGGGGGGGAAGACCCTCTCCGTGTCTCAGCT
    ACTGGAACCTCTCGGGGGGACCATCATCGGGGAGTCACGTTACATCCTCAGGTTCCCCATTTTTGTATGTCCCCCCCTTCTGGGAGAGGCACAGAGTCGA
117   ThrLeuGluSerProProGlySerSerProSerValGlnCysArgSerProArgGlyLysAsnIleGlnGlyGlyLysThrLeuSerValSerGlnLeu
```

FIG. 1B-1


bstXI
aluI
sacI bstNI
hgiAI
bsp1286
banII
601 GGAGCTCCAGGATAGTGGCACCTGGACATGCACTGTCTTGCAGAACCAGAAGAGGTGGAGTTCAAAATAGACATCGTGGTGCTAGCTTTCCAGAAGGCC
CCTCGAGGTCCTATCACCGTGGACCTGTACGTGACACGAACGTCTTGGTCTTCTTCCACCTCAAGTTTTATCTGTAGCACCACGATCGAAAGGTCTTCCGG
150 GluLeuGlnAspSerGlyThrTrpThrCysThrValLeuGlnAsnGlnLysLysValGluGluPheLysIleAspIleValValLeuAlaPheGlnLysAla
↑
scrFI
bstNI
nlaIV
banI
nlaIII
mboII
aluI
mnlI
701 TCCAGCATAGTCTATAAGAAAGAGGGGGAACAGGTGGAGTTCTCCTTCCCACTCGCCTTTACAGTTGAAAAGCTGACGGGCAGTGGCGAGCTGTGGTGGC
AGGTCGTATCAGATATTCTTTCTCCCCCTTGTCCACCTCAAGAGGGAAGGGTGAGCGGAAATGTCAACTTTTCGACTGCCCGTCACCGCTCGACACCACCG
183 SerSerIleValTyrLysLysGluGlyGluGlnValGluGluPheSerPheProLeuAlaPheThrValGluLysLysLeuThrGlySerGlyGluLeuTrpTrpGln
sau96I
nlaIV
avaII
ppuMI
scrFI
bstNI
aluI
bstEII ecoO ddeI
hphI
sau3AI
dpnI
mboII
mnlI pflMI alwI
mnlI mnlI
801 AGGCGGAGAGGGCTTCCTCCTCCAAGTCTTGGATCACCTTGACCTGAAGAACAAGGAAGTGTCTGTAAAACGGGTTACCCAGGACCCTAAGCTCCAGAT
TCCGCCTCTCCCGAAGGAGGTTCAGAACCTAGTGGAAACTGGACTTCTTGTTCCTTCACAGACATTTTGCCCAATGGGTCCTGGGATTCGAGGTCTA
217 AlaGluArgAlaSerSerSerLysSerTrpIleThrPheAspLeuLysAsnLysGluValSerValLysArgValThrGlnAspProLysLeuGlnMet

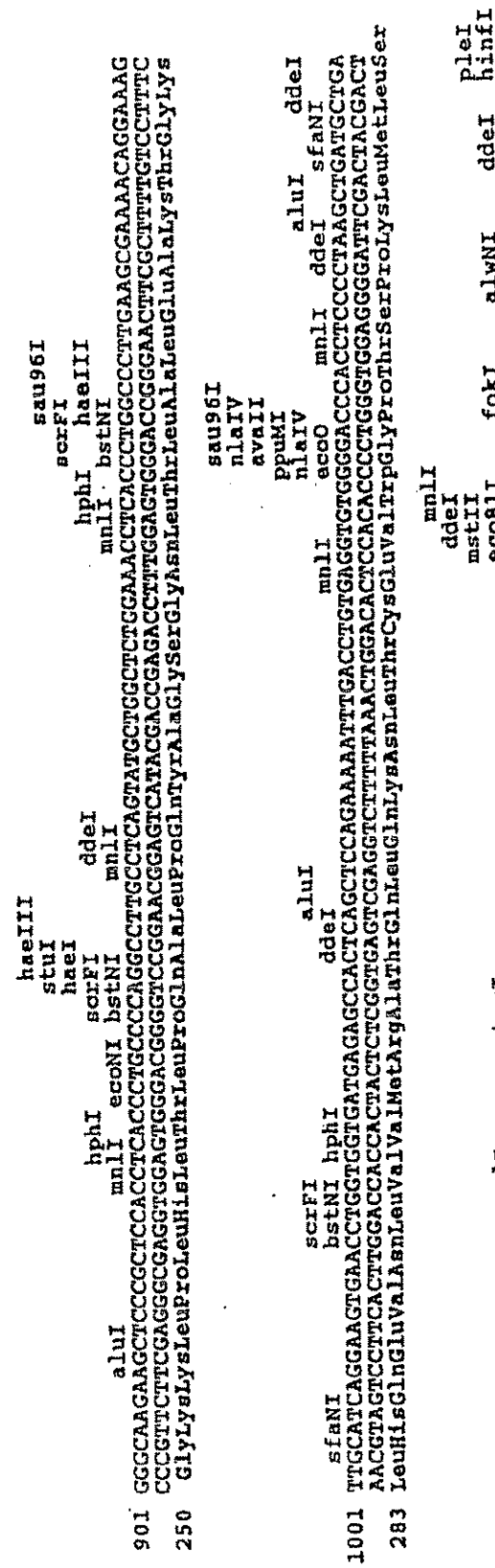
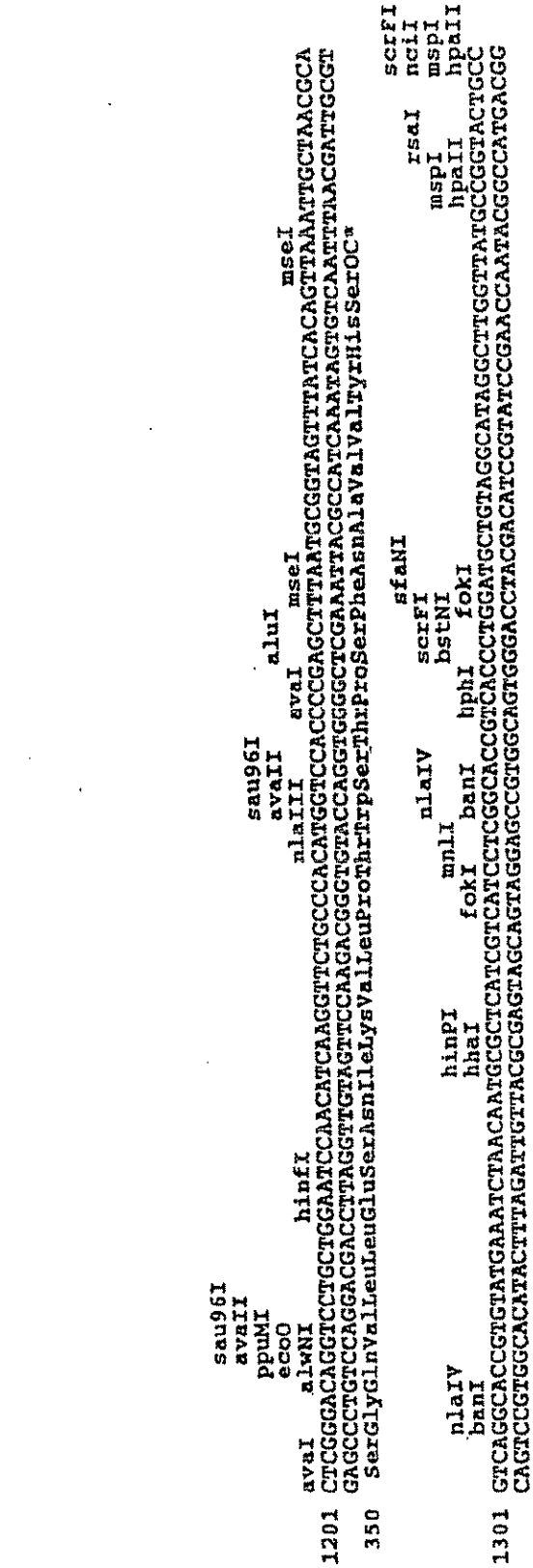

sau96I
scrFI
hphI haeIII
mnlI bstNI
ddeI
mnlI
haeIII
stuI
haeI
scrFI
bstNI
hphI
mnlI ecoNI bstNI
aluI
901 GGGCAAGAAGCTCCCGCTCCACCTCACCTGCCTTGCCTCAGTATGCTGGCTCTGGCAAACCTCACCCTGGCCCTTGAAGCGAAAACAGGAAAG
250 GlyLysLysLeuProLeuHisLeuThrLeuProGlnTyrAlaGlySerGlyAsnLeuThrLeuAlaLeuGluAlaLysThrGlyLysLys
sau96I
nlaIV
avaII
ppuMI
nlaIV
mnlI
aluI
ddeI
eco0
ddeI
sfaNI
aluI
mnlI
scrFI
bstNI hphI
ddeI
sfaNI
1001 TTGCATCAGGAAGTGAACCTGGTGGTGATGAGAGCCACTCAGCTCCAGAAAAATTTGACCTGTGAGGTGTGGGGACCCACCTCCCCTAAGCTGATGCTGA
283 LeuHisGlnGluValAsnLeuValValMetArgAlaThrGlnLeuGlnLysAsnLeuThrCysGluValTrpGlyProThrSerProLysLeuMetLeuSer
mnlI
ddeI
mstII
eco81I
fokI
alwNI
ddeI
pleI
hinfI
mnlI
taqI
1101 GTTTGAAACTGGAGAACAAGGAGGCAAAGGTCTCGAAGCGGGAGAAGGCGGTGTGGGTGCTGAACCCTGAGGCGGGGATGTGGCAGTGTCTGCTGAGTGA
317 LeuLysLeuGluAsnLysGluAlaLysValSerLysArgGluLysAlaValTrpValLeuAsnProGluAlaGlyMetTrpGlnCysLeuLeuSerAsp

FIG. 1B-2

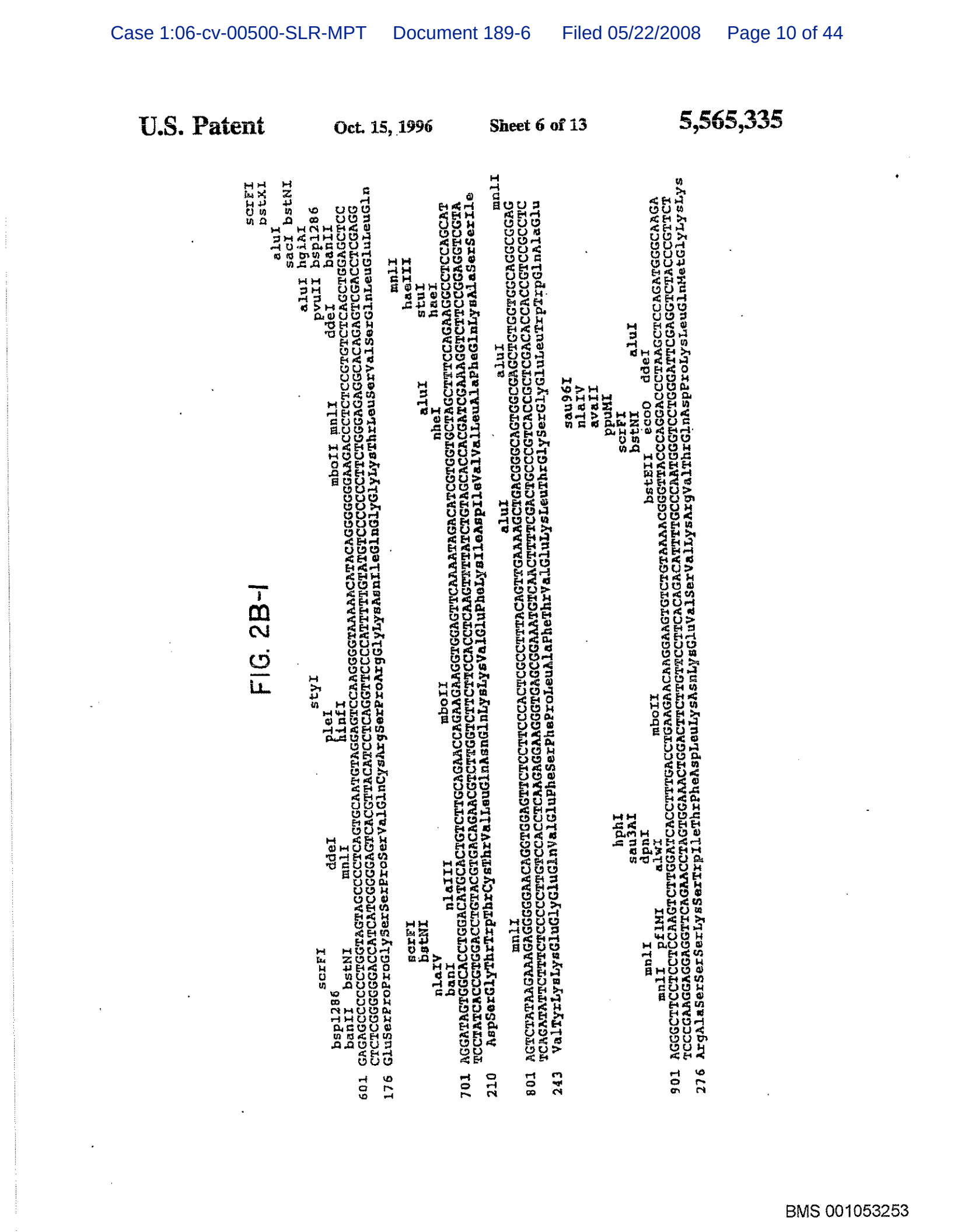


```
                sau96I
                avaII
               ppuMI                                    sau96I
               ecoO                                     avaII      aluI
      avaI   alwNI       hinfI                       nlaIII    avaI    mseI                    mseI
1201  CTCGGGACAGGTCCTGCTGGAATCCAACATCAAGGTTCTGCCCACATGGTCCACCCCGAGCTTTAATGCGGTAGTTTATCACAGTTAAATTGCTAACGCA
      GAGCCCTGTCCAGGACGACCTTAGGTTGTAGTTCCAAGACGGGTGTACCAGGTGGGGCTCGAAATTACGCCATCAAATAGTGTCAATTTAACGATTGCGT
 350   SerGlyGlnValLeuLeuGluSerAsnIleLysValLeuProThrTrpSerThrProSerPheAsnAlaValValTyrHisSerOC*
                                                                                                  scrFI
                                                                  sfaNI
                                                          nlaIV      scrFI                          rsaI  nciI
                                                                                                 mspI    mspI
          nlaIV                    hinPI         mnlI                bstNI                        hpaII   hpaII
          banI                     hhaI       fokI   banI   hphI   fokI
1301  GTCAGGCACCGTGTATGAAATCTAACAATGCGCTCATCGTCATCCTCGGCACCGTCACCCTGGATGCTGTAGGCATAGGCTTGGTTATGCCGGTACTGCC
      CAGTCCGTGGCACATACTTTAGATTGTTACGCGAGTAGCAGTAGGAGCCGTGGCAGTGGGACCTACGACATCCGTATCCGAACCAATACGGCCATGACGG

         mnlI
       haeIII
      sau96I
1401  GGGCCTCTTGCGGGAT
      CCCGGAGAACGCCCTA
```

FIG. 1C

FIG. 2A

FIG. 2B-1


scrFI
bstXI
aluI
sacI bstNI
aluI hgiAI
pvuII bsp1286
ddeI banII
mboII mnlI
mnlI
haeIII
stuI
haeI
aluI
nheI
aluI
mnlI
aluI
sau96I
nlaIV
avaII
ppuMI
scrFI
aluI
bstNI ddeI
bstEII ecoO
hphI
sau3AI
dpnI
mboII
mnlI pflMI alwI
scrFI
bstNI
nlaIV
banI
nlaIII
mboII
aluI
ddeI
pleI
hinfI
styI
bsp1286
banII bstNI
mnlI
scrFI
bstNI
mnlI
601 GAGAGCCCCCCCTGGTAGTAGCCCCCTCAGTGCAATGTAGGAGTCCAAGGGGTAAAAACATACAGGGGGGGAAGACCCTCTCCGTGTCTCAGCTGGAGCTCC
CTCTCGGGGGGACCATCATCGGGGAGTCACGTTACATCCTCAGGTTCCCCCATTTTTGTATGTCCCCCCCTCTGGGAGAGGCACAGAGTCGACCTCGAGG
176 GluSerProProGlySerSerProSerProArgGlyGlyLysAsnIleGlnGlyGlyLysThrLeuSerValSerGlnLeuGluLeuGlu
701 AGGATAGTGGCACCTGGACATGCACTGTCTTGCAGAACCAGAAGAGGTGGAGTTCAAAATAGACATCGTGGTGCTAGCTTTCCAGAAGGCCTCTCAGCAT
TCCTATCACCGTGGACCTGTACGTGACAGAACGTCTTGGTCTTCTCCACCTCAAGTTTTATCTGTAGCACCACGATCGAAAGGTCTTCCGGAGGTCGTA
210 AspSerGlyThrTrpThrCysThrValLeuGlnAsnGlnLysLysValGluPheLysIleAspIleValValLeuAlaPheGlnLysAlaSerSerIle
801 AGTCTATAAGAAAGAGGGGGAACAGGTGGAGTTCTCCTTCCCACTCGCCTTTACAGTTGAAAAGCTGACGGGCAGTGGCGAGCTGGTGGTGGCAGGCGGAG
TCAGATATTCTTTCTCCCCCTTGTCCACCTCAAGAGGAAGGGTGAGCGGAAATGTCAACTTTTCGACTGCCCGTCACCGCTCGACACCACCGTCCGCCTC
243 ValTyrLysLysGluGlyGluGlnValGluPheSerPheProLeuAlaPheThrValGluLysLeuThrGlySerGlyGluLeuTrpTrpGlnAlaGlu
901 AGGGCTTCCTCCTCCAAGTCTTGGATCACCTTTGACCTGAAGAACAAGGAAGTGTCTGTAAAACGGGTTACCCAGGACCCTAAGCCTCCAGATGGGCAAGA
TCCCGAAGGAGGAGGTTCAGAACCTAGTGGAAACTGGACTTCTTGTTCCCTTCACAGACATTTTGCCCAATGGGTCCTGGGATTCGAGGTCTACCCGTTCT
276 ArgAlaSerSerSerLysSerTrpIleThrPheAspLeuLysAsnLysGluValSerValLysArgValThrGlnAspProLysLeuGlnMetGlyLysLys

```
                                 haeIII
                                stuI                                      sau96I
                                haeI                                   scrFI
                    hphI      scrFI      ddeI                      hphI   haeIII
     aluI         mnlI  ecoNI bstNI     mnlI                    mnlI  bstNI                          sfaNI
1001 AGCTCCCGCTCCACCTCACCCTGCCCCAGGCCTTGCCTCAGTATGCTGGCTCTGGAAACCTCACCCTGGCCCTTGAAGCGAAAACAGGAAAGTTGCATCA
     TCGAGGGCGAGGTGGAGTGGGACGGGGTCCGGAACGGAGTCATACGACCGAGACCTTTGGAGTGGGACCGGGAACTTCGCTTTTGTCCTTTCAACGTAGT
 310  LeuProLeuHisLeuThrLeuProGlnAlaLeuProGlnTyrAlaGlySerGlyAsnLeuThrLeuAlaLeuGluAlaLysThrGlyLysLeuHisGln

                                                                      sau96I
                                                                      nlaIV
                                                                      avaII
                                                                     ppuMI
              scrFI                  aluI                            nlaIV            aluI    ddeI
              bstNI hphI           ddeI                     mnlI    ecoO    mnlI  ddeI   sfaNI
1101 GGAAGTGAACCTGGTGGTGATGAGAGCCACTCAGCTCCAGAAAAATTTGACCTGTGAGGTGTGGGGACCCACCTCCCCTAAGCTGATGCTGAGTTTGAAA
     CCTTCACTTGGACCACCACTACTCTCGGTGAGTCGAGGTCTTTTTAAACTGGACACTCCACACCCCTGGGTGGAGGGGATTCGACTACGACTCAAACTTT
 343  GluValAsnLeuValValMetArgAlaThrGlnLeuGlnLysAsnLeuThrCysGluValTrpGlyProThrSerProLysLeuMetLeuSerLeuLys
```

FIG. 2B-2

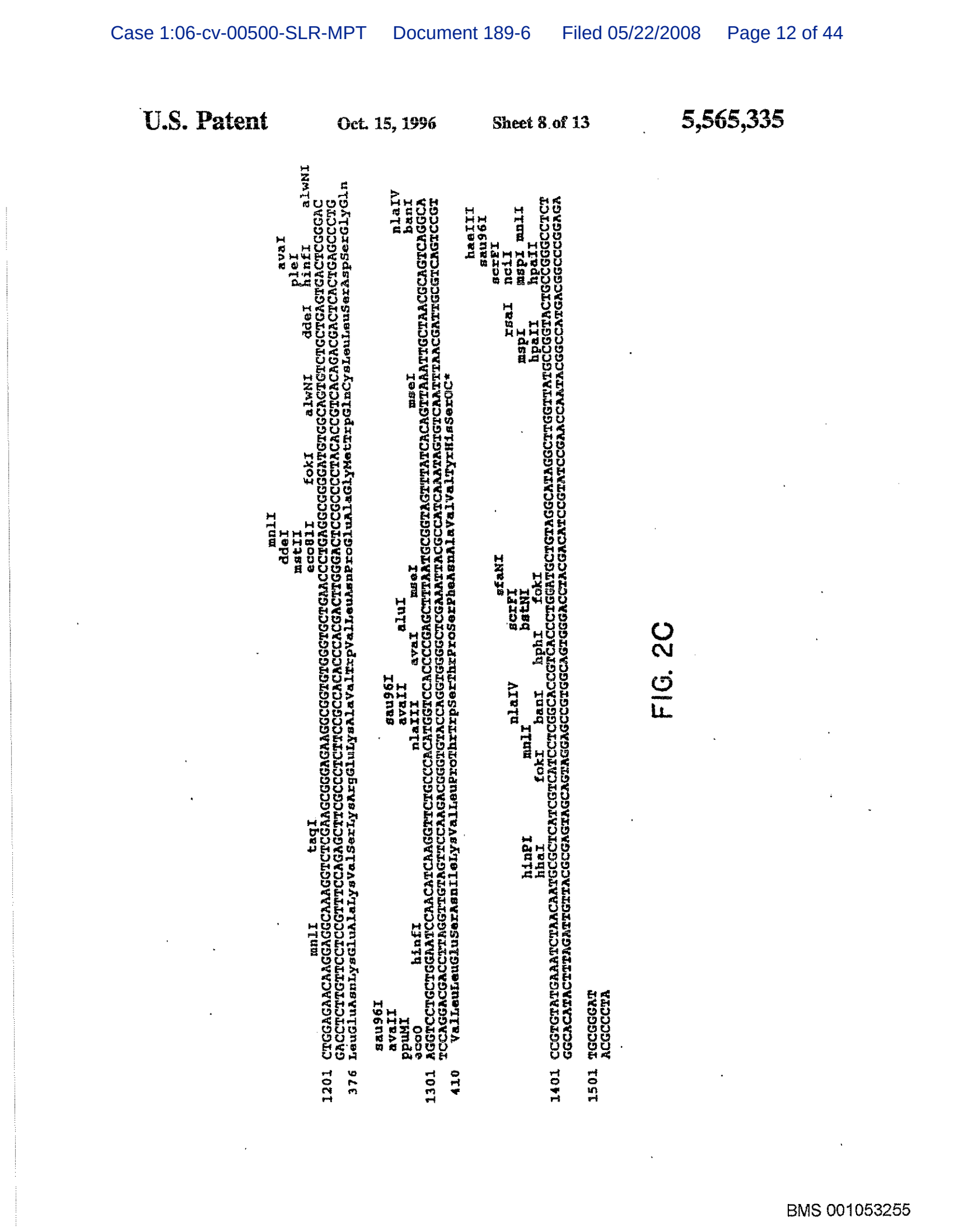


FIG. 2C

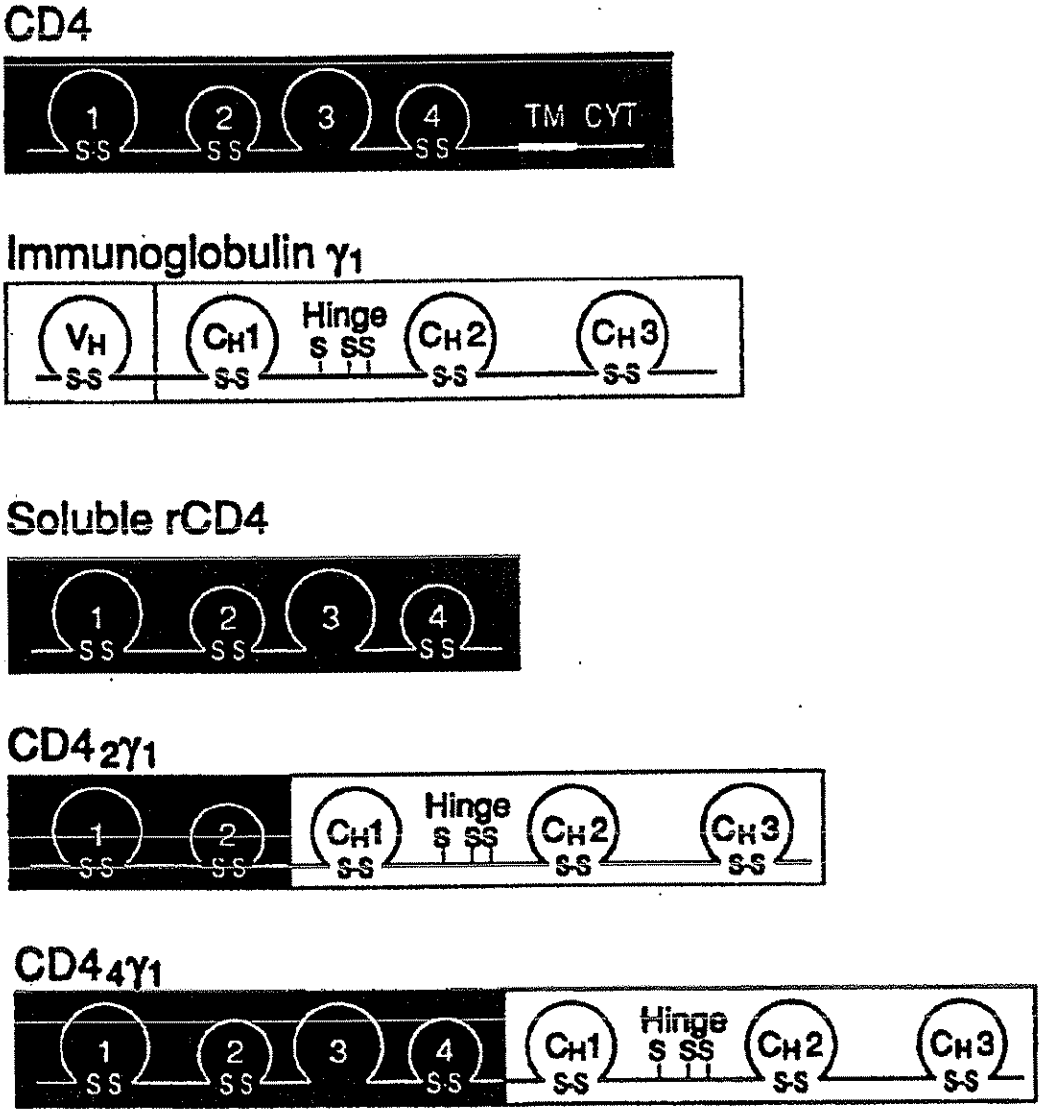

CD4
1
2
3
4
TM CYT
Immunoglobulin γ1
VH
CH1
Hinge
CH2
CH3
Soluble rCD4
CD4 2γ1
CD4 4γ1

FIG. 3

FIG. 4A

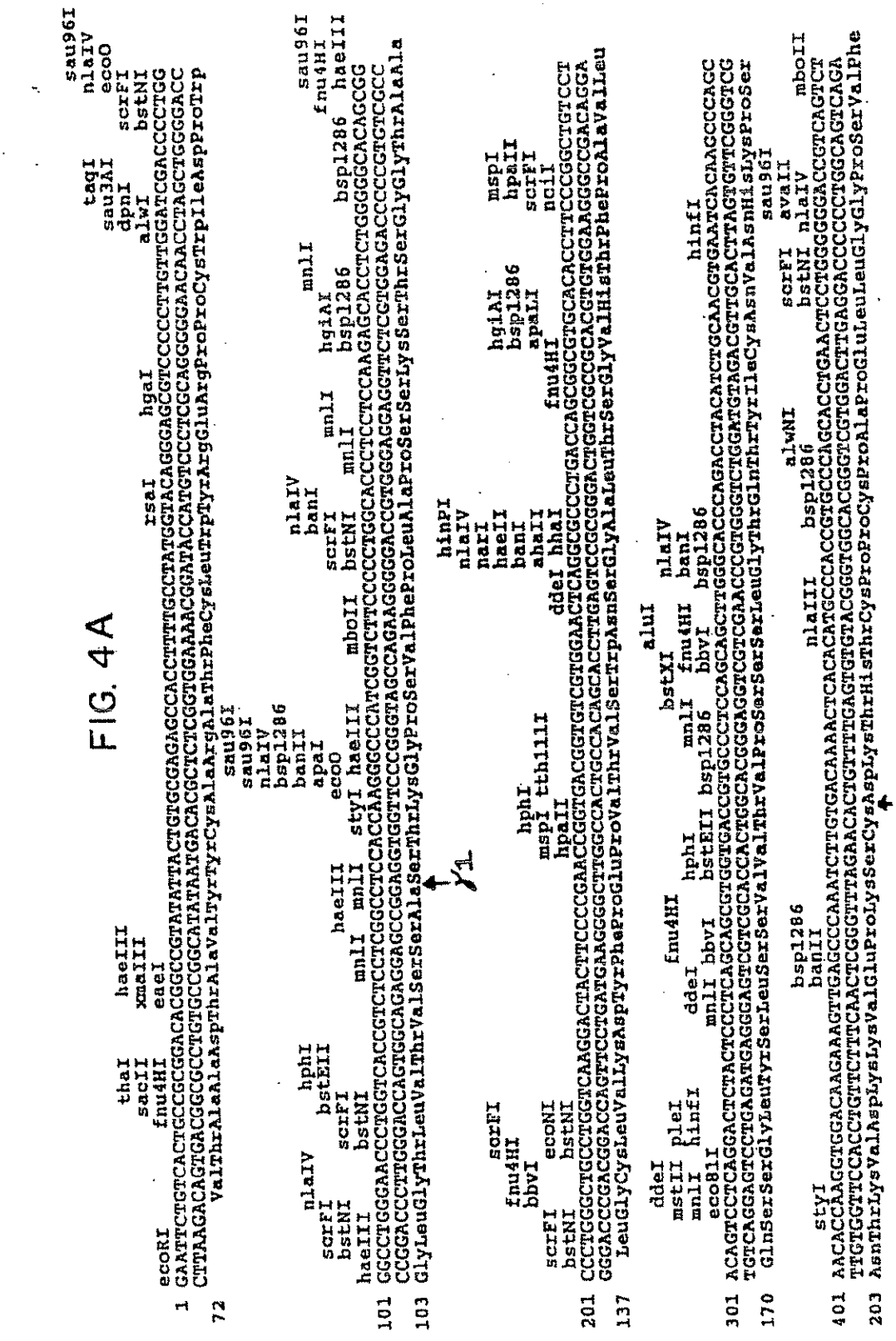


FIG. 4B-1

```
                                                sau96I
                                                nlaIV
                                               mspI
                                          sau3AI  avaII   mnlI                                                  mnlI
                                        nlaIII   hpaII   ddeI                                                 ddeI
                                       bspHI    scrFI   mstII                                                mstII
      mboII                           mnlI  dpnI nciI    eco81I    nlaIII                              mboII eco81I
     mnlI                styI                                                                                  mnlI
501 TCCTCTTCCCCCCAAAACCCAAGGACACCCTCATGATCTCCCGGACCCCTGAGGTCACATGCGTGGTGGTGGACGTGAGCCACGAAGACCCTGAGGTCAA
    AGGAGAAGGGGGGTTTTGGGTTCCTGTGGGAGTACTAGAGGGCCTGGGGACTCCAGTGTACGCACCACCACCTGCACTCGGTGCTTCTGGGACTCCAGTT
237   LeuPheProProLysProLysAspThrLeuMetIleSerArgThrProGluValThrCysValValValAspValSerHisGluAspProGluValLys

                                                                                                scrFI
                                                                                                nciI
                                                                                                mspI
                                                                thaI                            hpaII
                                                               sacII                                          mnlI
                                                              fnu4HI mnlI    rsaI          rsaI          hgaI  hphI
            rsaI            mnlI
601 GTTCAACTGGTACGTGGACGGCGTGGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAGCAGTACAACAGCACGTACCGGGTGGTCAGCGTCCTCACC
    CAAGTTGACCATGCACCTGCCGCACCTCCACGTATTACGGTTCTGTTTCGGCGCCCTCCTCGTCATGTTGTCGTGCATGGCCCACCAGTCGCAGGAGTGG
270  PheAsnTrpTyrValAspGlyValGluValHisAsnAlaLysThrLysProArgGluGluGlnTyrAsnSerThrTyrArgValValSerValLeuThr

         scrFI
      ecoNI bstNI                   rsaI                                  mnlI          taqI
701 GTCCTGCACCAGGACTGGCTGAATGGCAAGGAGTACAAGTGCAAGGTCTCCAACAAAGCCCTCCCAGCCCCCATCGAGAAAACCATCTCCAAAGCCAAAG
    CAGGACGTGGTCCTGACCGACTTACCGTTCCTCATGTTCACGTTCCAGAGGTTGTTTCGGGAGGGTCGGGGGTAGCTCTTTTGGTAGAGGTTTCGGTTTC
303 ValLeuHisGlnAspTrpLeuAsnGlyLysGluTyrLysCysLysValSerAsnLysAlaLeuProAlaProIleGluLysThrIleSerLysAlaLysGly
```

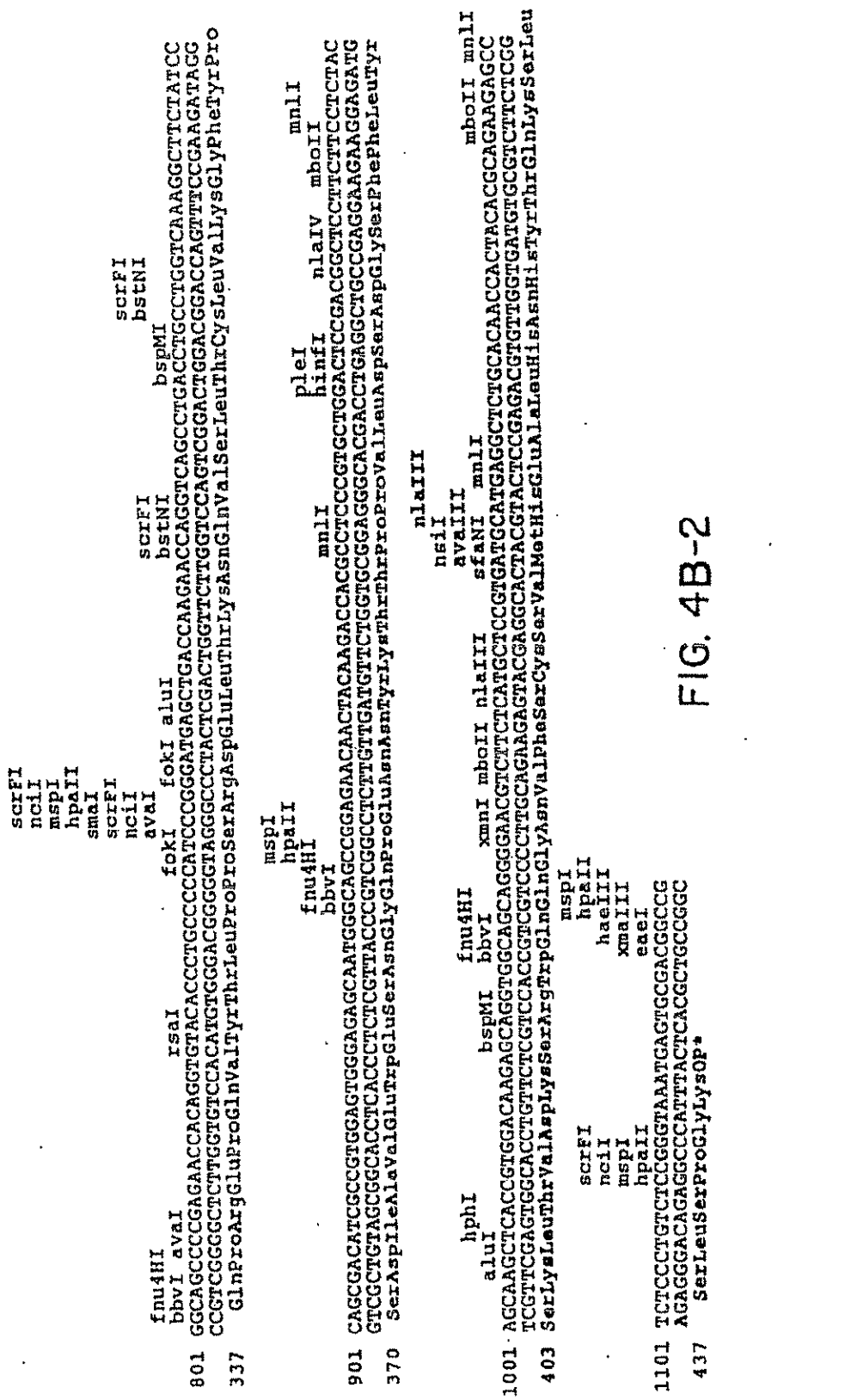

scrFI
ncil
mspI
hpaII
smaI
scrFI
ncil
aval
fokI
aluI
bstNI
bspMI
scrFI
bstNI
fnu4HI
bbvI aval
rsaI
801 GGCAGCCCCGAGAACCACAGGTGTACACCCTGCCCCCATCCCGGGATGAGCTGACCAAGAACCAGGTCAGCCTGACCTGCCTGGTCAAAGGCTTCTATCC
CCGTCGGGGGCTCTTGGTGTCCACATGTGGGACGGGGGTAGGGCCCTACTCGACTGGTTCTTGGTCCAGTCGGACTGGACGGGACCAGTTTCCGAAGATAGG
337 GlnProArgGluProGlnValTyrThrLeuProProSerArgAspGluLeuThrLysAsnGlnValSerLeuThrCysLeuValLysGlyPheTyrPro
mspI
hpaII
fnu4HI
bbvI
pleI
hinfI
mnlI
nlaIV mboII
mnlI
901 CAGCGACATCGCCGTGGAGTGGGAGAGCAATGGGCAGCCGGAGAACAACTACAAGACCACGCCTCCCGTGCTGGACTCCGACGGCTCCTTCTTCCTCTAC
GTCGCTGTAGCGGCACCTCACCCTCTCGTTACCCGGTCGGCCTCTTGTTGATGTTCTGGTGCGGAGGGCACGACCTGAGGCTGCCGAGGAAGAAGGAGATG
370 SerAspIleAlaValGluTrpGluSerAsnGlyGlnProGluAsnAsnTyrLysThrThrProProValLeuAspSerAspGlySerPhePheLeuTyr
nlaIII
nsiI
avaIII
fnu4HI
bbvI
sfaNI mnlI
mboII mnlI
1001 AGCAAGCTCACCGTGGACAAGAGCAGGTGGCAGCAGGGGAACGTCTTCTCATGCTCCGTGATGCATGAGGCTCTGCACAACCACTACACGCAGAAGAGCCC
TCGTTCGAGTGGCACCTGTTCTCGTCCACCGTCGTACAAGAGTACGAGGCACTACGTACTCCGAGACGTGTTGGTGATGTGCGTCTTCTCGG
403 SerLysLeuThrValAspLysSerArgTrpGlnGlnGlyAsnValPheSerCysSerValMetHisGluAlaLeuHisAsnHisTyrThrGlnLysSerLeu
hphI
aluI
bspMI
fnu4HI
bbvI
mspI
hpaII
haeIII
xmaIII
eaeI
scrFI
ncil
mspI
hpaII
1101 TCTCCCTGTCTCCGGGTAAATGAGTGCGACGGCCG
AGAGGGACAGAGGCCCATTTACTCACGCTGCCGGC
437 SerLeuSerProGlyLysOP*

FIG. 4B-2

sau96I
avaII
nlaIV
mnlI styI
fnu4HI
pstI
haeIII
ecoRI hphI fnu4HI bbvI mboII
1 GAATTCACTCTCACCATCAGCGGCCTGCAGCCTGAAGATTTTGCAACTTATTACTGCCAACAGTATAAGAGTTTGTCGCTCACTTTCGGCGGAGGGACCA
72 CTTAAGTGAGAGTGGTAGTCGCCGGACGTCGGACTTCTAAAACGTTGAATAATGACGGTTGTCATATTCTCAAACAGCGAGTGAAAGCCGCCTCCCTGGT
ThrLeuThrIleSerGlyLeuGlnProGluAspPheAlaThrTyrTyrCysGlnGlnTyrLysSerLeuSerLeuThrPheGlyGlyGlyThrLys
sau3AI fnu4HI mboII
dpnI bbvI mboII mnlI
101 AGGTGGAGATCAAACGAACTGTGGCTGCACCATCTGTCTTCATCTTCCCGCCATCTGATGAGCAGTTGAAATCTGGAACTGCCTCTGTTGTGTGCCTGCT
104 TCCACCTCTAGTTTGCTTGACACCGACGGTAGACACGAAGTAGAAGGGCGGTAGACTACTCGTCAACTTTAGACCTTGACGGAGACAACACACGGACGA
ValGluIleLysArgThrValAlaAlaProSerValPheIlePheProProSerAspGluGlnLeuLysSerGlyThrAlaSerValValCysLeuLeu
haeI bstNI scrFI
mnlI mnlI
xmnI rsaI
201 GAATAACTTCTATCCCAGAGAGGCCAAAGTACAGTGGAAGGTGGATAACGCCCTCCAATCGGGTAACTCCCAGGAGAGTGTCACAGAGCAGGACAGCAAG
137 AsnAsnPheTyrProArgGluAlaLysValGlnTrpLysValAspAsnAlaLeuGlnSerGlyAsnSerGlnGluSerValThrGluGlnAspSerLys
sacI
hgiAI
bsp1286
haeIII
sau96I aluI
ecoO banII
hphI alwNI ddeI
fnu4HI accI
ddeI ddeI
mnlI bbvI hgaI
301 GACAGCACCTACAGCCTCAGCAGCACCCTGACGCTGAGCAAAGCCAGACTACGAGAAACACAAAGTCTACGCCTGCGAAGTCACCCATCAGGGCCTGAGCT
170 AspSerThrTyrSerLeuSerSerThrLeuThrLeuSerLysAlaAspTyrGluLysHisLysValTyrAlaCysGluValThrHisGlnGlyLeuSerSer
ddeI
aluI mnlI bsp1286 bspMI mnlI
401 CGCCCGTCACAAAGAGCTTCAACAGGGGAGAGTGTTAGAGGGAGAAGTGCCCCCACCTGCTCCTCAGT
204 ProValThrLysSerPheAsnArgGlyGluCysAM*

FIG. 5

ADHESION VARIANTS

This application is a continuation of U.S. Ser. No. 07/936,190 filed 26 Aug. 1992, now U.S. Pat. No. 5,336,603, which is a continuation of U.S. Ser. No. 07/842,777 filed 18 Feb. 1992, now abandoned, which is a continuation of U.S. Ser. No. 07/250,785 filed 28 Sep. 1988, now abandoned, which is a continuation-in-part of U.S. Ser. No. 07/104,329 filed 2 Oct. 1987, now abandoned.

BACKGROUND OF THE INVENTION

This application relates to compositions for antiviral or immunomodulatory therapy. In particular, it relates to compositions useful in the treatment of Human Immunodeficiency Virus (HIV) infections.

The primary immunologic abnormality resulting from infection by HIV is the progressive depletion and functional impairment of T lymphocytes expressing the CD4 cell surface glycoprotein (H. Lane et al., Ann. Rev. Immunol. 3:477 [1985]). CD4 is a non-polymorphic glycoprotein with homology to the immunoglobulin gene superfamily (P. Maddon et al., Cell 42:93 [1985]). Together with the CD8 surface antigen, CD4 defines two distinct subsets of mature peripheral T cells (E. Reinherz et al., Cell 19:821 [1980]), which are distinguished by their ability to interact with nominal antigen targets in the context of class I and class II major histocompatibility complex (MHC) antigens, respectively (S. Swain, Proc. Natl. Acad. Sci. 78:7101 [1981]; E. Engleman et al., J. Immunol. 127:2124 [1981]; H. Spitz et al., J. Immunol. 129:1563 [1982]; W. Biddison et al., J. Exp. Med. 156:1065 [1982]; and D. Wilde et al., J. Immunol. 131:2178 [1983]). For the most part, CD4 T cells display the helper/inducer T cell phenotype (E. Reinherz, supra), although CD4 T cells characterized as cytotoxic/suppressor T cells have also been identified (Y. Thomas et al., J. Exp. Med. 154:459 [1981]; S. Meuer et al., Proc. Natl. Acad. Sci. USA 79:4395 [1982]; and A. Krensky et al., Proc. Natl. Acad. Sci. USA 79:2365 [1982]). The loss of CD4 helper/inducer T cell function probably underlies the profound defects in cellular and humoral immunity leading to the opportunistic infections and malignancies characteristic of the acquired immunodeficiency syndrome (AIDS) (H. Lane supra).

Studies of HIV-I infection of fractionated CD4 and CD8 T cells from normal donors and AIDS patients have revealed that depletion of CD4 T cells results from the ability of HIV-I to selectively infect, replicate in, and ultimately destroy this T lymphocyte subset (D. Klatzmann et al., Science 225:59 [1984]). The possibility that CD4 itself is an essential component of the cellular receptor for HIV-I was first indicated by the observation that monoclonal antibodies directed against CD4 block HIV-I infection and syncytia induction (A. Dalgleish et al., Nature [London] 312:767 [1984]; J. McDougal et al., J. Immunol. 135:3151 [1985]). This hypothesis has been confirmed by the demonstration that a molecular complex forms between CD4 and gp120, the major envelope glycoprotein of HIV-I (J. McDougal et al., Science 231:382 [1986]); and the finding that HIV-I tropism can be conferred upon ordinarily non-permissive human cells following the stable expression of a CD4 cDNA (P. Maddon et al., Cell 47:333 [1986]). Furthermore, the neurotropic properties of HIV-I, reflected by a high incidence of central nervous system dysfunction in HIV-I infected individuals (W. Snider et al., Ann. Neurol. 14:403 [1983]), and the ability to detect HIV-I in the brain tissue and cerebrospinal fluid of AIDS patients (G. Shaw et al., Science 227:177 [1985]; L. Epstein, AIDS Res. 1:447 [1985]; S. Koenig, Science 233:1089 [1986]; D. Ho et al., N. Engl. J. Med. 313:1498 [1985]; J. Levy et al., Lancet II:586 [1985]), appears to be explained by the expression of CD4 in cells of neuronal, glial and monocyte/macrophage origin (P. Maddon, Cell 47:444 [1986]; I. Funke et al., J. Exp. Med. 165:1230 [1986]; B. Tourvieille et al., Science 234:610 [1986]).

In addition to determining the susceptibility to HIV-I infection, the manifestation of cytopathic effects in the infected host cell appears to involve CD4. Antibody to CD4 was found to inhibit the fusion of uninfected CD4 T cells with HIV-I infected cells in vitro; moreover, the giant multinucleated cells produced by this event die shortly after being formed resulting in the depletion of the population of CD4 cells (J. Lifson et al., Science 232:1123 [1986]). Formation of syncytia also requires gp120 expression, and can be elicited by coculturing CD4-positive cell lines with cell lines expressing the HIV-I env gene in the absence of other vital structural or regulatory proteins (J. Sodroski et al., Nature 322:470 [1986]; J. Lifson et al., Nature 323:725 [1986]). Thus, in mediating both the initial infection by HIV-I as well as eventual cell death, the interaction between gp120 and CD4 constitutes one of several critical entry points in the vital life cycle amenable to therapeutic intervention (H. Mitsuya et al., Nature 325:773 [1987]).

The known sequence of the CD4 precursor predicts a hydrophobic signal peptide, an extracellular region of approximately 370 amino acids, a highly hydrophobic stretch with significant identity to the membrane-spanning domain of the class II MHC beta chain, and a highly charged intracellular sequence of 40 residues (P. Madden, Cell 42:93 [1985]). The extracellular domain of CD4 consists of four continuous regions each having amino acid and structural similarity to the variable and joining (V-J) domains of immunoglobulin light chains as well as related regions in other members of the immunoglobulin gene superfamily (a subclass of which are defined herein by the coined term "adhesons". These structurally similar regions of CD4 are termed the V_1, V_2, V_3 and V_4 domains (denominated 1–4 in FIG. 3).

A successful strategy in the development of drugs for the treatment of many receptor mediated abnormalities has been the identification of antagonists which block binding of the natural ligand. Since the CD4 adheson ordinarily binds to the recognition sites of the HIV envelope it would appear to be a candidate for therapeutically sequestering these HIV sites, thereby blocking vital infectivity. However, full length CD4 and other adhesons are cell membrane proteins which are anchored in the lipid bilayer of cells. The presence of membrane components will be undesirable from the standpoint of manufacturing and purification. In addition, since adhesons are normally present only on cell surfaces, it would be desirable to produce adhesons in a form which is more stable in the circulation. Additionally, even truncated, soluble CD4 adheson (generally referred to as CD4T) may not be optimally effective as a therapeutic since it possesses a relatively short biological half-life, binds to HIV no better than cell surface CD4, may not cross the placental or other biological barriers, and merely sequesters the HIV recognition sites without in itself bearing an infected-cell killing or virus killing functionality.

Accordingly, it is an object of this invention to produce soluble, secreted adhesons. It is another object to produce CD4 derivatives useful in the treatment of AIDS and related conditions, in a manner essentially unaffected by the extreme degree of genetic variation observed among various

HIV-I isolates and their respective env polypeptides (J. Coffin, Cell 46:1 [1986]). Still another object is to prepare adhesons fused to other polypeptides in order to provide molecules with novel functionalities such as those described above for therapeutic use, or diagnostic reagents for the in vitro assay of adhesons or their ligands. In particular, it is an objective to prepare molecules for directing toxins or effector molecules (for example the Fc domain of immunoglobulin) to cells bearing receptors for the adhesons, e.g, HIV gp120 in the case of CD4, and for use in facilitating purification of the adhesons. It is a further object to provide stable, highly purified adheson preparations.

SUMMARY

The objects of this invention are accomplished by providing nucleic acid encoding an amino acid sequence variant of an adheson, in particular a variant in which the transmembrane domain is modified so that it is no longer capable of becoming lodged in the cell membrane. In the case of CD4 such variants are termed soluble CD4.

Variant adhesons are produced by a method comprising (a) transforming a host cell with nucleic acid encoding an amino acid sequence variant of an adheson, (b) culturing the host cell and (c) recovering the variant adheson from the host cell culture media or from lysates of the host cell.

In specific embodiments, the objects of this invention are accomplished by providing an adheson variant selected from the group consisting of (a) an adheson amino acid sequence variant having an inactivated transmembrane domain and (b) a polypeptide comprising an adheson extracellular domain fused to the sequence of a polypeptide which is different from the adheson, this latter, for example, selected from a cytotoxin, an immunogen or a protein with a long plasma half life such as an immunoglobulin constant domain.

In a preferred embodiment a polypeptide comprising a gp120 binding domain of the CD4 adheson is fused at its C-terminus to an immunoglobulin constant domain, or is linked to a cytotoxic polypeptide such as ricin.

The CD4 adheson variants provided herein are purified and formulated in pharmacologically acceptable vehicles for administration to patients in need of antiviral, neuromodulatory or immunomodulatory therapy, in particular patients infected with HIV, and for use in the modulation of cell adhesion.

BRIEF DESCRIPTION OF THE DRAWINGS

FIGS. 1A, 1B-1, 1B-2 and 1C depict the amino acid and nucleotide sequences (SEQ ID NOS:1–3, respectively of a secreted form of the CD4 adheson. The signal processing site is designated with an arrow.

FIGS. 2, 2B-1, 2B-2, and 2C depict the amino acid and nucleotide sequences (SEQ ID NOS:4–6, respectively), of a fusion of the herpes gD leader and N-terminal 27 residues to the putative mature N-terminus of CD4T.

FIG. 3 depicts the structural elements of the native and soluble CD4 adheson, the native human IgG_1 (γ_1) heavy chain and two exemplary heavy chain-CD4 chimeras.

FIGS. 4a, 4B-1 and 4B-2 are a map of the linkered human IgG_1 (γ_1) chain fragment employed in the preparation of CD4 fusions (SEQ ID NOS: 7–9 for amino acid and nucleotide sequences, respectively). Insert sites are designated γ1 and Fc.

FIG. 5 is a map of the human κ light chain fragment useful for CD4 fusions at the arrow flanked by $V_\kappa J_\kappa$ (light variable and joining) and C_κ (light constant) (SEQ ID NOS: 10–12 for amino acid and nucleotide sequences, respectively).

DETAILED DESCRIPTION

Adhesons are cell surface polypeptides having an extracellular domain which is homologous to a member of the immunoglobulin gene superfamily, excluding, however, highly polymorphic members of this superfamily selected from the group of class I and class II major histocompatibility antigens, immunoglobulins and T-cell receptor α, β, γ and δ chains. Examples of adhesons include CD1, CD2, CD4, CD8, CD28, the γ, 67 and ε chains of CD3, OX-2, Thy-1, the intercellular or neural cell adhesion molecules (I-CAM or N-CAM), lymphocyte function associated antigen-3 (LFA-3), neurocytoplasmic protein (NCP-3), poly-Ig receptor, myelin-associated, glycoprotein (MAG), high affinity IgE receptor, the major glycoprotein of peripheral myelin (Po), platelet derived growth factor receptor, colony stimulating factor-1 receptor, macrophage Fc receptor, Fc gamma receptors and carcinoembryonic antigen. Homologous as defined herein means having the sequence of a member of the immunoglobulin gene superfamily or having a sequence therewithin which has substantially the same as (or a greater degree of) amino acid sequence homology to a known member of the superfamily as the specific examples given above have to the sequence of an immunoglobulin variable or constant domain. Preferred adhesons are CD4, CD8 and high affinity IgE Fc receptor.

This invention is particularly concerned with amino acid sequence variants of adhesons. Amino acid sequence variants of adhesons are prepared with various objectives in mind, including increasing the affinity of the adheson for its binding partner, facilitating the stability, purification and preparation of the adheson, increasing its plasma half life, improving therapeutic efficacy as described above in the background, and introducing additional functionalities and lessening the severity or occurrence of side effects during therapeutic use of the adheson. Amino acid sequence variants of adhesons fall into one or a combination of the following classes: insertional, substitutional or deletional variants.

Insertional amino acid sequence variants are those in which one or more amino acid residues extraneous to the adheson are introduced into a predetermined site in the adheson including the C- or N-termini. Such variants are referred to as fusions of the adheson and a different polypeptide. Such other polypeptides contain sequences other than those which are normally found in the adheson at the inserted position. Several groups of fusions are contemplated herein. Immunologically active adheson fusions comprise an adheson and a, polypeptide containing a non-adheson epitope. The non-adheson epitope is any immunologically competent polypeptide, i.e., any polypeptide which is capable of eliciting an immune response in the animal to which the fusion is to be administered or which is capable of being bound by an antibody raised against the non-adheson polypeptide. Typical non-adheson epitopes will be those which are borne by allergens, autoimmune epitopes, or other potent immunogens or antigens recognized by pre-existing antibodies in the fusion recipient, including bacterial polypeptides such as trpLE, beta-galactosidase, vital polypeptides such as herpes gD protein, and the like. Immunogenic fusions are produced by cross-linking in vitro or by recombinant cell culture transformed with

DNA encoding an immunogenic polypeptide. It is preferable that the immunogenic fusion be one in which the immunogenic sequence is Joined to or inserted into the adheson antigen or fragment thereof by a peptide bond(s). These products therefore consist of a linear polypeptide chain containing adheson epitopes and at least one epitope foreign to the adheson. It will be understood that it is within the scope of this invention to introduce the epitopes anywhere within the adheson molecule or fragment thereof. Such fusions are conveniently made in recombinant host cells or by the use of bifunctional cross-linking agents. The use of a cross-linking agent to fuse the adheson to the immunogenic polypeptide is not as desirable as a linear fusion because the cross-linked products are not as easily synthesized in structurally homogeneous form.

These immunogenic insertions are particularly useful when formulated into a pharmacologically acceptable carrier and administered to a subject in order to raise antibodies against the adheson, which antibodies in turn are useful in diagnostics or in purification of adheson by immunoaffinity techniques known per se. Alternatively, in the purification of adhesons, binding partners for the fused non-adheson polypeptide, e.g. antibodies, receptors or ligands, are used to adsorb the fusion from impure admixtures, after which the fusion is eluted and, if desired, the adheson is recovered from the fusion, e.g. by enzymatic cleavage.

Other fusions, which may or may not also be immunologically active, include fusions of the adheson sequence with a signal sequence heterologous to the adheson, fusions of transmembrane-modified CD4 adhesons, for example, to polypeptides having enhanced plasma half life (ordinarily >about 20 hours) such as immunoglobulin chains or fragments thereof, and fusions with cytotoxic functionalities. Signal sequence fusions are employed in order to more expeditiously direct the secretion of the adheson. The heterologous signal replaces the native adheson signal, and when the resulting fusion is recognized, i.e. processed and cleaved by the host cell, the adheson is secreted. Signals are selected based on the intended host cell, and may include bacterial yeast, mammalian and vital sequences. The herpes gD glycoprotein signal is suitable for use in mammalian expression systems.

Plasma proteins which have enhanced plasma half-life longer than that of transmembrane modified CD4 include serum albumin, immunoglobulins, apolipoproteins, and transferrin. Preferably, the adheson-plasma protein fusion is not significantly immunogenic in the animal in which it is used and the plasma protein does not cause undesirable side effects in patients by virtue of its normal biological activity.

In a specific embodiment the adheson immunoglobulin-like domain which may be homologous either to the constant or to the variable region domains is conjugated with an immunoglobulin constant region sequence. The resulting products are referred to herein as immunoadhesons. Immunoglobulins and certain variants thereof are known and many have been prepared in recombinant cell culture. For example, see U.S. Pat. No. 4,745,055; EP 256,654; Faulkner et al., Nature 298:286 (1982); EP 120,694; EP 125,023; Morrison, J. Immun. 123:793 (1979); Köhler et al., P.N.A.S. USA 77:2197 (1980); Raso et al., Cancer Res. 41:2073 (1981); Morrison et al., Ann. Rev. Immunol. 2:239 (1984); Morrison, Science 229:1202 (1985); Morrison et al., P.N.A.S. USA 81:6851 (1984); EP 255,694; EP 266,663; and WO 88/03559. Reassorted immunoglobulin chains also are known. See for example U.S. Pat. No. 4,444;878; WO 88/03565; and EP 68,763 and references cited therein.

Ordinarily, the domains of adhesons that are homologous to immunoglobulins and extracellular in their native environment are fused C-terminally to the N-terminus of the constant region of immunoglobulins in place of the variable region(s) thereof, retaining at least functionally active hinge, CH2 and CH3 domains of the constant region of an immunoglobulin heavy chain. This ordinarily is accomplished by constructing the appropriate DNA sequence and expressing it in recombinant cell culture. Immunoglobulins and other polypeptides having enhanced plasma half life are fused to the extracellular or ligand binding domains of other adhesons in the same fashion.

The boundary domains for the CD4 V-like regions (V1–V4) are, respectively, about 100–109, about 175–184, about 289–298, and about 360–369 (based on the precursor CD4 amino acid sequence in which the initiating met is –25; FIG. 1A). CD4 sequences containing any of the CD4 V domains are fused to the immunoglobulin sequence. It is preferable that the V1V2 or V1V2V3V4 be fused at their C-termini to the immunoglobulin constant region. The precise site at which the fusion is made is not critical; the boundary domains noted herein are for guidance only and other sites neighboring or within the V regions may be selected in order to optimize the secretion or-binding characteristics of the CD4. The optimal site will be determined by routine experimentation. In general, it has been found that the fusions are expressed intracellularly, but a great deal of variation is encountered in the degree of secretion of the fusions from recombinant hosts. For instance, the following table demonstrates the various immunoglobulin fusions that have been obtained by the method of this invention. In all examples of CD4 immunoadhesons, the CD4 signal was used to direct secretion from 293 cells. Lower case m represents murine origin, while the lower case h designates human origin. V and C are abbreviations for immunoglobulin variable and constant domains, respectively. The numerical subscripts indicate the number of parenthetical units found in the designated multimer. It will be understood that the chains of the multimers are believed to be disulfide bonded in the same fashion as native immunoglobulins. The CD4 immunoadhesons Typically contained either the first N-terminal 366 residues of CD4 ($CD4_4$) or the first 180 N-terminal residues of CD4 ($CD4_2$) linked at their C-terminus to the κ0 (light) chain or IgG1 heavy chain constant region (γ1).

TABLE 1

Transfected Gene	Secreted Product
$mV_\kappa C_\kappa$	$mV_\kappa C_\kappa$ and/or $(mV_\kappa C_\kappa)_2$
$mV_{\gamma 1} C_{\gamma 1}$	ND
$mV_\kappa C_\kappa + mV_{\gamma 1} C_{\gamma 1}$	$(mV_\kappa C_\kappa)_2(mV_{\gamma 1} C_{\gamma 1})_2$ + $mV_\kappa C_\kappa$ and/or $(mV_\kappa C_\kappa)_2$
$hCD4\text{-}mC_\kappa$	$hCD4\text{-}mC_\kappa$ and/or $(hCD4\text{-}mC_\kappa)_2$
$hCD4\text{-}mC_{\gamma 1}$	ND
$hCD4\text{-}mC_\kappa + hCD4\text{-}mC_{\gamma 1}$	$(hCD4\text{-}mC_\kappa)_2(hCD4\text{-}mC_{\gamma 1})_2$ + $hCD4\text{-}mC_\kappa$ and/or $(hCD4\text{-}mC_\kappa)_2$
$hCD4\text{-}hC_\kappa$	$hCD4\text{-}hC_\kappa$ and/or $(hCD4\text{-}hC_\kappa)_2$
$hCD4\text{-}hC_{\gamma 1}$	$(hCD4\text{-}hC_{\gamma 1})_2$
$hCD4\text{-}hC_\kappa + hCD4\text{-}hC_{\gamma 1}$	$(hCD4\text{-}hC_\kappa)_2(hCD4\text{-}hC_{\gamma 1})_2$ + $hCD4\text{-}hC_\kappa$ and/or $(hCD4\text{-}hC_\kappa)_2$
$mV_\kappa C_\kappa + hCD4\text{-}hC_{\gamma 1}$	$(mV_\kappa C_\kappa)_2(hCD4\text{-}hC_{\gamma 1})_2$ + $mV_\kappa C_\kappa$ and/or $(mV_\kappa C_\kappa)_2$

*ND = Not detected

It is interesting to observe from this table that the CD4-human heavy chain immunoadheson was secreted as a dimer whereas the analogous murine construction was not detected (this not excluding the intracellular accumulation of the protein, however). The ability of the hCD4-hCγ1 transformants to produce heavy chain dimer was unexpected since previous work had suggested that immunoglobulin heavy

chains are not secreted unless the hosts are cotransformed with nucleic acid encoding both heavy and light chain (Valle et al., Nature 241:338 [1981]). According to this invention, CD4-IgG immunoadheson chimeras are readily secreted wherein the CD4 epitope is present in heavy chain dimers, light chain monomers or dimers, and heavy and light chain heterotetramers wherein the CD4 epitope is present fused to one or more light or heavy chains, including heterotetramers wherein up to and including all four variable region-analogies are derived from CD4. Where light-heavy chain non-CD4 variable domain is present, a heterofunctional antibody thus is provided.

Various exemplary hereto-and chimeric immunoadheson antibodies produced in accordance with this invention are schematically diagrammed below. "A" means at least a portion of the extracellular domain of an adheson containing its ligand binding site; V_L, V_H, C_L and C_H represent light or heavy chain variable or constant domains of an immunoglobulin; n is an integer; and Y designates a covalent cross-linking moiety.

(a) AC_L;

(b) AC_L-AC_L;

(c) AC_H-[AC_H, AC_L-AC_H, AC_L-V_HC_H, V_LC_L-AC_H, or V_LC_L-V_HC_H];

(d) AC_L-AC_H-[AC_H, AC_L-AC_H, AC_L-V_HC_H, V_LC_L-AC_H, or V_LC_L-V_HC_H];

(e) AC_L-V_HC_H-[AC_H, AC_L-AC_H, AC_L-V_HC_H, V_LC_L-AC_H, or V_LC_L-V_HC_H];

(f) V_LC_L-AC_H-[AC_H, AC_L-AC_H, AC_L-V_HC_H, V_LC_L-AC_H, or V_LC_L-V_HC_H]; or

(g) $[A\text{-}Y]_n$-$[V_LC_L\text{-}V_HC_H]_2$.

The structures shown in this table show only key features, e. g. they do not show Joining (J) or other domains of the immunoglobulins, nor are disulfide bonds shown. These are omitted in the interests of brevity. However, where such domains are required for binding activity they shall be construed as being present in the ordinary locations which they occupy in the adheson, immunoadheson or immunoglobulin molecules as the case may be. These examples are representative of divalent antibodies; more complex structures would result by employing immunoglobulin heavy chain sequences from other classes, e.g. IgM. The immunoglobulin V_LV_H antibody combining site, also designated as the companion immunoglobulin, preferably is capable of binding to a predetermined antigen.

Suitable companion immunoglobulin combining sites and fusion partners are obtained from IgG-1, -2, -3, or -4 subtypes, IgA, IgE, IgD or IgM, but preferably IgG-1.

A preferred embodiment is a fusion of an N-terminal portion of CD4, which contains the binding site for the gp120 envelope protein of HIV, to the C-terminal F_c portion of an antibody, containing the effector functions of immunoglobulin G_1. There are two preferred embodiments of this sort; in one, the entire heavy chain constant region is fused to a portion of CD4; in another, a sequence beginning in the hinge.,region Just upstream of the pepsin cleavage site which defines IgG F_c chemically (residue 216, taking the first residue of heavy chain constant region to be 114 [Kobat et al., "Sequences of Proteins of Immunological Interest" 4th Ed., 1987], or analogous sites of other immunoglobulins) is fused to a portion of CD4. These embodiments are described in the examples.

More particularly, those variants in which one or more immunoglobulin-like domains of an adheson are substituted for the variable region of an immunoglobulin chain are believed to exhibit improved in vivo plasma half life. These chimeras are constructed in a fashion similar to chimeric antibodies in which a variable domain from an antibody of one species is substituted for the variable domain of another species. See, for example, EP 0 125 023; Munro, Nature 312: (13 Dec. 1984); Neuberger et al., Nature 312: (13 Dec. 1984); Sharon et al., Nature 309: (24 May 1984); Morrison et al., Proc. Natl. Acad. Sci. USA 81:6851–6855 (1984); Morrison et al. Science 229:1202–1207 (1985); and Boulianne et al. Nature 312:643–646 (13 Dec. 1984). The DNA encoding the adheson immunoglobulin-like domain(s) is cleaved by a restriction enzyme at or proximal to the 3' end of the DNA encoding the immunoglobulin-like domain(s) and at a point at or near the DNA encoding the N-terminal end of the mature adheson polypeptide (where use of a different leader is contemplated) or at or proximal to the N-terminal coding region for the adheson (where the native adheson signal is employed). This DNA fragment then is readily inserted into DNA encoding an immunoglobulin light or heavy chain constant region and, if, necessary, tailored by deletional mutagenesis. Preferably, this is a human immunoglobulin when the variant is intended for in vivo therapy for humans. DNA encoding immunoglobulin light or heavy chain constant regions is known or readily available from cDNA libraries or is synthesized. See, for example, Adams et al., Biochemistry 19:2711–2719 (1980); Gough et al., Biochemistry 19:2702–2710 (1980); Dolby et al., P.N.A.S. USA, 77:6027-6031 (1980); Rice et al., P.N.A.S. USA 79:7862–7865 (1982); Falkner et al., Nature 298:286–288 (1982); and Morrison et al., Ann. Rev. Immunol 2:239–256 (1984).

DNA encoding the immunoglobulin or immunoadheson chimeric chain(s) is transfected into a host cell for expression. If the host cell is producing an immunoglobulin prior to transfection then one need only transfect with the adheson fused to light or to heavy chain to produce a heteroantibody. The aforementioned immunoglobulins having one or more arms bearing the adheson domain and one or more arms bearing companion variable regions result in dual specificity for adheson ligand and for an antigen. These are produced by the above-described recombinant methods or by in vitro procedures. In the latter case, for example. $F(ab')_2$ fragments of the adheson fusion and an immunoglobulin are prepared, the F(ab')2 fragments converted to Fab' fragments by reduction under mild reducing conditions, and then reoxidized in each other's presence under acidic conditions in accord with methods known per se. See also U.S. Pat. No. 4,44,878.

Additionally, procedures are known for producing intact heteroantibodies from immunoglobulins having different specificities. These procedures are adopted for the in vitro production of heterochimeric antibodies by simply substituting the immunoadheson chains for one of the previously employed immunoglobulins.

In an alternative method for producing a heterofunctional antibody, host cells producing an adheson-immunoglobulin fusion, e.g. transfected myelomas, also are fused with B cells or hybridomas which secrete antibody having the desired companion specificity for an antigen. Heterobifunctional antibody is recovered from the culture medium of such hybridomas, and thus may be produced somewhat more conveniently than by conventional in vitro resorting methods (EP 68,763).

Another group of fusions are those in which an adheson is conjugated with a toxic substance, e.g. a polypeptide such as ricin (including deglycosylated ricin A chain), diptheria toxin A, or a non-peptidyl cytotoxin. Where the toxin is a polypeptide it is convenient to cross-link the polypeptide to the adheson or its transmembrane-deleted variant by con-

ventional in vitro protein cross-linking agents (for suitable methods for linking ricin A chain or deglycosylated A chain to CD4 see, for example, Duncan et al., "Analy. Biochem." 132:68–73 [1983]; Thorpe et al., "Cancer Res." 47:5924 [1987]; and Ghotie et al., "Cancer Res." 48:2610 [1988]) or by recombinant synthesis as a fusion (see for example, U.S. Pat. No. 4,765,382). Alternatively, where companion antibodies are anti-ricin antibody immunoglobulin variable domains, such immunoglobulin heteroantibodies are employed to deliver ricin to HIV infected cells following the general procedure of Raso et al., Cancer Research, 41:2073 (1981).

Another class of adheson variants are deletional variants. Deletions are characterized by the removal of one or more amino acid residues from an adheson sequence. Typically, the transmembrane and cytoplasmic domains of adhesons are deleted. In the case of CD4, at least residues 368 to 395 (the transmembrane region), and ordinarily 396–433 as well (the cytoplasmic domain), will be deleted to obtain secreted forms of this adheson. Parenthetically, the amino acid residues follow the numbers given for mature CD4 as noted, for example, in FIGS. 1A–1C. Thus, CD4T molecules generally will terminate in the vicinity of about residues 366–368, or at any other suitable site N-terminal thereto which preserves the gp120-binding capability of the CD4 variant.

Substitutional variants are those in which at least one residue in the adheson sequence has been removed and a different residue inserted in its place. The native N-terminal residue for mature CD4 is now known to be lysine. Thus, the sequence shown in FIG. 1, with an N-terminal asparagine, is an amino acid sequence variant of native mature CD4. Table II below describes substitutions which in general will result in fine modulation of the characteristics of the CD antigen.

TABLE II

Original Residue	Exemplary Substitutions
Ala	ser
Arg	lys
Asn	gln; his
Asp	glu
Cys	ser; ala
Gln	asn
Glu	asp
Gly	pro
His	asn; gln
Ile	leu; val
Leu	ile; val
Lys	arg; gln; glu
Met	leu; ile
Phe	met; leu; tyr
Ser	thr
Thr	ser
Trp	tyr
Tyr	trp; phe
Val	ile; leu

Substantial changes in function or immunological identity are made by selecting substitutions that are less conservative than those in Table 2, i.e., selecting residues that differ more significantly in their effect on maintaining (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a sheet or helical conformation, (b) the charge or hydrophobicity of the molecule at the target site or (c) the bulk of the side chain. The substitutions which in general are expected to produce the greatest changes in adheson properties will be those in which (a) a hydrophilic residue, e.g. seryl or threonyl, is substituted for (or by) a hydrophobic residue, e.g. leucyl, isoleucyl, phenylalanyl, valyl or alanyl; (b) a cysteinyl or prolyl is substituted for (or by) any other residue; (c) a residue having an electropositive side chain, e.g., lysyl, arginyl, or histidyl, is substituted for (or by) an electronegative residue, e.g., glutamyl or aspartyl; or (d) a residue having a bulky side chain, e.g., phenylalanyl, is substituted for (or by) one not having a side chain, e.g., glycyl.

A preferred class of substitutional or deletional variants are those involving the transmembrane region of the adheson. The transmembrane region of the adheson is a highly hydrophobic or lipophilic domain that is the proper size to span the lipid bilayer of the cellular membrane. It is believed to anchor the adheson in the cell membrane.

Deletion or substitution of the transmembrane domain will facilitate recovery and provide a soluble form of the adheson by reducing its cellular or membrane lipid affinity and improving its water solubility. If the transmembrane and cytoplasmic domains are deleted one avoids the introduction of potentially immunogenic epitopes, either by exposure of otherwise intracellular polypeptides that might be recognized by the body as foreign or by insertion of heterologous polypeptides that are potentially immunogenic. A principal advantage of the transmembrane deleted adheson is that it is secreted into the culture medium of recombinant hosts. This variant is water soluble and does not have an appreciable affinity for cell membrane lipids, thus considerably simplifying its recovery from recombinant cell culture.

It will be amply apparent from the foregoing discussion that substitutions, deletions, insertions or any combination thereof are introduced to arrive at a final construct. As a general proposition, all variants will not have a functional transmembrane domain and preferably will not have a functional cytoplasmic sequence. This is generally accomplished by deletion of the relevant domain, although adequate insertional or substitutional mutagens also can be effective for this purpose. For example, the transmembrane domain is substituted by any amino acid sequence, e.g. a random or homopolynucleic sequence of about 5 to 50 serine, threonine, lysine, arginine, glutamine, aspartic acid and like hydrophilic residues, which altogether exhibit a hydrophilic hydropathy profile, so that it is secreted into the culture medium of recombinant hosts. This variant should also be considered to be an adheson variant.

These variants ordinarily are prepared by site specific mutagenesis of nucleotides in the DNA encoding the adheson, thereby producing DNA encoding the variant, and thereafter expressing the DNA in recombinant cell culture. However, variant adhesons also are prepared by in vitro synthesis. Obviously, variations made in the DNA encoding the variant adhesons must not place the sequence out of reading frame and preferably will not create complementary regions that could produce secondary mRNA structure deleterious to expression (EP 75,444A). The CD4 variants typically exhibit the same gp120 binding activity as does the naturally-occurring prototype, although variants also are selected in order to modify the characteristics of the CD4 adheson as indicated above.

While the site for introducing an amino acid sequence variation is predetermined, the mutation per se need not be predetermined. For example, in order to optimize the performance of a mutation at a given site, random mutagenesis may be conducted at the target codon or region and the expressed adheson variants screened for the optimal combination of desired activities. Techniques for making substitution mutations at predetermined sites in DNA having a known sequence are well known, for example, M13 primer mutagenesis.

Adheson variants that are not capable of binding HIV gp120 are useful nonetheless as immunogens for raising

antibodies to the adheson or as immunoassay kit components (labelled, as a competitive reagent for gp120 assay, or unlabelled as a standard for an adheson assay) so long as at least one adheson epitope remains active.

The DNA encoding adhesons is obtained by known procedures. See Williams, Immunol. Today 8:298–303 (1987) and citations therein. In general, prokaryotes are used for cloning of CD4 variant DNA sequences. For example, *E. coli* strain SR101 (for propagating M13 phage, a λ-resistant strain of JM 101; Messing et al., Nucl. Acids. Res. 9(2):309–321 [1981]); and *E. coli* K12 strain 294 (ATCC No. 31446) are particularly useful. Other microbial strains which may be used include *E. coli* B, UM101 and *E. coli* $_{102}$ 1776 (ATCC No. 31537). These examples are illustrative rather than limiting.

DNA encoding the variant adhesons is inserted for expression into vectors containing promoters and control sequences which are derived from species compatible with the intended host cell. The vector ordinarily, but need not, carry a replication site as well as one or more marker sequences which are capable of providing phenotypic selection in transformed cells. For example, *E. coli* is typically transformed using a derivative of pBR322 which is a plasmid derived from an *E. coli* species (Bolivar, et al., Gene 2:95 [1977]). pBR322 contains genes for ampicillin and tetracycline resistance and thus provides easy means for identifying transformed cells. The pBR322 plasmid, or other microbial plasmid, must also contain or be modified to contain promoters and other control elements commonly used in recombinant DNA constructions.

Promoters suitable for use with prokaryotic hosts illustratively include the β-lactamase and lactose promoter systems (Chang et al., Nature, 275:615 [1978]; and Goeddel et al., Nature 281: 544 [1979]), alkaline phosphatase, the tryptophan (trp) promoter system (Goeddel, Nucleic Acids Res. 8: 4057 [1980] and EPO Appln. Publ. No. 36,776) and hybrid promoters such as the tac promoter (H. de Boer et al., Proc. Natl. Acad. Sci. USA80: 21–25 [1983]). However, other functional bacterial promoters are suitable. Their nucleotide sequences are generally known, thereby enabling a skilled worker operably to ligate them to DNA encoding the adheson variant using linkers or adaptors to supply any required restriction sites, (Siebenlist et al., Cell 20: 269 [1980]). Promoters for use in bacterial systems also will contain a Shine-Dalgarno (S.D.) sequence operably linked to the DNA encoding the antigen.

In addition to prokaryotes, eukaryotic microbes such as yeast cultures also are useful as cloning or expression hosts. *Saccharomyces cerevisiae*, or common baker's yeast, is the most commonly used eukaryotic microorganism, although a number of other strains are commonly available. For expression in Saccharomyces, the plasmid YRp7, for example, (Stinchcomb, et al., Nature 282:39 [1979]; Kingsman et al, Gene 7: 141 [1979]; Tschemper et al., Gene 10: 157 [1980]) is commonly used. This plasmid already contains the trp1 gene which provides a selection marker for a mutant strain of yeast lacking the ability to grow in tryptophan, for example ATCC no. 44076 or PEP4-1 (Jones, Genetics 85: 12 [1977]). The presence of the trp1 lesion as a characteristic of the yeast host cell genome then provides an effective means of selection by growth in the absence of tryptophan.

Suitable promoting sequences for use with yeast hosts include the promoters for 3-phosphoglycerate kinase (Hitzeman et al., J. Biol. Chem. 255: 2073 [1980]) or other glycolytic enzymes (Hess et al., J. Adv. Enzyme Reg. 7: 149 [1968]; and Holland, Biochemistry 17:4900 [1978]), such as enolase, glyceraldehyde-3-phosphate dehydrogenase, hexokinase, pyruvate decarboxylase, phosphofructokinase, glucose-6-phosphate isomerase, 3-phosphoglycerate mutase, pyruvate kinase, triosephosphate isomerase, phosphoglucose isomerase, and glucokinase.

Other yeast promoters, which are inducible promoters having the additional advantage-of transcription controlled by growth conditions, are the promoter regions for alcohol dehydrogenase 2, isocytochrome C, acid phosphatase, degradative enzymes associated with nitrogen metabolism, metallothionein, glyceraldehyde-3-phosphate dehydrogenase, and enzymes responsible for maltose and galactose utilization. Suitable vectors and promoters for use in yeast expression are further described in R. Hitzeman et al., European Patent Publication No. 73,657A. Yeast enhancers also are advantageously used with yeast promoters.

Promoters for controlling transcription from vectors in mammalian host cells may be obtained from various sources, for example, the genomes of viruses such as: polyoma, Simian Virus 40 (SV40), adenovirus, retroviruses, hepatitis-B virus and most preferably cytomegalovirus, or from heterologous mammalian promoters, e.g. the beta actin promoter. The early and late promoters of the SV40 virus are conveniently obtained as an SV40 restriction fragment which also contains the SV40 viral origin of replication. Fiefs et al., Nature, 273: 113 (1978). The immediate early promoter of the human cytomegalovirus is conveniently obtained as a HindIII E restriction fragment. Greenaway, P. J. et al., Gene 18:355–360 (1982). Of course, promoters from the host cell or related species also are useful herein.

DNA transcription in higher eukaryotes is increased by inserting an enhancer sequence into the vector. Enhancers are cis-acting elements of DNA, usually from about 10 to 300 bp, that act to increase the transcription initiation capability of a promoter. Enhancers are relatively orientation and position independent having been found 5' (Laimins, L. et al., Proc. Natl. Acad. Sci. 78:993 [1981]) and 3' (Lusky, M. L., et al., Mol. Cell Bio. 3: 1108 [1983]) to the transcription unit, within an intron (BanerJi, J. L. et al., Cell 33: 729 [1983]) as well as within the coding sequence itself (Osborne, T. F., et al. Mol. Cell Bio. 4: 1293 [1984]). Many enhancer sequences are now known from mammalian genes (globin, elastase, albumin, α-fetoprotein and insulin). Typically, however, one will use an enhancer from a eukaryotic cell virus. Examples include the SV40 enhancer on the late side of the replication origin (bp 100–270), the cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers.

Expression vectors used in eukaryotic host cells (yeast, fungi, insect, plant, animal, human or nucleated cells) may also contain sequences necessary for, the termination of transcription which may affect mRNA expression. These regions are transcribed as polyadenylated segments in the untranslated portion of the mRNA encoding the adheson.

Expression vector systems generally will contain a selection gene, also termed a selectable marker. Examples of suitable selectable markers for mammalian cells are dihydrofolate reductase (DHFR), thymidine kinase or neomycin. When such selectable markers are successfully transferred into a mammalian host cell, the transformed mammalian host cell can survive if placed under selective pressure. There are two widely used distinct categories of selective regimes. The first category is based on a cell's metabolism and the use of a mutant cell line which lacks the ability to grow independent of a supplemented medium. Two examples are: CHO DHFR$^-$ cells and mouse LTK$^-$ cells. These cells lack the ability to grow without the addition of

such nutrients as thymidine or hypoxanthine. Because these cells lack certain genes necessary for a complete nucleotide synthesis pathway, they cannot survive unless the missing nucleotides are provided in a supplemented medium. An alternative to supplementing the medium is to introduce an intact DHFR or TK gene into cells lacking the respective genes, thus altering their growth requirements. Individual cells which were not transformed with the DHFR or TK gene will not be capable of survival in non-supplemented media.

The second category is dominant selection, which refers to a selection scheme used in any cell type and does not require the use of a mutant cell line. These schemes typically use a drug to arrest growth of a host cell. Those cells which have a novel gene would express a protein conveying drug resistance and would survive the selection. Examples of such dominant selection use the drugs neomycin, (Southern P. and Berg, P., J. Molec. Appl. Genet 1: 327 [1982]), mycophenolic acid (Mulligan, R. C. and Berg, P. Science 209: 1422 [1980]) or hygromycin (Sugden, B. et al., Mol. Cell. Biol. 5: 410–413 [1985]). The three examples given above employ bacterial genes under eukaryotic control to convey resistance to the appropriate drug G418 or neomycin (geneticin), xgpt (mycophenolic acid) or hygromycin, respectively.

"Amplification" refers to the increase or replication of an isolated region within a cell's chromosomal DNA. Amplification is achieved using a selection agent e.g. methotrexate (MTX) which inactivates DHFR. Amplification or the making of successive copies of the DHFR gene results in greater amounts of DHFR being produced in the face of greater amounts of MTX. Amplification pressure is applied notwithstanding the presence of endogenous DHFR, by adding ever greater amounts of MTX to the media. Amplification of a desired gene can be achieved by cotransfecting a mammalian host cell with a plasmid having a DNA encoding a desired protein and the DHFR or amplification gene permitting cointegration. One ensures that the cell requires more DHFR, which requirement is met by replication of the selection gene, by selecting only for cells that can grow in the presence of ever-greater MTX concentration. So long as the gene encoding a desired heterologous protein has cointegrated with the selection gene, replication of this gene gives rise to replication of the gene encoding the desired protein. The result is that increased-copies of the gene, i.e. an amplified gene, encoding the desired heterologous protein express more of the desired heterologous protein.

Preferred host cells for expressing the CD antigen variants of this invention are mammalian cell lines, examples including: monkey kidney CV1 line transformed by SV40 (COS-7, ATCC CRL 1651); human embryonic kidney line (293, Graham, F. L. et al., J. Gen Virol. 36: 59 [1977] and 293S cells [293 subclones selected for better suspension growth]); baby hamster kidney cells (BHK, ATCC CCL 10); chinese hamster ovary-cells-DHFR (CHO, Urlaub and Chasin, Proc. Natl. Acad. Sci. (USA) 77: 4216, [1980]); mouse sertoli cells (TM4, Mather, J. P., Biol. Reprod. 23: 243–251 [1980]); monkey kidney cells (CV1 ATCC CCL 70); african green monkey kidney cells (VERO-76, ATCC CRL-1587); human cervical carcinoma cells (HELA, ATCC CCL 2); canine kidney cells (MDCK, ATCC CCL 34); buffalo rat liver cells (BRL 3A, ATCC CRL 1442); human lung cells (W138, ATCC CCL 75); human liver cells (Hep G2, HB 8065); mouse mammary tumor cells (MMT 060562, ATCC CCL51); and TRI cells (Mather, J. P. et al., Annals N.Y. Acad. Sci. 383: 44–68 [1982]).

"Transformation" means introducing DNA into an organism so that the DNA is replicable, either as an extrachromosomal element or by chromosomal integration. One suitable transformation of the host cells is the method of Graham, F. and van der Eb, A., Virology 52: 456–457 (1973). However, other methods for introducing DNA into cells such as by nuclear injection or by protoplast fusion may also be used. If prokaryotic cells or cells which contain substantial cell walls are used as hosts, the preferred method of transfection is calcium treatment using calcium chloride as described by Cohen, F. N. et al., Proc. Natl. Acad. Sci. (USA), 69:2110 (1972).

Construction of suitable vectors containing the desired coding and control sequences employs standard and manipulative ligation techniques. Isolated plasmids or DNA fragments are cleaved, tailored, and religated in the form desired to form the plasmids required. Suitable procedures are well known for the construction described herein. See, for example, Maniatis, T. et al., *Molecular Cloning*, 133–134 Cold Spring Harbor, (1982); "Current Protocols in Molecular Biology", edited by Ausubel et al., (1987) pub. by Greene Publishing Associates & Wiley-Interscience.

Correct plasmid sequences are confirmed by transforming *E. coli* K12 strain 294 (ATCC 31446) with ligation mixtures, successful transformants are selected by ampicillin or tetracycline resistance where appropriate, and plasmids from the transformants are prepared, and then analyzed by restriction enzyme digestion and/or sequenced by the method of Messing et al., Nucleic Acids Res. 9: 309 (1981) or by the method of Maxam et al., Methods in Enzymology 65: 499 (1980).

Host cells are transformed with the expression vectors of this invention. Thereafter they are cultured in appropriate culture media, e.g. containing substances for inducing promoters, selecting transformants or amplifying genes. The culture conditions, such as temperature, pH and the like, are those previously used with the host cell selected for expression, and will be apparent to the ordinarily skilled artisan.

The secreted adheson variants are recovered and purified from the culture supernatants or lysates of recombinant hosts. Typically, the supernatants are concentrated by ultrafiltration, contacted with a ligand affinity or immunoaffinity matrix so as to adsorb the adheson variant, and eluted from the matrix. Optionally, the adheson is purified by ion exchange chromatography.

Surprisingly, purification of soluble CD4 adheson from culture medium was unexpectedly difficult. Notwithstanding that the hydrophobic transmembrane region of the antigen had been deleted, the antigen exhibited a strong tendency to form aggregates that could be readily removed from suspension by centrifugation at 1000× g, and which avidly coat surfaces such as ultrafiltration membranes. This appears to result from the reduction in concentration of albumin or other serum protein (ordinarily present in the crude preparation) to a particular level, below which the truncated antigen no longer remains soluble. This phenomenon appears to be aggravated by exposure of the CD4 adheson to low pH (< about pH 4). As a result, separation procedures (particularly those that employ acid elution, such as immunoaffinity) should be modified so that the eluate is maintained at, or immediately returned to, about neutrality. Further, a surfactant, e.g. a detergent such as Tween 80, should be included with the antigen during the separation procedure. The final purified product will be stabilized with a predetermined protein such as albumin, and/or a detergent.

The purified adheson is formulated into conventional pharmacologically acceptable excipients.

It is administered to patients having HIV infection at a dosage capable of maintaining a concentration of greater than about 100 ng of soluble CD4 adheson/ml plasma. For

CD4 adheson variants having different molecular weights, about 2 picomoles of soluble receptor per ml of plasma will be initially evaluated clinically in order to establish a stoichiometric equivalence with native (membrane bound) and soluble receptor. The ordinary dosage of soluble CD4 is 100 μg/kg of patient weight/day.

The therapeutic CD4 variants are employed with other therapies and agents for the treatment of AIDS, including AZT, neutralizing antibodies and immunocytotoxins, gp120 fragments and vaccines.

In order to facilitate understanding of the following examples certain frequently occurring methods and/or terms will be described.

"Plasmids" are designated by a lower case p preceded and/or followed by capital letters and/or numbers. The starting plasmids herein are either commercially available, publicly available on an unrestricted basis, or can be constructed from available plasmids in accord with published procedures. In addition, equivalent plasmids to those described are known in the art and will be apparent to the ordinarily skilled artisan.

"Digestion" of DNA refers to catalytic cleavage of the DNA with a restriction enzyme that acts only at certain sequences in the DNA. The various restriction enzymes used herein are commercially available and their reaction conditions, cofactors and other requirements are used as would be known to the ordinarily skilled artisan. For analytical purposes, typically 1 μg of plasmid or DNA fragment is used with about 2 units of enzyme in about 20 μl of buffer solution. For the purpose of isolating DNA fragments for plasmid construction, typically 5 to 50 μg of DNA are digested with 20 to 250 units of enzyme in a larger volume. Appropriate buffers and substrate amounts for particular restriction enzymes are specified by the manufacturer. Incubation times of about 1 hour at 37° C. are ordinarily used, but may vary in accordance with the supplier's instructions. After digestion the reaction is electrophoresed directly on a polyacrylamide gel to isolate the desired fragment.

"Recovery" or "isolation" of a given fragment of DNA from a restriction digest means separation of the digest on polyacrylamide or agarose gel by electrophoresis, identification of the fragment of interest by comparison of its mobility versus that of marker DNA fragments of known molecular weight, removal of the gel section containing the desired fragment, and separation of the gel from DNA. This procedure is known generally (Lawn, R. et al., Nucleic Acids Res. 9: 6103–6114 [1981], and Goeddel, D. et al., Nucleic Acids Res. 9: 4057 [1980]).

"Dephosphorylation" refers to the removal of the terminal 5' phosphates by treatment with bacterial alkaline phosphatase (BAP). This procedure prevents the two restriction cleaved ends of a DNA fragment from "circularizing" or forming a closed loop that would impede insertion of another DNA fragment at the restriction site. Procedures and reagents for dephosphorylation and other recombinant manipulations are conventional. Reactions using BAP are carried out in 50 mM Tris at 68° C. to suppress the activity of any exonucleases which may be present in the enzyme preparations. Reactions are run for 1 hour. Following the reaction the DNA fragment is gel purified.

"Ligation" refers to the process of forming phosphodiester bonds between two double stranded nucleic acid fragments (Maniatis, T. et al., Id. at 146). Unless otherwise provided, ligation may be accomplished using known buffers and conditions with 10 units of T4 DNA ligase ("ligase") per 0.5 μg of approximately equimolar amounts of the DNA fragments to be ligated.

"Filling" or "blunting" refers to the procedures by which the single stranded end in the cohesive terminus of a restriction enzyme-cleaved nucleic acid is converted to a double strand. This eliminates the cohesive terminus and forms a blunt end. This process is a versatile tool for converting a restriction cut end that may be cohesive with the ends created by only one or a few other restriction enzymes into a terminus compatible with any blunt-cutting restriction endonuclease or other filled cohesive terminus. Typically, blunting is accomplished by incubating 2–15 μg of the target DNA in 10 mM $MgCl_2$, 1 mM dithiothreitol, 50 mM NaCl, 10 mM Tris (pH 7.5) buffer at about 37° C. in the presence of 8 units of the Klenow fragment of DNA polymerase I and 250 μM of each of the four deoxynucleoside triphosphates. The incubation generally is terminated after 30 min. and the reaction mixture is subjected to phenol and chloroform extraction and ethanol precipitation.

The following examples merely illustrate the best mode now contemplated for practicing the invention, but should not be construed to limit the invention. All literature citations herein are expressly incorporated by reference.

EXAMPLE 1

Construction of Vectors for the Expression of Native CD4 and Secreted Derivatives

Section 1

The plasmid used for recombinant synthesis of human CD4 was pSVeCD4DHFR. The plasmid was constructed as follows:

λCD4P1 containing most of the coding sequence of human CD4 (obtained from a human placental cDNA library using oligonucleotide probes based on the published sequence [Maddon et al. 1985]), supra was digested with EcoRI to produce the cDNA insert. This fragment was recovered by polyacrylamide gel electrophoresis (fragment 1).

pUC18 was digested with EcoRI and the single fragment recovered by polyacrylamide gel electrophoresis (fragment 2). Fragment 1 was ligated to fragment 2 and the ligation mixture transformed into *E. coli* strain 294. The transformed culture was plated on ampicillin media plates and resistant colonies selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct DNA fragments. This plasmid is referred to as pUCCD4.

pSVeE'DHFR (Muesing et al., Cell 48:691–701 [1987]) was digested with KpnI and BamHI and blunted with *E. coli* DNA polymerase I (Klenow fragment) and the four dNTPs. Fragment 3 containing the pML-Amp^r region, SV40 early promoter, the HIV LTR, and the mouse DHFR gene was recovered by gel electrophoresis, ligated and the ligation mixture transformed into *E. coli* strain 294. The transformed culture was plated on ampicillin media plates and resistant colonies were selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the BamHI restriction site and the absence of the KpnI restriction site. This plasmid is referred to as pSVeABKDHFR and allows EcoRI-BamHI fragments to be inserted after the SV40 early promoter and transcribed under its control, following transfection into an appropriate cell line.

Synthetic oligonucleotides (adaptors 1–8, below) were made to extend from 76 bp 5' of the initiation codon of CD4 translation to the RsaI restriction site at 121 bp 3' of the

initiator, with the sequence AATT at the 5' end of the sense strand to generate an end which could ligate to an EcoRI restriction fragment. These oligonucleotides were ligated and the 204 bp fragment containing the entire sequence was recovered by gel electrophoresis (fragment 4).

CD4 adaptor 1: AATTCAAGCCCAGAGCCCTGC-CATTTCTGTGGGCTCAGGTCCCT

CD4 adaptor 2: pACTGCTCAGCCCCTTCCTC-CCTCGGCAAGGCCACAATGAACCGGGGAGTC

CD4 adaptor 3: pCCTTTTAGGCACTTGCTTCTGGT-GCTGCAACTGGCGCTCCTCCCAGC (SEQ ID NO: 15)

CD4 adaptor 4: pAGCCACTCAGGGAAACAAAGTG-GTGCTGGGCAAAAAAGGGGATACAGTG-GAACTGACCTGT (SEQ ID NO: 16)

CD4 adaptor 5: pACAGGTCAGTTCCACTGTATC-CCCTTTTTTGCCCAGCACCACTTTGTTTCC (SEQ ID NO: 17)

CD4 adaptor 6: pCTGAGTGGCTGCTGGGAG-GAGCGCCAGTTGCAGCACCAGAAGCAAGT (SEQ ID NO: 18)

CD4 adaptor 7: pGCCTAAAAGGGACTCCCCGGT-TCATTGTGGCCTTGCCGAGGGAGGAAGGG (SEQ ID NO: 19)

CD4 adaptor 8: GCTGAGCAGTAGGGACCTGAGC-CCACAGAAATGGCAGGGCTCTGGGCTTG (SEQ ID NO: 20)

pUCCD4 was digested with RsaI and SstI and the 401 bp fragment containing part of the CD4 coding sequence recovered by gel electrophoresis (fragment 5). pUC18 was digested with EcoRI and SstI and the fragment comprising the bulk of the plasmid recovered by gel electrophoresis (fragment 6). Fragments 4 and 5 were ligated to fragment 6 and the ligation mixture was transformed into *E. coli* strain 294. The transformed culture was plated on ampicillin media plates and resistant colonies selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct fragment. The sequence of the inserted synthetic DNA was checked by excising the 605 bp EcoRI-SstI fragments from several transformants and ligating them to M13mp19 which had been digested with the same enzymes. After transformation into *E. coli* strain JM101, single-stranded DNA was prepared and sequenced. One plasmid which contained the correct sequence was selected, and is referred to as pCD4int.

pCD4int was digested with EcoRI and SstI and fragment 7 containing the 5' end of the CD4 coding region was recovered by gel electrophoresis. pUCCD4 was digested with SstI and BamHI and the 1139 bp fragment containing the remainder of the CD4 coding region (fragment 8) recovered by gel electrophoresis.

pSVeΔBKDHFR was digested with EcoRI and BamHI and fragment 9 comprising the bulk of the plasmid was isolated. Fragments 7, 8 and 9 were ligated and the ligation mixture transformed into *E. coli* strain 294. The transformed culture was plated on ampicillin media plates and the resistant colonies were selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct fragment. This plasmid is referred to as pSVeCD4DHFR, and was used to direct synthesis of recombinant intact CD4.

Section 2

A plasmid was constructed to direct the synthesis of a CD4 derivative lacking the putative transmembrane domain and most of the putative cytoplasmic domain (Maddon et al., Supra). This was done with the intention of creating a secreted form of CD4, based on the assumption that these domains anchor the CD4 glycoprotein to the cell membrane, and that their deletion would result in the secretion of the product. This plasmid is referred to as pSVeCD4ΔN1aDHFR and was constructed as follows:

pUCCD4 was digested with SstI and TaqI and the 531 bp fragment (fragment 10) recovered, pUCCD4 was digested with NlaIII and TaqI and the 112 bp fragment (fragment 11) recovered. pUCCD4 was digested with BamHI and NlaIII and the 301 bp fragment (fragment 12) recovered. pCD4int was digested with SstI and BamHI and fragment 13 comprising the bulk of the plasmid recovered. Fragments 10, 11, and 12 were ligated together with fragment 13 and the ligation mixture was transformed into *E. coli* strain 294. The transformed culture was plated on ampicillin media plates and resistant colonies were selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct fragment. Plasmid DNA from several transformants was sequenced to ensure that the 195 bp NlaIII fragment had been deleted and that the proper reading frame was restored. The resulting plasmid is referred to as pCD4ΔN1a.

pCD4ΔN1a was digested with EcoRI and BamHI and the 1541 bp fragment containing the sequence of a CD4 derivative lacking the transmembrane and cytoplasmic domains recovered (fragment 14) and ligated to fragment 9, and the ligation mixture was transformed into *E. coli* strain 294. The transformed culture was plated on ampicillin media plates and resistant colonies were selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct fragment. This plasmid is referred to as pSVeCD4ΔN1aDHFR.

Both pSVeCD4DHFR and pSVeCD4ΔN1aDHFR were transfected into CHO cells by the same method used to establish cell lines stably expressing HIV-I polypeptides (Muesing, Smith and Capon, Cell 48:6910701 [1987]). These cells were assayed for production by radioimmunoprecipitation as described below. While no product was detected in initial experiments, subsequent experiments showed that the above described coding segment could indeed direct the synthesis of a soluble CD4 adheson variant both in CHO and 293 cells.

Section 3

A different expression system was initially used for the synthesis and expression of a CD4 variant lacking completely the cytoplasmic and transmembrane domains. This system uses the cytomegalovirus promoter and can be used in cultured cells of human origin. The first plasmid constructed for use in this system contained the entire coding region for CD4 and was intended to function as a control in the following studies. It is referred to as pRKCD4, and was constructed as follows:

pSVeCD4DHFR was digested with EcoRI and BamHI and fragment 15 containing the entire CD4 coding region was isolated. pRK5 (U.S. Ser. No. 97,472, filed Sep. 11, 1987) now abandoned was digested with EcoRI and BamHI and fragment 16 comprising the bulk of the plasmid recovered by gel electrophoresis and ligated to fragment 15, and the ligation mixture was transformed into *E. coli* strain 294. The transformed culture was plated on ampicillin media plates and resistant colonies were selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct fragment. This plasmid is referred to as pRKCD4.

The next plasmid constructed was designed to direct the expression of the above-mentioned (Section 3) secreted derivative of CD4. The coding region of CD4 was fused

after amino acid residue 368 of mature CD4 to,a sequence from pBR322 which codes for 9 more residues before a translation termination codon. This removes the putative CD4 transmembrane and cytoplasmic domains, which are presumed to anchor CD4 to the cell surface. The plasmid is referred to as pRKCD4T (and produces a protein called CD4T), and was constructed as follows:

pSVeCD4DHFR was digested with HpaII, blunted with Klenow fragment and the four dNTPs, and digested with BstEII. The 382 bp fragment (fragment 17) containing part of the CD4 coding sequence was recovered by gel electrophoresis. pSVeCD4DHFR was digested with EcoRI and BstEII and the 874 bp fragment (fragment 18) recovered. pBR322 was digested with HindIII, blunted with Klenow fragment and the four dNTPs, and digested with EcoRI. Fragment 19 comprising the bulk of the plasmid was isolated and ligated to fragments 17 and 18 and the ligation mixture transformed into *E. coli* strain 294. The transformed culture was plated on ampicillin media plates and resistant colonies were selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct fragment. This plasmid is referred to as pCD4Tint.

pRK5 was digested with EcoRI and SmaI and fragment 20 comprising the bulk of the plasmid isolated. pCD4Tint was digested with EcoRI and EcoRV and the 1410 bp fragment containing the CD4 coding sequence to the HpaII site at 1176 bp 3' of the initiating codon and the 154 bp HindIII-EcoRV fragment of pBR322 was recovered (fragment 21). Fragments 20 and 21 were ligated and the ligation mixture was transformed into *E. coli* strain 294. The transformed culture was plated on ampicillin media plates and resistant colonies were selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct fragment. This plasmid is referred to as pRKCD4T.

Section 5a

In order to create a secreted form of CD4 which could be purified with an antibody directed to herpes virus type glycoprotein D, a plasmid was constructed to express a derivative of CD4T in which the region coding for the mature, processed CD4T polypeptide was fused to a sequence coding for the signal peptide and the first 27 residues of the mature type I Herpes Simplex Virus gD glycoprotein. This plasmid is referred to as pRKGDCD4T, and was constructed as follows:

pgDTrunc.DHFR was digested with EcoRI and PvuII and the fragment containing the coding region for the signal peptide and first 27 residues of the mature HSV I gD glycoprotein was isolated (fragment 22). pRKCD4T was digested with EcoRI and BstEII and fragment 23 containing the 3' end of the CD4 coding sequence and the pRK5 region was isolated.

Synthetic oligonucleotides GD (adaptors 1–2, below) containing the coding sequence of CD4 from the codon for the amino terminal residue of mature CD4 to the Rsa site at 121 bp 3' of translation initiation, and containing the sequence CTGCTCGAG at the 5' end of the sense strand were prepared (fragment 24). pRKCD4 was digested with RsaI and BstEII and the 665 bp fragment containing part of the coding region for CD4 was recovered (fragment 25) and ligated to fragment 24. After digestion with BstEII to ensure that only monomeric fragment was present, the 724 bp fragment containing both sequences was recovered by gel electrophoresis (fragment 26).

Fragments 22, 23 and 26 were ligated and the ligation mixture was transformed into *E. coli* strain 294. The transformed culture was plated on ampicillin media plates and resistant colonies were selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct fragment. The sequence of several transformants was checked to ensure that the synthetic insert was correct and that reading frame was preserved. This plasmid is referred to as pRKGDCD4T.

These pRK5 derived plasmids preferably were transfected into 293S cells for stable expression according to Mussing, et al. Cell 48:691 (1987), with the exception that in addition to the plasmid of interest a plasmid expressing the neomycin resistance gene pRSV neo (Gorman et al. Science 221:553–555 [1985]) was cotransfected. 293 cells also are used satisfactorily as host cells. 2 days after transfection, the cells were passaged into standard medium (1:1 F12/DME supplemented with L-glutamine, penicillin-streptomycin and 10% FBS) with 0.5 mg/ml G418 (Geneticin sulfate; Gibco) for selection of stable cell lines, rather than in media containing methotrexate as shown by Mussing et al. Cells were assayed for production of CD4 or CD4 analogs by radioimmunoprecipitation. Binding studies (section 5c) used conditioned supernatants from these cells in the 1:1 F12/DME medium. Materials used in infectivity assays (section 5b) were obtained as described in section 8 below.

gDCD4 adaptor 1: CTGCTCGAGCAGGGAAA-CAAAGTGGTGCTGGGCAAAAAAGGG-GATACAGTGGAACTGAC (SEQ ID NO: 21)

gDCD4 adaptor 2: pACAGGTCAGTTCCACTGTATC-CCCTTTTTTGCCCAGCAC-CACTTTGTTTCCCTGCTCGA (SEQ ID NO: 22)

Section 5b

The following constitutes a study of the neutralization of HIV-1 infectivity by soluble CD4 analogs. A modification of the neutralization procedure of Robert-Guroff et al. Nature 316:72 (1985) was followed. Equal volumes of inhibitor supernatant and virus (60 microliters) were,incubated at 4 degrees C. for 1 hour, then the same volume of H9 (Gallo et al., Science 224:500, [1984]) at 5×10^6/ml was added and incubation continued for 1 hour at 37 degrees C. Following absorption, 2.5×10^5 cells in 150 microliters were transferred to 2 ml of incubation media. After 4 days at 37 degrees C., the cultures were split 1:2 with fresh media and incubated for an additional 3 days. Cultures were harvested, reverse transcriptase activity was measured (Groopman et al., AIDS Research and Human Retroviruses 3:71, [1987]), and immunofluorescence reactivity with HIV-1 positive serum was determined as described (Poiesz et al., Proc. Acad. Nat. Sci. USA 77:7415, [1980]). Inhibitor supernatants were obtained from confluent plate cultures of 293S/CDT4cells, 293S/gDCD4T cells or untransfected 293S cells by replacing the growth medium with incubation media and harvesting the supernatants 24 hours later. Inhibitor supernatant replaced part or all of the incubation media during the first three days of culture as indicated in the second column of Table III. Challenge dose of virus was 100 $TCID_{50}$ (Groopman et al., supra) of HIV-1 strain HTLV-IIIB grown in H9 cells assayed in the same system. Incubation media consisted of RPMI 1640 media containing 2 mM L-glutamine, 100 units/ml penicillin, 100 micrograms/ml streptomycin, 2 micrograms/ml polybrene and 20% fetal calf serum (M.A. Bioproducts).

TABLE III

Inhibitor supernatant	Dilution of inhibitor supernatant	Indirect immunofluorescence (% positive cells)		Reverse transcriptase (cpm/ml × 10^5)	
mock-transfected	undil.; 1:4	65.3	65.5	21.8	23.9
mock-transfected	undil.; 1:4	61.2	61.1	18.5	28.1
CD4T	undil.; 1:4	0.4	18.0	0.11	5.94
CD4T	undil.; 1:4	0.8	16.1	0.15	3.72
gDCD4T	undil.; 1:4	6.4	26.8	0.14	9.92
gDCD4T	undil.; 1:4	1.4	36.1	0.23	11.3

Both forms of soluble CD4 virtually abolished the growth of HIV-1, when incubated with virus-infected cells without prior dilution (Table III). At a dilution of 1:4 the soluble CD4 preparations were only partially effective in inhibiting virus growth; however, the level of fluorescent-positive cells and reverse transcriptase was still significantly lower than cultures receiving mock-transfected cell supernatants (Table III). Since there was no significant difference in virus growth between diluted and undiluted control supernatants, nor did any of the supernatants affect the growth of uninfected H9 cells (data not shown), soluble CD4 proteins present in these supernatants were concluded to be responsible for the neutralization of HIV-1 infection of H9 cells.

Section 5c

To determine the affinity constant for interactions between gp120 and CD4 or CD4 variants, saturation binding analysis was carried out with soluble CD4 (supra) and detergent solubilized intact CD4 (Lasky et al. Cell 50:975 [1987]) employing radioiodinated gp120 labeled with lactoperoxidase. Binding reactions consisted of ^{125}I-gp120 (3 ng to 670 ng, 2.9 nCi/ng) incubated for 1 hour at 0 degrees C. with cell lysates containing intact CD4 (Lasky et al., supra.) or cell supernatants containing unlabeled CD4T or gDCD4T prepared as described in section 5a. Reactions (0.2 ml) had a final composition of 0.5× McDougal Lysis Buffer (McDLB) (1× McDLB contains 0.5% Nonidet NP-40, 0.2% Na deoxycholate, 0.12M NaCl, 0.02M Tris-HCl, pH 8.0) and were performed in duplicate, both in the presence or absence of 50 micrograms of unlabeled purified gp120 (74 fold or greater excess). Following incubation, bound gp120 was quantitated by immunoprecipitation and counted in a gamma counter. For immunoprecipitation, binding reaction solutions were preabsorbed with 5 microliters of normal rabbit serum for one hour at 0° C., and cleared with 40 microliters of Pansorbin (10% w/v, Calbiochem) for 30 minutes at 0 degrees C. Samples were then incubated overnight at 0 degrees C. with 2 microliters of normal serum or 5 microliters (0.25 microgram) of OKT4 monoclonal antibody (Ortho) followed by collection of immune complexes with 10 microliters of Pansorbin. Precipitates were washed twice in 1× McDLB and once in water, then eluted by eluting at 100 degrees C. for 2 minutes in sample buffer (0.12M Tris-HCl pH 6.8, 4% SDS, 0.7M mercaptoethanol, 20% glycerol, and 0.1% bromophenol blue). CD4 molecules were bound saturably by gp120, and yielded a simple mass action binding curve. Supernatants from mock-transfected cells gave a level of specifically bound gp120 less than 1% that found for supernatants containing soluble CD4. Scatchard analysis revealed a single class of binding sites on each molecule, with apparent dissociation constants (Kd) of 1.3×10^{-9}M, 0.83×10^{-9}M and 0.72×10^{-9}M for intact CD4, CD4T and gDCD4T, respectively. The values obtained for CD4-gp120 binding in solution are comparable to the affinity previously measured for gp120 binding to CD4 on whole cells (Kd-4.0×10^{-9}M. Lasky, Cell, supra);

Section 6

In order to produce secreted derivatives of CD4 which are free of extraneous amino acid residues, two plasmids were constructed for expression in 293 cells. The plasmids contain CD4 genes which have been truncated without the addition of extra residues, and are referred to as pRKCD4ΔNIa and pRKCD4TP (and produce proteins called CD4ΔNIa and CD4TP, respectively, and were constructed as follows:

Fragment 14 containing the CD4 gene with the 195 bp NIaIII restriction fragment deleted was ligated to fragment 16, which is pRK5 digested with EcoRI and BamHI. The ligation mixture was transformed into *E. coli* strain 294, the transformed culture plated on ampicillin media plates and resistant colonies were selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct fragment. The resulting plasmid is referred to as pRKCD4ΔNIa.

Synthetic DNA (5' CGT GAT AGA AGC TTT CTA GAG 3' (SEQ ID NO:23)) was made to attach to the HpaII site at 1176bp and so that when so attached it would terminate translation after amino acid residue 368 of mature CD4 (fragment 27). The other end of this fragment was designed to ligate to BamHI restriction fragments. pUCCD4 was digested with BstEII and HpaII and the 382 bp fragment containing part of the CD4 gene was recovered (fragment 28). Fragments 27 and 28 were ligated and then digested with BstEII to reduce dimerized fragments to monomers, and the resulting 401 bp fragment was recovered (fragment 29).

pRKCD4 was digested with BstII and BamHI and the fragment comprising the bulk of the plasmid (fragment 30) was isolated and ligated to fragment 29. The ligation mixture was transformed into *E. coli* strain 294, the transformed culture plated on ampicillin media plates and resistant colonies were selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct fragment. The resulting plasmid is referred to as pRKCD4TP. Both plasmids are transfected into 293 cells to generate stable variant CD4-expressing cell lines as described above.

Section 7

Two plasmids were constructed to direct the expression of secreted CD4 lacking extraneous amino acid residues in CHO cells. These are referred to as pSVeCD4ΔNIaSVDHFR and pSVeCD4TPSVDHFR (and encode proteins having the primary sequence of CD4ΔNIa and CD4TP), respectively, and were constructed as follows:

pE348HBV.E400D22 was digested with PvuI and EcoRI and the fragment containing the SV40 early promoter and part of the β-lactamase gene was recovered (fragment 31). pE348HBV.E400D22 was digested with PvuI and BamHI and the large fragment containing the balance of the β-lactamase gene as well as the SV40 early promoter and the DHFR gene was isolated (fragment 32).

Fragments 31 and 32 were ligated together with fragment 14 and transformed into *E. coli* strain 294. The transformed culture was plated on ampicillin media plates and resistant colonies were selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct fragment. The resulting plasmid is referred to as pSVECD4ΔNIaSVDHFR. This plasmid contains the same DNA fragment encoding the soluble CD4 molecule found in the above-mentioned plasmid pSVeCD4ΔNIaDHFR (Section 2).

pRKCD4TP was digested with EcoRI and BamHI and the fragment containing the truncated CD4 coding region was

isolated and ligated to fragments 31 and 32. The ligation mixture was transformed into *E. coli* strain 294, the transformed culture plated on ampicillin media plates and resistant colonies were selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct fragment. The resulting plasmid is referred to as pSVeCD4TPSVDHFR. Both of these plasmids are transfected into CHO cells and amplified transfectants selected by methotrexate using conventional procedures.

EXAMPLE 2

Fusions of the V region of the CD4 gene, which is homologous to the variable region of immunoglobulin genes Maddon et al. 1985) supra, to the constant (C) region of human immunoglobulin κ and γ2 chains are constructed as follows:

Synthetic DNA is made to code for the C region of human κ chain (residues 109–214) based on the sequence published by Morin et al., Proc. Natl. Acad. Sci. 82:7025–7029 (1985), with the addition at the 5' end of the coding strand of the sequence GGGG, which allows this fragment to be ligated to the BspMI site at the end of the putative V-like region of CD4. At the 3' end of the coding region, a translational stop codon is added as well as a sequence which allows this end to be ligated to BamHI restriction fragments. The synthetic DNA is made in 8 fragments, 4 for each strand, 70–90 bases long. These are then allowed to anneal and are ligated prior to isolation on a polyacrylamide gel (fragment 33).

pRKCD4 is digested with EcoRI and BspMI and the 478bp fragment containing the region coding for the putative V-like domain of CD4 is recovered (fragment 34). Fragments 33 and 34 are ligated together with fragment 16 (from the expression vector pRK5). The ligation mixture is transformed into *E. coli* strain 294, the transformed culture plated on ampicillin media plates and resistant colonies are selected. Plasmid DNA is prepared from transformants and checked by restriction analysis for the presence of the correct fragment. The resulting plasmid is referred to as pRKCD4Ck.

A plasmid encoding a fusion of the CD4 V-like domain to the human immunoglobulin Cγ2 region is constructed in a similar fashion, and is referred to as pRKCD4Cγ2. Both of these plasmids are transfected into 293 cells, myeloma cells or other competent cells in order to obtain cell lines expressing variant CD4 molecules as described above.

EXAMPLE 3

The gDCD4T secreted by the method of Example 1 was purified from cell culture fluid containing either 10% FBS (fetal bovine serum) or no added FBS. The conditioned cell culture fluid was first concentrated by ultrafiltration, then purified by immunoaffinity chromaography. The immunoaffinity column was produced by coupling murine monoclonal antibody 5B6 (whose epitope is on the HSV-1 gD portion of the gDCD4T molecule) to glyceryl coated controlled pore glass by the method of Roy et al., *J. Chromatogr.*, 303: 225–228 (1984). The concentrated cell culture fluid is applied directly to the column and the contaminating proteins are washed away with neutral pH buffer. The column is then washed with neutral buffer containing tetramethylammonium chloride followed by neutral buffer containing Tween 80. The bound gDCD4T is eluted from the column with buffer at pH3 containing Tween 80 (0.1% w/v) and is neutralized immediately as it is eluted. The eluted neutralized gDCD4T is then concentrated by ultrafiltration and dialyzed/diafiltered to exchange the buffer for a physiological salt solution containing Tween 80 at approximately 0.1% w/v.

If the detergent is not present the gDCD4T forms aggregates as evidenced by the ability of centrifugation at approximately 10,000 Xg for 2 minutes to remove the gDCD4T from the solution. Incubation of gDCD4T at 4° C. in 0.1M sodium acetate, 0.5M NaCl and 0.25M Tris at pH 7 together with BSA, Tween 80 or glycerol as candidate stabilizers showed that, in the absence of a stabilizer the gDCD4T gradually aggregated over the space of 12 days to the point where only about 60–70% of the protein was soluble. However, use of 0.1% w/v Tween 80 or (0.5 rag/ml BSA ensured that about 100% or 80%, respectively, of the gDCD4T remained soluble over this period. Surprisingly glycerol was ineffective as a stabilizer and produced results inferior even to the control at 8 days about 80% of the gDCD4T was aggregated when stored in the presence of glycerol.

EXAMPLE 4

Plasmids were constructed to direct the expression of proteins containing differing lengths of the amino-terminal, extracellular domain of CD4 fused to the constant region of human immunoglobulin γ1. These plasmids are referred to as $pRKCD4_{2\gamma1}$, $pRKCD4_{e4\ \gamma1}$, $pRKCD4_{2\gamma1}$, $pRKCD4_{e2\gamma1}$, $pRKCD4_{1\gamma1}$, and $pRKCD4_{e1\gamma1}$.

Plasmid $pRKCD4_{4\gamma1}$ contains the portion of the CD4 gene from the initiation codon to the fusion site after the codon for serine residue 366 of the mature CD4 polypeptide, immediately followed by the sequence coding for the constant region of human immunoglobulin γ1, starting at the codon for serine residue 114 of mature human immunoglobulin γ1 (Kabat et al. *Sequences of Proteins of Immunological Interest*, National Institute of Health, Bethesda, Md. [1987]).

Plasmid $pRKCD4_{e4\gamma1}$ contains the portion of the CD4 gene from the initiation codon to the fusion site after the codon for lysine residue 360 of the mature CD4 polypeptide immediately followed by the sequence coding for the constant region of human immunoglobulin γ1, starting at the codon for serine residue 114 of mature human immunoglobulin γ1 (Kabat et al. supra).

Plasmid $pRKCD4_{2\gamma1}$ contains the portion of the CD4 gene from the initiation codon to the fusion site after the codon for glutamine residue 180 of the mature CD4 polypeptide, immediately followed by the sequence coding for the constant region of human immunoglobulin γ1, starting at the codon for serine residue 114 of mature human immunoglobulin γ1 (Kabat et al.).

Plasmid $pRKCD4_{e}2\gamma1$ contains the portion of the CD4 gene from the initiation codon to the fusion site after the codon for leucine residue 177 of the mature CD4 polypeptide, immediately followed by the sequence coding for the constant region of human immunoglobulin γ1, starting at the codon for serine residue 114 of mature human immunoglobulin γ1 (Kabat et al., supra).

Plasmid $pRKCD4_{1\gamma1}$ contains the portion of the CD4 gene from the initiation codon to the fusion site after the codon for aspartic acid residue 105 of the mature CD4 polypeptide, immediately followed by the sequence coding for the constant region of human immunoglobulin γ1, starting at the codon for serine residue 114 of mature human immunoglobulin γ1 (Kabat et al., supra).

Plasmid pRKCD4$_{c1\gamma1}$ contains the portion of the CD4 gene from the initiation codon to the fusion sire after the codon for leucine residue 100 of the mature CD4 polypeptide, immediately followed by the sequence coding for the constant region of human immunoglobulin γ1, starting at the codon for serine residue 114 of mature human immunoglobulin γ1 (Kabat et al., supra).

Construction of these plasmids required the prior construction of plasmid pRKCD4TP/γ1. It was constructed as follows:

A cDNA clone coding for human immunoglobulin γ1 was obtained from a human spleen cDNA library (Clontech Laboratories, Inc.) using oligonucleotides based on the published sequence (Ellison et al., "Nucl. Acids Res." 10:4071–4079 [1982]), and an EcoRI-EagI fragment (the EcoRI site was contributed by a linker; see FIG. 4A,B) containing part of the variable and all of the constant region was obtained. This fragment was blunted with Klenow fragment, and recovered by gel electrophoresis (Fragment a1).

Plasmid pRKCD4TP-kk, encoding a substitutional variant of soluble CD4 (residues 1–368) containing a lysine residue instead of asparagine at position 1 of the mature polypeptide, was constructed from plasmid pRKCD4TP by site-directed mutagenesis. A synthetic oligonucleotide was made as a primer for a mutagenesis reaction to obtain the desired coding sequence. This was synthesized as a 51-mer which contained two silent mutations from the natural sequence in addition to the substitution mutation, and 21 bases on each side of the mutated codons:

(SEQ ID NO: 24)

5'- CCC TTT TTT GCC CAG CAC CAC CTT CTT GCC CTG AGT GGC TGC TGG GAG GAG-3'

Plasmid pRKCD4TP was transformed into *E. coli* strain SR101 and the transformed colonies were plated on ampicillin media plates. Resistant colonies were selected and grown in the presence of m13K07 helper bacteriophage to yield secreted, encapsidated single-stranded templates of pRKCD4TP. The single-stranded plasmid DNA was isolated and used as the template for mutagenesis reactions with the synthetic oligonucleotides described above as primers. The mutagenesis reaction was transformed into *E. coli* SR101 and the transformed culture plated on ampicillin media plates. Transformants were screened by colony hybridization (ref. Grunstein-Hogness) for the presence of the appropriate sequence, using the following 16 mer as the probe.

5'- C CAC CTT CTT GCC CTG -3' (SEQ ID NO: 25) The hybridization conditions chosen were sufficiently stringent that the probe only detects the correctly fused product. Colonies identified as positive were selected and plasmid DNA was isolated and transformed into *E. coli* strain SR101. The transformed cultures were plated on ampicillin media plates, and resistant colonies were selected and grown in the presence of m13K07 bacteriophage. Templates were prepared as above and screened by sequencing.

Plasmid pRKCD4TP-kk was digested with XbaI real and treated with Klenow Enzyme, and Fragment a2, containing the linearized plasmid, was recovered by gel electrophoresis, and ligated with fragment a1. The ligation mixture was transformed into *E. coli* strain 294, the transformed culture plated on ampicillin media plates and resistant colonies were selected. Plasmid DNA was prepared from the transformants and checked by restriction analysis for the presence of the correct fragment in the correct orientation (i.e., the immunoglobulin coding region in the same orientation as the CD4 coding region, and at the 3' end of the CD4 coding region). This plasmid is referred to as pRKCD4TP/γ1.

Synthetic oligonucleotides were made as primers for deletional mutagenesis reactions to fuse the appropriate coding sequences of IgG1 and CD4 as described above. These were synthesized as 48-mers comprising 24 nucleotides on each side of the desired fusion site (i.e., corresponding to the COOH-terminal 8 residues of the desired CD4 moiety, and the NH_2-terminal 8 residues of the desired immunoglobulin moiety). Plasmid pRKCD4TP/γ1 was transformed into *E. coli* strain SR101 and the transformed cultures plated on ampicillin media plates. Resistant colonies were selected and grown in the presence of m13K07 helper bacteriophage to yield secreted, encapsidated single-stranded templates of pRKCD4TP/γ1. The single-stranded plasmid DNA was isolated and used as the template for mutagenesis reactions with the synthetic oligonucleotides described above as primers. The mutagenesis reactions were transformed into *E. coli*. SR101 and the transformed culture was plated on ampicillin media plates. Transformants were screened by colony hybridization (ref. Grunstein-Hogness) for the presence of the appropriate fusion site, using 16 mers as probes. These 16 mers comprise 8 bases on either side of the fusion site, and the hybridization conditions chosen were sufficiently stringent that the probes only detect the correctly fused product. Colonies identified as positive were selected and plasmid DNA was isolated and transformed into *E. coli* strain SR101. The transformed cultures were placed on ampicillin media plates, and resistant colonies were selected and grown in the presence of m13K07 bacteriophage. Templates were prepared as above and screened by sequencing.

The plasmids were transfected into 293 cells using standard procedures and assayed for expression and production as described above.

	Expressed	Secreted
pRKCD4$_{1\gamma1}$	+	–
pRKCD4$_{\alpha2\gamma1}$	+	+
pRKCD4$_{2\gamma1}$	+	+
pRKCD4$_{\alpha4\gamma1}$	+	+
pRKCD4$_{4\gamma1}$	+	+

Plasmids also were constructed to direct the expression of fusion proteins containing differing lengths of the amino-terminal, extracellular domain of CD4 fused to the truncated portion of the constant region of human immunoglobulin γ1, comprising only the hinge region and constant domains CH_2 and CH_3.

Synthetic oligonucleotides were made as primers for mutagenesis reactions to delete the immunoglobulin sequence from Ser114 to Cys215 inclusive (Kabat et al., supra). These were synthesized as 48-mers comprising 24 nucleotides on each side of the desired fusion site (i.e., corresponding to the COOH-terminal 8 residues of the desired CD4 moiety, and the NH_2-terminal 8 residues of the desired immunoglobulin moiety). Plasmids pRKCD4$_{4\gamma1}$, pRKCD4 $_{2\gamma1}$ and pRKCD4$_{1\gamma1}$ were separately transformed into *E. coli* strain SR101 and the transformed culture was plated on ampicillin media plates. Resistant colonies were selected and grown in the presence of m13K07 helper bacteriophage to yield secreted, encapsidated single-stranded templates of these plasmids. The single-stranded plasmid DNA was isolated and used as the template for mutagenesis reactions with the synthetic oligonucleotides described above as primers. The mutagenesis reactions were transformed into *E. coli* SR101 and the transformed culture

was plated on ampicillin media plates. Transformants were screened by colony hybridization (Grunstein-Hogness) for the presence of the appropriate fusion site, using 16 mere as probes. These 16 mere comprise 8 bases on either side of the fusion site, and the hybridization conditions chosen were sufficiently stringent that the probes only detect the correctly fused product. Colonies identified as positive were selected and plasmid DNA was isolated and transformed into *E. coli* strain SR101. The transformed cultures were plated on ampicillin media plates, and resistant colonies were selected and grown in the presence of m13K07 bacteriophage. Templates were prepared as above and screened by sequencing.

The plasmid derived from plasmid pRKCD4$_{4\gamma1}$ is referred to as pRKCD4$_{4Fc1}$, that derived from plasmid pRKCD4$_{2\gamma1}$ is referred to as pRKCD4$_{2Fc1}$ and that derived from plasmid pRKCD4$_{1\gamma1}$ is referred to as pRKCD4$_{1Fc1}$.

pRKCD4$_{2Fc1}$, pPKCD4$_{1Fc1}$ and pRKCD4$_{4Fc1}$ are cultured in the same fashion as described above and CH1-deleted CD4 immunoadhesons recovered as described elsewhere herein.

Light Chain Fusions

Plasmids were constructed to direct the expression of proteins containing differing lengths of the amino terminal, extracellular domain of CD4 fused to the constant region of human immunoglobulin κ. These plasmids are referred to as pRKCD4$_{4\kappa}$ and pRKCD4$_{c4\kappa}$.

Plasmid pRKCD4$_{4\kappa}$ contains the portion of the CD4 gene from the initiation codon to the fusion site after the codon for serine residue 366 of the mature CD4 polypeptide, immediately followed by the sequence for the constant region of human immunoglobulin κ, starting at the codon for threonine residue 109 of the mature human immunoglobulin κ. (Kabat et al., supra)

Plasmid pRKCD4$_{c4\kappa}$ contains the portion of the CD4 gene from the initiation codon to the fusion site after the codon for lysine residue 360 of the mature CD4 polypeptide, immediately followed by the sequence for the constant region of human immunoglobulin κ, starting at the codon for threonine residue 109 of the mature human immunoglobulin κ. (Kabat et al., supra)

These plasmids were constructed in a manner analogous to plasmids pRKCD4$_{4\gamma1}$ and pRKCD4$_{c4\gamma1}$ described above, with the following exception:

The human immunoglobulin κ coding sequence (FIG. 5) was obtained from a human spleen cDNA library (Clontech Laboratories, Inc.) using oligonucleotides based on the published sequence (Hieter, P. A. et al., Cell 22:197–207 [1980]) and an EcoRI-BspMI fragment containing part of the variable region and the entire constant region was obtained (see FIG. 5). This fragment was blunted with Klenow fragment and the four dNTPs. This fragment was used instead of fragment a1, and was used to construct plasmid pRKCD4TP/hκ.

Expression in CHO Cells

Plasmids were or are constructed to direct the expression of the immunoadhesons described above in CHO cells. These are referred to as pSVeCD4$_{4\gamma1}$SVDHFR, pSVeCD4$_{2\gamma1}$SVDHFR, pSVeCD4$_{1\gamma1}$SVDHFR, pSVeCD4$_{c4\gamma1}$SVDHFR, pSVeCD4$_{c2\gamma1}$SVDHFR, pSVeCD4$_{c1\gamma1}$SVDHFR, pSVeCD4$_{Fc1}$SVDHFR, pSVeCD4$_{2Fc1}$SVDHFR, pSVeCD4$_{1Fc1}$SVDHFR, pSVeCD4$_{4\kappa}$SVDHFR and pSVeCD4$_{c4\kappa}$SVDHFR.

Fragment 31 was prepared as described above. Fragment 32a was prepared by digesting plasmid pE348HBV.E400 D22 with BamHI, blunting with Klenow fragment and the four dNTPs, then digesting with PvuI and isolating the large fragment containing the balance of the β-lactamase gene and the SV40 early promoter and the DHFR gene. Plasmids pRKCD4$_{4\gamma1}$, pRKCD4$_{2\gamma1}$, pRKCD4$_{1\gamma1}$, pRKCD4$_{c4\gamma1}$, PRKCD4$_{c2\gamma1}$, PRKCD4$_{c}$1γ1, PRKCD4$_{Fc1}$, PRKCD4$_{2Fc1}$, PRKCD4$_{1Fc1}$, pRKCD4$_{4\kappa}$ and pRKCD4$_{c4\kappa}$ were separately digested with HindIII, blunted with Klenow fragment and the four dNTPs, then digested with EcoRI and the fragments encoding the CD4-Ig fusion protein were isolated. The resulting DNA fragments were ligated together with fragments 31 and 32a and transformed into *E. coli* strain 294. Colonies were selected and checked for the presence of the correct plasmid as above, then transfected into CHO cells and amplified by methotrexate selection using conventional procedures.

EXAMPLE 5

Culture, Purification and Formulation of CD4 Variants

Plasmids encoding soluble CD4 adhesons such as CD4T, CD4TP, or soluble CD4 immunoadhesons were calcium phosphate transfected into CHO-DP7 (a proinsulin-transformed autocrine host cell derived from CHO; U.S. Ser. No. 97,472, supra) and the transformants grown in selective medium (1:1 HAM F12/DMEM GHT$^-$ containing 1–10% diafiltered or dialyzed bovine serum). Other suitable host cells are CHO cells or 293S human embryonic kidney cells. The transformants were amplified by methotrexate selection in the same medium but containing 500 nm methotrexate. A subclone capable of secreting CD4TP, CD4tp 500 b, was selected. CD4tp 500 b is cultured in a DMEM/HAM F12 medium at about 37° C. until CD4TP accumulates in the culture, after which the medium is separated from the cells and insoluble matter by centrifuging.

Culture fluid from CD4TP transformants was concentrated and diafiltered to lower the ionic strength. The concentrate was passed through a large volume of Q-Sepharose anion exchange resin (previously equilibrated with 25 mM NaCl, pH 8.5) in order to adsorb contaminants from the culture fluid. The isoelectric point of CD4TP is about 9.5, thus making it possible to discriminate between truncated forms of CD4 and most contaminants by alternate adsorption, respectively, on a cation exchange resin such as carboxymethyl or sulfonyl Sepharose, and an anion exchange resin such as quaternary ammonium Sepharose. In addition, since highly electropositive domains are present in the extracellular segment of CD4, any CD4-containing variant is purified in the same fashion as CD4TP. The unadsorbed culture fluid from the anion exchange resin step was then passed through a cation exchange resin (previously equilibrated with 25 mM NaCl at pH 8.5) whereby CD4TP was adsorbed to the resin. The CD4TP was eluted with a NaCl gradient at pH 8.5, this CD4 variant eluting at about 0.2M NaCl. Ammonium sulfate was added to the eluate to a concentration of 1.7M and the solution passed through a column of hydrophobic interaction chromatography resin (phenyl or butyl Sepharose). The CD4TP was eluted from the hydrophobic interaction column with a gradient of ammonium sulfate, the CD4TP emerging at about 0.7M ammonium sulfate. The eluate was concentrated and buffer exchanged on a G-25 column using phosphate buffered saline containing 0.02% (w/v) Tween 20 or Tween 80. The CD4TP was soluble and stable in this solution, which was sterile filtered and filled into vials as an aqueous formulation. Other polymeric nonionic surfactants are suitably used with the CD4 formulations, including Plutonic block copolymers or polyethylene glycol.

It is also possible to employ immunoaffinity purification of soluble CD4 wherein the CD4 is adsorbed onto an immobilized antibody against CD4. This method suffers from the disadvantage that elution of the soluble CD4 under acidic conditions leads to protein aggregation that is only thoroughly ameliorated at relatively higher levels of surfactant. The foregoing procedure permits the use of much lower quantities of surfactant, about from 0.01 to 0.10% (w/v) surfactant.

The procedure followed for the purification of CD4 fusions with immunoglobulin heavy chain was to concentrate recombinant supernatants by ultrafiltration and thereafter adsorb the fusion onto resin-immobilized Staphylococcal protein A. The fusion was eluted with 0.1M titrate buffer pH 3 with no salt or detergent. This preparation is buffered into Tris buffer at pH 7.5. The immunoglobulin fusions with CD4 V1–V4 optionally are further purified by the procedure described above for unfused CD4 variants. CD4 immunoglobulin fusions with CD4 V1–V2 also may be purified by the procedure above, except that it is not expected that the isoelectric point of this class of molecules will be as alkaline as that of species containing all four V regions of CD4.

EXAMPLE 6

The characteristics of several adheson variants were determined. As shown in table IV the immunoadhesons $CD4_{4Y1}$ and $CD4_{2Y1}$ show improved plasma half-life in rabbits, coupled with high-affinity gp120 binding and an affinity for Fc_γ receptor (determined with U937 cells) that is comparable to that of bulk human IgG1.

TABLE IV

	gp120 KD (nM)#	FcγR KD (nM)+	Plasma Half-Life++ In Rabbits (Hrs.)
CD4T§	2.3 ± 0.4	Not detected	0.25
$CD4_{4Y1}$	1.2 ± 0.1	2.83 ± 0.25	6.4
$CD4_{2Y1}$	1.4 ± 0.1	3.01 ± 0.68	40.6
human IgG1	ND**	3.52 ± 0.5	21 days*

*determined in humans
+ND was determined by the method of Anderson et al., "J. Immunol." 125:2735–2741 (1980).
#determined by the method of Smith et al., "Science" 238:1704–07 (1987).
§residues 1–368 only
++The adheson variant was injected intravenously into rabbits and samples of blood were collected periodically and assayed for the presence of the adheson variant.
**Not done.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(1 i i) NUMBER OF SEQUENCES: 25

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 402 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
Met Asn Arg Gly Val Pro Phe Arg His Leu Leu Leu Val Leu Gln
 1               5                  10                  15

Leu Ala Leu Leu Pro Ala Ala Thr Gln Gly Asn Lys Val Val Leu
            20                  25                  30

Gly Lys Lys Gly Asp Thr Val Glu Leu Thr Cys Thr Ala Ser Gln
        35                  40                  45

Lys Lys Ser Ile Gln Phe His Trp Lys Asn Ser Asn Gln Ile Lys
    50                  55                  60

Ile Leu Gly Asn Gln Gly Ser Phe Leu Thr Lys Gly Pro Ser Lys
65                  70                  75

Leu Asn Asp Arg Ala Asp Ser Arg Arg Ser Leu Trp Asp Gln Gly
            80                  85                  90

Asn Phe Pro Leu Ile Ile Lys Asn Leu Lys Ile Glu Asp Ser Asp
        95                  100                 105

Thr Tyr Ile Cys Glu Val Glu Asp Gln Lys Glu Glu Val Gln Leu
    110                 115                 120

Leu Val Phe Gly Leu Thr Ala Asn Ser Asp Thr His Leu Leu Gln
125                 130                 135

Gly Gln Ser Leu Thr Leu Thr Leu Glu Ser Pro Pro Gly Ser Ser
```

-continued

```
                    140                 145                 150

Pro Ser Val Gln Cys Arg Ser Pro Arg Gly Lys Asn Ile Glu Gly
                155                 160                 165

Gly Lys Thr Leu Ser Val Ser Gln Leu Glu Leu Gln Asp Ser Gly
                170                 175                 180

Thr Trp Thr Cys Thr Val Leu Gln Asn Gln Lys Lys Val Glu Phe
                185                 190                 195

Lys Ile Asp Ile Val Val Leu Ala Phe Gln Lys Ala Ser Ser Ile
                200                 205                 210

Val Tyr Lys Lys Glu Gly Glu Gln Val Glu Phe Ser Phe Pro Leu
                215                 220                 225

Ala Phe Thr Val Glu Lys Leu Thr Gly Ser Gly Glu Leu Trp Trp
                230                 235                 240

Gln Ala Glu Arg Ala Ser Ser Ser Lys Ser Trp Ile Thr Phe Asp
                245                 250                 255

Leu Lys Asn Lys Glu Val Ser Val Lys Arg Val Thr Gln Asp Pro
                260                 265                 270

Lys Leu Gln Met Gly Lys Lys Leu Pro Leu His Leu Thr Leu Pro
                275                 280                 285

Gln Ala Leu Pro Gln Tyr Ala Gly Ser Gly Asn Leu Thr Leu Ala
                290                 295                 300

Leu Glu Ala Lys Thr Gly Lys Leu His Gln Glu Val Asn Leu Val
                305                 310                 315

Val Met Arg Ala Thr Gln Leu Gln Lys Asn Leu Thr Cys Glu Val
                320                 325                 330

Trp Gly Pro Thr Ser Pro Lys Leu Met Leu Ser Leu Lys Leu Glu
                335                 340                 345

Asn Lys Glu Ala Lys Val Ser Lys Arg Glu Lys Ala Val Trp Val
                350                 355                 360

Leu Asn Pro Glu Ala Gly Met Trp Gln Cys Leu Leu Ser Asp Ser
                365                 370                 375

Gly Gln Val Leu Leu Glu Ser Asn Ile Lys Val Leu Pro Thr Trp
                380                 385                 390

Ser Thr Pro Ser Phe Asn Ala Val Val Tyr His Ser
                395                 400     402
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1416 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
AATTCAAGCC CAGAGCCCTG CCATTTCTGT GGGCTCAGGT CCCTACTGCT 50

CAGCCCCTTC CTCCCTCGGC AAGGCCACAA TGAACCGGGG AGTCCCTTTT 100

AGGCACTTGC TTCTGGTGCT GCAACTGGCG CTCCTCCCAG CAGCCACTCA 150

GGGAAACAAA GTGGTGCTGG GCAAAAAAGG GGATACAGTG GAACTGACCT 200

GTACAGCTTC CCAGAAGAAG AGCATACAAT TCCACTGGAA AAACTCCAAC 250

CAGATAAAGA TTCTGGGAAA TCAGGGCTCC TTCTTAACTA AAGGTCCATC 300

CAAGCTGAAT GATCGCGCTG ACTCAAGAAG AAGCCTTTGG GACCAAGGAA 350

ACTTTCCCCT GATCATCAAG AATCTTAAGA TAGAAGACTC AGATACTTAC 400
```

-continued

```
ATCTGTGAAG TGGAGGACCA GAAGGAGGAG GTGCAATTGC TAGTGTTCGG 450
ATTGACTGCC AACTCTGACA CCCACCTGCT TCAGGGGCAG AGCCTGACCC 500
TGACCTTGGA GAGCCCCCCT GGTAGTAGCC CCTCAGTGCA ATGTAGGAGT 550
CCAAGGGGTA AAAACATACA GGGGGGGAAG ACCCTCTCCG TGTCTCAGCT 600
GGAGCTCCAG GATAGTGGCA CCTGGACATG CACTGTCTTG CAGAACCAGA 650
AGAAGGTGGA GTTCAAAATA GACATCGTGG TGCTAGCTTT CCAGAAGGCC 700
TCCAGCATAG TCTATAAGAA AGAGGGGGAA CAGGTGGAGT TCTCCTTCCC 750
ACTCGCCTTT ACAGTTGAAA AGCTGACGGG CAGTGGCGAG CTGTGGTGGC 800
AGGCGGAGAG GGCTTCCTCC TCCAAGTCTT GGATCACCTT TGACCTGAAG 850
AACAAGGAAG TGTCTGTAAA ACGGGTTACC CAGGACCCTA AGCTCCAGAT 900
GGGCAAGAAG CTCCCGCTCC ACCTCACCCT GCCCCAGGCC TTGCCTCAGT 950
ATGCTGGCTC TGGAAACCTC ACCCTGGCCC TTGAAGCGAA AACAGGAAAG 1000
TTGCATCAGG AAGTGAACCT GGTGGTGATG AGAGCCACTC AGCTCCAGAA 1050
AAATTTGACC TGTGAGGTGT GGGGACCCAC CTCCCCTAAG CTGATGCTGA 1100
GTTTGAAACT GGAGAACAAG GAGGCAAAGG TCTCGAAGCG GGAGAAGGCG 1150
GTGTGGGTGC TGAACCCTGA GGCGGGGATG TGGCAGTGTC TGCTGAGTGA 1200
CTCGGGACAG GTCCTGCTGG AATCCAACAT CAAGGTTCTG CCCACATGGT 1250
CCACCCCGAG CTTTAATGCG GTAGTTTATC ACAGTTAAAT TGCTAACGCA 1300
GTCAGGCACC GTGTATGAAA TCTAACAATG CGCTCATCGT CATCCTCGGC 1350
ACCGTCACCC TGGATGCTGT AGGCATAGGC TTGGTTATGC CGGTACTGCC 1400
GGGCCTCTTG CGGGAT 1416
```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1416 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
TTAAGTTCGG GTCTCGGGAC GGTAAAGACA CCCGAGTCCA GGGATGACGA 50
GTCGGGGAAG GAGGGAGCCG TTCCGGTGTT ACTTGGCCCC TCAGGGAAAA 100
TCCGTGAACG AAGACCACGA CGTTGACCGC GAGGAGGGTC GTCGGTGAGT 150
CCCTTTGTTT CACCACGACC CGTTTTTTCC CCTATGTCAC CTTGACTGGA 200
CATGTCGAAG GGTCTTCTTC TCGTATGTTA AGGTGACCTT TTTGAGGTTG 250
GTCTATTTCT AAGACCCTTT AGTCCCGAGG AAGAATTGAT TTCCAGGTAG 300
GTTCGACTTA CTAGCGCGAC TGAGTTCTTC TTCGGAAACC CTGGTTCCTT 350
TGAAAGGGGA CTAGTAGTTC TTAGAATTCT ATCTTCTGAG TCTATGAATG 400
TAGACACTTC ACCTCCTGGT CTTCCTCCTC CACGTTAACG ATCACAAGCC 450
TAACTGACGG TTGAGACTGT GGGTGGACGA AGTCCCCGTC TCGGACTGGG 500
ACTCGAACCT CTCGGGGGGA CCATCATCGG GGAGTCACGT TACATCCTCA 550
GGTTCCCCAT TTTTGTATGT CCCCCCCTTC TGGGAGAGGC ACAGAGTCGA 600
CCTCGAGGTC CTATCACCGT GGACCTGTAC GTGACAGAAC GTCTTGGTCT 650
TCTTCCACCT CAAGTTTTAT CTGTAGCACC ACGATCGAAA GGTCTTCCGG 700
```

-continued

```
AGGTCGTATC AGATATTCTT TCTCCCCCTT GTCCACCTCA AGAGGAAGGG 750
TGAGCGGAAA TGTCAACTTT TCGACTGCCC GTCACCGCTC GACACCACCG 800
TCCGCCTCTC CCGAAGGAGG AGGTTCAGAA CCTAGTGGAA ACTGGACTTC 850
TTGTTCCTTC ACAGACATTT TGCCCAATGG GTCCTGGGAT TCGAGGTCTA 900
CCCGTTCTTC GAGGGCGAGG TGGAGTGGGA CGGGGTCCGG AACGGAGTCA 950
TACGACCGAG ACCTTTGGAG TGGGACCGGG AACTTCGCTT TTGTCCTTTC 1000
AACGTAGTCC TTCACTTGGA CCACCACTAC TCTCGGTGAG TCGAGGTCTT 1050
TTTAAACTGG ACACTCCACA CCCCTGGGTG GAGGGGATTC GACTACGACT 1100
CAAACTTTGA CCTCTTGTTC CTCCGTTTCC AGAGCTTCGC CCTCTTCCGC 1150
CACACCCACG ACTTGGGACT CCGCCCCTAC ACCGTCACAG ACGACTCACT 1200
GAGCCCTGTC CAGGACGACC TTAGGTTGTA GTTCCAAGAC GGGTGTACCA 1250
GGTGGGGCTC GAAATTACGC CATCAAATAG TGTCAATTTA ACGATTGCGT 1300
CAGTCCGTGG CACATACTTT AGATTGTTAC GCGAGTAGCA GTAGGAGCCG 1350
TGGCAGTGGG ACCTACGACA TCCGTATCCG AACCAATACG GCCATGACGG 1400
CCCGGAGAAC GCCCTA 1416
```

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 434 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```
Met Gly Gly Thr Ala Ala Arg Leu Gly Ala Val Ile Leu Phe Val
 1               5                  10                  15

Val Ile Val Gly Leu His Gly Val Arg Gly Lys Tyr Ala Leu Ala
            20                  25                  30

Asp Ala Ser Leu Lys Met Ala Asp Pro Asn Arg Phe Arg Gly Lys
        35                  40                  45

Asp Leu Pro Val Leu Asp Gln Leu Leu Glu Gln Gly Asn Lys Val
    50                  55                  60

Val Leu Gly Lys Lys Gly Asp Thr Val Glu Leu Thr Cys Thr Ala
65                  70                  75

Ser Gln Lys Lys Ser Ile Glu Phe His Trp Lys Asn Ser Asn Gln
                80                  85                  90

Ile Lys Ile Leu Gly Asn Gln Gly Ser Phe Leu Thr Lys Gly Pro
            95                  100                 105

Ser Lys Leu Asn Asp Arg Ala Asp Ser Arg Arg Ser Leu Trp Asp
        110                 115                 120

Gln Gly Asn Phe Pro Leu Ile Ile Lys Asn Leu Lys Ile Glu Asp
    125                 130                 135

Ser Asp Thr Tyr Ile Cys Glu Val Glu Asp Gln Lys Glu Glu Val
140                 145                 150

Gln Leu Leu Val Phe Gly Leu Thr Ala Asn Ser Asp Thr His Leu
                155                 160                 165

Leu Gln Gly Gln Ser Leu Thr Leu Thr Leu Glu Ser Pro Pro Gly
            170                 175                 180

Ser Ser Pro Ser Val Gln Cys Arg Ser Pro Arg Gly Lys Asn Ile
        185                 190                 195
```

-continued

```
Gln Gly Gly Lys Thr Leu Ser Val Ser Gln Leu Glu Leu Gln Asp
200                 205                 210

Ser Gly Thr Trp Thr Cys Thr Val Leu Gln Asn Gln Lys Lys Val
215                 220                 225

Glu Phe Lys Ile Asp Ile Val Val Leu Ala Phe Gln Lys Ala Ser
230                 235                 240

Ser Ile Val Tyr Lys Lys Glu Gly Glu Gln Val Glu Phe Ser Phe
245                 250                 255

Pro Leu Ala Phe Thr Val Glu Lys Leu Thr Gly Ser Gly Glu Leu
260                 265                 270

Trp Trp Gln Ala Glu Arg Ala Ser Ser Ser Lys Ser Trp Ile Thr
275                 280                 285

Phe Asp Leu Lys Asn Lys Glu Val Ser Val Lys Arg Val Thr Gln
290                 295                 300

Asp Pro Lys Leu Gln Met Gly Lys Lys Leu Pro Leu His Leu Thr
305                 310                 315

Leu Pro Gln Ala Leu Pro Gln Tyr Ala Gly Ser Gly Asn Leu Thr
320                 325                 330

Leu Ala Leu Glu Ala Lys Thr Gly Lys Leu His Gln Glu Val Asn
335                 340                 345

Leu Val Val Met Arg Ala Thr Gln Leu Gln Lys Asn Leu Thr Cys
350                 355                 360

Glu Val Trp Gly Pro Thr Ser Pro Lys Leu Met Leu Ser Leu Lys
365                 370                 375

Leu Glu Asn Lys Glu Ala Lys Val Ser Lys Arg Glu Lys Ala Val
380                 385                 390

Trp Val Leu Asn Pro Glu Ala Gly Met Trp Gln Cys Leu Leu Ser
395                 400                 405

Asp Ser Gly Gln Val Leu Leu Glu Ser Asn Ile Lys Val Leu Pro
410                 415                 420

Thr Trp Ser Thr Pro Ser Phe Asn Ala Val Val Tyr His Ser
425                 430                 434
```

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1508 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:5:

```
AAGCTTCAGC GCGAACGACC AACTACCCCG ATCATCAGTT ATCCTTAAGG  50
TCTCTTTTGT GTGGTGCGTT CCGGTATGGG GGGGACTGCC GCCAGGTTGG 100
GGGCCGTGAT TTTGTTTGTC GTCATAGTGG GCCTCCATGG GGTCCGCGGC 150
AAATATGCCT TGGCGGATGC CTCTCTCAAG ATGGCCGACC CCAATCGATT 200
TCGCGGCAAA GACCTTCCGG TCCTGGACCA GCTGCTCGAG CAGGGAAACA 250
AAGTGGTGCT GGGCAAAAAA GGGGATACAG TGGAACTGAC CTGTACAGCT 300
TCCCAGAAGA AGAGCATACA ATTCCACTGG AAAAACTCCA ACCAGATAAA 350
GATTCTGGGA AATCAGGGCT CCTTCTTAAC TAAAGGTCCA TCCAAGCTGA 400
ATGATCGCGC TGACTCAAGA AGAAGCCTTT GGGACCAAGG AAACTTTCCC 450
CTGATCATCA AGAATCTTAA GATAGAAGAC TCAGATACTT ACATCTGTGA 500
AGTGGAGGAC CAGAAGGAGG AGGTGCAATT GCTAGTGTTC GGATTGACTG 550
```

-continued

```
CCAACTCTGA CACCCACCTG CTTCAGGGGC AGAGCCTGAC CCTGACCTTG 600
GAGAGCCCCC CTGGTAGTAG CCCCTCAGTG CAATGTAGGA GTCCAAGGGG 650
TAAAAACATA CAGGGGGGGA AGACCCTCTC CGTGTCTCAG CTGGAGCTCC 700
AGGATAGTGG CACCTGGACA TGCACTGTCT TGCAGAACCA GAAGAAGGTG 750
GAGTTCAAAA TAGACATCGT GGTGCTAGCT TTCCAGAAGG CCTCCAGCAT 800
AGTCTATAAG AAAGAGGGGG AACAGGTGGA GTTCTCCTTC CCACTCGCCT 850
TTACAGTTGA AAAGCTGACG GGCAGTGGCG AGCTGTGGTG GCAGGCGGAG 900
AGGGCTTCCT CCTCCAAGTC TTGGATCACC TTTGACCTGA AGAACAAGGA 950
AGTGTCTGTA AAACGGGTTA CCCAGGACCC TAAGCTCCAG ATGGGCAAGA 1000
AGCTCCCGCT CCACCTCACC CTGCCCCAGG CCTTGCCTCA GTATGCTGGC 1050
TCTGGAAACC TCACCCTGGC CCTTGAAGCG AAAACAGGAA AGTTGCATCA 1100
GGAAGTGAAC CTGGTGGTGA TGAGAGCCAC TCAGCTCCAG AAAAATTTGA 1150
CCTGTCAGGT GTGGGGACCC ACCTCCCCTA AGCTGATGCT GAGTTTGAAA 1200
CTGGAGAACA AGGAGGCAAA GGTCTCGAAG CGGGAGAAGG CGGTGTGGGT 1250
GCTGAACCCT GAGGCGGGGA TGTGGCAGTG TCTGCTGAGT GACTCGGGAC 1300
AGGTCCTGCT GGAATCCAAC ATCAAGGTTC TGCCCACATG GTCCACCCCG 1350
AGCTTTAATG CGGTAGTTTA TCACAGTTAA ATTGCTAACG CAGTCAGGCA 1400
CCGTGTATGA AATCTAACAA TGCGCTCATC GTCATCCTCG GCACCGTCAC 1450
CCTGGATGCT GTAGGCATAG GCTTGGTTAT GCCGGTACTG CCGGGCCTCT 1500
TGCGGGAT 1508
```

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1508 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:6:

```
TTCGAAGTCG CGCTTGCTGG TTGATGGGGC TAGTAGTCAA TAGGAATTCC 50
AGAGAAAACA CACCACGCAA GGCCATACCC CCCCTGACGG CGGTCCAACC 100
CCCGGCACTA AAACAAACAG CAGTATCACC CGGAGGTACC CCAGGCGCCG 150
TTTATACGGA ACCGCCTACG GAGAGAGTTC TACCGGCTGG GGTTAGCTAA 200
AGCGCCGTTT CTGGAAGGCC AGGACCTGGT CGACGAGCTC GTCCCTTTGT 250
TTCACCACGA CCCGTTTTTT CCCCTATGTC ACCTTGACTG GACATGTCGA 300
AGGGTCTTCT TCTCGTATGT TAAGGTGACC TTTTTGAGGT TGGTCTATTT 350
CTAAGACCCT TTAGTCCCGA GGAAGAATTG ATTTCCAGGT AGGTTCGACT 400
TACTAGCGCG ACTGAGTTCT TCTTCGGAAA CCCTGGTTCC TTTGAAAGGG 450
GACTAGTAGT TCTTAGAATT CTATCTTCTG AGTCTATGAA TGTAGACACT 500
TCACCTCCTG GTCTTCCTCC TCCACGTTAA CGATCACAAG CCTAACTGAC 550
GGTTGAGACT GTGGGTGGAC GAAGTCCCCG TCTCGGACTG GGACTGGAAC 600
CTCTCGGGGG GACCATCATC GGGGAGTCAC GTTACATCCT CAGGTTCCCC 650
ATTTTTGTAT GTCCCCCCCT TCTGGGAGAG GCACAGAGTC GACCTCGAGG 700
```

-continued

```
TCCTATCACC GTGGACCTGT ACGTGACAGA ACGTCTTGGT CTTCTTCCAC 750
CTCAAGTTTT ATCTGTAGCA CCACGATCGA AAGGTCTTCC GGAGGTCGTA 800
TCAGATATTC TTTCTCCCCC TTGTCCACCT CAAGAGGAAG GGTGAGCGGA 850
AATGTCAACT TTTCGACTGC CCGTCACCGC TCGACACCAC CGTCCGCCTC 900
TCCCGAAGGA GGAGGTTCAG AACCTAGTGG AAACTGGACT TCTTGTTCCT 950
TCACAGACAT TTTGCCCAAT GGGTCCTGGG ATTCGAGGTC TACCCGTTCT 1000
TCGAGGGCGA GGTGGAGTGG GACGGGGTCC GGAACGGAGT CATACGACCG 1050
AGACCTTTGG AGTGGGACCG GGAACTTCGC TTTTGTCCTT TCAACGTAGT 1100
CCTTCACTTG GACCACCACT ACTCTCGGTG AGTCGAGGTC TTTTTAAACT 1150
GGACACTCCA CACCCCTGGG TGGAGGGGAT TCGACTACGA CTCAAACTTT 1200
GACCTCTTGT TCCTCCGTTT CCAGAGCTTC GCCCTCTTCC GCCACACCCA 1250
CGACTTGGGA CTCCGCCCCT ACACCGTCAC AGACGACTCA CTGAGCCCTG 1300
TCCAGGACGA CCTTAGGTTG TAGTTCCAAG ACGGGTGTAC CAGGTGGGGC 1350
TCGAAATTAC GCCATCAAAT AGTGTCAATT TAACGATTGC GTCAGTCCGT 1400
GGCACATACT TTAGATTGTT ACGCGAGTAG CAGTAGGAGC CGTGGCAGTG 1450
GGACCTACGA CATCCGTATC CGAACCAATA CGGCCATGAC GGCCCGGAGA 1500
ACGCCCTA 1508
```

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 371 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:7:

```
Val Thr Ala Ala Asp Thr Ala Val Tyr Tyr Cys Ala Arg Ala Thr
 1               5                  10                  15

Phe Cys Leu Trp Tyr Arg Glu Arg Pro Pro Cys Trp Ile Asp Pro
            20                  25                  30

Trp Gly Leu Gly Thr Leu Val Thr Val Ser Ser Ala Ser Thr Lys
        35                  40                  45

Gly Pro Ser Val Phe Pro Leu Ala Pro Ser Ser Lys Ser Thr Ser
    50                  55                  60

Gly Gly Thr Ala Ala Leu Gly Cys Leu Val Lys Asp Tyr Phe Pro
65                  70                  75

Glu Pro Val Thr Val Ser Trp Asn Ser Gly Ala Leu Thr Ser Gly
            80                  85                  90

Val His Thr Phe Pro Ala Val Leu Gln Ser Ser Gly Leu Tyr Ser
        95                  100                 105

Leu Ser Ser Val Val Thr Val Pro Ser Ser Ser Leu Gly Thr Gln
    110                 115                 120

Thr Tyr Ile Cys Asn Val Asn His Lys Pro Ser Asn Thr Lys Val
125                 130                 135

Asp Lys Lys Val Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys
140                 145                 150

Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe
155                 160                 165

Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr
170                 175                 180
```

-continued

```
Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro
                185                 190                 195

Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn
            200                 205                 210

Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg
        215                 220                 225

Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly
    230                 235                 240

Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro
245                 250                 255

Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro
            260                 265                 270

Gln Val Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn
        275                 280                 285

Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp
    290                 295                 300

Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr
305                 310                 315

Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu
            320                 325                 330

Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn
        335                 340                 345

Val Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr
    350                 355                 360

Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
365                 370 371
```

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1125 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:8:

```
GAATTCTGTC ACTGCCGCGG ACACGGCCGT ATATTACTGT GCGAGAGCCA 50
CCTTTTGCCT ATGGTACAGG GAGCGTCCCC CTTGTTGGAT CGACCCCTGG 100
GGCCTGGGAA CCCTGGTCAC CGTCTCCTCG GCCTCCACCA AGGGCCCATC 150
GGTCTTCCCC CTGGCACCCT CCTCCAAGAG CACCTCTGGG GGCACAGCGG 200
CCCTGGGCTG CCTGGTCAAG GACTACTTCC CCGAACCGGT GACGGTGTCG 250
TGGAACTCAG GCGCCCTGAC CAGCGGCGTG CACACCTTCC CGGCTGTCCT 300
ACAGTCCTCA GGACTCTACT CCCTCAGCAG CGTGGTGACC GTGCCCTCCA 350
GCAGCTTGGG CACCCAGACC TACATCTGCA ACGTGAATCA CAAGCCCAGC 400
AACACCAAGG TGGACAAGAA AGTTGAGCCC AAATCTTGTG ACAAAACTCA 450
CACATGCCCA CCGTGCCCAG CACCTGAACT CCTGGGGGGA CCGTCAGTCT 500
TCCTCTTCCC CCCAAAACCC AAGGACACCC TCATGATCTC CCGGACCCCT 550
GAGGTCACAT GCGTGGTGGT GGACGTGAGC CACGAAGACC CTGAGGTCAA 600
GTTCAAGTGG TACGTGGACG GCGTGGAGGT GCATAATGCC AAGACAAAGC 650
CGCGGGAGGA GCAGTACAAC AGCACGTACC GGGTGGTCAG CGTCCTCACC 700
GTCCTGCACC AGGACTGGCT GAATGGCAAG GAGTACAAGT GCAAGGTCTC 750
```

-continued

```
CAACAAAGCC CTCCCAGCCC CCATCGAGAA AACCATCTCC AAAGCCAAAG 800
GGCAGCCCCG AGAACCACAG GTGTACACCC TGCCCCCATC CCGGGATGAG 850
CTGACCAAGA ACCAGGTCAG CCTGACCTGC CTGGTCAAAG GCTTCTATCC 900
CAGCGACATC GCCGTGGAGT GGGAGAGCAA TGGGCAGCCG GAGAACAACT 950
ACAAGACCAC GCCTCCCGTG CTGGACTCCG ACGGCTCCTT CTTCCTCTAC 1000
AGCAAGCTCA CCGTGGACAA GAGCAGGTGG CAGCAGGGGA ACGTCTTCTC 1050
ATGCTCCGTG ATGCATGAGG CTCTGCACAA CCACTACACG CAGAAGAGCC 1100
TCTCCCTGTC TCCGGGTAAA TGAGTGCGAC GGCCG 1135
```

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1142 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:9:

```
CTTAAGACAG TGACGGCGCC TGTGCCGGCA TATAATGACA CGCTCTCGGT 50
GGAAAACGGA TACCATGTCC CTCGCAGGGG GAACAACCTA GCTGGGGACC 100
CCGGACCCTT GGGACCAGTG GCAGAGGAGC CGGAGGTGGT TCCCGGGTAG 150
CCAGAAGGGG GACCGTGGGA GGAGGTTCTC GTGGAGACCC CCGTGTCGCC 200
GGGACCCGAC GGACCAGTTC CTGATGAAGG GGCTTGGCCA CTGCCACAGC 250
ACCTTGAGTC CGCGGGACTG GTCGCCGCAC GTGTGGAAGG GCCGACAGGA 300
TGTCAGGAGT CCTGAGATGA GGGAGTCGTC GCACCACTGG CACGGGAGGT 350
CGTCGAACCC GTGGGTCTGG ATGTAGACGT TGCACTTAGT GTTCGGGTCG 400
TTGTGGTTCC ACCTGTTCTT TCAACTCGGG TTTAGAACAC TGTTTTGAGT 450
GTGTACGGGT GGCACGGGTC GTGGACTTGA GGACCCCCCT GGCAGTCAGA 500
AGGAGAAGGG GGGTTTTGGG TTCCTGTGGG AGTACTAGAG GGCCTGGGGA 550
CTCCAGTGTA CGCACCACCA CCTGCACTCG GTGCTTCTGG GACTCCAGTT 600
CAAGTTGACC ATGCACCTGC CGCACCTCCA CGTATTACGG TTCTGTTTCG 650
GCGCCCTCCT CGTCATGTTG TCGTGCATGG CCCACCAGTC GCAGGAGTGG 700
CAGGACGTGG TCCTGACCGA CTTACCGTTC CTCATGTTCA CGTTCCAGAG 750
GTTGTTTCGG GAGGGTCGGG GGTAGCTCTT TTGGTAGAGG TTTCGGTTTC 800
CCGTCGGGGC TCTTGGTGTC CACATGTGGG ACGGGGGTAG GGCCCTACTC 850
GACTGGTTCT TGGTCCAGTC GGACTGGACG GACCAGTTTC CGAAGATAGG 900
GTCGCTGTAG CGGCACCTCA CCCTCTCGTT ACCCGTCGGC CTCTTGTTGA 950
TGTTCTGGTG CGGAGGGCAC GACCTGAGGC TGCCGAGGAA GAAGGAGATG 1000
TCGTTCGAGT GGCACCTGTT CTCGTCCACC GTCGTCCCCT TGCAGAAGAG 1050
TACGAGGCAC TACGTACTAC GTACTCCGAG ACGTGTTGGT GATGTGCGTC 1100
TTCTCGGAGA GGGACAGAGG CCCATTTACT CACGCTGCCG GC 1142
```

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 143 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:10:

```
Thr Leu Thr Ile Ser Gly Leu Gln Pro Glu Asp Phe Ala Thr Tyr
 1               5                  10                  15

Tyr Cys Gln Gln Tyr Lys Ser Leu Ser Leu Thr Phe Gly Gly Gly
            20                  25                  30

Thr Lys Val Glu Ile Lys Arg Thr Val Ala Ala Pro Ser Val Phe
        35                  40                  45

Ile Phe Pro Pro Ser Asp Glu Gln Leu Lys Ser Gly Thr Ala Ser
    50                  55                  60

Val Val Cys Leu Leu Asn Asn Phe Tyr Pro Arg Glu Ala Leu Val
65                  70                  75

Gln Trp Leu Val Asp Asn Ala Leu Gln Ser Gly Asn Ser Gln Glu
            80                  85                  90

Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr Tyr Ser Leu Ser
        95                  100                 105

Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys His Lys Val
    110                 115                 120

Tyr Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser Pro Val Thr
125                 130                 135

Lys Ser Phe Asn Arg Gly Glu Cys
140                 143
```

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 468 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:11:

```
GAATTCACTC TCACCATCAG CGGCCTGCAG CCTGAAGATT TTGCAACTTA 50

TTACTGCCAA CAGTATAAGA GTTTGTCGCT CACTTTCGGC GGAGGGACCA 100

AGGTGGAGAT CAAACGAACT GTGGCTGCAC CATCTGTCTT CATCTTCCCG 150

CCATCTGATG AGCAGTTGAA ATCTGGAACT GCCTCTGTTG TGTGCCTGCT 200

GAATAACTTC TATCCCAGAG AGGCCAAAGT ACAGTGGAAG GTGGATAACG 250

CCCTCCAATC GGGTAACTCC CAGGAGAGTG TCACAGAGCA GGACAGCAAG 300

GACAGCACCT ACAGCCTCAG CAGCACCCTG ACGCTGAGCA AAGCAGACTA 350

CGAGAAACAC AAAGTCTACG CCTGCGAAGT CACCCATCAG GGCCTGAGCT 400

CGCCCGTCAC AAAGAGCTTC AACAGGGGAG AGTGTTAGAG GGAGAAGTGC 450

CCCCACCTGC TCCTCAGT 468
```

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 468 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:12:

```
CTTAAGTGAG AGTGGTAGTC GCCGGACGTC GGACTTCTAA AACGTTGAAT 50

AATGACGGTT GTCATATTCT CAAACAGCGA GTGAAAGCCG CCTCCCTGGT 100

TCCACCTCTA GTTTGCTTGA CACCGACGTG GTAGACAGAA GTAGAAGGGC 150
```

-continued

GGTAGACTAC TCGTCAACTT TAGACCTTGA CGGAGACAAC ACACGGACGA 200

CTTATTGAAG ATAGGGTCTC TCCGGTTTCA TGTCACCTTC CACCTATTGC 250

GGGAGGTTAG CCCATTGAGG GTCCTCTCAC AGTGTCTCGT CCTGTCGTTC 300

CTGTCGTGGA TGTCGGAGTC GTCGTGGGAC TGCGACTCGT TTCGTCTGAT 350

GCTCTTTGTG TTTCAGATGC GGACGCTTCA GTGGGTAGTC CCGGACTCGA 400

GCGGGCAGTG TTTCTCGAAG TTGTCCCCTC TCACAATCTC CCTCTTCACG 450

GGGGTGGACG AGGAGTCA 468

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 44 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:13:

AATTCAAGCC CAGAGCCCTG CCATTTCTGT GGGCTCAGGT CCCT 44

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 50 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:14:

ACTGGTCAGC CCCTTCCTCC CTCGGCAAGG CCACAATGAA CCGGGGAGTC 50

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 47 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:15:

CCTTTTAGGC ACTTGCTTCT GGTGCTGCAA CTGGCGCTCC TCCCAGC 47

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 61 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:16:

AGCCACTCAG GGAAACAAAG TGGTGCTGGG CAAAAAAGGG GATACAGTGG 50

AACTGACCTG T 61

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 51 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:17:

-continued

ACAGGTCAGT TCCACTGTAT CCCCTTTTTT GCCCAGCACC ACTTTGTTTC 50

C 51

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 47 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:18:

CTGAGTGGCT GCTGGGAGGA GCGCCAGTTG CAGCACCAGA AGCAAGT 47

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 50 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:19:

GCCTAAAAGG GACTCCCCGG TTCATTGTGG CCTTGCCGAG GGAGGAAGGG 50

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 50 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:20:

GCTGAGCAGT AGGGACCTGA GCCCACAGAA ATGGCAGGGC TCTGGGCTTG 50

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 59 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:21:

CTGCTCGAGC AGGGAAACAA AGTGGTGCTG GGCAAAAAAG GGGATACAGT 50

GGAACTGAC 59

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 59 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:22:

ACAGGTCAGT TCCACTGTAT CCCCTTTTTT GCCCAGCACC ACTTTGTTTC 50

CCTGCTCGA 59

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single

-continued

(D) TOPOLOGY: linear

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:23:

CGTGATAGAA GCTTTCTAGA G 21

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 51 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:24:

CCCTTTTTTG CCCAGCACCA CCTTCTTGCC CTGAGTGGCT GCTGGGAGGA 50

G 51

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:25:

CCACCTTCTT GCCCTG 16

We claim:

1. An immunoadheson comprising a fusion protein in which a polypeptide comprising an adheson variable (V) region is fused at its C-terminus to the N-terminus of a polypeptide comprising a constant region of an immunoglobulin chain.

2. The immunoadheson of claim 1 wherein the fusion protein is disulfide bonded to a companion immunoglobulin heavy chain-light chain pair bearing a V_LV_H antibody combining site capable of binding to a predetermined antigen.

3. The immunoadheson of claim 2 selected from the group consisting of AC_H-(V_LC_L-V_HC_H), (AC_LAC_H) - (V_LC_L-V_HC_H), (AC $_L$-V_HC_H) - (V_LC_L-V_HC_H), and (V_LC_L-AC_H) - (V_LC_L-V_HC_H), wherein A is an adheson V region, V_L and V_H are the variable domains of an immunoglobulin light and heavy chain, respectively, and C_L and C_H are the constant domains of an immunoglobulin light and heavy chain, respectively.

4. The immunoadheson of claim 2 wherein the V_LV_H antibody combining site is from an anti-ricin antibody.

5. The immunoadheson of claim 2 wherein the immunoglobulin sequences are from IgG-1 or IgG-3 subtypes.

6. The immunoadheson of claim 2 wherein the immunoglobulin sequences are from IgA or IgM.

7. The immunoadheson of claim 2 wherein the fusion protein comprises the V-J domains of the adheson.

8. The immunoadheson of claim 1 selected from the group consisting of AC_L, AC_L-AC_L, AC_H, AC_H-AC_H, AC_H-AC_LAC_H, AC_LAC_H-AC_LAC_H wherein A is an adheson variable region, and C_L and C_H are light and heavy chain constant regions respectively.

9. The immunoadheson of claim 8 wherein the immunoglobulin sequences are from IgG-1 or IgG-3 subtypes.

10. The immunoadheson of claim 8 wherein the immunoglobulin sequences are from IgA or IgM.

11. A composition comprising an immunoadheson comprising a fusion protein in which a polypeptide comprising an adheson variable (V) region is fused at its C-terminus to the N-terminus of a polypeptide comprising a constant region of an immunoglobulin chain.

12. The composition of claim 11 which is sterile and which further comprises a physiologically acceptable carrier.

13. The composition of claim 11 wherein said immunoadheson is disulfide bonded to a companion immunoglobulin heavy chain—light chain pair bearing a V_LV_H antibody combining site capable of binding to a predetermined antigen.

14. The composition of claim 13 which is sterile and which further comprises a physiologically acceptable carrier.

15. A method of preparing an immunoadheson comprising transfecting a host cell with nucleic acid encoding a fusion protein in which a polypeptide comprising an adheson variable region is fused at its C-terminus to the N-terminus of a polypeptide comprising a constant region of an immunoglobulin chain, and culturing the resulting transfected host cell under conditions which permit the expression of said immunoadheson by the transfected host cell.

16. The method of claim 15 wherein the heterofunctional immunoadheson is recovered from the host cell culture.

17. The method according to claim 15 wherein the host cell expresses nucleic acid encoding an immunoglobulin having a variable region directed against a predetermined antigen.

18. The method of claim 17 wherein the immunoadheson is recovered from the host cell culture.

* * * * *

EXHIBIT 15

08/11



US006117655A

United States Patent [19]

Capon et al.

[11] **Patent Number: 6,117,655**

[45] **Date of Patent: Sep. 12, 2000**

[54] **NUCLEIC ACID ENCODING ADHESION VARIANTS**

[75] Inventors: **Daniel J. Capon**, San Mateo; **Timothy J. Gregory**, Hillsborough, both of Calif.

[73] Assignee: **Genentech, Inc.**, South San Francisco, Calif.

[*] Notice: This patent is subject to a terminal disclaimer.

[21] Appl. No.: **08/457,918**

[22] Filed: **Jun. 1, 1995**

Related U.S. Application Data

[63] Continuation of application No. 08/236,311, May 2, 1994, Pat. No. 5,565,335, which is a continuation of application No. 07/936,190, Aug. 26, 1992, Pat. No. 5,336,603, which is a division of application No. 07/842,777, Feb. 18, 1992, abandoned, which is a continuation of application No. 07/250,785, Sep. 28, 1988, abandoned, which is a continuation-in-part of application No. 07/104,329, Oct. 2, 1987, abandoned.

[51] **Int. Cl.**[7] **C12N 15/13**; C12N 15/62; C12N 15/63

[52] **U.S. Cl.** **435/69.7**; 536/23.1; 536/23.4; 536/23.5; 536/23.53; 435/71.1; 435/71.2; 435/325; 435/252.3; 435/254.11; 435/320.1; 435/471

[58] **Field of Search** 435/69.7, 71.1, 435/71.2, 325, 252.3, 254.11, 320.1, 471; 536/23.4, 23.1, 23.5, 23.53

[56] **References Cited**

U.S. PATENT DOCUMENTS

3,959,080	5/1976	Orth et al.	435/179
4,002,531	1/1977	Royer	435/188
4,055,635	10/1977	Green et al.	424/400
4,301,144	11/1981	Iwashita	514/6
4,412,989	11/1983	Iwashita	530/402
4,431,739	2/1984	Riggs	435/257.3
4,444,878	4/1984	Paulus et al.	435/7.1
4,745,055	5/1988	Schenk	435/7.6
4,761,371	8/1988	Bell et al.	435/69.1
4,766,106	8/1988	Katre	514/12
4,816,567	3/1989	Cabilly et al.	530/387
4,879,211	11/1989	Wang et al.	435/5
5,116,964	5/1992	Capon et al.	435/64.7

FOREIGN PATENT DOCUMENTS

068763	1/1983	European Pat. Off. .
088695	9/1983	European Pat. Off. .
120694	10/1984	European Pat. Off. .
139416	5/1985	European Pat. Off. .
173494	3/1986	European Pat. Off. .
244221	11/1987	European Pat. Off. .
255694	2/1988	European Pat. Off. .
256654	2/1988	European Pat. Off. .
266663	5/1988	European Pat. Off. .
278776	8/1988	European Pat. Off. .
296786	12/1988	European Pat. Off. .
313377	4/1989	European Pat. Off. .
319815	6/1989	European Pat. Off. .
325262	7/1989	European Pat. Off. .
330227	8/1989	European Pat. Off. .
331356	9/1989	European Pat. Off. .
WO 85/03947	9/1985	WIPO .
WO 87/03600	6/1987	WIPO .
WO 88/01304	2/1988	WIPO .
WO 88/03559	5/1988	WIPO .
WO 88/07376	10/1988	WIPO .
WO 88/09344	12/1988	WIPO .
WO 89/01940	3/1989	WIPO .
WO 89/03221	4/1989	WIPO .

OTHER PUBLICATIONS

Miller et al., *Science* 244: 334–337, Apr. 21, 1989.
Blank et al, *Nature* 337: 187–189, Jan. 12, 1989.
Maddon et al., *Cell* 42: 93–104, Aug. 1985.
Treiger et al, *J. of Immunol.* 136(11): 4099–4105, Jun. 1, 1986.
Neuberger et al, *Nature* 312; 604–608, Dec. 13, 1987.
Gascoigne et al. "Secretion of a chimeric T-cell receptor-immunoglobulin protein". P.N.A.S. 84:2936–2940, May 1987.
Anderson et al., "Regulatory Interactions Between Members of the Immunoglobulin Superfamily" *Immunology Today* 9(7 & 8): 199–203 (1988).
Beauchamp et al., "A new procedure for the synthesis of polyethylene glycol–protein adducts; effects on function, receptor recognition, and clearance of superoxide dismutase, lactoferrin, and alpha 2–macroglobulin" *Analytical Biochemistry* 131(1): 25–33 (1983).
Blank et al., "Complete structure and expression in transfected cells of high affinity IgE receptor" *Nature* 337(6203): 187–189 (1989).
Boulianne, G. L. et al., "Production of functional chimaeric mouse/human antibody" *Nature* 312(5995): 643–646 (Dec. 1984).
Chaudhary et al., "Activity of a recombinant fusion protein between transforming growth factor type alpha and Pseudomonas toxin" *Proc. Natl. Acad. Sci. USA* 84:4538–4542 (1987).
Clark et al., "Peptide and nucleotide sequences for rat CD4 (W3/25) antigen: Evidence for derivation from a structure with four immunoglobulin–related domains" *Proc. Natl. Acad. Sci.* 84:1649–1653 (1987).

(List continued on next page.)

Primary Examiner—Prema Mertz
Attorney, Agent, or Firm—Jeffrey S. Kubinec

[57] **ABSTRACT**

Novel derivatives of cell surface proteins which are homologous to the immunoglobulin superfamily (adhesons) are provided. Amino acid sequence variations are introduced into adheson, the most noteworthy of which are those in which the transmembrane and, preferably, cytoplasmic domains are rendered functionally inactive, and in which adheson extracellular domains replace an immunoglobulin variable region. These variants are useful in therapy or diagnostics, in particular, CD4 variants are therapeutically useful in the treatment of HIV infections.

30 Claims, 13 Drawing Sheets

OTHER PUBLICATIONS

Estess et al., "Analysis of T-cell receptor structure and function using chimeric T-cell receptor/immunoglobulin molecules" *Journal of Cellular Biochemistry Supplement* 11(Part D, Abs. 331):258 (1987).

Falkner and Zachau, "Expression of mouse immunoglobulin genes in monkey cells" *Nature* 298:286–288 (1982).

Gascoigne et al., "Secretion of a chimeric T cell receptor–immunoglobulin protein" *Proc. Natl. Acad. Sci. USA* 84:2936–2940 (1987).

Gascoigne et al., "Secretion of chimeric T-cell receptor–immunoglobulin fusion proteins" *Journal of Cellular Biochemistry Supplement* 11(Part D, Abs. 333):259 (1987).

Goverman et al., "Chimeric T-cell receptor genes as tools in analyzing T-cell/target-cell interactions" *Journal of Cellular Biochemistry Supplement* 11(Part D, Abs. 334):259 (1987).

Hashimoto et al., "Rearrangement and expression of T cell receptor genes in pre-T cells" *Journal of Cellular Biochemistry Supplement* 11(Part D, Abs. 434):278 (1987).

Ivars et al., "Expression of cloned T cell receptor genes" *Journal of Cellular Biochemistry Supplement* 11(Part D, Abs. 435):278 (1987).

Johnson et al., "Correlation of T cell receptor structure and function" *Journal of Cellular Biochemistry Supplement* 11(Part D, Abs. 337):260 (1987).

Kohler, G., "Immunoglobulin chain loss in hybridoma lines" *Proc. Natl. Acad. Sci. USA* 77(4):2197–2199 (1980).

Littman DR, "The isolation and sequence of the gene encoding T8: a molecule defining functional classes of T lymphocytes" *Cell* 40(2):237–246 (1985).

Littman DR, "The structure of the CD4 and CD8 genes" *Annual Review of Immunology* 5:561–584 (1987).

Maddon et al., "The Isolation and Nucleotide Sequence of a cDNA Encoding the T Cell Surface Protein T4: A New Member of the Immunoglobulin Gene Family" *Cell* 42:93–104 (1985).

Maddon et al., "The T4 Gene Encodes the AIDS Virus Receptor and Is Expressed in the Immune System and the Brain" *Cell* 47:333–348 (1986).

Maddon P. et al., "Structure and expression of the human and mouse T4 genes" *Proc. Natl. Acad. Sci. USA* 84(24):9155–9159 (1987).

McDougal et al., "Binding of HTLV-III/LAV to T4$^+$ Cells by a Complex of the 110k Viral Protein and the T4 Molecule" *Science* 231:382–385 (1986).

McDougal et al., "The T4 Glycoprotein Is a Cell-Surface Receptor for the AIDS Virus" *Molecular Biology of Homosapiens* (Cold Spring Harbor Symposia on Quantitative Biology), New York:CSHL vol. LI (1986).

Miller, Larry et al., "Expression of high-affinity binding of human immunoglobulin E by transfected cells" *Science* 244:334–336 (1989).

Morrison, "Transfectomas Provide Novel Chimeric Antibodies" *Science* 229:1202–1207 (1985).

Morrison and Oi, "Transfer and expression of immunoglobulin genes" *Annual Review of Immunology* 2:239–256 (1984).

Morrison, J., "Sequentially derived mutants of the constant region of the heavy chain of murine immunoglobulins" *J. of Immunology* 123(2):793–800 (1979).

Morrison, S. L. et al., "Chimeric human antibody molecules: mouse antigen–binding domains with human constant region domains" *Proc. Natl. Acad. Sci. USA* 81(21):6851–6855 (Nov. 1984).

Murre et al., "Biochemical and functional analyses of a secreted H–2L^d molecule" *Molecular & Cellular Biology* 6:1315–1319 (1986).

Neuberger et al., "Recombinant antibodies possessing novel effector functions" *Nature* 312(5995):604–608 (Dec. 1984).

Oi, V.T. et al., "Chimeric Antibodies" *BioTechniques* 4(3):214–221 (1986).

Osborn, L. et al., "Direct expression cloning of vascular cell adhesion molecule 1, a cytokine–induced endothelial protein that binds to lymphocytes" *Cell* 59:1203–1211 (1989).

Peterson, A.S., "Genetic analysis of CD2/LFA–3 and CD4/HIV interactions" (thesis), Harvard University, Chapter 1, (1988).

Rose and Bergmann, "Expression from Cloned cDNA of Cell-Surface Secreted Forms of the Glycoprotein of Vesicular Stomatitis Virus in Eucaryotic Cells" *Cell* 30:753–762 (1982).

Rosenblum et al., "Modification of human leukocyte interferon pharmacology with a monclonal antibody" *Cancer Research* 45:2421–2424 (1985).

Salzawa et al., "The CD4 molecule is associated with the T-cell receptor" *Journal of Cellular Biochemistry Supplement* 11(Part D, Abs. 421):273 (1987).

Seed B., "An LFA–3 cDNA encodes a phospholipid-linked membrane protein homologous to its receptor CD2" *Nature* 329:840–842 (1987).

Sharon et al., "Expression of a V_hC_k Chimaeric Protein in Mouse Myeloma Cells" *Nature* 309:364–367 (1984).

Sleckman et al., "Expression and function of CD4 in a murine T-cell hybridoma" *Nature* 328(6128):351–353 (1987).

Smith et al., "Blocking of HIV-1 Infectivity by a Soluble, Secreted Form of the CD4 Antigen" *Science* 238:1704–1707 (1987).

Springer and Strominger, "Detergent-soluble HLA antigens contain a hydrophilic region at the COOH-terminus and a penultimate hydrophobic region" *Proc. Natl. Acad. Sci. USA* 73(7):2481–2485 (1976).

Terhorst C. et al., "Biochemical analysis of human T lymphocyte differentiation antigens T4 and T5" *Science* 209(4455):520–521 (1980).

Traunecker et al., "Highly efficient neutralization of HIV with recombinant CD4–immunoglobulin molecules" *Nature* 339:68–70 (1989).

Traunecker et al., "Soluble CD4 molecules neutralize human immunodeficiency virus" *Nature* 331:84–86 (1988).

Vitetta et al., "Redesigning nature's poisons to create anti-tumor reagents" *Science* 238:1098–1104 (1987).

Wills et al., "Mutations of the Rous Sarcoma Virus env Gene that affect the transport and subcellular location of the glycoprotein products" *Journal of Cell Biology* 99:2011–2023 (1984).

FIG. 1A

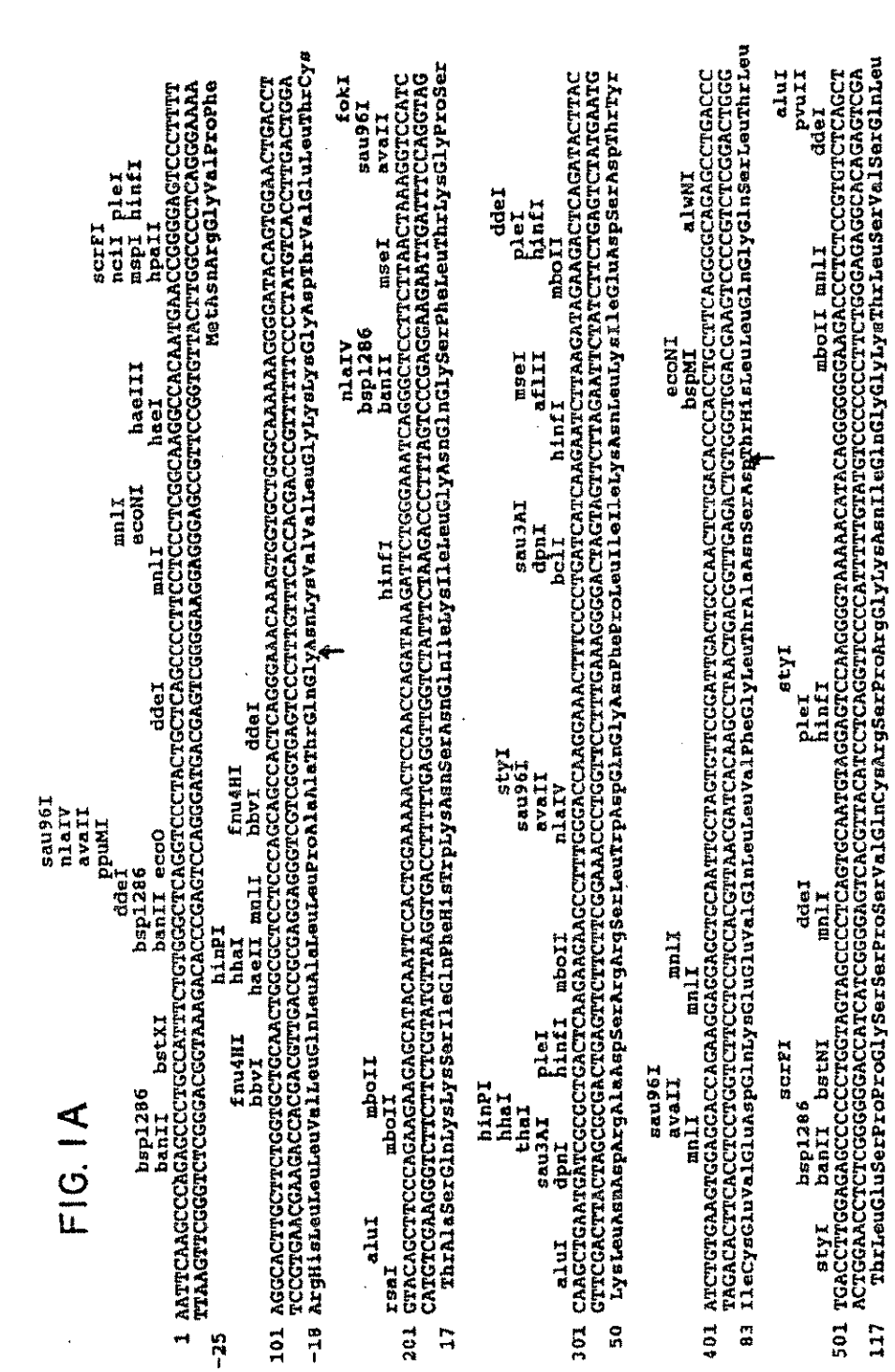
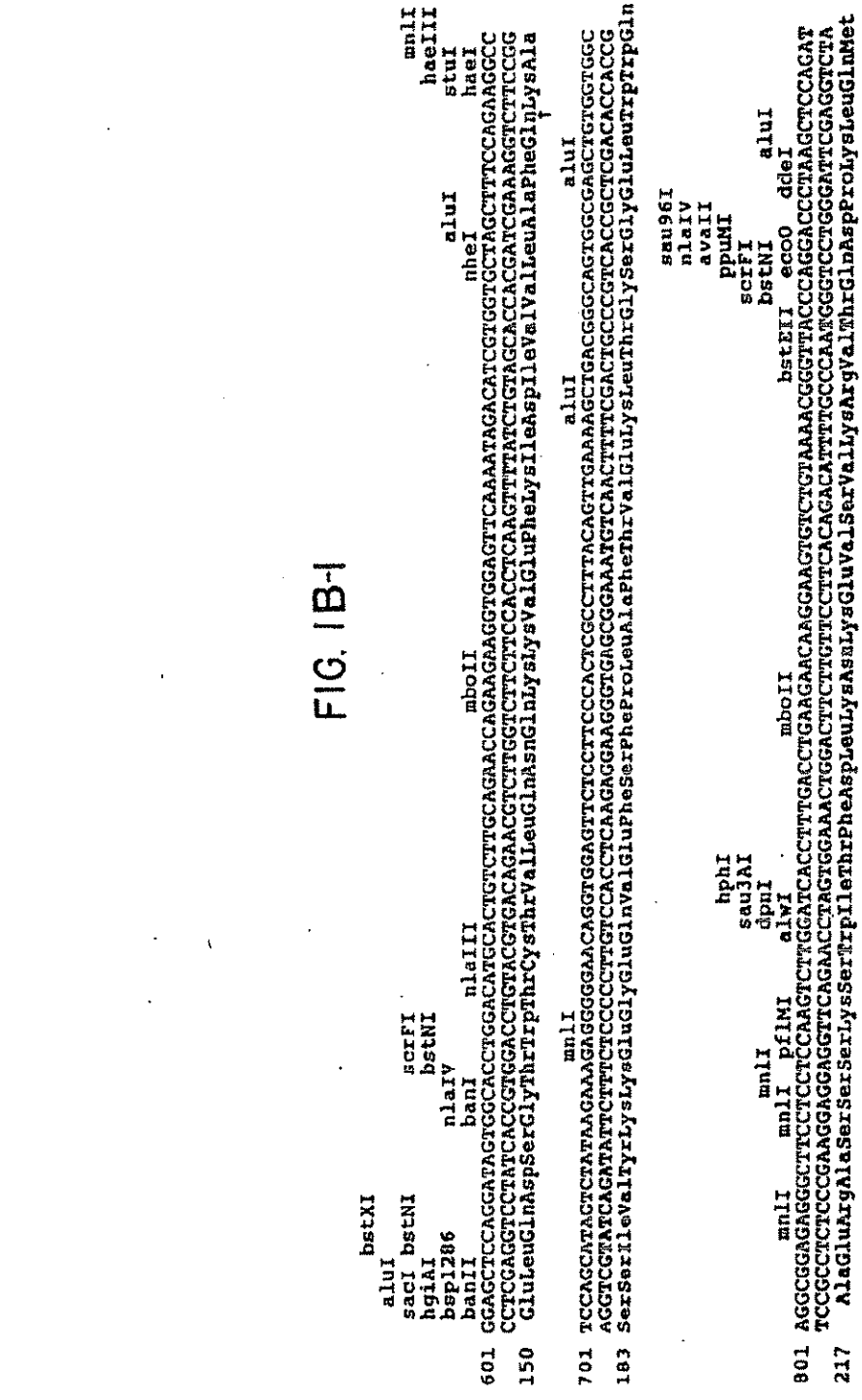
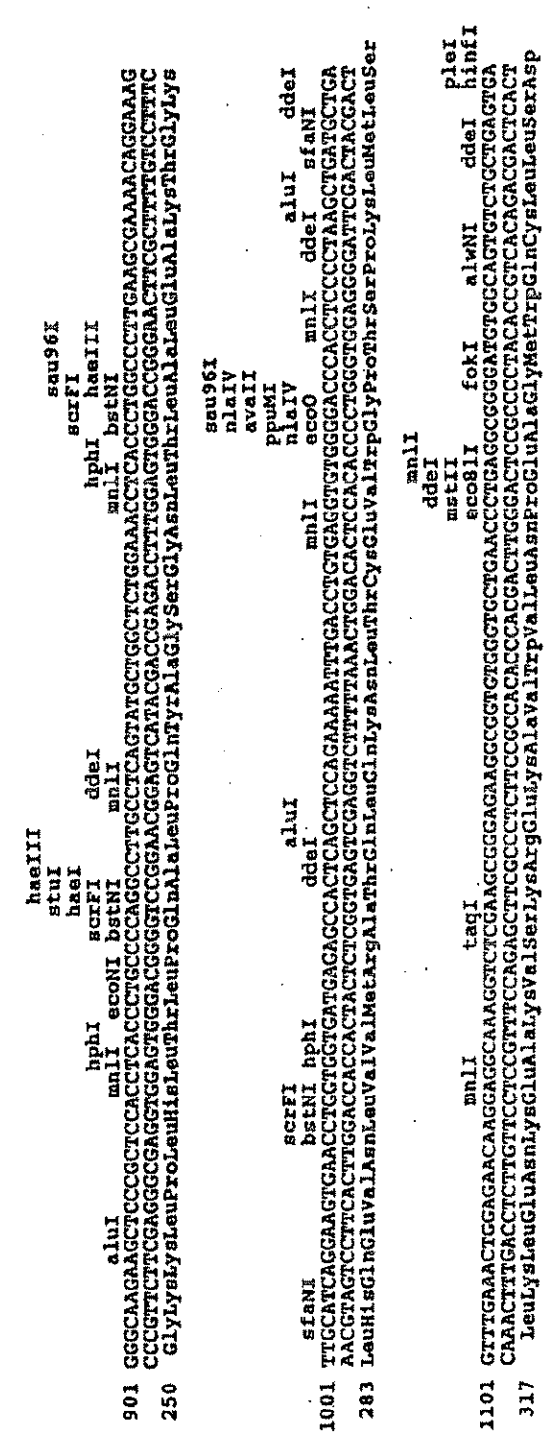

sau96I
nlaIV
ava II
ppuMI
ddeI
bsp1286
banII
bstXI
1 AATTCAAGCCCAGAGCCCTGCCATTTCTGTGGGCTCAGGTCCCTACTGCTCAGCCCCTTCCTCCCTCGGCAAGGCCACAATGAACCGGGGAGTCCCTTTT
TTAAGTTCGGGTCTCGGGACGGTAAAGACACCCGAGTCCAGGGATGACGAGTCGGGGAAGGAGGGAGCCGTTCCGGTGTTACTTGGCCCCTCAGGGAAAA
-25 Met hsnArgGlyValProPhe
hinPI
fnu4HI
hhaI
haeII
mnlI
bbvI
ddeI
fnu4HI
bbvI
101 AGGCACTTGCTTCTGGTGCTGCAACTGGCGCTCCTCCCAGCCACTCAGGCAAACAAAGTGGTGCTGGGCAAAAAGGGGATACAGTGGAACTGACCT
TCCGTGAACGAAGACCACGACGTTGACCGCGAGGAGGGTCGTCGGTGAGTCCCTTTGTTTCACCACGACCCGTTTTTCCCCTATGTCACCTTGACTGGA
-19 ArgHisLeuLeuLeuValLeuGlnLeuAlaLeuLeuProAlaAlaThrGlnGlyAsnLysValValLeuGlyLysLysGlyAspThrValGluLeuThrCys
nlaIV
bsp1286
banII
mseI
fokI
sau96I
avaII
hinfI
ddeI
pleI
hinfI
mboII
dpnI
bclI
hinfI
afIII
mboII
styI
sau96I
avaII
nlaIV
sau3AI
dpnI
bclI
mseI
afIII
hinfI
mboII
ddeI
pleI
hinfI
alwNI
bspMI
ecoNI
mnlI
ddeI
mnlI
mboII
hinfI
plaI
sau3AI
thaI
hhaI
hinPI
styI
sau96I
avaII
nlaIV
sau3AI
dpnI
bclI
mboII
hinfI
plaI
thaI
hhaI
hinPI
alwNI
bspMI
ecoNI
mnlI
ddeI
mnlI
mboII
hinfI
plaI
sau96I
avaII
mnlI
101 GTACAGCTTCCCAGAAGAAGAGCATACAATTCCACTGGAAAAACTCCAACCAGATAAAGATTCTGGGAAATCAGGGCTCCTTCTTAACTAAAGGTCCATC
CATGTCGAAGGGTCTTCTTCTCGTATGTTAAGGTGACCTTTTGAGGTTGGTCTATTTCTAAGACCCTTTAGTCCCGAGGAAGAATTGATTTCCAGGTAG
17 ThrAlaSerGlnLysLysSerIleGlnPheHisTrpLysAsnSerAsnGlnIleLysIleLeuGlyAsnGlnGlySerPheLeuThrLysGlyProSer
styI
sau96I
avaII
nlaIV
sau3AI
dpnI
bclI
mseI
afIII
hinfI
mboII
ddeI
plaI
hinfI
hinPI
hhaI
thaI
sau3AI
dpnI
mboII
hinfI
plaI
ddeI
mnlI
styI
bsp1286
banII
bstNI
mnlI
ddeI
plaI
hinfI
styI
mboII
mnlI
ddeI
aluI
pvuII
301 CAAGCTGAATGGATCGCGCCTGACTCAAGAAGAAGCCCTTTGGCACCAAGGAAACTTTCCCCTGATCATCAAGAATCTTAAGATACTGAGACTCAGATACTAC
GTTCGACTTACTAGCGCGACTGAGTTCTTCTTCGGAAACCCTGGTTCCTTTGAAAGGGGACTAGTAGTTCTTAGAATTCTATCTCTGAGTCTATGAATG
50 LysLeuAsnAspArgAlaAspSerArgArgSerLeuTrpAspGlnGlyAsnPheProLeuIleIleLysAsnLeuLysIleGluAspSerAspThrTyr
sau96I
avaII
mnlI
mnlI
ecoNI
bspMI
alwNI
401 ATCTGTGAAGTGGAGGACCAGAAGGAGGAGGTGCAATTGCTAGTGTTCGGATTGACTGCCAACTCTGACACCCACCTGCTTCAGGGGCAGAGCCTGACCC
TAGACACTTCACCTCCTGGTCTTCCTCCTCCACGTTAACGATCACAAGCCTAACTGACGGTTGAGACTGTGGGTGGACGAAGTCCCCGTCTCGGACTGGG
83 IleCysGluValGluAspGlnLysGluGluValGlnLeuLeuValLeuPheGlyLeuThrAlaAsnSerAspSerThrHisLeuLeuGlnGlyGlnSerLeuThrLeu
styI
sau96I
avaII
nlaIV
styI
aluI
pvuII
501 TGACCTTGGAGAGCCCCCCCTGGTAGTAGCCCCTCAGTGCAATGTAGGAGTCCAAGGGGTAAAAACATACAGGGGGGGAAGACCCTCTCCGTGTCTCAGCT
ACTGGAACCTCTCGGGGGGACCATCATCGGGGAGTCACGTTACATCCTCAGGTTCCCCCATTTTTGTATGTCCCCCCCTTCTGGGAGAGGCCACAGAGTCGA
117 ThrLeuGluSerProProGlySerSerProSerValGlnCysArgSerProArgGlyLysAsnIleGlnGlyGlyLysThrLeuSerValSerGlnLeu

FIG. 1B-1



```
          bstXI
       aluI
      sacI bstNI          scrFI                                                                                  mnlI
      hgiAI               bstNI                                                                                 haeIII
      bsp1286         nlaIV                                                                        aluI        stuI
      banII           banI      nlaIII                  mboII                                   nheI           haeI
601  GGAGCTCCAGGATAGTGGCACCTGGACATGCACTGTCTTGCAGAACCAGAAGAAGGTGGAGTTCAAAATAGACATCGTGGTGCTAGCTTTCCAGAAGGCC
     CCTCGAGGTCCTATCACCGTGGACCTGTACGTGACAGAACGTCTTGGTCTTCTTCCACCTCAAGTTTTATCTGTAGCACCACGATCGAAAGGTCTTCCGG
150   GluLeuGlnAspSerGlyThrTrpThrCysThrValLeuGlnAsnGlnLysLysValGluPheLysIleAspIleValValLeuAlaPheGlnLysAla
                                                                                                          ↑
                          mnlI                                                  aluI                 aluI
701  TCCAGCATAGTCTATAAGAAAGAGGGGGAACAGGTGGAGTTCTCCTTCCCACTCGCCTTTACAGTTGAAAAGCTGACGGGCAGTGGCGAGCTGTGGTGGC
     AGGTCGTATCAGATATTCTTTCTCCCCCTTGTCCACCTCAAGAGGAAGGGTGAGCGGAAATGTCAACTTTTCGACTGCCCGTCACCGCTCGACACCACCG
183  SerSerIleValTyrLysLysGluGlyGluGlnValGluPheSerPheProLeuAlaPheThrValGluLysLeuThrGlySerGlyGluLeuTrpTrpGln

                                                                                                 sau96I
                                                                                                 nlaIV
                                                                                                 avaII
                                        hphI                                                     ppuMI
                                      sau3AI                                                   scrFI
                        mnlI          dpnI                                                     bstNI      aluI
            mnlI    mnlI  pflMI    alwI            mboII                                 bstEII  ecoO  ddeI
801  AGGCGGAGAGGGCTTCCTCCTCCAAGTCTTGGATCACCTTTGACCTGAAGAACAAGGAAGTGTCTGTAAAACGGGTTACCCAGGACCCTAAGCTCCAGAT
     TCCGCCTCTCCCGAAGGAGGAGGTTCAGAACCTAGTGGAAACTGGACTTCTTGTTCCTTCACAGACATTTTGCCCAATGGGTCCTGGGATTCGAGGTCTA
217   AlaGluArgAlaSerSerSerLysSerTrpIleThrPheAspLeuLysAsnLysGluValSerValLysArgValThrGlnAspProLysLeuGlnMet
```



```
                                                haeIII
                                               stuI                                        sau96I
                                               haeI                                      scrFI
                                  hphI      scrFI      ddeI                          hphI   haeIII
                 aluI           mnlI  ecoNI bstNI     mnlI                         mnlI  bstNI
 901 GGGCAAGAAGCTCCCGCTCCACCTCACCCTGCCCCAGGCCTTGCCTCAGTATGCTGGCTCTGGAAACCTCACCCTGGCCCTTGAAGCGAAAACAGGAAAG
     CCCGTTCTTCGAGGGCGAGGTGGAGTGGGACGGGGTCCGGAACGGAGTCATACGACCGAGACCTTTGGAGTGGGACCGGGAACTTCGCTTTTGTCCTTTC
 250  GlyLysLysLeuProLeuHisLeuThrLeuProGlnAlaLeuProGlnTyrAlaGlySerGlyAsnLeuThrLeuAlaLeuGluAlaLysThrGlyLys

                                                                                      sau96I
                                                                                      nlaIV
                                                                                      avaII
                                                                                     ppuMI
                          scrFI                  aluI                                nlaIV            aluI    ddeI
       sfaNI              bstNI hphI           ddeI                        mnlI     ecoO    mnlI  ddeI   sfaNI
1001 TTGCATCAGGAAGTGAACCTGGTGGTGATGAGAGCCACTCAGCTCCAGAAAAATTTGACCTGTGAGGTGTGGGGACCCACCTCCCCTAAGCTGATGCTGA
     AACGTAGTCCTTCACTTGGACCACCACTACTCTCGGTGAGTCGAGGTCTTTTTAAACTGGACACTCCACACCCCTGGGTGGAGGGGATTCGACTACGACT
 283 LeuHisGlnGluValAsnLeuValValMetArgAlaThrGlnLeuGlnLysAsnLeuThrCysGluValTrpGlyProThrSerProLysLeuMetLeuSer

                                                                                 mnlI
                                                                                ddeI
                                                                               mstII                                pleI
                          mnlI        taqI                                     eco81I    fokI    alwNI    ddeI  hinfI
1101 GTTTGAAACTGGAGAACAAGGAGGCAAAGGTCTCGAAGCGGGAGAAGGCGGTGTGGGTGCTGAACCCTGAGGCGGGGATGTGGCAGTGTCTGCTGAGTGA
     CAAACTTTGACCTCTTGTTCCTCCGTTTCCAGAGCTTCGCCCTCTTCCGCCACACCCACGACTTGGGACTCCGCCCCTACACCGTCACAGACGACTCACT
 317   LeuLysLeuGluAsnLysGluAlaLysValSerLysArgGluLysAlaValTrpValLeuAsnProGluAlaGlyMetTrpGlnCysLeuLeuSerAsp
```

FIG. 1B-2



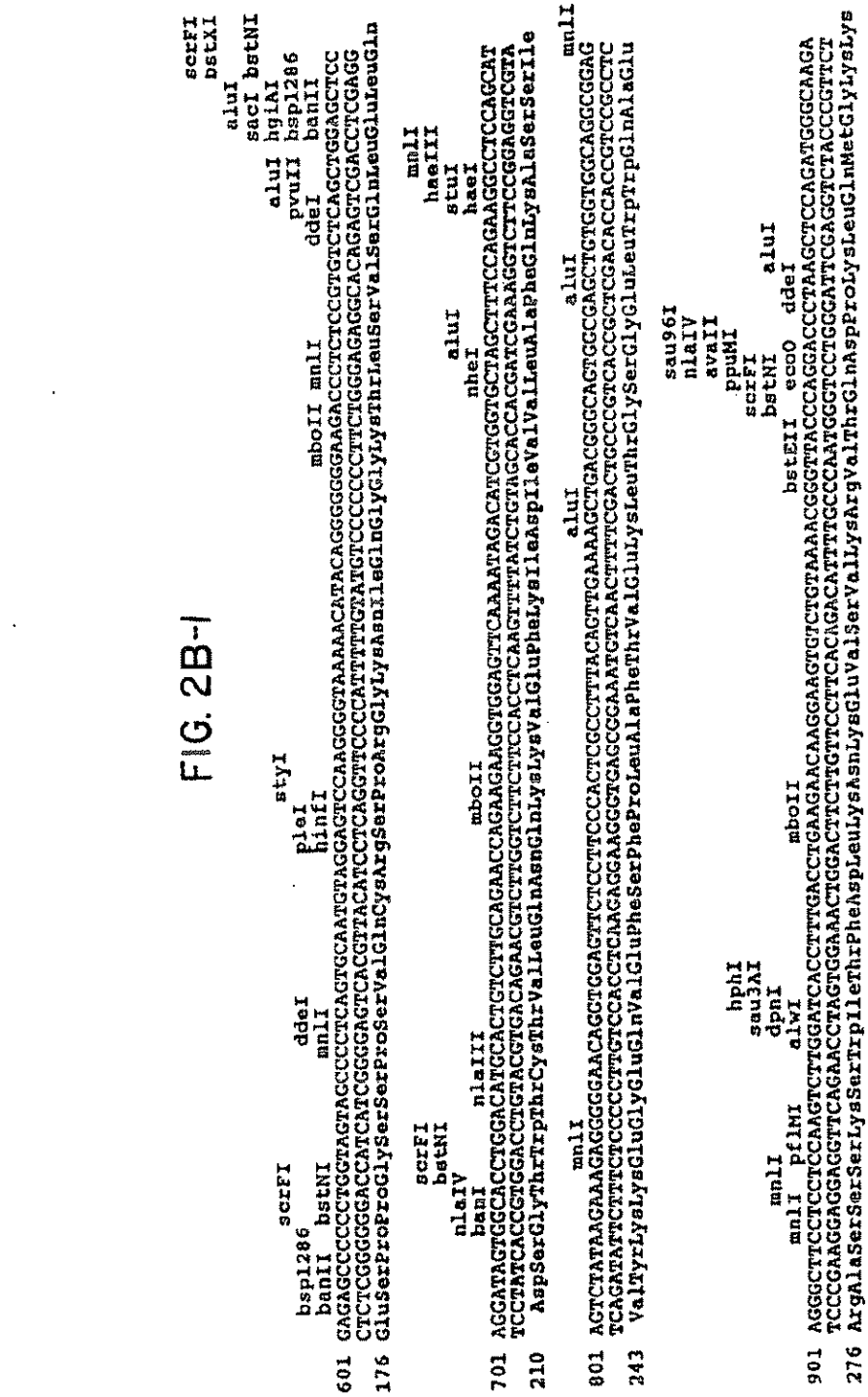
```
                sau96I
                avaII
               ppuMI                                  sau96I
               ecoO                                   avaII      aluI
      avaI   alwNI       hinfI                     nlaIII    avaI    mseI                    mseI
1201  CTCGGGACAGGTCCTGCTGGAATCCAACATCAAGGTTCTGCCCACATGGTCCACCCCGAGCTTTAATGCGGTAGTTTATCACAGTTAAATTGCTAACGCA
      GAGCCCTGTCCAGGACGACCTTAGGTTGTAGTTCCAAGACGGGTGTACCAGGTGGGGCTCGAAATTACGCCATCAAATAGTGTCAATTTAACGATTGCGT
 350   SerGlyGlnValLeuLeuGluSerAsnIleLysValLeuProThrTrpSerThrProSerPheAsnAlaValValTyrHisSerOC*

                                                                      sfaNI                                  scrFI
                                                      nlaIV      scrFI                                rsaI  nciI
          nlaIV                    hinPI         mnlI           bstNI                             mspI     mspI
          banI                     hhaI       fokI   banI   hphI   fokI                            hpaII    hpaII
1301  GTCAGGCACCGTGTATGAAATCTAACAATGCGCTCATCGTCATCCTCGGCACCGTCACCCTGGATGCTGTAGGCATAGGCTTGGTTATGCCGGTACTGCC
      CAGTCCGTGGCACATACTTTAGATTGTTACGCGAGTAGCAGTAGGAGCCGTGGCAGTGGGACCTACGACATCCGTATCCGAACCAATACGGCCATGACGG

         mnlI
       haeIII
      sau96I
1401  GGGCCTCTTGCGGGAT
      CCCGGAGAACGCCCTA
```

FIG. 1C

FIG. 2A

```
          thaI                                                                                  scrFI
   aluI   hinPI                sau3AI          mseI                          mspI                bstNI
  hindIII hhaI                 dpnI           aflII                         hpaII            fnu4HI      nlaIV
  1 AAGCTTCAGCGCGAACGACCAACTACCCCGATCATCAGTTATCCTTAAGGTCTCTTTTGTGTGGTGCGTTCCGGTATGGGGGGGACTGCCGCCAGGTTGG
    TTCGAAGTCGCGCTTGCTGGTTGATGGGGCTAGTAGTCAATAGGAATTCCAGAGAAAACACACCACGCAAGGCCATACCCCCCCTGACGGCGGTCCAACC
  1                                                                                  MetGlyGlyThrAlaAlaArgLeuGly
                                      styI      thaI
                                      ncoI  sau96I
                                    mnlI     avaII
    haeIII                         haeIII    nlaIV fnu4HI                 sfaNI            haeIII      taqI
   sau96I                         sau96I nlaIII  sacII          styI   fokI mnlI         eaeI        claI
101 GGGCCGTGATTTTGTTTGTCGTCATAGTGGGCCTCCATGGGGTCCGCGGCAAATATGCCTTGGCGGATGCCTCTCTCAAGATGGCCGACCCCAATCGATT
    CCCGGCACTAAAACAAACAGCAGTATCACCCGGAGGTACCCCAGGCGCCGTTTATACGGAACCGCCTACGGAGAGAGTTCTACCGGCTGGGGTTAGCTAA
 10   AlaValIleLeuPheValValIleValGlyLeuHisGlyValArgGlyLysTyrAlaLeuAlaAspAlaSerLeuLysMetAlaAspProAsnArgPhe

                        scrFI    fnu4HI
                      sau96I     aluI
                      avaII sau96I    xhoI
     fnu4HI         mspI bstNI  pvuII avaI                                                          aluI
    thaI            hpaII   avaII bbvI taqI                                                       rsaI
201 TCGCGGCAAAGACCTTCCGGTCCTGGACCAGCTGCTCGAGCAGGGAAACAAAGTGGTGCTGGGCAAAAAAGGGGATACAGTGGAACTGACCTGTACAGCT
    AGCGCCGTTTCTGGAAGGCCAGGACCTGGTCGACGAGCTCGTCCCTTTGTTTCACCACGACCCGTTTTTTCCCCTATGTCACCTTGACTGGACATGTCGA
 43  ArgGlyLysAspLeuProValLeuAspGlnLeuLeuGluGlnGlyAsnLysValValLeuGlyLysLysGlyAspThrValGluLeuThrCysThrAla

                                                                                  nlaIV                fokI
          mboII                                                                  bsp1286            sau96I
        mboII                                                  hinfI             banII     mseI     avaII     aluI
301 TCCCAGAAGAAGAGCATACAATTCCACTGGAAAAACTCCAACCAGATAAAGATTCTGGGAAATCAGGGCTCCTTCTTAACTAAAGGTCCATCCAAGCTGA
    AGGGTCTTCTTCTCGTATGTTAAGGTGACCTTTTTGAGGTTGGTCTATTTCTAAGACCCTTTAGTCCCGAGGAAGAATTGATTTCCAGGTAGGTTCGACT
 76 SerGlnLysLysSerIleGlnPheHisTrpLysAsnSerAsnGlnIleLysIleLeuGlyAsnGlnGlySerPheLeuThrLysGlyProSerLysLeuAsn

         hinPI
         hhaI                          styI                                                ddeI
        thaI                         sau96I                 sau3AI        mseI        pleI
      sau3AI   pleI                  avaII                  dpnI         aflII        hinfI
      dpnI     hinfI  mboII         nlaIV                  bclI      hinfI        mboII
401 ATGATCGCGCTGACTCAAGAAGAAGCCTTTGGGACCAAGGAAACTTTCCCCTGATCATCAAGAATCTTAAGATAGAAGACTCAGATACTTACATCTGTGA
    TACTAGCGCGACTGAGTTCTTCTTCGGAAACCCTGGTTCCTTTGAAAGGGGACTAGTAGTTCTTAGAATTCTATCTTCTGAGTCTATGAATGTAGACACT
110   AspArgAlaAspSerArgArgSerLeuTrpAspGlnGlyAsnPheProLeuIleIleLysAsnLeuLysIleGluAspSerAspThrTyrIleCysGlu

         sau96I
         avaII        mnlI                                                  ecoNI
        mnlI        mnlI                                                  bspMI        alwNI          styI
501 AGTGGAGGACCAGAAGGAGGAGGTGCAATTGCTAGTGTTCGGATTGACTGCCAACTCTGACACCCACCTGCTTCAGGGGCAGAGCCTGACCCTGACCTTG
    TCACCTCCTGGTCTTCCTCCTCCACGTTAACGATCACAAGCCTAACTGACGGTTGAGACTGTGGGTGGACGAAGTCCCCGTCTCGGACTGGGACTGGAAC
143  ValGluAspGlnLysGluGluValGlnLeuLeuValPheGlyLeuThrAlaAsnSerAspThrHisLeuLeuGlnGlyGlnSerLeuThrLeuThrLeu
```

FIG. 2B-I


scrFI
bstXI
aluI
sacI bstNI
aluI hgiAI
pvuII bsp1286
ddeI banII
mboII mnlI
hinfI
pleI
styI
ddeI
mnlI
bsp1286
banII bstNI
601 GAGAGCCCCCCCTGGTAGTACCCCCTCAGTGCAATGTAGGAGTCCAAGGGGTAAAAACATACAGGGGGGGAAGACCCTCTCCGTGTCTCAGCTGGAGCTCC
CTCTCGGGGGGACCATCATCGGGGAGTCACGTTACATCCTCAGGTTCCCCATTTTTGTATGTCCCCCCCTTCTGGGAGAGGCACAGAGTCGACCTCGAGG
176 GluSerProProGlySerSerProSerValGlnCysArgSerProArgGlyLysAsnIleGlnGlyGlyLysThrLeuSerValSerGlnLeuGluLeuGln
mnlI
haeIII
stuI
haeI
aluI
nheI
scrFI
bstNI
nlaIV
banI
nlaIII
mboII
aluI
mnlI
701 AGGATAGTGGCACCTGGACATGCACTGTCTTGCAGAACCAGAAGAGGTGGAGTTCAAAATAGACATCGTGGTGCTAGCTTTCCAGAAGGCCTCCAGCAT
TCCTATCACCGTGGACCTGTACGTGACAGAACGTCTTGGTCTTCTTCCACCTCAAGTTTTATCTGTAGCACCACGATCGAAAGGTCTTCCGGAGGTCGTA
210 AspSerGlyThrTrpThrCysThrValLeuGlnAsnGlnLysLysValGluPheLysIleAspIleValValLeuAlaPheGlnLysAlaSerSerIle
aluI
mnlI
801 AGTCTATAAGAAAGAGGGGGAACAGGTGGAGTTCTCCTTCCCACTCGCCTTTACAGTTGAAAAGCTGACGGGCAGTGGCGAGCCTGTGGTGGCAGGCGGAG
TCAGATATTCTTTCTCCCCCTTGTCCACCTCAAGAGGAAGGGTGAGCGGAAATGTCAACTTTTCGACTGCCCGTCACCGCGCTCGACACCACCGTCCGCCTC
243 ValTyrLysLysGluGlyGluGlnValGluPheSerPheProLeuAlaPheThrValGluLysLeuThrGlySerGlyGluLeuTrpTrpGlnAlaGlu
sau96I
nlaIV
avaII
ppuMI
scrFI
bstNI
aluI
bstEII ecoO ddeI
mboII
bphI
sau3AI
dpnI
mnlI pflMI alwI
mnlI
901 AGGGCTTCCTCCTCCAAGTCTTGGATCACCTTTGACCTGAAGAACAAGGAAGTGTCTGTAAAACGGGTACCCAGGACCCTAAGCTCCAGATGGGCAAGA
TCCCGAAGGAGGAGGTTCAGAACCTAGTGGAAACTGGACTTCTTGTTCCTTCACAGACATTTTGCCCAATGGGTCCTGGGATTCGAGGTCTACCCGTTCT
276 ArgAlaSerSerSerLysSerTrpIleThrPheAspLeuLysAsnLysGluValSerValLysArgValThrGlnAspProLysLeuGlnMetGlyLysLys

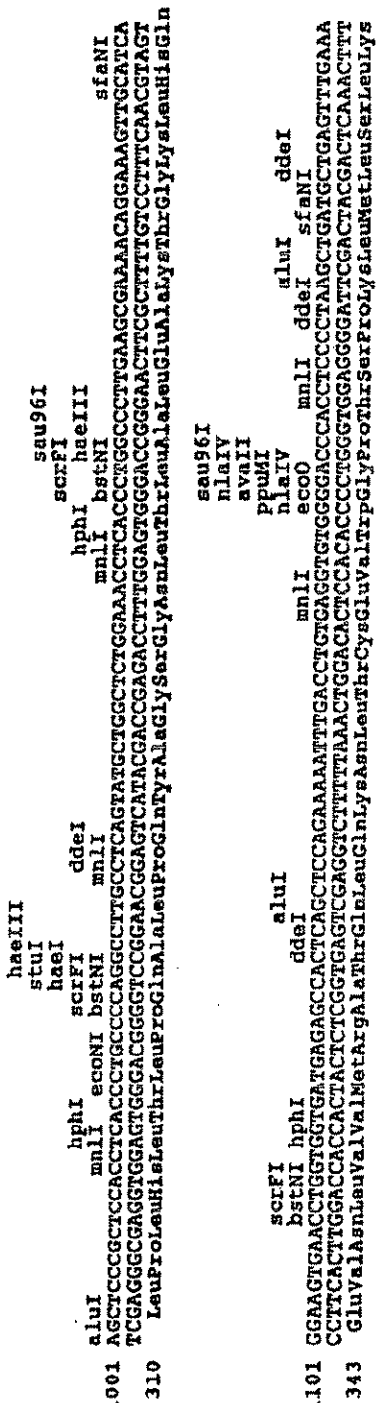
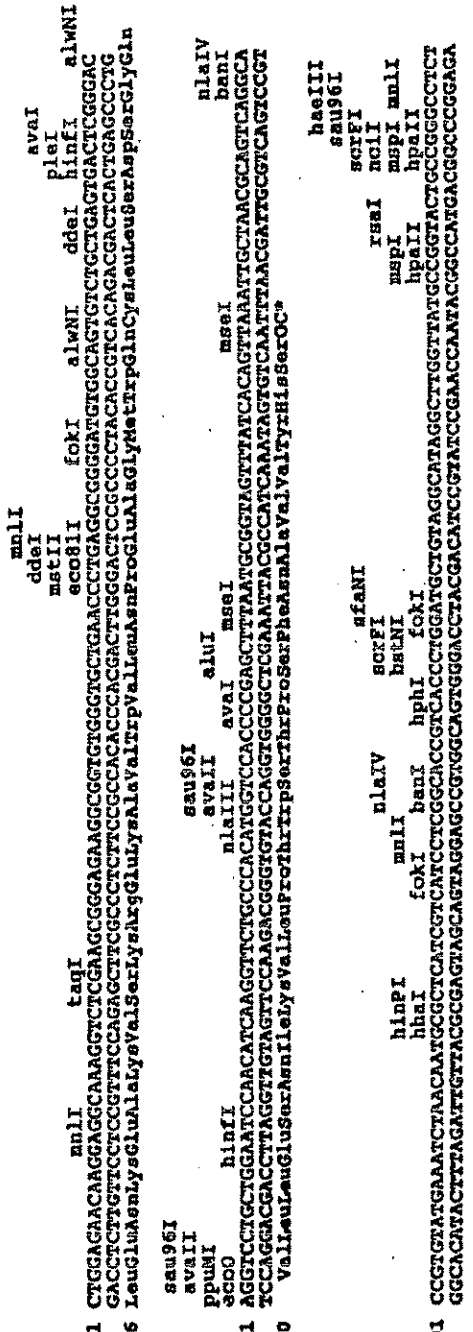

haeIII
stuI
haeI
scrFI
ddeI
hphI
mnlI
aluI
ecoNI
bstNI
sau96I
hphI
haeIII
mnlI
bstNI
sfaNI
sau96I
nlaIV
avaII
PpuMI
nlaIV
aluI
ddeI
mnlI
eco0
mnlI
ddeI
sfaNI
scrFI
bstNI
hphI
aluI
ddeI
mnlI
1001 AGCTCCCCGCTCCCACCTCACCCTGCCCCAGGCCTTGCCTCAGTATGCTGGCTCTGGGAAACCTCACCCTGGCCCTTGAAGCGAAAACAGGAAAGTTGCATCA
310 TCGAGGGCGAGGTGGGAGTGGGACGGGGTCCGGAACGGAGTCATACGACCGAGACCTTTGGAGTGGGACCGGGAACTTCGCTTTTGTCCTTTCAACGTAGT
LeuProLeuHisLeuThrLeuProGlnGlnTyrAlaGlySerGlyAsnLeuThrLeuAlaLeuGluAlaLeuGluAlaLysThrGlyLysLeuHisGln
1101 GGAAGTGAACCTGGTGGTGATGAGAGCCACTCAGCTCCAGAAAAATTTGACCTGTGAGGTGTGGGGACCCACCTCCCCTAAGCTGATGCTGAGTTTGAAA
343 CCTTCACTTGGACCACCACTACTCTCGGTGAGTCGTTTTTTAAACTGGACACTCCACACCCCTGGGTGGAGGGGATTCGACTACGACTCAAACTTT
GluValAsnLeuValValMetArgAlaThrGlnLeuGlnLysAsnLeuThrCysGluValTrpGlyProThrSerProLysLeuMetLeuSerLeuLys

FIG. 2B-2



```
                                                                                  mnlI
                                                                                ddeI                                    avaI
                                                                               mstII                                   pleI
                 mnlI        taqI                                          eco81I    fokI    alwNI    ddeI  hinfI    alwNI
1201 CTGGAGAACAAGGAGGCAAAGGTCTCGAAGCGGGAGAAGGCGGTGTGGGTGCTGAACCCTGAGGCGGGGATGTGGCAGTGTCTGCTGAGTGACTCGGGAC
     GACCTCTTGTTCCTCCGTTTCCAGAGCTTCGCCCTCTTCCGCCACACCCACGACTTGGGACTCCGCCCCTACACCGTCACAGACGACTCACTGAGCCCTG
 376 LeuGluAsnLysGluAlaLysValSerLysArgGluLysAlaValTrpValLeuAsnProGluAlaGlyMetTrpGlnCysLeuLeuSerAspSerGlyGln

      sau96I
      avaII                                  sau96I
     ppuMI                                   avaII      aluI                                                  nlaIV
     eco0       hinfI                      nlaIII    avaI    mseI                    mseI                      banI
1301 AGGTCCTGCTGGAATCCAACATCAAGGTTCTGCCCACATGGTCCACCCCGAGCTTTAATGCGGTAGTTTATCACAGTTAAATTGCTAACGCAGTCAGGCA
     TCCAGGACGACCTTAGGTTGTAGTTCCAAGACGGGTGTACCAGGTGGGGCTCGAAATTACGCCATCAAATAGTGTCAATTTAACGATTGCGTCAGTCCGT
 410   ValLeuLeuGluSerAsnIleLysValLeuProThrTrpSerThrProSerPheAsnAlaValValTyrHisSerOC*

                                                                                                                haeIII
                                                                                                               sau96I
                                                                                                              scrFI
                                                                  sfaNI                                   rsaI  ncil
                                                  nlaIV      scrFI                                       mspI     mspI mnlI
                          hinPI         mnlI                bstNI                                      hpaII    hpaII
                          hhaI       fokI   banI   hphI   fokI
1401 CCGTGTATGAAATCTAACAATGCGCTCATCGTCATCCTCGGCACCGTCACCCTGGATGCTGTAGGCATAGGCTTGGTTATGCCGGTACTGCCGGGCCTCT
     GGCACATACTTTAGATTGTTACGCGAGTAGCAGTAGGAGCCGTGGCAGTGGGACCTACGACATCCGTATCCGAACCAATACGGCCATGACGGCCCGGAGA

1501 TGCGGGAT
     ACGCCCTA
```

FIG. 2C

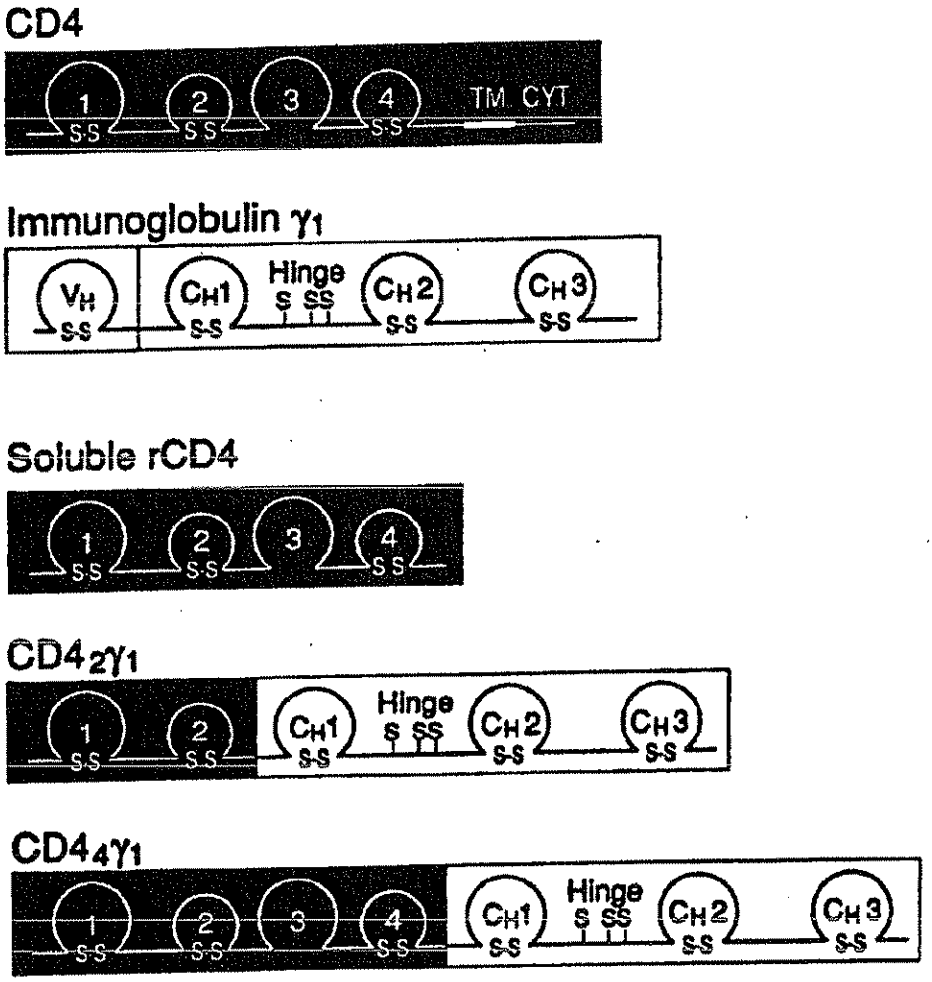

CD4
1
2
3
4
TM
CYT
S-S
Immunoglobulin γ1
VH
CH1
Hinge
S SS
CH2
CH3
Soluble rCD4
CD4 2γ1
CD4 4γ1

FIG. 3

FIG. 4A

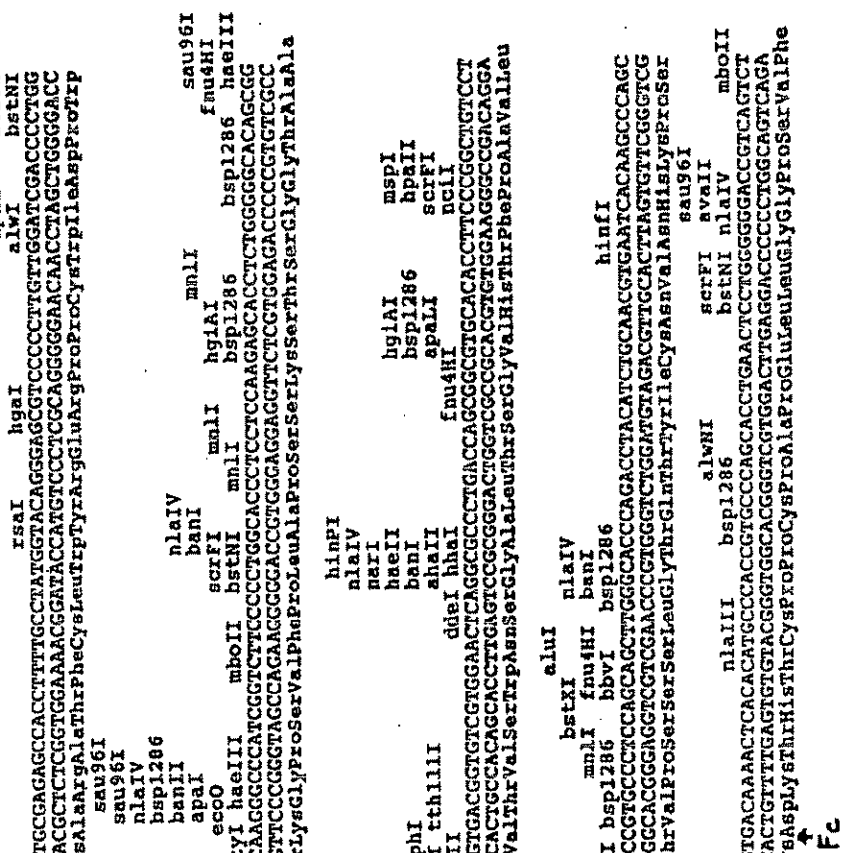
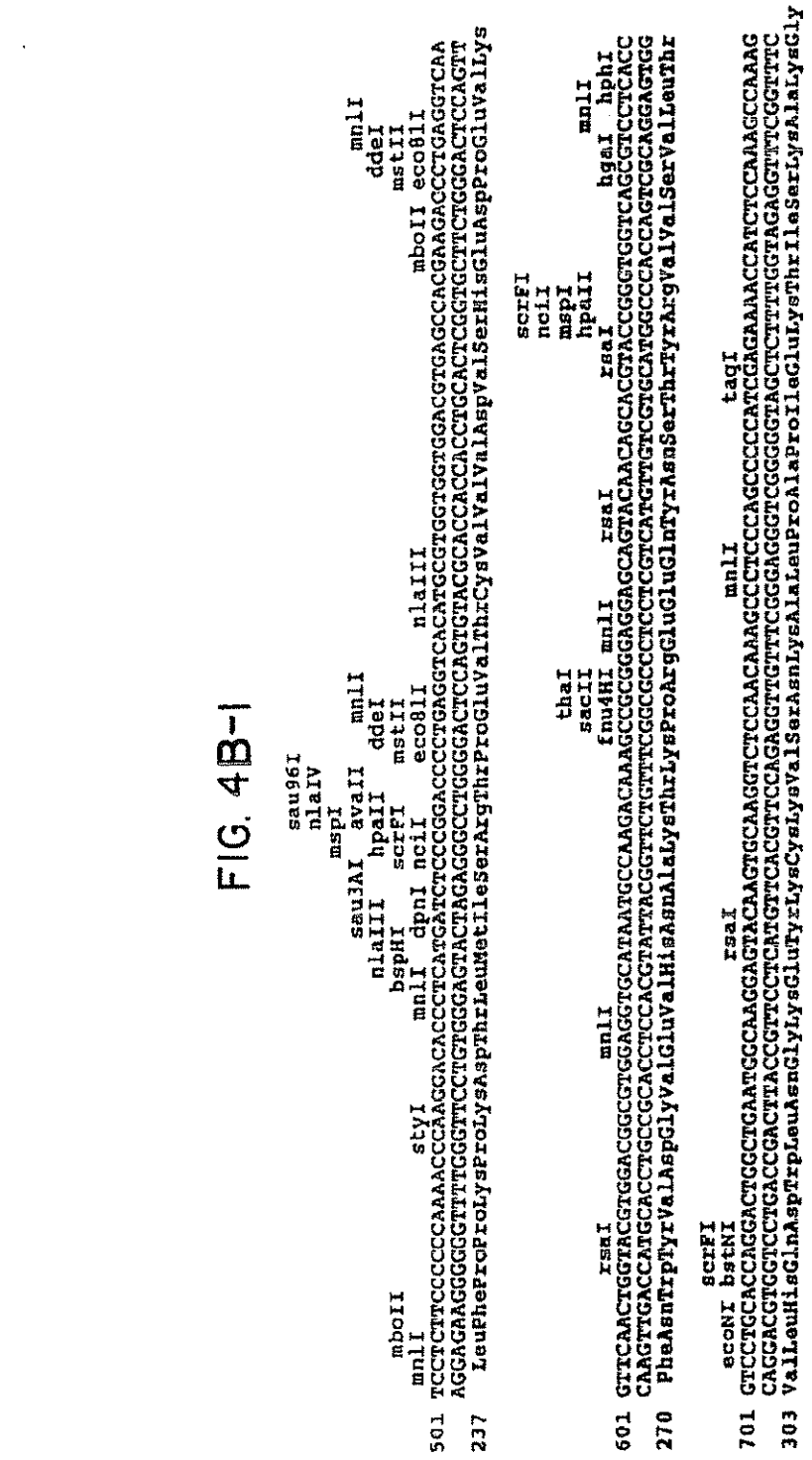

thaI haeIII sau96I
sacII xmaIII taqI nlaIV
fnu4HI eaeI rsaI hgaI sau3AI ecoO
ecoRI dpnI scrFI
alwI bstNI
1 GAATTCTGTCACTGCCGCCGGACACGGCCGTATATTACTGTGCCGAGAGCCACCTTTGCCTATGGTACAGGGAGCGTCCCCCTTGTTGGATCGACCCCTGG
CTTAAGACAGTGACGGCGGCCGGCATATAATGACACGCTCTCGGTGGAAAACGGATACCATGTCCCTCGCAGGGGGAACAACCTAGCTGGGGACC
72 ValThrAlaAlaAspThrAlaValTyrTyrCysAlaArgAlaThrPheCysLeuTrpTyrArgGluArgProProCysTrpIleAspProTrp
sau96I
sau96I
nlaIV
bsp1286
nlaIV banII hgiAI sau96I
hphI apaI mnlI fnu4HI
nlaIV bstEII scrFI banI bsp1286 haeIII
scrFI haeIII ecoO mnlI hgiAI bsp1286
bstNI mnlI styI haeIII mboII bstNI mnlI bsp1286 haeIII
101 GGCCTGGGAACCCTGGTCACCGTCTCCTCAGGCCCTCCACCAAGGGCCCATCGGTCTTCCCCCTGGCACCCTCCTCCAAGAGCACCTCTGGGGGCACAGCGG
CCGGACCCTTGGGACCAGTGGCAGAGGAGCCGGAGGTGGTTCCCGGGTAGCCACAGAAGGGGGACCGTGGGAGGAGGTTCTCGTGGAGACCCCCGTGTCGCC
103 GlyLeuGlyThrLeuValThrValSerSerAlaSerThrLysGlyProSerValPheProLeuAlaProSerSerLysSerThrSerGlyGlyThrAlaAla
hinPI
nlaIV
Y1 narI
scrFI haeII hglAI mspI
fnu4HI banI bsp1286 hpaII
bbvI hphI ahaII apaLI scrFI
msp1 tth111I fnu4HI nciI
scrFI ecoNI hpaII ddeI hhaI
bstNI bstNI
201 CCCTGGGCTGCCTGGTCAAGGACTACTTCCCCGAACCGGTGACGGTGTCGTGGAACTCAGGCGCCCTGACCAGCGGCGTGCACACCTTCCCGGCTGTCCT
GGGACCCGACGGACCAGTTCCTGATGAAGGGGCTTGGCCACTGCCACAGCACCTTGAGTCCGCCGGGACTGGTCGCCGCACAGTGTGGAAGGCCGACAGGA
137 LeuGlyCysLeuValLysAspTyrPheProGluProValThrValSerTrpAsnSerGlyAlaLeuThrSerGlyValHisThrPheProAlaValLeu
aluI
ddeI bstXI nlaIV
mstII pleI fnu4HI mnlI fnu4HI banI
mnlI hinfI hinfI
eco81I mnlI bbvI hphI bstEII bsp1286 bbvI bsp1286
301 ACAGTCCTCAGGACTCTACTCCCTCAGCAGCGTGGTGACCGTGCCCTCCAGCAGCTTGGGCACCCAGACCTACATCTGCAACGTGAATCACAAGCCCAGC
TGTCAGGAGTCCTGAGATGAGGGAGTCGTCGCACCACTGGCACGGGAGGTCGTCGAACCCGTGGGTCTGGATGTAGACGTTGCACTTAGTGTTCGGGTCG
170 GlnSerSerGlyLeuTyrSerLeuSerSerValValThrValProSerSerSerLeuGlyThrGlnThrTyrIleCysAsnValAsnHisLysProSer
sau96I
alwNI scrFI avaII
bsp1286 bstNI nlaIV mboII
bsp1286 banII nlaIII bsp1286
styI
401 AACACCAAGGTGGACAAGAAAGTTGAGCCCAAATCTTGTGACAAAACTCACACATGCCCACCGTGCCCAGCACCTGAACTCCTGGGGGGACCGTCAGTCT
TTGTGGTTCCACCTGTTCTTTCAACTCGGGTTTAGAACACTGTTTTGAGTGTACGGGTGGCACGGGTCGTGGACTTGAGGACCCCCCTGGCAGTCAGA
203 AsnThrLysValAspLysLysValGluProLysSerCysAspLysThrHisThrCysProProCysProAlaProGluLeuLeuGlyGlyProSerValPhe
Fc

FIG. 4B-1


sau96I
nlaIV
mspI
sau3AI avaII mnlI
nlaIII hpaII ddeI
bspHI scrFI mstII
mboII
mnlI styI mnlI dpnI ncil eco81I nlaIII mboII eco81I
501 TCCTCTTCCCCCCAAAACCCAAGGACACCCTCATGATCTCCCGGACCCCTGAGGTCACATGCGTGGTGGACGTGAGCCACGAAGACCCTGAGGTCAA
AGGAGAAGGGGGGTTTTGGGTTCCTGTGGGAGTACTAGAGGGCCTGGGGACTCCAGTGTACGCACCACCTGCACTCGGTGCTTCTGGGACTCCAGTT
237 LeuPheProProLysProLysAspThrLeuMetIleSerArgThrProGluValThrCysValValAspValSerHisGluAspProGluValLys
scrFI
ncil
mspI
hpaII
thaI
sacII
fnu4HI mnlI rsaI rsaI hgaI hphI
601 GTTCAACTGGTACGTGGACGGCGTGGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAGCAGTACAACAGCACGTACCGGGTGGTCAGCGTCCTCACC
CAAGTTGACCATGCACCTGCCGCACCTCCACGTATTACGGTTCTGTTTCGGCGCCCTCCTCGTCATGTTGTCGTGCATGGCCCACCAGTCGCAGGAGTGG
270 PheAsnTrpTyrValAspGlyValGluValHisAsnAlaLysThrLysProArgGluGluGlnTyrAsnSerThrTyrArgValValSerValLeuThr
rsaI mnlI mnlI taqI
scrFI
ecoNI bstNI rsaI
701 GTCCTGCACCAGGACTGGCTGAATGGCAAGGAGTACAAGTGCAAGGTCTCCAACAAAGCCCTCCCAGCCCCCATCGAGAAAACCATCTCCAAAGCCAAAG
CAGGACGTGGTCCTGACCGACTTACCGTTCCTCATGTTCACGTTCCAGAGGTTGTTTCGGGAGGGGTCGGGGGTAGCTCTTTGGTAGAGGTTTCGGTTTC
303 ValLeuHisGlnAspTrpLeuAsnGlyLysGluTyrLysCysLysValSerAsnLysAlaLeuProAlaProIleGluLysThrIleSerLysAlaLysGly

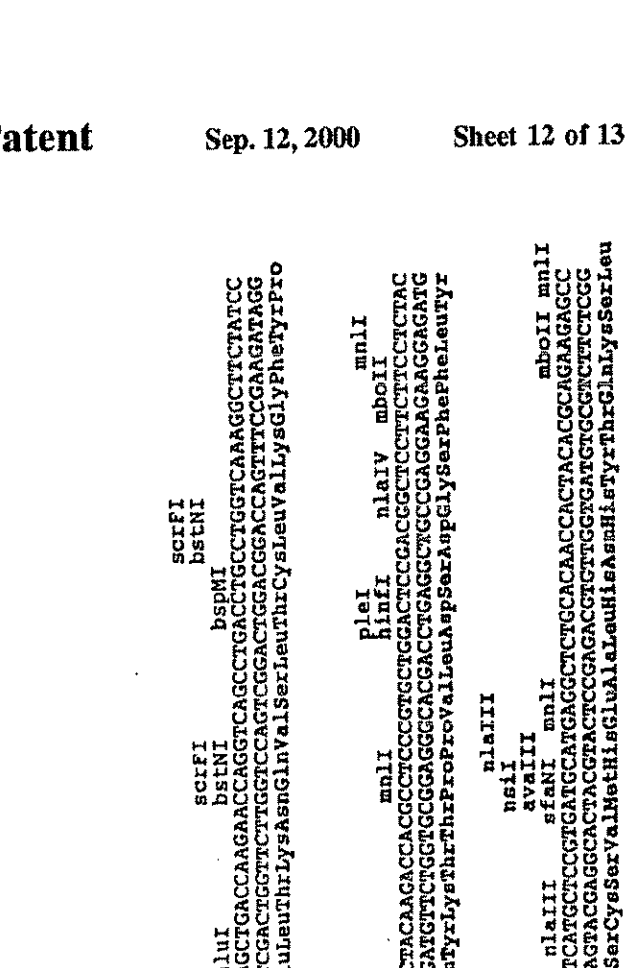
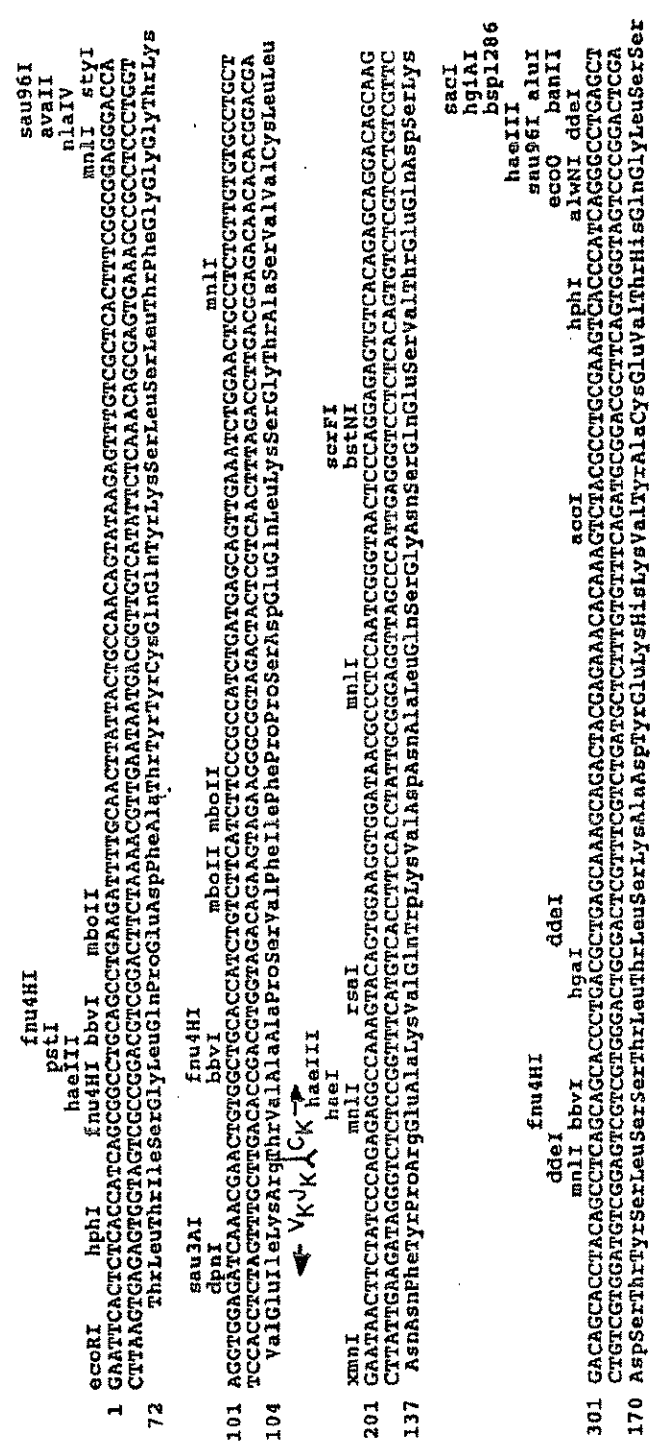

scrFI
ncil
mspI
hpaII
smaI
scrFI
ncil
aval
scrFI
bstNI
bspMI
fnu4HI
bbvI aval
rsaI
fokI fokI aluI
bstNI
801 GGCAGCCCCGAGAACCACAGGTGTACACCCTGCCCCCATCCCGGGATGAGCTGACCAAGAACCAGGTCAGCCTGACCTGCCTGGTCAAAGGCTTCTATCC
CCGTCGGGGCTCTTGGTGTCCACATGTGGGACGGGGGTAGGGCCCTACTCGACTGGTTCTTGGTCCAGTCGGACTGGACGGAACCAGTTTCCGAAGATAGG
337 GlnProArgGluProGlnValTyrThrLeuProProSerArgAspGluLeuThrLysAsnGlnValSerLeuThrCysLeuValLysGlyPheTyrPro
mspI
hpaII
fnu4HI
bbvI
mnlI
pleI
hinfI
nlaIV
mboII
mnlI
901 CAGCGACATCGCCGTGGAGTGGCAGAGCAATGGGCAGCCGGAGAACAACTACAAGACCACGCCTCCCGTGCTGGACTCCGACGGCTCCTTCTTCCTCTAC
GTCGCTGTAGCGGCACCTCACCCTCTCGTTACCCGTCGGCCTCTTGTTGATGTTCTGGTGCGGAGGGCACGACCTGAGGCTGCCGAGGAAGAAGGAGATG
370 SerAspIleAlaValGluTrpGluSerAsnGlyGlnProGluAsnAsnTyrLysThrThrProProValLeuAspSerAspGlySerPhePheLeuTyr
nlaIII
nsiI
avaIII
fnu4HI
bbvI
xmnI mboII nlaIII
sfaNI mnlI
mboII mnlI
1001 AGCAAGCTCACCCTGGACAAGAGCAGGTGGCAGCAGGGGAACGTCTTCTCATGCTCCGTGATGCATGAGGCTCTGCACAACCACTACACGCAGAAGAGCCC
TCGTTCGAGTGGCACCTGTTCTCGTCCACCGTCGTCCCCTTGCAGAAGAGTACGAGGCACTACGTACTCCGAGACGTGTTGGTGATGTGCGTCTTCTCGG
403 SerLysLeuThrValAspLysSerArgTrpGlnGlnGlyAsnValPheSerCysSerValMetHisGluAlaLeuHisAsnHisTyrThrGlnLysSerLeu
mspI
hphI
fnu4HI
bbvI
bspMI
aluI
scrFI
ncil
mspI
hpaII
hpaII
haeIII
xmaIII
eaeI
1101 TCTCCCTGTCTCCGGGTAAATGAGTGCGACGGCCG
AGAGGGACAGAGGCCCATTTACTCACGCTGCCGGC
437 SerLeuSerProGlyLysOP*

FIG. 4B-2



```
                                fnu4HI                                                                  sau96I
                               pstI                                                                     avaII
                            haeIII                                                                     nlaIV
    ecoRI     hphI     fnu4HI bbvI   mboII                                                          mnlI  styI
  1 GAATTCACTCTCACCATCAGCGGCCTGCAGCCTGAAGATTTTGCAACTTATTACTGCCAACAGTATAAGAGTTTGTCGCTCACTTTCGGCGGAGGGACCA
    CTTAAGTGAGAGTGGTAGTCGCCGGACGTCGGACTTCTAAAACGTTGAATAATGACGGTTGTCATATTCTCAAACAGCGAGTGAAAGCCGCCTCCCTGGT
 72       ThrLeuThrIleSerGlyLeuGlnProGluAspPheAlaThrTyrTyrCysGlnGlnTyrLysSerLeuSerLeuThrPheGlyGlyGlyThrLys

          sau3AI          fnu4HI
          dpnI            bbvI         mboII mboII                                          mnlI
101 AGGTGGAGATCAAACGAACTGTGGCTGCACCATCTGTCTTCATCTTCCCGCCATCTGATGAGCAGTTGAAATCTGGAACTGCCTCTGTTGTGTGCCTGCT
    TCCACCTCTAGTTTGCTTGACACCGACGTGGTAGACAGAAGTAGAAGGGCGGTAGACTACTCGTCAACTTTAGACCTTGACGGAGACAACACACGGACGA
104   ValGluIleLysArgThrValAlaAlaProSerValPheIlePheProProSerAspGluGlnLeuLysSerGlyThrAlaSerValValCysLeuLeu
                  ← VKJK CK →
                        haeIII
                        haeI                                                  scrFI
    xmnI                mnlI     rsaI                    mnlI                 bstNI
201 GAATAACTTCTATCCCAGAGAGGCCAAAGTACAGTGGAAGGTGGATAACGCCCTCCAATCGGGTAACTCCCAGGAGAGTGTCACAGAGCAGGACAGCAAG
    CTTATTGAAGATAGGGTCTCTCCGGTTTCATGTCACCTTCCACCTATTGCGGGAGGTTAGCCCATTGAGGGTCCTCTCACAGTGTCTCGTCCTGTCGTTC
137  AsnAsnPheTyrProArgGluAlaLysValGlnTrpLysValAspAsnAlaLeuGlnSerGlyAsnSerGlnGluSerValThrGluGlnAspSerLys

                                                                                                      sacI
                                                                                                      hgiAI
                                                                                                      bsp1286
                                                                                                 haeIII
                        fnu4HI                                                                   sau96I aluI
                    ddeI              ddeI                                                      ecoO   banII
                   mnlI bbvI      hgaI                                  accI            hphI    alwNI ddeI
301 GACAGCACCTACAGCCTCAGCAGCACCCTGACGCTGAGCAAAGCAGACTACGAGAAACACAAAGTCTACGCCTGCGAAGTCACCCATCAGGGCCTGAGCT
    CTGTCGTGGATGTCGGAGTCGTCGTGGGACTGCGACTCGTTTCGTCTGATGCTCTTTGTGTTTCAGATGCGGACGCTTCAGTGGGTAGTCCCGGACTCGA
170 AspSerThrTyrSerLeuSerSerThrLeuThrLeuSerLysAlaAspTyrGluLysHisLysValTyrAlaCysGluValThrHisGlnGlyLeuSerSer

                                                                        ddeI
                   aluI                    mnlI     bsp1286 bspMI  mnlI
401 CGCCCGTCACAAAGAGCTTCAACAGGGGAGAGTGTTAGAGGGAGAAGTGCCCCCACCTGCTCCTCAGT
    GCGGGCAGTGTTTCTCGAAGTTGTCCCCTCTCACAATCTCCCTCTTCACGGGGGTGGACGAGGAGTCA
204   ProValThrLysSerPheAsnArgGlyGluCysAM*
```

FIG. 5

NUCLEIC ACID ENCODING ADHESION VARIANTS

This application is a continuation of U.S. Ser. No. 08/236,311 filed May 2, 1994 now U.S. Pat. No. 5,565,335, which is a continuation of U.S. Ser. No. 07/936,190 filed Aug. 26, 1992, now U.S. Pat. No. 5,336,603, which is a divisional of 07/842,777 filed Feb. 18, 1992, now abandoned, which is a continuation of U.S. Ser. No. 07/250,785 filed Sep. 28, 1988, now abandoned, which is a continuation-in-part of U.S. Ser. No. 07/104,329 filed Oct. 2, 1987, now abandoned.

BACKGROUND OF THE INVENTION

This application relates to compositions for antiviral or immunomodulatory therapy. In particular, it relates to compositions useful in the treatment of Human Immunodeficiency Virus (HIV) infections.

The primary immunologic abnormality resulting from infection by HIV is the progressive depletion and functional impairment of T lymphocytes expressing the CD4 cell surface glycoprotein (H. Lane et al., Ann. Rev. Immunol. 3:477 [1985]). CD4 is a non-polymorphic glycoprotein with homology to the immunoglobulin gene superfamily (P. Maddon et al., Cell 42:93 [1985]). Together with the CD8 surface antigen, CD4 defines two distinct subsets of mature peripheral T cells (E. Reinherz et al., Cell 19:821 [1980]), which are distinguished by their ability to interact with nominal antigen targets in the context of class I and class II major histocompatibility complex (MHC) antigens, respectively (S. Swain, Proc. Natl. Acad. Sci. 78:7101 [1981]; E. Engleman et al., J. Immunol. 127:2124 [1981]; H. Spitz et al., J. Immunol. 129:1563 [1982]; W. Biddison et al., J. Exp. Med. 156:1065 [1982]; and D. Wilde et al., J. Immunol. 131:2178 [1983]). For the most part, CD4 T cells display the helper/inducer T cell phenotype (E. Reinherz, supra), although CD4 T cells characterized as cytotoxic/suppressor T cells have also been identified (Y. Thomas et al., J. Exp. Med. 154:459 [1981]; S. Meuer et al., Proc. Natl. Acad. Sci. USA 79:4395 [1982]; and A. Krensky et al., Proc. Natl. Acad. Sci. USA 79:2365 [1982]). The loss of CD4 helper/inducer T cell function probably underlies the profound defects in cellular and humoral immunity leading to the opportunistic infections and malignancies characteristic of the acquired immunodeficiency syndrome (AIDS) (H. Lane supra).

Studies of HIV-I infection of fractionated CD4 and CD8 T cells from normal donors and AIDS patients have revealed that depletion of CD4 T cells results from the ability of HIV-I to selectively infect, replicate in, and ultimately destroy this T lymphocyte subset (D. Klatzmann et al., Science 225:59 [1984]). The possibility that CD4 itself is an essential component of the cellular receptor for HIV-I was first indicated by the observation that monoclonal antibodies directed against CD4 block HIV-I infection and syncytia induction (A. Dalgleish et al., Nature [London] 312:767 [1984]; J. McDougal et al., J. Immunol. 135:3151 [1985]). This hypothesis has been confirmed by the demonstration that a molecular complex forms between CD4 and gp120, the major envelope glycoprotein of HIV-I (J. McDougal et al., Science 231:382 [1986]); and the finding that HIV-I tropism can be conferred upon ordinarily non-permissive human cells following the stable expression of a CD4 cDNA (P. Maddon et al., Cell 47:333 [1986]). Furthermore, the neurotropic properties of HIV-I, reflected by a high incidence of central nervous system dysfunction in HIV-I

infected individuals (W. Snider et al., Ann. Neurol. 14:403 [1983]), and the ability to detect HIV-I in the brain tissue and cerebrospinal fluid of AIDS patients (G. Shaw et al., Science 227:177 [1985]; L. Epstein, AIDS Res. 1:447 [1985]; S. Koenig, Science 233:1089 [1986]; D. Ho et al., N. Engl. J. Med. 313:1498 [1985]; J. Levy et al., Lancet II:586 [1985]), appears to be explained by the expression of CD4 in cells of neuronal, glial and monocyte/macrophage origin (P. Maddon, Cell 47:444 [1986]; I. Funke et al., J. Exp. Med. 165:1230 [1986]; B. Tourvieille et al., Science 234:610 [1986]).

In addition to determining the susceptibility to HIV-I infection, the manifestation of cytopathic effects in the infected host cell appears to involve CD4. Antibody to CD4 was found to inhibit the fusion of uninfected CD4 T cells with HIV-I infected cells in vitro; moreover, the giant multinucleated cells produced by this event die shortly after being formed, resulting in the depletion of the population of CD4 cells (J. Lifson et al., Science 232:1123 [1986]). Formation of syncytia also requires gp120 expression, and can be elicited by coculturing CD4-positive cell lines with cell lines expressing the HIV-I env gene in the absence of other viral structural or regulatory proteins (J. Sodroski et al., Nature 322:470 [1986]; J. Lifson et al., Nature 323:725 [1986]). Thus, in mediating both the initial infection by HIV-I as well as eventual cell death, the interaction between gp120 and CD4 constitutes one of several critical entry points in the viral life cycle amenable to therapeutic intervention (H. Mitsuya et al., Nature 325:773 [1987]).

The known sequence of the CD4 precursor predicts a hydrophobic signal peptide, an extracellular region of approximately 370 amino acids, a highly hydrophobic stretch with significant identity to the membrane-spanning domain of the class II MHC beta chain, and a highly charged intracellular sequence of 40 residues (P. Madden, Cell 42:93 [1985]). The extracellular domain of CD4 consists of four contiguous regions each having amino acid and structural similarity to the variable and joining (V-J) domains of immunoglobulin light chains as well as related regions in other members of the immunoglobulin gene superfamily (a subclass of which are defined herein by the coined term "adhesons". These structurally similar regions of CD4 are termed the V_1, V_2, V_3 and V_4 domains (denominated 1–4 in FIG. 3).

A successful strategy in the development of drugs for the treatment of many receptor mediated abnormalities has been the identification of antagonists which block binding of the natural ligand. Since the CD4 adheson ordinarily binds to the recognition sites of the HIV envelope it would appear to be a candidate for therapeutically sequestering these HIV sites, thereby blocking viral infectivity. However, full length CD4 and other adhesons are cell membrane proteins which are anchored in the lipid bilayer of cells. The presence of membrane components will be undesirable from the standpoint of manufacturing and purification. In addition, since adhesons are normally present only on cell surfaces, it would be desirable to produce adhesons in a form which is more stable in the circulation. Additionally, even truncated, soluble CD4 adheson (generally referred to as CD4T) may not be optimally effective as a therapeutic since it possesses a relatively short biological half-life, binds to HIV no better than cell surface CD4, may not cross the placental or other biological barriers, and merely sequesters the HIV recognition sites without in itself bearing an infected-cell killing or virus killing functionality.

Accordingly, it is an object of this invention to produce soluble, secreted adhesons. It is another object to produce

CD4 derivatives useful in the treatment of AIDS and related conditions, in a manner essentially unaffected by the extreme degree of genetic variation observed among various HIV-I isolates and their respective env polypeptides (J. Coffin, Cell 46:1 [1986]). Still another object is to prepare adhesons fused to other polypeptides in order to provide molecules with novel functionalities such as those described above for therapeutic use, or diagnostic reagents for the in vitro assay of adhesons or their ligands. In particular, it is an objective to prepare molecules for directing toxins or effector molecules (for example the Fc domain of immunoglobulin) to cells bearing receptors for the adhesons, e.g. HIV gp120 in the case of CD4, and for use in facilitating purification of the adhesons. It is a further object to provide stable, highly purified adheson preparations.

SUMMARY

The objects of this invention are accomplished by providing nucleic acid encoding an amino acid sequence variant of an adheson, in particular a variant in which the transmembrane domain is modified so that it is no longer capable of becoming lodged in the cell membrane. In the case of CD4 such variants are termed soluble CD4.

Variant adhesons are produced by a method comprising (a) transforming a host cell with nucleic acid encoding an amino acid sequence variant of an adheson, (b) culturing the host cell and (c) recovering the variant adheson from the host cell culture media or from lysates of the host cell.

In specific embodiments, the objects of this invention are accomplished by providing an adheson variant selected from the group consisting of (a) an adheson amino acid sequence variant having an inactivated transmembrane domain and (b) a polypeptide comprising an adheson extracellular domain fused to the sequence of a polypeptide which is different from the adheson, this latter, for example, selected from a cytotoxin, an immunogen or a protein with a long plasma half life such as an immunoglobulin constant domain.

In a preferred embodiment a polypeptide comprising a gp120 binding domain of the CD4 adheson is fused at its C-terminus to an immunoglobulin constant domain, or is linked to a cytotoxic polypeptide such as ricin.

The CD4 adheson variants provided herein are purified and formulated in pharmacologically acceptable vehicles for administration to patients in need of antiviral, neuromodulatory or immunomodulatory therapy, in particular patients infected with HIV, and for use in the modulation of cell adhesion.

BRIEF DESCRIPTION OF THE DRAWINGS

FIGS. 1A, 1B-1, 1B-2, and 1C depict the amino acid and nucleotide sequences (SEQ ID NOS:1–3, respectively) of a secreted form of the CD4 adheson. The signal processing site is designated with an arrow.

FIGS. 2A, 2B-1, 2B-2, and 2C depict the amino acid and nucleotide sequences (SEQ ID NOS:4–6, respectively) of a fusion of the herpes gD leader and N-terminal 27 residues to the putative mature N-terminus of CD4T.

FIG. 3 depicts the structural elements of the native and soluble CD4 adheson, the native human IgG_1 (γ_1) heavy chain and two exemplary heavy chain-CD4 chimeras.

FIGS. 4A, 4B-1, and 4B-2 are a map of the linkered human IgG_1 (γ_1) chain fragment employed in the preparation of CD4 fusions (SEQ ID NOS:7–9 for amino acid and nucleotide sequences, respectively). Insert sites are designated γ1 and Fc.

FIG. 5 is a map of the human κ light chain fragment useful for CD4 fusions at the arrow flanked by VκJκ (light variable and joining) and Cκ (light constant) (SEQ ID NOS:10–12 for amino acid and nucleotide sequences, respectively).

DETAILED DESCRIPTION

Adhesons are cell surface polypeptides having an extracellular domain which is homologous to a member of the immunoglobulin gene superfamily, excluding, however, highly polymorphic members of this superfamily selected from the group of class I and class II major histocompatibility antigens, immunoglobulins and T-cell receptor α, β, γ and δ chains. Examples of adhesons include CD1, CD2, CD4, CD8, CD28, the γ, δ and ε chains of CD3, OX-2, Thy-1, the intercellular or neural cell adhesion molecules (I-CAM or N-CAM), lymphocyte function associated antigen-3 (LFA-3), neurocytoplasmic protein (NCP-3), poly-Ig receptor, myelin-associated glycoprotein (MAG), high affinity IgE receptor, the major glycoprotein of peripheral myelin (Po), platelet derived growth factor receptor, colony stimulating factor-1 receptor, macrophage Fc receptor, Fc gamma receptors and carcinoembryonic antigen. Homologous as defined herein means having the sequence of a member of the immunoglobulin gene superfamily or having a sequence therewithin which has substantially the same as (or a greater degree of) amino acid sequence homology to a known member of the superfamily as the specific examples given above have to the sequence of an immunoglobulin variable or constant domain. Preferred adhesons are CD4, CD8 and high affinity IgE Fc receptor.

This invention is particularly concerned with amino acid sequence variants of adhesons. Amino acid sequence variants of adhesons are prepared with various objectives in mind, including increasing the affinity of the adheson for its binding partner, facilitating the stability, purification and preparation of the adheson, increasing its plasma half life, improving therapeutic efficacy as described above in the background, and introducing additional functionalities and lessening the severity or occurrence of side effects during therapeutic use of the adheson. Amino acid sequence variants of adhesons fall into one or a combination of the following classes: insertional, substitutional or deletional variants.

Insertional amino acid sequence variants are those in which one or more amino acid residues extraneous to the adheson are introduced into a predetermined site in the adheson including the C- or N-termini. Such variants are referred to as fusions of the adheson and a different polypeptide. Such other polypeptides contain sequences other than those which are normally found in the adheson at the inserted position. Several groups of fusions are contemplated herein. Immunologically active adheson fusions comprise an adheson and a polypeptide containing a non-adheson epitope. The non-adheson pitope is any immunologically competent polypeptide, i.e., any polypeptide which is capable of eliciting an immune response in the animal to which the fusion is to be administered or which is capable of being bound by an antibody raised against the non-adheson polypeptide. Typical non-adheson epitopes will be those which are borne by allergens, autoimmune epitopes, or other potent immunogens or antigens recognized by pre-existing antibodies in the fusion recipient, including bacterial polypeptides such as trpLE, beta-galactosidase, viral polypeptides such as herpes gD protein, and the like. Immunogenic fusions are produced by cross-linking in vitro or by recombinant cell culture transformed

with DNA encoding an immunogenic polypeptide. It is preferable that the immunogenic fusion be one in which the immunogenic sequence is joined to or inserted into the adheson antigen or fragment thereof by a peptide bond(s). These products therefore consist of a linear polypeptide chain containing adheson epitopes and at least one epitope foreign to the adheson. It will be understood that it is within the scope of this invention to introduce the epitopes anywhere within the adheson molecule or fragment thereof. Such fusions are conveniently made in recombinant host cells or by the use of bifunctional cross-linking agents. The use of a cross-linking agent to fuse the adheson to the immunogenic polypeptide is not as desirable as a linear fusion because the cross-linked products are not as easily synthesized in structurally homogeneous form.

These immunogenic insertions are particularly useful when formulated into a pharmacologically acceptable carrier and administered to a subject in order to raise antibodies against the adheson, which antibodies in turn are useful in diagnostics or in purification of adheson by immunoaffinity techniques known per se. Alternatively, in the purification of adhesons, binding partners for the fused non-adheson polypeptide, e.g. antibodies, receptors or ligands, are used to adsorb the fusion from impure admixtures, after which the fusion is eluted and, if desired, the adheson is recovered from the fusion, e.g. by enzymatic cleavage.

Other fusions, which may or may not also be immunologically active, include fusions of the adheson sequence with a signal sequence heterologous to the adheson, fusions of transmembrane-modified CD4 adhesons, for example, to polypeptides having enhanced plasma half life (ordinarily >about 20 hours) such as immunoglobulin chains or fragments thereof, and fusions with cytotoxic functionalities. Signal sequence fusions are employed in order to more expeditiously direct the secretion of the adheson. The heterologous signal replaces the native adheson signal, and when the resulting fusion is recognized, i.e. processed and cleaved by the host cell, the adheson is secreted. Signals are selected based on the intended host cell, and may include bacterial yeast, mammalian and viral sequences. The herpes gD glycoprotein signal is suitable for use in mammalian expression systems.

Plasma proteins which have enhanced plasma half-life longer than that of transmembrane modified CD4 include serum albumin, immunoglobulins, apolipoproteins, and transferrin. Preferably, the adheson-plasma protein fusion is not significantly immunogenic in the animal in which it is used and the plasma protein does not cause undesirable side effects in patients by virtue of its normal biological activity.

In a specific embodiment the adheson immunoglobulin-like domain which may be homologous either to the constant or to the variable region domains is conjugated with an immunoglobulin constant region sequence. The resulting products are referred to herein as immunoadhesons. Immunoglobulins and certain variants thereof are known and many have been prepared in recombinant cell culture. For example, see U.S. Pat. No. 4,745,055; EP 256,654; Faulkner et al., Nature 298:286 (1982); EP 120,694; EP 125,023; Morrison, J. Immun. 123:793 (1979); Kohler et al., P.N.A.S. USA 77:2197 (1980); Raso et al., Cancer Res. 41:2073 (1981); Morrison et al., Ann. Rev. Immunol. 2:239 (1984); Morrison, Science 229:1202 (1985); Morrison et al., P.N.A.S. USA 81:6851 (1984); EP 255,694; EP 266,663; and WO 88/03559. Reassorted immunoglobulin chains also are known. See for example U.S. Pat. No. 4,444,878; WO 88/03565; and EP 68,763 and references cited therein.

Ordinarily, the domains of adhesons that are homologous to immunoglobulins and extracellular in their native environment are fused C-terminally to the N-terminus of the constant region of immunoglobulins in place of the variable region(s) thereof, retaining at least functionally active hinge, CH2 and CH3 domains of the constant region of an immunoglobulin heavy chain. This ordinarily is accomplished by constructing the appropriate DNA sequence and expressing it in recombinant cell culture. Immunoglobulins and other polypeptides having enhanced plasma half life are fused to the extracellular or ligand binding domains of other adhesons in the same fashion.

The boundary domains for the CD4 V-like regions (V1–V4) are, respectively, about 100–109, about 175–184, about 289–298, and about 360–369 (based on the precursor CD4 amino acid sequence in which the initiating met is –25; FIG. 1A) . CD4 sequences containing any of the CD4 V domains are fused to the immunoglobulin sequence. It is preferable that the V1V2 or V1V2V3V4 be fused at their C-termini to the immunoglobulin constant region. The precise site at which the fusion is made is not critical; the boundary domains noted herein are for guidance only and other sites neighboring or within the V regions may be selected in order to optimize the secretion or binding characteristics of the CD4. The optimal site will be determined by routine experimentation. In general, it has been found that the fusions are expressed intracellularly, but a great deal of variation is encountered in the degree of secretion of the fusions from recombinant hosts. For instance, the following table demonstrates the various immunoglobulin fusions that have been obtained by the method of this invention. In all examples of CD4 immunoadhesons, the CD4 signal was used to direct secretion from 293 cells. Lower case m represents murine origin, while the lower case h designates human origin. V and C are abbreviations for immunoglobulin variable and constant domains, respectively. The numerical subscripts indicate the number of parenthetical units found in the designated multimer. It will be understood that the chains of the multimers are believed to be disulfide bonded in the same fashion as native immunoglobulins. The CD4 immunoadhesons typically contained either the first N-terminal 366 residues of CD4 ($CD4_4$) or the first 180 N-terminal residues of CD4 ($(CD4_2)$ linked at their C-terminus to the κ (light) chain or IgG1 heavy chain constant region (γ1).

TABLE I

Transfected Gene	Secreted Product
mV_KC_K	mV_KC_K and/or $(mV_KC_K)_2$
$mV_{\gamma 1}C_{\gamma 1}$	ND
$mV_KC_K + mV_{\gamma 1}C_{\gamma 1}$	$(mV_KC_K)_2(mV_{\gamma 1}C_{\gamma 1})_2$ + mV_KC_K and/or $(mV_KC_K)_2$
$hCD4\text{-}mC_K$	$hCD4\text{-}mC_K$ and/or $(hCD4\text{-}mC_K)_2$
$hCD4\text{-}mC_{\gamma 1}$	ND
$hCD4\text{-}mC_K + hCD4\text{-}mC_{\gamma 1}$	$(hCD4\text{-}mC_K)_2(hCD4\text{-}mC_{\gamma 1})_2$ + $hCD4\text{-}mC_K$ and/or $(hCD4\text{-}mC_K)_2$
$hCD4\text{-}hC_K$	$hCD4\text{-}hC_K$ and/or $(hCD4\text{-}hC_K)_2$
$hCD4\text{-}hC_{\gamma 1}$	$(hCD4\text{-}hC_{\gamma 1})_2$
$hCD4\text{-}hC_K + hCD4\text{-}hC_{\gamma 1}$	$(hCD4\text{-}hC_K)_2(hCD4\text{-}hC_{\gamma 1})_2$ + $hCD4\text{-}hC_K$ and/or $(hCD4\text{-}hC_K)_2$
$mV_KC_K + hCD4\text{-}hC_{\gamma 1}$	$(mV_KC_K)_2(hCD4\text{-}hC_{\gamma 1})_2$ + mV_KC_K and/or $(mV_KC_K)_2$

*ND = Not detected

It is interesting to observe from this table that the CD4-human heavy chain immunoadheson was secreted as a dimer whereas the analogous murine construction was not detected (this not excluding the intracellular accumulation of the protein, however). The ability of the $hCD4\text{-}hC_\gamma1$ transformants to produce heavy chain dimer was unexpected since

previous work had suggested that immunoglobulin heavy chains are not secreted unless the hosts are cotransformed with nucleic acid encoding both heavy and light chain (Valle et al., *Nature* 241:338 [1981]). According to this invention, CD4-IgG immunoadheson chimeras are readily secreted wherein the CD4 epitope is present in heavy chain dimers, light chain monomers or dimers, and heavy and light chain heterotetramers wherein the CD4 epitope is present fused to one or more light or heavy chains, including heterotetramers wherein up to and including all four variable region analogues are derived from CD4. Where light-heavy chain non-CD4 variable domain is present, a heterofunctional antibody thus is provided.

Various exemplary hetero-and chimeric immunoadheson antibodies produced in accordance with this invention are schematically diagrammed below. "A" means at least a portion of the extracellular domain of an adheson containing its ligand binding site; V_L, V_H, C_L and C_H represent light or heavy chain variable or constant domains of an immunoglobulin; n is an integer; and Y designates a covalent cross-linking moiety.

(a) AC_L;

(b) AC_L-AC_L;

(c) AC_H-[AC_H, AC_L-AC_H, AC_L-V_HC_H, V_LC_L-AC_H, or V_LC_L-V_HC_H];

(d) AC_L-AC_H-[AC_H, AC_L-AC_H, AC_L-V_HC_H, V_LC_L-AC_H, or V_LC_L-V_HC_H];

(e) AC_L-V_HC_H-[AC_H, AC_L-AC_H, AC_L-V_HC_H, V_LC_L-AC_H, or V_LC_L-V_HC_H];

(f) V_LC_L-AC_H-[AC_H, AC_L-AC_H, AC_L-V_HC_H, V_LC_L-AC_H, or V_LC_L-V_HC_H]; or

(g) [A-Y]$_n$-[V_LC_L-V_HC_H]$_2$.

The structures shown in this table show only key features, e.g. they do not show joining (J) or other domains of the immunoglobulins, nor are disulfide bonds shown. These are omitted in the interests of brevity. However, where such domains are required for binding activity they shall be construed as being present in the ordinary locations which they occupy in the adheson, immunoadheson or immunoglobulin molecules as the case may be. These examples are representative of divalent antibodies; more complex structures would result by employing immunoglobulin heavy chain sequences from other classes, e.g. IgM. The immunoglobulin V_LV_H antibody combining site, also designated as the companion immunoglobulin, preferably is capable of binding to a predetermined antigen.

Suitable companion immunoglobulin combining sites and fusion partners are obtained from IgG-1, -2, -3, or -4 subtypes, IgA, IgE, IgD or IgM, but preferably IgG-1.

A preferred embodiment is a fusion of an N-terminal portion of CD4, which contains the binding site for the gp120 envelope protein of HIV, to the C-terminal F_c portion of an antibody, containing the effector functions of immunoglobulin G_1. There are two preferred embodiments of this sort; in one, the entire heavy chain constant region is fused to a portion of CD4; in another, a sequence beginning in the hinge region just upstream of the papain cleavage site which defines IgG F, chemically (residue 216, taking the first residue of heavy chain constant region to be 114 [Kobat et al., "Sequences of Proteins of Immunological Interest" 4th Ed., 1987], or analogous sites of other immunoglobulins) is fused to a portion of CD4. These embodiments are described in the examples.

More particularly, those variants in which one or more immunoglobulin-like domains of an adheson are substituted for the variable region of an immunoglobulin chain are believed to exhibit improved in vivo plasma half life. These chimeras are constructed in a fashion similar to chimeric antibodies in which a variable domain from an antibody of one species is substituted for the variable domain of another species. See, for example, EP 0 125 023; Munro, Nature 312: (Dec. 13, 1984); Neuberger et al., Nature 312: (Dec. 13, 1984); Sharon et al., Nature 309: (May 24, 1984); Morrison et al., Proc. Natl. Acad. Sci. USA 81:6851–6855 (1984); Morrison et al. Science 229:1202–1207 (1985); and Boulianne et al., Nature 312:643–646 (Dec. 13, 1984). The DNA encoding the adheson immunoglobulin-like domain(s) is cleaved by a restriction enzyme at or proximal to the 3' end of the DNA encoding the immunoglobulin-like domain(s) and at a point at or near the DNA encoding the N-terminal end of the mature adheson polypeptide (where use of a different leader is contemplated) or at or proximal to the N-terminal coding region for the adheson (where the native adheson signal is employed). This DNA fragment then is readily inserted into DNA encoding an immunoglobulin light or heavy chain constant region and, if necessary, tailored by deletional mutagenesis. Preferably, this is a human immunoglobulin when the variant is intended for in vivo therapy for humans. DNA encoding immunoglobulin light or heavy chain constant regions is known or readily available from cDNA libraries or is synthesized. See, for example, Adams et al., Biochemistry 19:2711–2719 (1980); Gough et al., Biochemistry 19:2702–2710 (1980); Dolby et al., P.N.A.S. USA, 77:6027–6031 (1980); Rice et al., P.N.A.S. USA 79:7862–7865 (1982); Falkner et al., Nature 298:286–288 (1982); and Morrison et al., Ann. Rev. Immunol. 2:239–256 (1984).

DNA encoding the immunoglobulin or immunoadheson chimeric chain(s) is transfected into a host cell for expression. If the host cell is producing an immunoglobulin prior to transfection then one need only transfect with the adheson fused to light or to heavy chain to produce a heteroantibody. The aforementioned immunoglobulins having one or more arms bearing the adheson domain and one or more arms bearing companion variable regions result in dual specificity for adheson ligand and for an antigen. These are produced by the above-described recombinant methods or by in vitro procedures. In the latter case, for example, $F(ab')_2$ fragments of the adheson fusion and an immunoglobulin are prepared, the F(ab')2 fragments converted to Fab' fragments by reduction under mild reducing conditions, and then reoxidized in each other's presence under acidic conditions in accord with methods known per se. See also U.S. Pat. No. 4,444,878.

Additionally, procedures are known for producing intact heteroantibodies from immunoglobulins having different specificities. These procedures are adopted for the in vitro production of heterochimeric antibodies by simply substituting the immunoadheson chains for one of the previously employed immunoglobulins.

In an alternative method for producing a heterofunctional antibody, host cells producing an adheson-immunoglobulin fusion, e.g. transfected myelomas, also are fused with B cells or hybridomas which secrete antibody having the desired companion specificity for an antigen. Heterobifunctional antibody is recovered from the culture medium of such hybridomas, and thus may be produced somewhat more conveniently than by conventional in vitro resorting methods (EP 68,763).

Another group of fusions are those in which an adheson is conjugated with a toxic substance, e.g. a polypeptide such as ricin (including deglycosylated ricin A chain), diptheria toxin A, or a non-peptidyl cytotoxin. Where the toxin is a polypeptide it is convenient to cross-link the polypeptide to the adheson or its transmembrane-deleted variant by con-

ventional in vitro protein cross-linking agents (for suitable methods for linking ricin A chain or deglycosylated A chain to CD4 see, for example, Duncan et al., "Analy. Biochem." 132:68–73 [1983]; Thorpe et al., "Cancer Res." 47:5924 [1987]; and Ghotie et al., "Cancer Res." 48:2610 [1988]) or by recombinant synthesis as a fusion (see for example, U.S. Pat. No. 4,765,382). Alternatively, where companion antibodies are anti-ricin antibody immunoglobulin variable domains, such immunoglobulin heteroantibodies are employed to deliver ricin to HIV infected cells following the general procedure of Raso et al., Cancer Research, 41:2073 (1981).

Another class of adheson variants are deletional variants. Deletions are characterized by the removal of one or more amino acid residues from an adheson sequence. Typically, the transmembrane and cytoplasmic domains of adhesons are deleted. In the case of CD4, at least residues 368 to 395 (the transmembrane region), and ordinarily 396–433 as well (the cytoplasmic domain), will be deleted to obtain secreted forms of this adheson. Parenthetically, the amino acid residues follow the numbers given for mature CD4 as noted, for example, in FIGS. 1A–1C. Thus, CD4T molecules generally will terminate in the vicinity of about residues 366–368, or at any other suitable site N-terminal thereto which preserves the gp120-binding capability of the CD4 variant.

Substitutional variants are those in which at least one residue in the adheson sequence has been removed and a different residue inserted in its place. The native N-terminal residue for mature CD4 is now known to be lysine. Thus, the sequence shown in FIG. 1, with an N-terminal asparagine, is an amino acid sequence variant of native mature CD4. Table II below describes substitutions which in general will result in fine modulation of the characteristics of the CD antigen.

TABLE II

Original Residue	Exemplary Substitutions
Ala	ser
Arg	lys
Asn	gln; his
Asp	glu
Cys	ser; ala
Gln	asn
Glu	asp
Gly	pro
His	asn; gln
Ile	leu; val
Leu	ile; val
Lys	arg; gln; glu
Met	leu; ile
Phe	met; leu; tyr
Ser	thr
Thr	ser
Trp	tyr
Tyr	trp; phe
Val	ile; leu

Substantial changes in function or immunological identity are made by selecting substitutions that are less conservative than those in Table II, i.e., selecting residues that differ more significantly in their effect on maintaining (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a sheet or helical conformation, (b) the charge or hydrophobicity of the molecule at the target site or (c) the bulk of the side chain. The substitutions which in general are expected to produce the greatest changes in adheson properties will be those in which (a) a hydrophilic residue, e.g. seryl or threonyl, is substituted for (or by) a hydrophobic residue, e.g. leucyl, isoleucyl, phenylalanyl, valyl or alanyl; (b) a cysteinyl or prolyl is substituted for (or by) any other residue; (c) a residue having an electropositive side chain, e.g., lysyl, arginyl, or histidyl, is substituted for (or by) an electronegative residue, e.g., glutamyl or aspartyl; or (d) a residue having a bulky side chain, e.g., phenylalanyl, is substituted for (or by) one not having a side chain, e.g., glycyl.

A preferred class of substitutional or deletional variants are those involving the transmembrane region of the adheson. The transmembrane region of the adheson is a highly hydrophobic or lipophilic domain that is the proper size to span the lipid bilayer of the cellular membrane. It is believed to anchor the adheson in the cell membrane.

Deletion or substitution of the transmembrane domain will facilitate recovery and provide a soluble form of the adheson by reducing its cellular or membrane lipid affinity and improving its water solubility. If the transmembrane and cytoplasmic domains are deleted one avoids the introduction of potentially immunogenic epitopes, either by exposure of otherwise intracellular polypeptides that might be recognized by the body as foreign or by insertion of heterologous polypeptides that are potentially immunogenic. A principal advantage of the transmembrane deleted adheson is that it is secreted into the culture medium of recombinant hosts. This variant is water soluble and does not have an appreciable affinity for cell membrane lipids, thus considerably simplifying its recovery from recombinant cell culture.

It will be amply apparent from the foregoing discussion that substitutions, deletions, insertions or any combination thereof are introduced to arrive at a final construct. As a general proposition, all variants will not have a functional transmembrane domain and preferably will not have a functional cytoplasmic sequence. This is generally accomplished by deletion of the relevant domain, although adequate insertional or substitutional mutagens also can be effective for this purpose. For example, the transmembrane domain is substituted by any amino acid sequence, e.g. a random or homopolynucleic sequence of about 5 to 50 serine, threonine, lysine, arginine, glutamine, aspartic acid and like hydrophilic residues, which altogether exhibit a hydrophilic hydropathy profile, so that it is secreted into the culture medium of recombinant hosts. This variant should also be considered to be an adheson variant.

These variants ordinarily are prepared by site specific mutagenesis of nucleotides in the DNA encoding the adheson, thereby producing DNA encoding the variant, and thereafter expressing the DNA in recombinant cell culture. However, variant adhesons also are prepared by in vitro synthesis. Obviously, variations made in the DNA encoding the variant adhesons must not place the sequence out of reading frame and preferably will not create complementary regions that could produce secondary mRNA structure deleterious to expression (EP 75,444A). The CD4 variants typically exhibit the same gp120 binding activity as does the naturally-occurring prototype, although variants also are selected in order to modify the characteristics of the CD4 adheson as indicated above.

While the site for introducing an amino acid sequence variation is predetermined, the mutation per se need not be predetermined. For example, in order to optimize the performance of a mutation at a given site, random mutagenesis may be conducted at the target codon or region and the expressed adheson variants screened for the optimal combination of desired activities. Techniques for making substitution mutations at predetermined sites in DNA having a known sequence are well known, for example, M13 primer mutagenesis.

Adheson variants that are not capable of binding HIV gp120 are useful nonetheless as immunogens for raising

antibodies to the adheson or as immunoassay kit components (labelled, as a competitive reagent for gp120 assay, or unlabelled as a standard for an adheson assay) so long as at least one adheson epitope remains active.

The DNA encoding adhesons is obtained by known procedures. See Williams, Immunol. Today 8:298–303 (1987) and citations therein. In general, prokaryotes are used for cloning of CD4 variant DNA sequences. For example, *E. coli* strain SR101 (for propagating M13 phage, a λ-resistant strain of JM 101; Messing et al., Nucl. Acids. Res. 9(2) :309–321 [1981]); and *E. coli* K12 strain 294 (ATCC No. 31446) are particularly useful. Other microbial strains which may be used include *E. coli* B, UM101 and *E. coli* x1776 (ATCC No. 31537). These examples are illustrative rather than limiting.

DNA encoding the variant adhesons is inserted for expression into vectors containing promoters and control sequences which are derived from species compatible with the intended host cell. The vector ordinarily, but need not, carry a replication site as well as one or more marker sequences which are capable of providing phenotypic selection in transformed cells. For example, *E. coli* is typically transformed using a derivative of pBR322 which is a plasmid derived from an *E. coli* species (Bolivar, et al., Gene 2: 95 [1977]). pBR322 contains genes for ampicillin and tetracycline resistance and thus provides easy means for identifying transformed cells. The pBR322 plasmid, or other microbial plasmid, must also contain or be modified to contain promoters and other control elements commonly used in recombinant DNA constructions.

Promoters suitable for use with prokaryotic hosts illustratively include the β-lactamase and lactose promoter systems (Chang et al., Nature, 275: 615 [1978]; and Goeddel et al., Nature 281: 544 [1979]), alkaline phosphatase, the tryptophan (trp) promoter system (Goeddel, Nucleic Acids Res. 8: 4057 [1980] and EPO Appln. Publ. No. 36,776) and hybrid promoters such as the tac promoter (H. de Boer et al., Proc. Natl. Acad. Sci. USA 80: 21–25 [1983]). However, other functional bacterial promoters are suitable. Their nucleotide sequences are generally known, thereby enabling a skilled worker operably to ligate them to DNA encoding the adheson variant using linkers or adaptors to supply any required restriction sites (Siebenlist et al., Cell 20: 269 [1980]). Promoters for use in bacterial systems also will contain a Shine-Dalgarno (S.D.) sequence operably linked to the DNA encoding the antigen.

In addition to prokaryotes, eukaryotic microbes such as yeast cultures also are useful as cloning or expression hosts. *Saccharomyces cerevisiae*, or common baker's yeast, is the most commonly used eukaryotic microorganism, although a number of other strains are commonly available. For expression in Saccharomyces, the plasmid YRp7, for example, (Stinchcomb, et al., Nature 282: 39 [1979]; Kingsman et al, Gene 7: 141 [1979]; Tschemper et al., Gene 10: 157 [1980]) is commonly used. This plasmid already contains the trp1 gene which provides a selection marker for a mutant strain of yeast lacking the ability to grow in tryptophan, for example ATCC no. 44076 or PEP4-1 (Jones, Genetics 85: 12 [1977]). The presence of the trp1 lesion as a characteristic of the yeast host cell genome then provides an effective means of selection by growth in the absence of tryptophan.

Suitable promoting sequences for use with yeast hosts include the promoters for 3-phosphoglycerate kinase (Hitzeman et al., J. Biol. Chem. 255: 2073 [1980]) or other glycolytic enzymes (Hess et al., J. Adv. Enzyme Reg. 7: 149 [1968]; and Holland, Biochemistry 17: 4900 [1978]), such as enolase, glyceraldehyde-3-phosphate dehydrogenase, hexokinase, pyruvate decarboxylase, phosphofructokinase, glucose-6-phosphate isomerase, 3-phosphoglycerate mutase, pyruvate kinase, triosephosphate isomerase, phosphoglucose isomerase, and glucokinase.

Other yeast promoters, which are inducible promoters having the additional advantage of transcription controlled by growth conditions, are the promoter regions for alcohol dehydrogenase 2, isocytochrome C, acid phosphatase, degradative enzymes associated with nitrogen metabolism, metallothionein, glyceraldehyde-3-phosphate dehydrogenase, and enzymes responsible for maltose and galactose utilization. Suitable vectors and promoters for use in yeast expression are further described in R. Hitzeman et al., European Patent Publication No. 73,657A. Yeast enhancers also are advantageously used with yeast promoters.

Promoters for controlling transcription from vectors in mammalian host cells may be obtained from various sources, for example, the genomes of viruses such as: polyoma, Simian Virus 40 (SV40), adenovirus, retroviruses, hepatitis-B virus and most preferably cytomegalovirus, or from heterologous mammalian promoters, e.g. the beta actin promoter. The early and late promoters of the SV40 virus are conveniently obtained as an SV40 restriction fragment which also contains the SV40 viral origin of replication. Fiers et al., Nature, 273: 113 (1978). The immediate early promoter of the human cytomegalovirus is conveniently obtained as a HindIII E restriction fragment. Greenaway, P. J. et al., Gene 18: 355–360 (1982). Of course, promoters from the host cell or related species also are useful herein.

DNA transcription in higher eukaryotes is increased by inserting an enhancer sequence into the vector. Enhancers are cis-acting elements of DNA, usually from about 10 to 300 bp, that act to increase the transcription initiation capability of a promoter. Enhancers are relatively orientation and position independent having been found 5' (Laimins, L. et al., Proc. Natl. Acad. Sci. 78: 993 [1981]) and 3' (Lusky, M. L., et al., Mol. Cell Bio. 3: 1108 [1983]) to the transcription unit, within an intron (Banerji, J. L. et al., Cell 33: 729 [1983]) as well as within the coding sequence itself (Osborne, T. F., et al., Mol. Cell Bio. 4: 1293 [1984]). Many enhancer sequences are now known from mammalian genes (globin, elastase, albumin, α-fetoprotein and insulin). Typically, however, one will use an enhancer from a eukaryotic cell virus. Examples include the SV40 enhancer on the late side of the replication origin (bp 100–270), the cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers.

Expression vectors used in eukaryotic host cells (yeast, fungi, insect, plant, animal, human or nucleated cells) may also contain sequences necessary for the termination of transcription which may affect mRNA expression. These regions are transcribed as polyadenylated segments in the untranslated portion of the mRNA encoding the adheson.

Expression vector systems generally will contain a selection gene, also termed a selectable marker. Examples of suitable selectable markers for mammalian cells are dihydrofolate reductase (DHFR), thymidine kinase or neomycin. When such selectable markers are successfully transferred into a mammalian host cell, the transformed mammalian host cell can survive if placed under selective pressure. There are two widely used distinct categories of selective regimes. The first category is based on a cell's metabolism and the use of a mutant cell line which lacks the ability to grow independent of a supplemented medium. Two examples are: CHO DHFR-cells and mouse LTK-cells.

These cells lack the ability to grow without the addition of such nutrients as thymidine or hypoxanthine. Because these cells lack certain genes necessary for a complete nucleotide synthesis pathway, they cannot survive unless the missing nucleotides are provided in a supplemented medium. An alternative to supplementing the medium is to introduce an intact DHFR or TK gene into cells lacking the respective genes, thus altering their growth requirements. Individual cells which were not transformed with the DHFR or TK gene will not be capable of survival in non-supplemented media.

The second category is dominant selection, which refers to a selection scheme used in any cell type and does not require the use of a mutant cell line. These schemes typically use a drug to arrest growth of a host cell. Those cells which have a novel gene would express a protein conveying drug resistance and would survive the selection. Examples of such dominant selection use the drugs neomycin (Southern P. and Berg, P., J. Molec. Appl. Genet. 1: 327 [1982]), mycophenolic acid (Mulligan, R. C. and Berg, P. Science 209: 1422 [1980]) or hygromycin (Sugden, B. et al., Mol. Cell. Biol. 5: 410–413 [1985]). The three examples given above employ bacterial genes under eukaryotic control to convey resistance to the appropriate drug G418 or neomycin (geneticin), xgpt (mycophenolic acid) or bygromycin, respectively.

"Amplification" refers to the increase or replication of an isolated region within a cell's chromosomal DNA. Amplification is achieved using a selection agent e.g. methotrexate (MTX) which inactivates DHFR. Amplification or the making of successive copies of the DHFR gene results in greater amounts of DHFR being produced in the face of greater amounts of MTX. Amplification pressure is applied notwithstanding the presence of endogenous DHFR, by adding ever greater amounts of MTX to the media. Amplification of a desired gene can be achieved by cotransfecting a mammalian host cell with a plasmid having a DNA encoding a desired protein and the DHFR or amplification gene permitting cointegration. One ensures that the cell requires more DHFR, which requirement is met by replication of the selection gene, by selecting only for cells that can grow in the presence of ever-greater MTX concentration. So long as the gene encoding a desired heterologous protein has cointegrated with the selection gene, replication of this gene gives rise to replication of the gene encoding the desired protein. The result is that increased copies of the gene, i.e. an amplified gene, encoding the desired heterologous protein express more of the desired heterologous protein. Preferred host cells for expressing the CD antigen variants of this invention are mammalian cell lines, examples including: monkey kidney CV1 line transformed by SV40 (COS-7, ATCC CRL 1651); human embryonic kidney line (293, Graham, F. L. et al., J. Gen Virol. 36: 59 [1977] and 293S cells [293 subclones selected for better suspension growth]); baby hamster kidney cells (BHK, ATCC CCL 10); chinese hamster ovary-cells-DHFR (CHO, Urlaub and Chasin, Proc. Natl. Acad. Sci. (USA) 77: 4216, [1980]); mouse sertoli cells (TM4, Mather, J. P., Biol. Reprod. 23: 243–251 [1980]); monkey kidney cells (CV1 ATCC CCL 70); african green monkey kidney cells (VERO-76, ATCC CRL-1587); human cervical carcinoma cells (HELA, ATCC CCL 2); canine kidney cells (MDCK, ATCC CCL 34); buffalo rat liver cells (BRL 3A, ATCC CRL 1442); human lung cells (W138, ATCC CCL 75); human liver cells (Hep G2, HB 8065); mouse mammary tumor cells (MMT 060562, ATCC CCL51); and TRI cells (Mather, J.P. et al., Annals N.Y. Acad. Sci. 383: 44–68 [1982]).

"Transformation" means introducing DNA into an organism so that the DNA is replicable, either as an extrachromosomal element or by chromosomal integration. One suitable transformation of the host cells is the method of Graham, F. and van der Eb, A., Virology 52: 456–457 (1973). However, other methods for introducing DNA into cells such as by nuclear injection or by protoplast fusion may also be used. If prokaryotic cells or cells which contain substantial cell walls are used as hosts, the preferred method of transfection is calcium treatment using calcium chloride as described by Cohen, F. N. et al., Proc. Natl. Acad. Sci. (USA), 69: 2110 (1972).

Construction of suitable vectors containing the desired coding and control sequences employs standard and manipulative ligation techniques. Isolated plasmids or DNA fragments are cleaved, tailored, and religated in the form desired to form the plasmids required. Suitable procedures are well known for the construction described herein. See, for example, Maniatis, T. et al., *Molecular Cloning*, 133–134 Cold Spring Harbor, (1982); "Current Protocols in Molecular Biology", edited by Ausubel et al., (1987), pub. by Greene Publishing Associates & Wiley-Interscience.

Correct plasmid sequences are confirmed by transforming *E. coli* K12 strain 294 (ATCC 31446) with ligation mixtures, successful transformants are selected by ampicillin or tetracycline resistance where appropriate, and plasmids from the transformants are prepared, and then analyzed by restriction enzyme digestion and/or sequenced by the method of Messing et al., Nucleic Acids Res. 9: 309 (1981) or by the method of Maxam et al., Methods in Enzymology 65: 499 (1980).

Host cells are transformed with the expression vectors of this invention. Thereafter they are cultured in appropriate culture media, e.g. containing substances for inducing promoters, selecting transformants or amplifying genes. The culture conditions, such as temperature, pH and the like, are those previously used with the host cell selected for expression, and will be apparent to the ordinarily skilled artisan.

The secreted adheson variants are recovered and purified from the culture supernatants or lysates of recombinant hosts.

Typically, the supernatants are concentrated by ultrafiltration, contacted with a ligand affinity or immunoaffinity matrix so as to adsorb the adheson variant, and eluted from the matrix. Optionally, the adheson is purified by ion exchange chromatography.

Surprisingly, purification of soluble CD4 adheson from culture medium was unexpectedly difficult. Notwithstanding that the hydrophobic transmembrane region of the antigen had been deleted, the antigen exhibited a strong tendency to form aggregates that could be readily removed from suspension by centrifugation at 1000 × g, and which avidly coat surfaces such as ultrafiltration membranes. This appears to result from the reduction in concentration of albumin or other serum protein (ordinarily present in the crude preparation) to a particular level, below which the truncated antigen no longer remains soluble. This phenomenon appears to be aggravated by exposure of the CD4 adheson to low pH (<about pH 4). As a result, separation procedures (particularly those that employ acid elution, such as immunoaffinity) should be modified so that the eluate is maintained at, or immediately returned to, about neutrality. Further, a surfactant, e.g. a detergent such as Tween 80, should be included with the antigen during the separation procedure. The final purified product will be stabilized with a predetermined protein such as albumin, and/or a detergent.

The purified adheson is formulated into conventional pharmacologically acceptable excipients.

It is administered to patients having HIV infection at a dosage capable of maintaining a concentration of greater than about 100 ng of soluble CD4 adheson/ml plasma. For CD4 adheson variants having different molecular weights, about 2 picomoles of soluble receptor per ml of plasma will be initially evaluated clinically in order to establish a stoichiometric equivalence with native (membrane bound) and soluble receptor. The ordinary dosage of soluble CD4 is 100 μg/kg of patient weight/day.

The therapeutic CD4 variants are employed with other therapies and agents for the treatment of AIDS, including AZT, neutralizing antibodies and immunocytotoxins, gp120 fragments and vaccines.

In order to facilitate understanding of the following examples certain frequently occurring methods and/or terms will be described.

"Plasmids" are designated by a lower case p preceded and/or followed by capital letters and/or numbers. The starting plasmids herein are either commercially available, publicly available on an unrestricted basis, or can be constructed from available plasmids in accord with published procedures. In addition, equivalent plasmids to those described are known in the art and will be apparent to the ordinarily skilled artisan.

"Digestion" of DNA refers to catalytic cleavage of the DNA with a restriction enzyme that acts only at certain sequences in the DNA. The various restriction enzymes used herein are commercially available and their reaction conditions, cofactors and other requirements are used as would be known to the ordinarily skilled artisan. For analytical purposes, typically 1 μg of plasmid or DNA fragment is used with about 2 units of enzyme in about 20 μl of buffer solution. For the purpose of isolating DNA fragments for plasmid construction, typically 5 to 50 μg of DNA are digested with 20 to 250 units of enzyme in a larger volume. Appropriate buffers and substrate amounts for particular restriction enzymes are specified by the manufacturer. Incubation times of about 1 hour at 37° C. are ordinarily used, but may vary in accordance with the supplier's instructions. After digestion the reaction is electrophoresed directly on a polyacrylamide gel to isolate the desired fragment.

"Recovery" or "isolation" of a given fragment of DNA from a restriction digest means separation of the digest on polyacrylamide or agarose gel by electrophoresis, identification of the fragment of interest by comparison of its mobility versus that of marker DNA fragments of known molecular weight, removal of the gel section containing the desired fragment, and separation of the gel from DNA. This procedure is known generally (Lawn, R. et al., Nucleic Acids Res. 9: 6103–6114 [1981], and Goeddel, D. et al., Nucleic Acids Res. 8: 4057 [1980]).

"Dephosphorylation" refers to the removal of the terminal 5' phosphates by treatment with bacterial alkaline phosphatase (BAP). This procedure prevents the two restriction cleaved ends of a DNA fragment from "circularizing" or forming a closed loop that would impede insertion of another DNA fragment at the restriction site. Procedures and reagents for dephosphorylation and other recombinant manipulations are conventional. Reactions using BAP are carried out in 50 mM Tris at 68° C. to suppress the activity of any exonucleases which may be present in the enzyme preparations. Reactions are run for 1 hour. Following the reaction the DNA fragment is gel purified.

"Ligation" refers to the process of forming phosphodiester bonds between two double stranded nucleic acid fragments (Maniatis, T. et al., Id. at 146). Unless otherwise provided, ligation may be accomplished using known buffers and conditions with 10 units of T4 DNA ligase ("ligase") per 0.5 μg of approximately equimolar amounts of the DNA fragments to be ligated.

"Filling" or "blunting" refers to the procedures by which the single stranded end in the cohesive terminus of a restriction enzyme-cleaved nucleic acid is converted to a double strand. This eliminates the cohesive terminus and forms a blunt end. This process is a versatile tool for converting a restriction cut end that may be cohesive with the ends created by only one or a few other restriction enzymes into a terminus compatible with any blunt-cutting restriction endonuclease or other filled cohesive terminus. Typically, blunting is accomplished by incubating 2–15 μg of the target DNA in 10 mM $MgCl_2$, 1 mM dithiothreitol, 50 mM NaCl, 10 mM Tris (pH 7.5) buffer at about 37° C. in the presence of 8 units of the Klenow fragment of DNA polymerase I and 250 7μm of each of the four deoxynucleoside triphosphates. The incubation generally is terminated after 30 min. and the reaction mixture is subjected to phenol and chloroform extraction and ethanol precipitation.

The following examples merely illustrate the best mode now contemplated for practicing the invention, but should not be construed to limit the invention. All literature citations herein are expressly incorporated by reference.

EXAMPLE 1

Construction of Vectors for the Expression of Native CD4 and Secreted Derivatives

Section 1

The plasmid used for recombinant synthesis of human CD4 was pSVeCD4DHFR. The plasmid was constructed as follows:

λCD4P1 containing most of the coding sequence of human CD4 (obtained from a human placental cDNA library using oligonucleotide probes based on the published sequence Maddon et al. (1985) was digested with EcoRI to produce the cDNA insert. This fragment was recovered by polyacrylamide gel electrophoresis (fragment 1).

pUC18 was digested with EcoRI and the single fragment recovered by polyacrylamide gel electrophoresis (fragment 2). Fragment 1 was ligated to fragment 2 and the ligation mixture transformed into *E. coli* strain 294. The transformed culture was plated on ampicillin media plates and resistant colonies were selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct DNA fragments. This plasmid is referred to as pUCCD4.

pSVeE'DHFR (Muesing et al., Cell 48:691–701 [1987]) was digested with KpnI and BamHI and blunted with *E. coli* DNA polymerase I (Klenow fragment) and the four dNTPs. Fragment 3 containing the pML-Amp^r region, SV40 early promoter, the HIV LTR, and the mouse DHFR gene was recovered by gel electrophoresis, ligated and the ligation mixture transformed into *E. coli* strain 294. The transformed culture was plated on ampicillin media plates and resistant colonies selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the BamHI restriction site and the absence of the KpnI restriction site. This plasmid is referred to as pSVeΔBKDHFR and allows EcoRI-BamHI fragments to be inserted after the SV40 early promoter and transcribed under its control, following transfection into an appropriate cell line.

Synthetic oligonucleotides (adaptors 1–8, below) were made to extend from 76 bp 5' of the initiation codon of CD4 translation to the RsaI restriction site at 121 bp 3' of the initiator, with the sequence AATT at the 5' end of the sense

strand to generate an end which could ligate to an EcoRI restriction fragment. These oligonucleotides were ligated and the 204 bp fragment containing the entire sequence was recovered by gel electrophoresis (fragment 4).

CD4 adaptor 1: AATTCAAGCCCAGAGCCCTGC-CATTTCTGTGGGCTCAGGTCCCT (SEQ ID NO: 13)

CD4 adaptor 2: pACTGCTCAGCCCCTTCCTCCCTCG-GCAAGGCCACAATGAACCGGGGAGTC (SEQ ID NO: 14)

CD4 adaptor 3: pCCTTTTAGGCACTTGCTTCTGGT-GCTGCAACTGGCGCTCCTCCCAGC (SEQ ID NO: 15)

CD4 adaptor 4: pAGCCACTCAGGGAAACAAAGTGGT-GCTGGGCAAAAAAGGGGATACAGTG-GAACTGACCT GT (SEQ ID NO: 16)

CD4 adaptor 5: pACAGGTCAGTTCCACTGTATC-CCCTTTTTTGCCCAGCACCACTTTGTTTCC (SEQ ID NO: 17)

CD4 adaptor 6: pCTGAGTGGCTGCTGGGAGGAGCGC-CAGTTGCAGCACCAGAAGCAAGT (SEQ ID NO: 18)

CD4 adaptor 7: pGCCTAAAAGGGACTCCCCGGTTCAT-TGTGGCCTTGCCGAGGGAGGAAGGG (SEQ ID NO: 19)

CD4 adaptor 8: GCTGAGCAGTAGGGACCTGACCCCA-CAGAAATGGCAGGGCTCTGGGCTTG (SEQ ID NO: 20)

pUCCD4 was digested with RsaI and SstII and the 401 bp fragment containing part of the CD4 coding sequence recovered by gel electrophoresis (fragment 5). pUC18 was digested with EcoRI and SstI and the fragment comprising the bulk of the plasmid recovered by gel electrophoresis (fragment 6). Fragments 4 and 5 were ligated to fragment 6 and the ligation mixture was transformed into *E. coli* strain 294. The transformed culture was plated on ampicillin media plates and resistant colonies were selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct fragment. The sequence of the inserted synthetic DNA was checked by excising the 605 bp EcoRI-SstI fragments from several transformants and ligating them to M13mp19 which had been digested with the same enzymes. After transformation into *E. coli* strain JM101, single-stranded DNA was prepared and sequenced. One plasmid which contained the correct sequence was selected, and is referred to as pCD4int.

pCD4int was digested with EcoRI and SstII and fragment 7 containing the 5' end of the CD4 coding region was recovered by gel electrophoresis. pUCCD4 was digested with SstI and BamHI and the 1139 bp fragment containing the remainder of the CD4 coding region (fragment 8) recovered by gel electrophoresis.

pSVcΔBKDHFR was digested with EcoRI and BamHI and fragment 9 comprising the bulk of the plasmid was isolated. Fragments 7, 8 and 9 were ligated and the ligation mixture was transformed into *E. coli* strain 294. The transformed culture was plated on ampicillin media plates and the resistant colonies were selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct fragment. This plasmid is referred to as pSVeCD4DHFR, and was used to direct synthesis of recombinant intact CD4.

Section 2

A plasmid was constructed to direct the synthesis of a CD4 derivative lacking the putative transmembrane domain and most of the putative cytoplasmic domain (Maddon et al., supra). This was done with the intention of creating a secreted form of CD4, based on the assumption that these domains anchor the CD4 glycoprotein to the cell membrane, and that their deletion would result in the secretion of the product. This plasmid is referred to as pSVeCD4ΔNlaDHFR and was constructed as follows:

pUCCD4 was digested with SstI and TaqI and the 531 bp fragment (fragment 10) recovered. pUCCD4 was digested with NlaIII and TaqI and the 112 bp fragment (fragment 11) recovered. pUCCD4 was digested with BamHI and NlaIII and the 301 bp fragment (fragment 12) recovered. pCD4int was digested with SstI and BamHI and fragment 13 comprising the bulk of the plasmid recovered. Fragments 10, 11, and 12 were ligated together with fragment 13 and the ligation mixture was transformed into *E. coli* strain 294. The transformed culture was plated on ampicillin media plates and resistant colonies were selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct fragment. Plasmid DNA from several transformants was sequenced to ensure that the 195 bp NlaIII fragment had been deleted and that the proper reading frame was restored. The resulting plasmid is referred to as pCD4ΔNla. pCD4ΔNla was digested with EcoRI and BamHI and the 1541 bp fragment containing the sequence of a CD4 derivative lacking the transmembrane and cytoplasmic domains recovered (fragment 14) and ligated to fragment 9, and the ligation mixture was transformed into *E. coli* strain 294. The transformed culture was plated on ampicillin media plates and resistant colonies were selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct fragment. This plasmid is referred to as pSVeCD4DNlaDHFR.

Both pSVeCD4DHFR and pSVeCD4ΔNlaDHFR were transfected into CHO cells by the same method used to establish cell lines stably expressing HIV-I polypeptides (Muesing, Smith and Capon, Cell 48:6910701 [1987]). These cells were assayed for production by radioimmunoprecipitation as described below. While no product was detected in initial experiments, subsequent experiments showed that the above described coding segment could indeed direct the synthesis of a soluble CD4 adheson variant both in CHO and 293 cells.

Section 3

A different expression system was initially used for the synthesis and expression of a CD4 variant lacking completely the cytoplasmic and transmembrane domains. This system uses the cytomegalovirus promoter and can be used in cultured cells of human origin. The first plasmid constructed for use in this system contained the entire coding region for CD4 and was intended to function as a control in the following studies. It is referred to as pRKCD4, and was constructed as follows:

pSVeCD4DHFR was digested with EcoRI and BamHI and fragment 15 containing the entire CD4 coding region was isolated. pRK5 (U.S. Ser. No. 97,472, filed Sep. 11, 1987, now abandoned) was digested with EcoRI and BamHI and fragment 16 comprising the bulk of the plasmid recovered by gel electrophoresis, ligated to fragment 15, and the ligation mixture transformed into *E. coli* strain 294. The transformed culture was plated on ampicillin media plates and resistant colonies selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct fragment. This plasmid is referred to as pRKCD4.

Section 4

The next plasmid constructed was designed to direct the expression of the above-mentioned (Section 3) secreted derivative of CD4. The coding region of CD4 was fused

after amino acid residue 368 of mature CD4 to a sequence from pBR322 which codes for 9 more residues before a translation termination codon. This removes the putative CD4 transmembrane and cytoplasmic domains, which are presumed to anchor CD4 to the cell surface. The plasmid is referred to as pRKCD4T (and produces a protein called CD4T), and was constructed as follows:

pSVeCD4DHFR was digested with HpaII, blunted with Klenow fragment and the four dNTPs, and digested with BstEII. The 382 bp fragment (fragment 17) containing part of the CD4 coding sequence was recovered by gel electrophoresis. pSVeCD4DHFR was digested with EcoRI and BstEII and the 874 bp fragment (fragment 18) recovered. pBR322 was digested with HindIII, blunted with Klenow fragment and the four dNTPs, and digested with EcoRI. Fragment 19 comprising the bulk of the plasmid was isolated and ligated to fragments 17 and 18 and the ligation mixture transformed into *E. coli* strain 294. The transformed culture was plated on ampicillin media plates and resistant colonies were selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct fragment. This plasmid is referred to as pCD4Tint.

pRK5 was digested with EcoRI and SmaI and fragment 20 comprising the bulk of the plasmid isolated. pCD4Tint was digested with EcoRI and EcoRV and the 1410 bp fragment containing the CD4 coding sequence to the HpaII site at 1176 bp 3' of the initiating codon and the 154 bp HindIII-EcoRV fragment of pBR322 was recovered (fragment 21). Fragments and 21 were ligated and the ligation mixture was transformed into *E. coli* strain 294. The transformed culture was plated on ampicillin media plates and resistant colonies were selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct fragment. This plasmid is referred to as pRKCD4T.

Section 5a

In order to create a secreted form of CD4 which could be purified with an antibody directed to herpes virus type I glycoprotein D, a plasmid was constructed to express a derivative of CD4T in which the region coding for the mature, processed CD4T polypeptide was fused to a sequence coding for the signal peptide and the first 27 residues of the mature type I Herpes Simplex Virus gD glycoprotein. This plasmid is referred to as pRKGDCD4T, and was constructed as follows:

pgDTrunc.DHFR was digested with EcoRI and PvuII and the fragment containing the coding region for the signal peptide and first 27 residues of the mature HSV I gD glycoprotein was isolated (fragment 22). pRKCD4T was digested with EcoRI and BstEII and fragment 23 containing the 3' end of the CD4 coding sequence and the pRK5 region was isolated.

Synthetic oligonucleotides GD (adaptors 1–2, below) containing the coding sequence of CD4 from the codon for the amino terminal residue of mature CD4 to the RsaI site at 121 bp 3' of translation initiation, and containing the sequence CTGCTCGAG at the 5' end of the sense strand were prepared (fragment 24). pRKCD4 was digested with RsaI and BstEII and the 665 bp fragment containing part of the coding region for CD4 was recovered (fragment 25) and ligated to fragment 24. After digestion with BstEII to ensure that only monomeric fragment was present, the 724 bp fragment containing both sequences was recovered by gel electrophoresis (fragment 26).

Fragments 22, 23 and 26 were ligated and the ligation mixture was transformed into *E. coli* strain 294. The transformed culture was plated on ampicillin media plates and resistant colonies were selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct fragment. The sequence of several transformants was checked to ensure that the synthetic insert was correct and that reading frame was preserved. This plasmid is referred to as pRKGDCD4T.

These pRK5 derived plasmids preferably were transfected into 293S cells for stable expression according to Muesing, et al. Cell 48:691 (1987), with the exception that in addition to the plasmid of interest a plasmid expressing the neomycin resistance gene pRSV neo (Gorman et al. Science 221:553–555 [1985]) was cotransfected. 293 cells also are used satisfactorily as host cells. 2 days after transfection, the cells were passaged into standard medium (1:1 F12/DME supplemented with L-glutamine, penicillin-streptomycin and 10% FBS) with 0.5 mg/ml G418 (Geneticin sulfate; Gibco) for selection of stable cell lines, rather than in media containing methotrexate as shown by Muesing et al. Cells were assayed for production of CD4 or CD4 analogs by radioimmunoprecipitation. Binding studies (section 5c) used conditioned supernatants from these cells in the 1:1 F12/DME medium. Materials used in infectivity assays (section 5b) were obtained as described in section 8 below.

gDCD4 adaptor 1: CTGCTCGAGCAGGGAAA-CAAAGTGGTGCTGGGCAAAAAAGGG-GATACAGTGGAACTGAC (SEQ ID NO: 21)

gDCD4 adaptor 2: pACAGGTCAGTTCCACTGTATC-CCCTTTTTTGCCCAGCAC-CACTTTGTTTCCCTGCTCGA (SEQ ID NO: 22)

Section 5b

The following constitutes a study of the neutralization of HIV-1 infectivity by soluble CD4 analogs. A modification of the neutralization procedure of Robert-Guroff et al., Nature 316:72 (1985) was followed. Equal volumes of inhibitor supernatant and virus (60 microliters) were incubated at 4 degrees C for 1 hour, then the same volume of H9 (Gallo et al., Science 224:500, [1984]) at 5×10^6/ml was added and incubation continued for 1 hour at 37 degrees C. Following absorption, 2.5×10^5 cells in 150 microliters were transferred to 2 ml of incubation media. After 4 days at 37 degrees C, the cultures were split 1:2 with fresh media and incubated for an additional 3 days. Cultures were harvested, reverse transcriptase activity was measured (Groopman et al., AIDS Research and Human Retroviruses 3:71 [1987]), and immunofluorescence reactivity with HIV-1 positive serum was determined as described (Poiesz et al., Proc. Acad. Nat. Sci. USA 77:7415 [1980]). Inhibitor supernatants were obtained from confluent plate cultures of 293S/CDT4 cells, 293S/gDCD4T cells or untransfected 293S cells by replacing the growth medium with incubation media and harvesting the supernatants 24 hours later. Inhibitor supernatant replaced part or all of the incubation media during the first three days of culture as indicated in the second column of Table III. Challenge dose of virus was 100 $TCID_{50}$, (Groopman et al., supra) of HIV-1 strain HTLV-IIIB grown in H9 cells assayed in the same system. Incubation media consisted of RPMI 1640 media containing 2 mM L-glutamine, 100 units/ml penicillin, 100 micrograms/ml streptomycin, 2 micrograms/ml polybrene and 20% fetal calf serum (M.A. Bioproducts)c

TABLE III

Inhibitor supernatant	Dilution of inhibitor supernatant	Indirect immunofluorescence (% positive cells)		Reverse transcriptase (cpm/ml × 10^6)	
mock-transfected	undil.; 1:4	65.3	65.5	21.8	23.9
mock-transfected	undil.; 1:4	61.2	61.1	38.5	28.1
CD4T	undil.; 1:4	0.4	18.0	0.11	5.94
CD4T	undil.; 1:4	0.8	16.1	0.15	3.72
gDCD4T	undil.; 1:4	0.4	26.8	0.14	9.92
gDCD4T	undil.; 1:4	1.4	36.1	0.23	11.3

Both forms of soluble CD4 virtually abolished the growth of HIV-1, when incubated with virus-infected cells without prior dilution (Table III). At a dilution of 1:4 the soluble CD4 preparations were only partially effective in inhibiting virus growth; however, the level of fluorescent-positive cells and reverse transcriptase was still significantly lower than cultures receiving mock-transfected cell supernatants (Table III). Since there was no significant difference in virus growth between diluted and undiluted control supernatants, nor did any of the supernatants affect the growth of uninfected H9 cells (data not shown), soluble 1D4 proteins present in these supernatants were concluded to be responsible for the neutralization of HIV-1 infection of H9 cells.

Section 5c

To determine the affinity constant for interactions between gp120 and CD4 or CD4 variants, saturation binding analysis was carried out with soluble CD4 (supra) and detergent solubilized intact CD4 (Lasky et al. Cell 50:975 [1987]) employing radioiodinated gp120 labeled with lactoperoxidase. Binding reactions consisted of ^{125}I-gp120 (3 ng to 670 ng, 2.9 nCi/ng) incubated for 1 hour at 0 degrees C with cell lysates containing intact CD4 (Lasky et al., supra) or cell supernatants containing unlabeled CD4T or gDCD4T prepared as described in section 5a. Reactions (0.2 ml) had a final composition of 0.5× McDougal Lysis Buffer (McDLB) (1× McDLB contains 0.5% Nonidet NP-40, 0.2% Na deoxycholate, 0.12 M NaCl, 0.02 M Tris-HCl, pH 8.0) and were performed in duplicate, both in the presence or absence of 50 micrograms of unlabeled purified gp120 (74 fold or greater excess). Following incubation, bound gp120 was quantitated by immunoprecipitation and counted in a gamma counter. For immunoprecipitation, binding reaction solutions were preabsorbed with 5 microliters of normal rabbit serum for one hour at 0° C., and cleared with 40 microliters of Pansorbin (10% w/v, Calbiochem) for 30 minutes at 0 degrees C. Samples were then incubated overnight at 0 degrees C with 2 microliters of normal serum or 5 microliters (0.25 microgram) of OKT4 monoclonal antibody (Ortho) followed by collection of immune complexes with 10 microliters of Pansorbin. Precipitates were washed twice in 1× McDLB and once in water, then eluted by eluting at 100 degrees C for 2 minutes in sample buffer (0.12 M Tris-HCl pH 6.8, 4% SDS, 0.7 M mercaptoethanol, 20% glycerol, and 0.1% bromophenol blue). CD4 molecules were bound saturably by gp120, and yielded a simple mass action binding curve. Supernatants from mock-transfected cells gave a level of specifically bound gp120 less than 1% that found for supernatants containing soluble CD4. Scatchard analysis revealed a single class of binding sites on each molecule, with apparent dissociation constants (Kd) of 1.3×10^{-9} M, 0.83×10^{-9} M and 0.72×10^{-9} M for intact CD4, CD4T and gDCD4T, respectively. The values obtained for CD4-gp120 binding in solution are comparable to the affinity previously measured for gp120 binding to CD4 on whole cells (Kd=4.0×10^{-9} M; Lasky, Cell, supra).

Section 6

In order to produce secreted derivatives of CD4 which are free of extraneous amino acid residues, two plasmids were constructed for expression in 293 cells. The plasmids contain CD4 genes which have been truncated without the addition of extra residues, and are referred to as pRKCD4ΔNla and pRKCD4TP (and produce proteins called CD4ΔNla and CD4TP, respectively), and were constructed as follows:

Fragment 14 containing the CD4 gene with the 195 bp NlaIII restriction fragment deleted was ligated to fragment 16, which is pRK5 digested with EcoRI and BamHI. The ligation mixture was transformed into *E. coli* strain 294, the transformed culture plated on ampicillin media plates and resistant colonies were selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct fragment. The resulting plasmid is referred to as pRKCD4ΔNla.

Synthetic DNA (5' CGT GAT AGA AGC TTT CTA GAG 3') (SEQ ID NO: 23) was made to attach to the HpaII site at 1176 bp and so that when so attached it would terminate translation after amino acid residue 368 of mature CD4 (fragment 27). The other end of this fragment was designed to ligate to BamHI restriction fragments. pUCCD4 was digested with BstEII and HpaII and the 382 bp fragment containing part of the CD4 gene was recovered (fragment 28). Fragments 27 and 28 were ligated and then digested with BstEII to reduce dimerized fragments to monomers, and the resulting 401 bp fragment was recovered (fragment 29).

pRKCD4 was digested with BstII and BamHI and the fragment comprising the bulk of the plasmid (fragment 30) was isolated and ligated to fragment 29. The ligation mixture was transformed into *E. coli* strain 294, the transformed culture plated on ampicillin media plates and resistant colonies were selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct fragment. The resulting plasmid is referred to as pRKCD4TP. Both plasmids are transfected into 293 cells to generate stable variant CD4-expressing cell lines as described above.

Section 7

Two plasmids were constructed to direct the expression of secreted CD4 lacking extraneous amino acid residues in CHO cells. These are referred to as pSVeCD4ΔNlaSVDHFR and pSVeCD4TPSVDHFR (and encode proteins having the primary sequence of CD4ΔNla and CD4TP, respectively), and were constructed as follows:

pE348HBV.E400D22 was digested with PvuI and EcoRI and the fragment containing the SV40 early promoter and part of the β-lactamase gene was recovered (fragment 31). pE348HBV.E400D22 was digested with PvuI and BamHI and the large fragment containing the balance of the β-lactamase gene as well as the SV40 early promoter and the DHFR gene was isolated (fragment 32).

Fragments 31 and 32 were ligated together with fragment 14 and transformed into *E. coli* strain 294. The transformed culture was plated on ampicillin media plates and resistant colonies were selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct fragment. The resulting plasmid is referred to as pSVECD4ΔNlaSVDHFR. This plasmid contains the same DNA fragment encoding the soluble CD4 molecule found in the above-mentioned plasmid pSVeCD4DNlaDHFR (Section 2).

pRKCD4TP was digested with EcoRI and BamHI and the fragment containing the truncated CD4 coding region was

isolated and ligated to fragments 31 and 32. The ligation mixture was transformed into *E. coli* strain 294, the transformed culture plated on ampicillin media plates and resistant colonies were selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct fragment. The resulting plasmid is referred to as pSVeCD4TPSVDHFR. Both of these plasmids are transfected into CHO cells and amplified transfectants selected by methotrexate using conventional procedures.

EXAMPLE 2

Fusions of the V region of the CD4 gene, which is homologous to the variable region of immunoglobulin genes (Maddon et al. 1985, supra), to the constant (C) region of human immunoglobulin κ and γ2 chains are constructed as follows:

Synthetic DNA is made to code for the C region of human κ chain (residues 109–214) based on the sequence published by Morin et al., Proc. Natl. Acad. Sci. 82:7025–7029 (1985), with the addition at the 5' end of the coding strand of the sequence GGGG, which allows this fragment to be ligated to the BspMI site at the end of the putative V-like region of CD4. At the 3' end of the coding region, a translational stop codon is added as well as a sequence which allows this end to be ligated to BamHI restriction fragments. The synthetic DNA is made in 8 fragments, 4 for each strand, 70–90 bases long. These are then allowed to anneal and are ligated prior to isolation on a polyacrylamide gel (fragment 33).

pRKCD4 is digested with EcoRI and BspMI and the 478 bp fragment containing the region coding for the putative V-like domain of CD4 is recovered (fragment 34). Fragments 33 and 34 are ligated together with fragment 16 (from the expression vector pRK5). The ligation mixture is transformed into *E. coli* strain 294, the transformed culture plated on ampicillin media plates and resistant colonies are selected. Plasmid DNA is prepared from transformants and checked by restriction analysis for the presence of the correct fragment. The resulting plasmid is referred to as pRKCD4Cκ.

A plasmid encoding a fusion of the CD4 V-like domain to the human immunoglobulin Cγ2 region is constructed in a similar fashion, and is referred to as pRKCD4Cγ2. Both of these plasmids are transfected into 293 cells, myeloma cells or other competent cells in order to obtain cell lines expressing variant CD4 molecules as described above.

EXAMPLE 3

The gDCD4T secreted by the method of Example 1 was purified from cell culture fluid containing either 10% FBS (fetal bovine serum) or no added FBS. The conditioned cell culture fluid was first concentrated by ultrafiltration, then purified by immunoaffinity chromatography. The immunoaffinity column was produced by coupling murine monoclonal antibody 5B6 (whose epitope is on the HSV-1 gD portion of the gDCD4T molecule) to glyceryl coated controlled pore glass by the method of Roy et al., *J. Chromatography*, 303:225:228 (1984). The concentrated cell culture fluid is applied directly to the column and the contaminating proteins are washed away with neutral pH buffer. The column is then washed with neutral buffer containing tetramethylammonium chloride followed by neutral buffer containing Tween 80. The bound gDCD4T is eluted from the column with buffer at pH3 containing Tween 80 (0.1% w/v) and is neutralized immediately as it is eluted. The eluted neutralized gDCD4T is then concentrated by ultrafiltration and dialyzed/diafiltered to exchange the buffer for a physiological salt solution containing Tween 80 at approximately 0.1% w/v.

If the detergent is not present the gDCD4T forms aggregates as evidenced by the ability of centrifugation at approximately 10,000 × g for 2 minutes to remove the gDCD4T from the solution. Incubation of gDCD4T at 4° C. in 0.1M sodium acetate, 0.5M NaCl and 0.25M Tris at pH 7 together with BSA, Tween 80 or glycerol as candidate stabilizers showed that, in the absence of a stabilizer the gDCD4T gradually aggregated over the space of 12 days to the point where only about 60–70% of the protein was soluble. However, use of 0.1% w/v Tween 80 or 0.5 mg/ml BSA ensured that about 100% or 80%, respectively, of the gDCD4T remained soluble over this period. Surprisingly glycerol was ineffective as a stabilizer and produced results inferior even to the control—at 8 days about 80% of the gDCD4T was aggregated when stored in the presence of glycerol.

EXAMPLE 4

Plasmids were constructed to direct the expression of proteins containing differing lengths of the amino-terminal, extracellular domain of CD4 fused to the constant region of human immunoglobulin γ1. These plasmids are referred to as $pRKCD4_{2\gamma1}$, $pRKCD4_{e4\gamma1}$, $pRKCD4_{2\gamma1}$, $pRKCD4_{e2\gamma1}$, $pRKCD4_{1\gamma1}$, and $pRKCD4_{e1\gamma1}$.

Plasmid $pRKCD4_{4\gamma1}$ contains the portion of the CD4 gene from the initiation codon to the fusion site after the codon for serine residue 366 of the mature CD4 polypeptide, immediately followed by the sequence coding for the constant region of human immunoglobulin γ1, starting at the codon for serine residue 114 of mature human immunoglobulin γ1 (Kabat et al., Sequences of Proteins or Immunological Interest, National Institute of Health, Bethesda, Md. [1987]).

Plasmid $pRKCD4_{e4\gamma1}$ contains the portion of the CD4 gene from the initiation codon to the fusion site after the codon for lysine residue 360 of the mature CD4 polypeptide, immediately followed by the sequence coding for the constant region of human immunoglobulin γ1, starting at the codon for serine residue 114 of mature human immunoglobulin γ1 (Kabat et al., supra).

Plasmid $pRKCD4_{2\gamma1}$ contains the portion of the CD4 gene from the initiation codon to the fusion site after the codon for glutamine residue 180 of the mature CD4 polypeptide, immediately followed by the sequence coding for the constant region of human immunoglobulin γ1, starting at the codon for serine residue 114 of mature human immunoglobulin γ1 (Kabat et al.).

Plasmid $pRKCD4_{e2\gamma1}$ contains the portion of the CD4 gene from the initiation codon to the fusion site after the codon for leucine residue 177 of the mature CD4 polypeptide, immediately followed by the sequence coding for the constant region of human immunoglobulin γ1, starting at the codon for serine residue 114 of mature human immunoglobulin γ1 (Kabat et al., supra).

Plasmid $pRKCD4_{1\gamma1}$ contains the portion of the CD4 gene from the initiation codon to the fusion site after the codon for aspartic acid residue 105 of the mature CD4 polypeptide, immediately followed by the sequence coding for the constant region of human immunoglobulin γ1, starting at the codon for serine residue 114 of mature human immunoglobulin γ1 (Kabat et al., supra).

Plasmid $pRKCD4_{e1\gamma1}$ contains the portion of the CD4 gene from the initiation codon to the fusion site after the codon for leucine residue 100 of the mature CD4

polypeptide, immediately followed by the sequence coding for the constant region of human immunoglobulin γ1, starting at the codon for serine residue 114 of mature human immunoglobulin γ1 (Kabat et al., supra).

Construction of these plasmids required the prior construction of plasmid pRKCD4TP/γ1. It was constructed as follows:

A cDNA clone coding for human immunoglobulin γ1 was obtained from a human spleen cDNA library (Clontech Laboratories, Inc.) using oligonucleotides based on the published sequence (Ellison et al., "Nucl. Acids Res." 10:4071–4079 [1982]), and an EcoRI-EagI fragment (the EcoRI site was contributed by a linker; see FIGS. 4A, B) containing part of the variable and all of the constant region was obtained. This fragment was blunted with Klenow fragment, and recovered by gel electrophoresis (Fragment a1).

Plasmid pRKCD4TP-KK, encoding a substitutional variant of soluble CD4 (residues 1-368) containing a lysine residue instead of asparagine at position 1 of the mature polypeptide, was constructed from plasmid pRKCD4TP by site-directed mutagenesis. A synthetic oligonucleotide was made as a primer for a mutagenesis reaction to obtain the desired coding sequence. This was synthesized as a 51-mer which contained two silent mutations from the natural sequence in addition to the substitution mutation, and 21 bases on each side of the mutated codons:

5'-CCC TTT TTT GCC CAG CAC CAC CTT CTT GCC CTG AGT GGC TGC TGG GAG GAG -3' (SEQ ID NO: 24)

Plasmid pRKCD4TP was transformed into *E. coli* strain SR101 and the transformed colonies were plated on ampicillin media plates. Resistant colonies were selected and grown in the presence of m13K07 helper bacteriophage to yield secreted, encapsidated single-stranded templates of pRKCD4TP. The single-stranded plasmid DNA was isolated and used as the template for mutagenesis reactions with the synthetic oligonucleotides described above as primers. The mutagenesis reaction was transformed into *E. coli* SR101 and the transformed culture plated on ampicillin media plates. Transformants were screened by colony hybridization (ref. Grunstein-Hogness) for the presence of the appropriate sequence, using the following 16 mer as the probe.

5'-C CAC CTT CTT GCC CTG -3' (SEQ ID NO: 25) The hybridization conditions chosen were sufficiently stringent that the probe only detects the correctly fused product. Colonies identified as positive were selected and plasmid DNA was isolated and transformed into *E. coli* strain SR101. The transformed cultures were plated on ampicillin media plates, and resistant colonies were selected and grown in the presence of m13K07 bacteriophage. Templates were prepared as above and screened by sequencing.

Plasmid pRKCD4TP-KK was digested with XbaI and treated with Klenow Enzyme, and Fragment a2, containing the linearized plasmid, was recovered by gel electrophoresis, and ligated with fragment a1. The ligation mixture was transformed into *E. coli* strain 294, the transformed culture plated on ampicillin media plates and resistant colonies were selected. Plasmid DNA was prepared from the transformants and checked by restriction analysis for the presence of the correct fragment in the correct orientation (i.e., the immunoglobulin coding region in the same orientation as the CD4 coding region, and at the 3' end of the CD4 coding region). This plasmid is referred to as pRKCD4TP/γ1.

Synthetic oligonucleotides were made as primers for deletional mutagenesis reactions to fuse the appropriate coding sequences of IgG1 and CD4 as described above. These were synthesized as 48-mers comprising 24 nucleotides on each side of the desired fusion site (i.e., corresponding to the COOH-terminal 8 residues of the desired CD4 moiety, and the NH_2-terminal 8 residues of the desired immunoglobulin moiety). Plasmid pRKCD4TP/γ1 was transformed into *E. coli* strain SR101 and the transformed cultures plated on ampicillin media plates. Resistant colonies were selected and grown in the presence of m13K07 helper bacteriophage to yield secreted, encapsidated single-stranded templates of pRKCD4TP/γ1. The single-stranded plasmid DNA was isolated and used as the template for mutagenesis reactions with the synthetic oligonucleotides described above as primers. The mutagenesis reactions were transformed into *E. coli* SR101 and the transformed culture was plated on ampicillin media plates. Transformants were screened by colony hybridization (ref. Grunstein-Hogness) for the presence of the appropriate fusion site, using 16mers as probes. These 16mers comprise 8 bases on either side of the fusion site, and the hybridization conditions chosen were sufficiently stringent that the probes only detect the correctly fused product. Colonies identified as positive were selected and plasmid DNA was isolated and transformed into *E. coli* strain SR101.

The transformed cultures were plated on ampicillin media plates, and resistant colonies were selected and grown in the presence of m13KO7 bacteriophage. Templates were prepared as above and screened by sequencing.

The plasmids were transfected into 293 cells using standard procedures and assayed for expression and production as described above.

	Expressed	Secreted
$pRKCD4_{1\gamma1}$	+	–
$pRKCD4_{c2\gamma1}$	+	+
$pRKCD4_{2\gamma1}$	+	+
$pRKCD4_{c4\gamma1}$	+	+
$pRKCD4_{4\gamma1}$	+	+

Plasmids also were constructed to direct the expression of fusion proteins containing differing lengths of the amino-terminal, extracellular domain of CD4 fused to the truncated portion of the constant region of human immunoglobulin g1, comprising only the hinge region and constant domains CH_2 and CH_3.

Synthetic oligonucleotides were made as primers for mutagenesis reactions to delete the immunoglobulin sequence from Ser114 to Cys215 inclusive (Kabat et al., Supra). These were synthesized as 48-mers comprising 24 nucleotides on each side of the desired fusion site (i.e., corresponding to the COOH-terminal 8 residues of the desired CD4 moiety, and the NH_2-terminal 8 residues of the desired immunoglobulin moiety). Plasmids $pRKCD4_{4\gamma1}$, $pRKCD4_{2\gamma1}$ and $pRKCD4_{1\gamma1}$ were separately transformed into *E. coli* strain SR101 and the transformed culture was plated on ampicillin media plates. Resistant colonies were selected and grown in the presence of m13KO7 helper bacteriophage to yield secreted, encapsidated single-stranded templates of these plasmids. The single-stranded plasmid DNA was isolated and used as the template for mutagenesis reactions with the synthetic oligonucleotides described above as primers. The mutagenesis reactions were transformed into *E. coli* SR101 and the transformed culture was plated on ampicillin media plates. Transformants were screened by colony hybridization (Grunstein-Hogness) for the presence of the appropriate fusion site, using 16mers as probes. These 16mers comprise 8 bases on either side of the

fusion site, and the hybridization conditions chosen were sufficiently stringent that the probes only detect the correctly fused product. Colonies identified as positive were selected and plasmid DNA was isolated and transformed into *E. coli* strain SR101. The transformed cultures were plated on ampicillin media plates, and resistant colonies were selected and grown in the presence of m13KO7 bacteriophage. Templates were prepared as above and screened by sequencing.

The plasmid derived from plasmid pRKCD4$_{4\gamma1}$ is referred to as pRKCD4$_{4Fc1}$, that derived from plasmid pRKCD4$_{2\gamma1}$ is referred to as pRKCD4$_{2Fc1}$ and that derived from plasmid pRKCD4$_{1\gamma1}$ is referred to as pRKCD4$_{1Fc1}$.

pRKCD4$_{2Fc1}$, pRKCD4$_{2Fc1}$ and pRKCD4$_{4Fc1}$ are cultured in the same fashion as described above and CH1-deleted CD4 immunoadhesons recovered as described elsewhere herein.

Light Chain Fusions

Plasmids were constructed to direct the expression of proteins containing differing lengths of the amino terminal, extracellular domain of CD4 fused to the constant region of human immunoglobulin κ. These plasmids are referred to as pRKCD4$_{4\kappa}$, and pRKCD4$_{e4\kappa}$.

Plasmid pRKCD4$_{4\kappa}$ contains the portion of the CD4 gene from the initiation codon to the fusion site after the codon for serine residue 366 of the mature CD4 polypeptide, immediately followed by the sequence for the constant region of human immunoglobulin κ, starting at the codon for threonine residue 109 of the mature human immunoglobulin κ. (Kabat et al., supra)

Plasmid pRKCD4$_{e4\kappa}$ contains the portion of the CD4 gene from the initiation codon to the fusion site after the codon for lysine residue 360 of the mature CD4 polypeptide, immediately followed by the sequence for the constant region of human immunoglobulin κ, starting at the codon for threonine residue 109 of the mature human immunoglobulin κ. (Kabat et al., supra)

These plasmids were constructed in a manner analogous to plasmids pRKCD4$_{4\gamma1}$ and pRKCD4$_{e4\gamma1}$ described above, with the following exception:

The human immunoglobulin κ coding sequence (FIG. 5) was obtained from a human spleen cDNA library (Clontech Laboratories, Inc.) using oligonucleotides based on the published sequence (Hieter, P. A. et al., Cell 22:197–207 [1980]) and an EcoRI-BspMI fragment containing part of the variable region and the entire constant region was obtained (see FIG. 5). This fragment was blunted with Klenow fragment and the four dNTPs. This fragment was used instead of fragment a1, and was used to construct plasmid pRKCD4TP/hκ.

Expression in CHO Cells

Plasmids were or are constructed to direct the expression of the immunoadhesons described above in CHO cells. These are referred to as pSVeCD4$_{4\gamma1}$SVDHFR, pSVeCD4$_{2\gamma1}$SVDHFR, pSVeCD4$_{1\gamma1}$SVDHFR, pSVeCD4$_{e4\gamma1}$SVDHFR, pSVeCD4$_{e2\gamma1}$SVDHFR, pSVeCD4$_{e1\gamma1}$SVDHFR, pSVeCD4$_{4Fc1}$SVDHFR, pSVeCD4$_{2Fc1}$SVDHFR, pSVeCD4$_{1Fc1}$SVDHFR, pSVeCD4$_{4\kappa}$SVDHFR and pSVeCD4$_{2\kappa}$SVDHFR.

Fragment 31 was prepared as described above. Fragment 32a was prepared by digesting plasmid pE348HBV.E400 D22 with BamHI, blunting with Klenow fragment and the four dNTPs, then digesting with PvuI and isolating the large fragment containing the balance of the β-lactamase gene and the SV40 early promoter and the DHFR gene. Plasmids pRKCD4$_{4\gamma1}$, pRKCD4$_{2\gamma1}$, pRKCD4$_{1\gamma1}$, pRKCD4$_{e4\gamma1}$, pRKCD4$_{e2\gamma1}$, pRKCD4$_{e1\gamma1}$, pRKCD4$_{4Fc1}$, pRKCD4$_{2Fc1}$, pRKCD4$_{1Fc1}$, pRKCD4$_{4\kappa}$, and pRKCD4$_{e4\kappa}$ were separately digested with HindIII, blunted with Klenow fragment and the four dNTPs, then digested with EcoRI and the fragments encoding the CD4-Ig fusion protein were isolated. The resulting DNA fragments were ligated together with fragments 31 and 32a and transformed into *E. coli* strain 294. Colonies were selected and checked for the presence of the correct plasmid as above, then transfected into CHO cells and amplified by methotrexate selection using conventional procedures.

EXAMPLE 5

Culture, Purification and Formulation of CD4 Variants

Plasmids encoding soluble CD4 adhesons such as CD4T, CD4TP, or soluble CD4 immunoadhesons were calcium phosphate transfected into CHO-DP7 (a proinsulin-transformed autocrine host cell derived from CHO; U.S. Ser. No. 97,472, supra) and the transformants grown in selective medium (1:1 HAM F12/DMEM GHT-containing 1–10% diafiltered or dialyzed bovine serum). Other suitable host cells are CHO cells or 293S human embryonic kidney cells. The transformants were amplified by methotrexate selection in the same medium but containing 500 nm methotrexate. A subclone capable of secreting CD4TP, CD4tp 500 b, was selected. CD4tp 500 b is cultured in a DMEM/HAM F12 medium at about 37° C. until CD4TP accumulates in the culture, after which the medium is separated from the cells and insoluble matter by centrifuging.

Culture fluid from CD4TP transformants was concentrated and diafiltered to lower the ionic strength. The concentrate was passed through a large volume of Q-Sepharose anion exchange resin (previously equilibrated with mM NaCl, pH 8.5) in order to adsorb contaminants from the culture fluid. The isoelectric point of CD4TP is about 9.5, thus making it possible to discriminate between truncated forms of CD4 and most contaminants by alternate adsorption, respectively, on a cation exchange resin such as carboxymethyl or sulfonyl Sepharose, and an anion exchange resin such as quaternary ammonium Sepharose. In addition, since highly electropositive domains are present in the extracellular segment of CD4, any CD4-containing variant is purified in the same fashion as CD4TP. The unadsorbed culture fluid from the anion exchange resin step was then passed through a cation exchange resin (previously equilibrated with 25 mM NaCl at pH 8.5) whereby CD4TP was adsorbed to the resin. The CD4TP was eluted with a NaCl gradient at pH 8.5, this CD4 variant eluting at about 0.2 M NaCl. Ammonium sulfate was added to the eluate to a concentration of 1.7M and the solution passed through a column of hydrophobic interaction chromatography resin (phenyl or butyl Sepharose). The CD4TP was eluted from the hydrophobic interaction column with a gradient of ammonium sulfate, the CD4TP emerging at about 0.7M ammonium sulfate. The eluate was concentrated and buffer exchanged on a G-25 column using phosphate buffered saline containing 0.02% (w/v) Tween 20 or Tween 80. The CD4TP was soluble and stable in this solution, which was sterile filtered and filled into vials as an aqueous formulation. Other polymeric nonionic surfactants are suitably used with the CD4 formulations, including Pluronic block copolymers or polyethylene glycol.

It is also possible to employ immunoaffinity purification of soluble CD4 wherein the CD4 is adsorbed onto an immobilized antibody against CD4. This method suffers from the disadvantage that elution of the soluble CD4 under

acidic conditions leads to protein aggregation that is only thoroughly ameliorated at relatively higher levels of surfactant. The foregoing procedure permits the use of much lower quantities of surfactant, about from 0.01 to 0.10% (w/v) surfactant.

The procedure followed for the purification of CD4 fusions with immunoglobulin heavy chain was to concentrate recombinant supernatants by ultrafiltration and thereafter adsorb the fusion onto resin-immobilized Staphylococcal protein A. The fusion was eluted with 0.1M citrate buffer pH 3 with no salt or detergent. This preparation is buffered into Tris buffer at pH 7.5. The immunoglobulin fusions with CD4 V1–V4 optionally are further purified by the procedure described above for unfused CD4 variants. CD4 immunoglobulin fusions with CD4 V1–V2 also may be purified by the procedure above, except that it is not expected that the isoelectric point of this class of molecules will be as alkaline as that of species containing all four V regions of CD4.

EXAMPLE 6

The characteristics of several adheson variants were determined. As shown in Table IV the immunoadhesons $CD4_{4IgG1}$ and $CD4_{2IgG1}$ show improved plasma half-life in rabbits, coupled with high-affinity gp120 binding and an affinity for Fcγ receptor (determined with U937 cells) that is comparable to that of bulk human IgG1.

TABLE IV

	gp120 KD (nM)[100]	FcγR KD (nM)*	Plasma Half-life++ In Rabbits (Hrs.)
$CD4T^{517}$	2.3 ± 0.4	Not detected	0.25
$CD4_{4IgG1}$	1.2 ± 0.1	2.83 ± 0.25	6.4
$CD4_{2IgG1}$	1.4 ± 0.1	3.01 ± 0.68	40.6
human IgG1	ND**	3.52 ± 0.5	21 days*

*determined in humans

+determined by the method of Anderson et al., "J. Immunol." 125:2735–2741 (1980).

#determined by the method of Smith et al., "Science" 238:1704–07 (1987).

§residues 1–368 only

++The adheson variant was injected intravenously into rabbits and samples of blood were collected periodically and assayed for the presence of the adheson variant.

**Not done.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(iii) NUMBER OF SEQUENCES: 25

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 402 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
Met Asn Arg Gly Val Pro Phe Arg His Leu Leu Leu Val Leu Gln
 1               5                  10                  15

Leu Ala Leu Leu Pro Ala Ala Thr Gln Gly Asn Lys Val Val Leu
            20                  25                  30

Gly Lys Lys Gly Asp Thr Val Glu Leu Thr Cys Thr Ala Ser Gln
        35                  40                  45

Lys Lys Ser Ile Gln Phe His Trp Lys Asn Ser Asn Gln Ile Lys
    50                  55                  60

Ile Leu Gly Asn Gln Gly Ser Phe Leu Thr Lys Gly Pro Ser Lys
65                  70                  75

Leu Asn Asp Arg Ala Asp Ser Arg Arg Ser Leu Trp Asp Gln Gly
            80                  85                  90

Asn Phe Pro Leu Ile Ile Lys Asn Leu Lys Ile Glu Asp Ser Asp
        95                  100                 105

Thr Tyr Ile Cys Glu Val Glu Asp Gln Lys Glu Glu Val Gln Leu
    110                 115                 120

Leu Val Phe Gly Leu Thr Ala Asn Ser Asp Thr His Leu Leu Gln
125                 130                 135

Gly Gln Ser Leu Thr Leu Thr Leu Glu Ser Pro Pro Gly Ser Ser
            140                 145                 150
```

-continued

Pro Ser Val Gln Cys Arg Ser Pro Arg Gly Lys Asn Ile Gln Gly
155 160 165

Gly Lys Thr Leu Ser Val Ser Gln Leu Glu Leu Gln Asp Ser Gly
170 175 180

Thr Trp Thr Cys Thr Val Leu Gln Asn Gln Lys Lys Val Glu Phe
185 190 195

Lys Ile Asp Ile Val Val Leu Ala Phe Gln Lys Ala Ser Ser Ile
200 205 210

Val Tyr Lys Lys Glu Gly Glu Gln Val Glu Phe Ser Phe Pro Leu
215 220 225

Ala Phe Thr Val Glu Lys Leu Thr Gly Ser Gly Glu Leu Trp Trp
230 235 240

Gln Ala Glu Arg Ala Ser Ser Ser Lys Ser Trp Ile Thr Phe Asp
245 250 255

Leu Lys Asn Lys Glu Val Ser Val Lys Arg Val Thr Gln Asp Pro
260 265 270

Lys Leu Gln Met Gly Lys Lys Leu Pro Leu His Leu Thr Leu Pro
275 280 285

Gln Ala Leu Pro Gln Tyr Ala Gly Ser Gly Asn Leu Thr Leu Ala
290 295 300

Leu Glu Ala Lys Thr Gly Lys Leu His Gln Glu Val Asn Leu Val
305 310 315

Val Met Arg Ala Thr Gln Leu Gln Lys Asn Leu Thr Cys Glu Val
320 325 330

Trp Gly Pro Thr Ser Pro Lys Leu Met Leu Ser Leu Lys Leu Glu
335 340 345

Asn Lys Glu Ala Lys Val Ser Lys Arg Glu Lys Ala Val Trp Val
350 355 360

Leu Asn Pro Glu Ala Gly Met Trp Gln Cys Leu Leu Ser Asp Ser
365 370 375

Gly Gln Val Leu Leu Glu Ser Asn Ile Lys Val Leu Pro Thr Trp
380 385 390

Ser Thr Pro Ser Phe Asn Ala Val Val Tyr His Ser
395 400 402

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1416 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

AATTCAAGCC CAGAGCCCTG CCATTTCTGT GGGCTCAGGT CCCTACTGCT 50

CAGCCCCTTC CTCCCTCGGC AAGGCCACAA TGAACCGGGG AGTCCCTTTT 100

AGGCACTTGC TTCTGGTGCT GCAACTGGCG CTCCTCCCAG CAGCCACTCA 150

GGGAAACAAA GTGGTGCTGG GCAAAAAAGG GGATACAGTG GAACTGACCT 200

GTACAGCTTC CCAGAAGAAG AGCATACAAT TCCACTGGAA AAACTCCAAC 250

CAGATAAAGA TTCTGGGAAA TCAGGGCTCC TTCTTAACTA AAGGTCCATC 300

CAAGCTGAAT GATCGCGCTG ACTCAAGAAG AAGCCTTTGG GACCAAGGAA 350

ACTTTCCCCT GATCATCAAG AATCTTAAGA TAGAAGACTC AGATACTTAC 400

-continued

```
ATCTGTGAAG TGGAGGACCA GAAGGAGGAG GTGCAATTGC TAGTGTTCGG          450
ATTGACTGCC AACTCTGACA CCCACCTGCT TCAGGGGCAG AGCCTGACCC          500
TGACCTTGGA GAGCCCCCCT GGTAGTAGCC CCTCAGTGCA ATGTAGGAGT          550
CCAAGGGGTA AAAACATACA GGGGGGGAAG ACCCTCTCCG TGTCTCAGCT          600
GGAGCTCCAG GATAGTGGCA CCTGGACATG CACTGTCTTG CAGAACCAGA          650
AGAAGGTGGA GTTCAAAATA GACATCGTGG TGCTAGCTTT CCAGAAGGCC          700
TCCAGCATAG TCTATAAGAA AGAGGGGGAA CAGGTGGAGT TCTCCTTCCC          750
ACTCGCCTTT ACAGTTGAAA AGCTGACGGG CAGTGGCGAG CTGTGGTGGC          800
AGGCGGAGAG GGCTTCCTCC TCCAAGTCTT GGATCACCTT TGACCTGAAG          850
AACAAGGAAG TGTCTGTAAA ACGGGTTACC CAGGACCCTA AGCTCCAGAT          900
GGGCAAGAAG CTCCCGCTCC ACCTCACCCT GCCCCAGGCC TTGCCTCAGT          950
ATGCTGGCTC TGGAAACCTC ACCCTGGCCC TTGAAGCGAA AACAGGAAAG         1000
TTGCATCAGG AAGTGAACCT GGTGGTGATG AGAGCCACTC AGCTCCAGAA         1050
AAATTTGACC TGTGAGGTGT GGGGACCCAC CTCCCCTAAG CTGATGCTGA         1100
GTTTGAAACT GGAGAACAAG GAGGCAAAGG TCTCGAAGCG GGAGAAGGCG         1150
GTGTGGGTGC TGAACCCTGA GGCGGGGATG TGGCAGTGTC TGCTGAGTGA         1200
CTCGGGACAG GTCCTGCTGG AATCCAACAT CAAGGTTCTG CCCACATGGT         1250
CCACCCCGAG CTTTAATGCG GTAGTTTATC ACAGTTAAAT TGCTAACGCA         1300
GTCAGGCACC GTGTATGAAA TCTAACAATG CGCTCATCGT CATCCTCGGC         1350
ACCGTCACCC TGGATGCTGT AGGCATAGGC TTGGTTATGC CGGTACTGCC         1400
GGGCCTCTTG CGGGAT                                             1416
```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1416 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
TTAAGTTCGG GTCTCGGGAC GGTAAAGACA CCCGAGTCCA GGGATGACGA          50
GTCGGGGAAG GAGGGAGCCG TTCCGGTGTT ACTTGGCCCC TCAGGGAAAA         100
TCCGTGAACG AAGACCACGA CGTTGACCGC GAGGAGGGTC GTCGGTGAGT         150
CCCTTTGTTT CACCACGACC CGTTTTTTCC CCTATGTCAC CTTGACTGGA         200
CATGTCGAAG GGTCTTCTTC TCGTATGTTA AGGTGACCTT TTTGAGGTTG         250
GTCTATTTCT AAGACCCTTT AGTCCCGAGG AAGAATTGAT TTCCAGGTAG         300
GTTCGACTTA CTAGCGCGAC TGAGTTCTTC TTCGGAAACC CTGGTTCCTT         350
TGAAAGGGGA CTAGTAGTTC TTAGAATTCT ATCTTCTGAG TCTATGAATG         400
TAGACACTTC ACCTCCTGGT CTTCCTCCTC CACGTTAACG ATCACAAGCC         450
TAACTGACGG TTGAGACTGT GGGTGGACGA AGTCCCCGTC TCGGACTGGG         500
ACTGGAACCT CTCGGGGGGA CCATCATCGG GGAGTCACGT TACATCCTCA         550
GGTTCCCCAT TTTTGTATGT CCCCCCCTTC TGGGAGAGGC ACAGAGTCGA         600
CCTCGAGGTC CTATCACCGT GGACCTGTAC GTGACAGAAC GTCTTGGTCT         650
```

-continued

```
TCTTCCACCT CAAGTTTTAT CTGTAGCACC ACGATCGAAA GGTCTTCCGG          700

AGGTCGTATC AGATATTCTT TCTCCCCCTT GTCCACCTCA AGAGGAAGGG          750

TGAGCGGAAA TGTCAACTTT TCGACTGCCC GTCACCGCTC GACACCACCG          800

TCCGCCTCTC CCGAAGGAGG AGGTTCAGAA CCTAGTGGAA ACTGGACTTC          850

TTGTTCCTTC ACAGACATTT TGCCCAATGG GTCCTGGGAT TCGAGGTCTA          900

CCCGTTCTTC GAGGGCGAGG TGGAGTGGGA CGGGGTCCGG AACGGAGTCA          950

TACGACCGAG ACCTTTGGAG TGGGACCGGG AACTTCGCTT TTGTCCTTTC         1000

AACGTAGTCC TTCACTTGGA CCACCACTAC TCTCGGTGAG TCGAGGTCTT         1050

TTTAAACTGC ACACTCCACA CCCCTGGGTG GAGGGGATTC GACTACGACT         1100

CAAACTTTGA CCTCTTGTTC CTCCGTTTCC AGAGCTTCGC CCTCTTCCGC        1150

CACACCCACG ACTTGGGACT CCGCCCCTAC ACCGTCACAG ACGACTCACT        1200

GAGCCCTGTC CAGGACGACC TTAGGTTGTA GTTCCAAGAC GGGTGTACCA        1250

GGTGGGGCTC GAAATTACGC CATCAAATAG TGTCAATTTA ACGATTGCGT        1300

CAGTCCGTGG CACATACTTT AGATTGTTAC GCGAGTAGCA GTAGGAGCCG        1350

TGGCAGTGGG ACCTACGACA TCCGTATCCG AACCAATACG GCCATGACGG        1400

CCCGGAGAAC GCCCTA                                             1416
```

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 434 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```
Met Gly Gly Thr Ala Ala Arg Leu Gly Ala Val Ile Leu Phe Val
 1               5                  10                  15

Val Ile Val Gly Leu His Gly Val Arg Gly Lys Tyr Ala Leu Ala
            20                  25                  30

Asp Ala Ser Leu Lys Met Ala Asp Pro Asn Arg Phe Arg Gly Lys
        35                  40                  45

Asp Leu Pro Val Leu Asp Gln Leu Leu Glu Gln Gly Asn Lys Val
    50                  55                  60

Val Leu Gly Lys Lys Gly Asp Thr Val Glu Leu Thr Cys Thr Ala
65                  70                  75

Ser Gln Lys Lys Ser Ile Gln Phe His Trp Lys Asn Ser Asn Gln
                80                  85                  90

Ile Lys Ile Leu Gly Asn Gln Gly Ser Phe Leu Thr Lys Gly Pro
            95                  100                 105

Ser Lys Leu Asn Asp Arg Ala Asp Ser Arg Arg Ser Leu Trp Asp
        110                 115                 120

Gln Gly Asn Phe Pro Leu Ile Ile Lys Asn Leu Lys Ile Glu Asp
    125                 130                 135

Ser Asp Thr Tyr Ile Cys Glu Val Glu Asp Gln Lys Glu Glu Val
140                 145                 150

Gln Leu Leu Val Phe Gly Leu Thr Ala Asn Ser Asp Thr His Leu
155                 160                 165

Leu Gln Gly Gln Ser Leu Thr Leu Thr Leu Glu Ser Pro Pro Gly
            170                 175                 180

Ser Ser Pro Ser Val Gln Cys Arg Ser Pro Arg Gly Lys Asn Ile
```

-continued

```
                185                 190                 195

Gln Gly Gly Lys Thr Leu Ser Val Ser Gln Leu Glu Leu Gln Asp
                200                 205                 210

Ser Gly Thr Trp Thr Cys Thr Val Leu Gln Asn Gln Lys Lys Val
                215                 220                 225

Glu Phe Lys Ile Asp Ile Val Val Leu Ala Phe Gln Lys Ala Ser
                230                 235                 240

Ser Ile Val Tyr Lys Lys Glu Gly Glu Gln Val Glu Phe Ser Phe
                245                 250                 255

Pro Leu Ala Phe Thr Val Glu Lys Leu Thr Gly Ser Gly Glu Leu
                260                 265                 270

Trp Trp Gln Ala Glu Arg Ala Ser Ser Ser Lys Ser Trp Ile Thr
                275                 280                 285

Phe Asp Leu Lys Asn Lys Glu Val Ser Val Lys Arg Val Thr Gln
                290                 295                 300

Asp Pro Lys Leu Gln Met Gly Lys Lys Leu Pro Leu His Leu Thr
                305                 310                 315

Leu Pro Gln Ala Leu Pro Gln Tyr Ala Gly Ser Gly Asn Leu Thr
                320                 325                 330

Leu Ala Leu Glu Ala Lys Thr Gly Lys Leu His Gln Glu Val Asn
                335                 340                 345

Leu Val Val Met Arg Ala Thr Gln Leu Gln Lys Asn Leu Thr Cys
                350                 355                 360

Glu Val Trp Gly Pro Thr Ser Pro Lys Leu Met Leu Ser Leu Lys
                365                 370                 375

Leu Glu Asn Lys Glu Ala Lys Val Ser Lys Arg Glu Lys Ala Val
                380                 385                 390

Trp Val Leu Asn Pro Glu Ala Gly Met Trp Gln Cys Leu Leu Ser
                395                 400                 405

Asp Ser Gly Gln Val Leu Leu Glu Ser Asn Ile Lys Val Leu Pro
                410                 415                 420

Thr Trp Ser Thr Pro Ser Phe Asn Ala Val Val Tyr His Ser
                425                 430             434
```

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1508 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

```
AAGCTTCAGC GCGAACGACC AACTACCCCG ATCATCAGTT ATCCTTAAGG          50

TCTCTTTTGT GTGGTGCGTT CCGGTATGGG GGGGACTGCC GCCAGGTTGG         100

GGGCCGTGAT TTTGTTTGTC GTCATAGTGG GCCTCCATGG GGTCCGCGGC         150

AAATATGCCT TGGCGGATGC CTCTCTCAAG ATGGCCGACC CCAATCGATT         200

TCGCGGCAAA GACCTTCCGG TCCTGGACCA GCTGCTCGAG CAGGGAAACA         250

AAGTGGTGCT GGGCAAAAAA GGGGATACAG TGGAACTGAC CTGTACAGCT         300

TCCCAGAAGA AGAGCATACA ATTCCACTGG AAAAACTCCA ACCAGATAAA         350

GATTCTGGGA AATCAGGGCT CCTTCTTAAC TAAAGGTCCA TCCAAGCTGA         400

ATGATCGCGC TGACTCAAGA AGAAGCCTTT GGGACCAAGG AAACTTTCCC         450
```

-continued

```
CTGATCATCA AGAATCTTAA GATAGAAGAC TCAGATACTT ACATCTGTGA          500
AGTGGAGGAC CAGAAGGAGG AGGTGCAATT GCTAGTGTTC GGATTGACTG          550
CCAACTCTGA CACCCACCTG CTTCAGGGGC AGAGCCTGAC CCTGACCTTG          600
GAGAGCCCCC CTGGTAGTAG CCCCTCAGTG CAATGTAGGA GTCCAAGGGG          650
TAAAAACATA CAGGGGGGGA AGACCCTCTC CGTGTCTCAG CTGGAGCTCC          700
AGGATAGTGG CACCTGGACA TGCACTGTCT TGCAGAACCA GAAGAAGGTG          750
GAGTTCAAAA TAGACATCGT GGTGCTAGCT TTCCAGAAGG CCTCCAGCAT          800
AGTCTATAAG AAAGAGGGGG AACAGGTGGA GTTCTCCTTC CCACTCGCCT          850
TTACAGTTGA AAAGCTGACG GGCAGTGGCG AGCTGTGGTG GCAGGCGGAG          900
AGGGCTTCCT CCTCCAAGTC TTGGATCACC TTTGACCTGA AGAACAAGGA          950
AGTGTCTGTA AAACGGGTTA CCCAGGACCC TAAGCTCCAG ATGGGCAAGA         1000
AGCTCCCGCT CCACCTCACC CTGCCCCAGG CCTTGCCTCA GTATGCTGGC         1050
TCTGGAAACC TCACCCTGGC CCTTGAAGCG AAAACAGGAA AGTTGCATCA         1100
GGAAGTGAAC CTGGTGGTGA TGAGAGCCAC TCAGCTCCAG AAAAATTTGA         1150
CCTGTCAGGT GTGGGGACCC ACCTCCCCTA AGCTGATGCT GAGTTTGAAA         1200
CTGGAGAACA AGGAGGCAAA GGTCTCGAAG CGGGAGAAGG CGGTGTGGGT         1250
GCTGAACCCT GAGGCGGGGA TGTGGCAGTG TCTGCTGAGT GACTCGGGAC         1300
AGGTCCTGCT GGAATCCAAC ATCAAGGTTC TGCCCACATG GTCCACCCCG         1350
AGCTTTAATG CGGTAGTTTA TCACAGTTAA ATTGCTAACG CAGTCAGGCA         1400
CCGTGTATGA AATCTAACAA TGCGCTCATC GTCATCCTCG GCACCGTCAC         1450
CCTGGATGCT GTAGGCATAG GCTTGGTTAT GCCGGTACTG CCGGGCCTCT         1500
TGCGGGAT                                                      1508
```

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1508 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

```
TTCGAAGTCG CGCTTGCTGG TTGATGGGGC TAGTAGTCAA TAGGAATTCC           50
AGAGAAAACA CACCACGCAA GGCCATACCC CCCCTGACGG CGGTCCAACC          100
CCCGGCACTA AAACAAACAG CAGTATCACC CGGAGGTACC CCAGGCGCCG          150
TTTATACGGA ACCGCCTACG GAGAGAGTTC TACCGGCTGG GGTTAGCTAA          200
AGCGCCGTTT CTGGAAGGCC AGGACCTGGT CGACGAGCTC GTCCCTTTGT          250
TTCACCACGA CCCGTTTTTT CCCCTATGTC ACCTTGACTG GACATGTCGA          300
AGGGTCTTCT TCTCGTATGT TAAGGTGACC TTTTTGAGGT TGGTCTATTT          350
CTAAGACCCT TTAGTCCCGA GGAAGAATTG ATTTCCAGGT AGGTTCGACT          400
TACTAGCGCG ACTGAGTTCT TCTTCGGAAA CCCTGGTTCC TTTGAAAGGG          450
GACTAGTAGT TCTTAGAATT CTATCTTCTG AGTCTATGAA TGTAGACACT          500
TCACCTCCTG GTCTTCCTCC TCCACGTTAA CGATCACAAG CCTAACTGAC          550
GGTTGAGACT GTGGGTGGAC GAAGTCCCCG TCTCGGACTG GGACTGGAAC          600
```

-continued

```
CTCTCGGGGG GACCATCATC GGGGAGTCAC GTTACATCCT CAGGTTCCCC          650

ATTTTTGTAT GTCCCCCCCT TCTGGGAGAG GCACAGAGTC GACCTCGAGG          700

TCCTATCACC GTGGACCTGT ACGTGACAGA ACGTCTTGGT CTTCTTCCAC          750

CTCAAGTTTT ATCTGTAGCA CCACGATCGA AAGGTCTTCC GGAGGTCGTA          800

TCAGATATTC TTTCTCCCCC TTGTCCACCT CAAGAGGAAG GGTGAGCGGA          850

AATGTCAACT TTTCGACTGC CCGTCACCGC TCGACACCAC CGTCCGCCTC          900

TCCCGAAGGA GGAGGTTCAG AACCTAGTGG AAACTGGACT TCTTGTTCCT          950

TCACAGACAT TTTGCCCAAT GGGTCCTGGG ATTCGAGGTC TACCCGTTCT         1000

TCGAGGGCGA GGTGGAGTGG GACGGGGTCC GGAACGGAGT CATACGACCG         1050

AGACCTTTGG AGTGGGACCG GGAACTTCGC TTTTGTCCTT TCAACGTAGT         1100

CCTTCACTTG GACCACCACT ACTCTCGGTG AGTCGAGGTC TTTTTAAACT         1150

GGACACTCCA CACCCCTGGG TGGAGGGGAT TCGACTACGA CTCAAACTTT         1200

GACCTCTTGT TCCTCCGTTT CCAGAGCTTC GCCCTCTTCC GCCACACCCA         1250

CGACTTGGGA CTCCGCCCCT ACACCGTCAC AGACGACTCA CTGAGCCCTG         1300

TCCAGGACGA CCTTAGGTTG TAGTTCCAAG ACGGGTGTAC CAGGTGGGGC         1350

TCGAAATTAC GCCATCAAAT AGTGTCAATT TAACGATTGC GTCAGTCCGT         1400

GGCACATACT TTAGATTGTT ACGCGAGTAG CAGTAGGAGC CGTGGCAGTG         1450

GGACCTACGA CATCCGTATC CGAACCAATA CGGCCATGAC GGCCCGGAGA         1500

ACGCCCTA                                                      1508
```

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 371 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

```
Val Thr Ala Ala Asp Thr Ala Val Tyr Tyr Cys Ala Arg Ala Thr
 1               5                  10                  15

Phe Cys Leu Trp Tyr Arg Glu Arg Pro Pro Cys Trp Ile Asp Pro
            20                  25                  30

Trp Gly Leu Gly Thr Leu Val Thr Val Ser Ser Ala Ser Thr Lys
        35                  40                  45

Gly Pro Ser Val Phe Pro Leu Ala Pro Ser Ser Lys Ser Thr Ser
    50                  55                  60

Gly Gly Thr Ala Ala Leu Gly Cys Leu Val Lys Asp Tyr Phe Pro
65                  70                  75

Glu Pro Val Thr Val Ser Trp Asn Ser Gly Ala Leu Thr Ser Gly
            80                  85                  90

Val His Thr Phe Pro Ala Val Leu Gln Ser Ser Gly Leu Tyr Ser
        95                  100                 105

Leu Ser Ser Val Val Thr Val Pro Ser Ser Ser Leu Gly Thr Gln
    110                 115                 120

Thr Tyr Ile Cys Asn Val Asn His Lys Pro Ser Asn Thr Lys Val
125                 130                 135

Asp Lys Lys Val Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys
            140                 145                 150

Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe
```

-continued

```
                155                 160                 165

Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr
                170                 175                 180

Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro
                185                 190                 195

Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn
                200                 205                 210

Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg
                215                 220                 225

Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly
                230                 235                 240

Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro
                245                 250                 255

Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro
                260                 265                 270

Gln Val Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn
                275                 280                 285

Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp
                290                 295                 300

Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr
                305                 310                 315

Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu
                320                 325                 330

Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn
                335                 340                 345

Val Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr
                350                 355                 360

Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
                365                 370 371
```

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1135 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

```
GAATTCTGTC ACTGCCGCGG ACACGGCCGT ATATTACTGT GCGAGAGCCA          50

CCTTTTGCCT ATGGTACAGG GAGCGTCCCC CTTGTTGGAT CGACCCCTGG         100

GGCCTGGGAA CCCTGGTCAC CGTCTCCTCG GCCTCCACCA AGGGCCCATC         150

GGTCTTCCCC CTGGCACCCT CCTCCAAGAG CACCTCTGGG GGCACAGCGG         200

CCCTGGGCTG CCTGGTCAAG GACTACTTCC CCGAACCGGT GACGGTGTCG         250

TGGAACTCAG GCGCCCTGAC CAGCGGCGTG CACACCTTCC CGGCTGTCCT         300

ACAGTCCTCA GGACTCTACT CCCTCAGCAG CGTGGTGACC GTGCCCTCCA         350

GCAGCTTGGG CACCCAGACC TACATCTGCA ACGTGAATCA CAAGCCCAGC         400

AACACCAAGG TGGACAAGAA AGTTGAGCCC AAATCTTGTG ACAAAACTCA         450

CACATGCCCA CCGTGCCCAG CACCTGAACT CCTGGGGGGA CCGTCAGTCT         500

TCCTCTTCCC CCCAAAACCC AAGGACACCC TCATGATCTC CCGGACCCCT         550

GAGGTCACAT GCGTGGTGGT GGACGTGAGC CACGAAGACC CTGAGGTCAA         600
```

-continued

```
GTTCAAGTGG TACGTGGACG GCGTGGAGGT GCATAATGCC AAGACAAAGC          650

CGCGGGAGGA GCAGTACAAC AGCACGTACC GGGTGGTCAG CGTCCTCACC          700

GTCCTGCACC AGGACTGGCT GAATGGCAAG GAGTACAAGT GCAAGGTCTC          750

CAACAAAGCC CTCCCAGCCC CCATCGAGAA AACCATCTCC AAAGCCAAAG          800

GGCAGCCCCG AGAACCACAG GTGTACACCC TGCCCCCATC CCGGGATGAG          850

CTGACCAAGA ACCAGGTCAG CCTGACCTGC CTGGTCAAAG GCTTCTATCC          900

CAGCGACATC GCCGTGGAGT GGGAGAGCAA TGGGCAGCCG GAGAACAACT          950

ACAAGACCAC GCCTCCCGTG CTGGACTCCG ACGGCTCCTT CTTCCTCTAC         1000

AGCAAGCTCA CCGTGGACAA GAGCAGGTGG CAGCAGGGGA ACGTCTTCTC         1050

ATGCTCCGTG ATGCATGAGG CTCTGCACAA CCACTACACG CAGAAGAGCC         1100

TCTCCCTGTC TCCGGGTAAA TGAGTGCGAC GGCCG                         1135
```

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1142 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

```
CTTAAGACAG TGACGGCGCC TGTGCCGGCA TATAATGACA CGCTCTCGGT           50

GGAAAACGGA TACCATGTCC CTCGCAGGGG GAACAACCTA GCTGGGGACC          100

CCGGACCCTT GGGACCAGTG GCAGAGGAGC CGGAGGTGGT TCCCGGGTAG          150

CCAGAAGGGG GACCGTGGGA GGAGGTTCTC GTGGAGACCC CCGTGTCGCC          200

GGGACCCGAC GGACCAGTTC CTGATGAAGG GGCTTGGCCA CTGCCACAGC          250

ACCTTGAGTC CGCGGGACTG GTCGCCGCAC GTGTGGAAGG GCCGACAGGA          300

TGTCAGGAGT CCTGAGATGA GGGAGTCGTC GCACCACTGG CACGGGAGGT          350

CGTCGAACCC GTGGGTCTGG ATGTAGACGT TGCACTTAGT GTTCGGGTCG          400

TTGTGGTTCC ACCTGTTCTT TCAACTCGGG TTTAGAACAC TGTTTTGAGT          450

GTGTACGGGT GGCACGGGTC GTGGACTTGA GGACCCCCCT GGCAGTCAGA          500

AGGACAAGGG GCGTTTTGGG TTCCTGTGGG AGTACTAGAG GGCCTGGGGA          550

CTCCAGTGTA CGCACCACCA CCTGCACTCG GTGCTTCTGG GACTCCAGTT          600

CAAGTTGACC ATGCACCTGC CGCACCTCCA CGTATTACGG TTCTGTTTCG          650

GCGCCCTCCT CGTCATGTTG TCGTGCATGG CCCACCAGTC GCAGGAGTGG          700

CAGGACGTGG TCCTGACCGA CTTACCGTTC CTCATGTTCA CGTTCCAGAG          750

GTTGTTTCGG GAGGGTCGGG GGTAGCTCTT TTGGTAGAGG TTTCGGTTTC          800

CCGTCGGGGC TCTTGGTGTC CACATGTGGG ACGGGGGTAG GGCCCTACTC          850

GACTGGTTCT TGGTCCAGTC GGACTGGACG GACCAGTTTC CGAAGATAGG          900

GTCGCTGTAG CGGCACCTCA CCCTCTCGTT ACCCGTCGGC CTCTTGTTGA          950

TGTTCTGGTG CGGAGGGCAC GACCTGAGGC TGCCGAGGAA GAAGGAGATG         1000

TCGTTCGAGT GGCACCTGTT CTCGTCCACC GTCGTCCCCT TGCAGAAGAG         1050

TACGAGGCAC TACGTACTAC GTACTCCGAG ACGTGTTGGT GATGTGCGTC         1100

TTCTCGGAGA GGGACAGAGG CCCATTTACT CACGCTGCCG GC                 1142
```

-continued

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 143 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

```
Thr Leu Thr Ile Ser Gly Leu Gln Pro Glu Asp Phe Ala Thr Tyr
 1               5                  10                  15

Tyr Cys Gln Gln Tyr Lys Ser Leu Ser Leu Thr Phe Gly Gly Gly
            20                  25                  30

Thr Lys Val Glu Ile Lys Arg Thr Val Ala Ala Pro Ser Val Phe
        35                  40                  45

Ile Phe Pro Pro Ser Asp Glu Gln Leu Lys Ser Gly Thr Ala Ser
    50                  55                  60

Val Val Cys Leu Leu Asn Asn Phe Tyr Pro Arg Glu Ala Leu Val
65                  70                  75

Gln Trp Leu Val Asp Asn Ala Leu Gln Ser Gly Asn Ser Gln Glu
            80                  85                  90

Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr Tyr Ser Leu Ser
        95                  100                 105

Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys His Lys Val
    110                 115                 120

Tyr Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser Pro Val Thr
125                 130                 135

Lys Ser Phe Asn Arg Gly Glu Cys
    140             143
```

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 468 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

```
GAATTCACTC TCACCATCAG CGGCCTGCAG CCTGAAGATT TTGCAACTTA              50

TTACTGCCAA CAGTATAAGA GTTTGTCGCT CACTTTCGGC GGAGGGACCA             100

AGGTGGAGAT CAAACGAACT GTGGCTGCAC CATCTGTCTT CATCTTCCCG             150

CCATCTGATG AGCAGTTGAA ATCTGGAACT GCCTCTGTTG TGTGCCTGCT             200

GAATAACTTC TATCCCAGAG AGGCCAAAGT ACAGTGGAAG GTGGATAACG             250

CCCTCCAATC GGGTAACTCC CAGGAGAGTG TCACAGAGCA GGACAGCAAG             300

GACAGCACCT ACAGCCTCAG CAGCACCCTG ACGCTGAGCA AAGCAGACTA             350

CGAGAAACAC AAAGTCTACG CCTGCGAAGT CACCCATCAG GGCCTGAGCT             400

CGCCCGTCAC AAAGAGCTTC AACAGGGGAG AGTGTTAGAG GGAGAAGTGC             450

CCCCACCTGC TCCTCAGT                                                468
```

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 468 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

-continued

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

CTTAAGTGAG AGTGGTAGTC GCCGGACGTC GGACTTCTAA AACGTTGAAT 50

AATGACGGTT GTCATATTCT CAAACAGCGA GTGAAAGCCG CCTCCCTGGT 100

TCCACCTCTA GTTTGCTTGA CACCGACGTG GTAGACAGAA GTAGAAGGGC 150

GGTAGACTAC TCGTCAACTT TAGACCTTGA CGGAGACAAC ACACGGACGA 200

CTTATTGAAG ATAGGGTCTC TCCGGTTTCA TGTCACCTTC CACCTATTGC 250

GGGAGGTTAG CCCATTGAGG GTCCTCTCAC AGTGTCTCGT CCTGTCGTTC 300

CTGTCGTGGA TGTCGGAGTC GTCGTGGGAC TGCGACTCGT TTCGTCTGAT 350

GCTCTTTGTG TTTCAGATGC GGACGCTTCA GTGGGTAGTC CCGGACTCGA 400

CCGGGCAGTG TTTCTCGAAG TTGTCCCCTC TCACAATCTC CCTCTTCACG 450

GGGGTGGACG AGGAGTCA 468

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 44 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

AATTCAAGCC CAGAGCCCTG CCATTTCTGT GGGCTCAGGT CCCT 44

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 50 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

ACTGCTCAGC CCCTTCCTCC CTCGGCAAGG CCACAATGAA CCGGGGAGTC 50

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 47 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

CCTTTTAGGC ACTTGCTTCT GGTGCTGCAA CTGGCGCTCC TCCCAGC 47

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 61 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

AGCCACTCAG GGAAACAAAG TGGTGCTGGG CAAAAAAGGG GATACAGTGG 50

AACTGACCTG T 61

(2) INFORMATION FOR SEQ ID NO:17:

-continued

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 51 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

ACAGGTCAGT TCCACTGTAT CCCCTTTTTT GCCCAGCACC ACTTTGTTTC 50

C 51

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 47 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

CTGAGTGGCT GCTGGGAGGA GCGCCAGTTG CAGCACCAGA AGCAAGT 47

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 50 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

GCCTAAAAGG GACTCCCCGG TTCATTGTGG CCTTGCCGAG GGAGGAAGGG 50

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 50 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

GCTGAGCAGT AGGGACCTGA GCCCACAGAA ATGGCAGGGC TCTGGGCTTG 50

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 59 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

CTGCTCGAGC AGGGAAACAA AGTGGTGCTG GGCAAAAAAG GGGATACAGT 50

GGAACTGAC 59

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 59 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

ACAGGTCAGT TCCACTGTAT CCCCTTTTTT GCCCAGCACC ACTTTGTTTC 50

-continued

CCTGCTCGA 59

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

CGTGATAGAA GCTTTCTAGA C 21

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 51 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

CCCTTTTTTG CCCAGCACCA CCTTCTTGCC CTGAGTGGCT GCTGGGAGGA 50

G 51

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

CCACCTTCTT GCCCTG 16

We claim:

1. Nucleic acid encoding an immunoadhesn in which a polypeptide comprising an adhesn variable region is fused at its C-terminus to the N-terminus of a polypeptide comprising a constant region of an immunoglobulin.

2. The nucleic acid according to claim 1 wherein the immunoglobulin constant region is selected from the group consisting of IgG-1, IgG-2, IgG-3, IgG-4, IgA, IgE, IgD, and IgM.

3. The nucleic acid according to claim 2 wherein the immunoglobulin constant region is selected from the group consisting of IgG-1 and IgG-3.

4. The nucleic acid according to claim 2 wherein the immunoglobulin constant region is selected from the group consisting of IgA and IgM.

5. The nucleic acid of claim 1 wherein the immunoglobulin constant region is the constant region of the immunoglobulin light chain.

6. The nucleic acid of claim 1 wherein the immunoglobulin constant region is the constant region of an immunoglobulin heavy chain.

7. The nucleic acid of claim 1 wherein the adhesn variable region is the variable region of an adhesn selected from the group consisting of CD1, CD2, CD4, CD8, CD28, the γ, δ, or ε chains of CD3, an intercellular adhesion molecule, LFA-3 and the high affinity IgE receptor.

8. The nucleic acid of claim 7 wherein the adhesn variable region is a variable region of human CD4.

9. The nucleic acid of claim 8 wherein the human CD4 variable region comprises the V1 and V2 domains of human CD4.

10. The nucleic acid of claim 9 wherein the human CD4 variable region comprises the V1, V2, V3 and V4 domains of human CD4.

11. The nucleic acid of claim 9 wherein the human CD4 variable region comprises the sequence Asn-Lys-Val-Val-Leu-Gly-Lys-Lys.

12. The nucleic acid of claim 8 wherein the human CD4 variable region comprises the amino acid sequence extending from residues 24 to 117 of the precursor of human CD4 in which the initiating Met is 1.

13. The nucleic acid of claim 8 further comprising a nucleic acid encoding a signal sequence heterologous to the naturally occurring human CD4 gene.

14. The nucleic acid of claim 1 further comprising a nucleic acid encoding a signal sequence heterologous to the signal sequence naturally associated with the adhesn variable region.

15. A recombinant host cell comprising the nucleic acid of claim 1.

16. The recombinant host cell of claim 15 further comprising a nucleic acid encoding a companion immunoglobulin heavy chain-light chain pair bearing a $V_L V_H$ antibody combining site capable of binding to a predetermined antigen.

17. The host cell according to claim 16 wherein the $V_L V_H$ antibody combining site is from an anti-ricin antibody.

18. The host cell of claim 16 which is a mammalian cell.

19. A method of preparing an immunoadheson comprising transfecting a host cell with the nucleic acid of claim 1 and thereafter culturing the host under conditions suitable for the expression of the immunoadheson.

20. The method of claim 18 wherein the immunoadheson is recovered from the host cell culture.

21. A replicable vector comprising the nucleic acid of claim 1.

22. The nucleic acid of claim 1 wherein the adheson variable region is fused at its C-terminus to the N-terminus of the constant region of the immunoglobulin.

23. The nucleic acid of claim 22 wherein the immunoglobulin constant region is the constant region of an immunoglobulin heavy chain.

24. The nucleic acid of claim 23 wherein the immunoglobulin constant region is the constant region of an IgG.

25. The nucleic acid of claim 23 wherein the adheson variable region is the variable region of an adheson selected from the group consisting of CD1, CD2, CD4, CD8, CD28, the γ, δ, or ε chains of CD3, an intercellular adhesion molecule, LFA-3 and the high affinity IgE receptor.

26. The nucleic acid of claim 25 wherein the adheson is LFA-3.

27. The nucleic acid of claim 22 further comprising a nucleic acid encoding a signal sequence heterologous to the signal sequence naturally associated with the adheson variable region.

28. A recombinant host cell comprising the nucleic acid of claim 22.

29. The host cell of claim 28 which is a mammalian cell.

30. A method of preparing an immunoadheson comprising transfecting a host cell with the nucleic acid of claim 22 and thereafter culturing the host under conditions suitable for the expression of the immunoadheson.

* * * * *

EXHIBIT 16

10/07



US006710169B2

(12) **United States Patent**
Capon et al.

(10) **Patent No.: US 6,710,169 B2**
(45) **Date of Patent: *Mar. 23, 2004**

(54) **ADHESON VARIANTS**

(75) Inventors: **Daniel J. Capon**, San Mateo, CA (US); **Timothy J. Gregory**, Hillsborough, CA (US)

(73) Assignee: **Genentech, Inc.**, South San Francisco, CA (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

This patent is subject to a terminal disclaimer.

(21) Appl. No.: **10/157,408**

(22) Filed: **May 28, 2002**

(65) **Prior Publication Data**

US 2003/0104535 A1 Jun. 5, 2003

Related U.S. Application Data

(63) Continuation of application No. 09/641,554, filed on Aug. 17, 2000, now abandoned, which is a continuation of application No. 09/275,310, filed on Mar. 24, 1999, now abandoned, which is a continuation of application No. 08/457,918, filed on Jun. 1, 1995, now Pat. No. 6,117,655, which is a continuation of application No. 08/236,311, filed on May 2, 1994, now Pat. No. 5,565,335, which is a continuation of application No. 07/936,190, filed on Aug. 26, 1992, now Pat. No. 5,336,603, which is a division of application No. 07/842,777, filed on Feb. 18, 1992, now abandoned, which is a continuation of application No. 07/250,785, filed on Sep. 28, 1988, now abandoned, which is a continuation-in-part of application No. 07/104,329, filed on Oct. 2, 1987, now abandoned.

(51) **Int. Cl.**[7] **C07K 14/705**; C07K 14/73; C12N 15/62

(52) **U.S. Cl.** **530/387.3**; 530/350; 536/23.4; 435/69.7

(58) **Field of Search** 530/387.1, 387.3, 530/350; 536/23.4; 435/69.7

(56) **References Cited**

U.S. PATENT DOCUMENTS

3,959,080	A	5/1976	Orth et al.
4,002,531	A	1/1977	Royer
4,055,635	A	10/1977	Green et al.
4,301,144	A	11/1981	Iwashita et al.
4,412,989	A	11/1983	Iwashita et al.
4,444,878	A	4/1984	Paulus et al.
4,745,055	A	5/1988	Schenk
4,761,371	A	8/1988	Bell et al.
4,766,106	A	8/1988	Katre et al.
4,816,567	A	3/1989	Cabilly et al.
4,879,211	A	11/1989	Wang et al.
5,336,603	A	8/1994	Capon et al.

FOREIGN PATENT DOCUMENTS

EP	068763	1/1983
EP	088695	9/1983
EP	120694	10/1984
EP	139416	5/1985
EP	173494	3/1986
EP	244221	11/1987
EP	255694	2/1988
EP	256654	2/1988
EP	266663	5/1988
EP	278776	8/1988
EP	296786	12/1988
EP	313377	4/1989
EP	319815	6/1989
EP	325262	7/1989
EP	330227	8/1989
EP	331356	9/1989
WO	WO 85/03947	9/1985
WO	WO 87/03600	6/1987
WO	WO 88/01304	2/1988
WO	WO 88/03559	5/1988
WO	WO 88/07376	10/1988
WO	WO 88/09344	12/1988
WO	WO 89/01940	3/1989
WO	WO 89/03221	4/1989

OTHER PUBLICATIONS

Allaway et al., "Expression and Characterization of CD4–IgG$_2$, a Novel Heterotetramer That Neutralizes Primary HIV Type 1 Isolates" *Aids Research and Retroviruses* 11 (5) :533–539 (1995).

Amster et al., "Synthesis of part of a mouse immunoglobulin light chain in a bacterial clone" *Nucl. Acids Res.* 8(10) :2055–2065 (1980).

Anderson et al., "Regulatory Interactions Between Members of the Immunoglobulin Superfamily" *Immunology Today* 9 (7 & 8) :199–203 (1988).

Aruffo et al., "CD44 is the Principal Cell Surface Receptor for Hyaluronate" *Cell* 61:1303–1313 (1990).

Beauchamp et al., "A new procedure for the synthesis of polyethylene glycol–protein adducts; effects on function, receptor recognition, and clearance of superoxide dismutase, lactoferrin, and α_2–macroglobulin" *Analytical Biochemistry* 131(1) :25–33 (1983).

Blank et al., "Complete structure and expression in transfected cells of high affinity IgE receptor" *Nature* 337(6203) :187–189 (1989).

Boulianne et al., "Production of functional chimaeric mouse/human antibody" *Nature* 312:643–646 (Dec. 13, 1984).

Byrn et al., "Biological Properties of a CD4 Immunoadhesin" *Nature* 344:667–670 (Apr. 12, 1990).

(List continued on next page.)

Primary Examiner—Prema Mertz
(74) *Attorney, Agent, or Firm*—Jeffrey S. Kubinec

(57) **ABSTRACT**

Novel derivatives of cell surface proteins which are homologous to the immunoglobulin superfamily (adhesons) are provided. Amino acid sequence variations are introduced into adheson, the most noteworthy of which are those in which the transmembrane and, preferably, cytoplasmic domains are rendered functionally inactive, and in which adheson extracellular domains replace an immunoglobulin variable region. These variants are useful in therapy or diagnostics, in particular, CD4 variants are therapeutically useful in the treatment of HIV infections.

19 Claims, 13 Drawing Sheets

OTHER PUBLICATIONS

Capon et al., "Designing CD4 Immunoadhesins for AIDS Therapy" *Nature* 337:525–531 (1989).

Chamow et al., "A Humanized Bispecific Immunoadhesin–Antibody That Retargets CD3+ Effectors to Kill HIV–1–Infected Cells" *Journal of Immunology* 153:4268–4280 (1994).

Chaudhary et al., "Activity of a recombinant fusion protein between transforming growth factor type alpha and Pseudomonas toxin" *Proc. Natl. Acad. Sci. USA* 84:4538–4542 (1987).

Clark et al., "Peptide and nucleotide sequences for rat CD4 (W3/25) antigen: Evidence for derivation from a structure with four immunoglobulin–related domains" *Proc. Natl. Acad. Sci.* 84:1649–1653 (1987).

Clerici et al., "Effect of a Recombinant CD4–IgG on In Vitro T Helper Cell Function: Data From a Phase I/II Study of Patients with AIDS" *The Journal of Infectious Diseases* 168:1012–1016 (1993).

Estess et al., "Analysis of T–cell receptor structure and function using chimeric T–cell receptor/immunoglobulin molecules" *Journal of Cellular Biochemistry Supplement* 11 (Part D, Abs. 331) :258 (1987).

Falkner and Zachau, "Expression of mouse immunoglobulin genes in monkey cells" *Nature* 298:286–288 (1982).

Gascoigne et al., "Secretion of a Chimeric T–cell Receptor–immunoglobulin Protein" *Proc. Natl. Acad. Sci. USA* 84:2936–2940 (1987).

Gascoigne et al., "Secretion of chimeric T cell receptor–immunoglobulin fusion proteins" *Journal of Cellular Biochemistry Supplement* 11 (Part D, Abs. 333):259 (1987).

Goverman et al., "Chimeric T–cell receptor genes as tools in analyzing T–cell/target–cell interactions" *Journal of Cellular Biochemistry Supplement* 11 (Part D, Abs. 334) :259 (1987).

Haak–Frendscho et al., "Human IgE Receptor α chain–IgG Chimera Blocks Passive Cutaneous Anaphylaxis Reaction In Vivo" *J. Immunol.* 151(1) :351–358 (1993).

Hakimi et al., "The α subunit of the human IgE receptor (FcERI) is sufficient for high affinity IgE binding" *Journal of Biological Chemistry* 265(36):22079–22081 (1990).

Hashimoto et al., "Rearrangement and expression of T cell receptor genes in pre–T cells" *Journal of Cellular Biochemistry Supplement* 11 (Part D, Abs. 434):278 (1987).

"Human Immunodeficiency Virus Infection" *The Merck Manual of Diagnosis and Therapy*, Mark H. Beers and Robert Berkow, M.D., 17th edition, Whitehouse Station, N.J.:Merck Research Laboratories, Chapter 163, pp. 1312–1323 (1999).

Ivars et al., "Expression of cloned T cell receptor genes" *Journal of Cellular Biochemistry Supplement* 11 (Part D, Abs. 435) :278 (1987).

Johnson et al., "Correlation of T cell receptor structure and function" *Journal of Cellular Biochemistry Supplement* 11 (Part D, Abs. 337):260 (1987).

Kahn et al., "A Phase 1 Study of Recombinant Human CD4 Immunoglobulin G (rCD4–IgG) In Patients With HIV–Associated Immune Thrombocytopenic Purpura", (W.B. 2156) pp. 221 (1991).

Kohler, G., "Immunoglobulin chain loss in hybridoma lines" *Proc. Natl. Acad. Sci. USA* 77(4):2197–2199 (1980).

Linsley et al., "Binding of the B Cell Activation Antigen B7 to CD28 Costimulates T Cell Proliferation and Interleukin 2 mRNA Accumulation" *Journal of Experimental Medicine* 173:721–730 (1991).

Linsley et al., "CTLA–4 is a second receptor for the B cell activation antigen B7" *Journal of Experimental Medicine* 174:561–569 (1991).

Littman DR, "The isolation and sequence of the gene encoding T8: a molecule defining functional classes of T lymphocytes" *Cell* 40(2):237–246 (1985).

Littman DR, "The structure of the CD4 and CD8 genes" *Annual Review of Immunology* 5:561–584 (1987).

Maddon et al., "The isolation and Nucleotide Sequence of a cDNA Encoding the T Cell Surface Protein T4: A New Member of the Immunoglobulin Gene Family" *Cell* 42:93–104 (1985).

Maddon et al., "The T4 Gene Encodes the AIDS Virus Receptor and Is Expressed in the Immune System and the Brain" *Cell* 47:333–348 (1986).

Maddon P. et al., "Structure and expression of the human and mouse T4 genes" *Proc. Natl. Acad. Sci. USA* 84 (24):9155–9159 (1987).

McDougal et al., "Binding of HTLV–III/LAV to $T4^+$ Cells by a Complex of the 110k Viral Protein and the T4 Molecule" *Science* 231:382–385 (1986).

McDougal et al., "The T4 Glycoprotein Is a Cell–Surface Receptor for the AIDS Virus" *Molecular Biology of Homosapiens* (Cold Spring Harbor Symposia on Quantitative Biology), New York:CSHL vol. LI (1986).

Miller, Larry et al., "Expression of high–affinity binding of human immunoglobulin E by transfected cells" *Science* 244:334–336 (1989).

Morrison et al., "Chimeric Human Antibody Molecules: Mouse Antigen–Binding Domains with Human Constant Region Domains" *Proc. Natl. Acad. Sci. USA* 81:6851–6855 (Nov. 1984).

Morrison et al., "Genetically Engineered Antibody Molecules and Their Application" *Biological Approaches to the Controlled Delivery of Drugs*, Annals of the New York Academy of Sciences, New York, New York:The New York Academy of Sciences vol. 507:187–198 (1987).

Morrison et al., "Transfer and expression of immunoglobulin genes" *Annual Review of Immunology* 2:239–256 (1984).

Morrison, J., "Sequentially derived mutants of the constant region of the heavy chain of murine immunoglobulins" *J. of Immunology* 123(2):793–800 (1979).

Morrison, S. L., "Transfectomas Provide Novel Chimeric Antibodies" *Science* 229:1202–1207 (Sep. 20, 1985).

Murre et al., "Biochemical and functional analyses of a secreted H–$2L^d$ molecule" *Molecular & Cellular Biology* 6:1315–1319 (1986).

Neuberger et al., "Recombinant Antibodies Possessing Novel Effector Functions" *Nature* 312:604–608 (Dec. 13, 1984).

Oi et al., "Chimeric Antibodies" *BioTechniques* 4(3) :214–221 (1986).

Osborn, L. et al., "Direct expression cloning of vascular cell adhesion molecule 1, a cytokine–induced endothelial protein that binds to lymphocytes" *Cell* 59:1203–1211 (1989).

Peterson, A.S., "Genetic analysis of CD2/LFA–3 and CD4/HIV interactions" (thesis), Harvard University, Chapter 1, (1988).

Putney et al., "HTLV–III/LAV neutralizing antibodies to an *E.coli*" *Science* 234:1392–1395 (1986).

Rose and Bergmann, "Expression from Cloned cDNA of Cell-Surface Secreted Forms of the Glycoprotein of Vesicular Stomatitis Virus in Eucaryotic Cells" *Cell* 30:753–762 (1982).

Rosenblum et al., "Modification of human leukocyte interferon pharmacology with a monoclonal antibody" *Cancer Research* 45:2421–2424 (1985).

Salzawa et al., "The CD4 molecule is associated with the T-cell receptor" *Journal of Cellular Biochemistry Supplement* 11 (Part D, Abs. 421):273 (1987).

Schacker et al., "The Effects of High-Dose Recombinant Soluble CD4 on Human Immunodeficiency Virus Type 1 Viremia" *The Journal of Infectious Diseases* 169:37–40 (1994).

Seed B., "An LFA-3 cDNA encodes a phospholipid-linked membrane protein homologous to its receptor CD2" *Nature* 329:840–842 (1987).

Sharon et al., "Expression of a V_HC_K chimaeric protein in mouse myeloma cells" *Nature* 309:364–367 (1984).

Shearer et al., "Transport of Recombinant Human CD4-Immunoglobulin G Across the Human Placenta: Pharmacokinetics and Safety in Six Mother-Infant Pairs in AIDS Clinical Trial Group Protocol 146" *Clinical and Diagnostics Laboratory Immunology* pp. 281–285 (1995).

Sleckman et al., "Expression and function of CD4 in a murine T-cell hybridoma" *Nature* 328 (6128):351–353 (1987).

Smith et al., "Blocking of HIV-1 Infectivity by a Soluble, Secreted Form of the CD4 Antigen" *Science* 238:1704–1707 (1987).

Springer and Strominger, "Detergent-soluble HLA antigens contain a hydrophilic region at the COOH-terminus and a penultimate hydrophobic region" *Proc. Natl. Acad. Sci. USA* 73(7):2481–2485 (1976).

Stamenkovic et al., "The B Lymphocyte Adhesion Molecule CD22 Interacts with Leukocyte Common Antigen CD45RO on T Cells and α2-6 Sialyltransferase, CD75, on B Cells" *Cell* 66:1133–1144 (1991).

Terhorst C. et al., "Biochemical analysis of human T lymphocyte differentiation antigens T4 and T5" *Science* 209(4455):520–521 (1980).

Traunecker et al., "A novel approach for preparing anti-T cell receptor constant region antibodies" *European Journal of Immunology* 16:851–854 (1986).

Traunecker et al., "Highly Efficient Neutralization of HIV with Recombinant CD4-immunoglobulin Molecules" *Nature* 339:68–70 (1989).

Traunecker et al., "Soluble CD4 molecules neutralize human immunodeficiency virus" *Nature* 331:84–86 (1988).

Ushijima et al., "Synergistic Effect of Recombinant CD4-Immunoglobulin in Combination with Azidothymidine, Dideoxyinosine and 0.5 β-monoclonal Antibody on Human Immunodeficiency Virus Infection in vitro" *Letters in Applied Microbiology* 19:1–5 (1994).

Vitetta et al., "Redesigning Nature's Poisons to Create Anti-Tumor Reagents" *Science* 238:1098–1104 (1987).

Ward et al., "Prevention of HIV-1 IIIB Infection in Chimpanzees by CD4 Immunoadhesin" *Nature* 352:434–436 (1991).

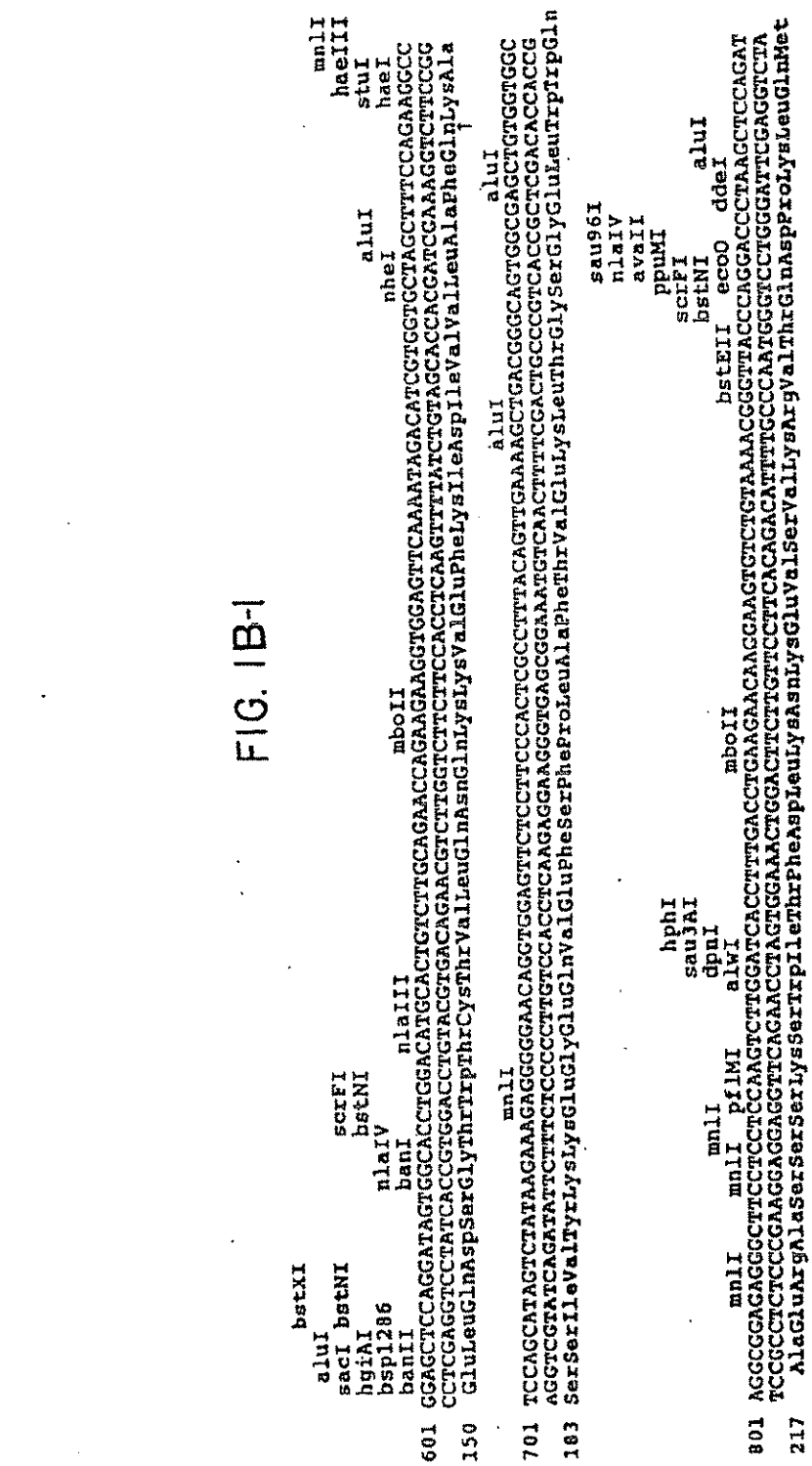
Wills et al., "Mutations of the Rous Sarcoma Virus env Gene that affect the transport and subcellular location of the glycoprotein products" *Journal of Cell Biology* 99:2011–2023 (1984).

FIG. 1A

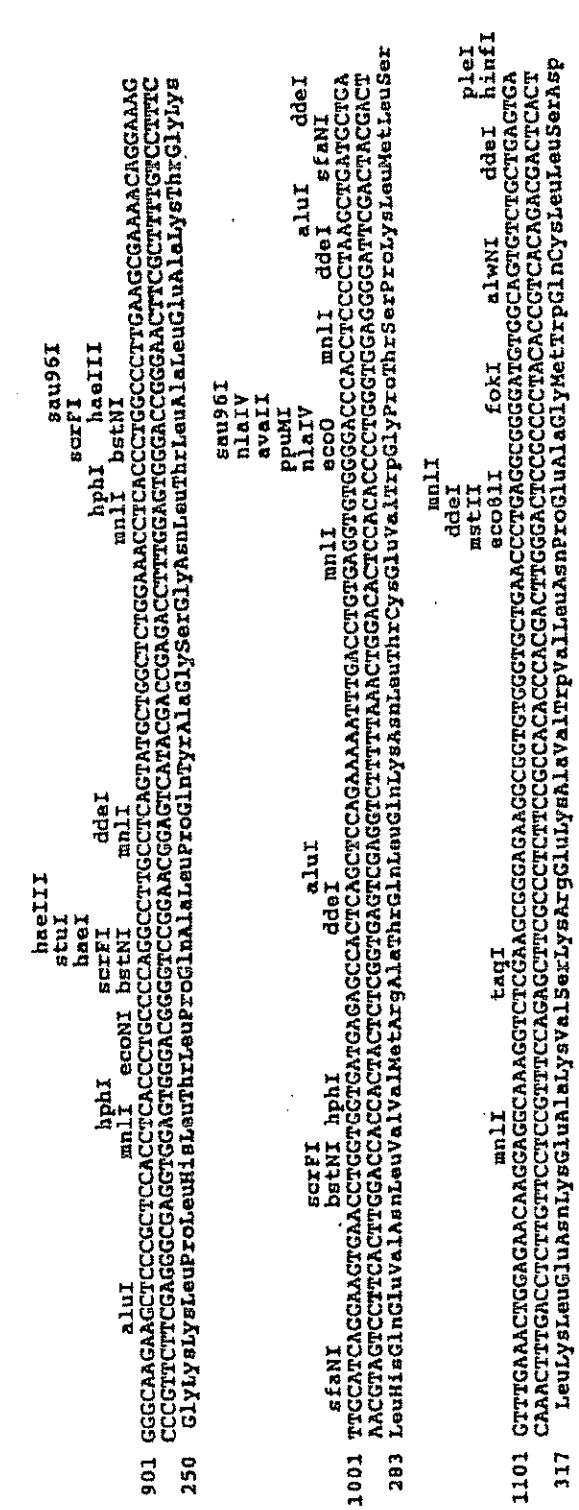


```
                                                sau96I
                                                nlaIV
                                                avaII
                                               ppuMI                                                  scrFI
                                          ddeI                                  mnlI                  nciI pleI
                bsp1286                 bsp1286                                 ecoNI    haeIII       mspI hinfI
                banII   bstXI           banII ecoO          ddeI        mnlI        haeI          hpaII
  1 AATTCAAGCCCAGAGCCCTGCCATTTCTGTGGGCTCAGGTCCCTACTGCTCAGCCCCTTCCTCCCTCGGCAAGGCCACAATGAACCGGGGAGTCCCTTTT
    TTAAGTTCGGGTCTCGGGACGGTAAAGACACCCGAGTCCAGGGATGACGAGTCGGGGAAGGAGGGAGCCGTTCCGGTGTTACTTGGCCCCTCAGGGAAAA
-25                                                                                  MetAsnArgGlyValProPhe
                                hinPI
                      fnu4HI    hhaI        fnu4HI
                      bbvI     haeII mnlI   bbvI   ddeI
101 AGGCACTTGCTTCTGGTGCTGCAACTGGCGCTCCTCCCAGCAGCCACTCAGGGAAACAAAGTGGTGCTGGGCAAAAAAGGGGATACAGTGGAACTGACCT
    TCCGTGAACGAAGACCACGACGTTGACCGCGAGGAGGGTCGTCGGTGAGTCCCTTTGTTTCACCACGACCCGTTTTTTCCCCTATGTCACCTTGACTGGA
-18 ArgHisLeuLeuLeuValLeuGlnLeuAlaLeuLeuProAlaAlaThrGlnGlyAsnLysValValLeuGlyLysLysGlyAspThrValGluLeuThrCys
                                                                                                  nlaIV                  fokI
                                                                                                bsp1286              sau96I
        aluI        mboII                                                                       banII     mseI       avaII
    rsaI          mboII                                                   hinfI
201 GTACAGCTTCCCAGAAGAAGAGCATACAATTCCACTGGAAAAACTCCAACCAGATAAAGATTCTGGGAAATCAGGGCTCCTTCTTAACTAAAGGTCCATC
    CATGTCGAAGGGTCTTCTTCTCGTATGTTAAGGTGACCTTTTTGAGGTTGGTCTATTTCTAAGACCCTTTAGTCCCGAGGAAGAATTGATTTCCAGGTAG
 17   ThrAlaSerGlnLysLysSerIleGlnPheHisTrpLysAsnSerAsnGlnIleLysIleLeuGlyAsnGlnGlySerPheLeuThrLysGlyProSer

                hinPI
                hhaI                                  styI                                                  ddeI
               thaI                                 sau96I                 sau3AI        mseI        pleI
             sau3AI   pleI                          avaII                  dpnI         aflII        hinfI
      aluI   dpnI     hinfI  mboII                 nlaIV                  bclI      hinfI        mboII
301 CAAGCTGAATGATCGCGCTGACTCAAGAAGAAGCCTTTGGGACCAAGGAAACTTTCCCCTGATCATCAAGAATCTTAAGATAGAAGACTCAGATACTTAC
    GTTCGACTTACTAGCGCGACTGAGTTCTTCTTCGGAAACCCTGGTTCCTTTGAAAGGGGACTAGTAGTTCTTAGAATTCTATCTTCTGAGTCTATGAATG
 50  LysLeuAsnAspArgAlaAspSerArgArgSerLeuTrpAspGlnGlyAsnPheProLeuIleIleLysAsnLeuLysIleGluAspSerAspThrTyr

                sau96I
                avaII        mnlI                                                                ecoNI
               mnlI        mnlI                                                                bspMI        alwNI
401 ATCTGTGAAGTGGAGGACCAGAAGGAGGAGGTGCAATTGCTAGTGTTCGGATTGACTGCCAACTCTGACACCCACCTGCTTCAGGGGCAGAGCCTGACCC
    TAGACACTTCACCTCCTGGTCTTCCTCCTCCACGTTAACGATCACAAGCCTAACTGACGGTTGAGACTGTGGGTGGACGAAGTCCCCGTCTCGGACTGGG
 83 IleCysGluValGluAspGlnLysGluGluValGlnLeuLeuValPheGlyLeuThrAlaAsnSerAspThrHisLeuLeuGlnGlyGlnSerLeuThrLeu

                      scrFI                                  styI                                                  aluI
                bsp1286                 ddeI                pleI                                                  pvuII
        styI   banII  bstNI            mnlI                hinfI                                  mboII mnlI       ddeI
501 TGACCTTGGAGAGCCCCCCTGGTAGTAGCCCCTCAGTGCAATGTAGGAGTCCAAGGGGTAAAAACATACAGGGGGGGAAGACCCTCTCCGTGTCTCAGCT
    ACTGGAACCTCTCGGGGGGACCATCATCGGGGAGTCACGTTACATCCTCAGGTTCCCCATTTTTGTATGTCCCCCCCTTCTGGGAGAGGCACAGAGTCGA
117   ThrLeuGluSerProProGlySerSerProSerValGlnCysArgSerProArgGlyLysAsnIleGlnGlyGlyLysThrLeuSerValSerGlnLeu
```

FIG. 1B-1

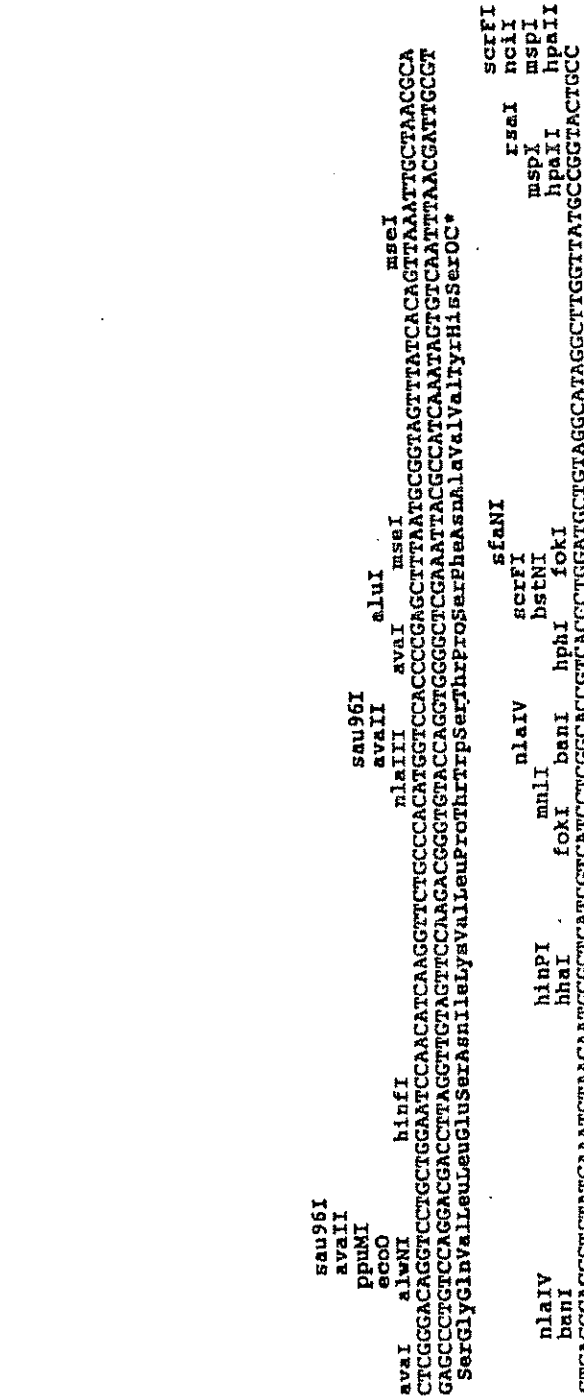


```
        bstXI
     aluI                                                                                                  mnlI
    sacI bstNI          scrFI                                                                             haeIII
    hgiAI               bstNI                                                                  aluI       stuI
    bsp1286         nlaIV                                                                     nheI        haeI
    banII           banI      nlaIII                  mboII
601 GGAGCTCCAGGATAGTGGCACCTGGACATGCACTGTCTTGCAGAACCAGAAGAAGGTGGAGTTCAAAATAGACATCGTGGTGCTAGCTTTCCAGAAGGCC
    CCTCGAGGTCCTATCACCGTGGACCTGTACGTGACAGAACGTCTTGGTCTTCTTCCACCTCAAGTTTTATCTGTAGCACCACGATCGAAAGGTCTTCCGG
150  GluLeuGlnAspSerGlyThrTrpThrCysThrValLeuGlnAsnGlnLysLysValGluPheLysIleAspIleValValLeuAlaPheGlnLysAla
                                                                                                  ↑
                                                                                  aluI                aluI
                        mnlI
701 TCCAGCATAGTCTATAAGAAAGAGGGGGAACAGGTGGAGTTCTCCTTCCCACTCGCCTTTACAGTTGAAAAGCTGACGGGCAGTGGCGAGCTGTGGTGGC
    AGGTCGTATCAGATATTCTTTCTCCCCCTTGTCCACCTCAAGAGGAAGGGTGAGCGGAAATGTCAACTTTTCGACTGCCCGTCACCGCTCGACACCACCG
183 SerSerIleValTyrLysLysGluGlyGluGlnValGluPheSerPheProLeuAlaPheThrValGluLysLeuThrGlySerGlyGluLeuTrpTrpGln

                                                                                                sau96I
                                                                                                nlaIV
                                                                                                avaII
                                                                                               ppuMI
                                    hphI                                                     scrFI
                                   sau3AI                                                    bstNI      aluI
                      mnlI         dpnI
            mnlI    mnlI  pflMI    alwI            mboII                          bstEII  ecoO  ddeI
801 AGGCGGAGAGGGCTTCCTCCTCCAAGTCTTGGATCACCTTTGACCTGAAGAACAAGGAAGTGTCTGTAAAACGGGTTACCCAGGACCCTAAGCTCCAGAT
    TCCGCCTCTCCCGAAGGAGGAGGTTCAGAACCTAGTGGAAACTGGACTTCTTGTTCCTTCACAGACATTTTGCCCAATGGGTCCTGGGATTCGAGGTCTA
217   AlaGluArgAlaSerSerSerLysSerTrpIleThrPheAspLeuLysAsnLysGluValSerValLysArgValThrGlnAspProLysLeuGlnMet
```



```
                                                haeIII
                                               stuI                                                  sau96I
                                               haeI                                               scrFI
                                  hphI      scrFI      ddeI                                  hphI   haeIII
                 aluI           mnlI  ecoNI bstNI     mnlI                                  mnlI  bstNI
 901 GGGCAAGAAGCTCCCGCTCCACCTCACCCTGCCCCAGGCCTTGCCTCAGTATGCTGGCTCTGGAAACCTCACCCTGGCCCTTGAAGCGAAAACAGGAAAG
     CCCGTTCTTCGAGGGCGAGGTGGAGTGGGACGGGGTCCGGAACGGAGTCATACGACCGAGACCTTTGGAGTGGGACCGGGAACTTCGCTTTTGTCCTTTC
 250  GlyLysLysLeuProLeuHisLeuThrLeuProGlnAlaLeuProGlnTyrAlaGlySerGlyAsnLeuThrLeuAlaLeuGluAlaLysThrGlyLys

                                                                                                sau96I
                                                                                                nlaIV
                                                                                                avaII
                                                                                               ppuMI
                                                                                               nlaIV                aluI    ddeI
                                scrFI                  aluI                                   mnlI   ecoO    mnlI  ddeI   sfaNI
       sfaNI                 bstNI hphI           ddeI                                 mnlI
1001 TTGCATCAGGAAGTGAACCTGGTGGTGATGAGAGCCACTCAGCTCCAGAAAAATTTGACCTGTGAGGTGTGGGGACCCACCTCCCCTAAGCTGATGCTGA
     AACGTAGTCCTTCACTTGGACCACCACTACTCTCGGTGAGTCGAGGTCTTTTTAAACTGGACACTCCACACCCCTGGGTGGAGGGGATTCGACTACGACT
 283 LeuHisGlnGluValAsnLeuValValMetArgAlaThrGlnLeuGlnLysAsnLeuThrCysGluValTrpGlyProThrSerProLysLeuMetLeuSer

                                                                                          mnlI
                                                                                         ddeI
                                                                                        mstII                                  pleI
                                  mnlI        taqI                                      eco81I    fokI    alwNI    ddeI  hinfI
1101 GTTTGAAACTGGAGAACAAGGAGGCAAAGGTCTCGAAGCGGGAGAAGGCGGTGTGGGTGCTGAACCCTGAGGCGGGGATGTGGCAGTGTCTGCTGAGTGA
     CAAACTTTGACCTCTTGTTCCTCCGTTTCCAGAGCTTCGCCCTCTTCCGCCACACCCACGACTTGGGACTCCGCCCCTACACCGTCACAGACGACTCACT
 317   LeuLysLeuGluAsnLysGluAlaLysValSerLysArgGluLysAlaValTrpValLeuAsnProGluAlaGlyMetTrpGlnCysLeuLeuSerAsp
```

FIG. 1B-2



```
                sau96I
                avaII
               ppuMI                                  sau96I
               ecoO                                   avaII      aluI
      avaI   alwNI       hinfI                     nlaIII    avaI    mseI                    mseI
1201  CTCGGGACAGGTCCTGCTGGAATCCAACATCAAGGTTCTGCCCACATGGTCCACCCCGAGCTTTAATGCGGTAGTTTATCACAGTTAAATTGCTAACGCA
      GAGCCCTGTCCAGGACGACCTTAGGTTGTAGTTCCAAGACGGGTGTACCAGGTGGGGCTCGAAATTACGCCATCAAATAGTGTCAATTTAACGATTGCGT
 350   SerGlyGlnValLeuLeuGluSerAsnIleLysValLeuProThrTrpSerThrProSerPheAsnAlaValValTyrHisSerOC*
                                                                                  sfaNI                                  scrFI
                                                                  nlaIV      scrFI                                  rsaI  nciI
          nlaIV                    hinPI          mnlI              bstNI                                  mspI     mspI
          banI                     hhaI       fokI   banI   hphI   fokI                                   hpaII    hpaII
1301  GTCAGGCACCGTGTATGAAATCTAACAATGCGCTCATCGTCATCCTCGGCACCGTCACCCTGGATGCTGTAGGCATAGGCTTGGTTATGCCGGTACTGCC
      CAGTCCGTGGCACATACTTTAGATTGTTACGCGAGTAGCAGTAGGAGCCGTGGCAGTGGGACCTACGACATCCGTATCCGAACCAATACGGCCATGACGG

         mnlI
       haeIII
      sau96I
1401  GGGCCTCTTGCGGGAT
      CCCGGAGAACGCCCTA
```

FIG. 1C

FIG. 2A

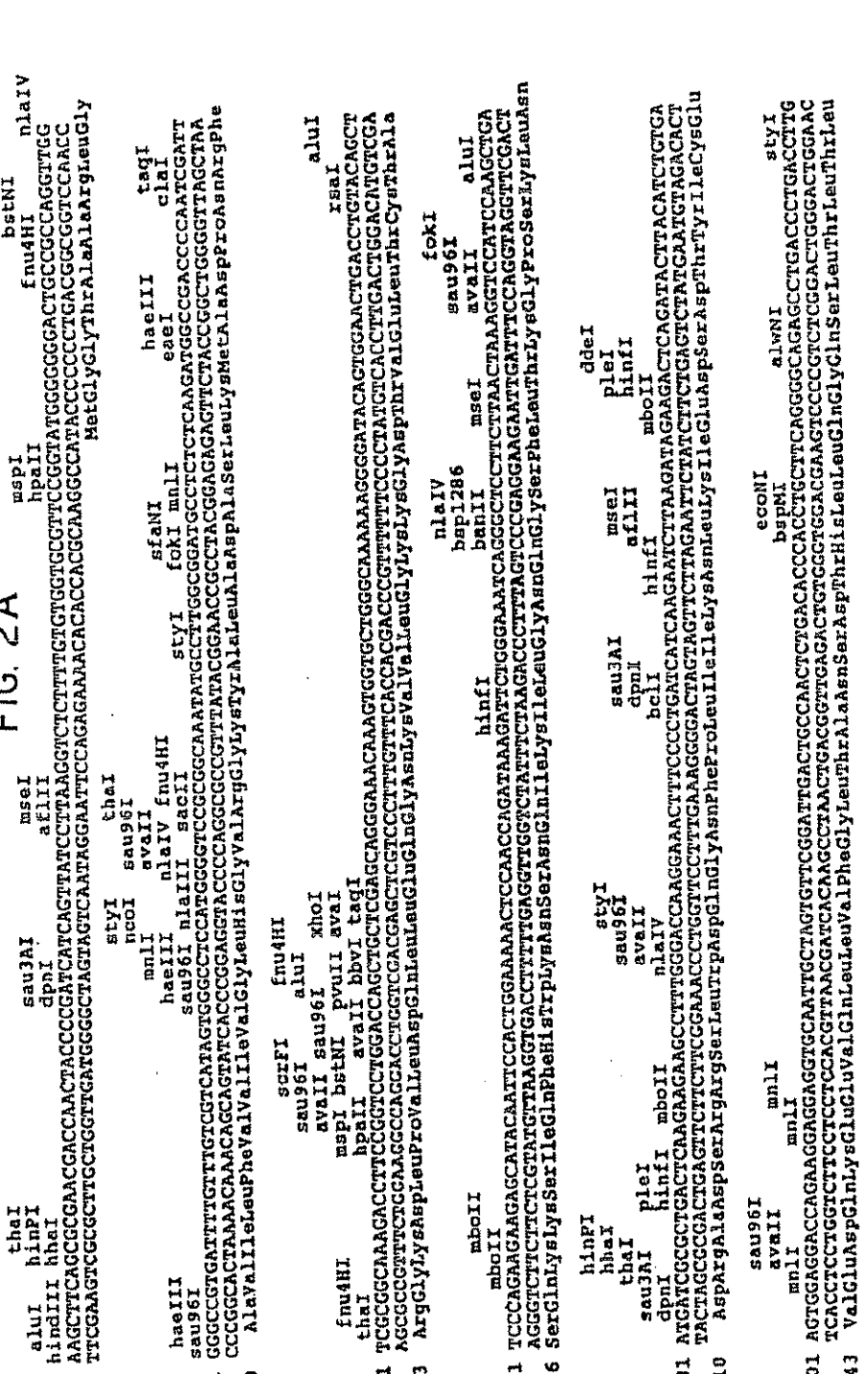
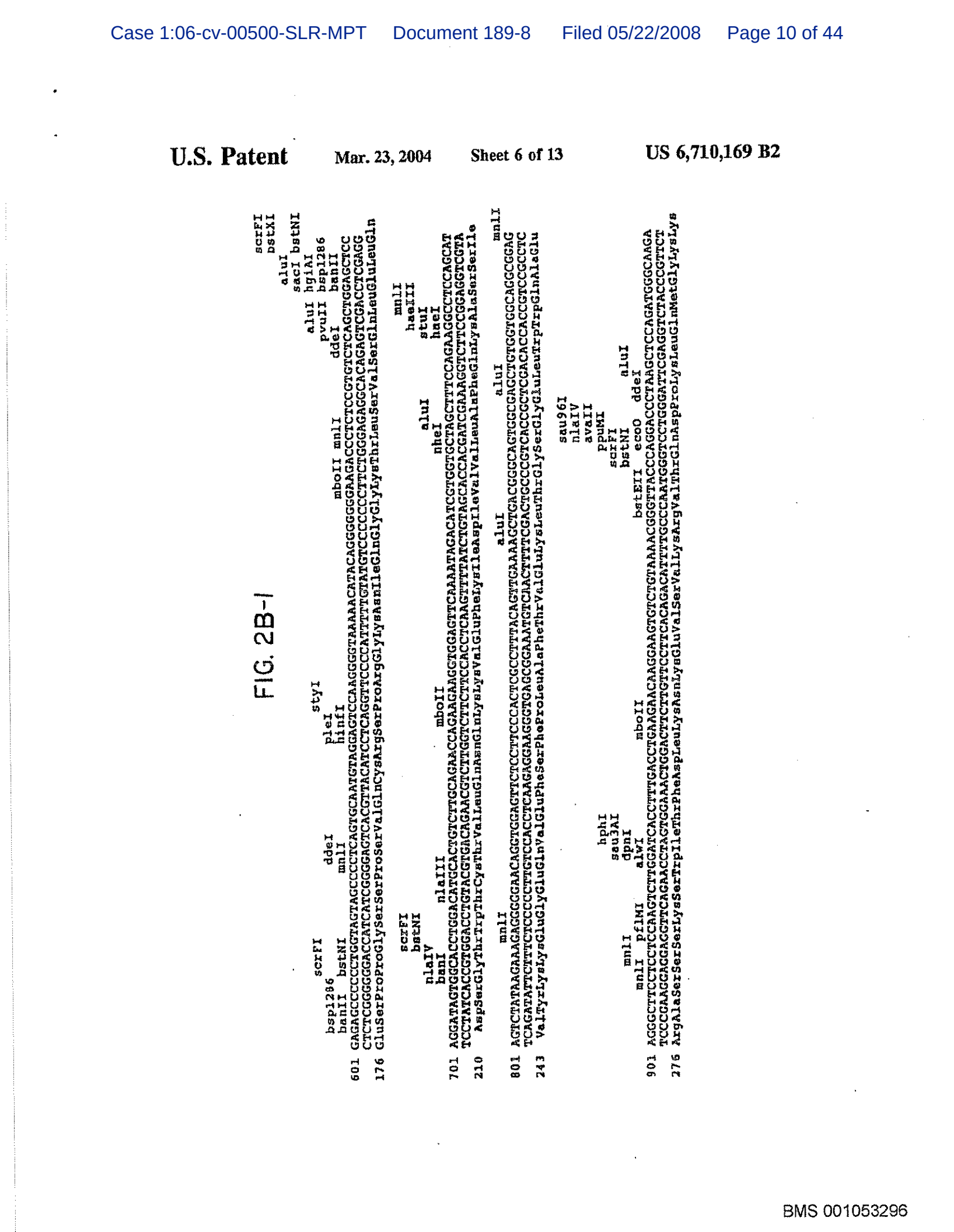
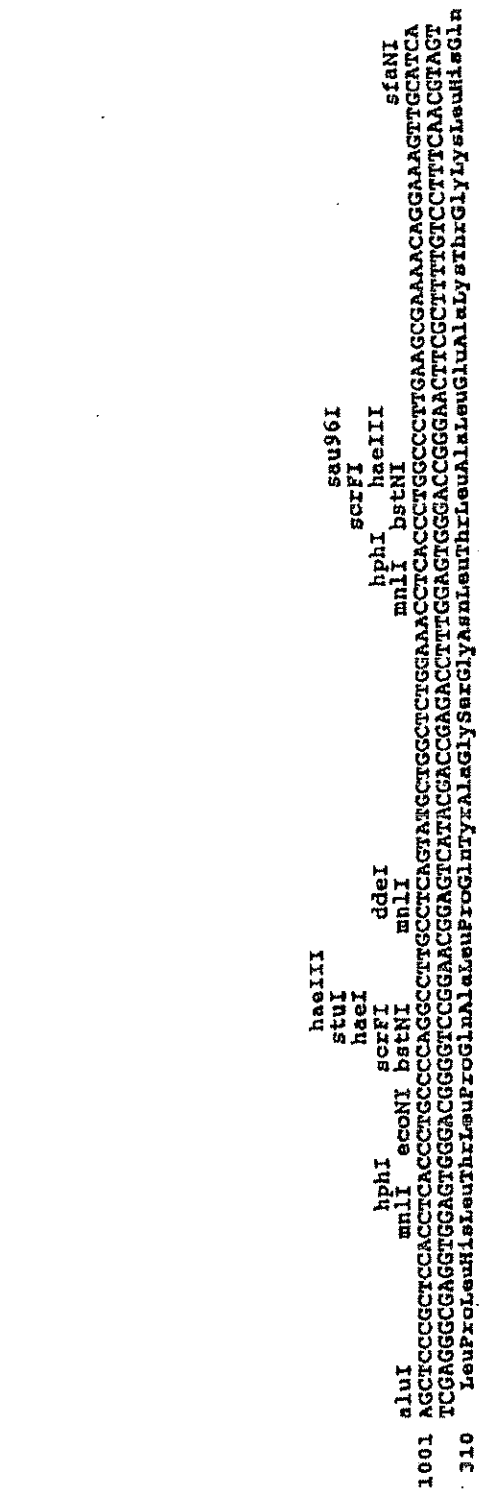

1 AAGCTTCAGCGCGAACCACCAACTACCCCGATCATCAGTTATCCTTAAGGTCTCTTTTGTGTGGTGCGTTCCGGTATGGGGGGGGACTGCCGCCCAGGTTGG
hindIII hhaI
aluI hinPI
thaI
101 GGGCCGTGATTTTGTTGTCGTCATAGTGGGCCTCCCATGGGGTCCGCGGCGGCAAATATGCCTTGGCGGATGCCTCTCTCAAGATGGCCGACCCCAATCGATT
sau96I
hae3II
10 AlaValIleLeuPheValIleValGlyValLysTyrAlaLeuAlaAspAlaSerLeuLysMetAlaAspProAsnArgPhe
1 MetGlyGlyThrAlaAlaArgLeuGly
thaI
styI sau96I
ncoI avaII
mnlI hae3III nlaIV fnu4HI styI fokI mnlI sfaNI haeIII taqI
sau96I nlaIII sacII eaeI claI
sau96I
101 CCCGGCACTAAAACAAACAGCAGTATCACCCCGAGGTACCCCAGGCGCCGTTTATATACGGAACCGCCTACGGGAGGAGCTTCTACCGGCTGGGGTTAGCTAA
fnu4HI
thaI
201 TCGCGGCAAGACCTTCCGGTCCTGGACCAGCTGCTCGAGCAGGGAAACAAGTGGTGCTGGGCAAAAAAGGGATACAGTGGATCTGACCTGTACAGCT
AGCGCCGTTCTGGCAAGGCCAGGACCTGGTCGACGAGCTCGTCCCTTTGTTTCACCACGACCCGTTTTTTCCCCTATGTCACCTTGACTGGACATGTCGA
43 ArgGlyLysAspLeuProValLeuAspGlnLeuLeuGluGlnGlyAsnLysValValIleLeuGlyLysGlyAspThrValGluGluLeuThrCysThrAla
hpaII avaII sau96I xhoI
fnu4HI pvuII avaI
mspI bstNI
hpaII avaII bbvI taqI
nlaIV fokI
sau96I
bsp1286 hinfI banII mseI avaII aluI
mboII
301 TCCCAGAAGAAGAGCATACAATTCCACTGGAAAAACTCCAACCAGATAAAGATTCTGGGAAATCAGGGCTCCTTCTTAACTAAAGGTCCATCCAAGCTCA
AGGGTCTTCTTCTCGTATGTTAAGGTGACCTTTTTGAGGTTGGTCTATTTCTAAGACCCTTTAGTCCCGAGGAAGAATTGATTTCCAGGTAGGTTCGACT
76 SerGlnLysLysSerIleAsnSerArgAsnGlnIleLeuLysIleLeuGlyAsnGlnGlySerPheLeuThrLysGlyProSerLysLeuAsn
ddeI
hinPI pleI
hhaI styI sau3AI mseI hinfI
thaI sau96I avaII dpnI aflII hinfI
sau3AI pleI avaII bclI mboII
dpnI hinfI mboII nlaIV hinfI
401 ATGATCGCGCTGACTCAAGAAGAAGCCCTTTGGGACCAAGGAAACTTTCCCCTGATCATCAAGAATCTTAAGATAGAAGACTCAGATACTTACATCTGTGA
TACTAGCGCGACTGAGTTCTTCTTCGGGAAACCCTGGTTCCTTTGAAAGGGGACTAGTAGTTCTTAGAATTCTATCTTCTGAGTCTTATGAATGTAGACACT
110 AspArgAlaAspSerArgArgSerLeuTrpAspGlnGlyAsnPheProLeuIleIleLysAsnLeuLysIleGluAspSerAspThrTyrIleCysGlu
ecoNI
sau96I bspMI alwNI styI
avaII mnlI
mnlI
501 AGTGGAGGACCAGAAGGAGGAGGTGCAATTGCTAGTGTTCGGATTGACTGCCAACTCTGACACCCACCTGCTTCAGGGGCAGAGCCCTGACCCTGACCTTG
TCACCTCCTGGTCTTCCTCCTCCACGTTAACGATCACAACGGCCTAACTGACGGTTGAGACTGTGGGTGGACGAAGTCCCCGTCTCGGACTGGGACTGGAAC
143 ValGluAspGlnLysGluGluValGlnLeuLeuValPheGlyLeuThrAlaAsnSerAspThrHisLeuLeuGlnGlyGlnSerLeuThrLeuThrLeu

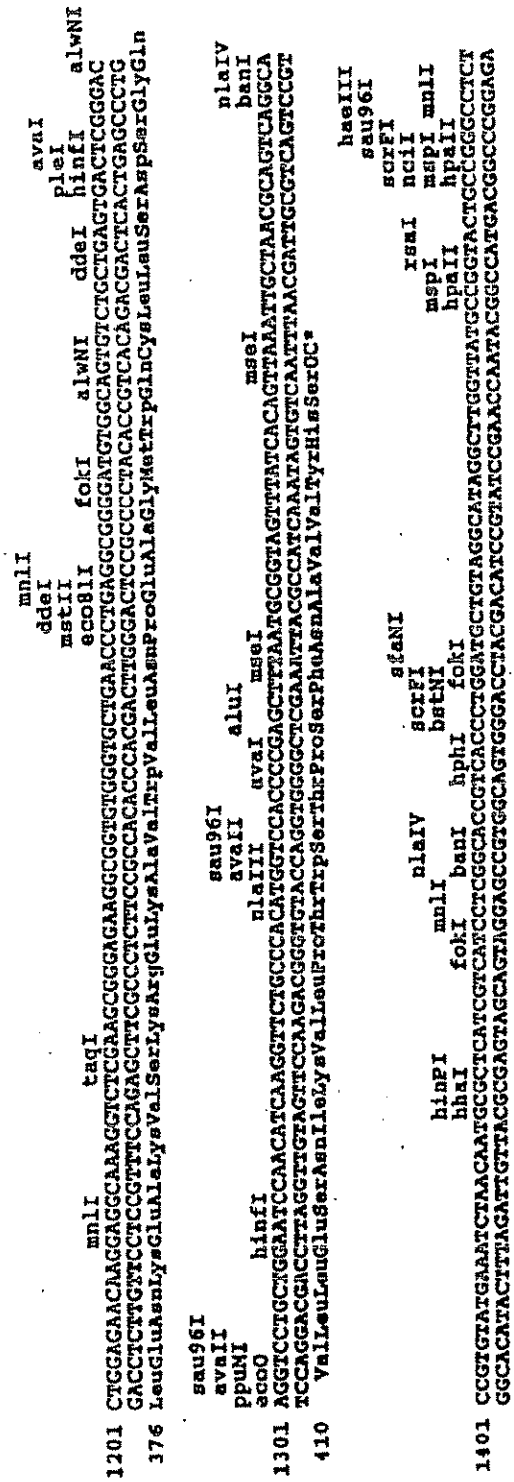
FIG. 2B-1


scrFI
bstXI
aluI
sacI bstNI
aluI hgiAI
pvuII bsp1286
ddeI banII
mboII mnlI
mnlI
haeIII
stuI
haeI
nheI
aluI
aluI
mnlI
mnlI
sau96I
nlaIV
avaII
ppuMI
scrFI
bstNI aluI
ecoO ddeI
bstEII
hphI
sau3AI
dpnI
mboII
mnlI pflMI alwI
scrFI
bstNI
nlaIV
banI
nlaIII
mboII
aluI
mnlI
styI
pleI
hinfI
ddeI
mnlI
bsp1286
banII bstNI
601 GAGAGCCCCCCCTGGTAGTAGCCCCTCAGTGCAATGTAGGAGTCCAAGGGGTAAAAACATACAGGGGGGGAAGACCCTCTCCGTGTCTCAGCTGGAGCTCC
CTCTCGGGGGGGACCATCATCGGGGAGTCACGTTACATCCTCAGGTTCCCCATTTTTGTATGTCCCCCCCTTCTGGGAGAGGCACAGAGTCGACCTCGAGG
176 GluSerProProGlySerSerProSerValGlnCysArgSerProArgGlyLysAsnIleGlnGlyGlyLysThrLeuSerValSerGlnLeuGluLeuGln
701 AGGATAGTGGCACCTGGACATGCACTGTCTTGCAGAACCAGAAGAAGGTGGAGTTCAAATAGACATCGTGGTGCTAGCTTTCCAGAAGGCCTCCAGCAT
TCCTATCACCGTGGACCTGTACGTGACAGAACGTCTTGGTCTTCTTCCACCTCAAGTTTTATCTGTAGCACCACGATCGAAAGGTCTTCCGGAGGTCGTA
210 AspSerGlyThrTrpThrCysThrValLeuGlnAsnGlnLysLysValGluGluPheLysIleAspIleValValLeuAlaPheGlnLysAlaSerSerIle
801 AGTCTATAAGAAAGAGGGCGAACAGGTGGAGTTCTCCTTCCCACTCGCCTTTACAGTTGAAAGCTGACGGCAGTGTGGTGGGCAGGCGGAG
TCAGATATTCTTTCTCCCCCTTGTCCACCTCAAGAGGGAAGGGTGAGCGGAAATGTCACTTTCGACTGCCCGTCACCGCCACACCGTCCGCCTC
243 ValTyrLysLysGluGlyGlnValGluGluPheSerProLeuAlaPheThrValGluLysLeuThrGlySerGlyGluLeuTrpTrpGlnAlaGlu
901 AGGGCTTCCTCCTCCAAGTCTTGGATCACCTTTGACCTGAAGAACAAGGAAGTGTCTGTAAAACGGGTTACCCCAGGACCCTAAGCTCCAGATGGGCAAGA
TCCCGAAGGAGGAGGTTCAGAACCTAGTGGAAACTGGACTTCTTGTTCCTTCACAGACATTTTGCCCAATGGGTCCTGGGATTCGAGGTCTACCCGTTCT
276 ArgAlaSerSerSerLysSerTrpIleThrPheAspLeuLysAsnLysGluValSerValLysArgValThrGlnAspProLysLeuGlnMetGlyLysLys



```
                                 haeIII
                                stuI                                      sau96I
                                haeI                                    scrFI
                    hphI      scrFI      ddeI                       hphI   haeIII
     aluI          mnlI  ecoNI bstNI     mnlI                     mnlI  bstNI                                  sfaNI
1001 AGCTCCCGCTCCACCTCACCCTGCCCCAGGCCTTGCCTCAGTATGCTGGCTCTGGAAACCTCACCCTGGCCCTTGAAGCGAAAACAGGAAAGTTGCATCA
     TCGAGGGCGAGGTGGAGTGGGACGGGGTCCGGAACGGAGTCATACGACCGAGACCTTTGGAGTGGGACCGGGAACTTCGCTTTTGTCCTTTCAACGTAGT
 310   LeuProLeuHisLeuThrLeuProGlnAlaLeuProGlnTyrAlaGlySerGlyAsnLeuThrLeuAlaLeuGluAlaLysThrGlyLysLeuHisGln

                                                                        sau96I
                                                                        nlaIV
                                                                        avaII
                                                                       ppuMI
                scrFI                  aluI                            nlaIV              aluI    ddeI
                bstNI hphI           ddeI                     mnlI    ecoO    mnlI  ddeI   sfaNI
1101 GGAAGTGAACCTGGTGGTGATGAGAGCCACTCAGCTCCAGAAAAATTTGACCTGTGAGGTGTGGGGACCCACCTCCCCTAAGCTGATGCTGAGTTTGAAA
     CCTTCACTTGGACCACCACTACTCTCGGTGAGTCGAGGTCTTTTTAAACTGGACACTCCACACCCCTGGGTGGAGGGGATTCGACTACGACTCAAACTTT
 343  GluValAsnLeuValValMetArgAlaThrGlnLeuGlnLysAsnLeuThrCysGluValTrpGlyProThrSerProLysLeuMetLeuSerLeuLys
```

FIG. 2B-2



```
                                                                                  mnlI
                                                                                ddeI                                    avaI
                                                                               mstII                                   pleI
                                                                               eco81I    fokI    alwNI    ddeI   hinfI    alwNI
                  mnlI        taqI
1201 CTGGAGAACAAGGAGGCAAAGGTCTCGAAGCGGGAGAAGGCGGTGTGGGTGCTGAACCCTGAGGCGGGGATGTGGCAGTGTCTGCTGAGTGACTCGGGAC
     GACCTCTTGTTCCTCCGTTTCCAGAGCTTCGCCCTCTTCCGCCACACCCACGACTTGGGACTCCGCCCCTACACCGTCACAGACGACTCACTGAGCCCTG
 376 LeuGluAsnLysGluAlaLysValSerLysArgGluLysAlaValTrpValLeuAsnProGluAlaGlyMetTrpGlnCysLeuLeuSerAspSerGlyGln

      sau96I
      avaII                                   sau96I
     ppuMI                                    avaII      aluI                                                  nlaIV
     eco0       hinfI                        nlaIII    avaI    mseI                    mseI                     banI
1301 AGGTCCTGCTGGAATCCAACATCAAGGTTCTGCCCACATGGTCCACCCCGAGCTTTAATGCGGTAGTTTATCACAGTTAAATTGCTAACGCAGTCAGGCA
     TCCAGGACGACCTTAGGTTGTAGTTCCAAGACGGGTGTACCAGGTGGGGCTCGAAATTACGCCATCAAATAGTGTCAATTTAACGATTGCGTCAGTCCGT
 410   ValLeuLeuGluSerAsnIleLysValLeuProThrTrpSerThrProSerPheAsnAlaValValTyrHisSerOC*

                                                                                                                haeIII
                                                                                                               sau96I
                                                                                                              scrFI
                                                                  sfaNI                                  rsaI  nciI
                                                  nlaIV          scrFI                                  mspI    mspI mnlI
                                hinPI          mnlI             bstNI                                  hpaII   hpaII
                                hhaI         fokI   banI   hphI   fokI
1401 CCGTGTATGAAATCTAACAATGCGCTCATCGTCATCCTCGGCACCGTCACCCTGGATGCTGTAGGCATAGGCTTGGTTATGCCGGTACTGCCGGGCCTCT
     GGCACATACTTTAGATTGTTACGCGAGTAGCAGTAGGAGCCGTGGCAGTGGGACCTACGACATCCGTATCCGAACCAATACGGCCATGACGGCCCGGAGA

1501 TGCGGGAT
     ACGCCCTA
```

FIG. 2C

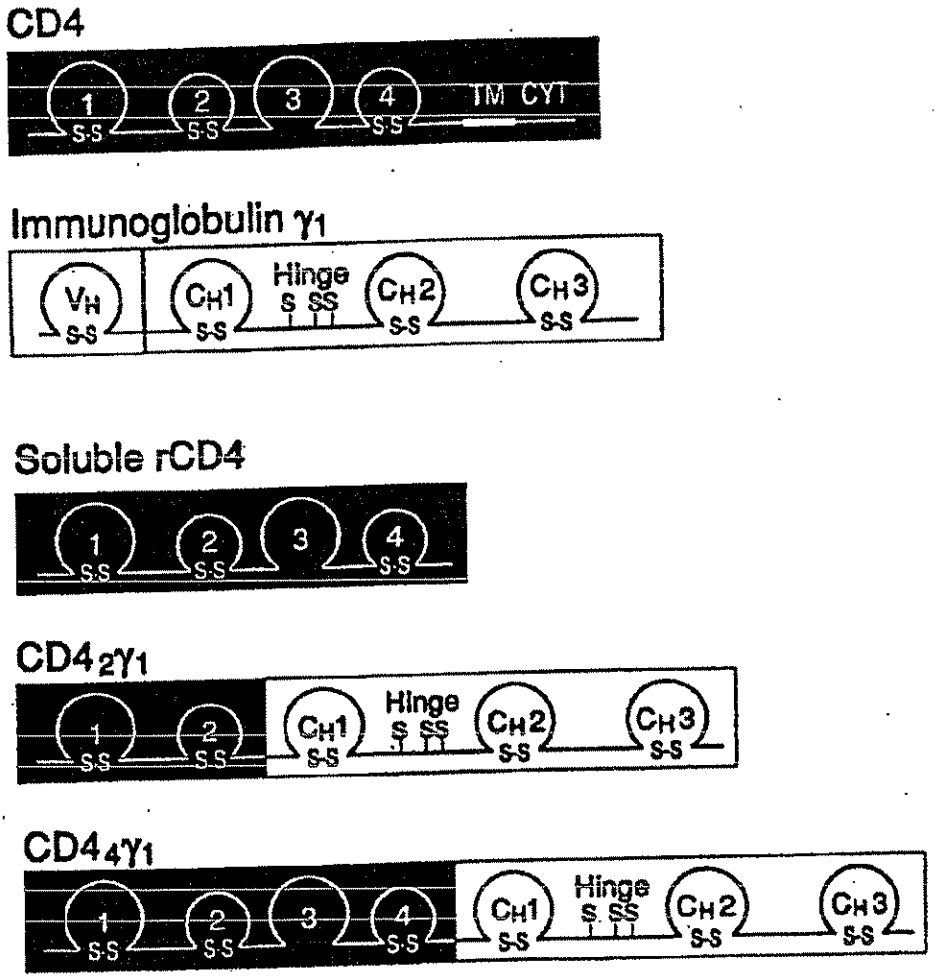
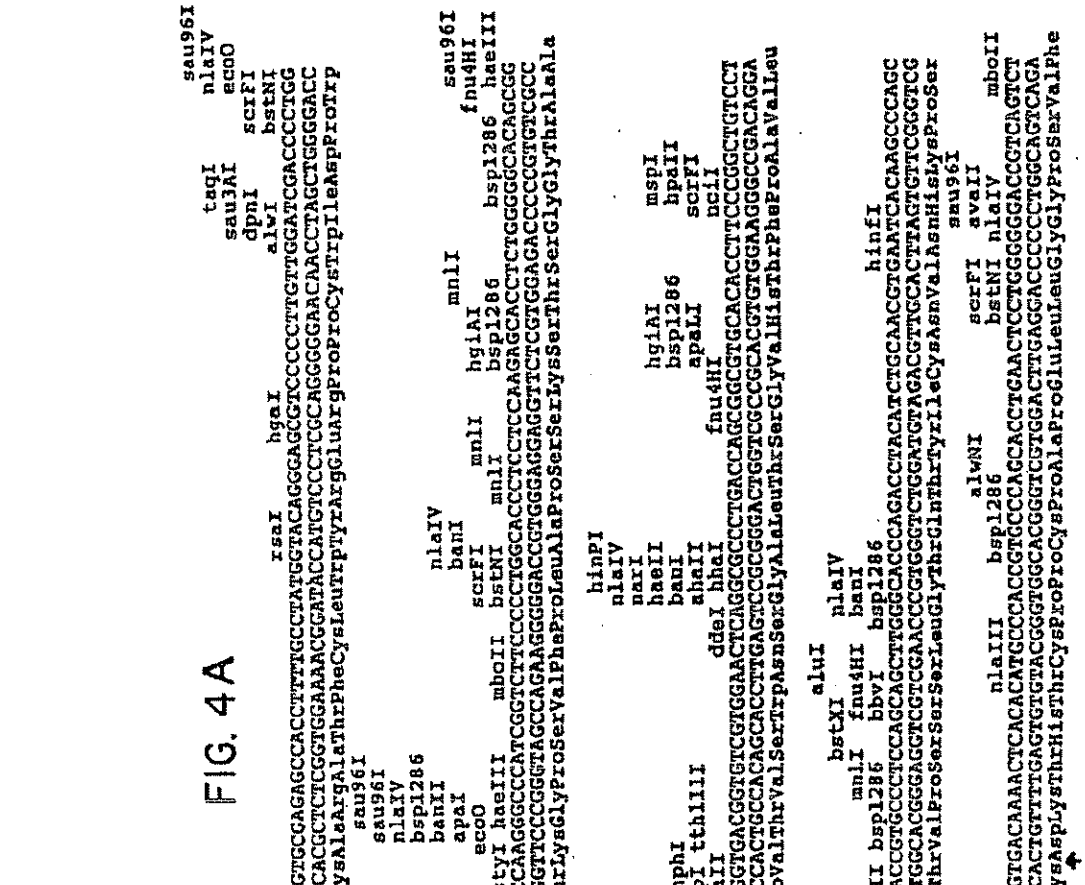

CD4
1
2
3
4
TM
CYT
S-S
Immunoglobulin γ1
VH
CH1
Hinge
S SS
CH2
CH3
Soluble rCD4
CD4 2γ1
CD4 4γ1

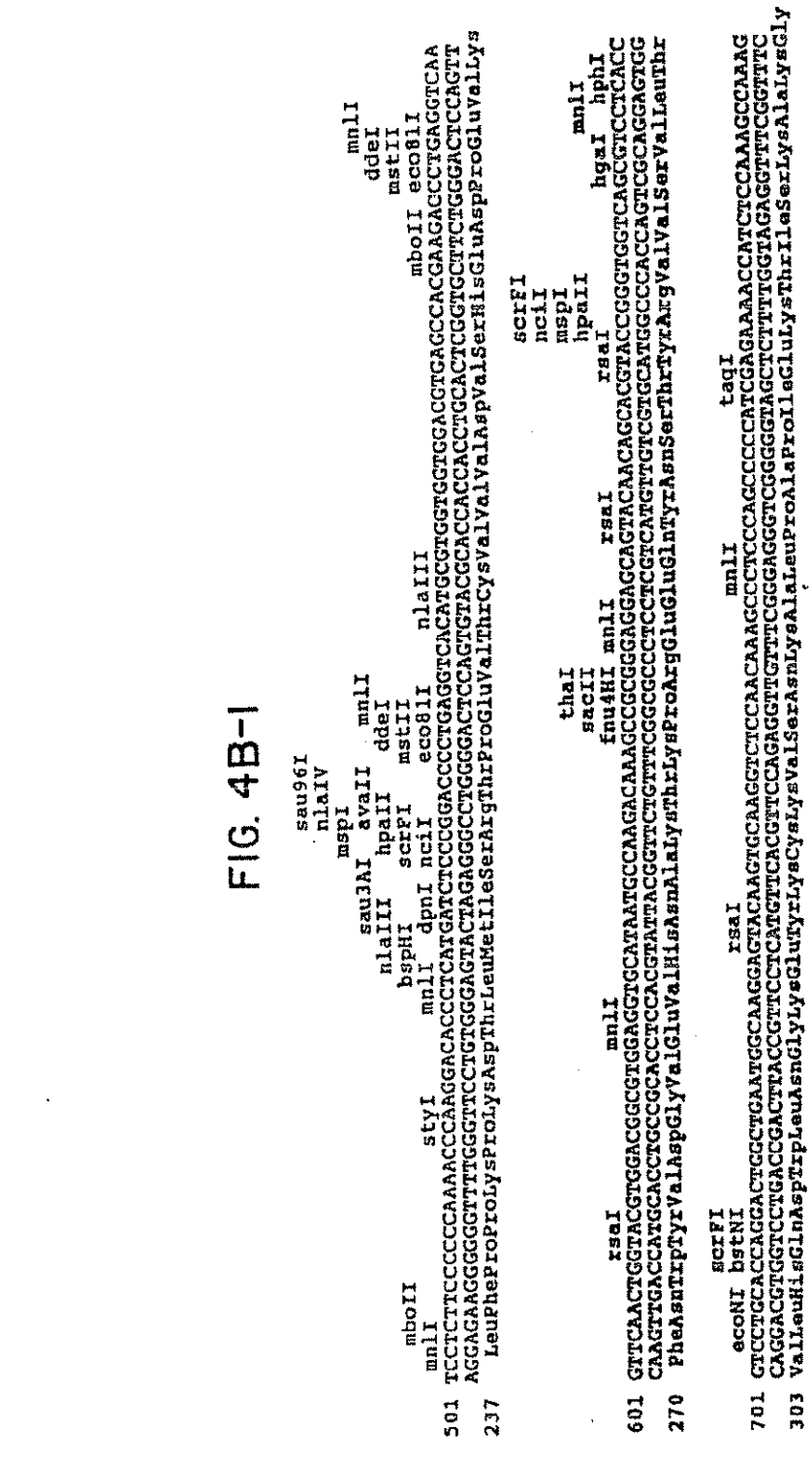
FIG. 3

FIG. 4A



```
                                                                                                                                  sau96I
                                                                                                                      taqI      nlaIV
                                                                                                                    sau3AI      ecoO
                  thaI     haeIII                                                                                    dpnI     scrFI
                 sacII    xmaIII                                                                                    alwI     bstNI
   ecoRI        fnu4HI    eaeI                                                  rsaI     hgaI
  1 GAATTCTGTCACTGCCGCGGACACGGCCGTATATTACTGTGCGAGAGCCACCTTTTGCCTATGGTACAGGGAGCGTCCCCCTTGTTGGATCGACCCCTGG
    CTTAAGACAGTGACGGCGCCTGTGCCGGCATATAATGACACGCTCTCGGTGGAAAACGGATACCATGTCCCTCGCAGGGGGAACAACCTAGCTGGGGACC
 72        ValThrAlaAlaAspThrAlaValTyrTyrCysAlaArgAlaThrPheCysLeuTrpTyrArgGluArgProProCysTrpIleAspProTrp
                                                  sau96I
                                                 sau96I
                                                 nlaIV
                                                 bsp1286
          nlaIV     hphI                          banII                  nlaIV
      scrFI       bstEII                          apaI                   banI                  mnlI                  sau96I
      bstNI     scrFI              haeIII        ecoO                  scrFI      mnlI    hgiAI                  fnu4HI
    haeIII      bstNI          mnlI  mnlI  styI haeIII     mboII  bstNI   mnlI       bsp1286      bsp1286  haeIII
101 GGCCTGGGAACCCTGGTCACCGTCTCCTCGGCCTCCACCAAGGGCCCATCGGTCTTCCCCCTGGCACCCTCCTCCAAGAGCACCTCTGGGGGCACAGCGG
    CCGGACCCTTGGGACCAGTGGCAGAGGAGCCGGAGGTGGTTCCCGGGTAGCCAGAAGGGGGACCGTGGGAGGAGGTTCTCGTGGAGACCCCCGTGTCGCC
103 GlyLeuGlyThrLeuValThrValSerSerAlaSerThrLysGlyProSerValPheProLeuAlaProSerSerLysSerThrSerGlyGlyThrAlaAla
                                      ↑
                                      γ1
                                                                        hinPI
                                                                       nlaIV
                scrFI                                                  narI
          fnu4HI                                                       haeII                hgiAI       mspI
          bbvI                                 hphI                    banI                 bsp1286     hpaII
     scrFI     ecoNI                         mspI tth111I              ahaII                apaLI      scrFI
     bstNI     bstNI                         hpaII                 ddeI hhaI        fnu4HI           nciI
201 CCCTGGGCTGCCTGGTCAAGGACTACTTCCCCGAACCGGTGACGGTGTCGTGGAACTCAGGCGCCCTGACCAGCGGCGTGCACACCTTCCCGGCTGTCCT
    GGGACCCGACGGACCAGTTCCTGATGAAGGGGCTTGGCCACTGCCACAGCACCTTGAGTCCGCGGGACTGGTCGCCGCACGTGTGGAAGGGCCGACAGGA
137   LeuGlyCysLeuValLysAspTyrPheProGluProValThrValSerTrpAsnSerGlyAlaLeuThrSerGlyValHisThrPheProAlaValLeu

          ddeI                                                  aluI
         mstII pleI            fnu4HI                  bstXI      nlaIV
         mnlI  hinfI       ddeI        hphI        mnlI  fnu4HI  banI
         eco81I           mnlI bbvI   bstEII bsp1286   bbvI   bsp1286                          hinfI
301 ACAGTCCTCAGGACTCTACTCCCTCAGCAGCGTGGTGACCGTGCCCTCCAGCAGCTTGGGCACCCAGACCTACATCTGCAACGTGAATCACAAGCCCAGC
    TGTCAGGAGTCCTGAGATGAGGGAGTCGTCGCACCACTGGCACGGGAGGTCGTCGAACCCGTGGGTCTGGATGTAGACGTTGCACTTAGTGTTCGGGTCG
170  GlnSerSerGlyLeuTyrSerLeuSerSerValValThrValProSerSerSerLeuGlyThrGlnThrTyrIleCysAsnValAsnHisLysProSer
                                                                                                  sau96I
                          bsp1286                                          alwNI        scrFI  avaII
        styI              banII                                  nlaIII    bsp1286          bstNI nlaIV      mboII
401 AACACCAAGGTGGACAAGAAAGTTGAGCCCAAATCTTGTGACAAAACTCACACATGCCCACCGTGCCCAGCACCTGAACTCCTGGGGGGACCGTCAGTCT
    TTGTGGTTCCACCTGTTCTTTCAACTCGGGTTTAGAACACTGTTTTGAGTGTGTACGGGTGGCACGGGTCGTGGACTTGAGGACCCCCCTGGCAGTCAGA
203 AsnThrLysValAspLysLysValGluProLysSerCysAspLysThrHisThrCysProProCysProAlaProGluLeuLeuGlyGlyProSerValPhe
                                          ↑
                                          Fc
```

FIG. 4B-1


sau96I
nlaIV
mspI
mnlI
sau3AI avaII mnlI
ddeI
nlaIII hpaII ddeI
mstII
bspHI scrFI mstII
mboII eco81I
mnlI dpnI nciI eco81I nlaIII
styI
501 TCCTCTTCCCCCCAAAACCCAAGGACACCCTCATGATCTCCCGGACCCCTGAGGTCACATGCGTGGTGGACGTGAGCCACGAAGACCCTGAGGTCAA
AGGAGAAGGGGGTTTTGGGTTCCTGTGGGAGTACTAGAGGGCCTGGGGACTCCAGTGTACGCACCACCACCTGCACTCGGTGCTTCTGGGACTCCAGTT
237 LeuPheProProLysProLysAspThrLeuMetIleSerArgThrProGluValThrCysValValValAspValSerHisGluAspProGluValLys
scrFI
nciI
mspI
hpaII
mnlI
thaI
sacII
rsaI
hgaI hphI
fnu4HI mnlI rsaI
601 GTTCAACTGGTACGTGGACGGCGTGGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAGCAGTACAACAGCACGTACCGGGTGGTCAGCGTCCTCACC
CAAGTTGACCATGCACCTGCCGCACCTCCACGTATTACGGTTCTGTTTCGGCGCCCTCCTCGTCATGTTGTCGTGCATGGCCCACCAGTCGCAGGAGTGG
270 PheAsnTrpTyrValAspGlyValGluValHisAsnAlaLysThrLysProArgGluGluGlnTyrAsnSerThrTyrArgValValSerValLeuThr
scrFI
rsaI
mnlI
taqI
ecoNI bstNI
701 GTCCTGCACCAGGACTGGCTGAATGGCAAGGAGTACAAGTGCAAGGTCTCCAACAAAGCCCTCCCAGCCCCCATCGAGAAAACCATCTCCAAAGCCAAAG
CAGGACGTGGTCCTGACCGACTTACCGTTCCTCATGTTCACGTTCCAGAGGTTGTTTCGGGAGGGGTCGGGGGTAGCTCTTTTGGTAGAGGTTTCGGTTTC
303 ValLeuHisGlnAspTrpLeuAsnGlyLysGluTyrLysCysLysValSerAsnLysAlaLeuProAlaProIleGluLysThrIleSerLysAlaLysGly

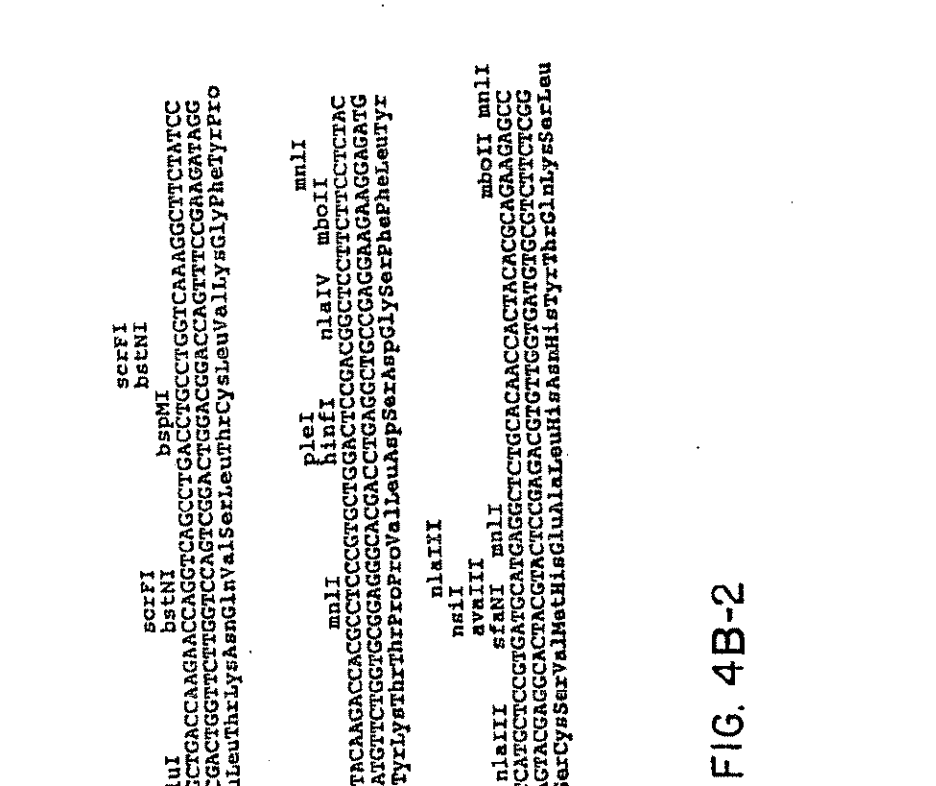

scrFI
ncil
mspI
hpaII
smaI
scrFI
ncil
avaI
fokI fokI aluI
rsaI
fnu4HI
bbvI avaI
801 GGCAGCCCCGAGAACCACAGGTGTACACCCTGCCCCCATCCCGGGATGAGCTGACCAAGAACCAGGTCAGCCTGACCTGCCTGGTCAAAGGCTTCTATCC
CCGTCGGGGCTCTTGGTGTCCACATGTGGGACGGGGGTAGGGCCCTACTCGACTGGTTCTTGGTCCAGTCGGACTGGACCGGACCAGTTTCCGAAGATAGG
337 GlnProArgGluProGlnValTyrThrLeuProProSerArgAspGluLeuThrLysAsnGlnValSerLeuThrCysLeuValLysGlyPheTyrPro
mspI
hpaII
fnu4HI
bbvI
901 CAGCGACATCGCCGTGGAGTGGGAGAGCAATGGGCAGCCGGAGAACAACTACAAGACCACGCCTCCCGTGCTGGACTCCGACGGCTCCTTCTTCCTCTAC
GTCGCTGTAGCGGCACCTCACCCTCTCGTTACCCGTCGGCCCTCTTGTTGATGTTCTGGGTGCGGAGGGCACGACCTGAGGCTGCCCGAGGAAGAAGGAGATG
370 SerAspIleAlaValGluTrpGluSerAsnGlyGlnProGluAsnAsnTyrLysThrThrProProValLeuAspSerAspGlySerPhePheLeuTyr
nlaIII
ns1I
ava1II
sfaNI mnlI
fnu4HI
bbvI xmnI mboII nlaIII
bspMI
aluI
1001 AGCAAGCTCACCGTGGACAAGAGCAGGTGGCAGCAGGGGAACGTCTTCTCATGCTCCGTGATGCATGAGGCTCTGCACAACCACTACACGCAGAAGAGCCC
TCGTTCGAGTGGCACCTGTTCTCGTCCACCGTCGTCCCCTTGCAGAAGAGTACGAGGCCACTACGTACTCCGAGACGTGTTGGTGATGTGCGTCTTCTCGG
403 SerLysLeuThrValAspLysSerArgTrpGlnGlnGlyAsnValPheSerCysSerValMetHisGluAlaLeuHisAsnHisTyrThrGlnLysSerLeu
mspI
hpaII
scrFI
ncil
mspI
hpaII
haeIII
xmaIII
eaeI
1101 TCTCCCTGTCTCCGGGTAAATGAGTGCGACGGCCG
AGAGGGACAGAGGCCCATTTACTCACGCTGCCGGC
437 SerLeuSerProGlyLysOP*

FIG. 4B-2

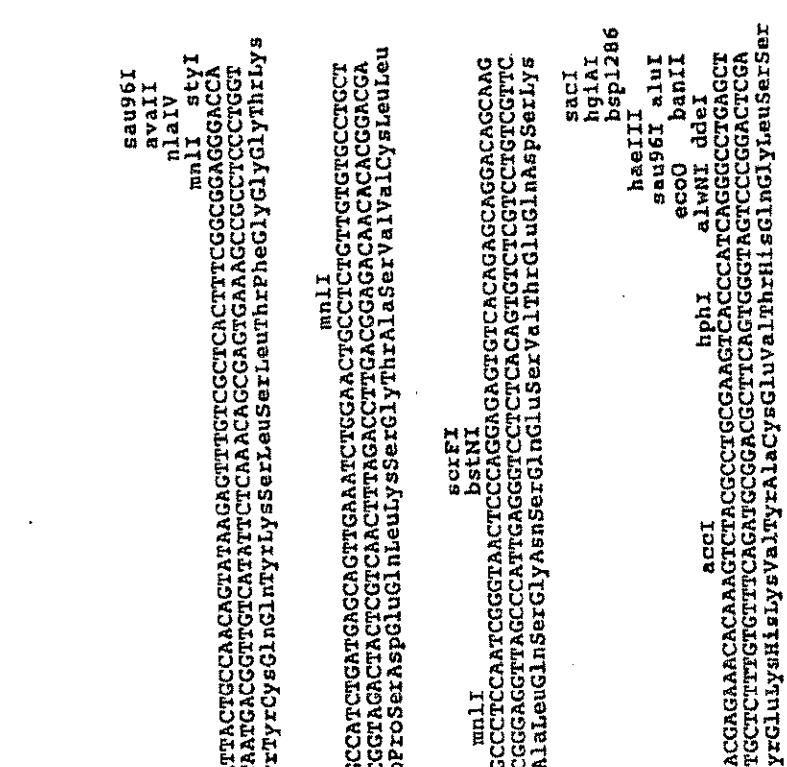

sau96I
avaII
nlaIV
mnlI styI
fnu4HI
pstI
haeIII
fnu4HI bbvI mboII
ecoRI hphI
1 GAATTCACTCTCACCATCAGCGGCCTGCAGCCTGAAGATTTTGCAACTTATTACTGCCAACAGTATAAGAGTTTGTCGCTCACTTTCGGCGGAGGGACCA
CTTAAGTGAGAGTGGTAGTCGCCGGACGTCGGACTTCTAAAACGTTGAATAATGACGGTGTCATATTCTCAAACAGCGAGTGAAAGCCGCCTCCCTGGT
72 ThrLeuThrIleSerGlyLeuGlnProGluAspPheAlaThrTyrTyrCysGlnGlnTyrLysSerLeuSerLeuThrPheGlyGlyGlyThrLys
sau3AI fnu4HI mnlI
dpnI bbvI mboII mboII
101 AGGTGGAGATCAAACGAACTGTGGCTGCACCATCTGTCTTCATCTTCCCGCCATCTGATGAGCAGTTGAAATCTGGAACTGCCTCTGTTGTGTGCCTGCT
TCCACCTCTAGTTTGCTTGACACCGACGTGGTAGACAGAAGTAGAAGGGCGGTAGACTACTCGTCAACTTTAGACCTTGACGGAGACAACACACGGACGA
104 ValGluIleLysArgThrValAlaAlaProSerValPheIlePheProProSerAspGluGlnLeuLysSerGlyThrAlaSerValValCysLeuLeu
scrFI
bstNI
mnlI
xmnI
201 GAATAACTTCTATCCCAGAGAGGCCAAAGTACAGTGGAAGGTGGATAACGCCCTCCAATCGGGTAACTCCCAGGAGAGTGTCACAGAGCAGGACAGCAAG
137 AsnAsnPheTyrProArgGluAlaLysValGlnTrpLysValAspAsnAlaLeuGlnSerGlyAsnSerGlnGluSerValThrGluGlnAspSerLys
saeI
hgiAI
bsp1286
haeIII
sau96I aluI
ecoO banII
hphI alwNI ddeI
fnu4HI
ddeI ddeI
mnlI bbvI hgaI
301 GACAGCACCTACAGCCTCAGCAGCACCCTGACGCTGAGCAAAGCAGACTACGAGAAACACAAAGTCTACGCCTGCGAAGTCACCCATCAGGGCCTGAGCT
CTGTCGTGGATGTCGGAGTCGTCGTGGGACTGCGACTCGTTTCGTCTGATGCTCTTTGTGTTTCAGATGCGGACGCTTCAGTGGGTAGTCCCGGACTCGA
170 AspSerThrTyrSerLeuSerSerThrLeuThrLeuSerLysAlaAspTyrGluLysHisLysValTyrAlaCysGluValThrHisGlnGlyLeuSerSer
ddeI
aluI mnlI bsp1286 bspMI mnlI
401 CGCCCGTCACAAAGAGCTTCAACAGGGGAGAGTGTTAGAGGGAGAAGTGCCCCCACCTGCTCCTCAGT
GCGGGCAGTGTTTCTCGAAGTTGTCCCCTCTCACAATCTCCCTCTTCACGGGGGTGGACGAGGAGTCA
204 ProValThrLysSerPheAsnArgGlyGluCysAM*

FIG. 5

ADHESON VARIANTS

This application is a continuation application under 37 C.F.R. §1.53(b) claiming priority to co-pending application Ser. No. 09/641,554, filed on Aug. 17, 2000 now abandoned, which is a continuation of application Ser. No. 09/275,310, filed on Mar. 24, 1999, now abandoned which is a continuation of application Ser. No. 08/457,918, filed on Jun. 1, 1995 (U.S. Pat. No. 6,117,655), which is a continuation of application Ser. No. 08/236,311, filed on May 2, 1994 (U.S. Pat. No. 5,565,335), which is a continuation of application Ser. No. 07/936,190, filed on Aug. 26, 1992 (U.S. Pat. No. 5,336,603), which is a divisional of application Ser. No. 07/842,777 filed on Feb. 18, 1992 (now abandoned), which is a continuation of application Ser. No. 07/250,785, filed Sep. 28, 1988 (now abandoned) which is a continuation-in-part of application Ser. No. 07/104,329, filed Oct. 2, 1987 (now abandoned) which applications are incorporated herein by reference.

BACKGROUND OF THE INVENTION

This application relates to compositions for antiviral or immunomodulatory therapy. In particular, it relates to compositions useful in the treatment of Human Immunodeficiency Virus (HIV) infections.

The primary immunologic abnormality resulting from infection by HIV is the progressive depletion and functional impairment of T lymphocytes expressing the CD4 cell surface glycoprotein (H. Lane et al., Ann. Rev. Immunol. 3:477 [1985]). CD4 is a non-polymorphic glycoprotein with homology to the immunoglobulin gene superfamily (P. Maddon et al., Cell 42:93 [1985]). Together with the CD8 surface antigen, CD4 defines two distinct subsets of mature peripheral T cells (E. Reinherz et al., Cell 19:821 [1980]), which are distinguished by their ability to interact with nominal antigen targets in the context of class I and class II major histocompatibility complex (MHC) antigens, respectively (S. Swain, Proc. Natl. Acad. Sci. 78:7101 [1981]; E. Engleman et al., J. Immunol. 127:2124 [1981]; H. Spitz et al., J. Immunol. 129:1563 [1982]; W. Biddison et al., J. Exp. Med. 156:1065 [1982]; and D. Wilde et al., J. Immunol. 131:2178 [1983]). For the most part, CD4 T cells display the helper/inducer T cell phenotype (E. Reinherz, supra), although CD4 T cells characterized as cytotoxic/suppressor T cells have also been identified (Y. Thomas et al., J. Exp. Med. 154:459 [1981]; S. Meuer et al., Proc. Natl. Acad. Sci. USA 79:4395 [1982]; and A. Krensky et al., Proc. Natl. Acad. Sci. USA 79:2365 [1982]). The loss of CD4 helper/inducer T cell function probably underlies the profound defects in cellular and humoral immunity leading to the opportunistic infections and malignancies characteristic of the acquired immunodeficiency syndrome (AIDS) (H. Lane supra).

Studies of HIV-I infection of fractionated CD4 and CD8 T cells from normal donors and AIDS patients have revealed that depletion of CD4 T cells results from the ability of HIV-I to selectively infect, replicate in, and ultimately destroy this T lymphocyte subset (D. Klatzmann et al., Science 225:59 [1984]). The possibility that CD4 itself is an essential component of the cellular receptor for HIV-I was first indicated by the observation that monoclonal antibodies directed against CD4 block HIV-I infection and syncytia induction (A. Dalgleish et al., Nature [London] 312:767 [1984]; J. McDougal et al., J. Immunol. 135:3151 [1985]). This hypothesis has been confirmed by the demonstration that a molecular complex forms between CD4 and gp120,

the major envelope glycoprotein of HIV-I (J. McDougal et al., Science 231:382 [1986]; and the finding that HIV-I tropism can be conferred upon ordinarily non-permissive human cells following the stable expression of a CD4 cDNA (P. Maddon et al., Cell 47:333 [1986]). Furthermore, the neurotropic properties of HIV-I, reflected by a high incidence of central nervous system dysfunction in HIV-I infected individuals (W. Snider et al., Ann. Neurol. 14:403 [1983]), and the ability to detect HIV-I in the brain tissue and cerebrospinal fluid of AIDS patients (G. Shaw et al., Science 227:177 [1985]; L. Epstein, AIDS Res. 1:447 [1985]; S. Koenig, Science 233:1089 [1986]; D. Ho et al., N. Engl. J. Med. 313:1498 [1985]; J. Levy et al., Lancet II:586 [1985]), appears to have its explanation in the expression of CD4 in cells of neuronal, glial and monocyte/macrophage origin (P. Maddon, Cell 47:444 [1986]; I. Funke et al., J. Exp. Med. 165:1230 [1986]; B. Tourvieille et al., Science 234:610 [1986]).

In addition to determining the susceptibility to HIV-I infection, the manifestation of cytopathic effects in the infected host cell appears to involve CD4. Antibody to CD4 was found to inhibit the fusion of uninfected CD4 T cells with HIV-I infected cells in vitro; moreover, the giant multinucleated cells produced by this event die shortly after being formed resulting in the depletion of the population of CD4 cells (J. Lifson et al., Science 232:1123 [1986]). Formation of syncytia also requires gp120 expression, and can be elicited by coculturing CD4-positive cell lines with cell lines expressing the HIV-I env gene in the absence of other viral structural or regulatory proteins (J. Sodroski et al., Nature 322:470 [1986]; J. Lifson et al., Nature 323:725 [1986]). Thus, in mediating both the initial infection by HIV-I as well as eventual cell death, the interaction between gp120 and CD4 constitutes one of several critical entry points in the viral life cycle amenable to therapeutic intervention (H. Mitsuya et al., Nature 325:773 [1987]).

The known sequence of the CD4 precursor predicts a hydrophobic signal peptide, an extracellular region of approximately 370 amino acids, a highly hydrophobic stretch with significant identity to the membrane-spanning domain of the class II MHC beta chain, and a highly charged intracellular sequence of 40 residues (P. Madden, Cell 42:93 [1985]). The extracellular domain of CD4 consists of four contiguous regions each having amino acid and structural similarity to the variable and joining (V–J) domains of immunoglobulin light chains as well as related regions in other members of the immunoglobulin gene superfamily (a subclass of which are defined herein by the coined term "adhesons". These structurally similar regions of CD4 are termed the V_1, V_2, V_3 and V_4 domains (denominated 1–4 in FIG. 3).

A successful strategy in the development of drugs for the treatment of many receptor mediated abnormalities has been the identification of antagonists which block binding of the natural ligand. Since the CD4 adheson ordinarily binds to the recognition sites of the HIV envelope it would appear to be a candidate for therapeutically sequestering these HIV sites, thereby blocking viral infectivity. However, full length CD4 and other adhesons are cell membrane proteins which are anchored in the lipid bilayer of cells. The presence of membrane components will be undesirable from the standpoint of manufacturing and purification. In addition, since adhesons are normally present only on cell surfaces, it would be desirable to produce adhesons in a form which is more stable in the circulation. Additionally, even truncated, soluble CD4 adheson (generally referred to as CD4T) may not be optimally effective as a therapeutic since it possesses

a relatively short biological half-life, binds to HIV no better than cell surface CD4, may not cross the placental or other biological barriers and since it merely sequesters the HIV recognition sites without in itself bearing an infected-cell killing or virus killing functionality.

Accordingly, it is an object of this invention to produce soluble, secreted adhesons. It is another object to produce CD4 derivatives useful in the treatment of AIDS and related conditions, in a manner essentially unaffected by the extreme degree of genetic variation observed among various HIV-I isolates and their respective env polypeptides (J. Coffin, Cell 46:1 [1986]). Still another object is to prepare adhesons fused to other polypeptides in order to provide molecules with novel functionalities such as those described above for therapeutic use, or diagnostic reagents for the in vitro assay of adhesons or their ligands. In particular, it is an objective to prepare molecules for directing toxins or effector molecules (for example the Fc domain of immunoglobulin) to cells bearing receptors for the adhesons, e.g. HIV gp120 in the case of CD4, and for use in facilitating purification of the adhesons. It is a further object to provide stable, highly purified adheson preparations.

SUMMARY

The objects of this invention are accomplished by providing nucleic acid encoding an amino acid sequence variant of an adheson, in particular a variant in which the transmembrane domain is modified so that it is no longer capable of becoming lodged in the cell membrane. In the case of CD4 such variants are termed soluble CD4.

Variant adhesons are produced by a method comprising (a) transforming a host cell with nucleic acid encoding an amino acid sequence variant of an adheson, (b) culturing the host cell and (c) recovering the variant adheson from the host cell culture media or from lysates of the host cell.

In specific embodiments, the objects of this invention are accomplished by providing an adheson variant selected from the group consisting of (a) an adheson amino acid sequence variant having an inactivated transmembrane domain and (b) a polypeptide comprising an adheson extracellular domain fused to the sequence of a polypeptide which is different from the adheson, this latter, for example, selected from a cytotoxin, an immunogen or a protein with a long plasma half life such as an immunoglobulin constant domain.

In a preferred embodiment a polypeptide comprising a gp120 binding domain of the CD4 adheson is fused at its C-terminus to an immunoglobulin constant domain, or is linked to a cytotoxic polypeptide such as ricin.

The CD4 adheson variants provided herein are purified and formulated in pharmacologically acceptable vehicles for administration to patients in need of antiviral, neuromodulatory or immunomodulatory therapy, in particular patients infected with HIV, and for use in the modulation of cell adhesion.

BRIEF DESCRIPTION OF THE DRAWINGS

FIGS. *1a–1c* depict the amino acid and nucleotide sequence of a secreted form of the CD4 adheson. The signal processing site is designated with an arrow.

FIGS. *2a–2c* depict the amino acid and nucleotide sequence of a fusion of the herpes gD leader and N-terminal 27 residues to the putative mature N-terminus of CD4T.

FIG. 3 depicts the structural elements of the native and soluble CD4 adheson, the native human IgG_1 (γ_1) heavy chain and two exemplary heavy chain-CD4 chimeras.

FIGS. 4A–4B–2 are a map of the linkered human IgG_1 (γ_1) chain fragment employed in the preparation of CD4 fusions. Insert sites are designated γ1 and Fc.

FIG. 5 is a map of the human κ light chain fragment useful for CD4 fusions at the arrow flanked by $V_\kappa J_\kappa$ (light variable and joining) and C_κ (light constant).

DETAILED DESCRIPTION

Adhesons are cell surface polypeptides having an extracellular domain which is homologous to a member of the immunoglobulin gene superfamily, excluding, however, highly polymorphic members of this superfamily selected from the group of class I and class II major histocompatibility antigens, immunoglobulins and T-cell receptor α, β, γ and δ chains. Examples of adhesons include CD1, CD2, CD4, CD8, CD28, the γ, δ and ε chains of CD3, OX-2, Thy-1, the intercellular or neural cell adhesion molecules (I-CAM or N-CAM), lymphocyte function associated antigen-3 (LFA-3), neurocytoplasmic protein (NCP-3), poly-Ig receptor, myelin-associated glycoprotein (MAG), high affinity IgE receptor, the major glycoprotein of peripheral myelin (Po), platelet derived growth factor receptor, colony stimulating factor-1 receptor, macrophage Fc receptor, Fc gamma receptors and carcinoembryonic antigen. Homologous as defined herein means having the sequence of a member of the immunoglobulin gene superfamily or having a sequence therewithin which has substantially the same as (or a greater degree of) amino acid sequence homology to a known member of the superfamily as the specific examples given above have to the sequence of an immunoglobulin variable or constant domain. Preferred adhesons are CD4, CD8 and high affinity IgE Fc receptor.

This invention is particularly concerned with amino acid sequence variants of adhesons. Amino acid sequence variants of adhesons are prepared with various objectives in mind, including increasing the affinity of the adheson for its binding partner, facilitating the stability, purification and preparation of the adheson, increasing its plasma half life, improving therapeutic efficacy as described above in the background, introducing additional functionalities and lessening the severity or occurrence of side effects during therapeutic use of the adheson. Amino acid sequence variants of adhesons fall into one or a combination of the following classes: insertional, substitutional or deletional variants.

Insertional amino acid sequence variants are those in which one or more amino acid residues extraneous to the adheson are introduced into a predetermined site in the adheson including the C or N termini. Such variants are referred to as fusions of the adheson and a different polypeptide. Such other polypeptides contain sequences other than those which are normally found in the adheson at the inserted position. Several groups of fusions are contemplated herein. Immunologically active adheson fusions comprise an adheson and a polypeptide containing a non-adheson epitope. The non-adheson epitope is any immunologically competent polypeptide, i.e., any polypeptide which is capable of eliciting an immune response in the animal to which the fusion is to be administered or which is capable of being bound by an antibody raised against the non-adheson polypeptide. Typical non-adheson epitopes will be those which are borne by allergens, autoimmune epitopes, or other potent immunogens or antigens recognized by pre-existing antibodies in the fusion recipient, including bacterial polypeptides such as trpLE, beta-

galactosidase, viral polypeptides such as herpes gD protein, and the like. Immunogenic fusions are produced by cross-linking in vitro or by recombinant cell culture transformed with DNA encoding an immunogenic polypeptide. It is preferable that the immunogenic fusion be one in which the immunogenic sequence is joined to or inserted into the adheson antigen or fragment thereof by a peptide bond(s). These products therefore consist of a linear polypeptide chain containing adheson epitopes and at least one epitope foreign to the adheson. It will be understood that it is within the scope of this invention to introduce the epitopes anywhere within the adheson molecule or fragment thereof. Such fusions are conveniently made in recombinant host cells or by the use of bifunctional cross-linking agents. The use of a cross-linking agent to fuse the adheson to the immunogenic polypeptide is not as desirable as a linear fusion because the cross-linked products are not as easily synthesized in structurally homogeneous form.

These immunogenic insertions are particularly useful when formulated into a pharmacologically acceptable carrier and administered to a subject in order to raise antibodies against the adheson, which antibodies in turn are useful in diagnostics or in purification of adheson by immunoaffinity techniques known per se. Alternatively, in the purification of adhesons, binding partners for the fused non-adheson polypeptide, e.g. antibodies, receptors or ligands, are used to adsorb the fusion from impure admixtures, after which the fusion is eluted and, if desired, the adheson is recovered from the fusion, e.g. by enzymatic cleavage.

Other fusions, which may or may not also be immunologically active, include fusions of the adheson sequence with a signal sequence heterologous to the adheson, fusions of transmembrane-modified CD4 adhesons, for example, to polypeptides having enhanced plasma half life (ordinarily>about 20 hours) such as immunoglobulin chains or fragments thereof, and fusions with cytotoxic functionalities. Signal sequence fusions are employed in order to more expeditiously direct the secretion of the adheson. The heterologous signal replaces the native adheson signal, and when the resulting fusion is recognized, i.e. processed and cleaved by the host cell, the adheson is secreted. Signals are selected based on the intended host cell, and may include bacterial yeast, mammalian and viral sequences. The herpes gD glycoprotein signal is suitable for use in mammalian expression systems.

Plasma proteins which have enhanced plasma half-life longer than that of transmembrane modified CD4 include serum albumin, immunoglobulins, apolipoproteins, and transferrin. Preferably, the adheson-plasma protein fusion is not significantly immunogenic in the animal in which it is used and the plasma protein does not cause undesirable side effects in patients by virtue of its normal biological activity.

In a specific embodiment the adheson immunoglobulin-like domain which may be homologous either to the constant or to the variable region domains is conjugated with an immunoglobulin constant region sequence. The resulting products are referred to herein as immunoadhesons. Immunoglobulins and certain variants thereof are known and many have been prepared in recombinant cell culture. For example, see U.S. Pat. No. 4,745,055; EP 256,654; Faulkner et al., Nature 298:286 (1982); EP 120,694; EP 125,023; Morrison, J. Immun. 123:793 (1979); Köhler et al., P.N.A.S. USA 77:2197 (1980); Raso et al., Cancer Res. 41:2073 (1981); Morrison et al., Ann. Rev. Immunol. 2:239 (1984); Morrison, Science 229:1202 (1985); Morrison et al., P.N.A.S. USA 81:6851 (1984); EP 255,694; EP 266,663; and WO 88/03559. Reassorted immunoglobulin chains also are known. See for example U.S. Pat. No. 4,444,878; WO 88/03565; and EP 68,763 and references cited therein.

Ordinarily, the domains of adhesons that are homologous to immunoglobulins and extracellular in their native environment are fused C-terminally to the N-terminus of the constant region of immunoglobulins in place of the variable region(s) thereof, retaining at least functionally active hinge, CH2 and CH3 domains of the constant region of an immunoglobulin heavy chain. This ordinarily is accomplished by constructing the appropriate DNA sequence and expressing it in recombinant cell culture. Immunoglobulins and other polypeptides having enhanced plasma half life are fused to the extracellular or ligand binding domains of other adhesons in the same fashion.

The boundary domains for the CD4 V-like regions (V1–V4) are, respectively, about 100–109, about 175–184, about 289–298, and about 360–369 (based on the precursor CD4 amino acid sequence in which the initiating met is –25; FIG. 1*a*). CD4 sequences containing any of the CD4 V domains are fused to the immunoglobulin sequence. It is preferable that the V1V2 or V1V2V3V4 be fused at their C-termini to the immunoglobulin constant region. The precise site at which the fusion is made is not critical; the boundary domains noted herein are for guidance only and other sites neighboring or within the V regions may be selected in order to optimize the secretion or binding characteristics of the CD4. The optimal site will be determined by routine experimentation. In general, it has been found that the fusions are expressed intracellularly, but a great deal of variation is encountered in the degree of secretion of the fusions from recombinant hosts. For instance, the following table demonstrates the various immunoglobulin fusions that have been obtained by the method of this invention. In all examples of CD4 immunoadhesons, the CD4 signal was used to direct secretion from 293 cells. Lower case m represents murine origin, while the lower case h designates human origin. V and C are abbreviations for immunoglobulin variable and constant domains respectively. The numerical subscripts indicate the number of parenthetical units found in the designated multimer. It will be understood that the chains of the multimers are believed to be disulfide bonded in the same fashion as native immunoglobulins. The CD4 immunoadhesons typically contained either the first N-terminal 366 residues of CD4 ($CD4_4$) or the first 180 N-terminal residues of CD4 ($CD4_2$) linked at their C-terminus to the κ (light) chain or IgG1 heavy chain constant region (γ1).

TABLE I

Transfected Gene	Secreted Product
$mV_\kappa C_\kappa$	$mV_\kappa C_\kappa$ and/or $(mV_\kappa C_\kappa)_2$
$mV_{\gamma 1}C_{\gamma 1}$	ND
$mV_\kappa C_\kappa + mV_{\gamma 1}C_{\gamma 1}$	$(mV_\kappa C_\kappa)_2(mV_{\gamma 1}C_{\gamma 1})_2$ + $mV_\kappa C_\kappa$ and/or $(mV_\kappa C_\kappa)_2$
$hCD4\text{-}mC_\kappa$	$hCD4\text{-}mC_\kappa$ and/or $(hCD4\text{-}mC_\kappa)_2$
$hCD4\text{-}mC_{\gamma 1}$	ND
$hCD4\text{-}mC_\kappa + hCD4\text{-}mC_{\gamma 1}$	$(hCD4\text{-}mC_\kappa)_2(hCD4\text{-}mC_{\gamma 1})_2$ + $hCD4\text{-}mC_\kappa$ and/or $(hCD4\text{-}mC_\kappa)_2$
$hCD4\text{-}hC_\kappa$	$hCD4\text{-}hC_\kappa$ and/or $(hCD4\text{-}hC_\kappa)_2$
$hCD4\text{-}hC_{\gamma 1}$	$(hCD4\text{-}hC_{\gamma 1})_2$
$hCD4\text{-}hC_\kappa + hCD4\text{-}hC_{\gamma 1}$	$(hCD4\text{-}hC_\kappa)_2(hCD4\text{-}hC_{\gamma 1})_2$ + $hCD4\text{-}hC_\kappa$ and/or $(hCD4\text{-}hC_\kappa)_2$
$mV_\kappa C_\kappa + hCD4\text{-}hC_{\gamma 1}$	$(mV_\kappa C_\kappa)_2(hCD4\text{-}hC_{\gamma 1})_2$ + $mV_\kappa C_\kappa$ and/or $(mV_\kappa C_\kappa)_2$

* ND = Not detected

It is interesting to observe from this table that the CD4-human heavy chain immunoadheson was secreted as a dimer

whereas the analogous murine construction was not detected (this not excluding the intracellular accumulation of the protein, however). The ability of the hCD4-hCγ1 transformants to produce heavy chain dimer was unexpected since previous work had suggested that immunoglobulin heavy chains are not secreted unless the hosts are cotransformed with nucleic acid encoding both heavy and light chain (Valle et al., *Nature* 241:338 [1981]). According to this invention, CD4-IgG immunoadheson chimeras are readily secreted wherein the CD4 epitope is present in heavy chain dimers, light chain monomers or dimers, and heavy and light chain heterotetramers wherein the CD4 epitope is present fused to one or more light or heavy chains, including heterotetramers wherein up to and including all four variable region analogues are derived from CD4. Where light-heavy chain non-CD4 variable domain is present, a heterofunctional antibody thus is provided.

Various exemplary hetero- and chimeric immunoadheson antibodies produced in accordance with this invention are schematically diagrammed below. "A" means at least a portion of the extracellular domain of an adheson containing its ligand binding site; V_L, V_H, C_L and C_H represent light or heavy chain variable or constant domains of an immunoglobulin; n is an integer; and Y designates a covalent cross-linking moiety.

(a) AC_L;

(b) AC_L-AC_L;

(c) AC_H-[AC_H, AC_L-AC_H, AC_L-V_HC_H, V_LC_L-AC_H, or V_LC_L-V_HC_H];

(d) AC_L-AC_H-[AC_H, AC_L-AC_H, AC_L-V_HC_H, V_LC_L-AC_H, or V_LC_L-V_HC_H];

(e) AC_L-V_HC_H-[AC_H, AC_L-AC_H, AC_L-V_HC_H, V_LC_L-AC_H, or V_LC_L-V_HC_H];

(f) V_LC_L-AC_H-[AC_H, AC_L-AC_H, AC_L-V_HC_H, V_LC_L-AC_H, or V_LC_L-V_HC_H]; or

(g) [A-Y]$_n$-[V_LC_L-V_HC_H]$_2$.

The structures shown in this table show only key features, e.g. they do not show joining (J) or other domains of the immunoglobulins, nor are disulfide bonds shown. These are omitted in the interests of brevity. However, where such domains are required for binding activity they shall be construed as being present in the ordinary locations which they occupy in the adheson, immunoadheson or immunoglobulin molecules as the case may be. These examples are representative of divalent antibodies; more complex structures would result by employing immunoglobulin heavy chain sequences from other classes, e.g. IgM. The immunoglobulin V_LV_H antibody combining site also designated as the companion immunoglobulin, preferably is capable of binding to a predetermined antigen.

Suitable companion immunoglobulin combining sites and fusion partners are obtained from IgG-1, -2, -3, or -4 subtypes, IgA, IgE, IgD or IgM, but preferably IgG-1.

A preferred embodiment is a fusion of an N-terminal portion of CD4, which contains the binding site for the gp120 envelope protein of HIV, to the C-terminal F_c portion of an antibody, containing the effector functions of immunoglobulin G_1. There are two preferred embodiments of this sort; in one, the entire heavy chain constant region is fused to a portion of CD4; in another, a sequence beginning in the hinge region just upstream of the papain cleavage site which defines IgG F_c chemically (residue 216, taking the first residue of heavy chain constant region to be 114 [Kobat et al., "Sequences of Proteins of Immunological Interest" 4th Ed., 1987], or analogous sites of other immunoglobulins) is fused to a portion of CD4. These embodiments are described in the examples.

More particularly, those variants in which one or more immunoglobulin-like domains of an adheson are substituted for the variable region of an immunoglobulin chain are believed to exhibit improved in vivo plasma half life. These chimeras are constructed in a fashion similar to chimeric antibodies in which a variable domain from an antibody of one species is substituted for the variable domain of another species. See, for example, EP 0 125 023; Munro, Nature 312: (Dec. 13, 1984); Neuberger et al., Nature 312: (Dec. 13, 1984); Sharon et al., Nature 309: (May 24, 1984); Morrison et al., Proc. Natl. Acad. Sci. USA 81:6851–6855 (1984); Morrison et al. Science 229:1202–1207 (1985); and Boulianne et al., Nature 312:643–646 (Dec. 13, 1984). The DNA encoding the adheson immunoglobulin-like domain(s) is cleaved by a restriction enzyme at or proximal to the 3' end of the DNA encoding the immunoglobulin-like domain(s) and at a point at or near the DNA encoding the N-terminal end of the mature adheson polypeptide (where use of a different leader is contemplated) or at or proximal to the N-terminal coding region for the adheson (where the native adheson signal is employed). This DNA fragment then is readily inserted into DNA encoding an immunoglobulin light or heavy chain constant region and, if necessary, tailored by deletional mutagenesis. Preferably, this is a human immunoglobulin when the variant is intended for in vivo therapy for humans. DNA encoding immunoglobulin light or heavy chain constant regions is known or readily available from cDNA libraries or is synthesized. See for example, Adams et al., Biochemistry 19:2711–2719 (1980); Gough et al., Biochemistry 19:2702–2710 (1980); Dolby et al., P.N.A.S. USA, 77:6027–6031 (1980); Rice et al., P.N.A.S. USA 79:7862–7865 (1982); Falkner et. al., Nature 298:286–288 (1982); and Morrison et al., Ann. Rev. Immunol. 2:239–256 (1984).

DNA encoding the immunoglobulin or immunoadheson chimeric chain(s) is transfected into a host cell for expression. If the host cell is producing an immunoglobulin prior to transfection then one need only transfect with the adheson fused to light or to heavy chain to produce a heteroantibody. The aforementioned immunoglobulins having one or more arms bearing the adheson domain and one or more arms bearing companion variable regions result in dual specificity for adheson ligand and for an antigen. These are produced by the above-described recombinant methods or by in vitro procedures. In the latter case, for example, $F(ab')_2$ fragments of the adheson fusion and an immunoglobulin are prepared, the $F(ab')_2$ fragments converted to Fab' fragments by reduction under mild reducing conditions, and then reoxidized in each other's presence under acidic conditions in accord with methods known per se. See also U.S. Pat. No. 4,444,878. Additionally, procedures are known for producing intact heteroantibodies from immunoglobulins having different specificities. These procedures are adopted for the in vitro production of heterochimeric antibodies by simply substituting the immunoadheson chains for one of the previously employed immunoglobulins.

In an alternative method for producing a heterofunctional antibody, host cells producing an adheson-immunoglobulin fusion, e.g. transfected myelomas, also are fused with B cells or hybridomas which secrete antibody having the desired companion specificity for an antigen. Heterobifunctional antibody is recovered from the culture medium of such hybridomas, and thus may be produced somewhat more conveniently than by conventional in vitro resorting methods (EP 68,763).

Another group of fusions are those in which an adheson is conjugated with a toxic substance, e.g. a polypeptide such

as ricin (including deglycosylated ricin A chain), diptheria toxin A, or a non-peptidyl cytotoxin. Where the toxin is a polypeptide it is convenient to cross-link the polypeptide to the adheson or its transmembrane-deleted variant by conventional in vitro protein cross-linking agents (for suitable methods for linking ricin A chain or deglycosylated A chain to CD4 see, for example, Duncan et al., "Analy. Biochem." 132:68–73 [1983]; Thorpe et al., "Cancer Res." 47:5924 [1987]; and Ghotie et al., "Cancer Res." 48:2610 [1988]) or by recombinant synthesis as a fusion (see for example, U.S. Pat. No. 4,765,382). Alternatively, where companion antibodies are anti-ricin antibody immunoglobulin variable domains, such immunoglobulin heteroantibodies are employed to deliver ricin to HIV infected cells following the general procedure of Raso et al., Cancer Research, 41:2073 (1981).

Another class of adheson variants are deletional variants. Deletions are characterized by the removal of one or more amino acid residues from a adheson sequence. Typically, the transmembrane and cytoplasmic domains of adhesons ar deleted. In the case of CD4, at least residues 368 to 395 (the transmembrane region), and ordinarily 396–433 as well (the cytoplasmic domain), will be deleted to obtain secreted forms of this adheson. Parenthetically, the amino acid residues follow the numbers given for mature CD4 as noted, for example, in FIGS. 1*a*–1*c*. Thus, CD4T molecules generally will terminate in the vicinity of about residues 366–368, or at any other suitable site N-terminal thereto which preserves the gp120-binding capability of the CD4 variant.

Substitutional variants are those in which at least one residue in the adheson sequence has been removed and a different residue inserted in its place. The native N-terminal residue for mature CD4 is now known to be lysine. Thus, the sequence shown in FIG. 1, with an N-terminal asparagine, is an amino acid sequence variant of native mature CD4. Table 2 below describes substitutions which in general will result in fine modulation of the characteristics of the CD antigen.

TABLE 2

Original Residue	Exemplary Substitutions
Ala	ser
Arg	lys
Asn	gln; his
Asp	glu
Cys	ser; ala
Gln	asn
Glu	asp
Gly	pro
His	asn; gln
Ile	leu; val
Leu	ile; val
Lys	arg; gln; glu
Met	leu; ile
Phe	met; leu; tyr
Ser	thr
Thr	ser
Trp	tyr
Tyr	trp; phe
Val	ile; leu

Substantial changes in function or immunological identity are made by selecting substitutions that are less conservative than those in Table 2, i.e., selecting residues that differ more significantly in their effect on maintaining (a) the structure of the polypeptide backbone in the area of the substitution, for example as a sheet or helical conformation, (b) the charge or hydrophobicity of the molecule at the target site or (c) the bulk of the side chain. The substitutions which in general are expected to produce the greatest changes in adheson properties will be those in which (a) a hydrophilic residue, e.g. seryl or threonyl, is substituted for (or by) a hydrophobic residue, e.g. leucyl, isoleucyl, phenylalanyl, valyl or alanyl; (b) a cysteinyl or prolyl is substituted for (or by) any other residue; (c) a residue having an electropositive side chain, e.g., lysyl, arginyl, or histidyl, is substituted for (or by) an electronegative residue, e.g., glutamyl or aspartyl; or (d) a residue having a bulky side chain, e.g., phenylalanyl, is substituted for (or by) one not having a side chain, e.g., glycyl.

A preferred class of substitutional or deletional variants are those involving the transmembrane region of the adheson. The transmembrane region of the adheson is a highly hydrophobic or lipophilic domain that is the proper size to span the lipid bilayer of the cellular membrane. It is believed to anchor the adheson in the cell membrane.

Deletion or substitution of the transmembrane domain will facilitate recovery and provide a soluble form of the adheson by reducing its cellular or membrane lipid affinity and improving its water solubility. If the transmembrane and cytoplasmic domains are deleted one avoids the introduction of potentially immunogenic epitopes, either by exposure of otherwise intracellular polypeptides that might be recognized by the body as foreign or by insertion of heterologous polypeptides that are potentially immunogenic. A principal advantage of the transmembrane deleted adheson is that it is secreted into the culture medium of recombinant hosts. This variant is water soluble and does not have an appreciable affinity for cell membrane lipids, thus considerably simplifying its recovery from recombinant cell culture.

It will be amply apparent from the foregoing discussion that substitutions, deletions, insertions or any combination thereof are introduced to arrive at a final construct. As a general proposition, all variants will not have a functional transmembrane domain and preferably will not have a functional cytoplasmic sequence. This is generally accomplished by deletion of the relevant domain, although adequate insertional or substitutional mutagens also can be effective for this purpose. For example, the transmembrane domain is substituted by any amino acid sequence, e.g. a random or homopolynucleic sequence of about 5 to 50 serine, threonine, lysine, arginine, glutamine, aspartic acid and like hydrophilic residues, which altogether exhibit a hydrophilic hydropathy profile, so that it is secreted into the culture medium of recombinant hosts. This variant should also be considered to be an adheson variant.

These variants ordinarily are prepared by site specific mutagenesis of nucleotides in the DNA encoding the adheson, thereby producing DNA encoding the variant, and thereafter expressing the DNA in recombinant cell culture. However, variant adhesons also are prepared by in vitro synthesis. Obviously, variations made in the DNA encoding the variant adhesons must not place the sequence out of reading frame and preferably will not create complementary regions that could produce secondary mRNA structure deleterious to expression (EP 75,444A). The CD4 variants typically exhibit the same gp120 binding activity as does the naturally-occurring prototype, although variants also are selected in order to modify the characteristics of the CD4 adheson as indicated above.

While the site for introducing an amino acid sequence variation is predetermined, the mutation per se need not be predetermined. For example, in order to optimize the performance of a mutation at a given site, random mutagenesis may be conducted at the target codon or region and the expressed adheson variants screened for the optimal combination of desired activities. Techniques for making sub-

stitution mutations at predetermined sites in DNA having a known sequence are well known, for example M13 primer mutagenesis.

Adheson variants that are not capable of binding HIV gp120 are useful nonetheless as immunogens for raising antibodies to the adheson or as immunoassay kit components (labelled, as a competitive reagent for gp120 assay, or unlabelled as a standard for an adheson assay) so long as at least one adheson epitope remains active.

The DNA encoding adhesons is obtained by known procedures. See Williams, Immunol. Today 8:298–303 (1987) and citations therein. In general, prokaryotes are used for cloning of CD4 variant DNA sequences. For example, *E. coli* strain SR101 (for propagating m13 phage, a λ-resistant strain of JM 101; Messing et al., Nucl. Acids. Res. 9(2) :309–321 [1981]); and *E. coli* K12 strain 294 (ATCC No. 31446) are particularly useful. Other microbial strains which may be used include *E. coli* B, UM101 and *E. coli* $_{\chi}$1776 (ATCC No. 31537). These examples are illustrative rather than limiting.

DNA encoding the variant adhesons are inserted for expression into vectors containing promoters and control sequences which are derived from species compatible with the intended host cell. The vector ordinarily, but need not, carry a replication site as well as one or more marker sequences which are capable of providing phenotypic selection in transformed cells. For example, *E. coli* is typically transformed using a derivative of pBR322 which is a plasmid derived from an *E. coli* species (Bolivar, et al., Gene 2: 95 [1977]). pBR322 contains genes for ampicillin and tetracycline resistance and thus provides easy means for identifying transformed cells. The pBR322 plasmid, or other microbial plasmid must also contain or be modified to contain promoters and other control elements commonly used in recombinant DNA constructions.

Promoters suitable for use with prokaryotic hosts illustratively include the β-lactamase and lactose promoter systems (Chang et al., Nature, 275: 615 [1978]; and Goeddel et al., Nature 281: 544 [1979]), alkaline phosphatase, the tryptophan (trp) promoter system (Goeddel, Nucleic Acids Res. 8: 4057 [1980] and EPO Appln. Publ. No. 36,776) and hybrid promoters such as the tac promoter (H. de Boer et al., Proc. Natl. Acad. Sci. USA 80: 21–25 [1983]). However, other functional bacterial promoters are suitable. Their nucleotide sequences are generally known, thereby enabling a skilled worker operably to ligate them to DNA encoding the adheson variant using linkers or adaptors to supply any required restriction sites (Siebenlist et al., Cell 20: 269 [1980]). Promoters for use in bacterial systems also will contain a Shine-Dalgarno (S.D.) sequence operably linked to the DNA encoding the antigen.

In addition to prokaryotes, eukaryotic microbes such as yeast cultures also are useful as cloning or expression hosts. *Saccharomyces cerevisiae*, or common baker's yeast is the most commonly used eukaryotic microorganism, although a number of other strains are commonly available. For expression in Saccharomyces, the plasmid YRp7, for example, (Stinchcomb, et al., Nature 282: 39 [1979]; Kingsman et al, Gene 7: 141 [1979]; Tschemper et al., Gene 10: 157 [1980]) is commonly used. This plasmid already contains the trp1 gene which provides a selection marker for a mutant strain of yeast lacking the ability to grow in tryptophan, for example ATCC no. 44076 or PEP4-1 (Jones, Genetics 85: 12 [1977]). The presence of the trp1 lesion as a characteristic of the yeast host cell genome then provides an effective means of selection by growth in the absence of tryptophan.

Suitable promoting sequences for use with yeast hosts include the promoters for 3-phosphoglycerate kinase (Hitzeman et al., J. Biol. Chem. 255: 2073 [1980]) or other glycolytic enzymes (Hess et al., J. Adv. Enzyme Reg. 7: 149 [1968]; and Holland, Biochemistry 17: 4900 [1978]), such as enolase, glyceraldehyde-3-phosphate dehydrogenase, hexokinase, pyruvate decarboxylase, phosphofructokinase, glucose-6-phosphate isomerase, 3-phosphoglycerate mutase, pyruvate kinase, triosephosphate isomerase, phosphoglucose isomerase, and glucokinase.

Other yeast promoters, which are inducible promoters having the additional advantage of transcription controlled by growth conditions, are the promoter regions for alcohol dehydrogenase 2, isocytochrome C, acid phosphatase, degradative enzymes associated with nitrogen metabolism, metallothionein, glyceraldehyde-3-phosphate dehydrogenase, and enzymes responsible for maltose and galactose utilization. Suitable vectors and promoters for use in yeast expression are further described in R. Hitzeman et al., European Patent Publication No. 73,657A. Yeast enhancers also are advantageously used with yeast promoters.

Promoters for controlling transcription from vectors in mammalian host cells may be obtained from various sources, for example, the genomes of viruses such as: polyoma, Simian Virus 40 (SV40), adenovirus, retroviruses, hepatitis-B virus and most preferably cytomegalovirus, or from heterologous mammalian promoters, e.g. the beta actin promoter. The early and late promoters of the SV40 virus are conveniently obtained as an SV40 restriction fragment which also contains the SV40 viral origin of replication. Fiers et al., Nature, 273: 113 (1978). The immediate early promoter of the human cytomegalovirus is conveniently obtained as a HindIII E restriction fragment. Greenaway, P. J. et al., Gene 18: 355–360 (1982). Of course, promoters from the host cell or related species also are useful herein.

DNA transcription in higher eukaryotes is increased by inserting an enhancer sequence into the vector. Enhancers are cis-acting elements of DNA, usually from about 10 to 300 bp, that act to increase the transcription initiation capability of a promoter. Enhancers are relatively orientation and position independent having been found 5' (Laimins, L. et al., Proc.Natl.Acad.Sci. 78: 993 [1981]) and 3' (Lusky, M. L., et al., Mol. Cell Bio. 3: 1108 [1983]) to the transcription unit, within an intron (Banerji, J. L. et al., Cell 33: 729 [1983]) as well as within the coding sequence itself (Osborne, T. F., et al., Mol. Cell Bio. 4: 1293 [1984]). Many enhancer sequences are now known from mammalian genes (globin, elastase, albumin, α-fetoprotein and insulin). Typically, however, one will use an enhancer from a eukaryotic cell virus. Examples include the SV40 enhancer on the late side of the replication origin (bp 100–270), the cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers.

Expression vectors used in eukaryotic host cells (yeast, fungi, insect, plant, animal, human or nucleated cells) may also contain sequences necessary for the termination of transcription which may affect mRNA expression. These regions are transcribed as polyadenylated segments in the untranslated portion of the mRNA encoding the adheson.

Expression vector systems generally will contain a selection gene, also termed a selectable marker. Examples of suitable selectable markers for mammalian cells are dihydrofolate reductase (DHFR), thymidine kinase or neomycin. When such selectable markers are successfully transferred into a mammalian host cell, the transformed mammalian host cell can survive if placed under selective pressure. There are two widely used distinct categories of selective

regimes. The first category is based on a cell's metabolism and the use of a mutant cell line which lacks the ability to grow independent of a supplemented medium. Two examples are: CHO DHFR$^-$ cells and mouse LTK$^-$ cells. These cells lack the ability to grow without the addition of such nutrients as thymidine or hypoxanthine. Because these cells lack certain genes necessary for a complete nucleotide synthesis pathway, they cannot survive unless the missing nucleotides are provided in a supplemented medium. An alternative to supplementing the medium is to introduce an intact DHFR or TK gene into cells lacking the respective genes, thus altering their growth requirements. Individual cells which were not transformed with the DHFR or TK gene will not be capable of survival in non supplemented media.

The second category is dominant selection which refers to a selection scheme used in any cell type and does not require the use of a mutant cell line. These schemes typically use a drug to arrest growth of a host cell. Those cells which have a novel gene would express a protein conveying drug resistance and would survive the selection. Examples of such dominant selection use the drugs neomycin, Southern P. and Berg, P., J. Molec. Appl. Genet. 1: 327 (1982), mycophenolic acid, Mulligan, R. C. and Berg, P. Science 209: 1422 (1980) or hygromycin, Sugden, B. et al., Mol. Cell. Biol. 5: 410–413 (1985). The three examples given above employ bacterial genes under eukaryotic control to convey resistance to the appropriate drug G418 or neomycin (geneticin), xgpt (mycophenolic acid) or hygromycin, respectively.

"Amplification" refers to the increase or replication of an isolated region within a cell's chromosomal DNA. Amplification is achieved using a selection agent e.g. methotrexate (MTX) which inactivates DHFR. Amplification or the making of successive copies of the DHFR gene results in greater amounts of DHFR being produced in the face of greater amounts of MTX. Amplification pressure is applied notwithstanding the presence of endogenous DHFR, by adding ever greater amounts of MTX to the media. Amplification of a desired gene can be achieved by cotransfecting a mammalian host cell with a plasmid having a DNA encoding a desired protein and the DHFR or amplification gene permitting cointegration. One ensures that the cell requires more DHFR, which requirement is met by replication of the selection gene, by selecting only for cells that can grow in the presence of ever-greater MTX concentration. So long as the gene encoding a desired heterologous protein has cointegrated with the selection gene replication of this gene gives rise to replication of the gene encoding the desired protein. The result is that increased copies of the gene, i.e. an amplified gene, encoding the desired heterologous protein express more of the desired heterologous protein.

Preferred host cells for expressing the CD antigen variants of this invention are mammalian cell lines, examples including: monkey kidney CV1 line transformed by SV40 (COS-7, ATCC CRL 1651); human embryonic kidney line (293, Graham, F. L. et al., J. Gen Virol. 36: 59 [1977] and 293S cells [293 subclones selected for better suspension growth]); baby hamster kidney cells (BHK, ATCC CCL 10); chinese hamster ovary-cells-DHFR(CHO, Urlaub and Chasin, Proc. Natl.Acad.Sci. (USA) 77: 4216, [1980]); mouse sertoli cells (TM4, Mather, J. P., Biol. Reprod. 23: 243–251 [1980]); monkey kidney cells (CV1 ATCC CCL 70); african green monkey kidney cells (VERO-76, ATCC CRL-1587); human cervical carcinoma cells (HELA, ATCC CCL 2); canine kidney cells (MDCK, ATCC CCL 34); buffalo rat liver cells (BRL 3A, ATCC CRL 1442); human lung cells (W138, ATCC CCL 75); human liver cells (Hep G2, HB 8065); mouse mammary tumor (MMT 060562, ATCC CCL51 cells); and TRI cells (Mather, J. P. et al., Annals N.Y. Acad. Sci. 383: 44–68 [1982]).

"Transformation" means introducing DNA into an organism so that the DNA is replicable, either as an extrachromosomal element or by chromosomal integration. One suitable for transformation of the host cells is the method of Graham, F. and van der Eb, A., Virology 52: 456–457 (1973). However, other methods for introducing DNA into cells such as by nuclear injection or by protoplast fusion may also be used. If prokaryotic cells or cells which contain substantial cell walls are used as hosts, the preferred method of transfection is calcium treatment using calcium chloride as described by Cohen, F. N. et al., Proc. Natl. Acad. Sci. (USA), 69: 2110 (1972).

Construction of suitable vectors containing the desired coding and control sequences employ standard and manipulative ligation techniques. Isolated plasmids or DNA fragments are cleaved, tailored, and religated in the form desired to form the plasmids required. Suitable procedures are well known for the construction described herein. See, for example, (Maniatis, T. et al., *Molecular Cloning*, 133–134 Cold Spring Harbor, [1982]; "Current Protocols in Molecular Biology", edited by Ausubel et al., [1987], pub. by Greene Publishing Associates & Wiley-Interscience).

Correct plasmid sequences are confirmed by transforming *E. coli* K12 strain 294 (ATCC 31446) with ligation mixtures, successful transformants selected by ampicillin or tetracycline resistance where appropriate, plasmids from the transformants prepared, and then analyzed by restriction enzyme digestion and/or sequenced by the method of Messing et al., Nucleic Acids Res. 9: 309 (1981) or by the method of Maxam et al., Methods in Enzymology 65: 499 (1980).

Host cells are transformed with the expression vectors of this invention. Thereafter they are cultured in appropriate culture media, e.g. containing substances for inducing promoters, selecting transformants or amplifying genes. The culture conditions, such as temperature, pH and the like, are those previously used with the host cell selected for expression, and will be apparent to the ordinarily skilled artisan.

The secreted adheson variants are recovered and purified from the culture supernatants or lysates of recombinant hosts. Typically, the supernatants are concentrated by ultrafiltration, contacted with a ligand affinity or immunoaffinity matrix so as to adsorb the adheson variant, and eluted from the matrix. Optionally, the adheson is purified by ion exchange chromatography.

Surprisingly, purification of soluble CD4 adheson from culture medium was unexpectedly difficult. Notwithstanding that the hydrophobic transmembrane region of the antigen had been deleted, the antigen exhibited a strong tendency to form aggregates that could be readily removed from suspension by centrifugation at 1000× g, and which avidly coat surfaces such as ultrafiltration membranes. This appears to result from the reduction in concentration of albumin or other serum protein (ordinarily present in the crude preparation) to a particular level, below which the truncated antigen no longer remains soluble. This phenomenon appears to be aggravated by exposure of the CD4 adheson to low pH (<about pH 4). As a result, separation procedures (particularly those that employ acid elution, such as immunoaffinity) should be modified so that the eluate is maintained at, or immediately returned to, about neutrality. Further, a surfactant, e.g. a detergent such as Tween 80, should be included with the antigen during the separation

procedure. The final purified product will be stabilized with a predetermined protein such as albumin, and/or a detergent.

The purified adheson is formulated into conventional pharmacologically acceptable excipients.

It is administered to patients having HIV infection at a dosage capable of maintaining a concentration of greater than about 100 ng of soluble CD4 adheson/ml plasma. For CD4 adheson variants having different molecular weights, about 2 picomoles of soluble receptor per ml of plasma will be initially evaluated clinically in order to establish a stoichiometric equivalence with native (membrane bound) and soluble receptor. The ordinary dosage of soluble CD4 is 100 μg/kg of patient weight/day.

The therapeutic CD4 variants are employed with other therapies and agents for the treatment of AIDS, including AZT, neutralizing antibodies and immunocytotoxins, gp120 fragments and vaccines.

In order to facilitate understanding of the following examples certain frequently occurring methods and/or terms will be described.

"Plasmids" are designated by a lower case p preceded and/or followed by capital letters and/or numbers. The starting plasmids herein are either commercially available, publicly available on an unrestricted basis, or can be constructed from available plasmids in accord with published procedures. In addition, equivalent plasmids to those described are known in the art and will be apparent to the ordinarily skilled artisan.

"Digestion" of DNA refers to catalytic cleavage of the DNA with a restriction enzyme that acts only at certain sequences in the DNA. The various restriction enzymes used herein are commercially available and their reaction conditions, cofactors and other requirements were used as would be known to the ordinarily skilled artisan. For analytical purposes, typically 1 μg of plasmid or DNA fragment is used with about 2 units of enzyme in about 20 μl of buffer solution. For the purpose of isolating DNA fragments for plasmid construction, typically 5 to 50 μg of DNA are digested with 20 to 250 units of enzyme in a larger volume. Appropriate buffers and substrate amounts for particular restriction enzymes are specified by the manufacturer. Incubation times of about 1 hour at 37° C. are ordinarily used, but may vary in accordance with the supplier's instructions. After digestion the reaction is electrophoresed directly on a polyacrylamide gel to isolate the desired fragment.

"Recovery" or "isolation" of a given fragment of DNA from a restriction digest means separation of the digest on polyacrylamide or agarose gel by electrophoresis, identification of the fragment of interest by comparison of its mobility versus that of marker DNA fragments of known molecular weight, removal of the gel section containing the desired fragment, and separation of the gel from DNA. This procedure is known generally (Lawn, R. et al., Nucleic Acids Res. 9: 6103–6114 [1981], and Goeddel, D. et al., Nucleic Acids Res. 8: 4057 [1980]).

"Dephosphorylation" refers to the removal of the terminal 5' phosphates by treatment with bacterial alkaline phosphatase (BAP). This procedure prevents the two restriction cleaved ends of a DNA fragment from "circularizing" or forming a closed loop that would impede insertion of another DNA fragment at the restriction site. Procedures and reagents for dephosphorylation and other recombinant manipulations are conventional. Reactions using BAP are carried out in 50 mM Tris at 68° C. to suppress the activity of any exonucleases which may be present in the enzyme preparations. Reactions were run for 1 hour. Following the reaction the DNA fragment is gel purified.

"Ligation" refers to the process of forming phosphodiester bonds between two double stranded nucleic acid fragments (Maniatis, T. et al., *Id.* at 146). Unless otherwise provided, ligation may be accomplished using known buffers and conditions with 10 units of T4 DNA ligase ("ligase") per 0.5 μg of approximately equimolar amounts of the DNA fragments to be ligated.

"Filling" or "blunting" refers to the procedures by which the single stranded end in the cohesive terminus of a restriction enzyme-cleaved nucleic acid is converted to a double strand. This eliminates the cohesive terminus and forms a blunt end. This process is a versatile tool for converting a restriction cut end that may be cohesive with the ends created by only one or a few other restriction enzymes into a terminus compatible with any blunt-cutting restriction endonuclease or other filled cohesive terminus. Typically, blunting is accomplished by incubating 2–15 μg of the target DNA in 10 mM $MgCl_2$, 1 mM dithiothreitol, 50 mM NaCl, 10 mM Tris (pH 7.5) buffer at about 37° C. in the presence of 8 units of the Klenow fragment of DNA polymerase 1 and 250 μM of each of the four deoxynucleoside triphosphates. The incubation generally is terminated after 30 min. phenol and chloroform extraction and ethanol precipitation.

The following examples merely illustrate the best mode now contemplated for practicing the invention, but should not be construed to limit the invention. All literature citations herein are expressly incorporated by reference.

EXAMPLE 1

Construction of Vectors for the Expression of Native CD4 and Secreted Derivatives

Section 1

The plasmid used for recombinant synthesis of human CD4 was pSVeCD4DHFR. The plasmid was constructed as follows:

λCD4P1 containing most of the coding sequence of human CD4 (obtained from a human placental cDNA library using oligonucleotide probes based on the published sequence [Maddon et al. 1985]) was digested with EcoRI to produce the cDNA insert. This fragment was recovered by polyacrylamide gel electrophoresis (fragment 1).

pUC18 was digested with EcoRI and the single fragment recovered by polyacrylamide gel electrophoresis (fragment 2). Fragment 1 was ligated to fragment 2 and the ligation mixture transformed into *E. coli* strain 294. The transformed culture was plated on ampicillin media plates and resistant colonies selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct DNA fragments. This plasmid is referred to as pUCCD4.

pSVeE'DHFR (Muesing et al., Cell 48:691–701 [1987]) was digested with KpnI and BamHI and blunted with *E. coli* DNA polymerase I (Klenow fragment) and the four dNTPs. Fragment 3 containing the pML-Amp^r region, SV40 early promoter, the HIV LTR, and the mouse DHFR gene was recovered by gel electrophoresis, ligated and the ligation mixture transformed into *E. coli* strain 294. The transformed culture was plated on ampicillin media plates and resistant colonies selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the BamHI restriction site and the absence of the KpnI restriction site. This plasmid is referred to as pSVeΔBKDHFR and allows EcoRI-BamHI fragments to be inserted after the SV40 early promoter and transcribed under its control, following transfection into an appropriate cell line.

Synthetic oligonucleotides (adaptors 1–8, below) were made to extend from 76 bp 5' of the initiation codon of CD4 translation to the RsaI restriction site at 121 bp 3' of the initiator, with the sequence AATT at the 5' end of the sense strand to generate an end which could ligate to an EcoRI restriction fragment. These oligonucleotides were ligated and the 204 bp fragment containing the entire sequence recovered by gel electrophoresis (fragment 4).

```
CD4 adaptor 1:
AATTCAAGCCCAGAGCCCTGCCATTTCTGTGGGCTCAGGTCCCT

CD4 adaptor 2:
pACTGCTCAGCCCCTTCCTCCCTCGGCAAGGCCACAATGAACCGGGGAGT
C

CD4 adaptor 3:
pCCTTTTAGGCACTTGCTTCTGGTGCTGCAACTGGCGCTCCTCCCAGC

CD4 adaptor 4:
pAGCCACTCAGGGAAACAAAGTGGTGGTGGGCAAAAAAGGGGATACAGTG
GAACTGACCTCT

CD4 adaptor 5:
pACAGGTCAGTTCCACTGTATCCCCTTTTTTGCCCAGCACCACTTTGTTT
CC

CD4 adaptor 6:
pCTGAGTGGCTGCTGGGAGGAGCGCCAGTTGCAGCACCAGAAGCAAGT

CD4 adaptor 7:
pGCCTAAAAGGGACTCCCCGGTTCATTGTGGCCTTGCCGAGGGAGGAAGG
G

CD4 adaptor 8:
GCTGAGCAGTAGGGACCTGAGCCCACAGAAATGGCAGGGCTCTGGGCTTG
```

pUCCD4 was digested with RsaI and SstI and the 401 bp fragment containing part of the CD4 coding sequence recovered by gel electrophoresis (fragment 5). pUC18 was digested with EcoRI and SstI and the fragment comprising the bulk of the plasmid recovered by gel electrophoresis (fragment 6). Fragments 4 and 5 were ligated to fragment 6 and the ligation mixture transformed into *E. coli* strain 294. The transformed culture was plated on ampicillin media plates and resistant colonies selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct fragment. The sequence of the inserted synthetic DNA was checked by excising the 605 bp EcoRI-SstI fragments from several transformants and ligating them to M13 mp19 which had been digested with the same enzymes. After transformation into *E. coli* strain JM101, single-stranded DNA was prepared and sequenced. One plasmid which contained the correct sequence was selected, and is referred to as pCD4int.

pCD4int was digested with EcoRI and SstI and fragment 7 containing the 5' end of the CD4 coding region was recovered by gel electrophoresis. pUCCD4 was digested with SstI and BamHI and the 1139 bp fragment containing the remainder of the CD4 coding region (fragment 8) recovered by gel electrophoresis.

pSVcΔBKDHFR was digested with EcoRI and BamHI and fragment 9 comprising the bulk of the plasmid was isolated. Fragments 7, 8 and 9 were ligated and the ligation mixture transformed into *E. coli* strain 294. The transformed culture was plated on ampicillin media plates and the resistant colonies selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct fragment. This plasmid is referred to as pSVeCD4DHFR, and was used to direct synthesis of recombinant intact CD4.

Section 2

A plasmid was constructed to direct the synthesis of a CD4 derivative lacking the putative transmembrane domain and most of the putative cytoplasmic domain (Maddon et al.). This was done with the intention of creating a secreted form of CD4, based on the assumption that these domains anchor the CD4 glycoprotein to the cell membrane, and that their deletion would result in the secretion of the product. This plasmid is referred to as pSVeCD4ΔNlaDHFR and was constructed as follows:

pUCCD4 was digested with SstI and TaqI and the 531 bp fragment (fragment 10) recovered. pUCCD4 was digested with NlaIII and TaqI and the 112 bp fragment (fragment 11) recovered. pUCCD4 was digested with BamHI and NlaIII and the 301 bp fragment (fragment 12) recovered. pCD4int was digested with SstI and BamHI and fragment 13 comprising the bulk of the plasmid recovered. Fragments 10, 11, and 12 were ligated together with fragment 13 and the ligation mixture transformed into *E. coli* strain 294. The transformed culture was plated on ampicillin media plates and resistant colonies selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct fragment. Plasmid DNA from several transformants was sequenced to ensure that the 195 bp NlaIII fragment had been deleted and that the proper reading frame was restored. The resulting plasmid is referred to as pCD4ΔNla.

pCD4ΔNla was digested with EcoRI and BamHI and the 1541 bp fragment containing the sequence of a CD4 derivative lacking the transmembrane and cytoplasmic domains recovered (fragment 14) and ligated to fragment 9 and the ligation mixture transformed into *E. coli* strain 294. The transformed culture was plated on ampicillin media plates and resistant colonies selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct fragment. This plasmid is referred to as pSVeCD4ΔNlaDHFR.

Both pSVeCD4DHFR and pSVeCD4ΔNlaDHFR were transfected into CHO cells by the same method used to establish cell lines stably expressing HIV-I polypeptides (Muesing, Smith and Capon, Cell 48:6910701 [1987]). These cells were assayed for production by radioimmunoprecipitation as described below. While no product was detected in initial experiments, subsequent experiments showed that the above described coding segment could indeed direct the synthesis of a soluble CD4 adheson variant both in CHO and 293 cells.

Section 3

A different expression system was initially used for the synthesis and expression of a CD4 variant lacking completely the cytoplasmic and transmembrane domains. This system uses the cytomegalovirus promoter and can be used in cultured cells of human origin. The first plasmid constructed for use in this system contained the entire coding region for CD4 and was intended to function as a control in the following studies. It is referred to as pRKCD4, and was constructed as follows:

pSVeCD4DHFR was digested with EcoRI and BamHI and fragment 15 containing the entire CD4 coding region was isolated. pRK5 (U.S. Ser. No. 97,472, filed Sep. 11, 1987) was digested with EcoRI and BamHI and fragment 16 comprising the bulk of the plasmid recovered by gel electrophoresis, ligated to fragment 15, and the ligation mixture transformed into *E. coli* strain 294. The transformed culture was plated on ampicillin media plates and resistant colonies selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct fragment. This plasmid is referred to as pRKCD4.

Section 4

The next plasmid constructed was designed to direct the expression of the above-mentioned (Section 3) secreted derivative of CD4. The coding region of CD4 was fused after amino acid residue 368 of mature CD4 to a sequence from pBR322 which codes for 9 more residues before a translation termination codon. This removes the putative CD4 transmembrane and cytoplasmic domains, which are presumed to anchor CD4 to the cell surface. The plasmid is referred to as pRKCD4T (and which produces protein called CD4T), and was constructed as follows:

pSVeCD4DHFR was digested with HpaII, blunted with Klenow fragment and the four dNTPs, and digested with BstEII. The 382 bp fragment (fragment 17) containing part of the CD4 coding sequence was recovered by gel electrophoresis. pSVeCD4DHFR was digested with EcoRI and BstEII and the 874 bp fragment (fragment 18) recovered. pBR322 was digested with HindIII, blunted with Klenow fragment and the four dNTPs, and digested with EcoRI. Fragment 19 comprising the bulk of the plasmid was isolated and ligated to fragments 17 and 18 and the ligation mixture transformed into *E. coli* strain 294. The transformed culture was plated on ampicillin media plates and resistant colonies selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct fragment. This plasmid is referred to as pCD4Tint.

pRK5 was digested with EcoRI and SmaI and fragment 20 comprising the bulk of the plasmid isolated. pCD4Tint was digested with EcoRI and EcoRV and the 1410 bp fragment containing the CD4 coding sequence to the HpaII site at 1176 bp 3' of the initiating codon and the 154 bp HindIII-EcoRV fragment of pBR322 was recovered (fragment 21). Fragments 20 and 21 were ligated and the ligation mixture transformed into *E. coli* strain 294. The transformed culture was plated on ampicillin media plates and resistant colonies selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct fragment. This plasmid is referred to as pRKCD4T.

Section 5a

In order to create a secreted form of CD4 which could be purified with an antibody directed to herpes virus type I glycoprotein D, a plasmid was constructed to express a derivative of CD4T in which the region coding for the mature, processed CD4T polypeptide was fused to a sequence coding for the signal peptide and the first 27 residues of the mature type I Herpes Simplex Virus gD glycoprotein. This plasmid is referred to as pRKGDCD4T, and was constructed as follows:

pgDTrunc.DHFR was digested with EcoRI and PvuII and the fragment containing the coding region for the signal peptide and first 27 residues of the mature HSV I gD glycoprotein was isolated (fragment 22). pRKCD4T was digested with EcoRI and BstEII and fragment 23 containing the 3' end of the CD4 coding sequence and the pRK5 region was isolated.

Synthetic oligonucleotides GD (adaptors 1–2, below) containing the coding sequence of CD4 from the codon for the amino terminal residue of mature CD4 to the RsaI site at 121 bp 3' of translation initiation, and containing the sequence CTGCTCGAG at the 5' end of the sense strand were prepared (fragment 24). pRKCD4 was digested with RsaI and BstEII and the 665 bp fragment containing part of the coding region for CD4 was recovered (fragment 25) and ligated to fragment 24. After digestion with BstEII to ensure that only monomeric fragment was present, the 724 bp fragment containing both sequences was recovered by gel electrophoresis (fragment 26).

Fragments 22, 23 and 26 were ligated and the ligation mixture transformed into *E. coli* strain 294. The transformed culture was plated on ampicillin media plates and resistant colonies selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct fragment. The sequence of several transformants was checked to ensure that the synthetic insert was correct and that reading frame was preserved. This plasmid is referred to as pRKGDCD4T.

These pRK5 derived plasmids preferably were transfected into 293S cells for stable expression according to Muesing, et al. Cell 48:691 (1987) with the exception that in addition to the plasmid of interest a plasmid expressing the neomycin resistance gene pRSV neo (Gorman et al. Science 221:553–555 (1985)) was cotransfected. 293 cells also are used satisfactorily as host cells. 2 days after transfection, the cells were passaged into standard medium (1:1 F12/DME supplemented with L-glutamine, penicillin-streptomycin and 10% FBS) with 0.5 mg/ml G418 (Geneticin sulfate; Gibco) for selection of stable cell lines, rather than in media containing methotrexate as shown by Muesing et al. Cells were assayed for production of CD4 or CD4 analogs by radioimmunoprecipitation. Binding studies (section 5c) used conditioned supernatants from these cells in the 1:1 F12/DME medium. Materials used in infectivity assays (section 5b) were obtained as described in section 8 below.

gDCD4 adaptor 1:

CTGCTCGAGCAGGGAAACAAAGTGGTGCTGGGCAAAAAGGGGATACAG7G
GAACTGAC

gDCD4 adaptor 2:

pACAGGTCAGTTCCACTGTATCCCCTTTTTTGCCCAGCACCACTTTGTTT
CCCTGCTCGA

Section 5b

The following constitutes a study of the neutralization of HIV-1 infectivity by soluble CD4 analogs. A modification of the neutralization procedure of Robert-Guroff et al., Nature 316:72 (1985) was followed. Equal volumes of inhibitor supernatant and virus (60 microliters) were incubated at 4 degrees C. for 1 hour, then the same volume of H9 (Gallo et al., Science 224:500, 1984) at 5×10^6/ml was added and incubation continued for 1 hour at 37 degrees C. Following absorption, 2.5×10^5 cells in 150 microliters were transferred to 2 ml of incubation media. After 4 days at 37 degrees C., the cultures were split 1:2 with fresh media and incubated for an additional 3 days. Cultures were harvested, reverse transcriptase activity was measured (Groopman et al., AIDS Research and Human Retroviruses 3:71, 1987), and immunofluorescence reactivity with HIV-1 positive serum was determined as described (Poiesz et al., Proc. Acad. Nat. Sci. USA 77:7415, 1980). Inhibitor supernatants were obtained from confluent plate cultures of 293S/CDT4, 293S/gDCD4T cells or untransfected 293S cells by replacing the growth medium incubation media and harvesting the supernatants 24 hours later. Inhibitor supernatant replaced part or all of the incubation media during the first three days of culture as indicated in the second column of Table 3. Challenge dose of virus was 100 $TCID_{50}$ (Groopman et al., supra) of HIV-1 strain HTLV-IIIB grown in H9 cells assayed in the same system. Incubation media consisted of RPMI 1640 media containing 2 mM L-glutamine, 100 units/ml penicillin, 100 micrograms/ml streptomycin, 2 micrograms/ml polybrene and 20% fetal calf serum (M.A. Bioproducts).

TABLE 3

Inhibitor supernatant	Dilution of Inhibitor supernatant	Indirect immunofluorescence (% positive cells		Reverse transcriptase (cpm/ml × 10^5)	
mock-transfected	undil.; 1:4	65.3	65.5	21.8	23.9
mock-transfected	undil.; 1:4	61.2	61.1	18.5	28.1
CD4T	undil.; 1:4	0.4	18.0	0.11	5.94
CD4T	undil.; 1:4	0.8	16.1	0.15	3.72
gDCD4T	undil.; 1:4	0.4	26.8	0.14	9.92
gDCD4T	undil.; 1:4	1.4	36.1	0.23	11.3

Both forms of soluble CD4 virtually abolished the growth of HIV-1, when incubated with virus-infected cells without prior dilution (Table 2). At a dilution of 1:4 the soluble CD4 preparations were only partially effective in inhibiting virus growth, however the level of fluorescent-positive cells and reverse transcriptase was still significantly lower than cultures receiving mock-transfected cell supernatants (Table 2). Since there was no significant difference in virus growth between diluted and undiluted control supernatants, nor did any of the supernatants affect the growth of uninfected H9 cells (data not shown), soluble CD4 proteins present in these supernatants were concluded to be responsible for the neutralization of HIV-1 infection of H9 cells.

Section 5c

To determine the affinity constant for interactions between gp120 and CD4 or CD4 variants, saturation binding analysis was carried out with soluble CD4 (supra) and detergent solubilized intact CD4 (Lasky et al. Cell 50:975 [1987]) employing radioiodinated gp120 labeled with lactoperoxidase. Binding reactions consisted of ^{125}I-gp120 (3 ng to 670 ng, 2.9 nCi/ng) incubated for 1 hour at 0 degrees C. with cell lysates containing intact CD4 (Laskey et al., op cit.) or cell supernatants containing unlabeled CD4T or gDCD4T prepared as described in section 5a. Reactions (0.2 ml) had a final composition of 0.5× McDougal Lysis Buffer (McDLB) (1× McDLB contains 0.5% Nonidet NP-40, 0.2% Na deoxycholate, 0.12 M NaCl, 0.02 M Tris-HCl, pH 8.0) and were performed in duplicate, both in the presence or absence of 50 micrograms of unlabeled purified gp120 (74 fold or greater excess). Following incubation, bound gp120 was quantitated by immunoprecipitation and counted in a gamma counter. For immunoprecipitation, binding reaction solutions were preabsorbed with 5 microliters of normal rabbit serum for one hour at 0° C., and cleared with 40 microliters of Pansorbin (10% w/v, Calbiochem) for 30 minutes at 0 degrees C. Samples were then incubated overnight at 0 degrees C. with 2 microliters of normal serum or 5 microliters (0.25 microgram) of OKT4 monoclonal antibody (Ortho) followed by collection of immune complexes with 10 microliters of Pansorbin. Precipitates were washed twice in 1× McDLB and once in water, then eluted by eluting at 100 degrees C. for 2 minutes in sample buffer (0.12 M Tris-HCl pH 6.8, 4% SDS, 0.7 M mercaptoethanol, 20% glycerol, and 0.1% bromophenol blue). CD4 molecules were bound saturably by gp120, and yielded a simple mass action binding curve. Supernatants from mock-transfected cells gave a level of specifically bound gp120 less than 1% that found for supernatants containing soluble CD4. Scatchard analysis revealed a single class of binding sites on each molecule, with apparent dissociation constants (Kd) of 1.3×10^{-9} M, 0.83×10^{-9} M and 0.72×10^{-9} M for intact CD4, CD4T and gDCD4T, respectively. The values obtained for CD4-gp120 binding in solution are comparable to the affinity previously measured for gp120 binding to CD4 on whole cells (Kd-4.0×10^{-9} M. Lasky, Cell, supra).

Section 6

In order to produce secreted derivatives of CD4 which are free of extraneous amino acid residues, two plasmids were constructed for expression in 293 cells. The plasmids contain CD4 genes which have been truncated without the addition of extra residues, and are referred to as pRKCD4ΔNla and pRKCD4TP (and which produce proteins called CD4TP and CD4ΔNla), and were constructed as follows:

Fragment 14 containing the CD4 gene with the 195 bp NlaIII restriction fragment deleted was ligated to fragment 16, which is pRK5 digested with EcoRI and BamHI. The ligation mixture was transformed into *E. coli* strain 294, the transformed culture plated on ampicillin media plates and resistant colonies selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct fragment. The resulting plasmid is referred to as pRKCD4ΔNla.

Synthetic DNA (5' CGT GAT AGA AGC TTT CTA GAG 3') was made to attach to the HpaII site at 1176 bp and which when so attached would terminate translation after amino acid residue 368 of mature CD4 (fragment 27). The other end of this fragment was designed to ligate to BamHI restriction fragments. pUCCD4 was digested with BstEII and HpaII and the 382 bp fragment containing part of the CD4 gene was recovered (fragment 28). Fragments 27 and 28 were ligated and then digested with BstEII to reduce dimerized fragments to monomers, and the resulting 401 bp fragment was recovered (fragment 29).

pRKCD4 was digested with BstII and BamHI and the fragment comprising the bulk of the plasmid (fragment 30) was isolated and ligated to fragment 29. The ligation mixture was transformed into *E. coli* strain 294, the transformed culture plated on ampicillin media plates and resistant colonies selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct fragment. The resulting plasmid is referred to as pRKCD4TP. Both plasmids are transfected into 293 cells to generate stable variant CD4-expressing cell lines as described above.

Section 7

Two plasmids were constructed to direct the expression of secreted CD4 lacking extraneous amino acid residues in CHO cells. These are referred to as pSVeCD4ΔNlaSVDHFR and pSVeCD4TPSVDHFR (and which encode proteins having the primary sequence of CD4ΔNla and CD4TP), and were constructed as follows:

pE348HBV.E400D22 was digested with PvuI and EcoRI and the fragment containing the SV40 early promoter and part of the β-lactamase gene was recovered (fragment 31). pE348HBV.E400D22 was digested with PvuI and BamHI and the large fragment containing the balance of the β-lactamase gene as well as the SV40 early promoter and the DHFR gene was isolated (fragment 32).

Fragments 31 and 32 were ligated together with fragment 14 and transformed into *E. coli* strain 294. The transformed culture was plated on ampicillin media plates and resistant colonies selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct fragment. The resulting plasmid is referred to as pSVECD4ΔNlaSVDHFR. This plasmid contains the same DNA fragment encoding the soluble CD4 molecule found in the above-mentioned plasmid pSVeCD4ΔNlaDHFR (Section 2).

pRKCD4TP was digested with EcoRI and BamHI and the fragment containing the truncated CD4 coding region was isolated and ligated to fragments 31 and 32. The ligation

mixture was transformed into *E. coli* strain 294, the transformed culture plated on ampicillin media plates and resistant colonies selected. Plasmid DNA was prepared from transformants and checked by restriction analysis for the presence of the correct fragment. The resulting plasmid is referred to as pSVeCD4TPSVDHFR. Both of these plasmids are transfected into CHO cells and amplified transfectants selected by methotrexate using conventional procedures.

EXAMPLE 2

Fusions of the V region of the CD4 gene, which is homologous to the variable region of immunoglobulin genes (ref Maddon et al. 1985), to the constant (C) region of human immunoglobulin κ and γ2 chains are constructed as follows:

Synthetic DNA is made to code for the C region of human κ chain (residues 109–214) based on the sequence published by Morin et al., Proc. Natl. Acad. Sci. 82:7025–7029, with the addition at the 5' end of the coding strand of the sequence GGGG, which allows this fragment to be ligated to the BspMI site at the end of the putative V-like region of CD4. At the 3' end of the coding region, a translational stop codon is added as well as a sequence which allows this end to be ligated to BamHI restriction fragments. The synthetic DNA is made in 8 fragments, 4 for each strand, 70–90 bases long. These are then allowed to anneal and ligated prior to isolation on a polyacrylamide gel (fragment 33).

pRKCD4 is digested with EcoRI and BspMI and the 478 bp fragment containing the region coding for the putative V-like domain of CD4 is recovered (fragment 34). Fragments 33 and 34 are ligated together with fragment 16 (from the expression vector pRK5). The ligation mixture is transformed into *E. coli* strain 294, the transformed culture plated on ampicillin media plates and resistant colonies selected. Plasmid DNA is prepared from transformants and checked by restriction analysis for the presence of the correct fragment. The resulting plasmid is referred to as pRKCD4Ck.

A plasmid encoding a fusion of the CD4 V-like domain to the human immunoglobulin Cγ2 region is constructed in a similar fashion, and is referred to as pRKCD4Cγ2. Both of these plasmids are transfected into 293 cells, myeloma cells or other competent cells in order to obtain cell lines expressing variant CD4 molecules as described above.

EXAMPLE 3

The gDCD4T secreted by the method of Example 1 was purified from cell culture fluid containing either 10% FBS (fetal bovine serum) or no added FBS. The conditioned cell culture fluid was first concentrated by ultrafiltration then purified by immunoaffinity chromatography. The immunoaffinity column was produced by coupling murine monoclonal antibody 5B6 (whose epitope is on the HSV-1 gD portion of the gDCD4T molecule) to glyceryl coated controlled pore glass by the method of Roy et al., 1984. The concentrated cell culture fluid is applied directly to the column and the contaminating proteins are washed away with neutral pH buffer. The column is then washed with neutral buffer containing tetramethylammonium chloride followed by neutral buffer containing Tween 80. The bound gDCD4T is eluted from the column with buffer at pH 3 containing Tween 80 (0.1% w/v) and is neutralized immediately as it is eluted. The eluted neutralized gDCD4T is then concentrated by ultrafiltration and dialyzed/diafiltered to exchange the buffer for a physiological salt solution containing Tween 80 at approximately 0.1% w/v.

If the detergent is not present the gDCD4T forms aggregates as evidenced by the ability of centrifugation at approximately 10,000× g for 2 minutes to remove the gDCD4T from the solution. Incubation of gDCD4T at 4° C. in 0.1M sodium acetate, 0.5M NaCl and 0.25M tris at pH 7 together with BSA, Tween 80 or glycerol as candidate stabilizers showed that, in the absence of a stabilizer the gDCD4T gradually aggregated over the space of 12 days to the point where only about 60–70% of the protein was soluble. However, use of 0.1% w/v Tween 80 or (0.5 mg/ml BSA ensured that about 100% or 80%, respectively, of the gDCD4T remained soluble over this period. Surprisingly glycerol was ineffective as a stabilizer and produced results inferior even to the control—at 8 days about 80% of the gDCD4T was aggregated when stored in the presence of glycerol.

EXAMPLE 4

Plasmids were constructed to direct the expression of proteins containing differing lengths of the amino-terminal, extracellular domain of CD4 fused to the constant region of human immunoglobulin γ1. These plasmids are referred to as pRKCD4$_{2\gamma1}$, pRKCD4$_{e4\gamma1}$, pRKCD4$_{2\gamma1}$, pRKCD4$_{e2\gamma1}$, pRKCD4$_{1\gamma1}$, and pRKCD4$_{e1\gamma1}$.

Plasmid pRKCD4$_{4\gamma1}$ contains the portion of the CD4 gene from the initiation codon to the fusion site after the codon for serine reside 366 of the mature CD4 polypeptide, immediately followed by the sequence coding for the constant region of human immunoglobulin γ1, starting at the codon for serine residue 114 of mature human immunoglobulin γ1 (Kabat et al.).

Plasmid pRKCD4$_{e4\gamma1}$ contains the portion of the CD4 gene from the initiation codon to the fusion site after the codon for lysine residue 360 of the mature CD4 polypeptide, immediately followed by the sequence coding for the constant region of human immunoglobulin γ1, starting at the codon for serine residue 114 of mature human immunoglobulin γ1 (Kabat et al.).

Plasmid pRKCD4$_{2\gamma1}$ contains the portion of the CD4 gene from the initiation codon to the fusion site after the codon for glutamine residue 180 of the mature CD4 polypeptide, immediately followed by the sequence coding for the constant region of human immunoglobulin γ1, starting at the codon for serine residue 114 of mature human immunoglobulin γ1 (Kabat et al.).

Plasmid pRKCD4$_{e2\gamma1}$ contains the portion of the CD4 gene from the initiation codon to the fusion site after the codon for leucine residue 177 of the mature CD4 polypeptide, immediately followed by the sequence coding for the constant region of human immunoglobulin γ1, starting at the codon for serine residue 114 of mature human immunoglobulin γ1 (Kabat et al.).

Plasmid pRKCD4$_{1\gamma1}$ contains the portion of the CD4 gene from the initiation codon to the fusion site after the codon for aspartic acid residue 105 of the mature CD4 polypeptide, immediately followed by the sequence coding for the constant region of human immunoglobulin γ1, starting at the codon for serine residue 114 of mature human immunoglobulin γ1 (Kabat et al.).

Plasmid pRKCD4$_{e1\gamma1}$ contains the portion of the CD4 gene from the initiation codon to the fusion site after the codon for leucine residue 100 of the mature CD4 polypeptide, immediately followed by the sequence coding for the constant region of human immunoglobulin γ1, starting at the codon for serine residue 114 of mature human immunoglobulin γ1 (Kabat et al.).

Construction of these plasmids required the prior construction of plasmid pRKCD4TP/γ1. It was constructed as follows:

A cDNA clone coding for human immunoglobulin γ1 was obtained from a human spleen cDNA library (Clontech Laboratories, Inc.) using oligonucleotides based on the published sequence (Ellison et al., "Nucl. Acids Res." 10:4071–4079 [1982]), and an EcoRI-EagI fragment (the EcoRI site was contributed by a linker; see FIGS. 4*a*, *b*) containing part of the variable and all of the constant region was obtained. This fragment was blunted with Klenow fragment, and recovered by gel electrophoresis (Fragment a1).

Plasmid pRKCD4TP-kk, encoding a substitutional variant of soluble CD4 (residues 1–368) containing a lysine residue instead of asparagine at position 1 of the mature polypeptide, was constructed from plasmid pRKCD4TP by site-directed mutagenesis. A synthetic oligonucleotide was made as a primer for a mutagenesis reaction to obtain the desired coding sequence. This was synthesized as a 51-mer which contained two silent mutations from the natural sequence in addition to the substitution mutation, and 21 bases on each side of the mutated codons:

```
5'-CCC TTT TTT GCC CAG CAC CAC CTT CTT GCC CTG-
AGT GGC TGC TGG GAG GAG-3'
```

Plasmid pRKCD4TP was transformed into *E. coli* strain SR101 and the transformed colonies plated on ampicillin media plates. Resistant colonies were selected and grown in the presence of m13K07 helper bacteriophage to yield secreted, encapsidated single-stranded templates of pRKCD4TP. The single-stranded plasmid DNA was isolated and used as the template for mutagenesis reactions with the synthetic oligonucleotides described above as primers. The mutagenesis reaction was transformed *E. coli* SR101 and the transformed culture plated on ampicillin media plates. Transformants were screened by colony hybridization (ref. Grunstein-Hogness) for the presence of the appropriate sequence, using the following 16 mer as the probe.

5'-C CAC CTT CTT GCC CTG-3'

The hybridization conditions chosen were sufficiently stringent that the probe only detects the correctly fused product. Colonies identified as positive were selected and plasmid DNA was isolated and transformed into *E. coli* strain SR101. The transformed cultures were plated on ampicillin media plates, and resistant colonies were selected and grown in the presence of m13K07 bacteriophage. Templates were prepared as above and screened by sequencing.

Plasmid pRKCD4TP-kk was digested with XbaI and treated with Klenow Enzyme, and Fragment a2, containing the linearized plasmid was recovered by gel electrophoresis, and ligated with fragment a1. The ligation mixture was transformed into *E. coli* strain 294, the transformed culture plated on ampicillin media plates and resistant colonies selected. Plasmid DNA was prepared from the transformants and checked by restriction analysis for the presence of the correct fragment in the correct orientation (i.e., the immunoglobulin coding region in the same orientation as the CD4 coding region, and at the 3' end of the CD4 coding region). This plasmid is referred to as pRKCD4TP/γ1.

Synthetic oligonucleotides were made as primers for deletional mutagenesis reactions to fuse the appropriate coding sequences of IgG1 and CD4 as described above. These were synthesized as 48-mers comprising 24 nucleotides on each side of the desired fusion site (i.e., corresponding to the COOH-terminal 8 residues of the desired CD4 moiety, and the NH_2-terminal 8 residues of the desired immunoglobulin moiety). Plasmid pRKCD4TP/γ1 was transformed into *E. coli* strain SR101 and the transformed

cultures plated on ampicillin media plates. Resistant colonies were selected and grown in the presence of m13K07 helper bacteriophage to yield secreted, encapsidated single-stranded templates of pRKCD4TP/γ1. The single-stranded plasmid DNA was isolated and used as the template for mutagenesis reactions with the synthetic oligonucleotides described above as primers. The mutagenesis reactions were transformed *E. coli* SR101 and the transformed culture plated on ampicillin media plates. Transformants were screened by colony hybridization (ref. Grunstein-Hogness) for the presence of the appropriate fusion site, using 16mers as probes. These 16mers comprise 8 bases on either side of the fusion site, and the hybridization conditions chosen were sufficiently stringent that the probes only detect the correctly fused product. Colonies identified as positive were selected and plasmid DNA was isolated and transformed into *E. coli* strain SR101. The transformed cultures were plated on ampicillin media plates, and resistant colonies were selected and grown in the presence of m13K07 bacteriophage. Templates were prepared as above and screened by sequencing.

The plasmids were transfected into 293 cells using standard procedures and assayed for expression and production as described above.

	Expressed	Secreted
$pRKCD4_{1\gamma1}$	+	–
$pRKCD4_{e2\gamma1}$	+	+
$pRKCD4_{2\gamma1}$	+	+
$pRKCD4_{e4\gamma1}$	+	+
$pRKCD4_{4\gamma1}$	+	+

Plasmids also were constructed to direct the expression of fusion proteins containing differing lengths of the amino-terminal, extracellular domain of CD4 fused to the truncated portion of the constant region of human immunoglobulin γ1, comprising only the hinge region and constant domains CH_2 and CH_3.

Synthetic oligonucleotides were made as primers for mutagenesis reactions to delete the immunoglobulin sequence from Ser114 to Cys215 inclusive (Kabat et al.). These were synthesized as 48-mers comprising 24 nucleotides on each side of the desired fusion site (i.e., corresponding to the COOH-terminal 8 residues of the desired CD4 moiety, and the NH_2-terminal 8 residues of the desired immunoglobulin moiety). Plasmids $pRKCD4_{4\gamma1}$, $pRKCD4_{2\gamma1}$ and $pRKCD4_{1\gamma1}$ were separately transformed into *E. coli* strain SR101 and the transformed culture plated on ampicillin media plates. Resistant colonies were selected and grown in the presence of m13K07 helper bacteriophage to yield secreted, encapsidated single-stranded templates of these plasmids. The single-stranded plasmid DNA was isolated and used as the template for mutagenesis reactions with the synthetic oligonucleotides described above as primers. The mutagenesis reactions were transformed *E. coli* SR101 and the transformed culture plated on ampicillin media plates. Transformants were screened by colony hybridization (Grunstein-Hogness) for the presence of the appropriate fusion site, using 16mers as probes. These 16mers comprise 8 bases on either side of the fusion site, and the hybridization conditions chosen were sufficiently stringent that the probes only detect the correctly fused product. Colonies identified as positive were selected and plasmid DNA was isolated and transformed into *E. coli* strain SR101. The transformed cultures were plated on ampicillin media plates, and resistant colonies were selected and grown in the presence of m13K07 bacteriophage. Templates were prepared as above and screened by sequencing.

The plasmid derived from plasmid pRKCD4$_{4\gamma1}$ is referred to as pRKCD4$_{4Fc1}$, that derived from plasmid pRKCD4$_{2\gamma1}$ is referred to as pRKCD4$_{2Fc1}$ and that derived from plasmid pRKCD4$_{1\gamma1}$ is referred to as pRKCD4$_{1Fc1}$.

pRKCD4$_{2Fc1}$, pRKCD4$_{1Fc1}$ and pRKCD4$_{4Fc1}$ are cultured in the same fashion as described above and CH1-deleted CD4 immunoadhesons recovered as described elsewhere herein.

Light Chain Fusions

Plasmids were constructed to direct the expression of proteins containing differing lengths of the amino terminal, extracellular domain of CD4 fused to the constant region of human immunoglobulin κ. These plasmids are referred to as pRKCD4$_{4\kappa}$, and pRKCD4$_{e4\kappa}$.

Plasmid pRKCD4$_{4\kappa}$ contains the portion of the CD4 gene from the initiation codon to the fusion site after the codon for serine residue 366 of the mature CD4 polypeptide, immediately followed by the sequence for the constant region of human immunoglobulin κ, starting at the codon for threonine residue 109 of the mature human immunoglobulin κ. (Kabat et al.)

Plasmid pRKCD4$_{e4\kappa}$ contains the portion of the CD4 gene from the initiation codon to the fusion site after the codon for lysine residue 360 of the mature CD4 polypeptide, immediately followed by the sequence for the constant region of human immunoglobulin κ, starting at the codon for threonine residue 109 of the mature human immunoglobulin κ. (Kabat et al.)

These plasmids were constructed in a manner analogous to plasmids pRKCD4$_{4\gamma1}$ and pRKCD4$_{e4\gamma1}$ described above, with the following exception:

The human immunoglobulin κ coding sequence (FIG. 5) was obtained from a human spleen cDNA library (Clontech Laboratories, Inc.) using oligonucleotides based on the published sequence (Hieter, P. A. et al., Cell 22:197–207 [1980]) and an EcoRI-BspMI fragment containing part of the variable region and the entire constant region was obtained (see FIG. 5). This fragment was blunted with Klenow fragment and the four dNTPs. This fragment was used instead of fragment a1, and was used to construct plasmid pRKCD4TP/hκ.

Expression in CHO Cells

Plasmids were or are constructed to direct the expression of the immunoadhesons described above in CHO cells. These are referred to as pSVeCD4$_{4\gamma1}$SVDHFR, pSVeCD4$_{2\gamma1}$SVDHFR, pSVeCD4$_{1\gamma1}$SVDHFR, pSVeCD4$_{e4\gamma1}$SVDHFR, pSVeCD4$_{e2\gamma1}$SVDHFR, pSVeCD4$_{e1\gamma1}$SVDHFR, pSVeCD4$_{4Fc1}$SVDHFR, pSVeCD4$_{2Fc1}$SVDHFR, pSVeCD4$_{1Fc1}$SVDHFR, pSVeCD4$_{4\kappa}$SVDHFR and pSVeCD4$_{2\kappa}$SVDHFR.

Fragment 31 was prepared as described above. Fragment 32a was prepared by digesting plasmid pE348HBV.E400 D22 with BamHI, blunting with Klenow fragment and the four dNTPs, then digesting with PvuI and isolating the large fragment containing the balance of the β-lactamase gene and the SV40 early promoter and the DHFR gene. Plasmids pRKCD4$_{4\gamma1}$, pRKCD4$_{2\gamma1}$, pRKCD4$_{1\gamma1}$, pRKCD4$_{e4\gamma1}$, pRKCD4$_{e2\gamma1}$, pRKCD4$_{e1\gamma1}$, pRKCD4$_{4Fc1}$, pRKCD4$_{2Fc1}$, pRKCD4$_{1Fc1}$, pRKCD4$_{4\kappa}$ and pRKCD4$_{2\kappa}$ were separately digested with HindIII, blunted with Klenow fragment and the four dNTPs, then digested with EcoRI and the fragments encoding the CD4-Ig fusion protein were isolated. The resulting DNA fragments were ligated together with fragments 31 and 32a and transformed into *E. coli* strain 294. Colonies were selected and checked for the presence of the correct plasmid as above, then transfected into CHO cells and amplified by methotrexate selection using conventional procedures.

EXAMPLE 5

Culture, Purification and Formulation of CD4 Variants

Plasmids encoding soluble CD4 adhesons such as CD4T, CD4TP, or soluble CD4 immunoadhesons were calcium phosphate transfected into CHO-DP7 (a proinsulin-transformed autocrine host cell derived from CHO; U.S. Ser. No. 97,472) and the transformants grown in selective medium (1:1 HAM F12/DMEM GHT™ containing 1–10% diafiltered or dialyzed bovine serum). Other suitable host cells are CHO cells or 293S human embryonic kidney cells. The transformants were amplified by methotrexate selection in the same medium but containing 500 nm methotrexate. A subclone capable of secreting CD4TP, CD4tp 500 b, was selected. CD4tp 500 b is cultured in a DMEM/HAM F12 medium at about 37° C. until CD4TP accumulates in the culture, after which the medium is separated from the cells and insoluble matter by centrifuging.

Culture fluid from CD4TP transformants was concentrated and diafiltered to lower the ionic strength. The concentrate was passed through a large volume of Q-Sepharose anion exchange resin (previously equilibrated with 25 mM NaCl, pH 8.5) in order to adsorb contaminants from the culture fluid. The isoelectric point of CD4TP is about 9.5, thus making it possible to discriminate between truncated forms of CD4 and most contaminants by alternate adsorption, respectively, on a cation exchange resin such as carboxymethyl or sulfonyl Sepharose, and an anion exchange resin such as quaternary ammonium Sepharose. In addition, since highly electropositive domains are present in the extracellular segment of CD4 any CD4-containing variant is purified in the same fashion as CD4TP. The unadsorbed culture fluid from the anion exchange resin step was then passed through a cation exchange resin (previously equilibrated with 25 mM NaCl at pH 8.5) whereby CD4TP was adsorbed to the resin. The CD4TP was eluted with a NaCl gradient at pH 8.5, this CD4 variant eluting at about 0.2 M NaCl. Ammonium sulfate was added to the eluate to a concentration of 1.7M and the solution passed through a column of hydrophobic interaction chromatography resin (phenyl or butyl Sepharose). The CD4TP was eluted from the hydrophobic interaction column with a gradient of ammonium sulfate, the CD4TP emerging at about 0.7M ammonium sulfate. The eluate was concentrated and buffer exchanged on a G-25 column using phosphate buffered saline containing 0.02% (w/v) Tween 20 or Tween 80. The CD4TP was soluble and stable in this solution, which was sterile filtered and filled into vials as an aqueous formulation. Other polymeric nonionic surfactants are suitably used with the CD4 formulations, including Pluronic block copolymers or polyethylene glycol.

It is also possible to employ immunoaffinity purification of soluble CD4 wherein the CD4 is adsorbed onto an immobilized antibody against CD4. This method suffers from the disadvantage that elution of the soluble CD4 under acidic conditions leads to protein aggregation that is only thoroughly ameliorated at relatively higher levels of surfactant. The foregoing procedure permits the use of much lower quantities of surfactant, about from 0.01 to 0.10% (w/v) surfactant.

The procedure followed for the purification of CD4 fusions with immunoglobulin heavy chain was to concentrate recombinant supernatants by ultrafiltration and thereafter adsorb the fusion onto resin-immobilized Staphylococcal protein A. The fusion was eluted with 0.1M citrate buffer pH 3 with no salt or detergent. This preparation is buffered into Tris buffer at pH 7.5. The immunoglobulin fusions with CD4 V1–V4 optionally are further purified by the procedure described above for unfused CD4 variants. CD4 immunoglobulin fusions with CD4 V1–V2 also may be purified by the procedure above, except that it is not expected that the isoelectric point of this class of molecules will be as alkaline as that of species containing all four V regions of CD4.

EXAMPLE 6

The characteristics of several adheson variants were determined. As shown in table 4 the immunoadhesons $CD4_{4\gamma1}$ and $CD4_{2\gamma1}$ show improved plasma half-life in rabbits, coupled with high-affinity gp120 binding and an affinity for Fcγ receptor (determined with U937 cells) that is comparable to that of bulk human IgG1.

TABLE 4

	gp120 KD (nM)*	FcγR KD (nM)*	Plasma Half-Life** In Rabbits (Hrs.)
CD4T§	2.3 ± 0.4	Not detected	0.25
$CD4_{4\gamma1}$	1.2 ± 0.1	2.83 ± 0.25	6.4
$CD4_{2\gamma1}$	1.4 ± 0.1	3.01 ± 0.68	40.6
human IgG1	ND**	3.52 ± 0.5	21 days*

*determined in humans

†KD was determined by the method of Anderson et al., "J. Immunol." 125: 2735–2741 (1980).

‡determined by the method of Smith et al., "Science" 238: 1704–07 (1987).

§residues 1–368 only

**The adheson variant was injected intravenously into rabbits and samples of blood were collected periodically and assayed for the presence of the adheson variant.

**Not done.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(iii) NUMBER OF SEQUENCES: 25

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 402 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

```
Met Asn Arg Gly Val Pro Phe Arg His Leu Leu Leu Val Leu Gln
 1               5                  10                  15

Leu Ala Leu Leu Pro Ala Ala Thr Gln Gly Asn Lys Val Val Leu
            20                  25                  30

Gly Lys Lys Gly Asp Thr Val Glu Leu Thr Cys Thr Ala Ser Gln
        35                  40                  45

Lys Lys Ser Ile Gln Phe His Trp Lys Asn Ser Asn Gln Ile Lys
    50                  55                  60

Ile Leu Gly Asn Gln Gly Ser Phe Leu Thr Lys Gly Pro Ser Lys
65                  70                  75

Leu Asn Asp Arg Ala Asp Ser Arg Arg Ser Leu Trp Asp Gln Gly
                80                  85                  90

Asn Phe Pro Leu Ile Ile Lys Asn Leu Lys Ile Glu Asp Ser Asp
            95                  100                 105

Thr Tyr Ile Cys Glu Val Glu Asp Gln Lys Glu Glu Val Gln Leu
        110                 115                 120

Leu Val Phe Gly Leu Thr Ala Asn Ser Asp Thr His Leu Leu Gln
    125                 130                 135

Gly Gln Ser Leu Thr Leu Thr Leu Glu Ser Pro Pro Gly Ser Ser
140                 145                 150
```

-continued

```
Pro Ser Val Gln Cys Arg Ser Pro Arg Gly Lys Asn Ile Gln Gly
                155                 160                 165

Gly Lys Thr Leu Ser Val Ser Gln Leu Glu Leu Gln Asp Ser Gly
                170                 175                 180

Thr Trp Thr Cys Thr Val Leu Gln Asn Gln Lys Lys Val Glu Phe
                185                 190                 195

Lys Ile Asp Ile Val Val Leu Ala Phe Gln Lys Ala Ser Ser Ile
                200                 205                 210

Val Tyr Lys Lys Glu Gly Glu Gln Val Glu Phe Ser Phe Pro Leu
                215                 220                 225

Ala Phe Thr Val Glu Lys Leu Thr Gly Ser Gly Glu Leu Trp Trp
                230                 235                 240

Gln Ala Glu Arg Ala Ser Ser Ser Lys Ser Trp Ile Thr Phe Asp
                245                 250                 255

Leu Lys Asn Lys Glu Val Ser Val Lys Arg Val Thr Gln Asp Pro
                260                 265                 270

Lys Leu Gln Met Gly Lys Lys Leu Pro Leu His Leu Thr Leu Pro
                275                 280                 285

Gln Ala Leu Pro Gln Tyr Ala Gly Ser Gly Asn Leu Thr Leu Ala
                290                 295                 300

Leu Glu Ala Lys Thr Gly Lys Leu His Gln Glu Val Asn Leu Val
                305                 310                 315

Val Met Arg Ala Thr Gln Leu Gln Lys Asn Leu Thr Cys Glu Val
                320                 325                 330

Trp Gly Pro Thr Ser Pro Lys Leu Met Leu Ser Leu Lys Leu Glu
                335                 340                 345

Asn Lys Glu Ala Lys Val Ser Lys Arg Glu Lys Ala Val Trp Val
                350                 355                 360

Leu Asn Pro Glu Ala Gly Met Trp Gln Cys Leu Leu Ser Asp Ser
                365                 370                 375

Gly Gln Val Leu Leu Glu Ser Asn Ile Lys Val Leu Pro Thr Trp
                380                 385                 390

Ser Thr Pro Ser Phe Asn Ala Val Val Tyr His Ser
                395                 400     402
```

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1416 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

```
AATTCAAGCC CAGAGCCCTG CCATTTCTGT GGGCTCAGGT CCCTACTGCT          50

CAGCCCCTTC CTCCCTCGGC AAGGCCACAA TGAACCGGGG AGTCCCTTTT         100

AGGCACTTGC TTCTGGTGCT GCAACTGGCG CTCCTCCCAG CAGCCACTCA         150

GGGAAACAAA GTGGTGCTGG GCAAAAAAGG GGATACAGTG GAACTGACCT         200

GTACAGCTTC CCAGAAGAAG AGCATACAAT TCCACTGGAA AAACTCCAAC         250

CAGATAAAGA TTCTGGGAAA TCAGGGCTCC TTCTTAACTA AAGGTCCATC         300

CAAGCTGAAT GATCGCGCTG ACTCAAGAAG AAGCCTTTGG GACCAAGGAA         350

ACTTTCCCCT GATCATCAAG AATCTTAAGA TAGAAGACTC AGATACTTAC         400

ATCTGTGAAG TGGAGGACCA GAAGGAGGAG GTGCAATTGC TAGTGTTCGG         450
```

-continued

ATTGACTGCC AACTCTGACA CCCACCTGCT TCAGGGGCAG AGCCTGACCC 500

TGACCTTGGA GAGCCCCCCT GGTAGTAGCC CCTCAGTGCA ATGTAGGAGT 550

CCAAGGGGTA AAAACATACA GGGGGGGAAG ACCCTCTCCG TGTCTCAGCT 600

GGAGCTCCAG GATAGTGGCA CCTGGACATG CACTGTCTTG CAGAACCAGA 650

AGAAGGTGGA GTTCAAAATA GACATCGTGG TGCTAGCTTT CCAGAAGGCC 700

TCCAGCATAG TCTATAAGAA AGAGGGGGAA CAGGTGGAGT TCTCCTTCCC 750

ACTCGCCTTT ACAGTTGAAA AGCTGACGGG CAGTGGCGAG CTGTGGTGGC 800

AGGCGGAGAG GGCTTCCTCC TCCAAGTCTT GGATCACCTT TGACCTGAAG 850

AACAAGGAAG TGTCTGTAAA ACGGGTTACC CAGGACCCTA AGCTCCAGAT 900

GGGCAAGAAG CTCCCGCTCC ACCTCACCCT GCCCCAGGCC TTGCCTCAGT 950

ATGCTGGCTC TGGAAACCTC ACCCTGGCCC TTGAAGCGAA AACAGGAAAG 1000

TTGCATCAGG AAGTGAACCT GGTGGTGATG AGAGCCACTC AGCTCCAGAA 1050

AAATTTGACC TGTGAGGTGT GGGGACCCAC CTCCCCTAAG CTGATGCTGA 1100

GTTTGAAACT GGAGAACAAG GAGGCAAAGG TCTCGAAGCG GGAGAAGGCG 1150

GTGTGGGTGC TGAACCCTGA GGCGGGGATG TGGCAGTGTC TGCTGAGTGA 1200

CTCGGGACAG GTCCTGCTGG AATCCAACAT CAAGGTTCTG CCCACATGGT 1250

CCACCCCGAG CTTTAATGCG GTAGTTTATC ACAGTTAAAT TGCTAACGCA 1300

GTCAGGCACC GTGTATGAAA TCTAACAATG CGCTCATCGT CATCCTCGGC 1350

ACCGTCACCC TGGATGCTGT AGGCATAGGC TTGGTTATGC CGGTACTGCC 1400

GGGCCTCTTG CGGGAT 1416

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1416 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

TTAAGTTCGG GTCTCGGGAC GGTAAAGACA CCCGAGTCCA GGGATGACGA 50

GTCGGGGAAG GAGGGAGCCG TTCCGGTGTT ACTTGGCCCC TCAGGGAAAA 100

TCCGTGAACG AAGACCACGA CGTTGACCGC GAGGAGGGTC GTCGGTGAGT 150

CCCTTTGTTT CACCACGACC CGTTTTTTCC CCTATGTCAC CTTGACTGGA 200

CATGTCGAAG GGTCTTCTTC TCGTATGTTA AGGTGACCTT TTTGAGGTTG 250

GTCTATTTCT AAGACCCTTT AGTCCCGAGG AAGAATTGAT TTCCAGGTAG 300

GTTCGACTTA CTAGCGCGAC TGAGTTCTTC TTCGGAAACC CTGGTTCCTT 350

TGAAAGGGGA CTAGTAGTTC TTAGAATTCT ATCTTCTGAG TCTATGAATG 400

TAGACACTTC ACCTCCTGGT CTTCCTCCTC CACGTTAACG ATCACAAGCC 450

TAACTGACGG TTGAGACTGT GGGTGGACGA AGTCCCCGTC TCGGACTGGG 500

ACTGGAACCT CTCGGGGGGA CCATCATCGG GGAGTCACGT TACATCCTCA 550

GGTTCCCCAT TTTTGTATGT CCCCCCCTTC TGGGAGAGGC ACAGAGTCGA 600

CCTCGAGGTC CTATCACCGT GGACCTGTAC GTGACAGAAC GTCTTGGTCT 650

TCTTCCACCT CAAGTTTTAT CTGTAGCACC ACGATCGAAA GGTCTTCCGG 700

-continued

```
AGGTCGTATC AGATATTCTT TCTCCCCCTT GTCCACCTCA AGAGGAAGGG          750

TGAGCGGAAA TGTCAACTTT TCGACTGCCC GTCACCGCTC GACACCACCG          800

TCCGCCTCTC CCGAAGGAGG AGGTTCAGAA CCTAGTGGAA ACTGGACTTC          850

TTGTTCCTTC ACAGACATTT TGCCCAATGG GTCCTGGGAT TCGAGGTCTA          900

CCCGTTCTTC GAGGGCGAGG TGGAGTGGGA CGGGGTCCGG AACGGAGTCA          950

TACGACCGAG ACCTTTGGAG TGGGACCGGG AACTTCGCTT TTGTCCTTTC         1000

AACGTAGTCC TTCACTTGGA CCACCACTAC TCTCGGTGAG TCGAGGTCTT         1050

TTTAAACTGG ACACTCCACA CCCCTGGGTG GAGGGGATTC GACTACGACT         1100

CAAACTTTGA CCTCTTGTTC CTCCGTTTCC AGAGCTTCGC CCTCTTCCGC         1150

CACACCCACG ACTTGGGACT CCGCCCCTAC ACCGTCACAG ACGACTCACT         1200

GAGCCCTGTC CAGGACGACC TTAGGTTGTA GTTCCAAGAC GGGTGTACCA         1250

GGTGGGGCTC GAAATTACGC CATCAAATAG TGTCAATTTA ACGATTGCGT         1300

CAGTCCGTGG CACATACTTT AGATTGTTAC GCGAGTAGCA GTAGGAGCCG         1350

TGGCAGTGGG ACCTACGACA TCCGTATCCG AACCAATACG GCCATGACGG         1400

CCCGGAGAAC GCCCTA                                              1416
```

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 434 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

```
Met Gly Gly Thr Ala Ala Arg Leu Gly Ala Val Ile Leu Phe Val
 1               5                  10                  15

Val Ile Val Gly Leu His Gly Val Arg Gly Lys Tyr Ala Leu Ala
            20                  25                  30

Asp Ala Ser Leu Lys Met Ala Asp Pro Asn Arg Phe Arg Gly Lys
        35                  40                  45

Asp Leu Pro Val Leu Asp Gln Leu Leu Glu Gln Gly Asn Lys Val
    50                  55                  60

Val Leu Gly Lys Lys Gly Asp Thr Val Glu Leu Thr Cys Thr Ala
65                  70                  75

Ser Gln Lys Lys Ser Ile Gln Phe His Trp Lys Asn Ser Asn Gln
            80                  85                  90

Ile Lys Ile Leu Gly Asn Gln Gly Ser Phe Leu Thr Lys Gly Pro
        95                  100                 105

Ser Lys Leu Asn Asp Arg Ala Asp Ser Arg Arg Ser Leu Trp Asp
    110                 115                 120

Gln Gly Asn Phe Pro Leu Ile Ile Lys Asn Leu Lys Ile Glu Asp
125                 130                 135

Ser Asp Thr Tyr Ile Cys Glu Val Glu Asp Gln Lys Glu Glu Val
            140                 145                 150

Gln Leu Leu Val Phe Gly Leu Thr Ala Asn Ser Asp Thr His Leu
        155                 160                 165

Leu Gln Gly Gln Ser Leu Thr Leu Thr Leu Glu Ser Pro Pro Gly
    170                 175                 180

Ser Ser Pro Ser Val Gln Cys Arg Ser Pro Arg Gly Lys Asn Ile
185                 190                 195
```

-continued

Gln Gly Gly Lys Thr Leu Ser Val Ser Gln Leu Glu Leu Gln Asp
200 205 210

Ser Gly Thr Trp Thr Cys Thr Val Leu Gln Asn Gln Lys Lys Val
215 220 225

Glu Phe Lys Ile Asp Ile Val Val Leu Ala Phe Gln Lys Ala Ser
230 235 240

Ser Ile Val Tyr Lys Lys Glu Gly Glu Gln Val Glu Phe Ser Phe
245 250 255

Pro Leu Ala Phe Thr Val Glu Lys Leu Thr Gly Ser Gly Glu Leu
260 265 270

Trp Trp Gln Ala Glu Arg Ala Ser Ser Ser Lys Ser Trp Ile Thr
275 280 285

Phe Asp Leu Lys Asn Lys Glu Val Ser Val Lys Arg Val Thr Gln
290 295 300

Asp Pro Lys Leu Gln Met Gly Lys Lys Leu Pro Leu His Leu Thr
305 310 315

Leu Pro Gln Ala Leu Pro Gln Tyr Ala Gly Ser Gly Asn Leu Thr
320 325 330

Leu Ala Leu Glu Ala Lys Thr Gly Lys Leu His Gln Glu Val Asn
335 340 345

Leu Val Val Met Arg Ala Thr Gln Leu Gln Lys Asn Leu Thr Cys
350 355 360

Glu Val Trp Gly Pro Thr Ser Pro Lys Leu Met Leu Ser Leu Lys
365 370 375

Leu Glu Asn Lys Glu Ala Lys Val Ser Lys Arg Glu Lys Ala Val
380 385 390

Trp Val Leu Asn Pro Glu Ala Gly Met Trp Gln Cys Leu Leu Ser
395 400 405

Asp Ser Gly Gln Val Leu Leu Glu Ser Asn Ile Lys Val Leu Pro
410 415 420

Thr Trp Ser Thr Pro Ser Phe Asn Ala Val Val Tyr His Ser
425 430 434

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1508 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

AAGCTTCAGC GCGAACGACC AACTACCCCG ATCATCAGTT ATCCTTAAGG 50

TCTCTTTTGT GTGGTGCGTT CCGGTATGGG GGGGACTGCC GCCAGGTIGG 100

GGGCCGTGAT TTTGTTTGTC GTCATAGTGG GCCTCCATGG GGTCCGCGGC 150

AAATATGCCT TGGCGGATGC CTCTCTCAAG ATGGCCGACC CCAATCGATT 200

TCGCGGCAAA GACCTTCCGG TCCTGGACCA GCTGCTCGAG CAGGGAAACA 250

AAGTGGTGCT GGGCAAAAAA GGGGATACAG TGGAACTGAC CTGTACAGCT 300

TCCCAGAAGA AGAGCATACA ATTCCACTGG AAAAACTCCA ACCAGATAAA 350

GATTCTGGGA AATCAGGGCT CCTTCTTAAC TAAAGGTCCA TCCAAGCTGA 400

ATGATCGCGC TGACTCAAGA AGAAGCCTTT GGGACCAAGG AAACTTTCCC 450

CTGATCATCA AGAATCTTAA GATAGAAGAC TCAGATACTT ACATCTGTGA 500

-continued

```
AGTGGAGGAC CAGAAGGAGG AGGTGCAATT GCTAGTGTTC GGATTGACTG          550
CCAACTCTGA CACCCACCTG CTTCAGGGGC AGAGCCTGAC CCTGACCTTG          600
GAGAGCCCCC CTGGTAGTAG CCCCTCAGTG CAATGTAGGA GTCCAAGGGG          650
TAAAAACATA CAGGGGGGGA AGACCCTCTC CGTGTCTCAG CTGGAGCTCC          700
AGGATAGTGG CACCTGGACA TGCACTGTCT TGCAGAACCA GAAGAAGGTG          750
GAGTTCAAAA TAGACATCGT GGTGCTAGCT TTCCAGAAGG CCTCCAGCAT          800
AGTCTATAAG AAAGAGGGGG AACAGGTGGA GTTCTCCTTC CCACTCGCCT          850
TTACAGTTGA AAAGCTGACG GGCAGTGGCG AGCTGTGGTG GCAGGCGGAG          900
AGGGCTTCCT CCTCCAAGTC TTGGATCACC TTTGACCTGA AGAACAAGGA          950
AGTGTCTGTA AAACGGGTTA CCCAGGACCC TAAGCTCCAG ATGGGCAAGA         1000
AGCTCCCGCT CCACCTCACC CTGCCCCAGG CCTTGCCTCA GTATGCTGGC         1050
TCTGGAAACC TCACCCTGGC CCTTGAAGCG AAAACAGGAA AGTTGCATCA         1100
GGAAGTGAAC CTGGTGGTGA TGAGAGCCAC TCAGCTCCAG AAAAATTTGA         1150
CCTGTCAGGT GTGGGGACCC ACCTCCCCTA AGCTGATGCT GAGTTTGAAA         1200
CTGGAGAACA AGGAGGCAAA GGTCTCGAAG CGGGAGAAGG CGGTGTGGGT         1250
GCTGAACCCT GAGGCGGGGA TGTGGCAGTG TCTGCTGAGT GACTCGGGAC         1300
AGGTCCTGCT GGAATCCAAC ATCAAGGTTC TGCCCACATG GTCCACCCCG         1350
AGCTTTAATG CGGTAGTTTA TCACAGTTAA ATTGCTAACG CAGTCAGGCA         1400
CCGTGTATGA AATCTAACAA TGCGCTCATC GTCATCCTCG GCACCGTCAC         1450
CCTGGATGCT GTAGGCATAG GCTTGGTTAT GCCGGTACTG CCGGGCCTCT         1500
TGCGGGAT                                                      1508
```

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1508 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

```
TTCGAAGTCG CGCTTGCTGG TTGATGGGGC TAGTAGTCAA TAGGAATTCC           50
AGAGAAAACA CACCACGCAA GGCCATACCC CCCCTGACGG CGGTCCAACC          100
CCCGGCACTA AAACAAACAG CAGTATCACC CGGAGGTACC CCAGGCGCCG          150
TTTATACGGA ACCGCCTACG GAGAGAGTTC TACCGGCTGG GGTTAGCTAA          200
AGCGCCGTTT CTGGAAGGCC AGGACCTGGT CGACGAGCTC GTCCCTTTGT          250
TTCACCACGA CCCGTTTTTT CCCCTATGTC ACCTTGACTG GACATGTCGA          300
AGGGTCTTCT TCTCGTATGT TAAGGTGACC TTTTTGAGGT TGGTCTATTT          350
CTAAGACCCT TTAGTCCCGA GGAAGAATTG ATTTCCAGGT AGGTTCGACT          400
TACTAGCGCG ACTGAGTTCT TCTTCGGAAA CCCTGGTTCC TTTGAAAGGG          450
GACTAGTAGT TCTTAGAATT CTATCTTCTG AGTCTATGAA TGTAGACACT          500
TCACCTCCTG GTCTTCCTCC TCCACGTTAA CGATCACAAG CCTAACTGAC          550
GGTTGAGACT GTGGGTGGAC GAAGTCCCCG TCTCGGACTG GGACTGGAAC          600
CTCTCGGGGG GACCATCATC GGGGAGTCAC GTTACATCCT CAGGTTCCCC          650
```

-continued

```
ATTTTTGTAT GTCCCCCCCT TCTGGGAGAG GCACAGAGTC GACCTCGAGG          700

TCCTATCACC GTGGACCTGT ACGTGACAGA ACGTCTTGGT CTTCTTCCAC          750

CTCAAGTTTT ATCTGTAGCA CCACGATCGA AAGGTCTTCC GGAGGTCGTA          800

TCAGATATTC TTTCTCCCCC TTGTCCACCT CAAGAGGAAG GGTGAGCGGA          850

AATGTCAACT TTTCGACTGC CCGTCACCGC TCGACACCAC CGTCCGCCTC          900

TCCCGAAGGA GGAGGTTCAG AACCTAGTGG AAACTGGACT TCTTGTTCCT          950

TCACAGACAT TTTGCCCAAT GGGTCCTGGG ATTCGAGGTC TACCCGTTCT         1000

TCGAGGGCGA GGTGGAGTGG GACGGGGTCC GGAACGGAGT CATACGACCG         1050

AGACCTTTGG AGTGGGACCG GGAACTTCGC TTTTGTCCTT TCAACGTAGT         1100

CCTTCACTTG GACCACCACT ACTCTCGGTG AGTCGAGGTC TTTTTAAACT         1150

GGACACTCCA CACCCCTGGG TGGAGGGGAT TCGACTACGA CTCAAACTTT         1200

GACCTCTTGT TCCTCCGTTT CCAGAGCTTC GCCCTCTTCC GCCACACCCA         1250

CGACTTGGGA CTCCGCCCCT ACACCGTCAC AGACGACTCA CTGAGCCCTG         1300

TCCAGGACGA CCTTAGGTTG TAGTTCCAAG ACGGGTGTAC CAGGTGGGGC         1350

TCGAAATTAC GCCATCAAAT AGTGTCAATT TAACGATTGC GTCAGTCCGT         1400

GGCACATACT TTAGATTGTT ACGCGAGTAG CAGTAGGAGC CGTGGCAGTG         1450

GGACCTACGA CATCCGTATC CGAACCAATA CGGCCATGAC GGCCCGGAGA         1500

ACGCCCTA                                                      1508
```

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 371 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

```
Val Thr Ala Ala Asp Thr Ala Val Tyr Tyr Cys Ala Arg Ala Thr
 1               5                  10                  15

Phe Cys Leu Trp Tyr Arg Glu Arg Pro Pro Cys Trp Ile Asp Pro
            20                  25                  30

Trp Gly Leu Gly Thr Leu Val Thr Val Ser Ser Ala Ser Thr Lys
        35                  40                  45

Gly Pro Ser Val Phe Pro Leu Ala Pro Ser Ser Lys Ser Thr Ser
    50                  55                  60

Gly Gly Thr Ala Ala Leu Gly Cys Leu Val Lys Asp Tyr Phe Pro
65                  70                  75

Glu Pro Val Thr Val Ser Trp Asn Ser Gly Ala Leu Thr Ser Gly
            80                  85                  90

Val His Thr Phe Pro Ala Val Leu Gln Ser Ser Gly Leu Tyr Ser
        95                  100                 105

Leu Ser Ser Val Val Thr Val Pro Ser Ser Ser Leu Gly Thr Gln
    110                 115                 120

Thr Tyr Ile Cys Asn Val Asn His Lys Pro Ser Asn Thr Lys Val
125                 130                 135

Asp Lys Lys Val Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys
            140                 145                 150

Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe
        155                 160                 165
```

-continued

```
Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr
                170                 175                 180

Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro
                185                 190                 195

Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn
                200                 205                 210

Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg
                215                 220                 225

Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly
                230                 235                 240

Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro
                245                 250                 255

Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro
                260                 265                 270

Gln Val Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn
                275                 280                 285

Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp
                290                 295                 300

Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr
                305                 310                 315

Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu
                320                 325                 330

Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn
                335                 340                 345

Val Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr
                350                 355                 360

Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
                365                 370 371
```

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1135 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

```
GAATTCTGTC ACTGCCGCGG ACACGGCCGT ATATTACTGT GCGAGAGCCA              50

CCTTTTGCCT ATGGTACAGG GAGCGTCCCC CTTGTTGGAT CGACCCCTGG             100

GGCCTGGGAA CCCTGGTCAC CGTCTCCTCG GCCTCCACCA AGGGCCCATC             150

GGTCTTCCCC CTGGCACCCT CCTCCAAGAG CACCTCTGGG GGCACAGCGG             200

CCCTGGGCTG CCTGGTCAAG GACTACTTCC CCGAACCGGT GACGGTGTCG             250

TGGAACTCAG GCGCCCTGAC CAGCGGCGTG CACACCTTCC CGGCTGTCCT             300

ACAGTCCTCA GGACTCTACT CCCTCAGCAG CGTGGTGACC GTGCCCTCCA             350

GCAGCTTGGG CACCCAGACC TACATCTGCA ACGTGAATCA CAAGCCCAGC             400

AACACCAAGG TGGACAAGAA AGTTGAGCCC AAATCTTGTG ACAAAACTCA             450

CACATGCCCA CCGTGCCCAG CACCTGAACT CCTGGGGGGA CCGTCAGTCT             500

TCCTCTTCCC CCCAAAACCC AAGGACACCC TCATGATCTC CCGGACCCCT             550

GAGGTCACAT GCGTGGTGGT GGACGTGAGC CACGAAGACC CTGAGGTCAA             600

GTTCAAGTGG TACGTGGACG GCGTGGAGGT GCATAATGCC AAGACAAAGC             650
```

-continued

```
CGCGGGAGGA GCAGTACAAC AGCACGTACC GGGTGGTCAG CGTCCTCACC          700

GTCCTGCACC AGGACTGGCT GAATGGCAAG GAGTACAAGT GCAAGGTCTC          750

CAACAAAGCC CTCCCAGCCC CCATCGAGAA AACCATCTCC AAAGCCAAAG          800

GGCAGCCCCG AGAACCACAG GTGTACACCC TGCCCCCATC CCGGGATGAG          850

CTGACCAAGA ACCAGGTCAG CCTGACCTGC CTGGTCAAAG GCTTCTATCC          900

CAGCGACATC GCCGTGGAGT GGGAGAGCAA TGGGCAGCCG GAGAACAACT          950

ACAAGACCAC GCCTCCCGTG CTGGACTCCG ACGGCTCCTT CTTCCTCTAC         1000

AGCAAGCTCA CCGTGGACAA GAGCAGGTGG CAGCAGGGGA ACGTCTTCTC         1050

ATGCTCCGTG ATGCATGAGG CTCTGCACAA CCACTACACG CAGAAGAGCC         1100

TCTCCCTGTC TCCGGGTAAA TGAGTGCGAC GGCCG                         1135
```

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1142 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

```
CTTAAGACAG TGACGGCGCC TGTGCCGGCA TATAATGACA CGCTCTCGGT           50

GGAAAACGGA TACCATGTCC CTCGCAGGGG GAACAACCTA GCTGGGGACC          100

CCGGACCCTT GGGACCAGTG GCAGAGGAGC CGGAGGTGGT TCCCGGGTAG          150

CCAGAAGGGG GACCGTGGGA GGAGGTTCTC GTGGAGACCC CCGTGTCGCC          200

GGGACCCGAC GGACCAGTTC CTGATGAAGG GGCTTGGCCA CTGCCACAGC          250

ACCTTGAGTC CGCGGGACTG GTCGCCGCAC GTGTGGAAGG GCCGACAGGA          300

TGTCAGGAGT CCTGAGATGA GGGAGTCGTC GCACCACTGG CACGGGAGGT          350

CGTCGAACCC GTGGGTCTGG ATGTAGACGT TGCACTTAGT GTTCGGGTCG          400

TTGTGGTTCC ACCTGTTCTT TCAACTCGGG TTTAGAACAC TGTTTTGAGT          450

GTGTACGGGT GGCACGGGTC GTGGACTTGA GGACCCCCCT GGCAGTCAGA          500

AGGAGAAGGG GGGTTTTGGG TTCCTGTGGG AGTACTAGAG GGCCTGGGGA          550

CTCCAGTGTA CGCACCACCA CCTGCACTCG GTGCTTCTGG GACTCCAGTT          600

CAAGTTGACC ATGCACCTGC CGCACCTCCA CGTATTACGG TTCTGTTTCG          650

GCGCCCTCCT CGTCATGTTG TCGTGCATGG CCCACCAGTC GCAGGAGTGG          700

CAGGACGTGG TCCTGACCGA CTTACCGTTC CTCATGTTCA CGTTCCAGAG          750

GTTGTTTCGG GAGGGTCGGG GGTAGCTCTT TTGGTAGAGG TTTCGGTTTC          800

CCGTCGGGGC TCTTGGTGTC CACATGTGGG ACGGGGGTAG GGCCCTACTC          850

GACTGGTTCT TGGTCCAGTC GGACTGGACG GACCAGTTTC CGAAGATAGG          900

GTCGCTGTAG CGGCACCTCA CCCTCTCGTT ACCCGTCGGC CTCTTGTTGA          950

TGTTCTGGTG CGGAGGGCAC GACCTGAGGC TGCCGAGGAA GAAGGAGATG         1000

TCGTTCGAGT GGCACCTGTT CTCGTCCACC GTCGTCCCCT TGCAGAAGAG         1050

TACGAGGCAC TACGTACTAC GTACTCCGAG ACGTGTTGGT GATGTGCGTC         1100

TTCTCGGAGA GGGACAGAGG CCCATTTACT CACGCTGCCG GC                 1142
```

-continued

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 143 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

Thr Leu Thr Ile Ser Gly Leu Gln Pro Glu Asp Phe Ala Thr Tyr
1 5 10 15

Tyr Cys Gln Gln Tyr Lys Ser Leu Ser Leu Thr Phe Gly Gly Gly
20 25 30

Thr Lys Val Glu Ile Lys Arg Thr Val Ala Ala Pro Ser Val Phe
35 40 45

Ile Phe Pro Pro Ser Asp Glu Gln Leu Lys Ser Gly Thr Ala Ser
50 55 60

Val Val Cys Leu Leu Asn Asn Phe Tyr Pro Arg Glu Ala Leu Val
65 70 75

Gln Trp Leu Val Asp Asn Ala Leu Gln Ser Gly Asn Ser Gln Glu
80 85 90

Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr Tyr Ser Leu Ser
95 100 105

Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys His Lys Val
110 115 120

Tyr Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser Pro Val Thr
125 130 135

Lys Ser Phe Asn Arg Gly Glu Cys
140 143

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 468 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

GAATTCACTC TCACCATCAG CGGCCTGCAG CCTGAAGATT TTGCAACTTA 50

TTACTGCCAA CAGTATAAGA GTTTGTCGCT CACTTTCGGC GGAGGGACCA 100

AGGTGGAGAT CAAACGAACT GTGGCTGCAC CATCTGTCTT CATCTTCCCG 150

CCATCTGATG AGCAGTTGAA ATCTGGAACT GCCTCTGTTG TGTGCCTGCT 200

GAATAACTTC TATCCCAGAG AGGCCAAAGT ACAGTGGAAG GTGGATAACG 250

CCCTCCAATC GGGTAACTCC CAGGAGAGTG TCACAGAGCA GGACAGCAAG 300

GACAGCACCT ACAGCCTCAG CAGCACCCTG ACGCTGAGCA AAGCAGACTA 350

CGAGAAACAC AAAGTCTACG CCTGCGAAGT CACCCATCAG GGCCTGAGCT 400

CGCCCGTCAC AAAGAGCTTC AACAGGGGAG AGTGTTAGAG GGAGAAGTGC 450

CCCCACCTGC TCCTCAGT 468

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 468 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

-continued

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

CTTAAGTGAG AGTGGTAGTC GCCGGACGTC GGACTTCTAA AACGTTGAAT 50

AATGACGGTT GTCATATTCT CAAACAGCGA GTGAAAGCCG CCTCCCTGGT 100

TCCACCTCTA GTTTGCTTGA CACCGACGTG GTAGACAGAA GTAGAAGGGC 150

GGTAGACTAC TCGTCAACTT TAGACCTTGA CGGAGACAAC ACACGGACGA 200

CTTATTGAAG ATAGGGTCTC TCCGGTTTCA TGTCACCTTC CACCTATTGC 250

GGGAGGTTAG CCCATTGAGG GTCCTCTCAC AGTGTCTCGT CCTGTCGTTC 300

CTGTCGTGGA TGTCGGAGTC GTCGTCGGAC TGCGACTCGT TTCGTCTGAT 350

GCTCTTTGTG TTTCAGATGC GGACGCTTCA GTGGGTAGTC CCGGACTCGA 400

GCGGGCAGTG TTTCTCGAAG TTGTCCCCTC TCACAATCTC CCTCTTCACG 450

GGGGTGGACG AGGAGTCA 468

(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 44 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

AATTCAAGCC CAGAGCCCTG CCATTTCTGT GGGCTCAGGT CCCT 44

(2) INFORMATION FOR SEQ ID NO: 14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 50 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

ACTGCTCAGC CCCTTCCTCC CTCGGCAAGG CCACAATGAA CCGGGGAGTC 50

(2) INFORMATION FOR SEQ ID NO: 15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 47 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

CCTTTTAGGC ACTTGCTTCT GGTGCTGCAA CTGGCGCTCC TCCCAGC 47

(2) INFORMATION FOR SEQ ID NO: 16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 61 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

AGCCACTCAG GGAAACAAAG TGGTGCTGGG CAAAAAAGGG GATACAGTGG 50

AACTGACCTG T 61

(2) INFORMATION FOR SEQ ID NO: 17:

-continued

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 51 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

ACAGGTCAGT TCCACTGTAT CCCCTTTTTT GCCCAGCACC ACTTTGTTTC 50

C 51

(2) INFORMATION FOR SEQ ID NO: 18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 47 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

CTGAGTGGCT GCTGGGAGGA GCGCCAGTTG CAGCACCAGA AGCAAGT 47

(2) INFORMATION FOR SEQ ID NO: 19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 50 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

GCCTAAAAGG GACTCCCCGG TTCATTGTGG CCTTGCCGAG GGAGGAAGGG 50

(2) INFORMATION FOR SEQ ID NO: 20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 50 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

GCTGAGCAGT AGGGACCTGA GCCCACAGAA ATGGCAGGGC TCTGGGCTTG 50

(2) INFORMATION FOR SEQ ID NO: 21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 59 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

CTGCTCGAGC AGGGAAACAA AGTGGTGCTG GGCAAAAAAG GGGATACAGT 50

GGAACTGAC 59

(2) INFORMATION FOR SEQ ID NO: 22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 59 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

ACAGGTCAGT TCCACTGTAT CCCCTTTTTT GCCCAGCACC ACTTTGTTTC 50

-continued

CCTGCTCGA 59

(2) INFORMATION FOR SEQ ID NO: 23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

CGTGATAGAA GCTTTCTAGA G 21

(2) INFORMATION FOR SEQ ID NO: 24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 51 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

CCCTTTTTTG CCCAGCACCA CCTTCTTGCC CTGAGTGGCT GCTGGGAGGA 50

G 51

(2) INFORMATION FOR SEQ ID NO: 25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

CCACCTTCTT GCCCTG 16

What is claimed is:

1. A dimer of two polypeptides wherein each polypeptide of the dimer has the formula AC_H wherein A is at least one immunoglobulin like domain of an adheson and C_H is a constant region of an immunoglobulin heavy chain and A is fused at its C-terminus to the N-terminus of C_H.

2. The polypeptide dimer of claim 1 which is a homodimer.

3. The polypeptide dimer of claim 1 wherein A comprises more than one immunoglobulin-like domain(s) of the adheson.

4. The polypeptide dimer of claim 1 wherein A contains the ligand binding site of an adheson.

5. The polypeptide dimer of claim 1 wherein the constant region of the immunoglobulin is selected from constant regions of IgG, IgA, IgE, IgD and IgM.

6. The polypeptide dimer of claim 2 wherein the constant region of the immunoglobulin is selected from constant regions of IgG, IgA, IgE, IgD and IgM.

7. The polypeptide of claim 5 wherein the constant region of the immunoglobulin is the constant region of an IgG.

8. The polypeptide of claim 6 wherein the constant region of the immunoglobulin is the constant region of an IgG.

9. The polypeptide of claim 7 wherein the IgG is an IgG1 subtype.

10. The polypeptide of claim 8 wherein the IgG is an IgG1 subtype.

11. The polypeptide of claim 6 wherein A is at least one immunoglobulin like domain of an adheson selected from the group consisting of CD1, CD2, CD4, CD8, CD28, OX-2, I-CAM, N-CAM, LFA-3 and the high affinity IgE receptor.

12. The polypeptide of claim 8 wherein A is an immunoglobulin like domain of an adheson selected from the group consisting of CD1, CD2, CD4, CD8, CD28, OX-2, I-CAM, N-CAM, LFA-3 and the high affinity IgE receptor.

13. The polypeptide of claim 6 wherein the adheson is CD4.

14. The polypeptide of claim 8 wherein the adheson is CD4.

15. The polypeptide of claim 6 wherein the adheson is LFA-3.

16. The polypeptide of claim 8 wherein the adheson is LFA-3.

17. The polypeptide of claim 9 wherein the constant region of the IgG1 retains at least functionally active hinge, CH2 and CH3 domains.

18. The polypeptide of claim 17 wherein the adheson is CD4.

19. The polypeptide of claim 17 wherein the adheson is LFA-3.

* * * * *