US007842458B2

(12) **United States Patent**
Kanner et al.

(10) **Patent No.: US 7,842,458 B2**
(45) **Date of Patent: Nov. 30, 2010**

(54) **NUCLEIC ACIDS AND CORRESPONDING PROTEINS ENTITLED 58P1D12 USEFUL IN TREATMENT AND DETECTION OF CANCER**

(75) Inventors: **Steven B. Kanner**, Santa Monica, CA (US); **Aya Jakobovits**, Beverly Hills, CA (US); **Pia M. Challita-Eid**, Encino, CA (US); **Wangmao Ge**, Los Angeles, CA (US); **Arthur B. Raitano**, Los Angeles, CA (US)

(73) Assignee: **Agensys, Inc.**, Santa Monica, CA (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 1247 days.

(21) Appl. No.: **10/919,654**

(22) Filed: **Aug. 16, 2004**

(65) **Prior Publication Data**

US 2005/0136435 A1 Jun. 23, 2005

Related U.S. Application Data

(60) Continuation-in-part of application No. 10/460,512, filed on Jun. 11, 2003, now Pat. No. 7,153,940, which is a division of application No. 09/638,203, filed on Aug. 11, 2000, now Pat. No. 6,602,501.

(60) Provisional application No. 60/148,935, filed on Aug. 12, 1999.

(51) **Int. Cl.**

C12Q 1/68	(2006.01)
G01N 33/53	(2006.01)
G01N 33/574	(2006.01)
C07H 21/04	(2006.01)
C07K 14/435	(2006.01)

(52) **U.S. Cl.** **435/6**; 435/7.1; 435/7.23; 436/63; 436/64; 436/501

(58) **Field of Classification Search** .... None
See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

5,681,923	A	10/1997	Platt
6,117,977	A	9/2000	Lasky et al.
6,602,501	B1	8/2003	Afar et al.
7,141,652	B1 *	11/2006	Goddard et al. 530/387.1
2002/0119130	A1	8/2002	Eaton et al.
2002/0182638	A1	12/2002	Eaton et al.
2002/0183493	A1	12/2002	Eaton et al.
2002/0183494	A1	12/2002	Eaton et al.
2002/0192751	A1	12/2002	Desnoyers et al.
2003/0004311	A1	1/2003	Baker et al.
2003/0008348	A1	1/2003	Desnoyers et al.
2003/0009012	A1	1/2003	Eaton et al.
2003/0009013	A1	1/2003	Eaton et al.
2003/0013855	A1	1/2003	Eaton et al.
2003/0017563	A1	1/2003	Baker et al.
2003/0018168	A1	1/2003	Eaton et al.
2003/0018173	A1	1/2003	Eaton et al.
2003/0018183	A1	1/2003	Eaton et al.
2003/0022239	A1	1/2003	Baker et al.
2003/0022328	A1	1/2003	Baker et al.
2003/0022331	A1	1/2003	Baker et al.
2003/0023042	A1	1/2003	Eaton et al.
2003/0027212	A1	2/2003	Eaton et al.
2003/0027249	A1	2/2003	Desnoyers et al.
2003/0027270	A1	2/2003	Baker et al.
2003/0027275	A1	2/2003	Baker et al.
2003/0027276	A1	2/2003	Baker et al.
2003/0027986	A1	2/2003	Eaton et al.
2003/0027992	A1	2/2003	Eaton et al.
2003/0027993	A1	2/2003	Eaton et al.
2003/0032061	A1	2/2003	Desnoyers et al.
2003/0032155	A1	2/2003	Baker et al.
2003/0032156	A1	2/2003	Baker et al.
2003/0036114	A1	2/2003	Desnoyers et al.
2003/0036179	A1	2/2003	Baker et al.
2003/0036180	A1	2/2003	Baker et al.
2003/0036634	A1	2/2003	Eaton et al.
2003/0044842	A1	3/2003	Desnoyers et al.
2003/0044945	A1	3/2003	Baker et al.
2003/0045684	A1	3/2003	Eaton et al.
2003/0049733	A1	3/2003	Desnoyers et al.
2003/0049734	A1	3/2003	Desnoyers et al.
2003/0049735	A1	3/2003	Eaton et al.
2003/0049816	A1	3/2003	Baker et al.
2003/0049817	A1	3/2003	Baker et al.
2003/0050462	A1	3/2003	Eaton et al.
2003/0050465	A1	3/2003	Eaton et al.
2003/0054516	A1	3/2003	Baker et al.
2003/0054517	A1	3/2003	Baker et al.

(Continued)

FOREIGN PATENT DOCUMENTS

EP	1074617	* 2/2001

(Continued)

OTHER PUBLICATIONS

Singh et al (Glycobiology, 2001, vol. 11, pp. 587-592).*

(Continued)

Primary Examiner—Karen A Canella
(74) *Attorney, Agent, or Firm*—Morrison & Foerster LLP

(57) **ABSTRACT**

A novel gene 58P1D12 and its encoded protein, and variants thereof, are described wherein 58P1D12 exhibits tissue specific expression in normal adult tissue, and is aberrantly expressed in the cancers listed in Table 1. Consequently, 58P1D12 provides a diagnostic, prognostic, prophylactic and/or therapeutic target for cancer. The 58P1D12 gene or fragment thereof, or its encoded protein, or variants thereof, or a fragment thereof, can be used to elicit a humoral or cellular immune response; antibodies or T cells reactive with 58P1D12 can be used in active or passive immunization.

11 Claims, 52 Drawing Sheets

U.S. PATENT DOCUMENTS

2003/0059909	A1	3/2003	Baker et al.
2003/0060600	A1	3/2003	Eaton et al.
2003/0060601	A1	3/2003	Eaton et al.
2003/0060602	A1	3/2003	Eaton et al.
2003/0065143	A1	4/2003	Eaton et al.
2003/0065161	A1	4/2003	Eaton et al.
2003/0068793	A1	4/2003	Baker et al.
2003/0068794	A1	4/2003	Baker et al.
2003/0068795	A1	4/2003	Baker et al.
2003/0068796	A1	4/2003	Baker et al.
2003/0068797	A1	4/2003	Baker et al.
2003/0068798	A1	4/2003	Baker et al.
2003/0069394	A1	4/2003	Eaton et al.
2003/0073210	A1	4/2003	Baker et al.
2003/0073211	A1	4/2003	Baker et al.
2003/0073212	A1	4/2003	Baker et al.
2003/0073213	A1	4/2003	Baker et al.
2003/0073214	A1	4/2003	Baker et al.
2003/0073215	A1	4/2003	Baker et al.
2003/0073216	A1	4/2003	Baker et al.
2003/0077710	A1	4/2003	Baker et al.
2003/0078387	A1	4/2003	Eaton et al.
2003/0082686	A1	5/2003	Baker et al.
2003/0082687	A1	5/2003	Baker et al.
2003/0082689	A1	5/2003	Baker et al.
2003/0082690	A1	5/2003	Baker et al.
2003/0082691	A1	5/2003	Baker et al.
2003/0082692	A1	5/2003	Baker et al.
2003/0082693	A1	5/2003	Baker et al.
2003/0082694	A1	5/2003	Baker et al.
2003/0082695	A1	5/2003	Baker et al.
2003/0082696	A1	5/2003	Baker et al.
2003/0082697	A1	5/2003	Baker et al.
2003/0082698	A1	5/2003	Baker et al.
2003/0082699	A1	5/2003	Baker et al.
2003/0082700	A1	5/2003	Baker et al.
2003/0082701	A1	5/2003	Baker et al.
2003/0082702	A1	5/2003	Baker et al.
2003/0082703	A1	5/2003	Baker et al.
2003/0082704	A1	5/2003	Baker et al.
2003/0082705	A1	5/2003	Baker et al.
2003/0082706	A1	5/2003	Baker et al.
2003/0082708	A1	5/2003	Baker et al.
2003/0082709	A1	5/2003	Baker et al.
2003/0082710	A1	5/2003	Baker et al.
2003/0082711	A1	5/2003	Baker et al.
2003/0082712	A1	5/2003	Baker et al.
2003/0082715	A1	5/2003	Baker et al.
2003/0082759	A1	5/2003	Baker et al.
2003/0082760	A1	5/2003	Baker et al.
2003/0082761	A1	5/2003	Baker et al.
2003/0082762	A1	5/2003	Baker et al.
2003/0082763	A1	5/2003	Baker et al.
2003/0082764	A1	5/2003	Baker et al.
2003/0082765	A1	5/2003	Baker et al.
2003/0082766	A1	5/2003	Baker et al.
2003/0083473	A1	5/2003	Eaton et al.
2003/0087344	A1	5/2003	Baker et al.
2003/0087345	A1	5/2003	Baker et al.
2003/0087346	A1	5/2003	Baker et al.
2003/0087347	A1	5/2003	Baker et al.
2003/0087349	A1	5/2003	Baker et al.
2003/0087350	A1	5/2003	Baker et al.
2003/0087351	A1	5/2003	Baker et al.
2003/0087352	A1	5/2003	Baker et al.
2003/0087353	A1	5/2003	Baker et al.
2003/0087354	A1	5/2003	Baker et al.
2003/0087355	A1	5/2003	Baker et al.
2003/0087357	A1	5/2003	Baker et al.
2003/0087358	A1	5/2003	Baker et al.
2003/0087359	A1	5/2003	Baker et al.
2003/0087360	A1	5/2003	Baker et al.
2003/0087361	A1	5/2003	Baker et al.
2003/0087362	A1	5/2003	Baker et al.
2003/0087363	A1	5/2003	Baker et al.
2003/0087364	A1	5/2003	Baker et al.
2003/0087365	A1	5/2003	Baker et al.
2003/0087366	A1	5/2003	Baker et al.
2003/0087367	A1	5/2003	Baker et al.
2003/0092063	A1	5/2003	Desnoyers et al.
2003/0092103	A1	5/2003	Baker et al.
2003/0092104	A1	5/2003	Baker et al.
2003/0092105	A1	5/2003	Baker et al.
2003/0092106	A1	5/2003	Baker et al.
2003/0092107	A1	5/2003	Baker et al.
2003/0092108	A1	5/2003	Baker et al.
2003/0092110	A1	5/2003	Baker et al.
2003/0092111	A1	5/2003	Baker et al.
2003/0092113	A1	5/2003	Baker et al.
2003/0092115	A1	5/2003	Baker et al.
2003/0092147	A1	5/2003	Baker et al.
2003/0096386	A1	5/2003	Baker et al.
2003/0100087	A1	5/2003	Baker et al.
2003/0100497	A1	5/2003	Baker et al.
2003/0180837	A1	9/2003	Eaton et al.
2006/0051779	A1	3/2006	Eaton et al.
2009/0068098	A1	3/2009	Kanner et al.

FOREIGN PATENT DOCUMENTS

JP	2003527074		9/2003
JP	2004 187620		7/2004
WO	WO-00/06698		2/2000
WO	WO-00/39296		7/2000
WO	WO00/56889	*	9/2000
WO	WO-00/56889		9/2000
WO	WO-01/12811		2/2001
WO	WO-01/16318		3/2001
WO	WO-01/40466		6/2001
WO	WO-01/55320		8/2001
WO	WO-01/57272		8/2001
WO	WO-01/57273		8/2001
WO	WO-01/57274		8/2001
WO	WO-01/77291		10/2001
WO	WO-02/00690		1/2002
WO	WO-02/08284		1/2002
WO	WO-02/10449		2/2002
WO	WO-02/70539		9/2002
WO	WO 2006/022722		3/2006
WO	WO 2009/033094		3/2009

OTHER PUBLICATIONS

Zwhalen et al (International Journal of Cancer, 2000, vol. 88, pp. 66-70).*

Nagase et al (DNA Research, 1995, vol. 2, pp. 37-43).*

Cameron, Mol. Biol. (1997) 7:253-265.

Hammer et al., Cell (1990) 63:1099-1112.

Houdebine, J. Biotech. (1994) 34:269-287.

Kappell et al., Current Opinions in Biotechnology (1992) 3:548-553.

Mullins et al., EMBO J. (1989) 8:4065-4072.

Mullins et al., Hypotension (1993) 22:630-633.

Mullins et al., J. Clin. Invest. (1996) 98:S37-S40.

Mullins et al., Nature (1990) 344:541-544.

Niemann, Transg. Res. (1997) 7:73-75.

Overbeek in: Transgenic Animal Technology, A Laboratory Handbook, Pinkert (ed.), Academic Press, Inc. (1994) pp. 96-98.

Taurog et al., J. Immunol. (1988) 141:4020-4023.

Wall, Theriogenology (1996) 45:57-68.

International Search Report for PCT/US04/26829, mailed on Jul. 14, 2005, 3 pages.

Bowie et al., Science (1990) 247:1306-1310.

Burgess et al., J. Cell Biol. (1990) 111:2129-2138.

Lazar et al., Mol. Cell. Biol. (1988) 8(3):1247-1252.

Ngo et al., in: "The Protein Folding Problem and Tediary Structure Prediction," Merz et al. (eds.), Birkhauser, Boston, MA (1994) pp. 433 and 492-495.
Skolnick and Fetrow, Tibtech (2000) 18:34-39.
Borowsky et al., The Journal of Cell Biology (1998) 143:429-442.
Ellerhorst et al., Proc. Amer. Assoc. Cancer Res. (1993) 34:95.
Gokudan et al., PNAS USA (1999) 96:10086-10091.
Idikio, Int. J. Of Oncology (1998) 12:1287-1290.
Kierszenbaum et al., Molecular Biology of the Cell (1999) 10 (Suppl):217a.
Kierszenbaum et al., Prostate (2000) 43:175-183.
Kobayashi et al., Biochimica et Biophysica Act (1989) 1009:244-250.
Matsui et al., Eur. J. Biochem. (1994) 219:449-454.
Nordsiek et al., EMBL No. H5H2137, 1999, 60 pgs.
Schlom, Molecular Foundations of Oncology (1991) pp. 95-134.
Shilayama et al., Med. Journal Kinki Univ. (1992)17:265-272.
Spertini et al., Leukemia (1991) 5(4):300-308.
Takahashi et al., The Journal of Biological Chemistry (1985) 260(22):12228-12233.
Tedder, Am. J. Respir. Cell Mol. Biol. (1991) 5:305-306.
U.S. Appl. No. 60/125,774 filed by Watanabe et al. on Mar. 23, 1999.
U.S. Appl. No. 60/125,778 filed by Wood on Mar. 23, 1999.
U.S. Appl. No. 60/128,826 filed by Wood on Mar. 24, 1999.
U.S. Appl. No. 60/127,035 filed by Watanabe and Wood on Mar. 31, 1999.
U.S. Appl. No. 60/127,706 filed by Wood on Apr. 5, 1999.
U.S. Appl. No. 60/130,359 filed by Watanabe and Wood on Apr. 21,1999.
U.S. Appl. No. 60/131,270 filed by Watanabe et al. on Apr. 27, 1999.
U.S. Appl. No. 60/131,272 filed by Watanabe and Wood on Apr. 27, 1999.
U.S. Appl. No. 60/131,291 filed by Wood on Apr. 27, 1999.
U.S. Appl. No. 60/132,371 filed by Watanabe and Wood on May 4, 1999.
U.S. Appl. No. 60/132,379 filed by Watanabe and Wood on May 4, 1999.
U.S. Appl. No. 60/132,383 filed by Watanabe and Wood on May 4, 1999.
U.S. Appl. No. 60/135,750 filed by Watanabe and Wood on May 25, 1999.
U.S. Appl. No. 60/138,166 filed by Wood on Jun. 8, 1999.
U.S. Appl. No. 60/144,791 filed by Wood on Jul. 20, 1999.
U.S. Appl. No. 60/146,970 filed by Wood on Aug. 3, 1999.
Blok et al., Prostate (1995) 26(4):213-224.
Ellerhorst et al., Int. J. Oncol. (Feb. 1999) 14(2):217-224.
EMBL Database, Accession No. AL078474.
Hubert et al., PNAS USA (1999) 96(25):14523-14528.
Klein et al., Nature Medicine (1997) 3(4):402-408.
Nemeth et al., J. Mol. Endocrinol. (Feb. 1998) 20(1):151-156.
Reiter et al., PNAS (Feb. 1998) 95:1735-1740.
Su et al., PNAS USA (1996) 93:7252-7257.
Colozza et al., Annals Oncol. (2005) 16:1723-1739.
Coradini et al., Curr. Opin. Obst. Gyn. (2004) 16:49-55.
Krajewska et al., Prostate (2006) 66:801-810.
Non-Final Office Action for U.S. Appl. No. 11/546,459, mailed on May 5, 2008, 15 pages.
Tang et al., J. Clin. Oncol. (1999) 17(6):1710-1719.
Turner et al., J. Clin. Oncol. (2001) 19(4):992-1000.
European Search Report for EP 08163095.6, mailed Dec. 9, 2008, 9 pages.
Supplementary European Search Report for EP 04816811.6, mailed Jul. 14, 2008, 4 pages.
Weng et al., The Journal of Biological Chemistry (2003) 278(21):19164-19170.
International Search Report and Written Opinion for PCT/US2009/054530, mailed Nov. 5, 2009, 22 pages.

* cited by examiner

FIGURE 1: 58P1D12 SSH SEQUENCE OF 427 NUCLEOTIDES. (SEQ ID NO: 1).

```
  1 CAGAATTACA AAATTCTGGA ACCAAGTCAA TGAGTTATTA TACTTCCATG CCACTTTCTT
 61 TTCTGGTACT CTTTGAAATC CACAGTGTAG ACTGGTTTGG ACTAGTTTTT GTTCTTCCTT
121 TACTTTTATG CAGCATCTGG AAACAACAGG TTCCAAAAGC AACCAGTATC AGTAAGAGCA
181 GGGGTATTGT TGGTATAACA ACATAAATTA GATTGGGAAT TATACCTGCT TCAGTAACAA
241 CCACATTCTG ATGGGTGTCT CCTGGTTGAT TTGTAAGATA AGGCTTTTCT ACAGGGGCTG
301 TTGGATTAAT CTCTGGTTCA TACTTGCAAA TATAATTGTG CTTCATGTTA CACCTGTCAT
361 CATTCCACTG GTAAAGGTAG GGACCCCCAA GGCCAGGATT GGCAGTTGGT TGGTGATCAT
421 CACAACA
```

Figure 2A. The cDNA (SEQ ID NO:2) and amino acid sequence (SEQ ID NO:3) of 58P1D12 v.1. The start methionine is underlined. The open reading frame extends from nucleic acid 380-1201 including the stop codon.

```
   1 atccaggaccagggcgcaccggctcagcctctcacttgtcagaggccggggaagagaagc
  61 aaagcgcaacggtgtggtccaagccggggcttctgcttcgcctctaggacatacacggga
 121 ccccctaacttcagtcccccaaacgcgcaccctcgaagtcttgaactccagccccgcaca
 181 tccacgcgcggcacaggcgcggcaggcggcaggtcccggccgaaggcgatgcgcgcaggg
 241 ggtcgggcagctgggctcgggcggcgggagtagggcccggcagggaggcagggaggctgc
 301 agagtcagagtcgcgggctgcgccctgggcagaggccgccctcgctccacgcaacacctg
   1                     M  S  R  V  V  S  L  L  L  G  A  A  L  L
 361 ctgctgccaccgcgccgcgATGAGCCGCGTGGTCTCGCTGCTGCTGGGCGCCGCGCTGCT
  15  C  G  H  G  A  F  C  R  R  V  V  S  G  Q  K  V  C  F  A  D
 421 CTGCGGCCACGGAGCCTTCTGCCGCCGCGTGGTCAGCGGCCAAAAGGTGTGTTTTGCTGA
  35  F  K  H  P  C  Y  K  M  A  Y  F  H  E  L  S  S  R  V  S  F
 481 CTTCAAGCATCCCTGCTACAAAATGGCCTACTTCCATGAACTGTCCAGCCGAGTGAGCTT
  55  Q  E  A  R  L  A  C  E  S  E  G  G  V  L  L  S  L  E  N  E
 541 TCAGGAGGCACGCCTGGCTTGTGAGAGTGAGGGAGGAGTCCTCCTCAGCCTTGAGAATGA
  75  A  E  Q  K  L  I  E  S  M  L  Q  N  L  T  K  P  G  T  G  I
 601 AGCAGAACAGAAGTTAATAGAGAGCATGTTGCAAAACCTGACAAAACCCGGGACAGGGAT
  95  S  D  G  D  F  W  I  G  L  W  R  N  G  D  G  Q  T  S  G  A
 661 TTCTGATGGTGATTTCTGGATAGGGCTTTGGAGGAATGGAGATGGGCAAACATCTGGTGC
 115  C  P  D  L  Y  Q  W  S  D  G  S  N  S  Q  Y  R  N  W  Y  T
 721 CTGCCCAGATCTCTACCAGTGGTCTGATGGAAGCAATTCCCAGTACCGAAACTGGTACAC
 135  D  E  P  S  C  G  S  E  K  C  V  V  M  Y  H  Q  P  T  A  N
 781 AGATGAACCTTCCTGCGGAAGTGAAAAGTGTGTTGTGATGTATCACCAACCAACTGCCAA
 155  P  G  L  G  G  P  Y  L  Y  Q  W  N  D  D  R  C  N  M  K  H
 841 TCCTGGCCTTGGGGGTCCCTACCTTTACCAGTGGAATGATGACAGGTGTAACATGAAGCA
 175  N  Y  I  C  K  Y  E  P  E  I  N  P  T  A  P  V  E  K  P  Y
 901 CAATTATATTTGCAAGTATGAACCAGAGATTAATCCAACAGCCCCTGTAGAAAAGCCTTA
 195  L  T  N  Q  P  G  D  T  H  Q  N  V  V  V  T  E  A  G  I  I
 961 TCTTACAAATCAACCAGGAGACACCCATCAGAATGTGGTTGTTACTGAAGCAGGTATAAT
 215  P  N  L  I  Y  V  V  I  P  T  I  P  L  L  L  L  I  L  V  A
1021 TCCCAATCTAATTTATGTTGTTATACCAACAATACCCCTGCTCTTACTGATACTGGTTGC
 235  F  G  T  C  C  F  Q  M  L  H  K  S  K  G  R  T  K  T  S  P
1081 TTTTGGAACCTGTTGTTTCCAGATGCTGCATAAAAGTAAAGGAAGAACAAAAACTAGTCC
 255  N  Q  S  T  L  W  I  S  K  S  T  R  K  E  S  G  M  E  V  *
1141 AAACCAGTCTACACTGTGGATTTCAAAGAGTACCAGAAAAGAAAGTGGCATGGAAGTATA
1201 Ataactcattgacttggttccagaattttgtaattctggatctgtataaggaatggcatc
1261 agaacaatagcttggaatggcttgaaatcacaaaggatctgcaagatgaactgtaagctc
1321 ccccttgaggcaaatattaaagtaatttttatatgtctattatttcatttaaagaatatg
1381 ctgtgctaataatggagtgagacatgcttattttgctaaaggatgcacccaaacttcaaa
```

Figure 2A (continued)

```
1441 cttcaagcaaatgaaatggacaatgcagataaagttgttatcaacacgtcgggagtatgt
1501 gtgttagaagcaattccttttatttctttcacctttcataagttgttatctagtcaatgt
1561 aatgtatattgtattgaaatttacagtgtgcaaaagtattttacctttgcataagtgttt
1621 gataaaaatgaactgttctaatatttatttttatggcatctcatttttcaatacatgctc
1681 ttttgattaaagaaacttattactgttgtcaactgaattcacacacacacaaatatagta
1741 ccatagaaaaagtttgttttctcgaaataattcatctttcagcttctctgcttttggtca
1801 atgtctaggaaatctcttcagaaataagaagctatttcattaagtgtgatataaacctcc
1861 tcaaacattttacttagaggcaaggattgtctaatttcaattgtgcaagacatgtgcctt
1921 ataattattttagcttaaaattaaacagattttgtaataatgtaactttgttaataggt
1981 gcataaacactaatgcagtcaatttgaacaaaagaagtgacatacacaatataaatcata
2041 tgtcttcacacgttgcctatataatgagaagcagctctctgagggttctgaaatcaatgt
2101 ggtccctctcttgcccactaaacaaagatggttgttcggggtttgggattgacactggag
2161 gcagatagttgcaaagttagtctaaggtttccctagctgtatttagcctctgactatatt
2221 agtatacaaagaggtcatgtggttgagaccaggtgaatagtcactatcagtgtggagaca
2281 agcacagcacacagacattttaggaaggaaaggaactacgaaatcgtgtgaaaatgggtt
2341 ggaacccatcagtgatcgcatattcattgatgagggtttgcttgagatagaaaatggtgg
2401 ctcctttctgtcttatctcctagtttcttcaatgcttacgccttgttcttctcaagagaa
2461 agttgtaactctctggtcttcatatgtccctgtgctccttttaaccaaataaagagttct
2521 tgtttctgaagaaaaaaaaaaaaaaaaaaa
```

Figure 2B. The cDNA (SEQ ID NO:4) and amino acid sequence (SEQ ID NO:5) of 58P1D12 v.2.

The start methionine is underlined. The open reading frame extends from nucleic acid 388-1086 including the stop codon.

```
  1 ctgtggtgtttttcccccgctcctctggctgccttcctgatggatctctgtggtcccagg
 61 caggaatggcctgcttggggacccagcgagctcccaaggcctttcctgctgcttcctcta
121 tccctgtgttttgcttggctctctaaattgactcagctccaggacatcaggaccccaggt
181 tctctggtcttgggactctgagacttgcaccaggaatcctgcccaggctctcaggccttt
241 ggactcagactgagctacttcactggctttcctggttctccagcttgaagatggcagatc
301 gtgggacttctcagcctccataattgagtgagccaattccctggccaaaggtgtgtttt
  1                            M  A  Y  F  H  E  L  S  S  R  V
361 gctgacttcaagcatccctgctacaaaATGGCCTACTTCCATGAACTGTCCAGCCGAGTG
 12 S  F  Q  E  A  R  L  A  C  E  S  E  G  G  V  L  L  S  L  E
421 AGCTTTCAGGAGGCACGCCTGGCTTGTGAGAGTGAGGGAGGAGTCCTCCTCAGCCTTGAG
 32 N  E  A  E  Q  K  L  I  E  S  M  L  Q  N  L  T  K  P  G  T
481 AATGAAGCAGAACAGAAGTTAATAGAGAGCATGTTGCAAAACCTGACAAAACCCGGGACA
 52 G  I  S  D  G  D  F  W  I  G  L  W  R  N  G  D  G  Q  T  S
541 GGGATTTCTGATGGTGATTTCTGGATAGGGCTTTGGAGGAATGGAGATGGGCAAACATCT
```

Figure 2B (continued)

```
  72 G   A   C   P   D   L   Y   Q   W   S   D   G   S   N   S   Q   Y   R   N   W
 601 GGTGCCTGCCCAGATCTCTACCAGTGGTCTGATGGAAGCAATTCCCAGTACCGAAACTGG
  92 Y   T   D   E   P   S   C   G   S   E   K   C   V   V   M   Y   H   Q   P   T
 661 TACACAGATGAACCTTCCTGCGGAAGTGAAAAGTGTGTTGTGATGTATCACCAACCAACT
 112 A   N   P   G   L   G   G   P   Y   L   Y   Q   W   N   D   D   R   C   N   M
 721 GCCAATCCTGGCCTTGGGGGTCCCTACCTTTACCAGTGGAATGATGACAGGTGTAACATG
 132 K   H   N   Y   I   C   K   Y   E   P   E   I   N   P   T   A   P   V   E   K
 781 AAGCACAATTATATTTGCAAGTATGAACCAGAGATTAATCCAACAGCCCCTGTAGAAAAG
 152 P   Y   L   T   N   Q   P   G   D   T   H   Q   N   V   V   V   T   E   A   G
 841 CCTTATCTTACAAATCAACCAGGAGACACCCATCAGAATGTGGTTGTTACTGAAGCAGGT
 172 I   I   P   N   L   I   Y   V   V   I   P   T   I   P   L   L   L   L   I   L
 901 ATAATTCCCAATCTAATTTATGTTGTTATACCAACAATACCCCTGCTCTTACTGATACTG
 192 V   A   F   G   T   C   C   F   Q   M   L   H   K   S   K   G   R   T   K   T
 961 GTTGCTTTTGGAACCTGTTGTTTCCAGATGCTGCATAAAAGTAAAGGAAGAACAAAAACT
 212 S   P   N   Q   S   T   L   W   I   S   K   S   T   R   K   E   S   G   M   E
1021 AGTCCAAACCAGTCTACACTGTGGATTTCAAAGAGTACCAGAAAAGAAAGTGGCATGGAA
 232 V   *
1081 GTATAAtaactcattgacttggttccagaattttgtaattctggatctgtataaggaatg
1141 gcatcagaacaatagcttggaatggcttgaaatcacaaaggatctgcaagatgaactgta
1201 agctccccttgaggcaaatattaaagtaatttttatatgtctattatttcatttaaaga
1261 atatgctgtgctaataatggagtgagacatgcttattttgctaaaggatgcacccaaact
1321 tcaaacttcaagcaaatgaaatggacaatgcagataaagttgttatcaacacgtcgggag
1381 tatgtgtgttagaagcaattcctttttatttctttcacctttcataagttgttatctagtc
1441 aatgtaatgtatattgtattgaaatttacagtgtgcaaaagtattttacctttgcataag
1501 tgtttgataaaaatgaactgttctaatatttatttttatggcatctcatttttcaataca
1561 tgctcttttgattaaagaaacttattactgttgtcaactgaattcacacacacacaaata
1621 tagtaccatagaaaaagtttgttttctcgaaataattcatctttcagcttctctgctttt
1681 ggtcaatgtctaggaaatctcttcagaaataagaagctatttcattaagtgtgatataaa
1741 cctcctcaaacattttacttagaggcaaggattgtctaatttcaattgtgcaagacatgt
1801 gccttataattatttttagcttaaaattaaacagattttgtaataatgtaactttgttaa
1861 taggtgcataaacactaatgcagtcaatttgaacaaaagaagtgacatacacaatataaa
1921 tcatatgtcttcacacgttgcctatataatgagaagcagctctctgagggttctgaaatc
1981 aatgtggtccctctcttgcccactaaacaaagatggttgttcggggtttgggattgacac
2041 tggaggcagatagttgcaaagttagtctaaggtttccctagctgtatttagcctctgact
2101 atattagtatacaaagaggtcatgtggttgagaccaggtgaatagtcactatcagtgtgg
2161 agacaagcacagcacacagacatttaggaaggaaaggaactacgaaatcgtgtgaaaat
2221 gggttggaacccatcagtgatcgcatattcattgatgagggtttgcttgagatagaaaat
2281 ggtggctcctttctgtcttatctcctagtttcttcaatgcttacgccttgttcttctcaa
2341 gagaaagttgtaactctctggtcttcatatgtccctgtgctcctttaaccaaataaaga
2401 gttcttgtttctgaagaa
```

Figure 2C. The cDNA (SEQ ID NO:6) and amino acid sequence (SEQ ID NO:7) of 58P1D12 v.3. The start methionine is underlined. The open reading frame extends from nucleic acid 206-904 including the stop codon.

```
   1 gcttgaagatggcagatcgtgggacttctcagcctccataattgagtgagccaattccct
  61 gaatacaacaagaagatggccatctgtagaccaggaggtggtccctccccagaaactgga
 121 tgggccagcacgttgattctgagcttctagcctccagaactgccaaaaggtgtgttttgc
   1                           M  A  Y  F  H  E  L  S  S  R  V  S
 181 tgacttcaagcatccctgctacaaaATGGCCTACTTCCATGAACTGTCCAGCCGAGTGAG
  13   F  Q  E  A  R  L  A  C  E  S  E  G  G  V  L  L  S  L  E  N
 241 CTTTCAGGAGGCACGCCTGGCTTGTGAGAGTGAGGGAGGAGTCCTCCTCAGCCTTGAGAA
  33   E  A  E  Q  K  L  I  E  S  M  L  Q  N  L  T  K  P  G  T  G
 301 TGAAGCAGAACAGAAGTTAATAGAGAGCATGTTGCAAAACCTGACAAAACCCGGGACAGG
  53   I  S  D  G  D  F  W  I  G  L  W  R  N  G  D  G  Q  T  S  G
 361 GATTTCTGATGGTGATTTCTGGATAGGGCTTTGGAGGAATGGAGATGGGCAAACATCTGG
  73   A  C  P  D  L  Y  Q  W  S  D  G  S  N  S  Q  Y  R  N  W  Y
 421 TGCCTGCCCAGATCTCTACCAGTGGTCTGATGGAAGCAATTCCCAGTACCGAAACTGGTA
  93   T  D  E  P  S  C  G  S  E  K  C  V  V  M  Y  H  Q  P  T  A
 481 CACAGATGAACCTTCCTGCGGAAGTGAAAAGTGTGTTGTGATGTATCACCAACCAACTGC
 113   N  P  G  L  G  G  P  Y  L  Y  Q  W  N  D  D  R  C  N  M  K
 541 CAATCCTGGCCTTGGGGGTCCCTACCTTTACCAGTGGAATGATGACAGGTGTAACATGAA
 133   H  N  Y  I  C  K  Y  E  P  E  I  N  P  T  A  P  V  E  K  P
 601 GCACAATTATATTTGCAAGTATGAACCAGAGATTAATCCAACAGCCCCTGTAGAAAAGCC
 153   Y  L  T  N  Q  P  G  D  T  H  Q  N  V  V  V  T  E  A  G  I
 661 TTATCTTACAAATCAACCAGGAGACACCCATCAGAATGTGGTTGTTACTGAAGCAGGTAT
 173   I  P  N  L  I  Y  V  V  I  P  T  I  P  L  L  L  L  I  L  V
 721 AATTCCCAATCTAATTTATGTTGTTATACCAACAATACCCCTGCTCTTACTGATACTGGT
 193   A  F  G  T  C  C  F  Q  M  L  H  K  S  K  G  R  T  K  T  S
 781 TGCTTTTGGAACCTGTTGTTTCCAGATGCTGCATAAAAGTAAAGGAAGAACAAAAACTAG
 213   P  N  Q  S  T  L  W  I  S  K  S  T  R  K  E  S  G  M  E  V
 841 TCCAAACCAGTCTACACTGTGGATTTCAAAGAGTACCAGAAAAGAAAGTGGCATGGAAGT
 233   *
 901 ATAAtaactcattgacttggttccagaattttgtaattctggatctgtataaggaatggc
 961 atcagaacaatagcttggaatggcttgaaatcacaaaggatctgcaagatgaactgtaag
1021 ctccccctt gaggcaaatattaaagtaattttatatgtctattatttcatttaaagaat
1081 atgctgtgctaataatggagtgagacatgcttattttgctaaaggatgcacccaaacttc
1141 aaacttcaagcaaatgaaatggacaatgcagataaagttgttatcaacacgtcgggagta
1201 tgtgtgttagaagcaattcctttatttctttcacctttcataagttgttatctagtcaa
1261 tgtaatgtatattgtattgaaatttacagtgtgcaaaagtattttacctttgcataagtg
1321 tttgataaaaatgaactgttctaatatttatttttatggcatctcatttttcaatacatg
1381 ctcttttgattaaagaaacttattactgttgtcaactgaattcacacacacacaaatata
1441 gtaccatagaaaaagtttgttttctcgaaataattcatctttcagcttctctgcttttgg
```

Figure 2C (continued)

```
1501 tcaatgtctaggaaatctcttcagaaataagaagctatttcattaagtgtgatataaacc
1561 tcctcaaacattttacttagaggcaaggattgtctaatttcaattgtgcaagacatgtgc
1621 cttataattatttttagcttaaaattaaacagattttgtaataatgtaactttgttaata
1681 ggtgcataaacactaatgcagtcaatttgaacaaaagaagtgacatacacaatataaatc
1741 atatgtcttcacacgttgcctatataatgagaagcagctctctgagggttctgaaatcaa
1801 tgtggtccctctcttgcccactaaacaaagatggttgttcggggtttgggattgacactg
1861 gaggcagatagttgcaaagttagtctaaggtttccctagctgtatttagcctctgactat
1921 attagtatacaaagaggtcatgtggttgagaccaggtgaatagtcactatcagtgtggag
1981 acaagcacagcacacagacattttaggaaggaaaggaactacgaaatcgtgtgaaaatgg
2041 gttggaacccatcagtgatcgcatattcattgatgagggtttgcttgagatagaaaatgg
2101 tggctcctttctgtcttatctcctagtttcttcaatgcttacgccttgttcttctcaaga
2161 gaaagttgtaactctctggtcttcatatgtccctgtgctccttttaaccaaataaagagt
2221 tcttgtttctgaagaa
```

Figure 2D. The cDNA (SEQ ID NO:8) and amino acid sequence (SEQ ID NO:9) of 58P1D12 v.4.
The start methionine is underlined. The open reading frame extends from nucleic acid 206-916 including the stop codon.

```
   1 gcttgaagatggcagatcgtgggacttctcagcctccataattgagtgagccaattccct
  61 gaatacaacaagaagatggccatctgtagaccaggaggtggtccctccccagaaactgga
 121 tgggccagcacgttgattctgagcttctagcctccagaactgccaaaaggtgtgttttgc
   1                           M  A  Y  F  H  E  L  S  S  R  V  S
 181 tgacttcaagcatccctgctacaaaATGGCCTACTTCCATGAACTGTCCAGCCGAGTGAG
  13  F  Q  E  A  R  L  A  C  E  S  E  G  G  V  L  L  S  L  E  N
 241 CTTTCAGGAGGCACGCCTGGCTTGTGAGAGTGAGGGAGGAGTCCTCCTCAGCCTTGAGAA
  33  E  A  E  Q  K  L  I  E  S  M  L  Q  N  L  T  K  P  G  T  G
 301 TGAAGCAGAACAGAAGTTAATAGAGAGCATGTTGCAAAACCTGACAAAACCCGGGACAGG
  53  I  S  D  G  D  F  W  I  G  L  W  R  N  G  D  G  Q  T  S  G
 361 GATTTCTGATGGTGATTTCTGGATAGGGCTTTGGAGGAATGGAGATGGGCAAACATCTGG
  73  A  C  P  D  L  Y  Q  W  S  D  G  S  N  S  Q  Y  R  N  W  Y
 421 TGCCTGCCCAGATCTCTACCAGTGGTCTGATGGAAGCAATTCCCAGTACCGAAACTGGTA
  93  T  D  E  P  S  C  G  S  E  K  C  V  V  M  Y  H  Q  P  T  A
 481 CACAGATGAACCTTCCTGCGGAAGTGAAAAGTGTGTTGTGATGTATCACCAACCAACTGC
 113  N  P  G  L  G  G  P  Y  L  Y  Q  W  N  D  D  R  C  N  M  K
 541 CAATCCTGGCCTTGGGGGTCCCTACCTTTACCAGTGGAATGATGACAGGTGTAACATGAA
 133  H  N  Y  I  C  K  Y  E  P  E  I  N  P  T  A  P  V  E  K  P
 601 GCACAATTATATTTGCAAGTATGAACCAGAGATTAATCCAACAGCCCCTGTAGAAAAGCC
 153  Y  L  T  N  Q  P  G  D  T  H  Q  N  V  V  V  T  E  A  V  K
 661 TTATCTTACAAATCAACCAGGAGACACCCATCAGAATGTGGTTGTTACTGAAGCAGTAAA
 173  E  E  Q  K  L  V  Q  T  S  L  H  C  G  F  Q  R  V  P  E  K
 721 GGAAGAACAAAAACTAGTCCAAACCAGTCTACACTGTGGATTTCAAAGAGTACCAGAAAA
 193  K  V  A  W  K  Y  N  N  S  L  T  W  F  Q  N  F  V  I  L  D
 781 GAAAGTGGCATGGAAGTATAATAACTCATTGACTTGGTTCCAGAATTTTGTAATTCTGGA
 213  L  Y  K  E  W  H  Q  N  N  S  L  E  W  L  E  I  T  K  D  L
 841 TCTGTATAAGGAATGGCATCAGAACAATAGCTTGGAATGGCTTGAAATCACAAAGGATCT
 233  Q  D  E  L  *
 901 GCAAGATGAACTGTAAgctccccccttgaggcaaatattaaagtaatttttatatgtctat
 961 tatttcatttaaagaatatgctgtgctaataatggagtgagacatgcttattttgctaaa
1021 ggatgcacccaaacttcaaacttcaagcaaatgaaatggacaatgcagataaagttgtta
1081 tcaacacgtcgggagtatgtgtgttagaagcaattcctttttatttctttcacctttcata
1141 agttgttatctagtcaatgtaatgtatattgtattgaaatttacagtgtgcaaaagtatt
1201 ttacctttgcataagtgtttgataaaaatgaactgttctaatatttatttttatggcatc
1261 tcatttttcaatacatgctcttttgattaaagaaacttattactgttgtcaactgaattc
1321 acacacacacaaatatagtaccatagaaaaagtttgttttctcgaaataattcatctttc
1381 agcttctctgcttttggtcaatgtctaggaaatctcttcagaaataagaagctatttcat
1441 taagtgtgatataaacctcctcaaacattttacttagaggcaaggattgtctaatttcaa
```

Figure 2D (continued)

```
1501 ttgtgcaagacatgtgccttataattatttttagcttaaaattaaacagattttgtaata
1561 atgtaactttgttaataggtgcataaacactaatgcagtcaatttgaacaaaagaagtga
1621 catacacaatataaatcatatgtcttcacacgttgcctatataatgagaagcagctctct
1681 gagggttctgaaatcaatgtggtccctctcttgcccactaaacaaagatggttgttcggg
1741 gtttgggattgacactggaggcagatagttgcaaagttagtctaaggtttccctagctgt
1801 atttagcctctgactatattagtatacaaagaggtcatgtggttgagaccaggtgaatag
1861 tcactatcagtgtggagacaagcacagcacacagacattttaggaaggaaaggaactacg
1921 aaatcgtgtgaaaatgggttggaacccatcagtgatcgcatattcattgatgagggtttg
1981 cttgagatagaaaatggtggctcctttctgtcttatctcctagtttcttcaatgcttacg
2041 ccttgttcttctcaagagaaagttgtaactctctggtcttcatatgtccctgtgctcctt
2101 ttaaccaaataaagagttcttgtttctgaagaa
```

Figure 2E. The cDNA (SEQ ID NO:10) and amino acid sequence (SEQ ID NO:11) of 58P1D12 v.5. The start methionine is underlined. The open reading frame extends from nucleic acid 106-816 including the stop codon.

```
   1 gcttgaagatggcagatcgtgggacttctcagcctccataattgagtgagccaattccct
   1                                              M  A  Y  F  H 
  61 ggccaaaaggtgtgttttgctgacttcaagcatccctgctacaaaATGGCCTACTTCCAT
   6 E  L  S  S  R  V  S  F  Q  E  A  R  L  A  C  E  S  E  G  G 
 121 GAACTGTCCAGCCGAGTGAGCTTTCAGGAGGCACGCCTGGCTTGTGAGAGTGAGGGAGGA
  26 V  L  L  S  L  E  N  E  A  E  Q  K  L  I  E  S  M  L  Q  N 
 181 GTCCTCCTCAGCCTTGAGAATGAAGCAGAACAGAAGTTAATAGAGAGCATGTTGCAAAAC
  46 L  T  K  P  G  T  G  I  S  D  G  D  F  W  I  G  L  W  R  N 
 241 CTGACAAAACCCGGGACAGGGATTTCTGATGGTGATTTCTGGATAGGGCTTTGGAGGAAT
  66 G  D  G  Q  T  S  G  A  C  P  D  L  Y  Q  W  S  D  G  S  N 
 301 GGAGATGGGCAAACATCTGGTGCCTGCCCAGATCTCTACCAGTGGTCTGATGGAAGCAAT
  86 S  Q  Y  R  N  W  Y  T  D  E  P  S  C  G  S  E  K  C  V  V 
 361 TCCCAGTACCGAAACTGGTACACAGATGAACCTTCCTGCGGAAGTGAAAAGTGTGTTGTG
 106 M  Y  H  Q  P  T  A  N  P  G  L  G  G  P  Y  L  Y  Q  W  N 
 421 ATGTATCACCAACCAACTGCCAATCCTGGCCTTGGGGGTCCCTACCTTTACCAGTGGAAT
 126 D  D  R  C  N  M  K  H  N  Y  I  C  K  Y  E  P  E  I  N  P 
 481 GATGACAGGTGTAACATGAAGCACAATTATATTTGCAAGTATGAACCAGAGATTAATCCA
 146 T  A  P  V  E  K  P  Y  L  T  N  Q  P  G  D  T  H  Q  N  V 
 541 ACAGCCCCTGTAGAAAAGCCTTATCTTACAAATCAACCAGGAGACACCCATCAGAATGTG
 166 V  V  T  E  A  V  K  E  E  Q  K  L  V  Q  T  S  L  H  C  G 
 601 GTTGTTACTGAAGCAGTAAAGGAAGAACAAAAACTAGTCCAAACCAGTCTACACTGTGGA
 186 F  Q  R  V  P  E  K  K  V  A  W  K  Y  N  N  S  L  T  W  F 
 661 TTTCAAAGAGTACCAGAAAAGAAAGTGGCATGGAAGTATAATAACTCATTGACTTGGTTC
 206 Q  N  F  V  I  L  D  L  Y  K  E  W  H  Q  N  N  S  L  E  W 
 721 CAGAATTTTGTAATTCTGGATCTGTATAAGGAATGGCATCAGAACAATAGCTTGGAATGG
```

Figure 2E (continued)

```
 226 L  E  I  T  K  D  L  Q  D  E  L  *
 781 CTTGAAATCACAAAGGATCTGCAAGATGAACTGTAAgctccccccttgaggcaaatattaa
 841 agtaattttatatgtctattatttcatttaaagaatatgctgtgctaataatggagtga
 901 gacatgcttattttgctaaaggatgcacccaaacttcaaacttcaagcaaatgaaatgga
 961 caatgcagataaagttgttatcaacacgtcgggagtatgtgtgttagaagcaattccttt
1021 tatttctttcacctttcataagttgttatctagtcaatgtaatgtatattgtattgaaat
1081 ttacagtgtgcaaaagtattttacctttgcataagtgtttgataaaaatgaactgttcta
1141 atatttatttttatggcatctcatttttcaatacatgctcttttgattaaagaaacttat
1201 tactgttgtcaactgaattcacacacacacaaatatagtaccatagaaaaagtttgtttt
1261 ctcgaaataattcatctttcagcttctctgcttttggtcaatgtctaggaaatctcttca
1321 gaaataagaagctatttcattaagtgtgatataaacctcctcaaacattttacttagagg
1381 caaggattgtctaatttcaattgtgcaagacatgtgccttataattatttttagcttaaa
1441 attaaacagattttgtaataatgtaactttgttaataggtgcataaacactaatgcagtc
1501 aatttgaacaaaagaagtgacatacacaatataaatcatatgtcttcacacgttgcctat
1561 ataatgagaagcagctctctgagggttctgaaatcaatgtggtccctctcttgcccacta
1621 aacaaagatggttgttcggggtttgggattgacactggaggcagatagttgcaaagttag
1681 tctaaggtttccctagctgtatttagcctctgactatattagtatacaaagaggtcatgt
1741 ggttgagaccaggtgaatagtcactatcagtgtggagacaagcacagcacacagacattt
1801 taggaaggaaaggaactacgaaatcgtgtgaaaatgggttggaacccatcagtgatcgca
1861 tattcattgatgagggtttgcttgagatagaaaatggtggctcctttctgtcttatctcc
1921 tagtttcttcaatgcttacgccttgttcttctcaagagaaagttgtaactctctggtctt
1981 catatgtccctgtgctccttttaaccaaataaagagttcttgtttctgaagaa
```

Figure 2F. 58P1D12 v.6 through v.15, SNP variants of 58P1D12 v.1.

Variants 58P1D12 v.6 through v.15 are variants with single nucleotide difference from 58P1D12 v.1.

Though these SNP variants are shown separately, they can also occur in any combinations and in any of the transcript variants listed above in Figures 2A-2E.

Variant	Nucleic acid position	Nucleic Acid Variation
58P1D12 v.6	1764	A/C
58P1D12 v.7	1987	G/A
58P1D12 v.8	2045	A/G
58P1D12 v.9	2066	T/C
58P1D12 v.10	2134	T/A
58P1D12 v.11	2350	G/T
58P1D12 v.12	2435	G/T
58P1D12 v.13	302	G/T
58P1D12 v.14	304	G/T
58P1D12 v.15	1533	C/T

Figure 3A. **Amino acid sequence 58P1D12 v.1 (SEQ ID. NO.: 12).** The 58P1D12 v.1 protein has 273 amino acids.

1 MSRVVSLLLG AALLCGHGAF CRRVVSGQKV CFADFKHPCY KMAYFHELSS RVSFQEARLA
61 CESEGGVLLS LENEAEQKLI ESMLQNLTKP GTGISDGDFW IGLWRNGDGQ TSGACPDLYQ
121 WSDGSNSQYR NWYTDEPSCG SEKCVVMYHQ PTANPGLGGP YLYQWNDDRC NMKHNYICKY
181 EPEINPTAPV EKPYLTNQPG DTHQNVVVTE AGIIPNLIYV VIPTIPLLLL ILVAFGTCCF
241 QMLHKSKGRT KTSPNQSTLW ISKSTRKESG MEV

Figure 3B. Amino acid sequence 58P1D12 v.2 (SEQ ID. NO.: 13). The 58P1D12 v.2 protein has 232 amino acids.

1 MAYFHELSSR VSFQEARLAC ESEGGVLLSL ENEAEQKLIE SMLQNLTKPG TGISDGDFWI
61 GLWRNGDGQT SGACPDLYQW SDGSNSQYRN WYTDEPSCGS EKCVVMYHQP TANPGLGGPY
121 LYQWNDDRCN MKHNYICKYE PEINPTAPVE KPYLTNQPGD THQNVVVTEA GIIPNLIYVV
181 IPTIPLLLLI LVAFGTCCFQ MLHKSKGRTK TSPNQSTLWI SKSTRKESGM EV

Figure 3C. Amino acid sequence 58P1D12 v.4 (SEQ ID. NO.: 14). The 58P1D12 v.4 protein has 236 amino acids.

1 MAYFHELSSR VSFQEARLAC ESEGGVLLSL ENEAEQKLIE SMLQNLTKPG TGISDGDFWI
61 GLWRNGDGQT SGACPDLYQW SDGSNSQYRN WYTDEPSCGS EKCVVMYHQP TANPGLGGPY
121 LYQWNDDRCN MKHNYICKYE PEINPTAPVE KPYLTNQPGD THQNVVVTEA VKEEQKLVQT
181 SLHCGFQRVP EKKVAWKYNN SLTWFQNFVI LDLYKEWHQN NSLEWLEITK DLQDEL

Figure 4: Intentionally Omitted.

Figure 5a: 58P1D12 variant 1 Hydrophilicity profile

(Hopp T.P., Woods K.R., 1981.
Proc. Natl. Acad. Sci. U.S.A. 78:3824-3828)

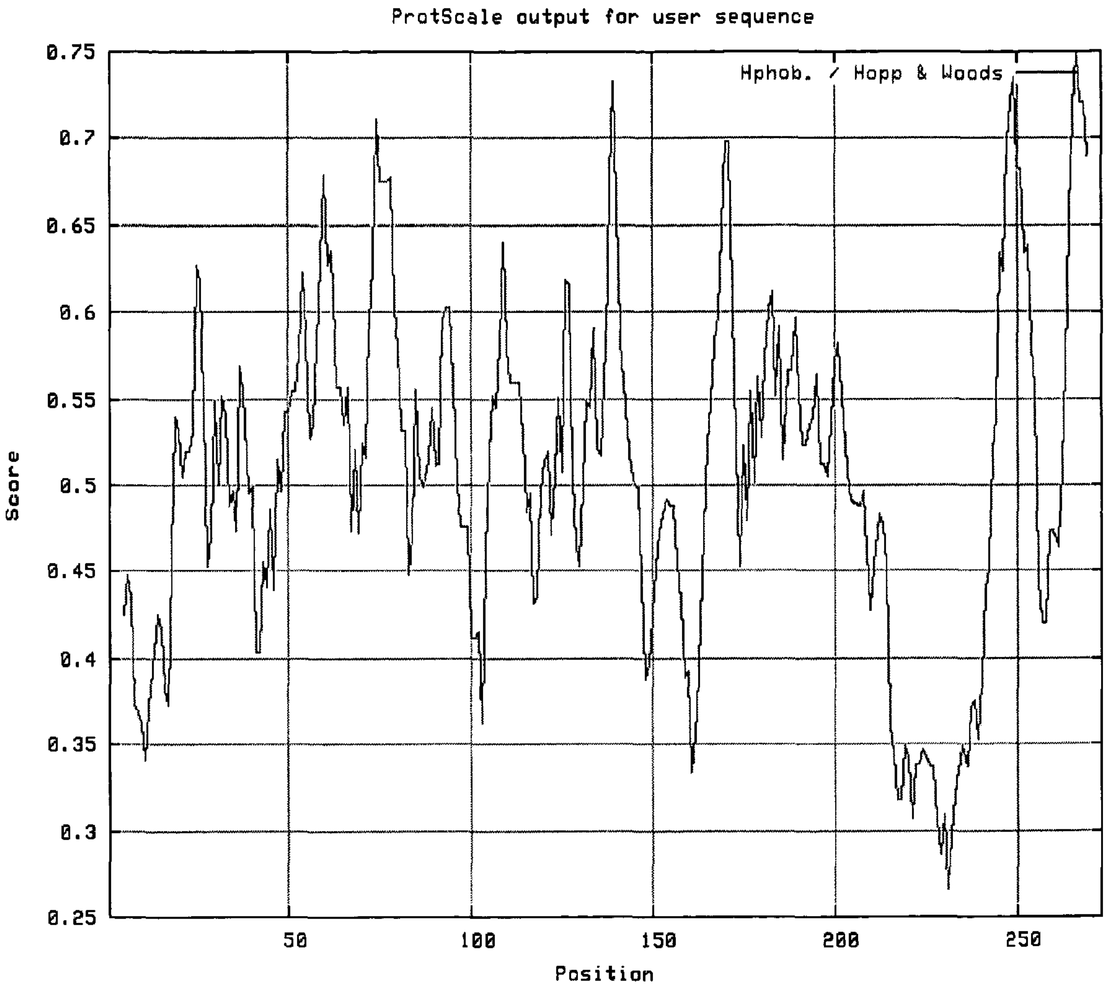


Figure 5b: 58P1D12 variant 2
Hydrophilicity profile
(Hopp T.P., Woods K.R., 1981.
Proc. Natl. Acad. Sci. U.S.A. 78:3824-3828)

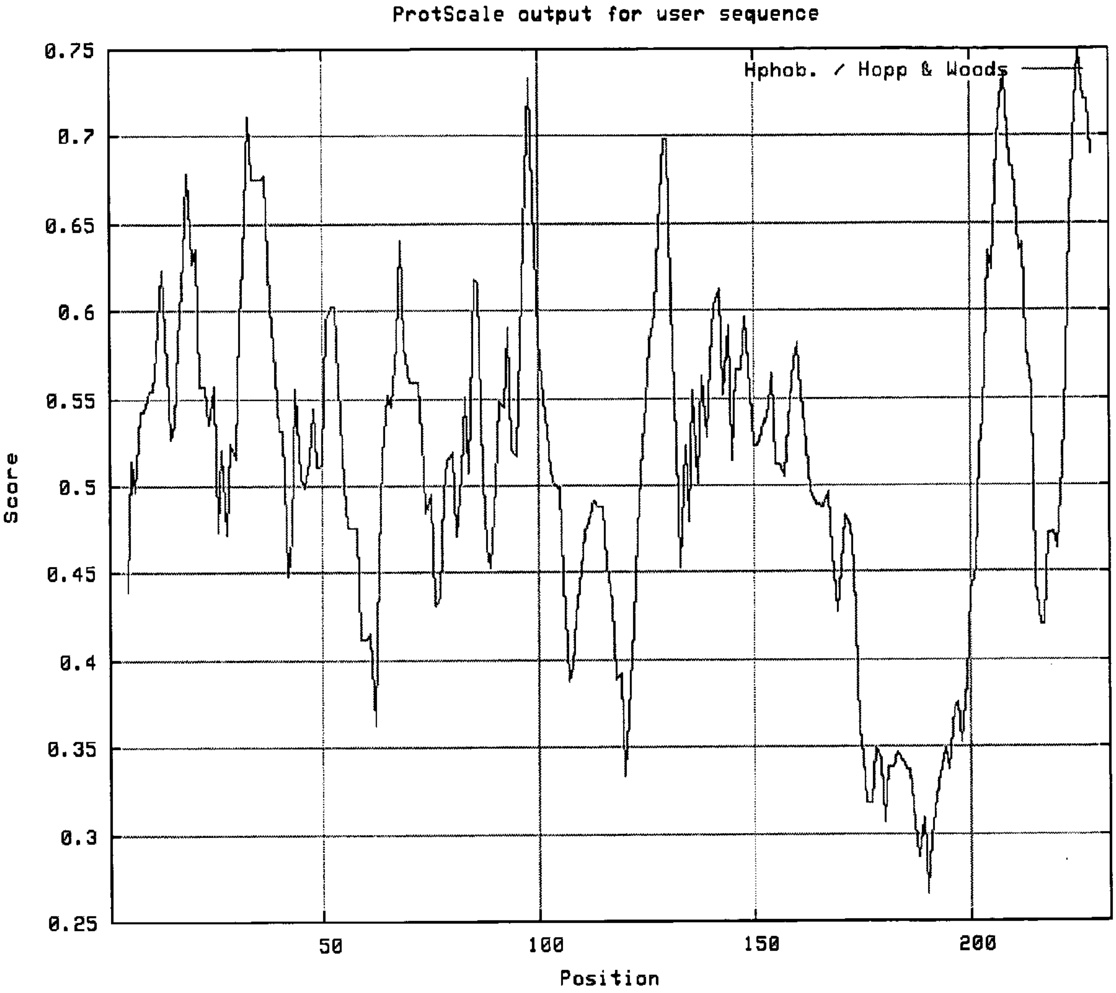


Figure 5c: 58P1D12 variant 3

Hydrophilicity profile

(Hopp T.P., Woods K.R., 1981.
Proc. Natl. Acad. Sci. U.S.A. 78:3824-3828)

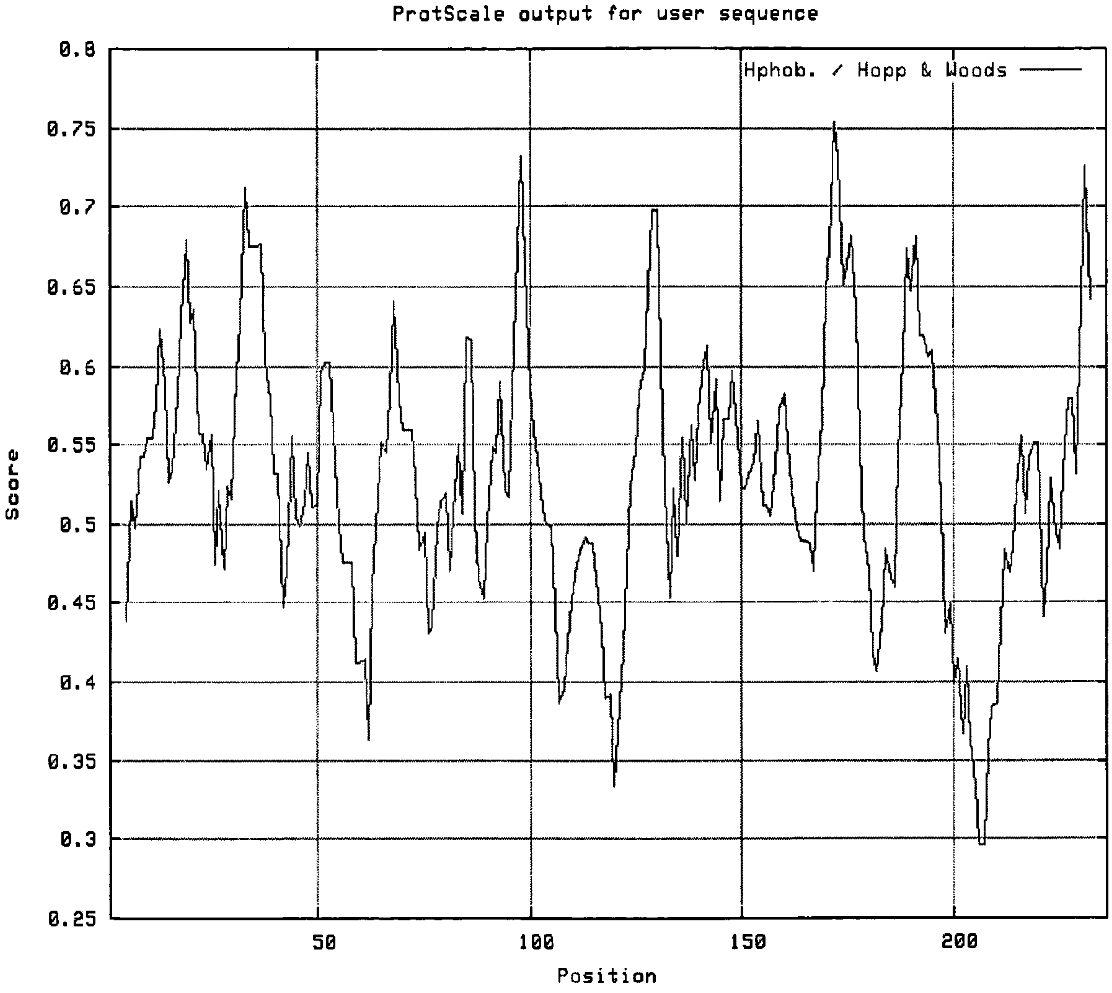


Figure 6a: 58P1D12 variant 1
Hydropathicity Profile
(Kyte J., Doolittle R.F., 1982. J. Mol. Biol. 157:105-132)

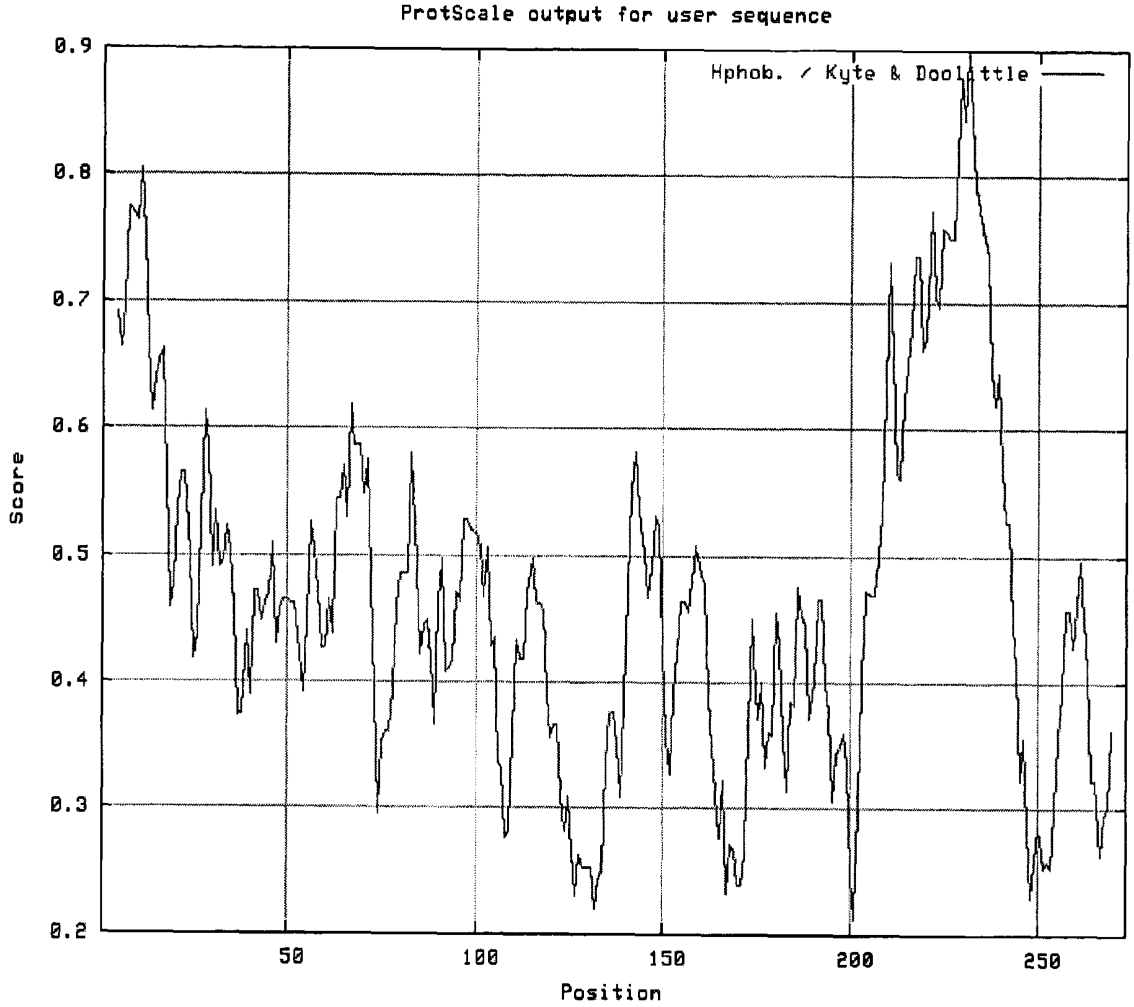


Figure 6b: 58P1D12 variant 2 Hydropathicity Profile

(Kyte J., Doolittle R.F., 1982. J. Mol. Biol. 157:105-132)

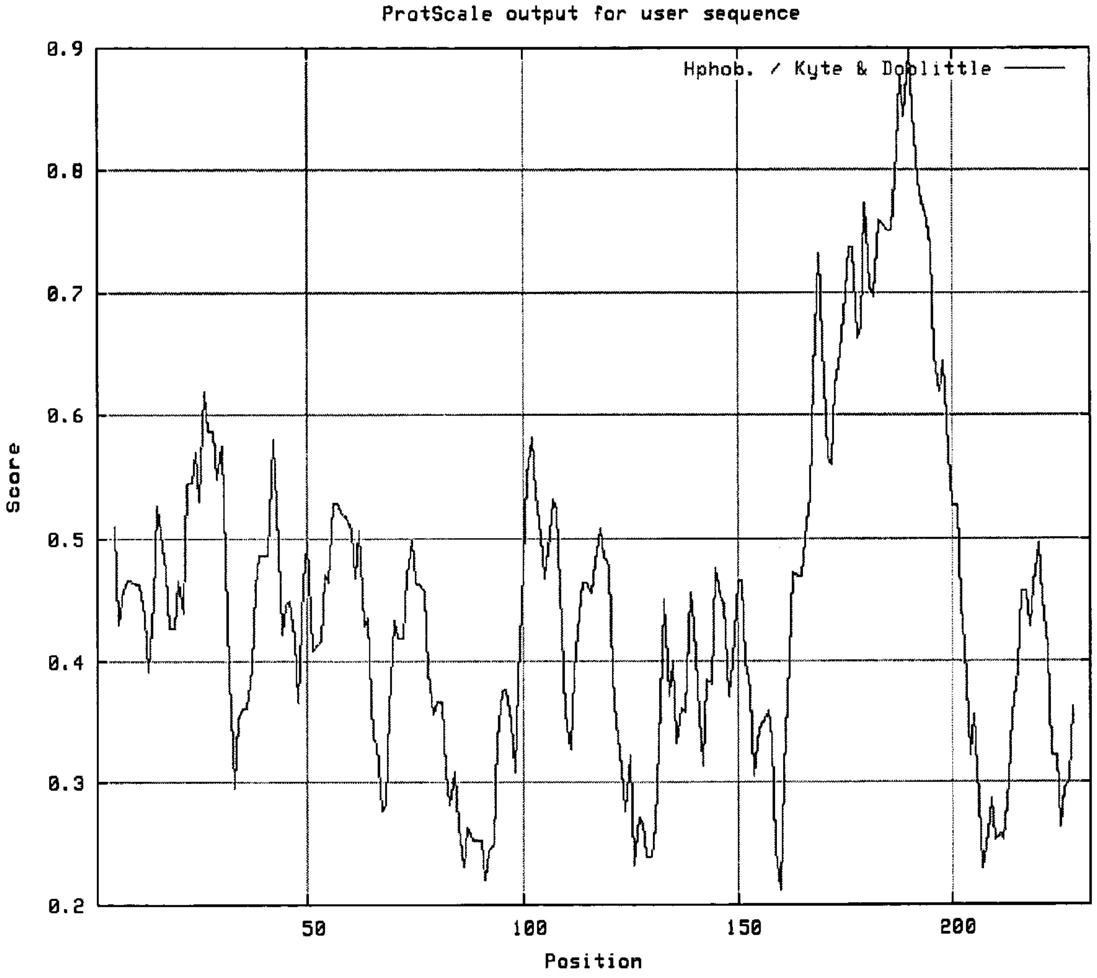


Figure 6c: 58P1D12 variant 3 Hydropathicity Profile

(Kyte J., Doolittle R.F., 1982. J. Mol. Biol. 157:105-132)

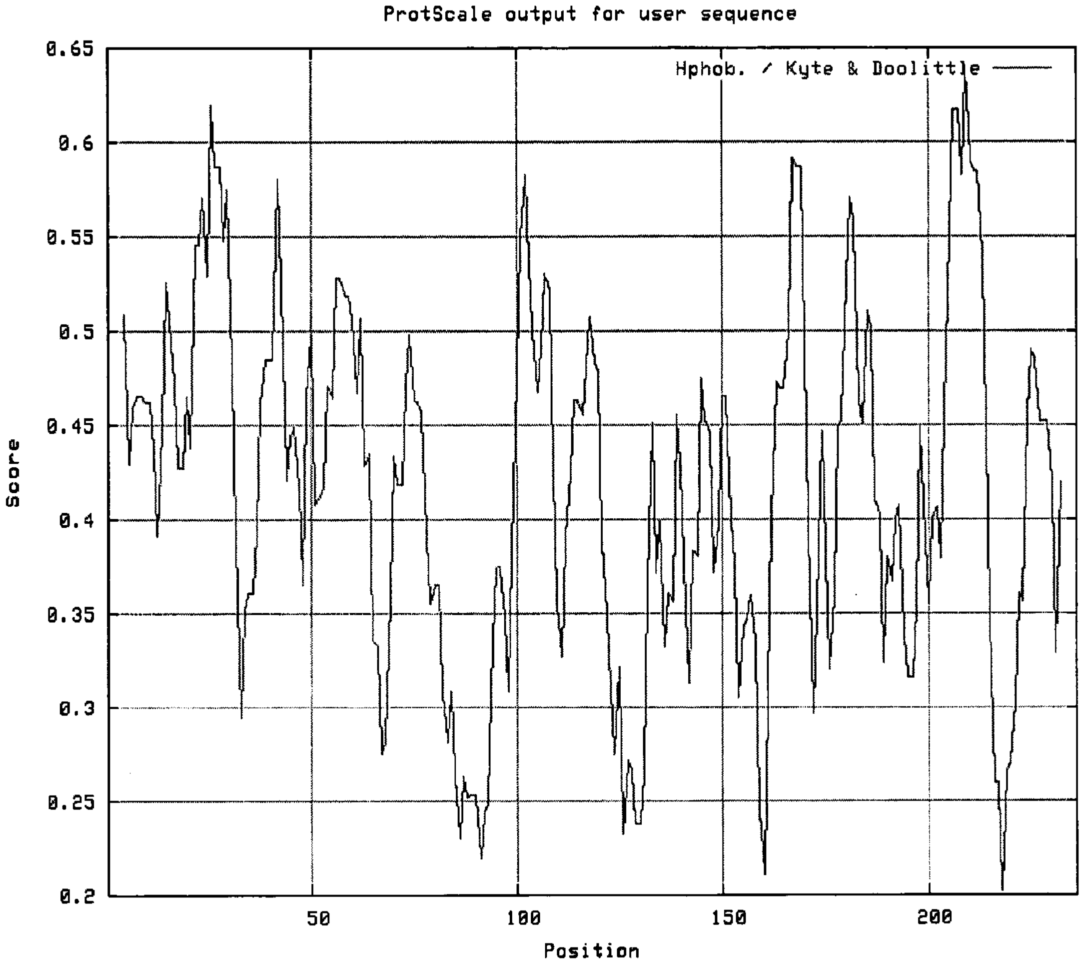


Figure 7a:58P1D12 variant 1
% Accessible Residues Profile
(Janin J., 1979. Nature 277:491-492)

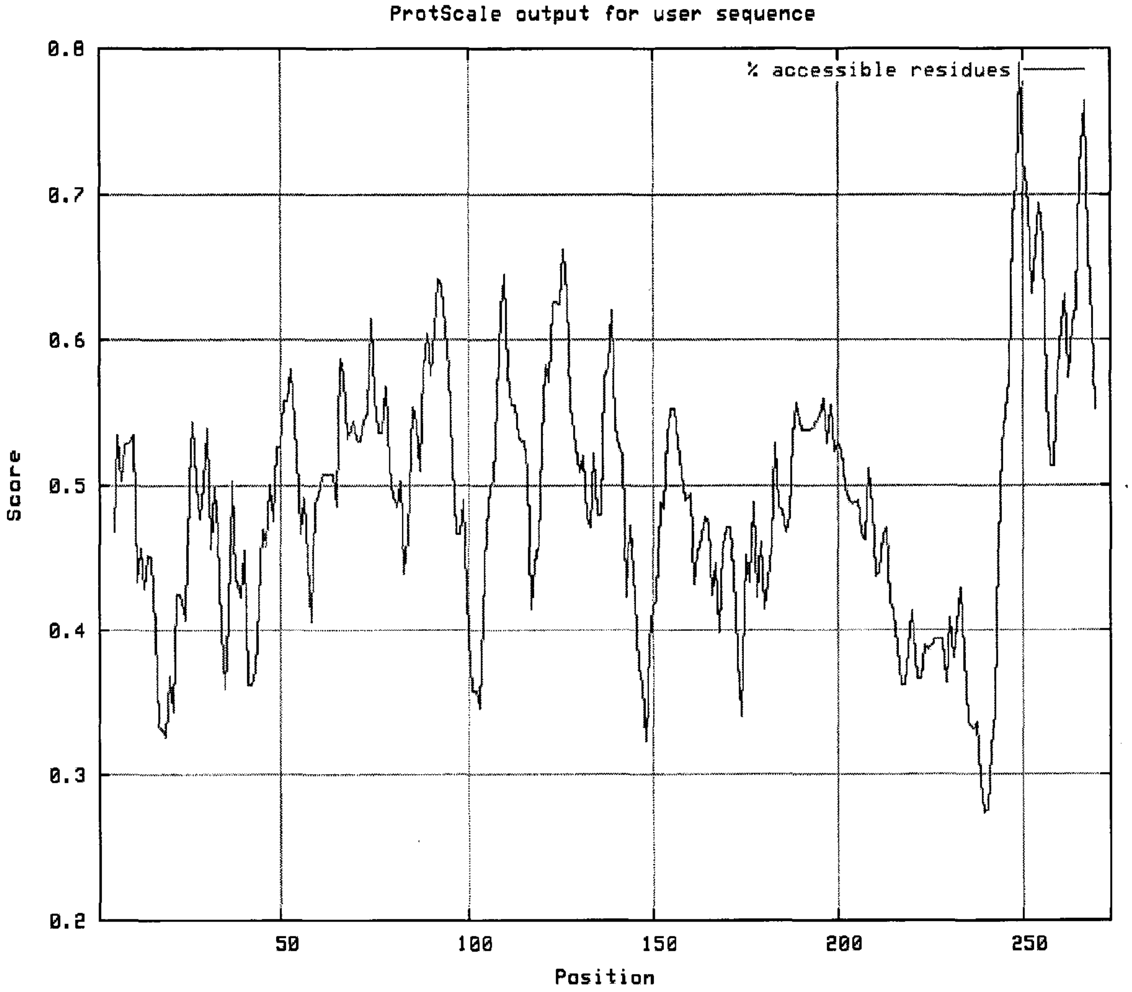


Figure 7b: 58P1D12 variant 2 % Accessible Residues Profile

(Janin J., 1979. Nature 277:491-492)

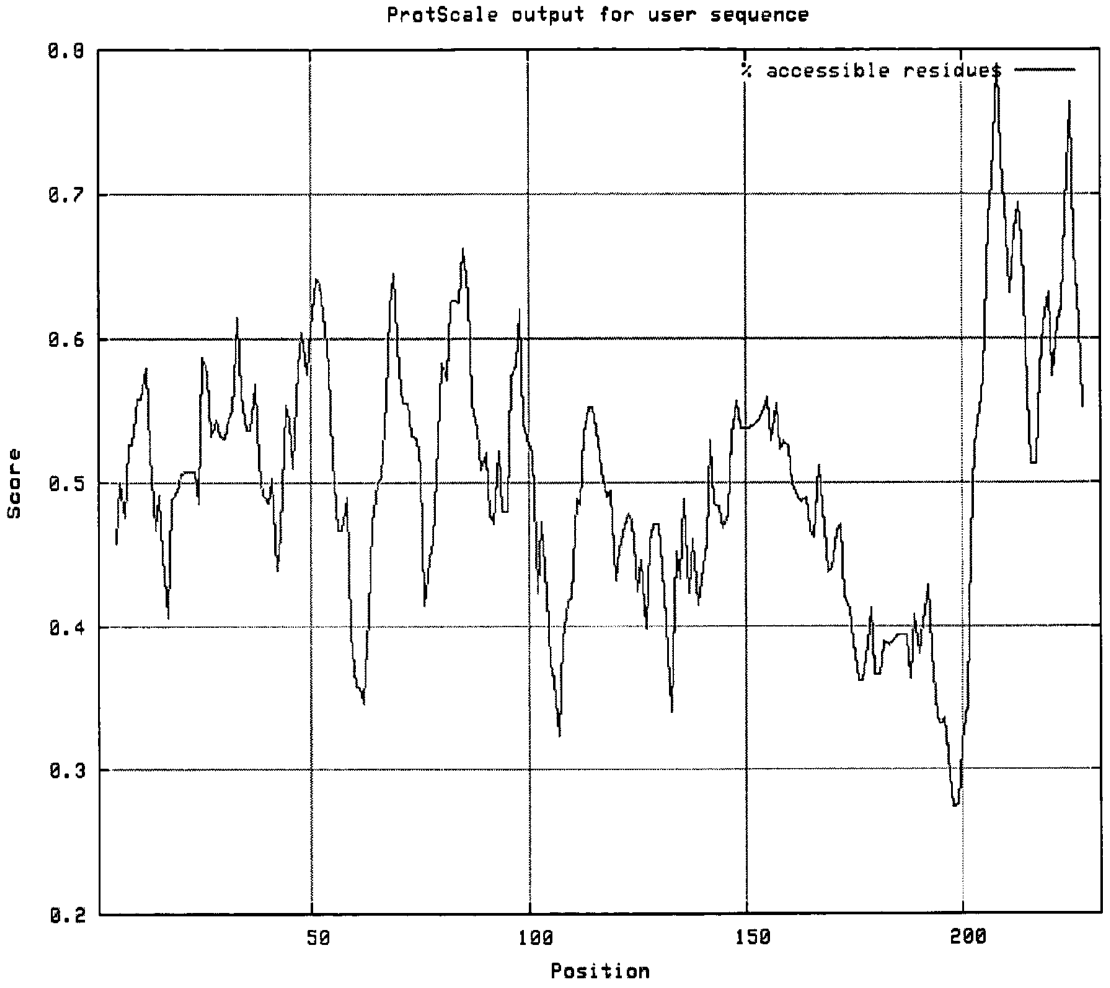


Figure 7c: 58P1D12 variant 3 % Accessible Residues Profile

(Janin J., 1979. Nature 277:491-492)

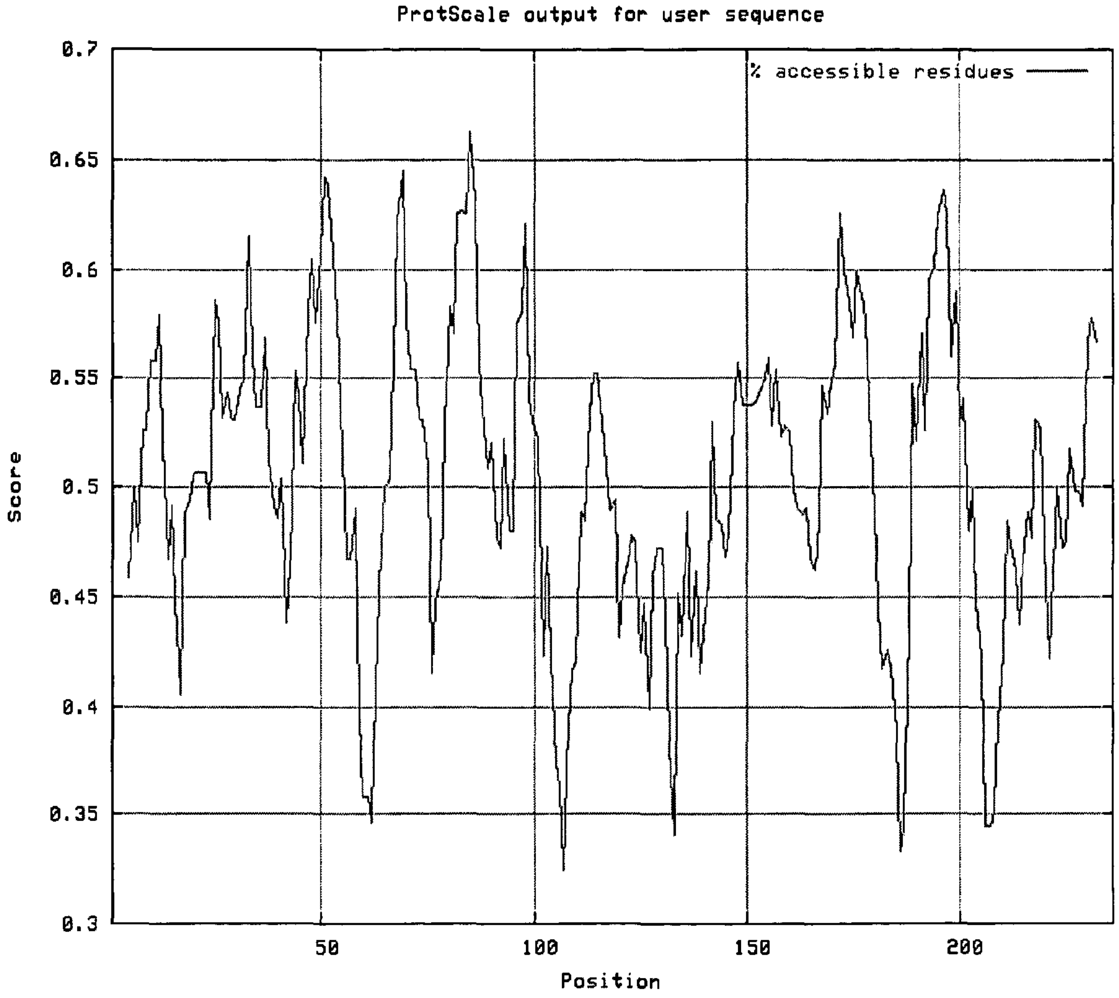


Figure 8a: 58P1D12 variant 1 Average Flexibility Profile

(Bhaskaran R., Ponnuswamy P.K., 1988.
Int. J. Pept. Protein Res. 32:242-255)

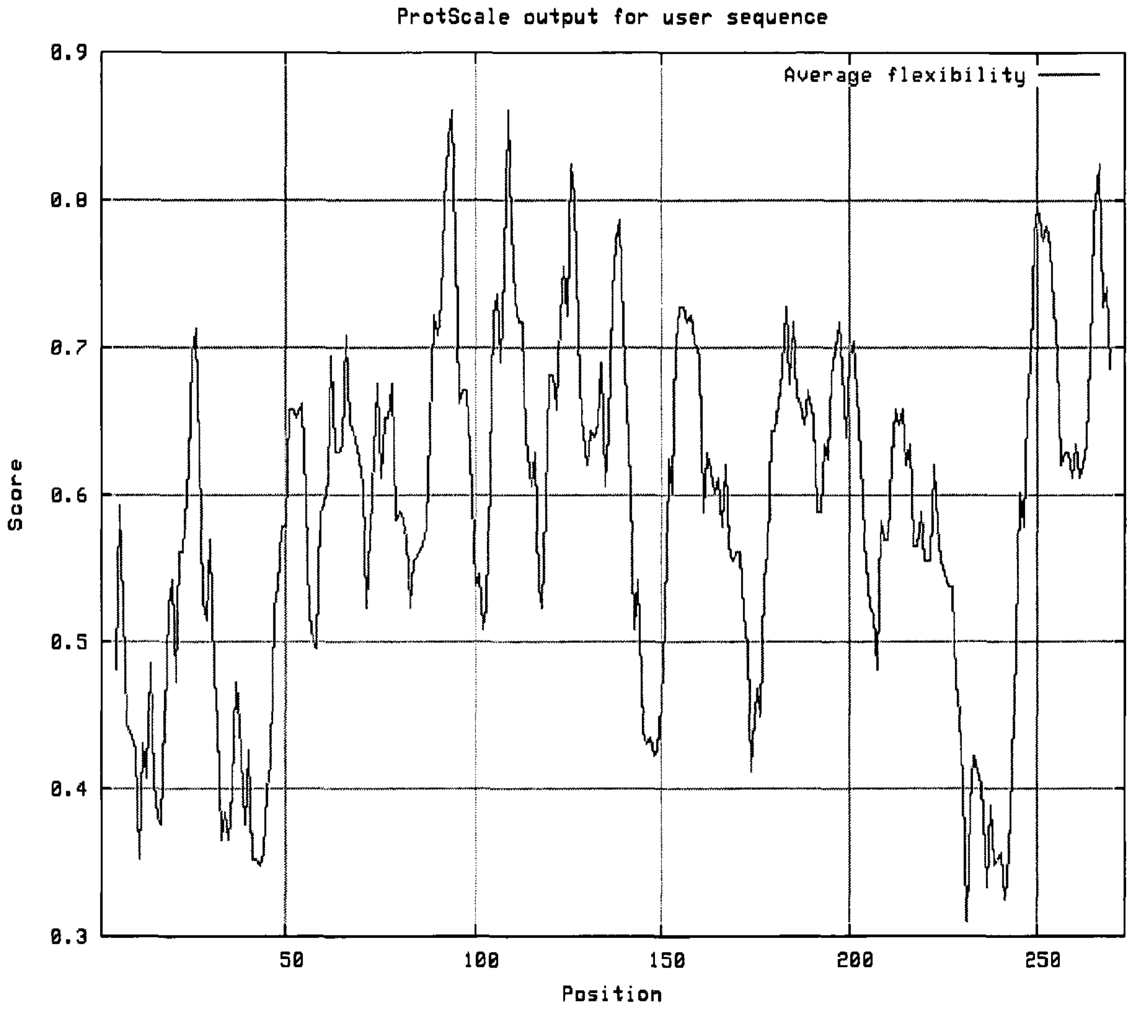


Figure 8b: 58P1D12 variant 2 Average Flexibility Profile

(Bhaskaran R., Ponnuswamy P.K., 1988. Int. J. Pept. Protein Res. 32:242-255)

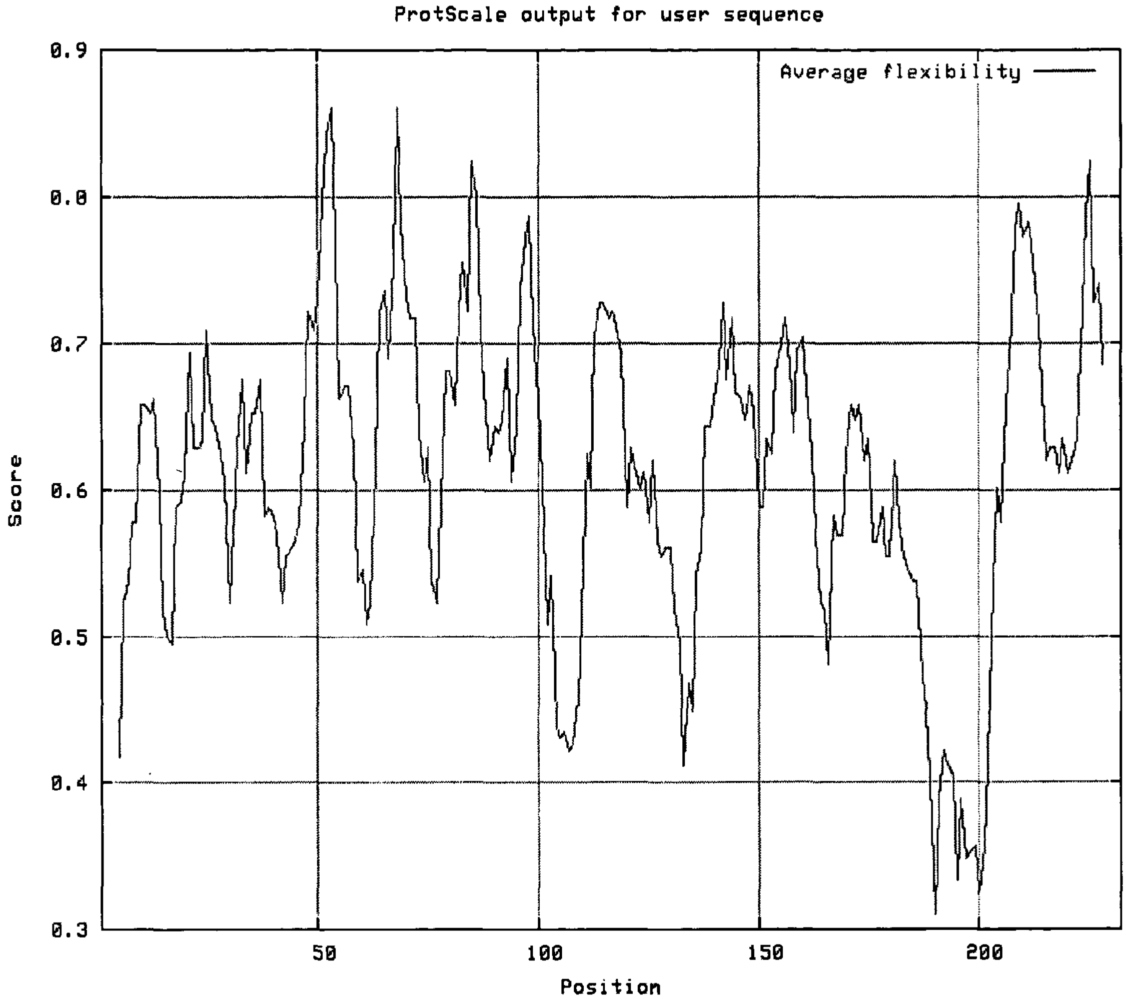


Figure 8c: 58P1D12 variant 3 Average Flexibility Profile

(Bhaskaran R., Ponnuswamy P.K., 1988. Int. J. Pept. Protein Res. 32:242-255)

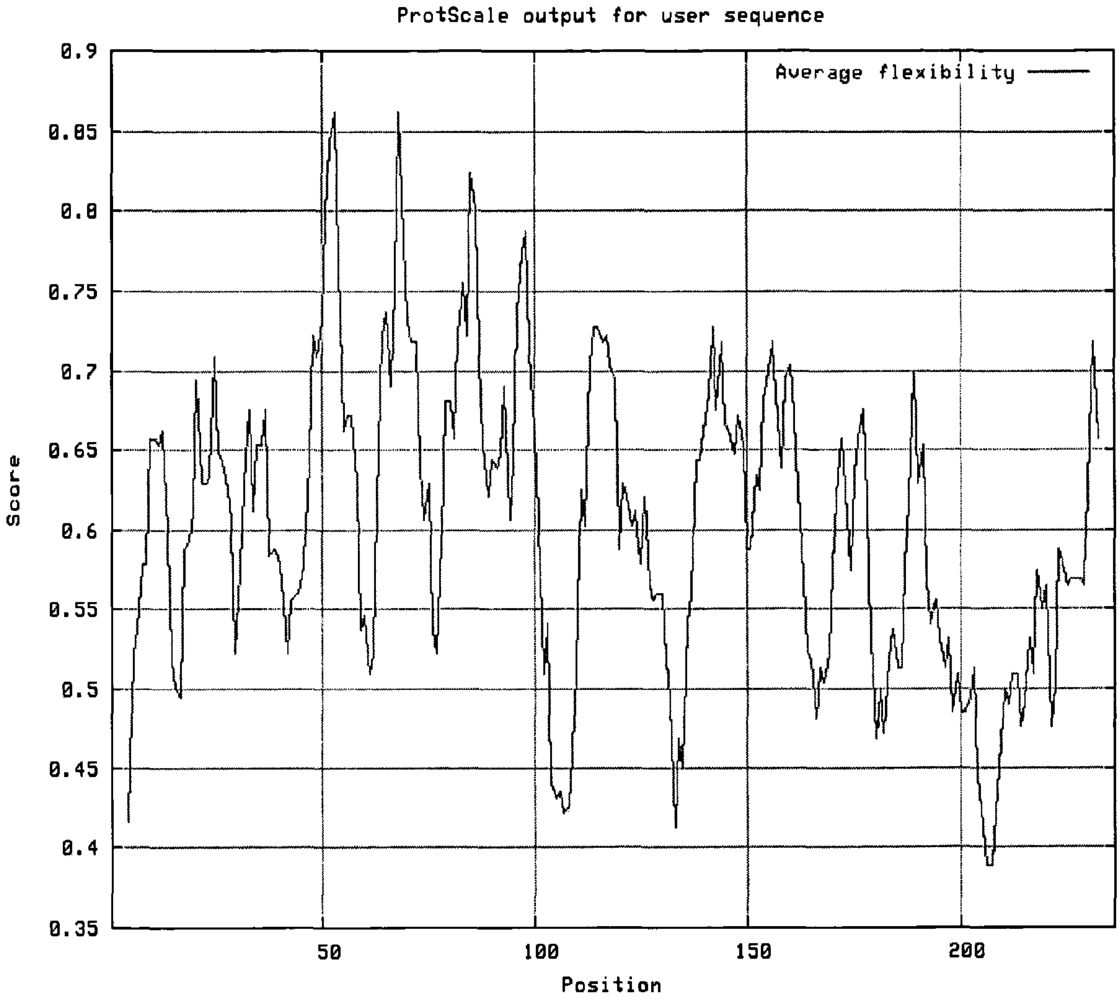


Figure 9a: 58P1D12 variant 1
Beta-turn Profile
(Deleage, G., Roux B. 1987. Protein Engineering 1:289-294)

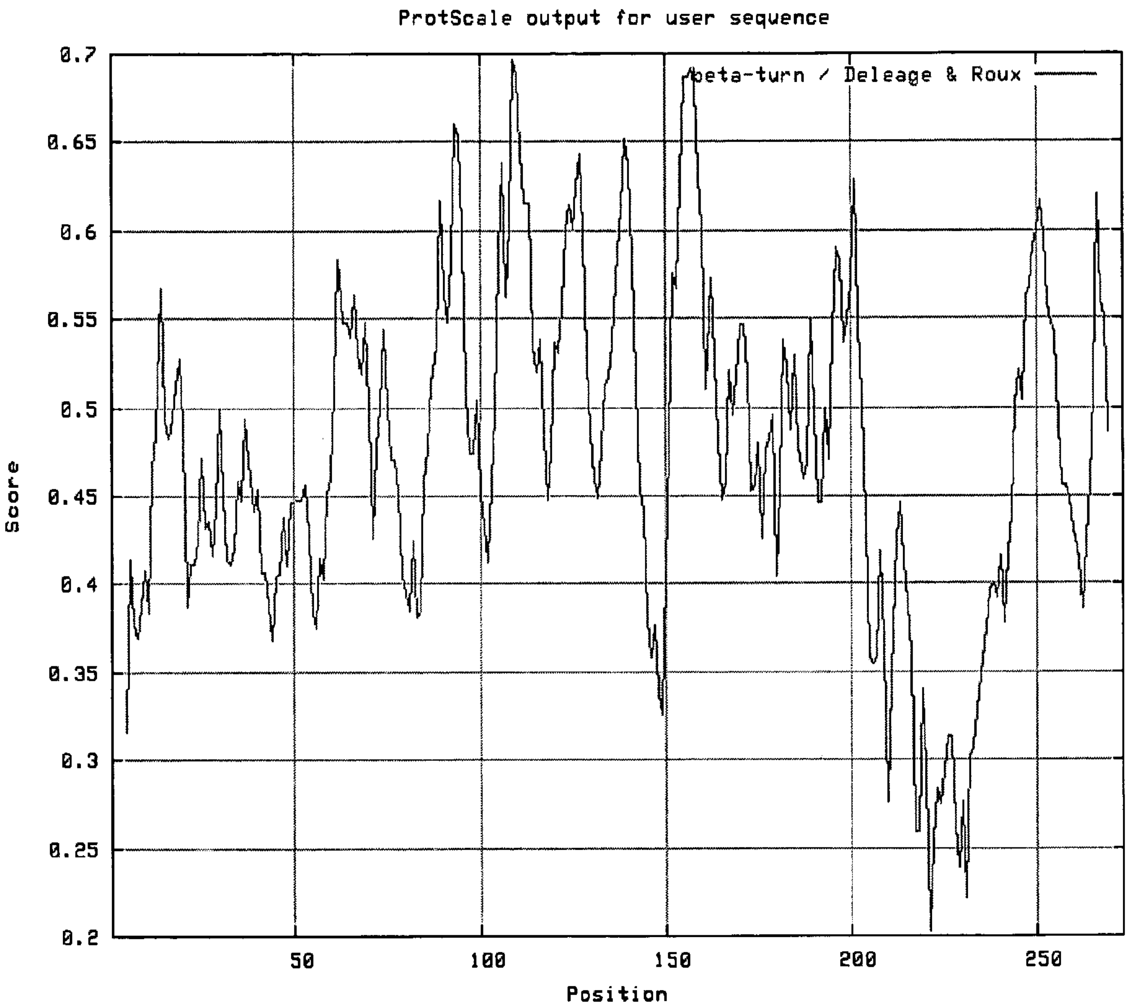


Figure 9b: 58P1D12 variant 2
Beta-turn Profile
(Deleage, G., Roux B. 1987. Protein Engineering 1:289-294)

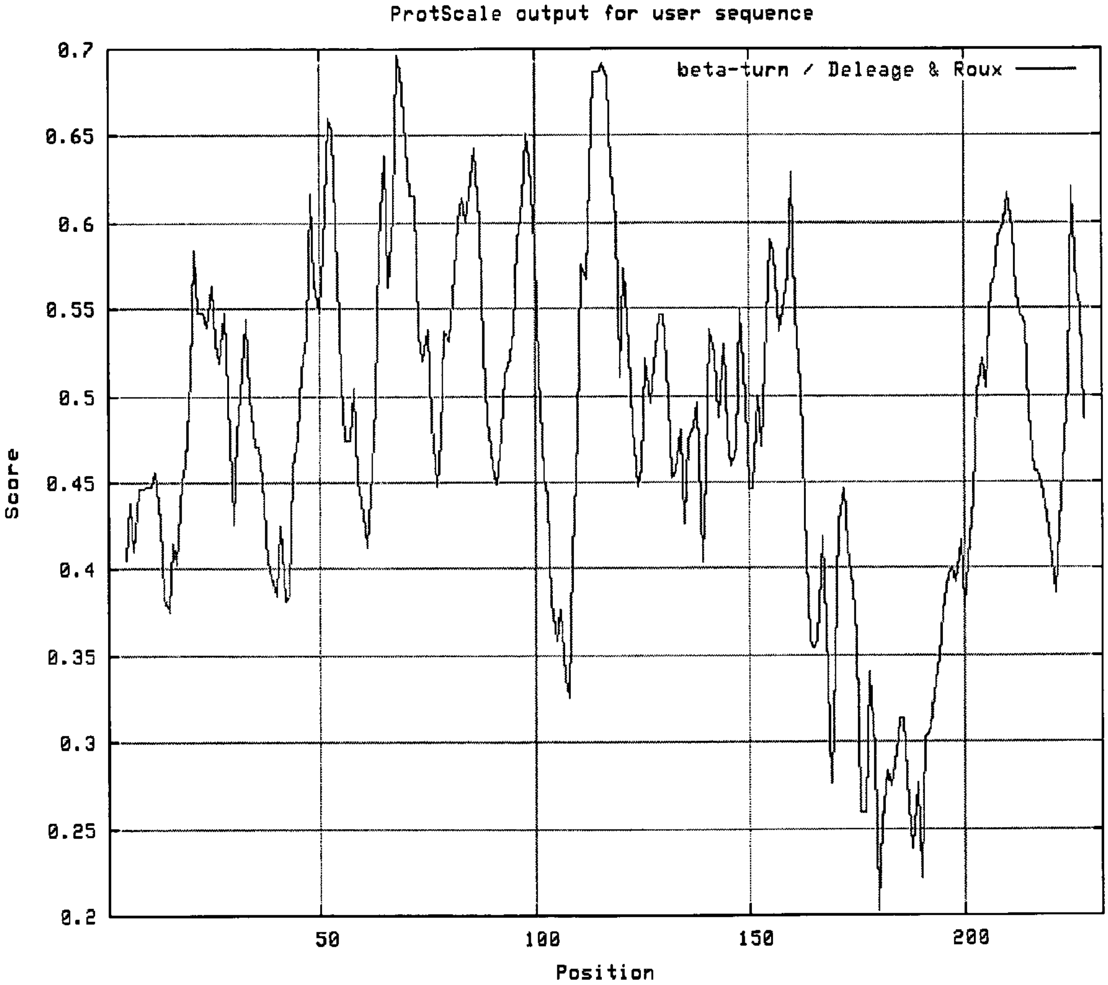


Figure 9c: 58P1D12 variant 3
Beta-turn Profile
(Deleage, G., Roux B. 1987. Protein Engineering 1:289-294)

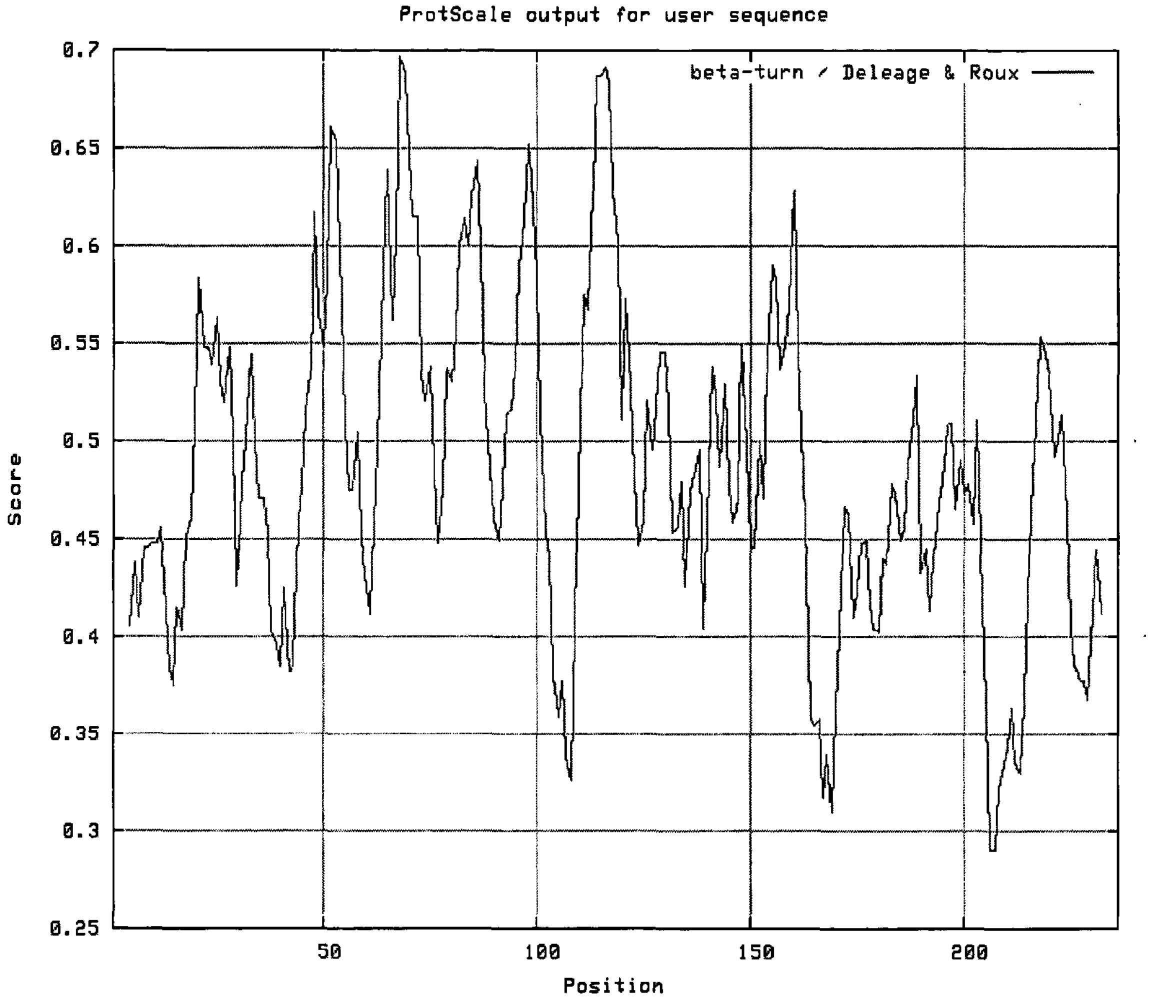
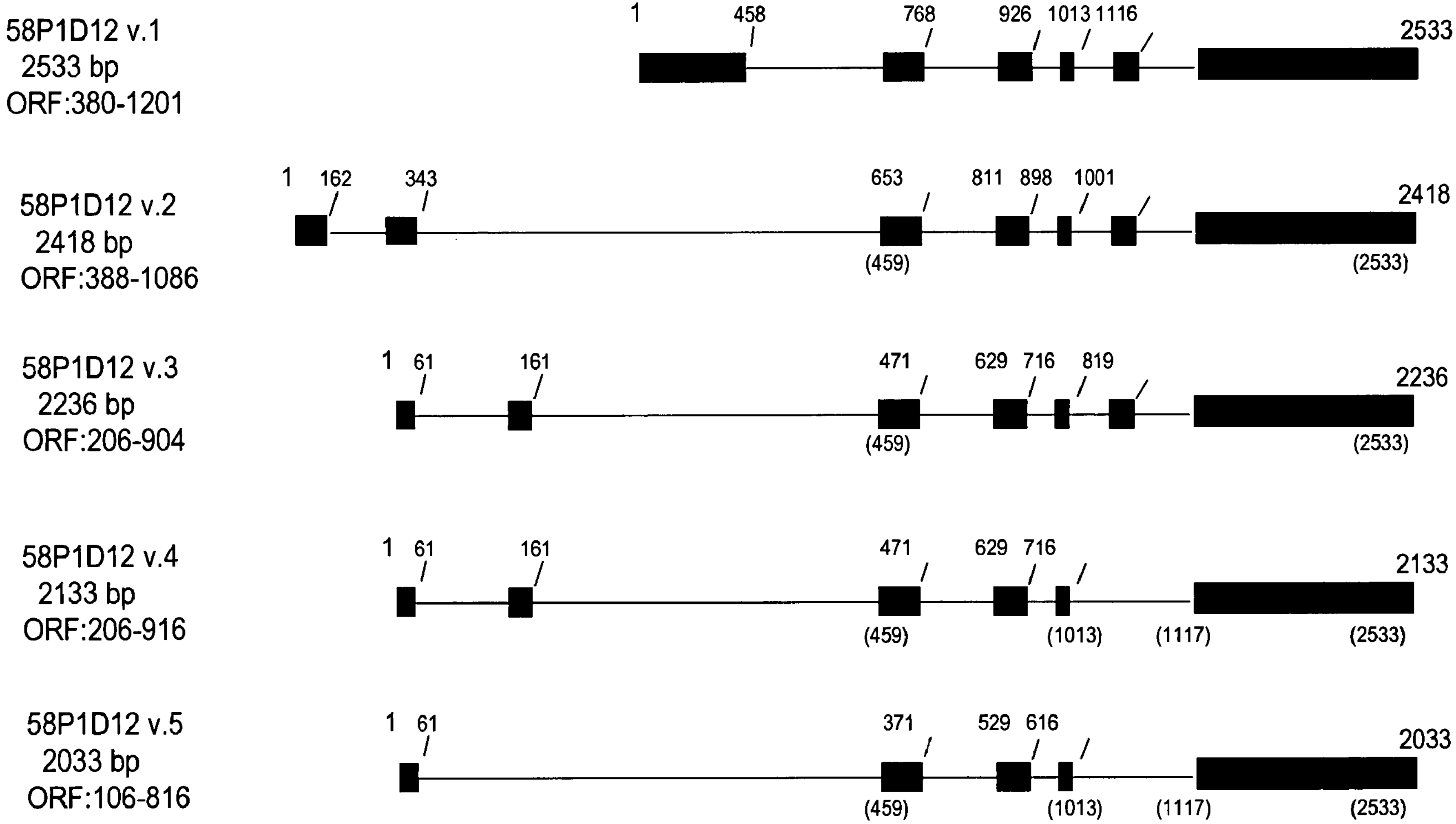


Figure 10


58P1D12 v.1
2533 bp
ORF:380-1201
1
458
768
926
1013
1116
2533
58P1D12 v.2
2418 bp
ORF:388-1086
1
162
343
653
811
898
1001
2418
(459)
(2533)
58P1D12 v.3
2236 bp
ORF:206-904
1
61
161
471
629
716
819
2236
(459)
(2533)
58P1D12 v.4
2133 bp
ORF:206-916
1
61
161
471
629
716
2133
(459)
(1013)
(1117)
(2533)
58P1D12 v.5
2033 bp
ORF:106-816
1
61
371
529
616
2033
(459)
(1013)
(1117)
(2533)

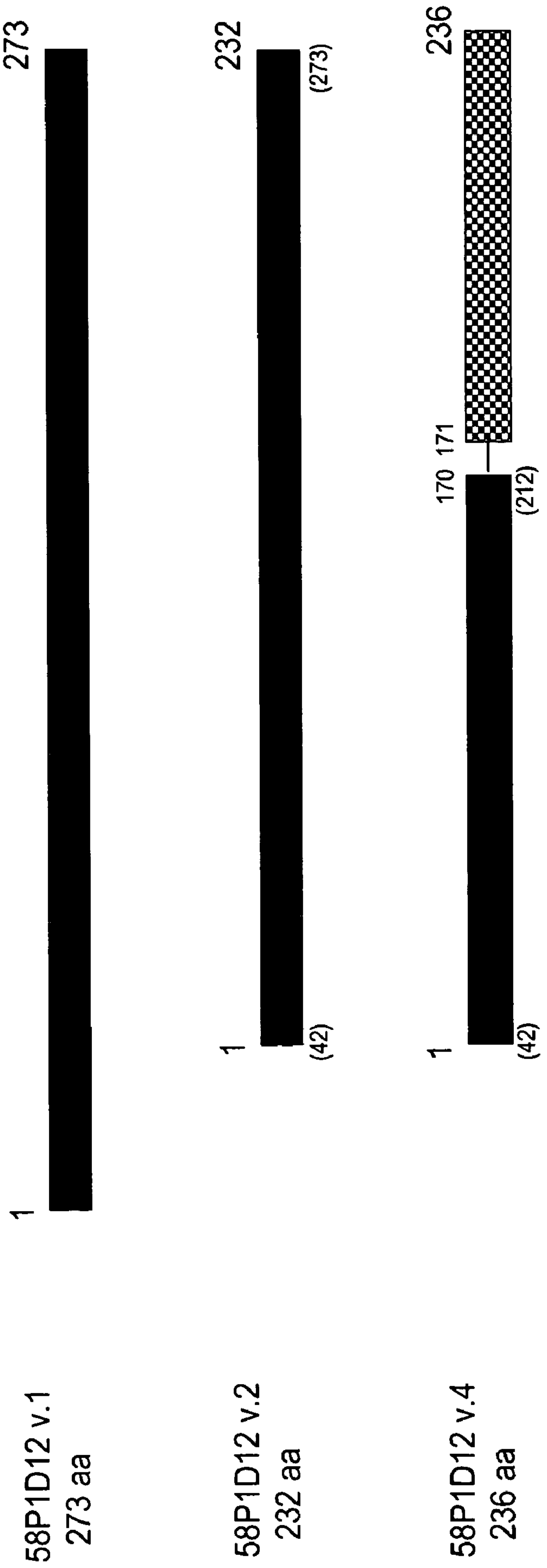

58P1D12 v.1
273 aa
1
273
58P1D12 v.2
232 aa
1
(42)
232
(273)
58P1D12 v.4
236 aa
1
(42)
170
(212)
171
236

Figure 11

Figure 12

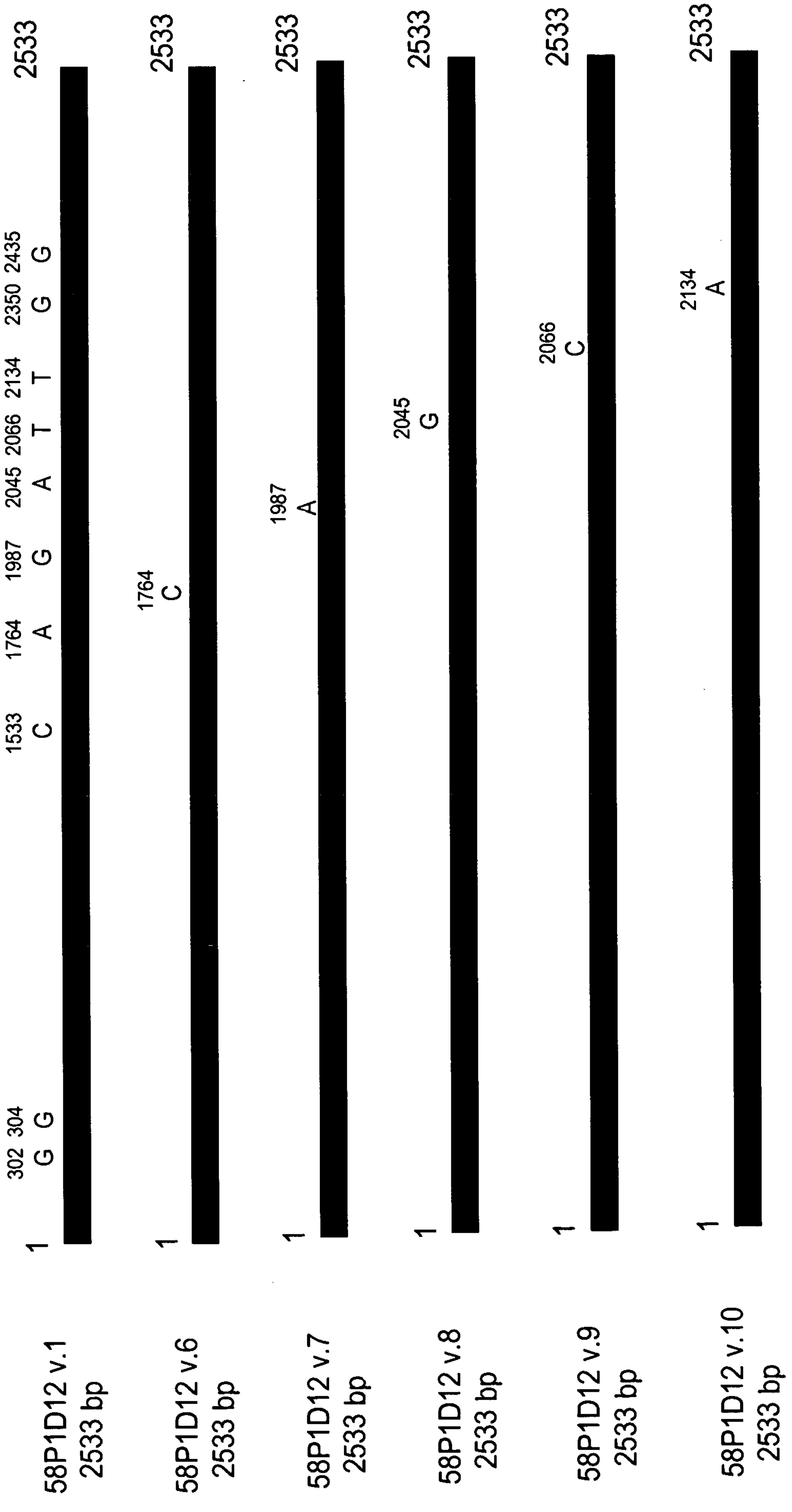

58P1D12 v.1
2533 bp
1
302 G
304 G
1533 C
1764 A
1987 G
2045 A
2066 T
2134 T
2350 G
2435 G
2533
58P1D12 v.6
2533 bp
1
1764 C
2533
58P1D12 v.7
2533 bp
1
1987 A
2533
58P1D12 v.8
2533 bp
1
2045 G
2533
58P1D12 v.9
2533 bp
1
2066 C
2533
58P1D12 v.10
2533 bp
1
2134 A
2533

Figure 12 (con'd)

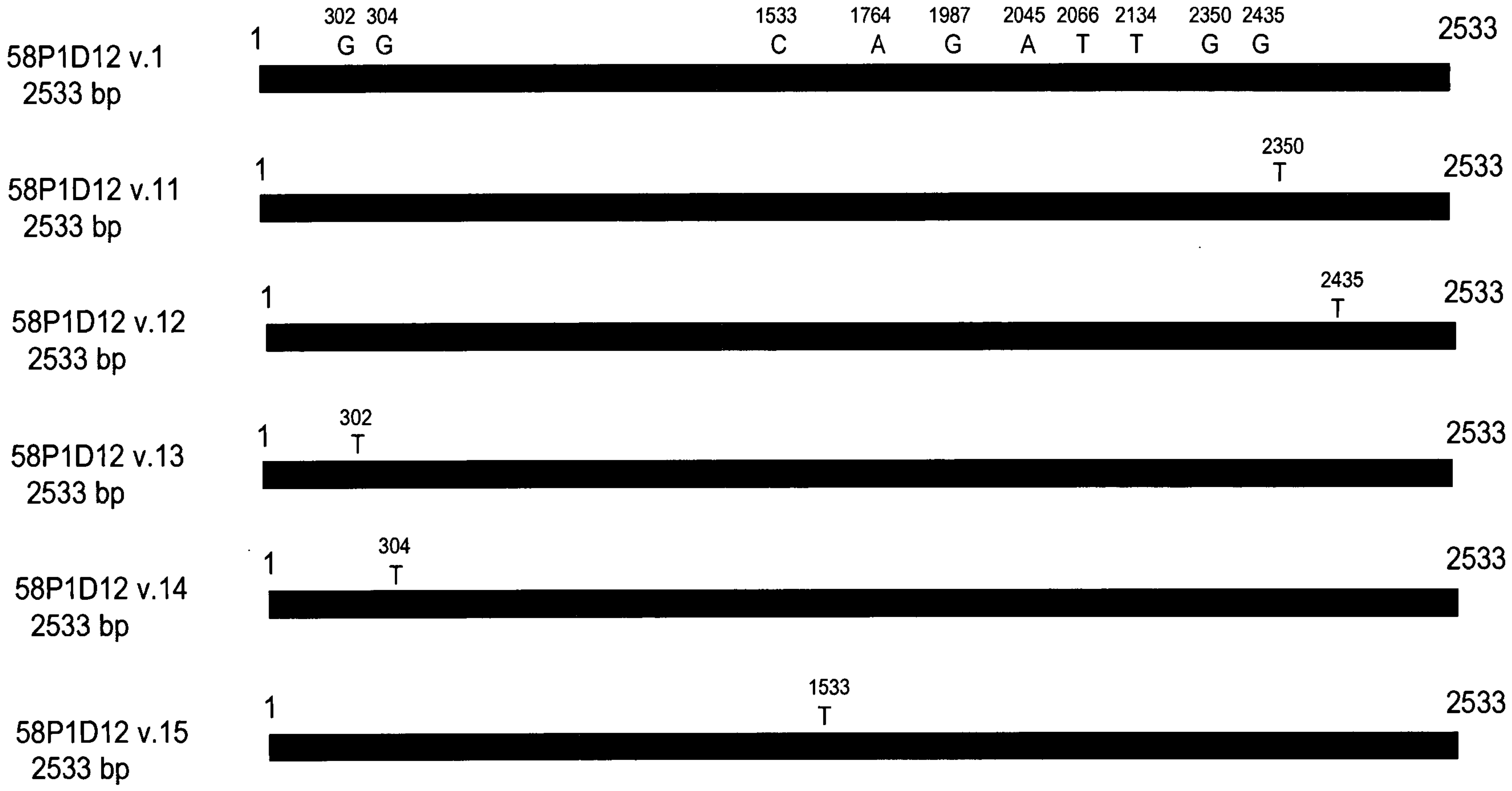


Figure 13a: Secondary structure prediction of 58P1D12 variant 1

```
         10        20        30        40        50        60        70
          |         |         |         |         |         |         |
MSRVVSLLLGAALLCGHGAFCRRVVSGQKVCFADFKHPCYKMAYFHELSSRVSFQEARLACESEGGVLLS
chhhhhhhhhhhhhhcccchheeeeccceeeeecccccccehhhhhcccccchhhhhhhhhhhccceeee
LENEAEQKLIESMLQNLTKPGTGISDGDFWIGLWRNGDGQTSGACPDLYQWSDGSNSQYRNWYTDEPSCG
echhhhhhhhhhhhhhhcccccccccccceeeehccccccccccccccccccccccccccccccccccccc
SEKCVVMYHQPTANPGLGGPYLYQWNDDRCNMKHNYICKYEPEINPTAPVEKPYLTNQPGDTHQNVVVTE
cceeeeeeeccccccccccceeeecccccccccccceeeecccccccccccccccccccccccccceeeeee
AGIIPNLIYVVIPTIPLLLLILVAFGTCCFQMLHKSKGRTKTSPNQSTLWISKSTRKESGMEV
ccecccceeeeeccchhhhhhhhhhhcchhhhhhhcccccccccccccceeeeeccccccccccc
```

```
Alpha helix     (h): 24.18%
Extended strand (e): 18.68%
Random coil     (c): 57.14%
```

Figure 13b: Secondary structure prediction of 58P1D12 variant 2

```
               10         20         30         40         50         60         70
                |          |          |          |          |          |          |
      MAYFHELSSRVSFQEARLACESEGGVLLSLENEAEQKLIESMLQNLTKPGTGISDGDFWIGLWRNGDGQT
      cceheccccchhhhhhhhhhhccceeeeechhhhhhhhhhhhhhhccccccccccceeeehcccccccc
      SGACPDLYQWSDGSNSQYRNWYTDEPSCGSEKCVVMYHQPTANPGLGGPYLYQWNDDRCNMKHNYICKYE
      ccccccccccccccccccccccccccccccceeeeeeecccccccccccceeeecccccccccccceeeecc
      PEINPTAPVEKPYLTNQPGDTHQNVVVTEAGIIPNLIYVVIPTIPLLLLILVAFGTCCFQMLHKSKGRTK
      ccccccccccccccccccccccceeeeeecceccceeeeeccchhhhhhhhhhhcchhhhhhhccccccc
      TSPNQSTLWISKSTRKESGMEV
      cccccceeeeeccccccccccc


                  Alpha helix     (h): 19.83%
                  Extended strand (e): 18.97%
                  Random coil     (c): 61.21%
```

Figure 13c: Secondary structure prediction of 58P1D12 variant 3

```
         10        20        30        40        50        60        70
          |         |         |         |         |         |         |
MAYFHELSSRVSFQEARLACESEGGVLLSLENEAEQKLIESMLQNLTKPGTGISDGDFWIGLWRNGDGQT
cceehecccccchhhhhhhhhhccceeeeechhhhhhhhhhhhhhhccccccccccceeeehcccccccc
SGACPDLYQWSDGSNSQYRNWYTDEPSCGSEKCVVMYHQPTANPGLGGPYLYQWNDDRCNMKHNYICKYE
ccccccccccccccccccccccccccccccceeeeeeeccccccccccceeeeccccccccccceeeecc
PEINPTAPVEKPYLTNQPGDTHQNVVVTEAVKEEQKLVQTSLHCGFQRVPEKKVAWKYNNSLTWFQNFVI
cccccccccccccccccccccceeeeeehhhhhhhhhhhhhhhhcchhhhchheeeccccchhhhhhhhh
LDLYKEWHQNNSLEWLEITKDLQDEL
hhhhhhhccccchhhhhhhhhhhhcc
```

```
Alpha helix     (h): 32.20%
Extended strand (e): 15.25%
Random coil     (c): 52.54%
```

Transmembrane prediction for 58P1D12 variant 1

Figure 13D

1 transmembrane domain predicted

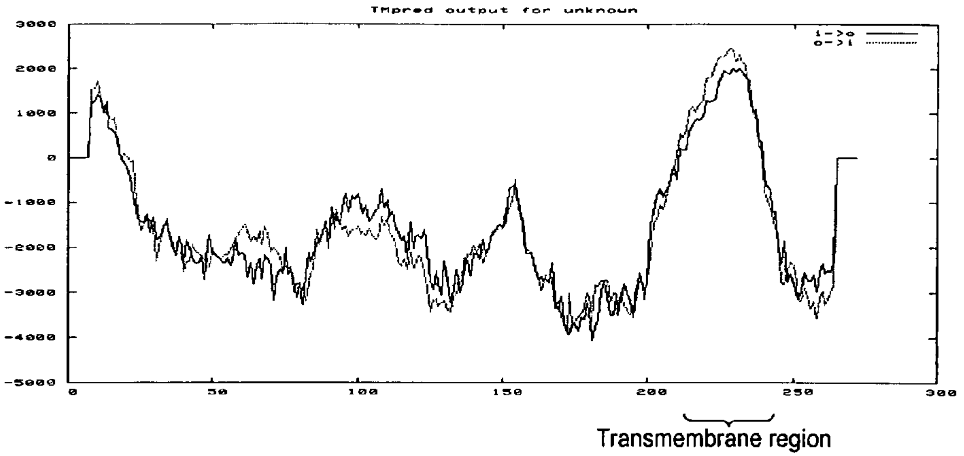
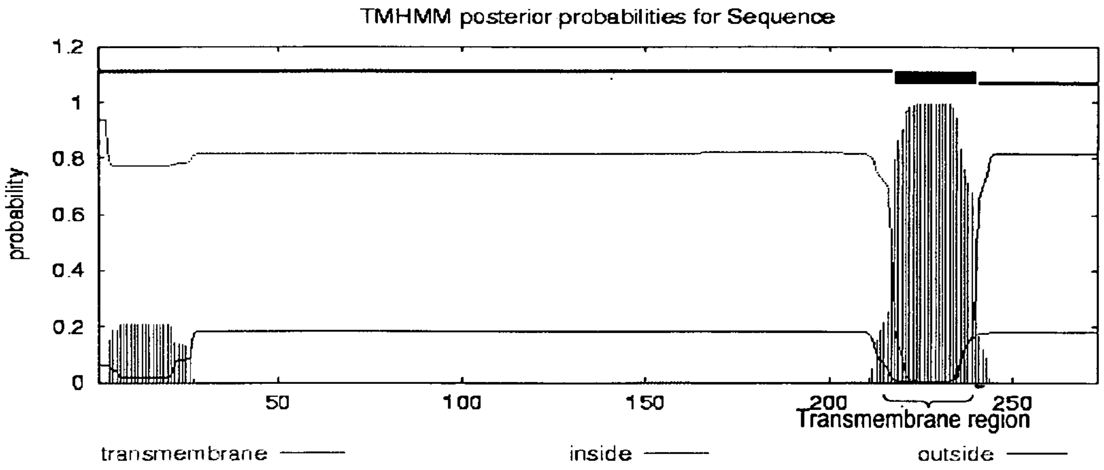


Figure 13E

1 transmembrane domain predicted



Transmembrane prediction for 58P1D12 variant 2

Figure 13F

1 transmembrane domain predicted

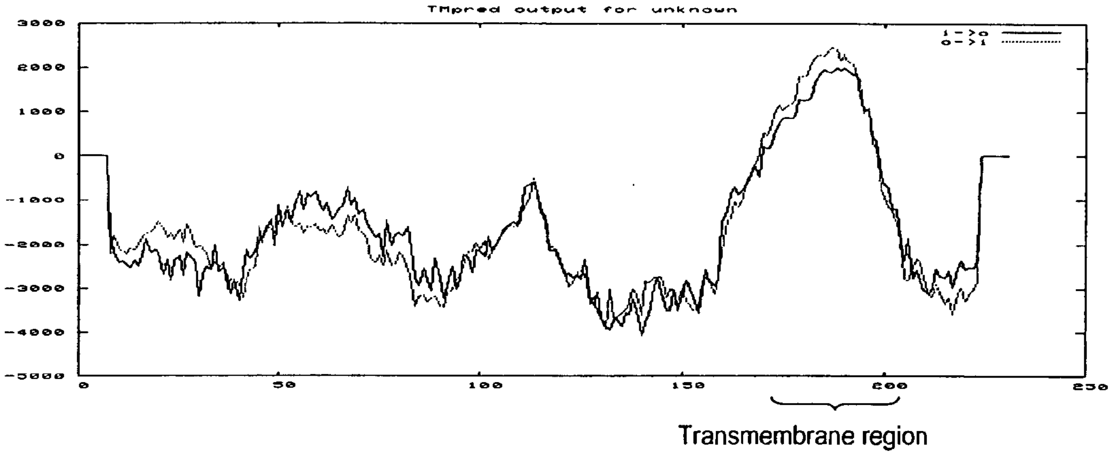
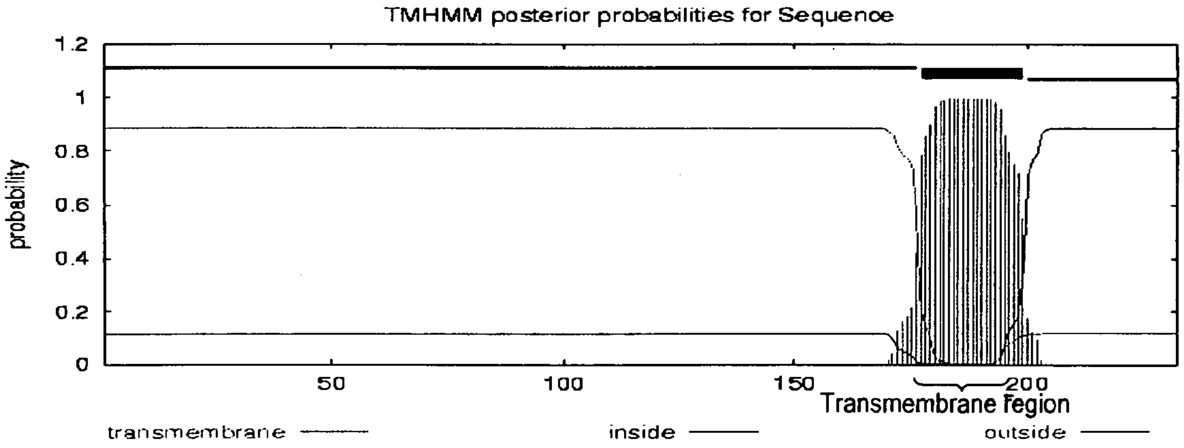


Figure 13G

1 transmembrane domain predicted



Transmembrane prediction for 58P1D12 variant 3

Figure 13H

No transmembrane domain predicted

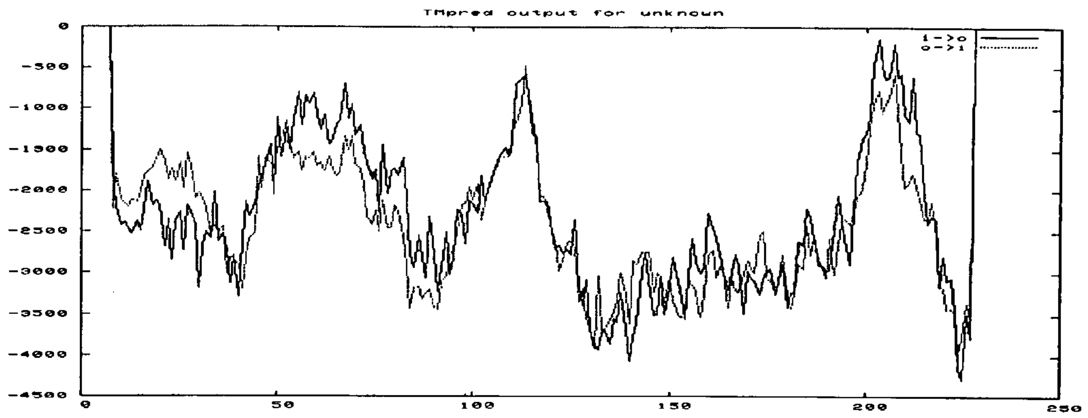


Figure 13I

No transmembrane domain predicted

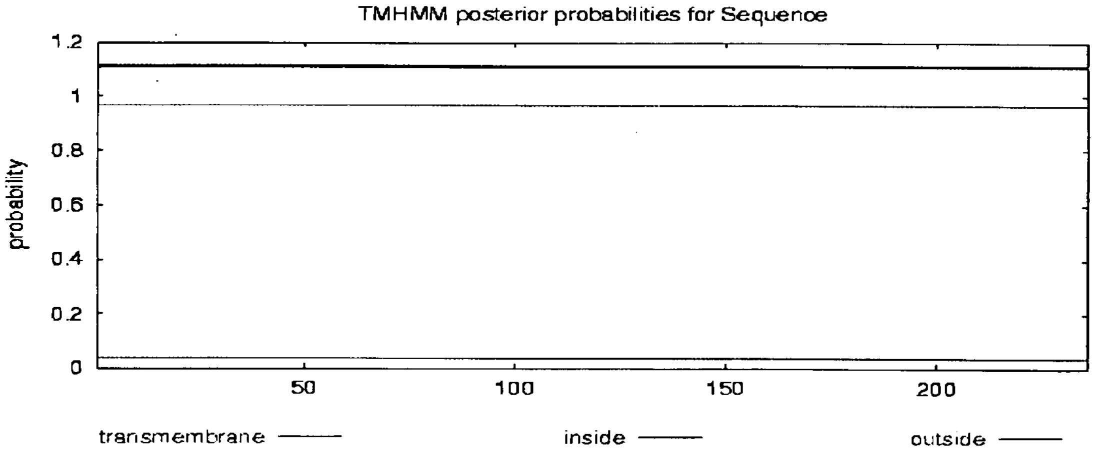
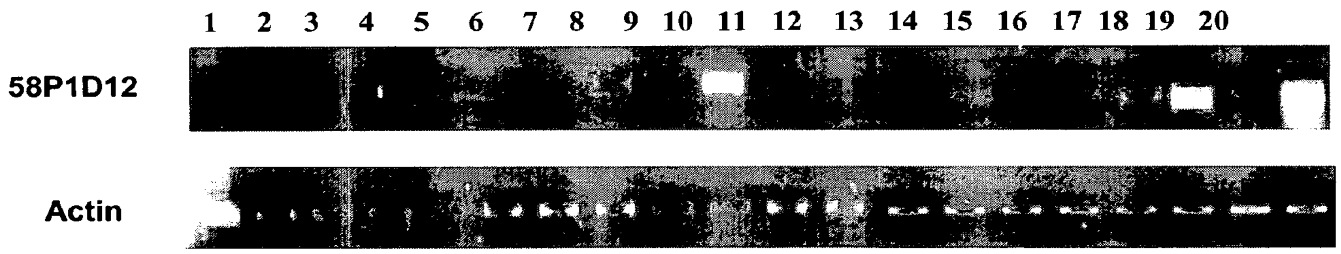
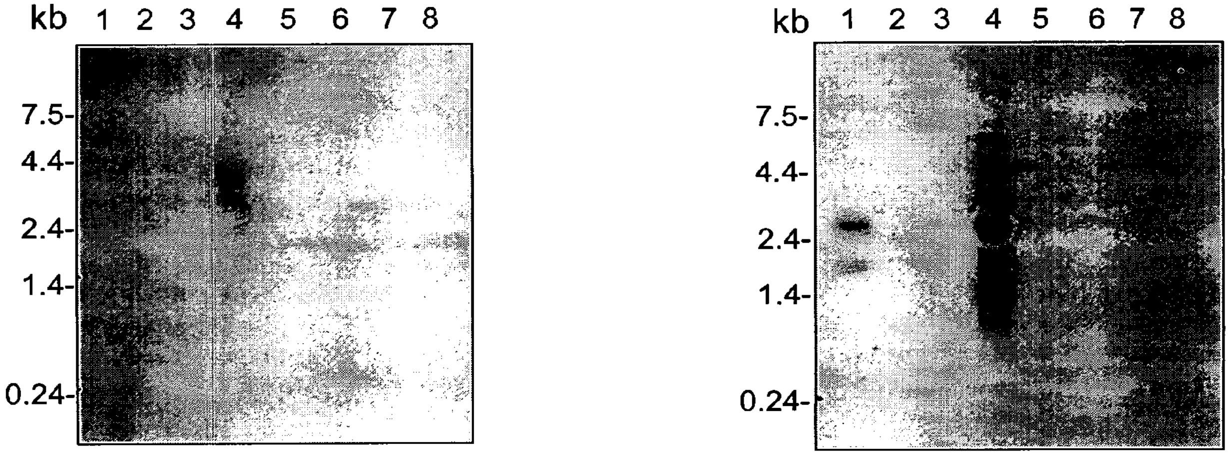


Figure 14 58P1D12 Expression in Normal and Patient Cancer Tissues



1. Bladder
2. Brain
3. Heart
4. Kidney
5. Liver
6. Lung
7. Prostate
8. Spleen
9. Skeletal Muscle
10. Testis
11. Pancreas
12. Colon
13. Stomach
14. Prostate cancer
15. Bladder cancer
16. Lung cancer
17. Ovary cancer
18. LAPC prostate xenograft pool
19. Bone cancer/Melanoma
20. Cervical cancer

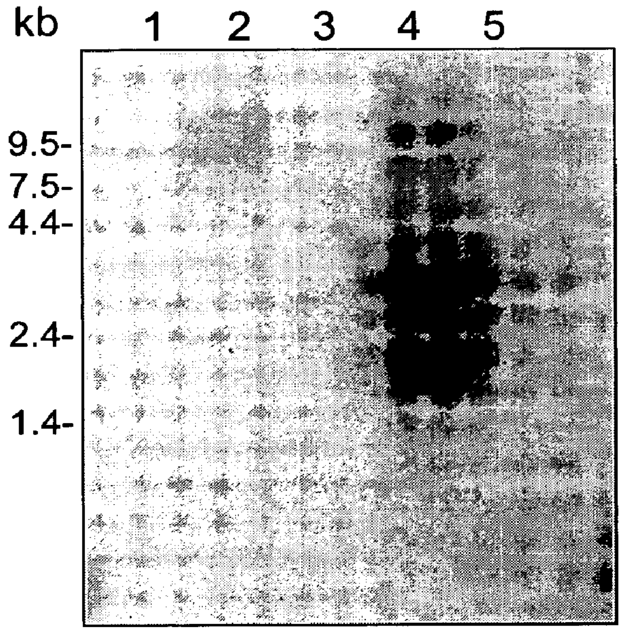
Figure 15 58P1D12 Expression in Normal Human Tissues



1. Heart
2. Brain
3. Placenta
4. Lung
5. Liver
6. Skeletal Muscle
7. Kidney
8. Pancreas

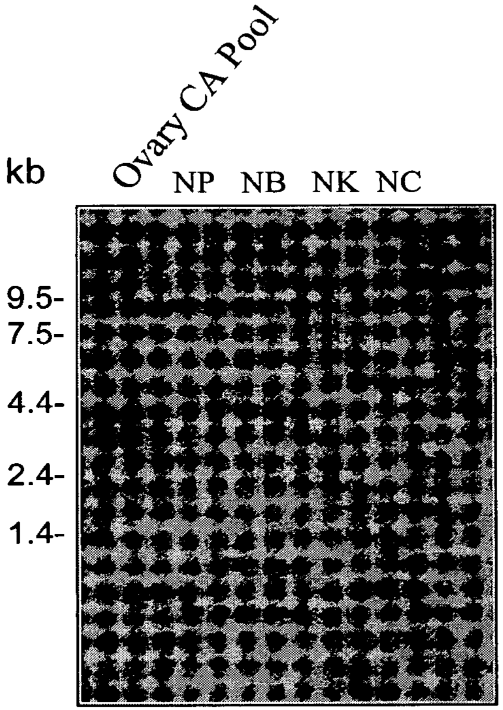
1. Spleen
2. Thymus
3. Prostate
4. Testis
5. Uterus
6. Small Intestine
7. Colon
8. Leukocytes

Figure 16 58P1D12 Expression in Prostate Cancer Xenograft tissues



1. Normal Prostate
2. LAPC-4 AD
3. LAPC-4 AI
4. LAPC-9 AD
5. LAPC-9 AI

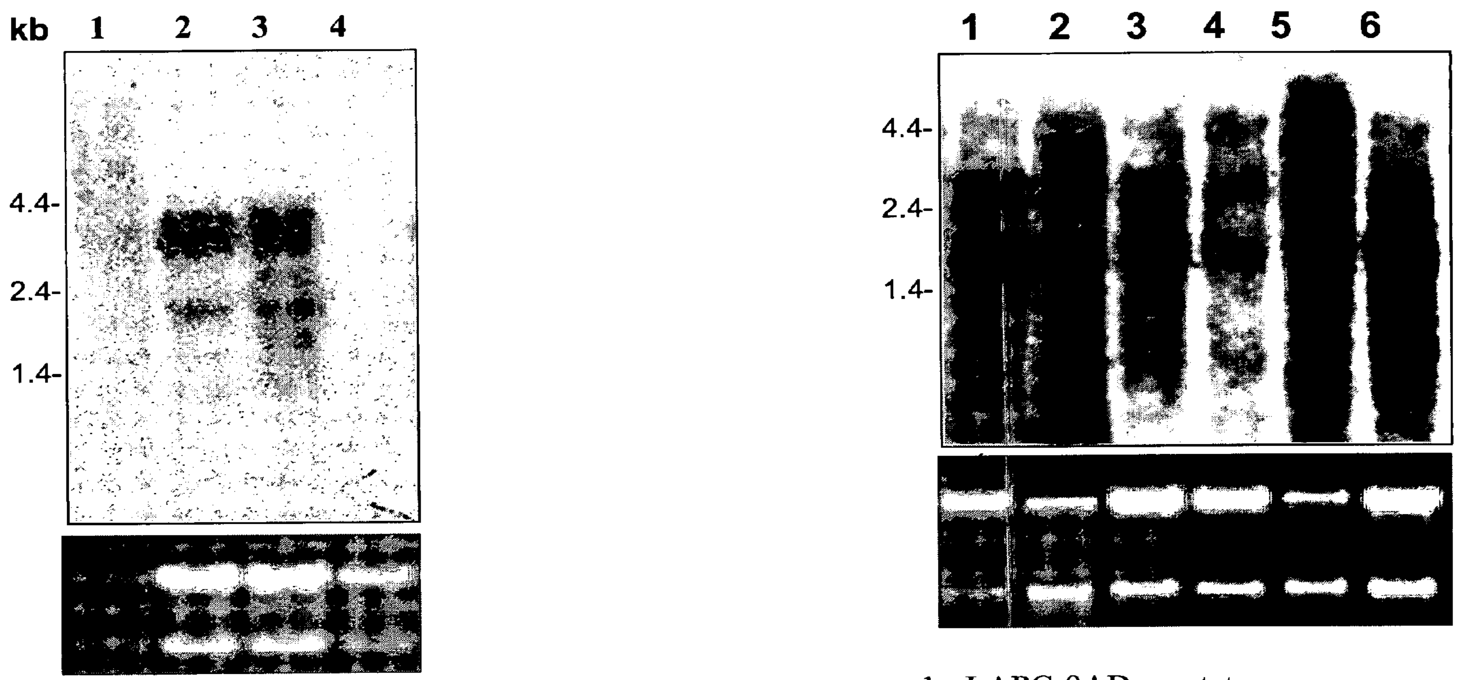
Figure 17 58P1D12 Expression in Normal and Patient Cancer Specimens



Ovary CA Pool = 10μg total RNA/per lane from a pool of 3 tumors as follows: *grade 2, 2, 3*

NP = Normal Prostate
NB = Normal Bladder
NK = Normal Kidney
NC = Normal Colon

Figure 18 58P1D12 Expression in Ovarian Cancer Patient Specimens



1. Normal Ovary
2. Ovary T - Papillary serous CA
3. Ovary T - Endometrioid Adenocarcinoma
4. Ovary T - Mucinous Cystadenocarcinoma

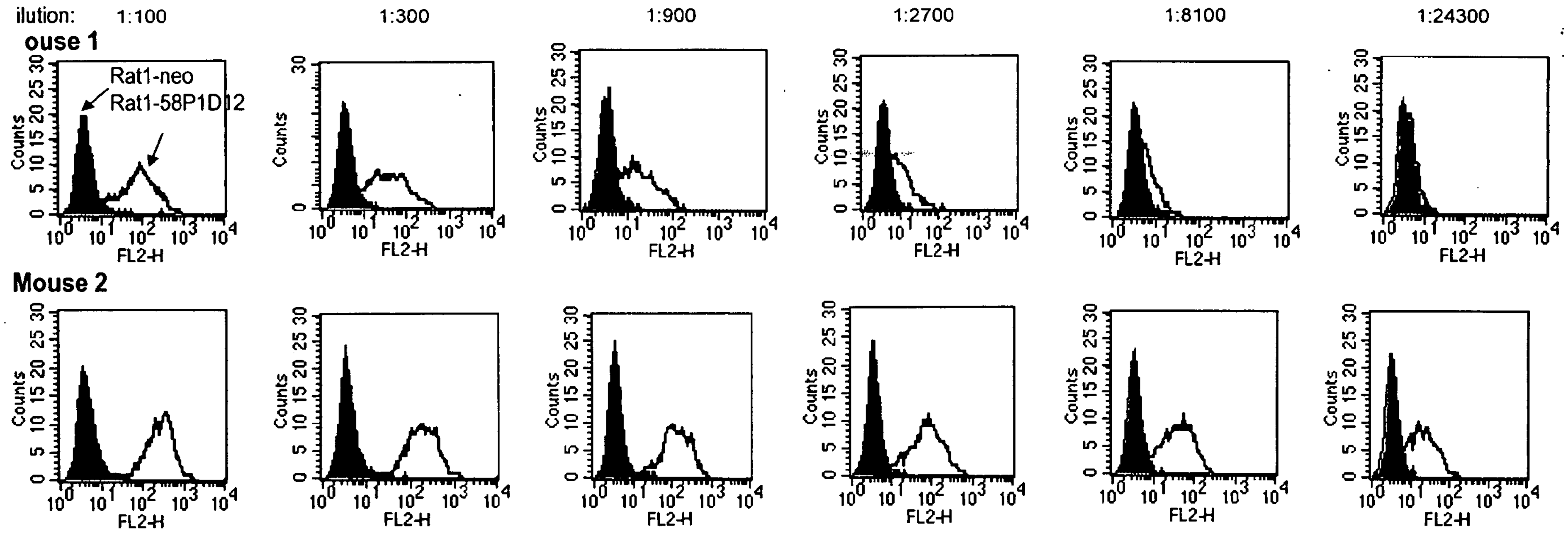
1. LAPC-9AD prostate cancer xenograft
2. Ovary T - Endometrioid Adenocarcinoma
3. Ovary T - Papillary Serous carcinoma
4. Ovary T - Mixed malignant, mullerian tumor
5. Ovary T - Granulosa, cystic variant
6. Ovary T - Papillary Serous carcinoma

Figure 19 58P1D12 Expression in a Large Panel of Ovarian Cancer Patient Specimens

Panel#	Diagnosis	Grade	Ovary CA expression
1	N.Ovary		□
1b	N.Ovary		□
2	Adeno CA Met to Omentum		░
3	AdenoCA	poorly diff	▓
4	AdenoCA	Grade 2	▓
5	AdenoCA mixed endometrioid & papillary		■
6	Clear cell CA	High grade	□
7	Clear cell CA	Grade 1	░
8	Endometrioid AdenoCA	2	▓
9	Endometrioid AdenoCA		░
10	Endometrioid AdenoCA		□
11	Endometrioid Adenocarcinoma		□
12	Endometrioid CA	Mod diff	□
13	Endometrioid CA	Grade 2	□
14	Epithelial neoplasm	poorly diff	■
15	Granulosa	Figo Ia1	□
16	Granulosa		░
17	Met to fallopian tube		□
18	Met to Intestine	Undiff.	▓
19	Met to Intestine		░
20	Met to Intestine		▓
21	Met to Intestine		▓
22	Met to Intestine		□
23	Met to Intestine		░
24	Mixed malignant mullerian tumor		□
25	Mucinous CystadenoCA		□
26	Mucinous CystadenoCA		□
27	Mucinous CystadenoCA	2	□
28	Mucinous intraglandular CA		░
29	Papillary Serous Adenocarcinoma		░
30	Papillary serous CA	3	■
31	Papillary serous CA	3	░
32	Papillary serous CA	Figo III	■
33	Papillary serous CA	Poor Diff.	░
34	Papillary Serous CA	Mod diff	░
35	Papillary Serous CA		□
36	Papillary Serous CA		░
37	Papillary serous CA	poorly diff	░
38	Papillary serous CA	mod to poorly diff	▓
39	Papillary serous CA	Grade IIIc	░
40	Papillary serous met		░
41	Papillary serous met to omentum		▓
42	Serous CA	3	□
43	Serous cystadenocarcinoma	Grade 1	░
44	Sertoli Leydig		░
45	Small cell carcinoma	Undiff.	░
46	Squamous	3	░
47	Yolk sac tumor	Stage IIC	□
48	ND		▓
49	ND		░
50	ND		░
Percent			75.5%

□	*Negative expression*
░	*Weak expression*
▓	*Medium expression*
■	*Strong expression*

Figure 20: Specific cell surface staining of Rat1-58P1D12 cells with serum from mice immunized with recombinant 58P1D12 protein



Filled histogram: Rat1-neo cells
Open histogram: Rat1-58P1D12 cells

Figure 21: Specific cell surface staining of Rat1-58P1D12 cells with monoclonal antibody hybridoma supernatants

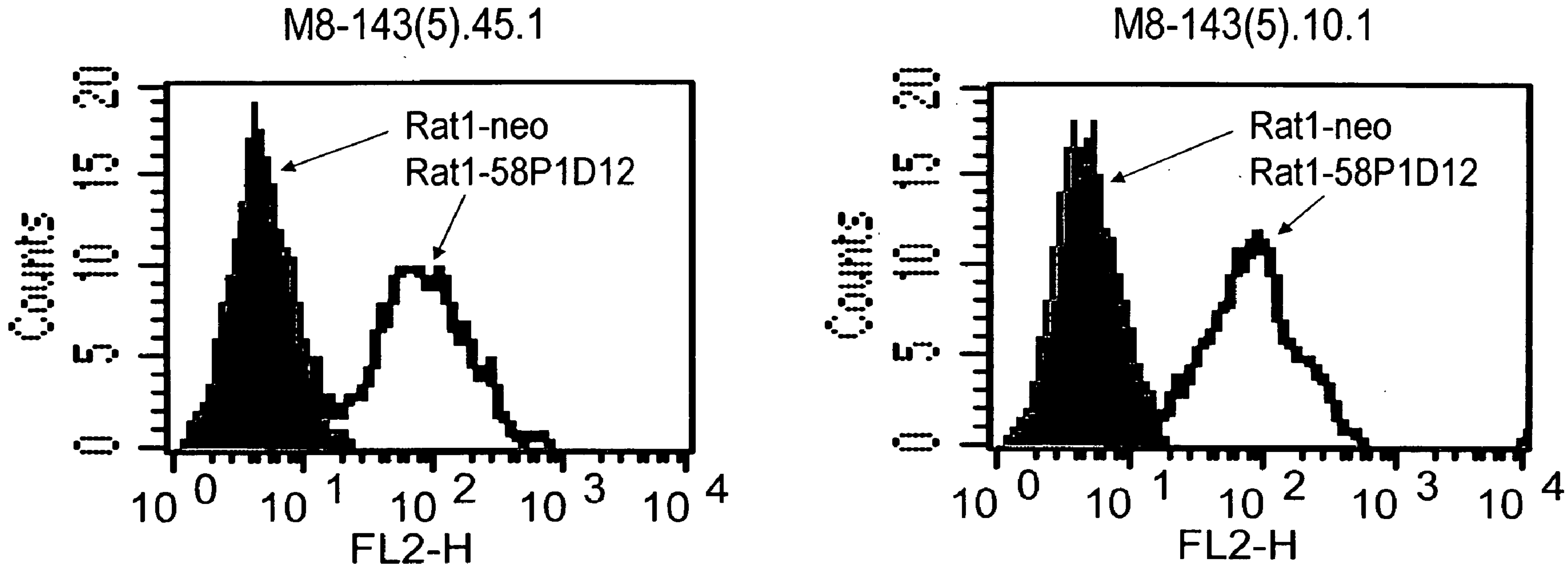


Figure 22: Affinity determination of MAb M8-143(5).10.1 by FACS-based equilibrium binding analysis

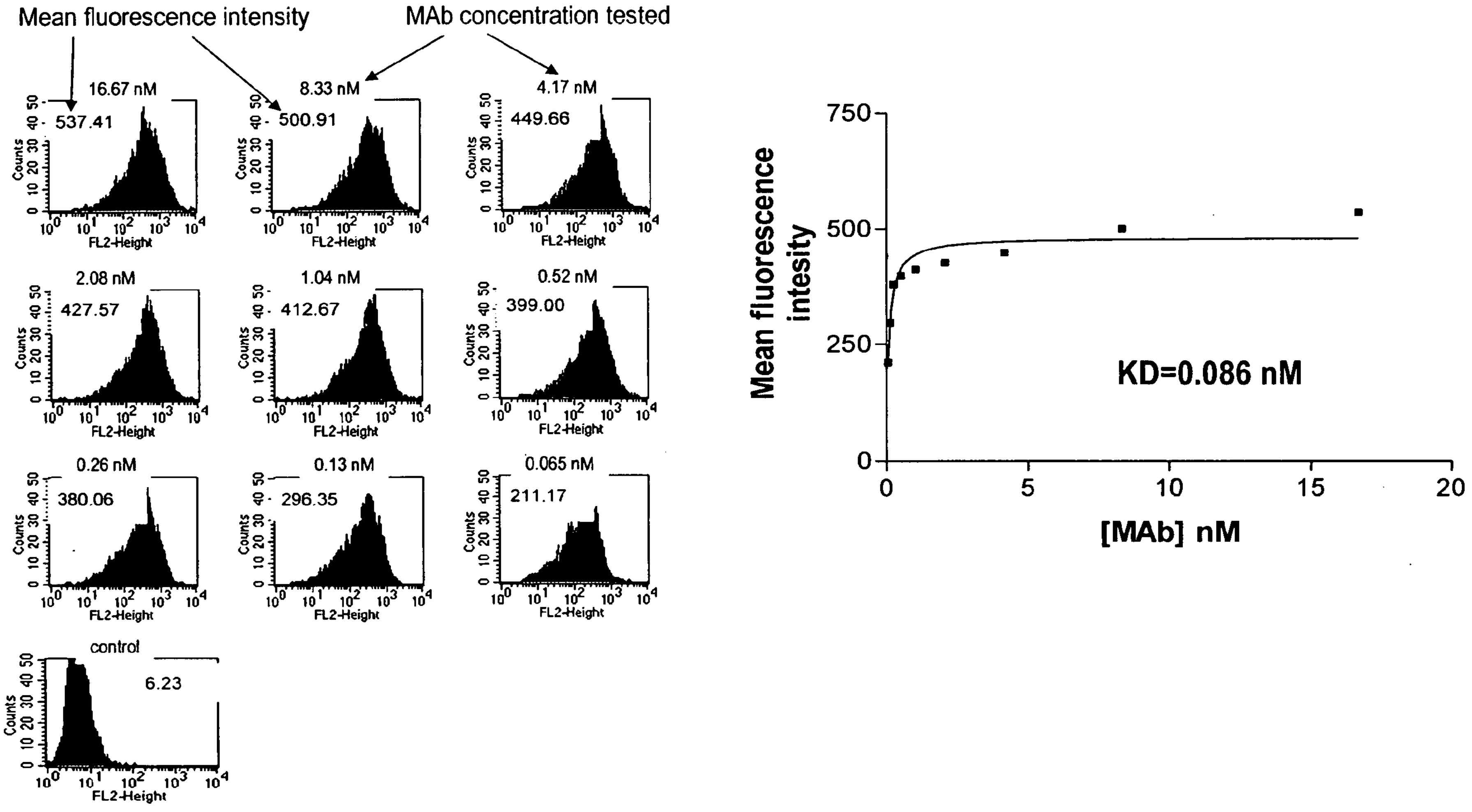
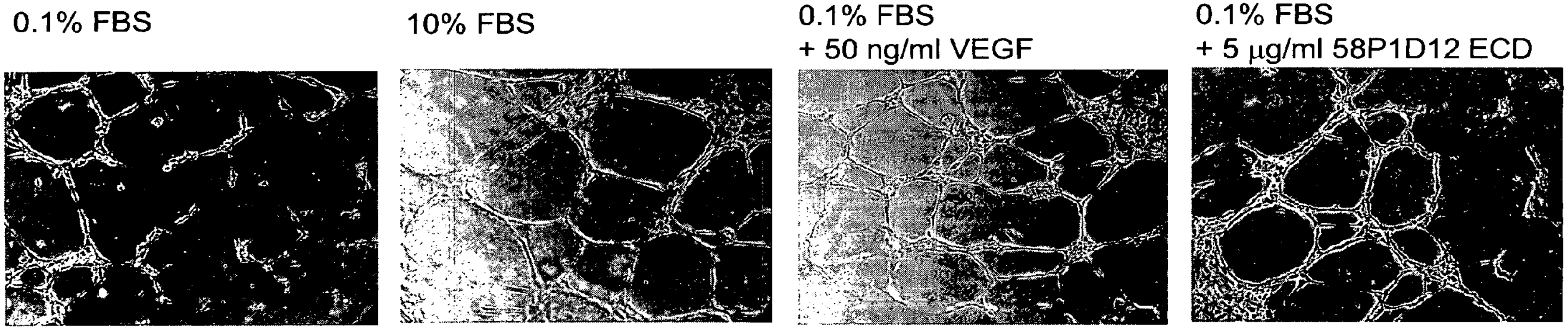


Figure 23: 58P1D12 ECD induces angiogenesis (tube formation) in umbilical vein endothelial cells (HUVEC)



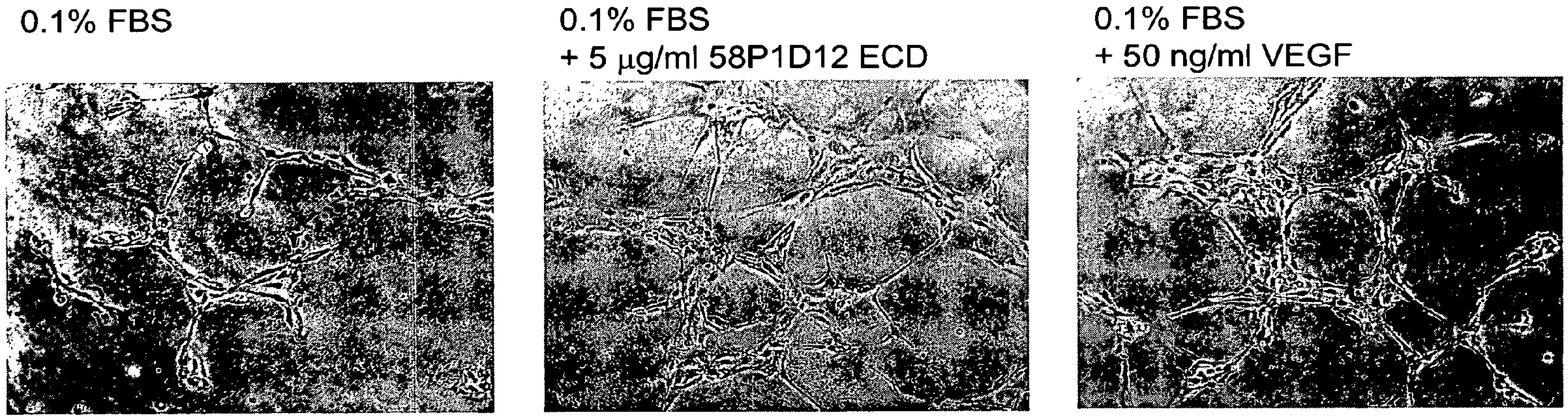
0.1% FBS

10% FBS

0.1% FBS
+ 50 ng/ml VEGF

0.1% FBS
+ 5 μg/ml 58P1D12 ECD

Figure 24: 58P1D12 ECD induces angiogenesis (tube formation) in lung microvascular endothelium (HMVEC)



0.1% FBS

0.1% FBS
+ 5 µg/ml 58P1D12 ECD

0.1% FBS
+ 50 ng/ml VEGF

Figure 25: Monoclonal antibody to 58P1D12 inhibits angiogenesis (tube formation) in umbilical vein endothelial cells (HUVEC)

Treatment	Number of tubes
FBS (0.1%)	7±2
FBS (5%)	114±9
58P1D12 ECD	38±2
58P1D12 ECD + M8-143(5)10	24±3
58P1D12 ECD + M3-47.24	34±2

Figure 26: 58P1D12 ECD enhances survival of HUVEC

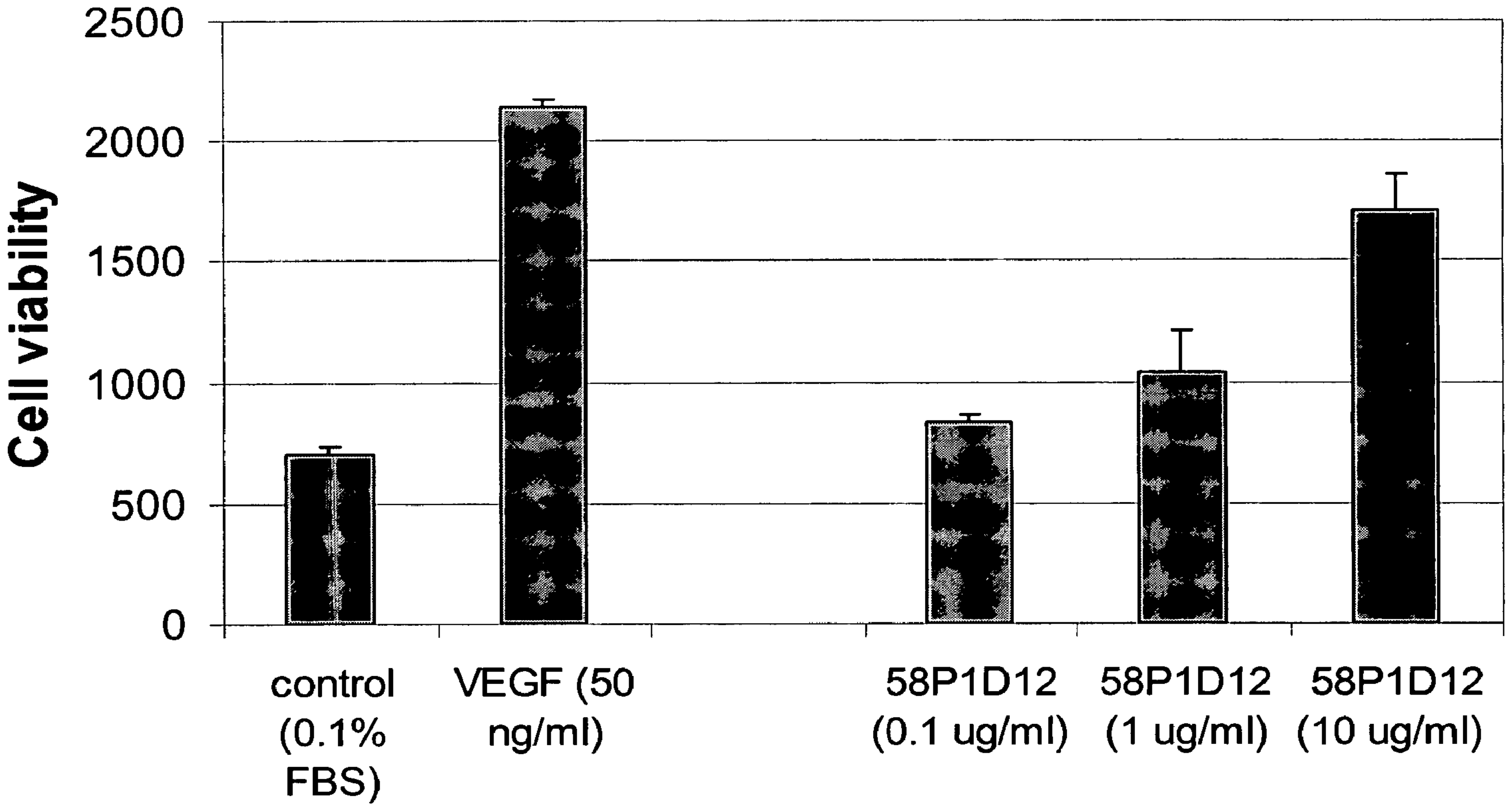


Figure 27: 58P1D12 ECD enhances proliferation of HMVEC

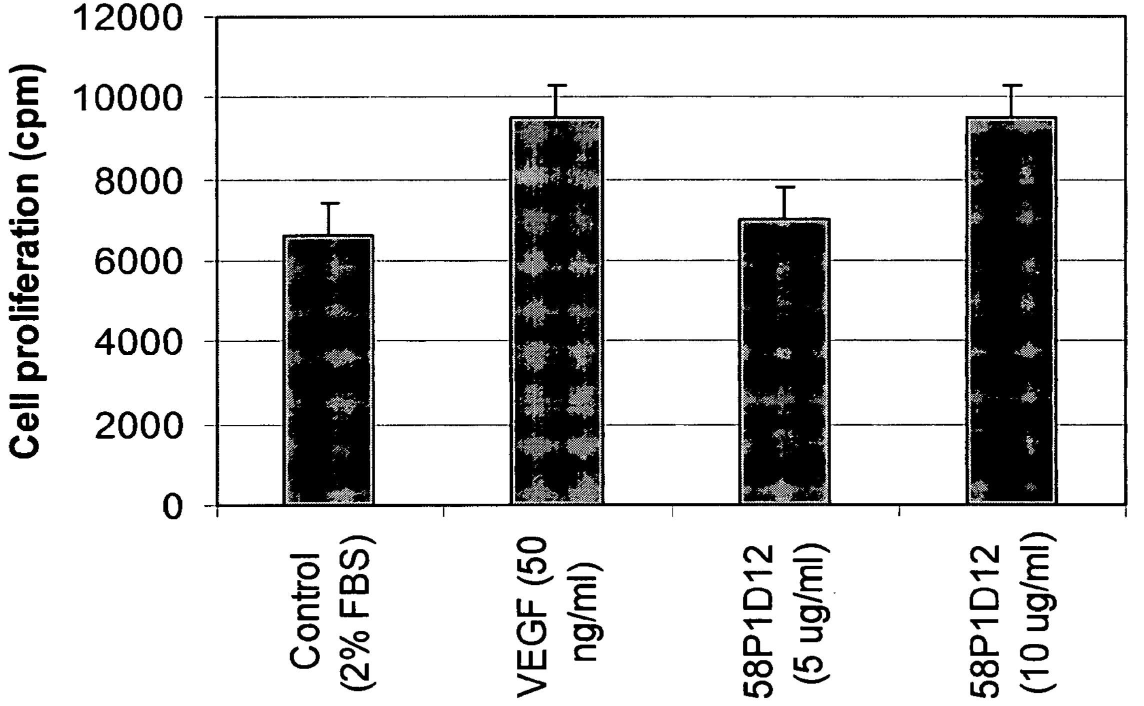
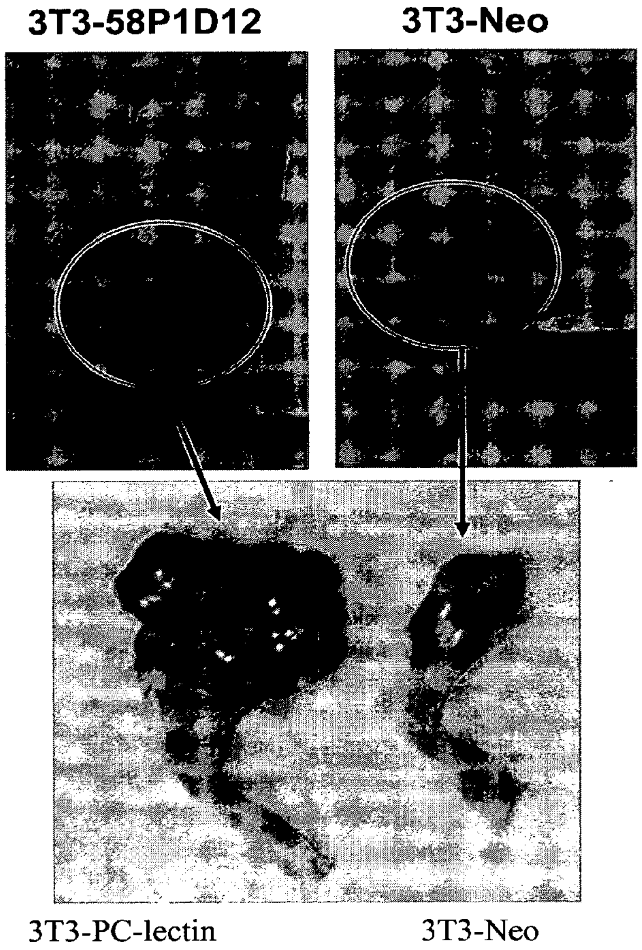


Figure 28: Intratibial tumor growth of 3T3-58P1D12 cells



NUCLEIC ACIDS AND CORRESPONDING PROTEINS ENTITLED 58P1D12 USEFUL IN TREATMENT AND DETECTION OF CANCER

CROSS-REFERENCE TO RELATED APPLICATIONS

This application is a continuation-in-part of pending U.S. utility patent application Ser. No. 10/460,512, filed 11 Jun. 2003 which is a divisional of issued U.S. Pat. No. 6,602,501, filed 11 Aug. 2000 which claims the benefit of priority of U.S. provisional patent application Ser. No. 60/148,935, filed 12 Aug. 1999, the contents of each are hereby incorporated by reference in their entirety.

STATEMENT OF RIGHTS TO INVENTIONS MADE UNDER FEDERALLY SPONSORED RESEARCH

Not applicable.

REFERENCE TO SEQUENCE LISTING SUBMITTED ON COMPACT DISC

This application incorporates by reference in its entirety the Sequence Listing that is provided in duplicate on compact discs (CD-Rs) that accompany the application. Each CD-R contains the following file: 51158-20020.20, having a date of creation of Jul. 13, 2004 and a file size of 116,736 bytes. All the material on the CD-Rs is hereby expressly incorporated by reference into the present application.

REFERENCE TO LIST OF TABLES (APPENDIX) SUBMITTED ON COMPACT DISC

The Compact Disc Appendix, which is a part of the present disclosure, is provided in duplicate on compact discs (CD-Rs). Each contains the following file: 511582002020 Tables, having a date of creation of Jul. 12, 2004 and a file size of 161,792 bytes. All the material on the compact discs is hereby expressly incorporated by reference into the present application.

FIELD OF THE INVENTION

The invention described herein relates to genes and their encoded proteins, termed 58P1D12 and variants thereof, expressed in certain cancers, and to diagnostic and therapeutic methods and compositions useful in the management of cancers that express 58P1D12.

BACKGROUND OF THE INVENTION

Cancer is the second leading cause of human death next to coronary disease. Worldwide, millions of people die from cancer every year. In the United States alone, as reported by the American Cancer Society, cancer causes the death of well over a half-million people annually, with over 1.2 million new cases diagnosed per year. While deaths from heart disease have been declining significantly, those resulting from cancer generally are on the rise. In the early part of the next century, cancer is predicted to become the leading cause of death.

Worldwide, several cancers stand out as the leading killers. In particular, carcinomas of the lung, prostate, breast, colon, pancreas, and ovary represent the primary causes of cancer death. These and virtually all other carcinomas share a common lethal feature. With very few exceptions, metastatic disease from a carcinoma is fatal. Moreover, even for those cancer patients who initially survive their primary cancers, common experience has shown that their lives are dramatically altered. Many cancer patients experience strong anxieties driven by the awareness of the potential for recurrence or treatment failure. Many cancer patients experience physical debilitations following treatment. Furthermore, many cancer patients experience a recurrence.

Worldwide, prostate cancer is the fourth most prevalent cancer in men. In North America and Northern Europe, it is by far the most common cancer in males and is the second leading cause of cancer death in men. In the United States alone, well over 30,000 men die annually of this disease—second only to lung cancer. Despite the magnitude of these figures, there is still no effective treatment for metastatic prostate cancer. Surgical prostatectomy, radiation therapy, hormone ablation therapy, surgical castration and chemotherapy continue to be the main treatment modalities. Unfortunately, these treatments are ineffective for many and are often associated with undesirable consequences.

On the diagnostic front, the lack of a prostate tumor marker that can accurately detect early-stage, localized tumors remains a significant limitation in the diagnosis and management of this disease. Although the serum prostate specific antigen (PSA) assay has been a very useful tool, however its specificity and general utility is widely regarded as lacking in several important respects.

Progress in identifying additional specific markers for prostate cancer has been improved by the generation of prostate cancer xenografts that can recapitulate different stages of the disease in mice. The LAPC (Los Angeles Prostate Cancer) xenografts are prostate cancer xenografts that have survived passage in severe combined immune deficient (SCID) mice and have exhibited the capacity to mimic the transition from androgen dependence to androgen independence (Klein et al., 1997, Nat. Med. 3:402). More recently identified prostate cancer markers include PCTA-1 (Su et al., 1996, Proc. Natl. Acad. Sci. USA 93: 7252), prostate-specific membrane (PSM) antigen (Pinto et al., Clin Cancer Res 1996 Sep. 2 (9): 1445-51), STEAP (Hubert, et al., Proc Natl Acad Sci USA. 1999 Dec. 7; 96(25): 14523-8) and prostate stem cell antigen (PSCA) (Reiter et al., 1998, Proc. Natl. Acad. Sci. USA 95: 1735).

While previously identified markers such as PSA, PSM, PCTA and PSCA have facilitated efforts to diagnose and treat prostate cancer, there is need for the identification of additional markers and therapeutic targets for prostate and related cancers in order to further improve diagnosis and therapy.

Renal cell carcinoma (RCC) accounts for approximately 3 percent of adult malignancies. Once adenomas reach a diameter of 2 to 3 cm, malignant potential exists. In the adult, the two principal malignant renal tumors are renal cell adenocarcinoma and transitional cell carcinoma of the renal pelvis or urethras. The incidence of renal cell adenocarcinoma is estimated at more than 29,000 cases in the United States, and more than 11,600 patients died of this disease in 1998. Transitional cell carcinoma is less frequent, with an incidence of approximately 500 cases per year in the United States.

Surgery has been the primary therapy for renal cell adenocarcinoma for many decades. Until recently, metastatic disease has been refractory to any systemic therapy. With recent developments in systemic therapies, particularly immunotherapies, metastatic renal cell carcinoma may be approached aggressively in appropriate patients with a possibility of durable responses. Nevertheless, there is a remaining need for effective therapies for these patients.

Of all new cases of cancer in the United States, bladder cancer represents approximately 5 percent in men (fifth most common neoplasm) and 3 percent in women (eighth most common neoplasm). The incidence is increasing slowly, concurrent with an increasing older population. In 1998, there was an estimated 54,500 cases, including 39,500 in men and 15,000 in women. The age-adjusted incidence in the United States is 32 per 100,000 for men and eight per 100,000 in women. The historic male/female ratio of 3:1 may be decreasing related to smoking patterns in women. There were an estimated 11,000 deaths from bladder cancer in 1998 (7,800 in men and 3,900 in women). Bladder cancer incidence and mortality strongly increase with age and will be an increasing problem as the population becomes more elderly.

Most bladder cancers recur in the bladder. Bladder cancer is managed with a combination of transurethral resection of the bladder (TUR) and intravesical chemotherapy or immunotherapy. The multifocal and recurrent nature of bladder cancer points out the limitations of TUR. Most muscle-invasive cancers are not cured by TUR alone. Radical cystectomy and urinary diversion is the most effective means to eliminate the cancer but carry an undeniable impact on urinary and sexual function. There continues to be a significant need for treatment modalities that are beneficial for bladder cancer patients.

An estimated 130,200 cases of colorectal cancer occurred in 2000 in the United States, including 93,800 cases of colon cancer and 36,400 of rectal cancer. Colorectal cancers are the third most common cancers in men and women. Incidence rates declined significantly during 1992-1996 (−2.1% per year). Research suggests that these declines have been due to increased screening and polyp removal, preventing progression of polyps to invasive cancers. There were an estimated 56,300 deaths (47,700 from colon cancer, 8,600 from rectal cancer) in 2000, accounting for about 11% of all U.S. cancer deaths.

At present, surgery is the most common form of therapy for colorectal cancer, and for cancers that have not spread, it is frequently curative. Chemotherapy, or chemotherapy plus radiation, is given before or after surgery to most patients whose cancer has deeply perforated the bowel wall or has spread to the lymph nodes. A permanent colostomy (creation of an abdominal opening for elimination of body wastes) is occasionally needed for colon cancer and is infrequently required for rectal cancer. There continues to be a need for effective diagnostic and treatment modalities for colorectal cancer.

There were an estimated 164,100 new cases of lung and bronchial cancer in 2000, accounting for 14% of all U.S. cancer diagnoses. The incidence rate of lung and bronchial cancer is declining significantly in men, from a high of 86.5 per 100,000 in 1984 to 70.0 in 1996. In the 1990s, the rate of increase among women began to slow. In 1996, the incidence rate in women was 42.3 per 100,000.

Lung and bronchial cancer caused an estimated 156,900 deaths in 2000, accounting for 28% of all cancer deaths. During 1992-1996, mortality from lung cancer declined significantly among men (−1.7% per year) while rates for women were still significantly increasing (0.9% per year). Since 1987, more women have died each year of lung cancer than breast cancer, which, for over 40 years, was the major cause of cancer death in women. Decreasing lung cancer incidence and mortality rates most likely resulted from decreased smoking rates over the previous 30 years; however, decreasing smoking patterns among women lag behind those of men. Of concern, although the declines in adult tobacco use have slowed, tobacco use in youth is increasing again.

Treatment options for lung and bronchial cancer are determined by the type and stage of the cancer and include surgery, radiation therapy, and chemotherapy. For many localized cancers, surgery is usually the treatment of choice. Because the disease has usually spread by the time it is discovered, radiation therapy and chemotherapy are often needed in combination with surgery. Chemotherapy alone or combined with radiation is the treatment of choice for small cell lung cancer; on this regimen, a large percentage of patients experience remission, which in some cases is long lasting. There is however, an ongoing need for effective treatment and diagnostic approaches for lung and bronchial cancers.

An estimated 182,800 new invasive cases of breast cancer were expected to occur among women in the United States during 2000. Additionally, about 1,400 new cases of breast cancer were expected to be diagnosed in men in 2000. After increasing about 4% per year in the 1980s, breast cancer incidence rates in women have leveled off in the 1990s to about 110.6 cases per 100,000.

In the U.S. alone, there were an estimated 41,200 deaths (40,800 women, 400 men) in 2000 due to breast cancer. Breast cancer ranks second among cancer deaths in women. According to the most recent data, mortality rates declined significantly during 1992-1996 with the largest decreases in younger women, both white and black. These decreases were probably the result of earlier detection and improved treatment.

Taking into account the medical circumstances and the patient's preferences, treatment of breast cancer may involve lumpectomy (local removal of the tumor) and removal of the lymph nodes under the arm; mastectomy (surgical removal of the breast) and removal of the lymph nodes under the arm; radiation therapy; chemotherapy; or hormone therapy. Often, two or more methods are used in combination. Numerous studies have shown that, for early stage disease, long-term survival rates after lumpectomy plus radiotherapy are similar to survival rates after modified radical mastectomy. Significant advances in reconstruction techniques provide several options for breast reconstruction after mastectomy. Recently, such reconstruction has been done at the same time as the mastectomy.

Local excision of ductal carcinoma in situ (DCIS) with adequate amounts of surrounding normal breast tissue may prevent the local recurrence of the DCIS. Radiation to the breast and/or tamoxifen may reduce the chance of DCIS occurring in the remaining breast tissue. This is important because DCIS, if left untreated, may develop into invasive breast cancer. Nevertheless, there are serious side effects or sequelae to these treatments. There is, therefore, a need for efficacious breast cancer treatments.

There were an estimated 23,100 new cases of ovarian cancer in the United States in 2000. It accounts for 4% of all cancers among women and ranks second among gynecologic cancers. During 1992-1996, ovarian cancer incidence rates were significantly declining. Consequent to ovarian cancer, there were an estimated 14,000 deaths in 2000. Ovarian cancer causes more deaths than any other cancer of the female reproductive system.

Surgery, radiation therapy, and chemotherapy are treatment options for ovarian cancer. Surgery usually includes the removal of one or both ovaries, the fallopian tubes (salpingo-oophorectomy), and the uterus (hysterectomy). In some very early tumors, only the involved ovary will be removed, especially in young women who wish to have children. In advanced disease, an attempt is made to remove all intra-abdominal disease to enhance the effect of chemotherapy.

There continues to be an important need for effective treatment options for ovarian cancer.

There were an estimated 28,300 new cases of pancreatic cancer in the United States in 2000. Over the past 20 years, rates of pancreatic cancer have declined in men. Rates among women have remained approximately constant but may be beginning to decline. Pancreatic cancer caused an estimated 28,200 deaths in 2000 in the United States. Over the past 20 years, there has been a slight but significant decrease in mortality rates among men (about −0.9% per year) while rates have increased slightly among women.

Surgery, radiation therapy, and chemotherapy are treatment options for pancreatic cancer. These treatment options can extend survival and/or relieve symptoms in many patients but are not likely to produce a cure for most. There is a significant need for additional therapeutic and diagnostic options for pancreatic cancer.

SUMMARY OF THE INVENTION

The present invention relates to a gene, designated 58P1D12, that has now been found to be over-expressed in the cancer(s) listed in Table I. Northern blot expression analysis of 58P1D12 gene expression in normal tissues shows a restricted expression pattern in adult tissues. The nucleotide (FIG. **2**) and amino acid (FIG. **2**, and FIG. **3**) sequences of 58P1D12 are provided. The tissue-related profile of 58P1D12 in normal adult tissues, combined with the over-expression observed in the tissues listed in Table I, shows that 58P1D12 is aberrantly over-expressed in at least some cancers, and thus serves as a useful diagnostic, prophylactic, prognostic, and/or therapeutic target for cancers of the tissue(s) such as those listed in Table I.

The invention provides polynucleotides corresponding or complementary to all or part of the 58P1D12 genes, mRNAs, and/or coding sequences, preferably in isolated form, including polynucleotides encoding 58P1D12-related proteins and fragments of 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, or more than 25 contiguous amino acids; at least 30, 35, 40, 45, 50, 55, 60, 65, 70, 80, 85, 90, 95, 100 or more than 100 contiguous amino acids of a 58P1D12-related protein, as well as the peptides/proteins themselves; DNA, RNA, DNA/RNA hybrids, and related molecules, polynucleotides or oligonucleotides complementary or having at least a 90% homology to the 58P1D12 genes or mRNA sequences or parts thereof, and polynucleotides or oligonucleotides that hybridize to the 58P1D12 genes, mRNAs, or to 58P1D12-encoding polynucleotides. Also provided are means for isolating cDNAs and the genes encoding 58P1D12. Recombinant DNA molecules containing 58P1D12 polynucleotides, cells transformed or transduced with such molecules, and host-vector systems for the expression of 58P1D12 gene products are also provided. The invention further provides antibodies that bind to 58P1D12 proteins and polypeptide fragments thereof, including polyclonal and monoclonal antibodies, murine and other mammalian antibodies, chimeric antibodies, humanized and fully human antibodies, and antibodies labeled with a detectable marker or therapeutic agent. In certain embodiments, there is a proviso that the entire nucleic acid sequence of FIG. **2** is not encoded and/or the entire amino acid sequence of FIG. **2** is not prepared. In certain embodiments, the entire nucleic acid sequence of FIG. **2** is encoded and/or the entire amino acid sequence of FIG. **2** is prepared, either of which are in respective human unit dose forms.

The invention further provides methods for detecting the presence and status of 58P1D12 polynucleotides and proteins in various biological samples, as well as methods for identifying cells that express 58P1D12. A typical embodiment of this invention provides methods for monitoring 58P1D12 gene products in a tissue or hematology sample having or suspected of having some form of growth dysregulation such as cancer.

The invention further provides various immunogenic or therapeutic compositions and strategies for treating cancers that express 58P1D12 such as cancers of tissues listed in Table I, including therapies aimed at inhibiting the transcription, translation, processing or function of 58P1D12 as well as cancer vaccines. In one aspect, the invention provides compositions, and methods comprising them, for treating a cancer that expresses 58P1D12 in a human subject wherein the composition comprises a carrier suitable for human use and a human unit dose of one or more than one agent that inhibits the production or function of 58P1D12. Preferably, the carrier is a uniquely human carrier. In another aspect of the invention, the agent is a moiety that is immunoreactive with 58P1D12 protein. Non-limiting examples of such moieties include, but are not limited to, antibodies (such as single chain, monoclonal, polyclonal, humanized, chimeric, or human antibodies), functional equivalents thereof (whether naturally occurring or synthetic), and combinations thereof. The antibodies can be conjugated to a diagnostic or therapeutic moiety. In another aspect, the agent is a small molecule as defined herein.

In another aspect, the agent comprises one or more than one peptide which comprises a cytotoxic T lymphocyte (CTL) epitope that binds an HLA class I molecule in a human to elicit a CTL response to 58P1D12 and/or one or more than one peptide which comprises a helper T lymphocyte (HTL) epitope which binds an HLA class II molecule in a human to elicit an HTL response. The peptides of the invention may be on the same or on one or more separate polypeptide molecules. In a further aspect of the invention, the agent comprises one or more than one nucleic acid molecule that expresses one or more than one of the CTL or HTL response stimulating peptides as described above. In yet another aspect of the invention, the one or more than one nucleic acid molecule may express a moiety that is immunologically reactive with 58P1D12 as described above. The one or more than one nucleic acid molecule may also be, or encodes, a molecule that inhibits production of 58P1D12. Non-limiting examples of such molecules include, but are not limited to, those complementary to a nucleotide sequence essential for production of 58P1D12 (e.g. antisense sequences or molecules that form a triple helix with a nucleotide double helix essential for 58P1D12 production) or a ribozyme effective to lyse 58P1D12 mRNA.

Note that to determine the starting position of any peptide set forth in Tables VIII-XXI and XXII to XLIX (collectively HLA Peptide Tables) respective to its parental protein, e.g., variant 1, variant 2, etc., reference is made to three factors: the particular variant, the length of the peptide in an HLA Peptide Table, and the Search Peptides in Table VII. Generally, a unique Search Peptide is used to obtain HLA peptides of a particular for a particular variant. The position of each Search Peptide relative to its respective parent molecule is listed in Table VII. Accordingly, if a Search Peptide begins at position "X", one must add the value "X−1" to each position in Tables VIII-XXI and XXII to XLIX to obtain the actual position of the HLA peptides in their parental molecule. For example, if a particular Search Peptide begins at position 150 of its parental molecule, one must add 150−1, i.e., 149 to each HLA peptide amino acid position to calculate the position of that amino acid in the parent molecule.

One embodiment of the invention comprises an HLA peptide, that occurs at least twice in Tables VIII-XXI and XXII to XLIX collectively, or an oligonucleotide that encodes the HLA peptide. Another embodiment of the invention comprises an HLA peptide that occurs at least once in Tables VIII-XXI and at least once in tables XXII to XLIX, or an oligonucleotide that encodes the HLA peptide.

Another embodiment of the invention is antibody epitopes, which comprise a peptide regions, or an oligonucleotide encoding the peptide region, that has one two, three, four, or five of the following characteristics:

i) a peptide region of at least 5 amino acids of a particular peptide of FIG. **3**, in any whole number increment up to the full length of that protein in FIG. **3**, that includes an amino acid position having a value equal to or greater than 0.5, 0.6, 0.7, 0.8, 0.9, or having a value equal to 1.0, in the Hydrophilicity profile of FIG. **5**;

ii) a peptide region of at least 5 amino acids of a particular peptide of FIG. **3**, in any whole number increment up to the full length of that protein in FIG. **3**, that includes an amino acid position having a value equal to or less than 0.5, 0.4, 0.3, 0.2, 0.1, or having a value equal to 0.0, in the Hydropathicity profile of FIG. **6**;

iii) a peptide region of at least 5 amino acids of a particular peptide of FIG. **3**, in any whole number increment up to the full length of that protein in FIG. **3**, that includes an amino acid position having a value equal to or greater than 0.5, 0.6, 0.7, 0.8, 0.9, or having a value equal to 1.0, in the Percent Accessible Residues profile of FIG. **7**;

iv) a peptide region of at least 5 amino acids of a particular peptide of FIG. **3**, in any whole number increment up to the full length of that protein in FIG. **3**, that includes an amino acid position having a value equal to or greater than 0.5, 0.6, 0.7, 0.8, 0.9, or having a value equal to 1.0, in the Average Flexibility profile of FIG. **8**; or

v) a peptide region of at least 5 amino acids of a particular peptide of FIG. **3**, in any whole number increment up to the full length of that protein in FIG. **3**, that includes an amino acid position having a value equal to or greater than 0.5, 0.6, 0.7, 0.8, 0.9, or having a value equal to 1.0, in the Beta-turn profile of FIG. **9**[!!! The Figure descriptions need to be revised when using this as a template]

BRIEF DESCRIPTION OF THE FIGURES

FIG. **1**. The 58P1D12 SSH sequence of 427 nucleotides.

FIG. **2**A) The cDNA and amino acid sequence of 58P1D12 variant 1 (also called "58P1D12 v.1" or "58P1D12 variant 1") is shown in FIG. **2**A. The start methionine is underlined. The open reading frame extends from nucleic acid 380-1201 including the stop codon.

FIG. **2**B) The cDNA and amino acid sequence of 58P1D12 variant 2 (also called "58P1D12 v.2") is shown in FIG. **2**B. The codon for the start methionine is underlined. The open reading frame extends from nucleic acid 388-1086 including the stop codon.

FIG. **2**C) The cDNA and amino acid sequence of 58P1D12 variant 3 (also called "58P1D12 v.3") is shown in FIG. **2**C. The codon for the start methionine is underlined. The open reading frame extends from nucleic acid 206-904 including the stop codon.

FIG. **2**D) The cDNA and amino acid sequence of 58P1D12 variant 4 (also called "58P1D12 v.4") is shown in FIG. **2**D. The codon for the start methionine is underlined. The open reading frame extends from nucleic acid 206-916 including the stop codon.

FIG. **2**E) The cDNA and amino acid sequence of 58P1D12 variant 5 (also called "58P1D12 v.5") is shown in FIG. **2**E. The codon for the start methionine is underlined. The open reading frame extends from nucleic acid 106-816 including the stop codon.

Variants 58P1D12 v.6 through v.15 are variants with single nucleotide difference from 58P1D12 v.1. Though these SNP variants are shown separately, they can also occur in any combinations and in any of the transcript variants listed above in FIGS. **2**A-**2**F.

FIG. **3**A) The amino acid sequence of 58P1D12 v.1 is shown in FIG. **3**A; it has 273 amino acids.

FIG. **3**B) The amino acid sequence of 58P1D12 v.2 is shown in FIG. **3**B; it has 232 amino acids.

FIG. **3**C) The amino acid sequence of 58P1D12 v.3 is shown in FIG. **3**C; it has 236 amino acids.

As used herein, a reference to 58P1D12 includes all variants thereof, including those shown in FIGS. **2**, **3**, **10**, **11**, and **12** unless the context clearly indicates otherwise.

FIG. **4**. Intentionally Omitted.

FIGS. **5**(*a*)-(*c*). Hydrophilicity amino acid profile of 58P1D12 v.1-v.3 determined by computer algorithm sequence analysis using the method of Hopp and Woods (Hopp T. P., Woods K. R., 1981. Proc. Natl. Acad. Sci. U.S.A. 78:3824-3828) accessed on the Protscale website located on the World Wide Web at (expasy.ch/cgi-bin/protscale.pl) through the ExPasy molecular biology server.

FIGS. **6**(*a*)-(*c*). Hydropathicity amino acid profile of 58P1D12 v.1-v.9 determined by computer algorithm sequence analysis using the method of Kyte and Doolittle (Kyte J., Doolittle R. F., 1982. J. Mol. Biol. 157:105-132) accessed on the ProtScale website located on the World Wide Web at (.expasy.ch/cgi-bin/protscale.pl) through the ExPasy molecular biology server.

FIGS. **7**(*a*)-(*c*). Percent accessible residues amino acid profile of 58P1D12 v.1-v.9 determined by computer algorithm sequence analysis using the method of Janin (Janin J., 1979 Nature 277:491-492) accessed on the ProtScale website located on the World Wide Web at (.expasy.ch/cgi-bin/protscale.pl) through the ExPasy molecular biology server.

FIGS. **8**(*a*)-(*c*). Average flexibility amino acid profile of 58P1D12 v.1-v.9 determined by computer algorithm sequence analysis using the method of Bhaskaran and Ponnuswamy (Bhaskaran R., and Ponnuswamy P. K., 1988. Int. J. Pept. Protein Res. 32:242-255) accessed on the ProtScale website located on the World Wide Web at (.expasy.ch/cgi-bin/protscale.pl) through the ExPasy molecular biology server.

FIG. **9**(*a*)-(*c*). Beta-turn amino acid profile of 58P1D12 v.1-v.9 determined by computer algorithm sequence analysis using the method of Deleage and Roux (Deleage, G., Roux B. 1987 Protein Engineering 1:289-294) accessed on the ProtScale website located on the World Wide Web at (.expasy.ch/cgi-bin/protscale.pl) through the ExPasy molecular biology server.

FIG. **10**. Structures of transcript variants of 58P1D12. Variant 58P1D12 v.2 through v.5 were transcript variants of 58P1D12 v.1. Variants 58P1D12 v.2 and v.3, each had two different exons in the 5' end, shared five exons in the 3' end with v.1. Variants v.4 and v.5 were similar to v.3, with deletions of one or two exons from v.3. All the exons in different transcript variants were aligned in relative order as on the genome. Poly A tails and SNP are not shown here. Numbers in "( )" underneath the boxes correspond to those of 58P1D12 v.1. Lengths of introns and exons are not proportional.

FIG. **11**. Schematic alignment of protein variants of 58P1D12. Protein variants corresponded to nucleotide variants. Nucleotide variants 58P1D12 v.3 coded for the same protein as variant v.2, and variant v.5 coded for the same protein as v.4. SNP variants v.6 through v.15 coded for the same protein as v.1. Nucleotide variants 58P1D12 v.6 through v.15 each had one alternative allele of a SNP as compared with variant v.1, as shown in FIG. **12**. Proteins of v.2 and v.3 were portion of v.1 protein, while variants v.4 and v.5 had a portion that was different from any port of v.1. Numbers underneath the box correspond to 58P1D12 v.1. Black boxes represent the same sequence as 58P1D12 v.1.

FIG. **12**. Schematic alignment of SNP variants of 58P1D12. Variants 58P1D12 v.6 through v.12 were variants with single nucleotide differences as compared to variant v.1. Though these SNP variants were shown separately, they could also occur in any combinations (called haplotype), and in any transcript variants that contained the base pairs, such as v.2 shown in FIG. **10**. Numbers correspond to those of 58P1D12 v.1. Black box shows the same sequence as 58P1D12 v.1. SNPs are indicated above the box.

FIG. **13**. Secondary structure and transmembrane domains prediction for 58P1D12 protein variants.

FIG. **13**A-**13**C: The secondary structures of 58P1D12 protein variants 1 (SEQ ID NO: 15), variants 2 (SEQ ID NO: 16), variant 3 (SEQ ID NO: 17), respectively, were predicted using the HNN—Hierarchical Neural Network method (NPS@: Network Protein Sequence Analysis TIBS 2000 March Vol. 25, No 3 [291]:147-150 Combet C., Blanchet C., Geourjon C. and Deléage G., http://pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_nn.html), accessed from the ExPasy molecular biology server located on the World Wide Web at (.expasy.ch/tools/). This method predicts the presence and location of alpha helices, extended strands, and random coils from the primary protein sequence. The percent of the protein variant in a given secondary structure is also listed.

FIG. **13**D, FIG. **13**F, FIG. **13**H: Schematic representation of the probability of existence of transmembrane regions of 58P1D12 protein variants 1, 2, 3, respectively, based on the TMpred algorithm of Hofmann and Stoffel which utilizes TMBASE (K. Hofmann, W. Stoffel. TMBASE—A database of membrane spanning protein segments Biol. Chem. Hoppe-Seyler 374:166, 1993).

FIG. **13**E, FIG. **13**G, FIG. **13**I: Schematic representation of the probability of the existence of transmembrane regions of 58P1D12 variants 1, 2, 3, respectively, based on the TMHMM algorithm of Sonnhammer, von Heijne, and Krogh (Erik L. L. Sonnhammer, Gunnar von Heijne, and Anders Krogh: A hidden Markov model for predicting transmembrane helices in protein sequences. In Proc. of Sixth Int. Conf. on Intelligent Systems for Molecular Biology, p 175-182 Ed J. Glasgow, T. Littlejohn, F. Major, R. Lathrop, D. Sankoff, and C. Sensen Menlo Park, Calif.: AAAI Press, 1998). The TMpred and TMHMM algorithms are accessed from the ExPasy molecular biology server located on the World Wide Web at (.expasy.ch/tools/).

FIG. **14**. 58P1D12 Expression in Normal and Patient Cancer Tissues. First strand cDNA was prepared from normal tissues (bladder, brain, heart, kidney, liver, lung, prostate, spleen, skeletal muscle, testis, pancreas, colon and stomach), and from pools of patient cancer specimens (prostate cancer pool, bladder cancer pool, lung cancer pool, ovary cancer pool, prostate cancer xenograft pool, bone and melanoma cancer pool, and cervical cancer pool). Normalization was performed by PCR using primers to actin. Semi-quantitative PCR, using primers to 58P1D12, was performed at 30 cycles of amplification. Results expression of 58P1D12 predominantly in testis amongst the normal tissues tested. 58P1D12 is also expressed in prostate cancer, bladder cancer, lung cancer, ovary cancer, prostate cancer xenograft, bone/melanoma cancer pool, and cervical cancer.

FIG. **15**. Expression of 58P1D12 in normal tissues. Two multiple tissue northern blots (Clontech) both with 2 ug of mRNA/lane were probed with the 58P1D12 sequence. Size standards in kilobases (kb) are indicated on the side. Results show predominant expression of 58P1D12 transcript in normal testis. A significantly weaker signal is detected in spleen, lung, kidney and pancreas.

FIG. **16**. 58P1D12 Expression in Prostate Cancer Xenograft tissues. RNA was extracted from LAPC prostate cancer xenografts, LAPC-4AD, LAPC-4AI, LAPC-9AD and LAPC-9AI, as well as from normal prostate. Northern blots with 10 ug of total RNA were probed with the 58P1D12 sequence. Size standards in kilobases are on the side. Results show expression of 58P1D12 in LAPC-9AD but not in the other tissues tested.

FIG. **17**. 58P1D12 Expression in Normal and Patient Cancer Specimens. RNA was extracted from a pool of 3 ovary tumor patient specimens, normal prostate (NP), normal bladder (NB), normal kidney (NK), and normal colon (NC). Northern blots with 10 ug of total RNA were probed with the 58P1D12 sequence. Size standards in kilobases are on the side. Results show expression of 58P1D12 in ovary tumor patient specimens but not in the normal tissues tested.

FIG. **18**. 58P1D12 Expression in Ovarian Cancer Patient Specimens. RNA was extracted from normal ovary, LAPC-9AD prostate cancer xenograft, and a panel of ovary tumor patient specimens (Ovary T). Northern blots with 10 ug of total RNA were probed with the 58P1D12 sequence. Size standards in kilobases are on the side. Results show expression of 58P1D12 in most of the patient specimens tested as well as in the prostate xenograft.

FIG. **19**. 58P1D12 Expression in a Large Panel of Ovarian Cancer Patient Specimens. First strand cDNA was prepared from a panel of ovary patient cancer specimens as well as two different normal ovary tissues. Normalization was performed by PCR using primers to actin. Semi-quantitative PCR, using primers to 58P1D12, was performed at 26 and 30 cycles of amplification. Samples were run on an agarose gel, and PCR products were quantitated using the AlphaImager software. Expression was recorded as negative, weak, medium or strong. Results show expression of 58P1D12 in the majority of patient cancer specimens tested (75%), but not in normal ovary.

FIG. **20**. Specific cell surface staining of Rat1-58P1D12 cells with serum from mice immunized with recombinant 58P1D12 protein. Rat1-58P1D12 and control Rat1-neo cells (~100,000 cells) were incubated with various dilutions of serum obtained from 2 different mice immunized with recombinant Tag5-58P1D12 protein. Cells were washed and then incubated with goat anti-mouse IgG-PE secondary Ab (1:100 dilution in PBS+2% FBS) on ice for 1 hour. Cells were washed again, fixed in 2% paraformaldehyde in PBS, and subjected to FACS analysis. The Rat1-58P1D12 cells, but not the Rat1-neo cells, show a shift in fluorescent intensity over a large dilution range demonstrating specific binding of anti-58P1D12 Abs present in the serum from the 2 mice.

FIG. **21**. Specific cell surface staining of Rat1-58P1D12 cells with monoclonal antibody hybridoma supernatants. Rat1-58P1D12 and control Rat1-neo cells (~100,000 cells) were incubated with supernatants derived from hybridomas M8-143(5).10.1 and M8-143(5).45.1. on ice for 1 hour. Cells were washed and then incubated with goat anti-mouse IgG-PE secondary Ab (1:100 dilution in PBS+2% FBS) on ice for 1 hour. Cells were washed again, fixed in 2% paraformaldehyde in PBS, and subjected to FACS analysis. The Rat1-58P1D12 cells, but not the Rat1-neo cells, show a shift in fluorescent intensity demonstrating specific binding of the MAb to 58P1D12 protein on the cell surface.

FIG. **22**. Affinity determination of MAb M8-143(5).10.1 by FACS-based equilibrium binding analysis. 3T3-58P1D12 cells were incubated at 4 C for 2 hours with various concentrations of MAb M8-143(5).10.1 in PBS+2% FBS+0.04% NaN3. Cells were washed, and then incubated at 4 C for 1 hour with a 1:100 dilution of goat anti-mouse IgG-PE secondary Ab. After washing, cells were fixed in 2% paraformaldehyde in PBS and subjected to FACS analysis. Shown on the left are the histograms of cells stained with various concentrations of MAb. On the right is a graphical representation of the data plotting MAb concentration versus the mean fluorescent intensity of the stained populations. Non-linear regression of the binding data was carried out with the GraphPad Prism software (Version 3.00 for Windows, GraphPad Software, San Diego Calif. USA) to determine the dissociation constant (KD). The analysis shows that MAb M8-143(5).10.1 is a high affinity MAb with an apparent KD of 0.086 nM on 3T3-58P1D12 cells.

FIG. **23**. 58P1D12 induces angiogenesis (tube formation) in umbilical vein endothelial cells (HUVEC). Primary umbilical vein endothelial cells were cultured in 10% FBS, then removed with EDTA and mixed with media containing either 10% FBS, 0.1% FBS, 0.1% FBS+50 ng/ml VEGF or 0.1% FBS+5 μg/ml 58P1D12 ECD protein. The mixtures were plated onto a 200 μl layer of Matrigel® and then incubated at 37° C. for 16 hours. The plates were then analyzed for the formation of tubes by microscopy. Images of representative fields are shown. The data show that HUVEC cells form tubes in 10% FBS but not in 0.1% FBS, thereby allowing the assessment of 58P1D12-induced angiogenesis in low serum conditions. At 5 μg/ml, 58P1D12 ECD exhibited the capacity to induce the formation of endothelial tubes similarly as the positive control VEGF.

FIG. **24**. 58P1D12 induces angiogenesis (tube formation) in lung microvascular endothelial cells (HMVEC). Primary lung endothelial cells were cultured in 10% FBS, then removed with EDTA and mixed with media containing either 0.1% FBS, 0.1% FBS+50 ng/ml VEGF or 0.1% FBS+5 μg/ml 58P1D12 ECD protein. The mixtures were plated onto a 200 μl layer of Matrigel® and then incubated at 37° C. for 16 hours. The plates were then analyzed for the formation of tubes by microscopy. Images of representative fields are shown. The data show that HMVEC cells do not form tubes in 0.1% FBS, thereby allowing the assessment of 58P1D12-induced angiogenesis in low serum conditions. At 5 μg/ml, 58P1D12 ECD exhibited the capacity to induce the formation of endothelial tubes similarly as the positive control VEGF.

FIG. **25**. Monoclonal antibody to 58P1D12 inhibits angiogenesis (tube formation) in umbilical vein endothelial cells (HUVEC). Primary umbilical vein endothelial cells were cultured in 10% FBS, then removed with EDTA and mixed with media containing either 0.1% FBS, 5% FBS, 0.1% FBS+3 μg/ml 58P1D12 ECD protein, 0.1% FBS+3 μg/ml 58P1D12 ECD protein+30 μg/ml MAb M8-143(5)10 or 0.1% FBS+3 μg/ml 58P1D12 ECD protein+30 μg/ml MAb M3-47.24. The mixtures were plated onto a 200 μl layer of Matrigel® and then incubated at 37° C. for 16 hours. The plates were then analyzed for the formation of tubes by microscopy. The total tube counts are shown. The data show that HUVEC cells form tubes in 5% FBS but not in 0.1% FBS, thereby allowing the assessment of 58P1D12-induced angiogenesis in low serum conditions. At 3 μg/ml, 58P1D12 ECD exhibited the capacity to induce the formation of endothelial tubes, and that incubation with MAb M8-143(5)10, but not MAb M3-47.24, inhibited the tube formation by 45%.

FIG. **26**. 58P1D12 ECD enhances survival of HUVEC. Primary umbilical vein endothelial cells were cultured in 10% FBS. Cells were trypsinized and 3000 cells per well were plated onto 96-well dishes for 72 hours. Cell viability was assayed by Alamar blue staining for 3 hours, followed by fluorescence measurement. The assay was performed in triplicate with standard deviations represented. In low serum conditions (0.1% FBS), cell viability was low, but was enhanced by increasing concentrations of 58P1D12 ECD. For positive control, the cells were treated with 50 ng/ml VEGF. Cell viability at 10 μg/ml 58P1D12 ECD was similar to that observed using 50 ng/ml VEGF (a known survival signal for endothelium).

FIG. **27**. 58P1D12 ECD enhances proliferation of HMVEC. Primary lung microvascular endothelial cells (HMVEC) were cultured in 10% FBS. Cells were trypsinized and 3000 cells per well were plated onto 96-well dishes in 2% FBS. The cells were either left untreated, or were incubated with 50 ng/ml VEGF, 5 μg/ml 58P1D12 ECD or 10 μg/ml 58P1D12 ECD. After 48 hours, the cells were incubated with 3H-thymidine and allowed to grow overnight. Incorporation of 3H-thymidine into cell DNA was assayed by standard protocols. The assay was performed in triplicate with standard deviations represented. The data indicate that basal proliferation could be measured in 2% FBS, where increased proliferation was observed with 58P1D12 ECD. Similar levels of enhanced proliferation were seen with 10 μg/ml 58P1D12 ECD compared with the VEGF positive control.

FIG. **28**. Intratibial tumor growth of 3T3-58P1D12 cells. 3T3 fibroblasts were infected with empty vector control amphotropic virus (Neo) or virus containing the 58P1D12 gene. Stable transfectants were selected in G418 (geneticin) and then grown in 10% FBS. For each cell type, 3T3-Neo and 3T3-58P1D12, 1×10^6 cells were mixed with Matrigel® and injected into the flanks of 7 male SCID mice. After 31 days, the mice were sacrificed, and the tumors were extracted with the hind quarters. The data are representative of the 7 mice treated with the two cell types. The data indicate that the expression of 58P1D12 in the non-tumor forming 3T3 line has increased tumor growth potential in vivo.

DETAILED DESCRIPTION OF THE INVENTION

Outline of Sections

- I.) Definitions
- II.) 58P1D12 Polynucleotides
  - II.A.) Uses of 58P1D12 Polynucleotides
    - II.A.1.) Monitoring of Genetic Abnormalities
    - II.A.2.) Antisense Embodiments
    - II.A.3.) Primers and Primer Pairs
    - II.A.4.) Isolation of 58P1D12-Encoding Nucleic Acid Molecules
    - II.A.5.) Recombinant Nucleic Acid Molecules and Host-Vector Systems
- III.) 58P1D12-related Proteins
  - III.A.) Motif-bearing Protein Embodiments
  - III.B.) Expression of 58P1D12-related Proteins
  - III.C.) Modifications of 58P1D12-related Proteins
  - III.D.) Uses of 58P1D12-related Proteins
- IV.) 58P1D12 Antibodies
- V.) 58P1D12 Cellular Immune Responses
- VI.) 58P1D12 Transgenic Animals
- VII.) Methods for the Detection of 58P1D12

I.) Definitions:

Unless otherwise defined, all terms of art, notations and other scientific terms or terminology used herein are intended to have the meanings commonly understood by those of skill in the art to which this invention pertains. In some cases, terms with commonly understood meanings are defined herein for clarity and/or for ready reference, and the inclusion of such definitions herein should not necessarily be construed to represent a substantial difference over what is generally understood in the art. Many of the techniques and procedures described or referenced herein are well understood and commonly employed using conventional methodology by those skilled in the art, such as, for example, the widely utilized molecular cloning methodologies described in Sambrook et al., Molecular Cloning: A Laboratory Manual 2nd. edition (1989) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. As appropriate, procedures involving the use of commercially available kits and reagents are generally carried out in accordance with manufacturer defined protocols and/or parameters unless otherwise noted.

The terms "advanced prostate cancer", "locally advanced prostate cancer", "advanced disease" and "locally advanced disease" mean prostate cancers that have extended through the prostate capsule, and are meant to include stage C disease under the American Urological Association (AUA) system, stage C1-C2 disease under the Whitmore-Jewett system, and stage T3-T4 and N+ disease under the TNM (tumor, node, metastasis) system. In general, surgery is not recommended for patients with locally advanced disease, and these patients have substantially less favorable outcomes compared to patients having clinically localized (organ-confined) prostate cancer. Locally advanced disease is clinically identified by palpable evidence of induration beyond the lateral border of the prostate, or asymmetry or induration above the prostate base. Locally advanced prostate cancer is presently diagnosed pathologically following radical prostatectomy if the tumor invades or penetrates the prostatic capsule, extends into the surgical margin, or invades the seminal vesicles.

"Altering the native glycosylation pattern" is intended for purposes herein to mean deleting one or more carbohydrate moieties found in native sequence 58P1D12 (either by removing the underlying glycosylation site or by deleting the glycosylation by chemical and/or enzymatic means), and/or adding one or more glycosylation sites that are not present in the native sequence 58P1D12. In addition, the phrase includes qualitative changes in the glycosylation of the native proteins, involving a change in the nature and proportions of the various carbohydrate moieties present.

The term "analog" refers to a molecule which is structurally similar or shares similar or corresponding attributes with another molecule (e.g. a 58P1D12-related protein). For example, an analog of a 58P1D12 protein can be specifically bound by an antibody or T cell that specifically binds to 58P1D12.

The term "antibody" is used in the broadest sense. Therefore, an "antibody" can be naturally occurring or man-made such as monoclonal antibodies produced by conventional hybridoma technology. Anti-58P1D12 antibodies comprise monoclonal and polyclonal antibodies as well as fragments containing the antigen-binding domain and/or one or more complementarity determining regions of these antibodies.

An "antibody fragment" is defined as at least a portion of the variable region of the immunoglobulin molecule that binds to its target, i.e., the antigen-binding region. In one embodiment it specifically covers single anti-58P1D12 antibodies and clones thereof (including agonist, antagonist and neutralizing antibodies) and anti-58P1D12 antibody compositions with polyepitopic specificity.

The term "codon optimized sequences" refers to nucleotide sequences that have been optimized for a particular host species by replacing any codons having a usage frequency of less than about 20%. Nucleotide sequences that have been optimized for expression in a given host species by elimination of spurious polyadenylation sequences, elimination of exon/intron splicing signals, elimination of transposon-like repeats and/or optimization of GC content in addition to codon optimization are referred to herein as an "expression enhanced sequences."

A "combinatorial library" is a collection of diverse chemical compounds generated by either chemical synthesis or biological synthesis by combining a number of chemical "building blocks" such as reagents. For example, a linear combinatorial chemical library, such as a polypeptide (e.g., mutein) library, is formed by combining a set of chemical building blocks called amino acids in every possible way for a given compound length (i.e., the number of amino acids in a polypeptide compound). Numerous chemical compounds are synthesized through such combinatorial mixing of chemical building blocks (Gallop et al., J. Med. Chem. 37(9):1233-1251 (1994)).

Preparation and screening of combinatorial libraries is well known to those of skill in the art. Such combinatorial chemical libraries include, but are not limited to, peptide libraries (see, e.g., U.S. Pat. No. 5,010,175, Furka, Pept. Prot. Res. 37:487-493 (1991), Houghton et al., Nature, 354:84-88 (1991)), peptoids (PCT Publication No WO 91/19735), encoded peptides (PCT Publication WO 93/20242), random bio-oligomers (PCT Publication WO 92/00091), benzodiazepines (U.S. Pat. No. 5,288,514), diversomers such as hydantoins, benzodiazepines and dipeptides (Hobbs et al., Proc.

Nat. Acad. Sci. USA 90:6909-6913 (1993)), vinylogous polypeptides (Hagihara et al., J. Amer. Chem. Soc. 114:6568 (1992)), nonpeptidal peptidomimetics with a Beta-D-Glucose scaffolding (Hirschmann et al., J. Amer. Chem. Soc. 114:9217-9218 (1992)), analogous organic syntheses of small compound libraries (Chen et al., J. Amer. Chem. Soc. 116:2661 (1994)), oligocarbamates (Cho, et al., Science 261: 1303 (1993)), and/or peptidyl phosphonates (Campbell et al., J. Org. Chem. 59:658 (1994)). See, generally, Gordon et al., J. Med. Chem. 37:1385 (1994), nucleic acid libraries (see, e.g., Stratagene, Corp.), peptide nucleic acid libraries (see, e.g., U.S. Pat. No. 5,539,083), antibody libraries (see, e.g., Vaughn et al., Nature Biotechnology 14(3): 309-314 (1996), and PCT/US96/10287), carbohydrate libraries (see, e.g., Liang et al., Science 274:1520-1522 (1996), and U.S. Pat. No. 5,593,853), and small organic molecule libraries (see, e.g., benzodiazepines, Baum, C&EN, January 18, page 33 (1993); isoprenoids, U.S. Pat. No. 5,569,588; thiazolidinones and metathiazanones, U.S. Pat. No. 5,549,974; pyrrolidines, U.S. Pat. Nos. 5,525,735 and 5,519,134; morpholino compounds, U.S. Pat. No. 5,506,337; benzodiazepines, U.S. Pat. No. 5,288,514; and the like).

Devices for the preparation of combinatorial libraries are commercially available (see, e.g., 357 NIPS, 390 NIPS, Advanced Chem Tech, Louisville Ky.; Symphony, Rainin, Woburn, Mass.; 433A, Applied Biosystems, Foster City, Calif.; 9050, Plus, Millipore, Bedford, NIA). A number of well-known robotic systems have also been developed for solution phase chemistries. These systems include automated workstations such as the automated synthesis apparatus developed by Takeda Chemical Industries, LTD. (Osaka, Japan) and many robotic systems utilizing robotic arms (Zymate H, Zymark Corporation, Hopkinton, Mass.; Orca, Hewlett-Packard, Palo Alto, Calif.), which mimic the manual synthetic operations performed by a chemist. Any of the above devices are suitable for use with the present invention. The nature and implementation of modifications to these devices (if any) so that they can operate as discussed herein will be apparent to persons skilled in the relevant art. In addition, numerous combinatorial libraries are themselves commercially available (see, e.g., ComGenex, Princeton, N.J.; Asinex, Moscow, RU; Tripos, Inc., St. Louis, Mo.; ChemStar, Ltd, Moscow, RU; 3D Pharmaceuticals, Exton, Pa.; Martek Biosciences, Columbia, Md.; etc.).

The term "cytotoxic agent" refers to a substance that inhibits or prevents the expression activity of cells, function of cells and/or causes destruction of cells. The term is intended to include radioactive isotopes chemotherapeutic agents, and toxins such as small molecule toxins or enzymatically active toxins of bacterial, fungal, plant or animal origin, including fragments and/or variants thereof. Examples of cytotoxic agents include, but are not limited to auristatins, auromycins, maytansinoids, yttrium, bismuth, ricin, ricin A-chain, combrestatin, duocarmycins, dolostatins, doxorubicin, daunorubicin, taxol, cisplatin, cc1065, ethidium bromide, mitomycin, etoposide, tenoposide, vincristine, vinblastine, colchicine, dihydroxy anthracin dione, actinomycin, diphtheria toxin, *Pseudomonas* exotoxin (PE) A, PE40, abrin, abrin A chain, modeccin A chain, alpha-sarcin, gelonin, mitogellin, retstrictocin, phenomycin, enomycin, curicin, crotin, calicheamicin, *Sapaonaria officinalis* inhibitor, and glucocorticoid and other chemotherapeutic agents, as well as radioisotopes such as At211, I131, I125, Y90, Re186, Re188, Sm153, Bi212 or 213, P32 and radioactive isotopes of Lu including Lu177. Antibodies may also be conjugated to an anti-cancer pro-drug activating enzyme capable of converting the pro-drug to its active form.

The "gene product" is sometimes referred to herein as a protein or mRNA. For example, a "gene product of the invention" is sometimes referred to herein as a "cancer amino acid sequence", "cancer protein", "protein of a cancer listed in Table I", a "cancer mRNA", "mRNA of a cancer listed in Table I", etc. In one embodiment, the cancer protein is encoded by a nucleic acid of FIG. **2**. The cancer protein can be a fragment, or alternatively, be the full-length protein to the fragment encoded by the nucleic acids of FIG. **2**. In one embodiment, a cancer amino acid sequence is used to determine sequence identity or similarity. In another embodiment, the sequences are naturally occurring allelic variants of a protein encoded by a nucleic acid of FIG. **2**. In another embodiment, the sequences are sequence variants as further described herein.

"High throughput screening" assays for the presence, absence, quantification, or other properties of particular nucleic acids or protein products are well known to those of skill in the art. Similarly, binding assays and reporter gene assays are similarly well known. Thus, e.g., U.S. Pat. No. 5,559,410 discloses high throughput screening methods for proteins; U.S. Pat. No. 5,585,639 discloses high throughput screening methods for nucleic acid binding (i.e., in arrays); while U.S. Pat. Nos. 5,576,220 and 5,541,061 disclose high throughput methods of screening for ligand/antibody binding.

In addition, high throughput screening systems are commercially available (see, e.g., Amersham Biosciences, Piscataway, N.J.; Zymark Corp., Hopkinton, Mass.; Air Technical Industries, Mentor, Ohio; Beckman Instruments, Inc. Fullerton, Calif.; Precision Systems, Inc., Natick, Mass.; etc.). These systems typically automate entire procedures, including all sample and reagent pipetting, liquid dispensing, timed incubations, and final readings of the microplate in detector(s) appropriate for the assay. These configurable systems provide high throughput and rapid start up as well as a high degree of flexibility and customization. The manufacturers of such systems provide detailed protocols for various high throughput systems. Thus, e.g., Zymark Corp. provides technical bulletins describing screening systems for detecting the modulation of gene transcription, ligand binding, and the like.

The term "homolog" refers to a molecule which exhibits homology to another molecule, by for example, having sequences of chemical residues that are the same or similar at corresponding positions.

"Human Leukocyte Antigen" or "HLA" is a human class I or class II Major Histocompatibility Complex (MHC) protein (see, e.g., Stites, et al., Immunology, 8th Ed., Lange Publishing, Los Altos, Calif. (1994).

The terms "hybridize", "hybridizing", "hybridizes" and the like, used in the context of polynucleotides, are meant to refer to conventional hybridization conditions, preferably such as hybridization in 50% formamide/6×SSC/0.1% SDS/100 μg/ml ssDNA, in which temperatures for hybridization are above 37 degrees C. and temperatures for washing in 0.1×SSC/0.1% SDS are above 55 degrees C.

The phrases "isolated" or "biologically pure" refer to material which is substantially or essentially free from components which normally accompany the material as it is found in its native state. Thus, isolated peptides in accordance with the invention preferably do not contain materials normally associated with the peptides in their in situ environment. For example, a polynucleotide is said to be "isolated" when it is substantially separated from contaminant polynucleotides that correspond or are complementary to genes other than the 58P1D12 genes or that encode polypeptides other than

58P1D12 gene product or fragments thereof. A skilled artisan can readily employ nucleic acid isolation procedures to obtain an isolated 58P1D12 polynucleotide. A protein is said to be "isolated," for example, when physical, mechanical or chemical methods are employed to remove the 58P1D12 proteins from cellular constituents that are normally associated with the protein. A skilled artisan can readily employ standard purification methods to obtain an isolated 58P1D12 protein. Alternatively, an isolated protein can be prepared by chemical means.

The term "mammal" refers to any organism classified as a mammal, including mice, rats, rabbits, dogs, cats, cows, horses and humans. In one embodiment of the invention, the mammal is a mouse. In another embodiment of the invention, the mammal is a human.

The terms "metastatic prostate cancer" and "metastatic disease" mean prostate cancers that have spread to regional lymph nodes or to distant sites, and are meant to include stage D disease under the AUA system and stage TxNxM+ under the TNM system. As is the case with locally advanced prostate cancer, surgery is generally not indicated for patients with metastatic disease, and hormonal (androgen ablation) therapy is a preferred treatment modality. Patients with metastatic prostate cancer eventually develop an androgen-refractory state within 12 to 18 months of treatment initiation. Approximately half of these androgen-refractory patients die within 6 months after developing that status. The most common site for prostate cancer metastasis is bone. Prostate cancer bone metastases are often osteoblastic rather than osteolytic (i.e., resulting in net bone formation). Bone metastases are found most frequently in the spine, followed by the femur, pelvis, rib cage, skull and humerus. Other common sites for metastasis include lymph nodes, lung, liver and brain. Metastatic prostate cancer is typically diagnosed by open or laparoscopic pelvic lymphadenectomy, whole body radionuclide scans, skeletal radiography, and/or bone lesion biopsy.

The term "modulator" or "test compound" or "drug candidate" or grammatical equivalents as used herein describe any molecule, e.g., protein, oligopeptide, small organic molecule, polysaccharide, polynucleotide, etc., to be tested for the capacity to directly or indirectly alter the cancer phenotype or the expression of a cancer sequence, e.g., a nucleic acid or protein sequences, or effects of cancer sequences (e.g., signaling, gene expression, protein interaction, etc.) In one aspect, a modulator will neutralize the effect of a cancer protein of the invention. By "neutralize" is meant that an activity of a protein is inhibited or blocked, along with the consequent effect on the cell. In another aspect, a modulator will neutralize the effect of a gene, and its corresponding protein, of the invention by normalizing levels of said protein. In preferred embodiments, modulators alter expression profiles, or expression profile nucleic acids or proteins provided herein, or downstream effector pathways. In one embodiment, the modulator suppresses a cancer phenotype, e.g. to a normal tissue fingerprint. In another embodiment, a modulator induced a cancer phenotype. Generally, a plurality of assay mixtures is run in parallel with different agent concentrations to obtain a differential response to the various concentrations. Typically, one of these concentrations serves as a negative control, i.e., at zero concentration or below the level of detection.

Modulators, drug candidates or test compounds encompass numerous chemical classes, though typically they are organic molecules, preferably small organic compounds having a molecular weight of more than 100 and less than about 2,500 Daltons. Preferred small molecules are less than 2000, or less than 1500 or less than 1000 or less than 500 D. Candidate agents comprise functional groups necessary for structural interaction with proteins, particularly hydrogen bonding, and typically include at least an amine, carbonyl, hydroxyl or carboxyl group, preferably at least two of the functional chemical groups. The candidate agents often comprise cyclical carbon or heterocyclic structures and/or aromatic or polyaromatic structures substituted with one or more of the above functional groups. Modulators also comprise biomolecules such as peptides, saccharides, fatty acids, steroids, purines, pyrimidines, derivatives, structural analogs or combinations thereof. Particularly preferred are peptides. One class of modulators are peptides, for example of from about five to about 35 amino acids, with from about five to about 20 amino acids being preferred, and from about 7 to about 15 being particularly preferred. Preferably, the cancer modulatory protein is soluble, includes a non-transmembrane region, and/or, has an N-terminal Cys to aid in solubility. In one embodiment, the C-terminus of the fragment is kept as a free acid and the N-terminus is a free amine to aid in coupling, i.e., to cysteine. In one embodiment, a cancer protein of the invention is conjugated to an immunogenic agent as discussed herein. In one embodiment, the cancer protein is conjugated to BSA. The peptides of the invention, e.g., of preferred lengths, can be linked to each other or to other amino acids to create a longer peptide/protein. The modulatory peptides can be digests of naturally occurring proteins as is outlined above, random peptides, or "biased" random peptides. In a preferred embodiment, peptide/protein-based modulators are antibodies, and fragments thereof, as defined herein.

Modulators of cancer can also be nucleic acids. Nucleic acid modulating agents can be naturally occurring nucleic acids, random nucleic acids, or "biased" random nucleic acids. For example, digests of prokaryotic or eukaryotic genomes can be used in an approach analogous to that outlined above for proteins.

The term "monoclonal antibody" refers to an antibody obtained from a population of substantially homogeneous antibodies, i.e., the antibodies comprising the population are identical except for possible naturally occurring mutations that are present in minor amounts.

A "motif", as in biological motif of a 58P1D12-related protein, refers to any pattern of amino acids forming part of the primary sequence of a protein, that is associated with a particular function (e.g. protein-protein interaction, protein-DNA interaction, etc) or modification (e.g. that is phosphorylated, glycosylated or amidated), or localization (e.g. secretory sequence, nuclear localization sequence, etc.) or a sequence that is correlated with being immunogenic, either humorally or cellularly. A motif can be either contiguous or capable of being aligned to certain positions that are generally correlated with a certain function or property. In the context of HLA motifs, "motif" refers to the pattern of residues in a peptide of defined length, usually a peptide of from about 8 to about 13 amino acids for a class I HLA motif and from about 6 to about 25 amino acids for a class II HLA motif, which is recognized by a particular HLA molecule. Peptide motifs for HLA binding are typically different for each protein encoded by each human HLA allele and differ in the pattern of the primary and secondary anchor residues.

A "pharmaceutical excipient" comprises a material such as an adjuvant, a carrier, pH-adjusting and buffering agents, tonicity adjusting agents, wetting agents, preservative, and the like.

"Pharmaceutically acceptable" refers to a non-toxic, inert, and/or composition that is physiologically compatible with humans or other mammals.

The term “polynucleotide” means a polymeric form of nucleotides of at least 10 bases or base pairs in length, either ribonucleotides or deoxynucleotides or a modified form of either type of nucleotide, and is meant to include single and double stranded forms of DNA and/or RNA. In the art, this term if often used interchangeably with “oligonucleotide”. A polynucleotide can comprise a nucleotide sequence disclosed herein wherein thymidine (T), as shown for example in FIG. **2**, can also be uracil (U); this definition pertains to the differences between the chemical structures of DNA and RNA, in particular the observation that one of the four major bases in RNA is uracil (U) instead of thymidine (T).

The term “polypeptide” means a polymer of at least about 4, 5, 6, 7, or 8 amino acids. Throughout the specification, standard three letter or single letter designations for amino acids are used. In the art, this term is often used interchangeably with “peptide” or “protein”.

An HLA “primary anchor residue” is an amino acid at a specific position along a peptide sequence which is understood to provide a contact point between the immunogenic peptide and the HLA molecule. One to three, usually two, primary anchor residues within a peptide of defined length generally defines a “motif” for an immunogenic peptide. These residues are understood to fit in close contact with peptide binding groove of an HLA molecule, with their side chains buried in specific pockets of the binding groove. In one embodiment, for example, the primary anchor residues for an HLA class I molecule are located at position 2 (from the amino terminal position) and at the carboxyl terminal position of a 8, 9, 10, 11, or 12 residue peptide epitope in accordance with the invention. Alternatively, in another embodiment, the primary anchor residues of a peptide binds an HLA class II molecule are spaced relative to each other, rather than to the termini of a peptide, where the peptide is generally of at least 9 amino acids in length. The primary anchor positions for each motif and supermotif are set forth in Table IV. For example, analog peptides can be created by altering the presence or absence of particular residues in the primary and/or secondary anchor positions shown in Table IV. Such analogs are used to modulate the binding affinity and/or population coverage of a peptide comprising a particular HLA motif or supermotif.

“Radioisotopes” include, but are not limited to the following (non-limiting exemplary uses are also set forth):

Examples of Medical Isotopes

Isotope	Description of use
Actinium-225 (AC-225)	See Thorium-229 (Th-229)
Actinium-227 (AC-227)	Parent of Radium-223 (Ra-223) which is an alpha emitter used to treat metastases in the skeleton resulting from cancer (i.e., breast and prostate cancers), and cancer radioimmunotherapy
Bismuth-212 (Bi-212)	See Thorium-228 (Th-228)
Bismuth-213 (Bi-213)	See Thorium-229 (Th-229)
Cadmium-109 (Cd-109)	Cancer detection
Cobalt-60 (Co-60)	Radiation source for radiotherapy of cancer, for food irradiators, and for sterilization of medical supplies
Copper-64 (Cu-64)	A positron emitter used for cancer therapy and SPECT imaging
Copper-67 (Cu-67)	Beta/gamma emitter used in cancer radioimmunotherapy and diagnostic studies (i.e., breast and colon cancers, and lymphoma)
Dysprosium-166 (Dy-166)	Cancer radioimmunotherapy
Erbium-169 (Er-169)	Rheumatoid arthritis treatment, particularly for the small joints associated with fingers and toes
Europium-152 (Eu-152)	Radiation source for food irradiation and for sterilization of medical supplies
Europium-154 (Eu-154)	Radiation source for food irradiation and for sterilization of medical supplies
Gadolinium-153 (Gd-153)	Osteoporosis detection and nuclear medical quality assurance devices
Gold-198 (Au-198)	Implant and intracavity therapy of ovarian, prostate, and brain cancers
Holmium-166 (Ho-166)	Multiple myeloma treatment in targeted skeletal therapy, cancer radioimmunotherapy, bone marrow ablation, and rheumatoid arthritis treatment
Iodine-125 (I-125)	Osteoporosis detection, diagnostic imaging, tracer drugs, brain cancer treatment, radiolabeling, tumor imaging, mapping of receptors in the brain, interstitial radiation therapy, brachytherapy for treatment of prostate cancer, determination of glomerular filtration rate (GFR), determination of plasma volume, detection of deep vein thrombosis of the legs
Iodine-131 (I-131)	Thyroid function evaluation, thyroid disease detection, treatment of thyroid cancer as well as other non-malignant thyroid diseases (i.e., Graves disease, goiters, and hyperthyroidism), treatment of leukemia, lymphoma, and other forms of cancer (e.g., breast cancer) using radioimmunotherapy
Iridium-192 (Ir-192)	Brachytherapy, brain and spinal cord tumor treatment, treatment of blocked arteries (i.e., arteriosclerosis and restenosis), and implants for breast and prostate tumors
Lutetium-177 (Lu-177)	Cancer radioimmunotherapy and treatment of blocked arteries (i.e., arteriosclerosis and restenosis)
Molybdenum-99 (Mo-99)	Parent of Technetium-99m (Tc-99m) which is used for imaging the brain, liver, lungs, heart, and other organs. Currently, Tc-99m is the most widely used radioisotope used for diagnostic imaging of various cancers and diseases involving the brain, heart, liver, lungs; also used in detection of deep vein thrombosis of the legs
Osmium-194 (Os-194)	Cancer radioimmunotherapy
Palladium-103 (Pd-103)	Prostate cancer treatment
Platinum-195m (Pt-195m)	Studies on biodistribution and metabolism of cisplatin, a chemotherapeutic drug
Phosphorus-32 (P-32)	Polycythemia rubra vera (blood cell disease) and leukemia treatment, bone cancer diagnosis/treatment; colon, pancreatic, and liver cancer treatment; radiolabeling nucleic acids for in vitro research, diagnosis of superficial tumors, treatment of blocked arteries (i.e., arteriosclerosis and restenosis), and intracavity therapy
Phosphorus-33 (P-33)	Leukemia treatment, bone disease diagnosis/treatment, radiolabeling, and treatment of blocked arteries (i.e., arteriosclerosis and restenosis)

-continued

Isotope	Description of use
Radium-223 (Ra-223)	See Actinium-227 (Ac-227)
Rhenium-186 (Re-186)	Bone cancer pain relief, rheumatoid arthritis treatment, and diagnosis and treatment of lymphoma and bone, breast, colon, and liver cancers using radioimmunotherapy
Rhenium-188 (Re-188)	Cancer diagnosis and treatment using radioimmunotherapy, bone cancer pain relief, treatment of rheumatoid arthritis, and treatment of prostate cancer
Rhodium-105 (Rh-105)	Cancer radioimmunotherapy
Samarium-145 (Sm-145)	Ocular cancer treatment
Samarium-153 (Sm-153)	Cancer radioimmunotherapy and bone cancer pain relief
Scandium-47 (Sc-47)	Cancer radioimmunotherapy and bone cancer pain relief
Selenium-75 (Se-75)	Radiotracer used in brain studies, imaging of adrenal cortex by gamma-scintigraphy, lateral locations of steroid secreting tumors, pancreatic scanning, detection of hyperactive parathyroid glands, measure rate of bile acid loss from the endogenous pool
Strontium-85 (Sr-85)	Bone cancer detection and brain scans
Strontium-89 (Sr-89)	Bone cancer pain relief, multiple myeloma treatment, and osteoblastic therapy
Technetium-99m (Tc-99m)	See Molybdenum-99 (Mo-99)
Thorium-228 (Th-228)	Parent of Bismuth-212 (Bi-212) which is an alpha emitter used in cancer radioimmunotherapy
Thorium-229 (Th-229)	Parent of Actinium-225 (Ac-225) and grandparent of Bismuth-213 (Bi-213) which are alpha emitters used in cancer radioimmunotherapy
Thulium-170 (Tm-170)	Gamma source for blood irradiators, energy source for implanted medical devices
Tin-117m (Sn-117m)	Cancer immunotherapy and bone cancer pain relief
Tungsten-188 (W-188)	Parent for Rhenium-188 (Re-188) which is used for cancer diagnostics/treatment, bone cancer pain relief, rheumatoid arthritis treatment, and treatment of blocked arteries (i.e., arteriosclerosis and restenosis)
Xenon-127 (Xe-127)	Neuroimaging of brain disorders, high resolution SPECT studies, pulmonary function tests, and cerebral blood flow studies
Ytterbium-175 (Yb-175)	Cancer radioimmunotherapy
Yttrium-90 (Y-90)	Microseeds obtained from irradiating Yttrium-89 (Y-89) for liver cancer treatment
Yttrium-91 (Y-91)	A gamma-emitting label for Yttrium-90 (Y-90) which is used for cancer radioimmunotherapy (i.e., lymphoma, breast, colon, kidney, lung, ovarian, prostate, pancreatic, and inoperable liver cancers)

By "randomized" or grammatical equivalents as herein applied to nucleic acids and proteins is meant that each nucleic acid and peptide consists of essentially random nucleotides and amino acids, respectively. These random peptides (or nucleic acids, discussed herein) can incorporate any nucleotide or amino acid at any position. The synthetic process can be designed to generate randomized proteins or nucleic acids, to allow the formation of all or most of the possible combinations over the length of the sequence, thus forming a library of randomized candidate bioactive proteinaceous agents.

In one embodiment, a library is "fully randomized," with no sequence preferences or constants at any position. In another embodiment, the library is a "biased random" library. That is, some positions within the sequence either are held constant, or are selected from a limited number of possibilities. For example, the nucleotides or amino acid residues are randomized within a defined class, e.g., of hydrophobic amino acids, hydrophilic residues, sterically biased (either small or large) residues, towards the creation of nucleic acid binding domains, the creation of cysteines, for cross-linking, prolines for SH-3 domains, serines, threonines, tyrosines or histidines for phosphorylation sites, etc., or to purines, etc.

A "recombinant" DNA or RNA molecule is a DNA or RNA molecule that has been subjected to molecular manipulation in vitro.

Non-limiting examples of small molecules include compounds that bind or interact with 58P1D12, ligands including hormones, neuropeptides, chemokines, odorants, phospholipids, and functional equivalents thereof that bind and preferably inhibit 58P1D12 protein function. Such non-limiting small molecules preferably have a molecular weight of less than about 10 kDa, more preferably below about 9, about 8, about 7, about 6, about 5 or about 4 kDa. In certain embodiments, small molecules physically associate with, or bind, 58P1D12 protein; are not found in naturally occurring metabolic pathways; and/or are more soluble in aqueous than non-aqueous solutions.

"Stringency" of hybridization reactions is readily determinable by one of ordinary skill in the art, and generally is an empirical calculation dependent upon probe length, washing temperature, and salt concentration. In general, longer probes require higher temperatures for proper annealing, while shorter probes need lower temperatures. Hybridization generally depends on the ability of denatured nucleic acid sequences to reanneal when complementary strands are present in an environment below their melting temperature. The higher the degree of desired homology between the probe and hybridizable sequence, the higher the relative temperature that can be used. As a result, it follows that higher relative temperatures would tend to make the reaction conditions more stringent, while lower temperatures less so. For additional details and explanation of stringency of hybridization reactions, see Ausubel et al., Current Protocols in Molecular Biology, Wiley Interscience Publishers, (1995).

"Stringent conditions" or "high stringency conditions", as defined herein, are identified by, but not limited to, those that: (1) employ low ionic strength and high temperature for washing, for example 0.015 M sodium chloride/0.0015 M sodium citrate/0.1% sodium dodecyl sulfate at 50° C.; (2) employ during hybridization a denaturing agent, such as formamide, for example, 50% (v/v) formamide with 0.1% bovine serum albumin/0.1% Ficoll/0.1% polyvinylpyrrolidone/50 mM sodium phosphate buffer at pH 6.5 with 750 mM sodium chloride, 75 mM sodium citrate at 42° C.; or (3) employ 50% formamide, 5×SSC (0.75 M NaCl, 0.075 M sodium citrate), 50 mM sodium phosphate (pH 6.8), 0.1% sodium pyrophosphate, 5×Denhardt's solution, sonicated salmon sperm DNA (50 μg/ml), 0.1% SDS, and 10% dextran sulfate at 42° C., with washes at 42° C. in 0.2×SSC (sodium chloride/sodium citrate) and 50% formamide at 55° C., followed by a high-stringency wash consisting of 0.1×SSC containing EDTA at 55° C. "Moderately stringent conditions" are described by, but not limited to, those in Sambrook et al., Molecular Cloning: A Laboratory Manual, New York: Cold Spring Harbor Press, 1989, and include the use of washing solution and hybridization conditions (e.g., temperature, ionic strength and % SDS) less stringent than those described above. An example of moderately stringent conditions is overnight incubation at 37° C. in a solution comprising: 20% formamide, 5×SSC (150 mM NaCl, 15 mM trisodium citrate), 50 mM sodium phosphate (pH 7.6), 5×Denhardt's solution, 10% dextran sulfate, and 20 mg/mL denatured sheared salmon sperm DNA, followed by washing the filters in 1×SSC at about 37-50° C. The skilled artisan will recognize how to adjust the temperature, ionic strength, etc. as necessary to accommodate factors such as probe length and the like.

An HLA "supermotif" is a peptide binding specificity shared by HLA molecules encoded by two or more HLA alleles. Overall phenotypic frequencies of HLA-supertypes in different ethnic populations are set forth in Table IV (F). The non-limiting constituents of various supetypes are as follows:

A2: A*0201, A*0202, A*0203, A*0204, A*0205, A*0206, A*6802, A*6901, A*0207

A3: A3, A11, A31, A*3301, A*6801, A*0301, A*1101, A*3101

B7: B7, B*3501-03, B*51, B*5301, B*5401, B*5501, B*5502, B*5601, B*6701, B*7801, B*0702, B*5101, B*5602

B44: B*3701, B*4402, B*4403, B*60 (B*4001), B61 (B*4006)

A1: A*0102, A*2604, A*3601, A*4301, A*8001

A24: A*24, A*30, A*2403, A*2404, A*3002, A*3003

B27: B*1401-02, B*1503, B*1509, B*1510, B*1518, B*3801-02, B*3901, B*3902, B*3903-04, B*4801-02, B*7301, B*2701-08

B58: B*1516, B*1517, B*5701, B*5702, B58

B62: B*4601, B52, B*1501 (B62), B*1502 (B75), B*1513 (B77)

Calculated population coverage afforded by different HLA-supertype combinations are set forth in Table IV (G).

As used herein "to treat" or "therapeutic" and grammatically related terms, refer to any improvement of any consequence of disease, such as prolonged survival, less morbidity, and/or a lessening of side effects which are the byproducts of an alternative therapeutic modality; full eradication of disease is not required.

A "transgenic animal" (e.g., a mouse or rat) is an animal having cells that contain a transgene, which transgene was introduced into the animal or an ancestor of the animal at a prenatal, e.g., an embryonic stage. A "transgene" is a DNA that is integrated into the genome of a cell from which a transgenic animal develops.

As used herein, an HLA or cellular immune response "vaccine" is a composition that contains or encodes one or more peptides of the invention. There are numerous embodiments of such vaccines, such as a cocktail of one or more individual peptides; one or more peptides of the invention comprised by a polyepitopic peptide; or nucleic acids that encode such individual peptides or polypeptides, e.g., a minigene that encodes a polyepitopic peptide. The "one or more peptides" can include any whole unit integer from 1-150 or more, e.g., at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, or 150 or more peptides of the invention. The peptides or polypeptides can optionally be modified, such as by lipidation, addition of targeting or other sequences. HLA class 1 peptides of the invention can be admixed with, or linked to, HLA class 11 peptides, to facilitate activation of both cytotoxic T lymphocytes and helper T lymphocytes. HLA vaccines can also comprise peptide-pulsed antigen presenting cells, e.g., dendritic cells.

The term "variant" refers to a molecule that exhibits a variation from a described type or norm, such as a protein that has one or more different amino acid residues in the corresponding position(s) of a specifically described protein (e.g. the 58P1D12 protein shown in FIG. **2** or FIG. **3**. An analog is an example of a variant protein. Splice isoforms and single nucleotides polymorphisms (SNPs) are further examples of variants.

The "58P1D12-related proteins" of the invention include those specifically identified herein, as well as allelic variants, conservative substitution variants, analogs and homologs that can be isolated/generated and characterized without undue experimentation following the methods outlined herein or readily available in the art. Fusion proteins that combine parts of different 58P1D12 proteins or fragments thereof, as well as fusion proteins of a 58P1D12 protein and a heterologous polypeptide are also included. Such 58P1D12 proteins are collectively referred to as the 58P1D12-related proteins, the proteins of the invention, or 58P1D12. The term "58P1D12-related protein" refers to a polypeptide fragment or a 58P1D12 protein sequence of 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, or more than 25 amino acids; or, at least 30, 35, 40, 45, 50, 55, 60, 65, 70, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 525, 550, 575, or 576 or more amino acids.

11.) 58P1D12 Polynucleotides

One aspect of the invention provides polynucleotides corresponding or complementary to all or part of a 58P1D12 gene, mRNA, and/or coding sequence, preferably in isolated form, including polynucleotides encoding a 58P1D12-related protein and fragments thereof, DNA, RNA, DNA/RNA hybrid, and related molecules, polynucleotides or oligonucleotides complementary to a 58P1D12 gene or mRNA sequence or a part thereof, and polynucleotides or oligonucleotides that hybridize to a 58P1D12 gene, mRNA, or to a 58P1D12 encoding polynucleotide (collectively, "58P1D12 polynucleotides"). In all instances when referred to in this section, T can also be U in FIG. **2**.

Embodiments of a 58P1D12 polynucleotide include: a 58P1D12 polynucleotide having the sequence shown in FIG. **2**, the nucleotide sequence of 58P1D12 as shown in FIG. **2** wherein T is U; at least 10 contiguous nucleotides of a polynucleotide having the sequence as shown in FIG. **2**; or, at least 10 contiguous nucleotides of a polynucleotide having the sequence as shown in FIG. **2** where T is U. For example, embodiments of 58P1D12 nucleotides comprise, without limitation:

(1) a polynucleotide comprising, consisting essentially of, or consisting of a sequence as shown in FIG. **2**, wherein T can also be U;

(11) a polynucleotide comprising, consisting essentially of, or consisting of the sequence as shown in FIG. **2**A, from

nucleotide residue number 380 through nucleotide residue number 1201, including the stop codon, wherein T can also be U;

(III) a polynucleotide comprising, consisting essentially of, or consisting of the sequence as shown in FIG. **2**B, from nucleotide residue number 388 through nucleotide residue number 1086, including the stop codon, wherein T can also be U;

(IV) a polynucleotide comprising, consisting essentially of, or consisting of the sequence as shown in FIG. **2**C, from nucleotide residue number 206 through nucleotide residue number 904, including the a stop codon, wherein T can also be U;

(V) a polynucleotide comprising, consisting essentially of, or consisting of the sequence as shown in FIG. **2**D, from nucleotide residue number 206 through nucleotide residue number 916, including the stop codon, wherein T can also be U;

(VI) a polynucleotide comprising, consisting essentially of, or consisting of the sequence as shown in FIG. **2**E, from nucleotide residue number 106 through nucleotide residue number 816, including the stop codon, wherein T can also be U;

(VII) a polynucleotide comprising, consisting essentially of, or consisting of the sequence as set forth in FIG. **2**F, wherein T can also be U;

(VIII) a polynucleotide that encodes a 58P1D12-related protein that is at least 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% homologous to an entire amino acid sequence shown in FIG. **2**A-F;

(IX) a polynucleotide that encodes a 58P1D12-related protein that is at least 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% identical to an entire amino acid sequence shown in FIG. **2**A-F;

(X) a polynucleotide that encodes at least one peptide set forth in Tables VIII-XXI and XXII-XLIX;

(XI) a polynucleotide that encodes a peptide region of at least 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35 amino acids of a peptide of FIG. **3**A in any whole number increment up to 273 that includes at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35 amino acid position(s) having a value greater than 0.5 in the Hydrophilicity profile of FIG. **5**;

(XII) a polynucleotide that encodes a peptide region of at least 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35 amino acids of a peptide of FIG. **3**A in any whole number increment up to 273 that includes 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35 amino acid position(s) having a value less than 0.5 in the Hydropathicity profile of FIG. **6**;

(XIII) a polynucleotide that encodes a peptide region of at least 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35 amino acids of a peptide of FIG. **3**A in any whole number increment up to 273 that includes 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35 amino acid position(s) having a value greater than 0.5 in the Percent Accessible Residues profile of FIG. **7**;

(XIV) a polynucleotide that encodes a peptide region of at least 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35 amino acids of a peptide of FIG. **3**A in any whole number increment up to 273 that includes 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35 amino acid position(s) having a value greater than 0.5 in the Average Flexibility profile of FIG. **8**;

(XV) a polynucleotide that encodes a peptide region of at least 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35 amino acids of a peptide of FIG. **3**A in any whole number increment up to 273 that includes 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35 amino acid position(s) having a value greater than 0.5 in the Beta-turn profile of FIG. **9**;

(XVI) a polynucleotide that encodes a peptide region of at least 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35 amino acids of a peptide of FIGS. **3**B, and/or **3**C in any whole number increment up to 232 and/or 236 respectively that includes 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35 amino acid position(s) having a value greater than 0.5 in the Hydrophilicity profile of FIG. **5**;

(XVII) a polynucleotide that encodes a peptide region of at least 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35 amino acids of a peptide of FIGS. **3**B and/or **3**C in any whole number increment up to 232 and/or 236 respectively that includes 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35 amino acid position(s) having a value less than 0.5 in the Hydropathicity profile of FIG. **6**;

(XVIII) a polynucleotide that encodes a peptide region of at least 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35 amino acids of a peptide of FIGS. **3**B and/or **3**C in any whole number increment up to 232 and/or 236 respectively that includes 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35 amino acid position(s) having a value greater than 0.5 in the Percent Accessible Residues profile of FIG. **7**;

(XIX) a polynucleotide that encodes a peptide region of at least 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35 amino acids of a peptide of FIGS. **3**B and/or **3**C in any whole number increment up to 232 and/or 236 respectively that includes 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35 amino acid position(s) having a value greater than 0.5 in the Average Flexibility profile of FIG. **8**;

(XX) a polynucleotide that encodes a peptide region of at least 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35 amino acids of a peptide of FIGS. **3**B and/or **3**C in any whole number increment up to 232 and/or 236 respectively that includes 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35 amino acid position(s) having a value greater than 0.5 in the Beta-turn profile of FIG. **9**;

(XXI) a polynucleotide that is fully complementary to a polynucleotide of any one of (I)-(XX);

(XXII) a polynucleotide that is fully complementary to a polynucleotide of any one of (I)-(XXI);

(XXIII) a peptide that is encoded by any of (I) to (XXII); and;

(XXIV) a composition comprising a polynucleotide of any of (I)-(XXII) or peptide of (XXIII) together with a pharmaceutical excipient and/or in a human unit dose form;

(XXV) a method of using a polynucleotide of any (I)-(XXII) or peptide of (XXIII) or a composition of (XXIV) in a method to modulate a cell expressing 58P1D12;

(XXVl) a method of using a polynucleotide of any (l)-(XXll) or peptide of (XXlll) or a composition of (XXlV) in a method to diagnose, prophylax, prognose, or treat an individual who bears a cell expressing 58P1D12;

(XXVll) a method of using a polynucleotide of any (l)-(XXll) or peptide of (XXlll) or a composition of (XXlV) in a method to diagnose, prophylax, prognose, or treat an individual who bears a cell expressing 58P1D12, said cell from a cancer of a tissue listed in Table 1;

(XXVlll) a method of using a polynucleotide of any (l)-(XXll) or peptide of (XXlll) or a composition of (XXlV) in a method to diagnose, prophylax, prognose, or treat a a cancer;

(XXlX) a method of using a polynucleotide of any (l)-(XXll) or peptide of (XXlll) or a composition of (XXlV) in a method to diagnose, prophylax, prognose, or treat a a cancer of a tissue listed in Table l; and;

(XXX) a method of using a polynucleotide of any (l)-(XXll) or peptide of (XXlll) or a composition of (XXlV) in a method to identify or characterize a modulator of a cell expressing 58P1D12.

As used herein, a range is understood to disclose specifically all whole unit positions thereof.

Typical embodiments of the invention disclosed herein include 58P1D12 polynucleotides that encode specific portions of 58P1D12 mRNA sequences (and those which are complementary to such sequences) such as those that encode the proteins and/or fragments thereof, for example: [!!! revise to include up to the appropriate length].

(a) 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 225, 250, 255, 260, 265, 270, and 273 or more contiguous amino acids of 58P1D12 variant 1; the maximal lengths relevant for other variants are: variant 2, 232 amino acids; and variant 3, 236 amino acids.

For example, representative embodiments of the invention disclosed herein include: polynucleotides and their encoded peptides themselves encoding about amino acid 1 to about amino acid 10 of the 58P1D12 protein shown in FlG. **2** or FlG. **3**, polynucleotides encoding about amino acid 10 to about amino acid 20 of the 58P1D12 protein shown in FlG. **2** or FlG. **3**, polynucleotides encoding about amino acid 20 to about amino acid 30 of the 58P1D12 protein shown in FlG. **2** or FlG. **3**, polynucleotides encoding about amino acid 30 to about amino acid 40 of the 58P1D12 protein shown in FlG. **2** or FlG. **3**, polynucleotides encoding about amino acid 40 to about amino acid 50 of the 58P1D12 protein shown in FlG. **2** or FlG. **3**, polynucleotides encoding about amino acid 50 to about amino acid 60 of the 58P1D12 protein shown in FlG. **2** or FlG. **3**, polynucleotides encoding about amino acid 60 to about amino acid 70 of the 58P1D12 protein shown in FlG. **2** or FlG. **3**, polynucleotides encoding about amino acid 70 to about amino acid 80 of the 58P1D12 protein shown in FlG. **2** or FlG. **3**, polynucleotides encoding about amino acid 80 to about amino acid 90 of the 58P1D12 protein shown in FlG. **2** or FlG. **3**, polynucleotides encoding about amino acid 90 to about amino acid 100 of the 58P1D12 protein shown in FlG. **2** or FlG. **3**, in increments of about 10 amino acids, ending at the carboxyl terminal amino acid set forth in FlG. **2** or FlG. **3**. Accordingly, polynucleotides encoding portions of the amino acid sequence (of about 10 amino acids), of amino acids, 100 through the carboxyl terminal amino acid of the 58P1D12 protein are embodiments of the invention. Wherein it is understood that each particular amino acid position discloses that position plus or minus five amino acid residues.

Polynucleotides encoding relatively long portions of a 58P1D12 protein are also within the scope of the invention. For example, polynucleotides encoding from about amino acid 1 (or 20 or 30 or 40 etc.) to about amino acid 20, (or 30, or 40 or 50 etc.) of the 58P1D12 protein "or variant" shown in FlG. **2** or FlG. **3** can be generated by a variety of techniques well known in the art. These polynucleotide fragments can include any portion of the 58P1D12 sequence as shown in FlG. **2**.

Additional illustrative embodiments of the invention disclosed herein include 58P1D12 polynucleotide fragments encoding one or more of the biological motifs contained within a 58P1D12 protein "or variant" sequence, including one or more of the motif-bearing subsequences of a 58P1D12 protein "or variant" set forth in Tables Vlll-XXl and XXll-XLlX. ln another embodiment, typical polynucleotide fragments of the invention encode one or more of the regions of 58P1D12 protein or variant that exhibit homology to a known molecule. ln another embodiment of the invention, typical polynucleotide fragments can encode one or more of the 58P1D12 protein or variant N-glycosylation sites, cAMP and cGMP-dependent protein kinase phosphorylation sites, casein kinase ll phosphorylation sites or N-myristoylation site and amidation sites.

Note that to determine the starting position of any peptide set forth in Tables Vlll-XXl and Tables XXll to XLlX (collectively HLA Peptide Tables) respective to its parental protein, e.g., variant 1, variant 2, etc., reference is made to three factors: the particular variant, the length of the peptide in an HLA Peptide Table, and the Search Peptides listed in Table Vll. Generally, a unique Search Peptide is used to obtain HLA peptides for a particular variant. The position of each Search Peptide relative to its respective parent molecule is listed in Table Vll. Accordingly, if a Search Peptide begins at position "X", one must add the value "X minus 1" to each position in Tables Vlll-XXl and Tables XXll-lL to obtain the actual position of the HLA peptides in their parental molecule. For example if a particular Search Peptide begins at position 150 of its parental molecule, one must add 150−1, i.e., 149 to each HLA peptide amino acid position to calculate the position of that amino acid in the parent molecule.

ll.A.) Uses of 58P1D12 Polynucleotides

ll.A.1.) Monitoring of Genetic Abnormalities

The polynucleotides of the preceding paragraphs have a number of different specific uses. The human 58P1D12 gene maps to the chromosomal location set forth in the Example entitled "Chromosomal Mapping of 58P1D12." For example, because the 58P1D12 gene maps to this chromosome, polynucleotides that encode different regions of the 58P1D12 proteins are used to characterize cytogenetic abnormalities of this chromosomal locale, such as abnormalities that are identified as being associated with various cancers. ln certain genes, a variety of chromosomal abnormalities including rearrangements have been identified as frequent cytogenetic abnormalities in a number of different cancers (see e.g. Krajinovic et al., Mutat. Res. 382(3-4): 81-83 (1998); Johansson et al., Blood 86(10): 3905-3914 (1995) and Finger et al., P.N.A.S. 85(23): 9158-9162 (1988)). Thus, polynucleotides encoding specific regions of the 58P1D12 proteins provide new tools that can be used to delineate, with greater precision than previously possible, cytogenetic abnormalities in the chromosomal region that encodes 58P1D12 that may contribute to the malignant phenotype. ln this context, these polynucleotides satisfy a need in the art for expanding the sensitivity of chromosomal screening in order to identify more subtle and less common chromosomal abnormalities (see e.g. Evans et al., Am. J. Obstet. Gynecol 171(4): 1055-1057 (1994)).

Furthermore, as 58P1D12 was shown to be highly expressed in prostate and other cancers, 58P1D12 polynucleotides are used in methods assessing the status of 58P1D12 gene products in normal versus cancerous tissues. Typically, polynucleotides that encode specific regions of the 58P1D12 proteins are used to assess the presence of perturbations (such as deletions, insertions, point mutations, or alterations resulting in a loss of an antigen etc.) in specific regions of the 58P1D12 gene, such as regions containing one or more motifs. Exemplary assays include both RT-PCR assays as well as single-strand conformation polymorphism (SSCP) analysis (see, e.g., Marrogi et al., J. Cutan. Pathol. 26(8): 369-378 (1999), both of which utilize polynucleotides encoding specific regions of a protein to examine these regions within the protein.

II.A.2.) Antisense Embodiments

Other specifically contemplated nucleic acid related embodiments of the invention disclosed herein are genomic DNA, cDNAs, ribozymes, and antisense molecules, as well as nucleic acid molecules based on an alternative backbone, or including alternative bases, whether derived from natural sources or synthesized, and include molecules capable of inhibiting the RNA or protein expression of 58P1D12. For example, antisense molecules can be RNAs or other molecules, including peptide nucleic acids (PNAs) or non-nucleic acid molecules such as phosphorothioate derivatives that specifically bind DNA or RNA in a base pair-dependent manner. A skilled artisan can readily obtain these classes of nucleic acid molecules using the 58P1D12 polynucleotides and polynucleotide sequences disclosed herein.

Antisense technology entails the administration of exogenous oligonucleotides that bind to a target polynucleotide located within the cells. The term “antisense” refers to the fact that such oligonucleotides are complementary to their intracellular targets, e.g., 58P1D12. See for example, Jack Cohen, Oligodeoxynucleotides, Antisense Inhibitors of Gene Expression, CRC Press, 1989; and Synthesis 1:1-5 (1988). The 58P1D12 antisense oligonucleotides of the present invention include derivatives such as S-oligonucleotides (phosphorothioate derivatives or S-oligos, see, Jack Cohen, supra), which exhibit enhanced cancer cell growth inhibitory action. S-oligos (nucleoside phosphorothioates) are isoelectronic analogs of an oligonucleotide (O-oligo) in which a nonbridging oxygen atom of the phosphate group is replaced by a sulfur atom. The S-oligos of the present invention can be prepared by treatment of the corresponding O-oligos with 3H-1,2-benzodithiol-3-one-1,1-dioxide, which is a sulfur transfer reagent. See, e.g., Iyer, R. P. et al., J. Org. Chem. 55:4693-4698 (1990); and Iyer, R. P. et al., J. Am. Chem. Soc. 112:1253-1254 (1990). Additional 58P1D12 antisense oligonucleotides of the present invention include morpholino antisense oligonucleotides known in the art (see, e.g., Partridge et al., 1996, Antisense & Nucleic Acid Drug Development 6: 169-175).

The 58P1D12 antisense oligonucleotides of the present invention typically can be RNA or DNA that is complementary to and stably hybridizes with the first 100 5' codons or last 100 3' codons of a 58P1D12 genomic sequence or the corresponding mRNA. Absolute complementarity is not required, although high degrees of complementarity are preferred. Use of an oligonucleotide complementary to this region allows for the selective hybridization to 58P1D12 mRNA and not to mRNA specifying other regulatory subunits of protein kinase. In one embodiment, 58P1D12 antisense oligonucleotides of the present invention are 15 to 30-mer fragments of the antisense DNA molecule that have a sequence that hybridizes to 58P1D12 mRNA. Optionally, 58P1D12 antisense oligonucleotide is a 30-mer oligonucleotide that is complementary to a region in the first 10 5' codons or last 10 3' codons of 58P1D12. Alternatively, the antisense molecules are modified to employ ribozymes in the inhibition of 58P1D12 expression, see, e.g., L. A. Couture & D. T. Stinchcomb; Trends Genet 12: 510-515 (1996).

II.A.3.) Primers and Primer Pairs

Further specific embodiments of these nucleotides of the invention include primers and primer pairs, which allow the specific amplification of polynucleotides of the invention or of any specific parts thereof, and probes that selectively or specifically hybridize to nucleic acid molecules of the invention or to any part thereof. Probes can be labeled with a detectable marker, such as, for example, a radioisotope, fluorescent compound, bioluminescent compound, a chemiluminescent compound, metal chelator or enzyme. Such probes and primers are used to detect the presence of a 58P1D12 polynucleotide in a sample and as a means for detecting a cell expressing a 58P1D12 protein.

Examples of such probes include polypeptides comprising all or part of the human 58P1D12 cDNA sequence shown in FIG. **2**. Examples of primer pairs capable of specifically amplifying 58P1D12 mRNAs are also described in the Examples. As will be understood by the skilled artisan, a great many different primers and probes can be prepared based on the sequences provided herein and used effectively to amplify and/or detect a 58P1D12 mRNA.

The 58P1D12 polynucleotides of the invention are useful for a variety of purposes, including but not limited to their use as probes and primers for the amplification and/or detection of the 58P1D12 gene(s), mRNA(s), or fragments thereof; as reagents for the diagnosis and/or prognosis of prostate cancer and other cancers; as coding sequences capable of directing the expression of 58P1D12 polypeptides; as tools for modulating or inhibiting the expression of the 58P1D12 gene(s) and/or translation of the 58P1D12 transcript(s); and as therapeutic agents.

The present invention includes the use of any probe as described herein to identify and isolate a 58P1D12 or 58P1D12 related nucleic acid sequence from a naturally occurring source, such as humans or other mammals, as well as the isolated nucleic acid sequence per se, which would comprise all or most of the sequences found in the probe used.

II.A.4.) Isolation of 58P1D12-Encoding Nucleic Acid Molecules

The 58P1D12 cDNA sequences described herein enable the isolation of other polynucleotides encoding 58P1D12 gene product(s), as well as the isolation of polynucleotides encoding 58P1D12 gene product homologs, alternatively spliced isoforms, allelic variants, and mutant forms of a 58P1D12 gene product as well as polynucleotides that encode analogs of 58P1D12-related proteins. Various molecular cloning methods that can be employed to isolate full length cDNAs encoding a 58P1D12 gene are well known (see, for example, Sambrook, J. et al., Molecular Cloning: A Laboratory Manual, 2d edition, Cold Spring Harbor Press, New York, 1989; Current Protocols in Molecular Biology. Ausubel et al., Eds., Wiley and Sons, 1995). For example, lambda phage cloning methodologies can be conveniently employed, using commercially available cloning systems (e.g., Lambda ZAP Express, Stratagene). Phage clones containing 58P1D12 gene cDNAs can be identified by probing with a labeled 58P1D12 cDNA or a fragment thereof. For example, in one embodiment, a 58P1D12 cDNA (e.g., FIG. **2**) or a portion thereof can be synthesized and used as a probe to retrieve overlapping and full-length cDNAs corresponding to a 58P1D12 gene. A 58P1D12 gene itself can be isolated by screening genomic DNA libraries, bacterial artificial chromosome libraries (BACs), yeast artificial chromosome libraries (YACs), and the like, with 58P1D12 DNA probes or primers.

II.A.5.) Recombinant Nucleic Acid Molecules and Host-Vector Systems

The invention also provides recombinant DNA or RNA molecules containing a 58P1D12 polynucleotide, a fragment, analog or homologue thereof, including but not limited to phages, plasmids, phagemids, cosmids, YACs, BACs, as well as various viral and non-viral vectors well known in the art, and cells transformed or transfected with such recombinant DNA or RNA molecules. Methods for generating such molecules are well known (see, for example, Sambrook et al., 1989, supra).

The invention further provides a host-vector system comprising a recombinant DNA molecule containing a 58P1D12 polynucleotide, fragment, analog or homologue thereof within a suitable prokaryotic or eukaryotic host cell. Examples of suitable eukaryotic host cells include a yeast cell, a plant cell, or an animal cell, such as a mammalian cell or an insect cell (e.g., a baculovirus-infectible cell such as an Sf9 or HighFive cell). Examples of suitable mammalian cells include various prostate cancer cell lines such as DU145 and TsuPr1, other transfectable or transducible prostate cancer cell lines, primary cells (PrEC), as well as a number of mammalian cells routinely used for the expression of recombinant proteins (e.g., COS, CHO, 293, 293T cells). More particularly, a polynucleotide comprising the coding sequence of 58P1D12 or a fragment, analog or homolog thereof can be used to generate 58P1D12 proteins or fragments thereof using any number of host-vector systems routinely used and widely known in the art.

A wide range of host-vector systems suitable for the expression of 58P1D12 proteins or fragments thereof are available, see for example, Sambrook et al., 1989, supra; Current Protocols in Molecular Biology, 1995, supra). Preferred vectors for mammalian expression include but are not limited to pcDNA 3.1 myc-His-tag (Invitrogen) and the retroviral vector pSRαtkneo (Muller et al., 1991, MCB 11:1785). Using these expression vectors, 58P1D12 can be expressed in several prostate cancer and non-prostate cell lines, including for example 293, 293T, rat-1, NIH 3T3 and TsuPr1. The host-vector systems of the invention are useful for the production of a 58P1D12 protein or fragment thereof. Such host-vector systems can be employed to study the functional properties of 58P1D12 and 58P1D12 mutations or analogs.

Recombinant human 58P1D12 protein or an analog or homolog or fragment thereof can be produced by mammalian cells transfected with a construct encoding a 58P1D12-related nucleotide. For example, 293T cells can be transfected with an expression plasmid encoding 58P1D12 or fragment, analog or homolog thereof, a 58P1D12-related protein is expressed in the 293T cells, and the recombinant 58P1D12 protein is isolated using standard purification methods (e.g., affinity purification using anti-58P1D12 antibodies). In another embodiment, a 58P1D12 coding sequence is subcloned into the retroviral vector pSRαMSVtkneo and used to infect various mammalian cell lines, such as NIH 3T3, TsuPr1, 293 and rat-1 in order to establish 58P1D12 expressing cell lines. Various other expression systems well known in the art can also be employed. Expression constructs encoding a leader peptide joined in frame to a 58P1D12 coding sequence can be used for the generation of a secreted form of recombinant 58P1D12 protein.

As discussed herein, redundancy in the genetic code permits variation in 58P1D12 gene sequences. In particular, it is known in the art that specific host species often have specific codon preferences, and thus one can adapt the disclosed sequence as preferred for a desired host. For example, preferred analog codon sequences typically have rare codons (i.e., codons having a usage frequency of less than about 20% in known sequences of the desired host) replaced with higher frequency codons. Codon preferences for a specific species are calculated, for example, by utilizing codon usage tables available on the INTERNET such as at URL dna.affrc.go.jp/~nakamura/codon.html.

Additional sequence modifications are known to enhance protein expression in a cellular host. These include elimination of sequences encoding spurious polyadenylation signals, exon/intron splice site signals, transposon-like repeats, and/or other such well-characterized sequences that are deleterious to gene expression. The GC content of the sequence is adjusted to levels average for a given cellular host, as calculated by reference to known genes expressed in the host cell. Where possible, the sequence is modified to avoid predicted hairpin secondary mRNA structures. Other useful modifications include the addition of a translational initiation consensus sequence at the start of the open reading frame, as described in Kozak, Mol. Cell Biol., 9:5073-5080 (1989). Skilled artisans understand that the general rule that eukaryotic ribosomes initiate translation exclusively at the 5' proximal AUG codon is abrogated only under rare conditions (see, e.g., Kozak PNAS 92(7): 2662-2666, (1995) and Kozak NAR 15(20): 8125-8148 (1987)).

III.) 58P1D12-Related Proteins

Another aspect of the present invention provides 58P1D12-related proteins. Specific embodiments of 58P1D12 proteins comprise a polypeptide having all or part of the amino acid sequence of human 58P1D12 as shown in FIG. **2** or FIG. **3**. Alternatively, embodiments of 58P1D12 proteins comprise variant, homolog or analog polypeptides that have alterations in the amino acid sequence of 58P1D12 shown in FIG. **2** or FIG. **3**.

Embodiments of a 58P1D12 polypeptide include: a 58P1D12 polypeptide having a sequence shown in FIG. **2**, a peptide sequence of a 58P1D12 as shown in FIG. **2** wherein T is U; at least 10 contiguous nucleotides of a polypeptide having the sequence as shown in FIG. **2**; or, at least 10 contiguous peptides of a polypeptide having the sequence as shown in FIG. **2** where T is U. For example, embodiments of 58P1D12 peptides comprise, without limitation:

(I) a protein comprising, consisting essentially of, or consisting of an amino acid sequence as shown in FIG. **2**A-F or FIGS. **3**A-C;

(II) a 58P1D12-related protein that is at least 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% homologous to an entire amino acid sequence shown in FIG. **2**A-F or **3**A-C;

(III) a 58P1D12-related protein that is at least 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% identical to an entire amino acid sequence shown in FIG. **2**A-F or **3**A-C;

(IV) a protein that comprises at least one peptide set forth in Tables VIII to XLIX, optionally with a proviso that it is not an entire protein of FIG. **2**;

(V) a protein that comprises at least one peptide set forth in Tables VIII-XXI, collectively, which peptide is also set forth in Tables XXII to XLIX, collectively, optionally with a proviso that it is not an entire protein of FIG. **2**;

(VI) a protein that comprises at least two peptides selected from the peptides set forth in Tables VIII-XLIX, optionally with a proviso that it is not an entire protein of FIG. **2**;

(VII) a protein that comprises at least two peptides selected from the peptides set forth in Tables VIII to XLIX collectively, with a proviso that the protein is not a contiguous sequence from an amino acid sequence of FIG. **2**;

(VIII) a protein that comprises at least one peptide selected from the peptides set forth in Tables VIII-XXI; and at least one peptide selected from the peptides set forth in Tables XXII to XLIX, with a proviso that the protein is not a contiguous sequence from an amino acid sequence of FIG. **2**;

(IX) a polypeptide comprising at least 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35 amino acids of a protein of FIGS. **3**A, **3**B, and/or **3**C in any whole number increment up to 273, 232, and/or 236 respectively that includes at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35 amino acid position(s) having a value greater than 0.5 in the Hydrophilicity profile of FIG. **5**;

(X) a polypeptide comprising at least 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35 amino acids of a protein of FIGS. **3**A, **3**B, and/or **3**C, in any whole number increment up to 273, 236, and/or 232 respectively that includes at least at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35 amino acid position(s) having a value less than 0.5 in the Hydropathicity profile of FIG. **6**;

(XI) a polypeptide comprising at least 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35 amino acids of a protein of FIGS. **3**A, **3**B and/or **3**C, in any whole number increment up to 273, 232 and/or 236 respectively that includes at least at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35 amino acid position(s) having a value greater than 0.5 in the Percent Accessible Residues profile of FIG. **7**;

(XII) a polypeptide comprising at least 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35 amino acids of a protein of FIGS. **3**A, **3**B and/or **3**C, in any whole number increment up to 273, 232 and/or 236 respectively that includes at least at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35 amino acid position(s) having a value greater than 0.5 in the Average Flexibility profile of FIG. **8**;

(XIII) a polypeptide comprising at least 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, amino acids of a protein of FIGS. **3**A, **3**B, and/or **3**C in any whole number increment up to 272, 232 and/or 236 respectively that includes at least at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35 amino acid position(s) having a value greater than 0.5 in the Beta-turn profile of FIG. **9**;

(XIV) a peptide that occurs at least twice in Tables VIII-XXI and XXII to XLIX, collectively;

(XV) a peptide that occurs at least three times in Tables VIII-XXI and XXII to XLIX, collectively;

(XVI) a peptide that occurs at least four times in Tables VIII-XXI and XXII to XLIX, collectively;

(XVII) a peptide that occurs at least five times in Tables VIII-XXI and XXII to XLIX, collectively;

(XVIII) a peptide that occurs at least once in Tables VIII-XXI, and at least once in tables XXII to XLIX;

(XIX) a peptide that occurs at least once in Tables VIII-XXI, and at least twice in tables XXII to XLIX;

(XX) a peptide that occurs at least twice in Tables VIII-XXI, and at least once in tables XXII to XLIX;

(XXI) a peptide that occurs at least twice in Tables VIII-XXI, and at least twice in tables XXII to XLIX;

(XXII) a peptide which comprises one two, three, four, or five of the following characteristics, or an oligonucleotide encoding such peptide:

i) a region of at least 5 amino acids of a particular peptide of FIG. **3**, in any whole number increment up to the full length of that protein in FIG. **3**, that includes an amino acid position having a value equal to or greater than 0.5, 0.6, 0.7, 0.8, 0.9, or having a value equal to 1.0, in the Hydrophilicity profile of FIG. **5**;

ii) a region of at least 5 amino acids of a particular peptide of FIG. **3**, in any whole number increment up to the full length of that protein in FIG. **3**, that includes an amino acid position having a value equal to or less than 0.5, 0.4, 0.3, 0.2, 0.1, or having a value equal to 0.0, in the Hydropathicity profile of FIG. **6**;

iii) a region of at least 5 amino acids of a particular peptide of FIG. **3**, in any whole number increment up to the full length of that protein in FIG. **3**, that includes an amino acid position having a value equal to or greater than 0.5, 0.6, 0.7, 0.8, 0.9, or having a value equal to 1.0, in the Percent Accessible Residues profile of FIG. **7**;

iv) a region of at least 5 amino acids of a particular peptide of FIG. **3**, in any whole number increment up to the full length of that protein in FIG. **3**, that includes an amino acid position having a value equal to or greater than 0.5, 0.6, 0.7, 0.8, 0.9, or having a value equal to 1.0, in the Average Flexibility profile of FIG. **8**; or,

v) a region of at least 5 amino acids of a particular peptide of FIG. **3**, in any whole number increment up to the full length of that protein in FIG. **3**, that includes an amino acid position having a value equal to or greater than 0.5, 0.6, 0.7, 0.8, 0.9, or having a value equal to 1.0, in the Beta-turn profile of FIG. **9**;

(XXIII) a composition comprising a peptide of (I)-(XXII) or an antibody or binding region thereof together with a pharmaceutical excipient and/or in a human unit dose form.

(XXIV) a method of using a peptide of (I)-(XXII), or an antibody or binding region thereof or a composition of (XXIII) in a method to modulate a cell expressing 58P1D12;

(XXV) a method of using a peptide of (I)-(XXII) or an antibody or binding region thereof or a composition of (XXIII) in a method to diagnose, prophylax, prognose, or treat an individual who bears a cell expressing 58P1D12;

(XXVI) a method of using a peptide of (I)-(XXII) or an antibody or binding region thereof or a composition (XXIII) in a method to diagnose, prophylax, prognose, or treat an individual who bears a cell expressing 58P1D12, said cell from a cancer of a tissue listed in Table I;

(XXVII) a method of using a peptide of (I)-(XXII) or an antibody or binding region thereof or a composition of (XXIII) in a method to diagnose, prophylax, prognose, or treat a cancer;

(XXVIII) a method of using a peptide of (I)-(XXII) or an antibody or binding region thereof or a composition of (XXIII) in a method to diagnose, prophylax, prognose, or treat a a cancer of a tissue listed in Table I;

(XXIX) a method of using a a peptide of (I)-(XXII) or an antibody or binding region thereof or a composition and;

(XXIII) in a method to identify or characterize a modulator of a cell expressing 58P1D12.

As used herein, a range is understood to specifically disclose all whole unit positions thereof.

Typical embodiments of the invention disclosed herein include 58P1D12 polynucleotides that encode specific portions of 58P1D12 mRNA sequences (and those which are complementary to such sequences) such as those that encode the proteins and/or fragments thereof, for example:

(a) 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 225, 250, 255, 260, 265, 270 and 273 or more contiguous amino acids of 58P1D12 variant 1; the maximal lengths relevant for other variants are: variant 2, 232 amino acids; and variant 4, 236 amino acids.

In general, naturally occurring allelic variants of human 58P1D12 share a high degree of structural identity and homology (e.g., 90% or more homology). Typically, allelic variants of a 58P1D12 protein contain conservative amino acid substitutions within the 58P1D12 sequences described herein or contain a substitution of an amino acid from a corresponding position in a homologue of 58P1D12. One class of 58P1D12 allelic variants are proteins that share a high degree of homology with at least a small region of a particular 58P1D12 amino acid sequence, but further contain a radical departure from the sequence, such as a non-conservative substitution, truncation, insertion or frame shift. In comparisons of protein sequences, the terms, similarity, identity, and homology each have a distinct meaning as appreciated in the field of genetics. Moreover, orthology and paralogy can be important concepts describing the relationship of members of a given protein family in one organism to the members of the same family in other organisms.

Amino acid abbreviations are provided in Table II. Conservative amino acid substitutions can frequently be made in a protein without altering either the conformation or the function of the protein. Proteins of the invention can comprise 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 conservative substitutions. Such changes include substituting any of isoleucine (I), valine (V), and leucine (L) for any other of these hydrophobic amino acids; aspartic acid (D) for glutamic acid (E) and vice versa; glutamine (Q) for asparagine (N) and vice versa; and serine (S) for threonine (T) and vice versa. Other substitutions can also be considered conservative, depending on the environment of the particular amino acid and its role in the three-dimensional structure of the protein. For example, glycine (G) and alanine (A) can frequently be interchangeable, as can alanine (A) and valine (V). Methionine (M), which is relatively hydrophobic, can frequently be interchanged with leucine and isoleucine, and sometimes with valine. Lysine (K) and arginine (R) are frequently interchangeable in locations in which the significant feature of the amino acid residue is its charge and the differing pK's of these two amino acid residues are not significant. Still other changes can be considered "conservative" in particular environments (see, e.g. Table III herein; pages 13-15 "Biochemistry" 2nd ED. Lubert Stryer ed (Stanford University); Henikoff et al., PNAS 1992 Vol 89 10915-10919; Lei et al., J Biol Chem 1995 May 19; 270(20):11882-6).

Embodiments of the invention disclosed herein include a wide variety of art-accepted variants or analogs of 58P1D12 proteins such as polypeptides having amino acid insertions, deletions and substitutions. 58P1D12 variants can be made using methods known in the art such as site-directed mutagenesis, alanine scanning, and PCR mutagenesis. Site-directed mutagenesis (Carter et al., Nucl. Acids Res., 13:4331 (1986); Zoller et al., Nucl. Acids Res., 10:6487 (1987)), cassette mutagenesis (Wells et al., Gene, 34:315 (1985)), restriction selection mutagenesis (Wells et al., Philos. Trans. R. Soc. London SerA, 317:415 (1986)) or other known techniques can be performed on the cloned DNA to produce the 58P1D12 variant DNA.

Scanning amino acid analysis can also be employed to identify one or more amino acids along a contiguous sequence that is involved in a specific biological activity such as a protein-protein interaction. Among the preferred scanning amino acids are relatively small, neutral amino acids. Such amino acids include alanine, glycine, serine, and cysteine. Alanine is typically a preferred scanning amino acid among this group because it eliminates the side-chain beyond the beta-carbon and is less likely to alter the main-chain conformation of the variant. Alanine is also typically preferred because it is the most common amino acid. Further, it is frequently found in both buried and exposed positions (Creighton, The Proteins, (W.H. Freeman & Co., N.Y.); Chothia, J. Mol. Biol., 150:1 (1976)). If alanine substitution does not yield adequate amounts of variant, an isosteric amino acid can be used.

As defined herein, 58P1D12 variants, analogs or homologs, have the distinguishing attribute of having at least one epitope that is "cross reactive" with a 58P1D12 protein having an amino acid sequence of FIG. **3**. As used in this sentence, "cross reactive" means that an antibody or T cell that specifically binds to a 58P1D12 variant also specifically binds to a 58P1D12 protein having an amino acid sequence set forth in FIG. **3**. A polypeptide ceases to be a variant of a protein shown in FIG. **3**, when it no longer contains any epitope capable of being recognized by an antibody or T cell that specifically binds to the starting 58P1D12 protein. Those skilled in the art understand that antibodies that recognize proteins bind to epitopes of varying size, and a grouping of the order of about four or five amino acids, contiguous or not, is regarded as a typical number of amino acids in a minimal epitope. See, e.g., Nair et al., J. Immunol 2000 165(12): 6949-6955; Hebbes et al., Mol Immunol (1989) 26(9):865-73; Schwartz et al., J Immunol (1985) 135(4):2598-608.

Other classes of 58P1D12-related protein variants share 70%, 75%, 80%, 85% or 90% or more similarity with an amino acid sequence of FIG. **3**, or a fragment thereof. Another specific class of 58P1D12 protein variants or analogs comprises one or more of the 58P1D12 biological motifs described herein or presently known in the art. Thus, encompassed by the present invention are analogs of 58P1D12 fragments (nucleic or amino acid) that have altered functional (e.g. immunogenic) properties relative to the starting fragment. It is to be appreciated that motifs now or which become part of the art are to be applied to the nucleic or amino acid sequences of FIG. **2** or FIG. **3**.

As discussed herein, embodiments of the claimed invention include polypeptides containing less than the full amino acid sequence of a 58P1D12 protein shown in FIG. **2** or FIG. **3**. For example, representative embodiments of the invention comprise peptides/proteins having any 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 or more contiguous amino acids of a 58P1D12 protein shown in FIG. **2** or FIG. **3**.

Moreover, representative embodiments of the invention disclosed herein include polypeptides consisting of about amino acid 1 to about amino acid 10 of a 58P1D12 protein shown in FIG. **2** or FIG. **3**, polypeptides consisting of about amino acid 10 to about amino acid 20 of a 58P1D12 protein shown in FIG. **2** or FIG. **3**, polypeptides consisting of about amino acid 20 to about amino acid 30 of a 58P1D12 protein shown in FIG. **2** or FIG. **3**, polypeptides consisting of about amino acid 30 to about amino acid 40 of a 58P1D12 protein shown in FIG. **2** or FIG. **3**, polypeptides consisting of about amino acid 40 to about amino acid 50 of a 58P1D12 protein shown in FIG. **2** or FIG. **3**, polypeptides consisting of about amino acid 50 to about amino acid 60 of a 58P1D12 protein shown in FIG. **2** or FIG. **3**, polypeptides consisting of about amino acid 60 to about amino acid 70 of a 58P1D12 protein shown in FIG. **2** or FIG. **3**, polypeptides consisting of about amino acid 70 to about amino acid 80 of a 58P1D12 protein shown in FIG. **2** or FIG. **3**, polypeptides consisting of about amino acid 80 to about amino acid 90 of a 58P1D12 protein shown in FIG. **2** or FIG. **3**, polypeptides consisting of about amino acid 90 to about amino acid 100 of a 58P1D12 protein shown in FIG. **2** or FIG. **3**, etc. throughout the entirety of a 58P1D12 amino acid sequence. Moreover, polypeptides consisting of about amino acid 1 (or 20 or 30 or 40 etc.) to about amino acid 20, (or 130, or 140 or 150 etc.) of a 58P1D12 protein shown in FIG. **2** or FIG. **3** are embodiments of the invention. It is to be appreciated that the starting and stopping positions in this paragraph refer to the specified position as well as that position plus or minus 5 residues.

58P1D12-related proteins are generated using standard peptide synthesis technology or using chemical cleavage methods well known in the art. Alternatively, recombinant methods can be used to generate nucleic acid molecules that encode a 58P1D12-related protein. In one embodiment, nucleic acid molecules provide a means to generate defined fragments of a 58P1D12 protein (or variants, homologs or analogs thereof).

III.A.) Motif-Bearing Protein Embodiments

Additional illustrative embodiments of the invention disclosed herein include 58P1D12 polypeptides comprising the amino acid residues of one or more of the biological motifs contained within a 58P1D12 polypeptide sequence set forth in FIG. **2** or FIG. **3**. Various motifs are known in the art, and a protein can be evaluated for the presence of such motifs by a number of publicly available Internet sites (see, e.g., URL addresses: pfam.wustl.edu/; searchlauncher.bcm.tmc.edu/seq-search/struc-predict.html; psort.ims.u-tokyo.ac.jp/; cbs-.dtu.dk/; ebi.ac.uk/interpro/scan.html; expasy.ch/tools/scnpsit1.html; Epimatrix™ and Epimer™, Brown University, brown.edu/Research/TB-HIV_Lab/epimatrix/epimatrix.html; and BIMAS, bimas.dcrt.nih.gov/.).

Motif bearing subsequences of all 58P1D12 variant proteins are set forth and identified in Tables VIII-XXI and XXII-XLIX.

Table V sets forth several frequently occurring motifs based on pfam searches (see URL address pfam.wustl.edu/). The columns of Table V list (1) motif name abbreviation, (2) percent identity found amongst the different member of the motif family, (3) motif name or description and (4) most common function; location information is included if the motif is relevant for location.

Polypeptides comprising one or more of the 58P1D12 motifs discussed above are useful in elucidating the specific characteristics of a malignant phenotype in view of the observation that the 58P1D12 motifs discussed above are associated with growth dysregulation and because 58P1D12 is overexpressed in certain cancers (See, e.g., Table I). Casein kinase II, cAMP and camp-dependent protein kinase, and Protein Kinase C, for example, are enzymes known to be associated with the development of the malignant phenotype (see e.g. Chen et al., Lab Invest., 78(2): 165-174 (1998); Gaiddon et al., Endocrinology 136(10): 4331-4338 (1995); Hall et al., Nucleic Acids Research 24(6): 1119-1126 (1996); Peterziel et al., Oncogene 18(46): 6322-6329 (1999) and O'Brian, Oncol. Rep. 5(2): 305-309 (1998)). Moreover, both glycosylation and myristoylation are protein modifications also associated with cancer and cancer progression (see e.g. Dennis et al., Biochem. Biophys. Acta 1473(1):21-34 (1999); Raju et al., Exp. Cell Res. 235(1): 145-154 (1997)). Amidation is another protein modification also associated with cancer and cancer progression (see e.g. Treston et al., J. Natl. Cancer Inst. Monogr. (13): 169-175 (1992)).

In another embodiment, proteins of the invention comprise one or more of the immunoreactive epitopes identified in accordance with art-accepted methods, such as the peptides set forth in Tables VIII-XXI and XXII-XLIX. CTL epitopes can be determined using specific algorithms to identify peptides within a 58P1D12 protein that are capable of optimally binding to specified HLA alleles (e.g., Table IV; Epimatrix™ and Epimer™, Brown University, URL brown.edu/Research/TB-HIV_Lab/epimatrix/epimatrix.html; and BIMAS, URL bimas.dcrt.nih.gov/.) Moreover, processes for identifying peptides that have sufficient binding affinity for HLA molecules and which are correlated with being immunogenic epitopes, are well known in the art, and are carried out without undue experimentation. In addition, processes for identifying peptides that are immunogenic epitopes, are well known in the art, and are carried out without undue experimentation either in vitro or in vivo.

Also known in the art are principles for creating analogs of such epitopes in order to modulate immunogenicity. For example, one begins with an epitope that bears a CTL or HTL motif (see, e.g., the HLA Class I and HLA Class II motifs/supermotifs of Table IV). The epitope is analoged by substituting out an amino acid at one of the specified positions, and replacing it with another amino acid specified for that position. For example, on the basis of residues defined in Table IV, one can substitute out a deleterious residue in favor of any other residue, such as a preferred residue; substitute a less-preferred residue with a preferred residue; or substitute an originally-occurring preferred residue with another preferred residue. Substitutions can occur at primary anchor positions or at other positions in a peptide; see, e.g., Table IV.

A variety of references reflect the art regarding the identification and generation of epitopes in a protein of interest as well as analogs thereof. See, for example, WO 97/33602 to Chesnut et al.; Sette, Immunogenetics 1999 50(3-4): 201-212; Sette et al., J. Immunol. 2001 166(2): 1389-1397; Sidney et al., Hum. Immunol. 1997 58(1):12-20; Kondo et al., Immunogenetics 1997 45(4):249-258; Sidney et al., J. Immunol. 1996 157(8): 3480-90; and Falk et al., Nature 351:290-6 (1991); Hunt et al., Science 255:1261-3 (1992); Parker et al., J. Immunol. 149:3580-7 (1992); Parker et al., J. Immunol. 152:163-75 (1994)); Kast et al., 1994 152(8): 3904-12; Borras-Cuesta et al., Hum. Immunol. 2000 61(3): 266-278; Alexander et al., J. Immunol. 2000 164(3); 164(3): 1625-1633; Alexander et al., PMID: 7895164, UI: 95202582; O'Sullivan et al., J. Immunol. 1991 147(8): 2663-2669; Alexander et al., Immunity 1994 1(9): 751-761 and Alexander et al., Immunol. Res. 1998 18(2): 79-92.

Related embodiments of the invention include polypeptides comprising combinations of the different motifs set forth in Table VI, and/or, one or more of the predicted CTL epitopes of Tables VIII-XXI and XXII-XLIX, and/or, one or more of the predicted HTL epitopes of Tables XLVI-XLIX, and/or, one or more of the T cell binding motifs known in the art. Preferred embodiments contain no insertions, deletions or substitutions either within the motifs or within the intervening sequences of the polypeptides. In addition, embodiments which include a number of either N-terminal and/or C-terminal amino acid residues on either side of these motifs may be desirable (to, for example, include a greater portion of the polypeptide architecture in which the motif is located). Typically, the number of N-terminal and/or C-terminal amino acid residues on either side of a motif is between about 1 to about 100 amino acid residues, preferably 5 to about 50 amino acid residues.

58P1D12-related proteins are embodied in many forms, preferably in isolated form. A purified 58P1D12 protein molecule will be substantially free of other proteins or molecules that impair the binding of 58P1D12 to antibody, T cell or other ligand. The nature and degree of isolation and purification will depend on the intended use. Embodiments of a 58P1D12-related proteins include purified 58P1D12-related proteins and functional, soluble 58P1D12-related proteins. In one embodiment, a functional, soluble 58P1D12 protein or fragment thereof retains the ability to be bound by antibody, T cell or other ligand.

The invention also provides 58P1D12 proteins comprising biologically active fragments of a 58P1D12 amino acid sequence shown in FIG. **2** or FIG. **3**. Such proteins exhibit properties of the starting 58P1D12 protein, such as the ability to elicit the generation of antibodies that specifically bind an epitope associated with the starting 58P1D12 protein; to be bound by such antibodies; to elicit the activation of HTL or CTL; and/or, to be recognized by HTL or CTL that also specifically bind to the starting protein.

58P1D12-related polypeptides that contain particularly interesting structures can be predicted and/or identified using various analytical techniques well known in the art, including, for example, the methods of Chou-Fasman, Garnier-Robson, Kyte-Doolittle, Eisenberg, Karplus-Schultz or Jameson-Wolf analysis, or based on immunogenicity. Fragments that contain such structures are particularly useful in generating subunit-specific anti-58P1D12 antibodies or T cells or in identifying cellular factors that bind to 58P1D12. For example, hydrophilicity profiles can be generated, and immunogenic peptide fragments identified, using the method of Hopp, T. P. and Woods, K. R., 1981, Proc. Natl. Acad. Sci. U.S.A. 78:3824-3828. Hydropathicity profiles can be generated, and immunogenic peptide fragments identified, using the method of Kyte, J. and Doolittle, R. F., 1982, J. Mol. Biol. 157:105-132. Percent (%) Accessible Residues profiles can be generated, and immunogenic peptide fragments identified, using the method of Janin J., 1979, Nature 277:491-492. Average Flexibility profiles can be generated, and immunogenic peptide fragments identified, using the method of Bhaskaran R., Ponnuswamy P. K., 1988, Int. J. Pept. Protein Res. 32:242-255. Beta-turn profiles can be generated, and immunogenic peptide fragments identified, using the method of Deleage, G., Roux B., 1987, Protein Engineering 1:289-294.

CTL epitopes can be determined using specific algorithms to identify peptides within a 58P1D12 protein that are capable of optimally binding to specified HLA alleles (e.g., by using the SYFPEITHI site at World Wide Web URL syfpeithi.bmi-heidelberg.com/; the listings in Table IV(A)-(E); Epimatrix™ and Epimer™, Brown University, URL (brown.edu/Research/TB-HIV_Lab/epimatrix/epimatrix.html); and BIMAS, URL bimas.dcrt.nih.gov/). Illustrating this, peptide epitopes from 58P1D12 that are presented in the context of human MHC Class I molecules, e.g., HLA-A1, A2, A3, A11, A24, B7 and B35 were predicted (see, e.g., Tables VIII-XXI, XXII-XLIX). Specifically, the complete amino acid sequence of the 58P1D12 protein and relevant portions of other variants, i.e., for HLA Class I predictions 9 flanking residues on either side of a point mutation or exon juction, and for HLA Class II predictions 14 flanking residues on either side of a point mutation or exon junction corresponding to that variant, were entered into the HLA Peptide Motif Search algorithm found in the Bioinformatics and Molecular Analysis Section (BIMAS) web site listed above; in addition to the site SYFPEITHI, at URL syfpeithi.bmi-heidelberg.com/.

The HLA peptide motif search algorithm was developed by Dr. Ken Parker based on binding of specific peptide sequences in the groove of HLA Class I molecules, in particular HLA-A2 (see, e.g., Falk et al., Nature 351: 290-6 (1991); Hunt et al., Science 255:1261-3 (1992); Parker et al., J. Immunol. 149:3580-7 (1992); Parker et al., J. Immunol. 152:163-75 (1994)). This algorithm allows location and ranking of 8-mer, 9-mer, and 10-mer peptides from a complete protein sequence for predicted binding to HLA-A2 as well as numerous other HLA Class I molecules. Many HLA class I binding peptides are 8-, 9-, 10 or 11-mers. For example, for Class I HLA-A2, the epitopes preferably contain a leucine (L) or methionine (M) at position 2 and a valine (V) or leucine (L) at the C-terminus (see, e.g., Parker et al., J. Immunol. 149: 3580-7 (1992)). Selected results of 58P1D12 predicted binding peptides are shown in Tables VIII-XXI and XXII-XLIX herein. In Tables VIII-XXI and XXII-XLVII, selected candidates, 9-mers and 10-mers, for each family member are shown along with their location, the amino acid sequence of each specific peptide, and an estimated binding score. In Tables XLVI-XLIX, selected candidates, 15-mers, for each family member are shown along with their location, the amino acid sequence of each specific peptide, and an estimated binding score. The binding score corresponds to the estimated half time of dissociation of complexes containing the peptide at 37° C. at pH 6.5. Peptides with the highest binding score are predicted to be the most tightly bound to HLA Class I on the cell surface for the greatest period of time and thus represent the best immunogenic targets for T-cell recognition.

Actual binding of peptides to an HLA allele can be evaluated by stabilization of HLA expression on the antigen-processing defective cell line T2 (see, e.g., Xue et al., Prostate 30:73-8 (1997) and Peshwa et al., Prostate 36:129-38 (1998)). Immunogenicity of specific peptides can be evaluated in vitro by stimulation of CD8+ cytotoxic T lymphocytes (CTL) in the presence of antigen presenting cells such as dendritic cells.

It is to be appreciated that every epitope predicted by the BIMAS site, Epimer™ and Epimatrix™ sites, or specified by the HLA class I or class II motifs available in the art or which become part of the art such as set forth in Table IV (or determined using World Wide Web site URL syfpeithi.bmi-heidelberg.com/, or BIMAS, bimas.dcrt.nih.gov/) are to be "applied" to a 58P1D12 protein in accordance with the invention. As used in this context "applied" means that a 58P1D12 protein is evaluated, e.g., visually or by computer-based patterns finding methods, as appreciated by those of skill in the relevant art. Every subsequence of a 58P1D12 protein of 8, 9, 10, or 11 amino acid residues that bears an HLA Class I motif, or a subsequence of 9 or more amino acid residues that bear an HLA Class II motif are within the scope of the invention.

III.B.) Expression of 58P1D12-Related Proteins

In an embodiment described in the examples that follow, 58P1D12 can be conveniently expressed in cells (such as 293T cells) transfected with a commercially available expression vector such as a CMV-driven expression vector encoding 58P1D12 with a C-terminal 6×His and MYC tag (pcDNA3.1/mycHIS, Invitrogen or Tag5, GenHunter Corporation, Nashville Tenn.). The Tag5 vector provides an IgGK secretion signal that can be used to facilitate the production of a secreted 58P1D12 protein in transfected cells. The secreted HIS-tagged 58P1D12 in the culture media can be purified, e.g., using a nickel column using standard techniques.

III.C.) Modifications of 58P1D12-Related Proteins

Modifications of 58P1D12-related proteins such as covalent modifications are included within the scope of this invention. One type of covalent modification includes reacting targeted amino acid residues of a 58P1D12 polypeptide with an organic derivatizing agent that is capable of reacting with selected side chains or the N- or C-terminal residues of a 58P1D12 protein. Another type of covalent modification of a 58P1D12 polypeptide included within the scope of this invention comprises altering the native glycosylation pattern of a protein of the invention. Another type of covalent modification of 58P1D12 comprises linking a 58P1D12 polypeptide to one of a variety of nonproteinaceous polymers, e.g., polyethylene glycol (PEG), polypropylene glycol, or polyoxyalkylenes, in the manner set forth in U.S. Pat. Nos. 4,640,835; 4,496,689; 4,301,144; 4,670,417; 4,791,192 or 4,179,337.

The 58P1D12-related proteins of the present invention can also be modified to form a chimeric molecule comprising 58P1D12 fused to another, heterologous polypeptide or amino acid sequence. Such a chimeric molecule can be synthesized chemically or recombinantly. A chimeric molecule can have a protein of the invention fused to another tumor-associated antigen or fragment thereof. Alternatively, a protein in accordance with the invention can comprise a fusion of fragments of a 58P1D12 sequence (amino or nucleic acid) such that a molecule is created that is not, through its length, directly homologous to the amino or nucleic acid sequences shown in FIG. **2** or FIG. **3**. Such a chimeric molecule can comprise multiples of the same subsequence of 58P1D12. A chimeric molecule can comprise a fusion of a 58P1D12-related protein with a polyhistidine epitope tag, which provides an epitope to which immobilized nickel can selectively bind, with cytokines or with growth factors. The epitope tag is generally placed at the amino- or carboxyl-terminus of a 58P1D12 protein. In an alternative embodiment, the chimeric molecule can comprise a fusion of a 58P1D12-related protein with an immunoglobulin or a particular region of an immunoglobulin. For a bivalent form of the chimeric molecule (also referred to as an "immunoadhesin"), such a fusion could be to the Fc region of an IgG molecule. The Ig fusions preferably include the substitution of a soluble (transmembrane domain deleted or inactivated) form of a 58P1D12 polypeptide in place of at least one variable region within an Ig molecule. In a preferred embodiment, the immunoglobulin fusion includes the hinge, CH2 and CH3, or the hinge, CH1, CH2 and CH3 regions of an IgG1 molecule. For the production of immunoglobulin fusions see, e.g., U.S. Pat. No. 5,428,130 issued Jun. 27, 1995.

III.D.) Uses of 58P1D12-Related Proteins

The proteins of the invention have a number of different specific uses. As 58P1D12 is highly expressed in prostate and other cancers, 58P1D12-related proteins are used in methods that assess the status of 58P1D12 gene products in normal versus cancerous tissues, thereby elucidating the malignant phenotype. Typically, polypeptides from specific regions of a 58P1D12 protein are used to assess the presence of perturbations (such as deletions, insertions, point mutations etc.) in those regions (such as regions containing one or more motifs). Exemplary assays utilize antibodies or T cells targeting 58P1D12-related proteins comprising the amino acid residues of one or more of the biological motifs contained within a 58P1D12 polypeptide sequence in order to evaluate the characteristics of this region in normal versus cancerous tissues or to elicit an immune response to the epitope. Alternatively, 58P1D12-related proteins that contain the amino acid residues of one or more of the biological motifs in a 58P1D12 protein are used to screen for factors that interact with that region of 58P1D12.

58P1D12 protein fragments/subsequences are particularly useful in generating and characterizing domain-specific antibodies (e.g., antibodies recognizing an extracellular or intracellular epitope of a 58P1D12 protein), for identifying agents or cellular factors that bind to 58P1D12 or a particular structural domain thereof, and in various therapeutic and diagnostic contexts, including but not limited to diagnostic assays, cancer vaccines and methods of preparing such vaccines.

Proteins encoded by the 58P1D12 genes, or by analogs, homologs or fragments thereof, have a variety of uses, including but not limited to generating antibodies and in methods for identifying ligands and other agents and cellular constituents that bind to a 58P1D12 gene product. Antibodies raised against a 58P1D12 protein or fragment thereof are useful in diagnostic and prognostic assays, and imaging methodologies in the management of human cancers characterized by expression of 58P1D12 protein, such as those listed in Table 1. Such antibodies can be expressed intracellularly and used in methods of treating patients with such cancers. 58P1D12-related nucleic acids or proteins are also used in generating HTL or CTL responses.

Various immunological assays useful for the detection of 58P1D12 proteins are used, including but not limited to various types of radioimmunoassays, enzyme-linked immunosorbent assays (ELISA), enzyme-linked immunofluorescent assays (ELIFA), immunocytochemical methods, and the like. Antibodies can be labeled and used as immunological imaging reagents capable of detecting 58P1D12-expressing cells (e.g., in radioscintigraphic imaging methods). 58P1D12 proteins are also particularly useful in generating cancer vaccines, as further described herein.

IV.) 58P1D12 Antibodies

Another aspect of the invention provides antibodies that bind to 58P1D12-related proteins. Preferred antibodies specifically bind to a 58P1D12-related protein and do not bind (or bind weakly) to peptides or proteins that are not 58P1D12-related proteins under physiological conditions. In this context, examples of physiological conditions include: 1) phosphate buffered saline; 2) Tris-buffered saline containing 25 mM Tris and 150 mM NaCl; or normal saline (0.9% NaCl); 4) animal serum such as human serum; or, 5) a combination of any of 1) through 4); these reactions preferably taking place at pH 7.5, alternatively in a range of pH 7.0 to 8.0, or alternatively in a range of pH 6.5 to 8.5; also, these reactions taking place at a temperature between 4° C. to 37° C. For example, antibodies that bind 58P1D12 can bind 58P1D12-related proteins such as the homologs or analogs thereof.

58P1D12 antibodies of the invention are particularly useful in cancer (see, e.g., Table 1) diagnostic and prognostic assays, and imaging methodologies. Similarly, such antibodies are useful in the treatment, diagnosis, and/or prognosis of other cancers, to the extent 58P1D12 is also expressed or overexpressed in these other cancers. Moreover, intracellularly expressed antibodies (e.g., single chain antibodies) are therapeutically useful in treating cancers in which the expression of 58P1D12 is involved, such as advanced or metastatic prostate cancers.

The invention also provides various immunological assays useful for the detection and quantification of 58P1D12 and mutant 58P1D12-related proteins. Such assays can comprise one or more 58P1D12 antibodies capable of recognizing and binding a 58P1D12-related protein, as appropriate. These assays are performed within various immunological assay formats well known in the art, including but not limited to various types of radioimmunoassays, enzyme-linked immunosorbent assays (ELISA), enzyme-linked immunofluorescent assays (ELIFA), and the like.

Immunological non-antibody assays of the invention also comprise T cell immunogenicity assays (inhibitory or stimulatory) as well as major histocompatibility complex (MHC) binding assays.

In addition, immunological imaging methods capable of detecting prostate cancer and other cancers expressing 58P1D12 are also provided by the invention, including but not limited to radioscintigraphic imaging methods using labeled 58P1D12 antibodies. Such assays are clinically useful in the detection, monitoring, and prognosis of 58P1D12 expressing cancers such as prostate cancer.

58P1D12 antibodies are also used in methods for purifying a 58P1D12-related protein and for isolating 58P1D12 homologues and related molecules. For example, a method of purifying a 58P1D12-related protein comprises incubating a 58P1D12 antibody, which has been coupled to a solid matrix, with a lysate or other solution containing a 58P1D12-related protein under conditions that permit the 58P1D12 antibody to bind to the 58P1D12-related protein; washing the solid matrix to eliminate impurities; and eluting the 58P1D12-related protein from the coupled antibody. Other uses of 58P1D12 antibodies in accordance with the invention include generating anti-idiotypic antibodies that mimic a 58P1D12 protein.

Various methods for the preparation of antibodies are well known in the art. For example, antibodies can be prepared by immunizing a suitable mammalian host using a 58P1D12-related protein, peptide, or fragment, in isolated or immunoconjugated form (Antibodies: A Laboratory Manual, CSH Press, Eds., Harlow, and Lane (1988); Harlow, Antibodies, Cold Spring Harbor Press, NY (1989)). In addition, fusion proteins of 58P1D12 can also be used, such as a 58P1D12 GST-fusion protein. In a particular embodiment, a GST fusion protein comprising all or most of the amino acid sequence of FIG. **2** or FIG. **3** is produced, then used as an immunogen to generate appropriate antibodies. In another embodiment, a 58P1D12-related protein is synthesized and used as an immunogen.

In addition, naked DNA immunization techniques known in the art are used (with or without purified 58P1D12-related protein or 58P1D12 expressing cells) to generate an immune response to the encoded immunogen (for review, see Donnelly et al., 1997, Ann. Rev. Immunol. 15: 617-648).

The amino acid sequence of a 58P1D12 protein as shown in FIG. **2** or FIG. **3** can be analyzed to select specific regions of the 58P1D12 protein for generating antibodies. For example, hydrophobicity and hydrophilicity analyses of a 58P1D12 amino acid sequence are used to identify hydrophilic regions in the 58P1D12 structure. Regions of a 58P1D12 protein that show immunogenic structure, as well as other regions and domains, can readily be identified using various other methods known in the art, such as Chou-Fasman, Garnier-Robson, Kyte-Doolittle, Eisenberg, Karplus-Schultz or Jameson-Wolf analysis. Hydrophilicity profiles can be generated using the method of Hopp, T. P. and Woods, K. R., 1981, Proc. Natl. Acad. Sci. U.S.A. 78:3824-3828. Hydropathicity profiles can be generated using the method of Kyte, J. and Doolittle, R. F., 1982, J. Mol. Biol. 157:105-132. Percent (%) Accessible Residues profiles can be generated using the method of Janin J., 1979, Nature 277:491-492. Average Flexibility profiles can be generated using the method of Bhaskaran R., Ponnuswamy P. K., 1988, Int. J. Pept. Protein Res. 32:242-255. Beta-turn profiles can be generated using the method of Deleage, G., Roux B., 1987, Protein Engineering 1:289-294. Thus, each region identified by any of these programs or methods is within the scope of the present invention. Methods for the generation of 58P1D12 antibodies are further illustrated by way of the examples provided herein. Methods for preparing a protein or polypeptide for use as an immunogen are well known in the art. Also well known in the art are methods for preparing immunogenic conjugates of a protein with a carrier, such as BSA, KLH or other carrier protein. In some circumstances, direct conjugation using, for example, carbodiimide reagents are used; in other instances linking reagents such as those supplied by Pierce Chemical Co., Rockford, Ill., are effective. Administration of a 58P1D12 immunogen is often conducted by injection over a suitable time period and with use of a suitable adjuvant, as is understood in the art. During the immunization schedule, titers of antibodies can be taken to determine adequacy of antibody formation.

58P1D12 monoclonal antibodies can be produced by various means well known in the art. For example, immortalized cell lines that secrete a desired monoclonal antibody are prepared using the standard hybridoma technology of Kohler and Milstein or modifications that immortalize antibody-producing B cells, as is generally known. Immortalized cell lines that secrete the desired antibodies are screened by immunoassay in which the antigen is a 58P1D12-related protein. When the appropriate immortalized cell culture is identified, the cells can be expanded and antibodies produced either from in vitro cultures or from ascites fluid.

The antibodies or fragments of the invention can also be produced, by recombinant means. Regions that bind specifically to the desired regions of a 58P1D12 protein can also be produced in the context of chimeric or complementarity-determining region (CDR) grafted antibodies of multiple species origin. Humanized or human 58P1D12 antibodies can also be produced, and are preferred for use in therapeutic contexts. Methods for humanizing murine and other non-human antibodies, by substituting one or more of the non-human antibody CDRs for corresponding human antibody sequences, are well known (see for example, Jones et al., 1986, Nature 321: 522-525; Riechmann et al., 1988, Nature 332: 323-327; Verhoeyen et al., 1988, Science 239:1534-1536). See also, Carter et al., 1993, Proc. Natl. Acad. Sci. USA 89: 4285 and Sims et al., 1993, J. Immunol. 151: 2296.

Methods for producing fully human monoclonal antibodies include phage display and transgenic methods (for review, see Vaughan et al., 1998, Nature Biotechnology 16: 535-539). Fully human 58P1D12 monoclonal antibodies can be generated using cloning technologies employing large human Ig gene combinatorial libraries (i.e., phage display) (Griffiths and Hoogenboom, Building an in vitro immune system: human antibodies from phage display libraries. In: Protein Engineering of Antibody Molecules for Prophylactic and Therapeutic Applications in Man, Clark, M. (Ed.), Nottingham Academic, pp 45-64 (1993); Burton and Barbas, Human Antibodies from combinatorial libraries. Id., pp 65-82). Fully human 58P1D12 monoclonal antibodies can also be produced using transgenic mice engineered to contain human immunoglobulin gene loci as described in PCT Patent Application WO98/24893, Kucherlapati and Jakobovits et al., published Dec. 3, 1997 (see also, Jakobovits, 1998, Exp. Opin. Invest. Drugs 7(4): 607-614; U.S. Pat. No. 6,162,963 issued 19 Dec. 2000; U.S. Pat. No. 6,150,584 issued 12 Nov. 2000; and, U.S. Pat. No. 6,114,598 issued 5 Sep. 2000). This method avoids the in vitro manipulation required with phage display technology and efficiently produces high affinity authentic human antibodies.

Reactivity of 58P1D12 antibodies with a 58P1D12-related protein can be established by a number of well known means, including Western blot, immunoprecipitation, ELISA, and FACS analyses using, as appropriate, 58P1D12-related proteins, 58P1D12-expressing cells or extracts thereof. A 58P1D12 antibody or fragment thereof can be labeled with a detectable marker or conjugated to a second molecule. Suitable detectable markers include, but are not limited to, a radioisotope, a fluorescent compound, a bioluminescent compound, chemiluminescent compound, a metal chelator or an enzyme. Further, bi-specific antibodies specific for two or more 58P1D12 epitopes are generated using methods generally known in the art. Homodimeric antibodies can also be generated by cross-linking techniques known in the art (e.g., Wolff et al., Cancer Res. 53: 2560-2565).

V.) 58P1D12 Cellular Immune Responses

The mechanism by which T cells recognize antigens has been delineated. Efficacious peptide epitope vaccine compositions of the invention induce a therapeutic or prophylactic immune responses in very broad segments of the world-wide population. For an understanding of the value and efficacy of compositions of the invention that induce cellular immune responses, a brief review of immunology-related technology is provided.

A complex of an HLA molecule and a peptidic antigen acts as the ligand recognized by HLA-restricted T cells (Buus, S. et al., Cell 47:1071, 1986; Babbitt, B. P. et al., Nature 317: 359, 1985; Townsend, A. and Bodmer, H., Annu. Rev. Immunol. 7:601, 1989; Germain, R. N., Annu. Rev. Immunol. 11:403, 1993). Through the study of single amino acid substituted antigen analogs and the sequencing of endogenously bound, naturally processed peptides, critical residues that correspond to motifs required for specific binding to HLA antigen molecules have been identified and are set forth in Table IV (see also, e.g., Southwood, et al., J. Immunol. 160: 3363, 1998; Rammensee, et al., Immunogenetics 41:178, 1995; Rammensee et al., SYFPEITHI, access via World Wide Web at URL (134.2.96.221/scripts.hlaserver.dll/home.htm); Sette, A. and Sidney, J. Curr. Opin. Immunol. 10:478, 1998; Engelhard, V. H., Curr. Opin. Immunol. 6:13, 1994; Sette, A. and Grey, H. M., Curr. Opin. Immunol. 4:79, 1992; Sinigaglia, F. and Hammer, J. Curr. Biol. 6:52, 1994; Ruppert et al., Cell 74:929-937, 1993; Kondo et al., J. Immunol. 155: 4307-4312, 1995; Sidney et al., J. Immunol. 157:3480-3490, 1996; Sidney et al., Human Immunol. 45:79-93, 1996; Sette, A. and Sidney, J. Immunogenetics 1999 November; 50(3-4): 201-12, Review).

Furthermore, x-ray crystallographic analyses of HLA-peptide complexes have revealed pockets within the peptide binding cleft/groove of HLA molecules which accommodate, in an allele-specific mode, residues borne by peptide ligands; these residues in turn determine the HLA binding capacity of the peptides in which they are present. (See, e.g., Madden, D. R. Annu. Rev. Immunol. 13:587, 1995; Smith, et al., Immunity 4:203, 1996; Fremont et al., Immunity 8:305, 1998; Stern et al., Structure 2:245, 1994; Jones, E. Y. Curr. Opin. Immunol. 9:75, 1997; Brown, J. H. et al., Nature 364:33, 1993; Guo, H. C. et al., Proc. Natl. Acad. Sci. USA 90:8053, 1993; Guo, H. C. et al., Nature 360:364, 1992; Silver, M. L. et al., Nature 360:367, 1992; Matsumura, M. et al., Science 257:927, 1992; Madden et al., Cell 70:1035, 1992; Fremont, D. H. et al., Science 257:919, 1992; Saper, M. A., Bjorkman, P. J. and Wiley, D. C., J. Mol. Biol. 219:277, 1991.)

Accordingly, the definition of class I and class II allele-specific HLA binding motifs, or class I or class II supermotifs allows identification of regions within a protein that are correlated with binding to particular HLA antigen(s).

Thus, by a process of HLA motif identification, candidates for epitope-based vaccines have been identified; such candidates can be further evaluated by HLA-peptide binding assays to determine binding affinity and/or the time period of association of the epitope and its corresponding HLA molecule. Additional confirmatory work can be performed to select, amongst these vaccine candidates, epitopes with preferred characteristics in terms of population coverage, and/or immunogenicity.

Various strategies can be utilized to evaluate cellular immunogenicity, including:

1) Evaluation of primary T cell cultures from normal individuals (see, e.g., Wentworth, P. A. et al., Mol. Immunol. 32:603, 1995; Celis, E. et al., Proc. Natl. Acad. Sci. USA 91:2105, 1994; Tsai, V. et al., J. Immunol. 158:1796, 1997; Kawashima, I. et al., Human Immunol. 59:1, 1998). This procedure involves the stimulation of peripheral blood lymphocytes (PBL) from normal subjects with a test peptide in the presence of antigen presenting cells in vitro over a period of several weeks. T cells specific for the peptide become activated during this time and are detected using, e.g., a lymphokine- or 51Cr-release assay involving peptide sensitized target cells.

2) Immunization of HLA transgenic mice (see, e.g., Wentworth, P. A. et al., J. Immunol. 26:97, 1996; Wentworth, P. A. et al., Int. Immunol. 8:651, 1996; Alexander, J. et al., J. Immunol. 159:4753, 1997). For example, in such methods peptides in incomplete Freund's adjuvant are administered subcutaneously to HLA transgenic mice. Several weeks following immunization, splenocytes are removed and cultured in vitro in the presence of test peptide for approximately one week. Peptide-specific T cells are detected using, e.g., a 51Cr-release assay involving peptide sensitized target cells and target cells expressing endogenously generated antigen.

3) Demonstration of recall T cell responses from immune individuals who have been either effectively vaccinated and/or from chronically ill patients (see, e.g., Rehermann, B. et al., J. Exp. Med. 181:1047, 1995; Doolan, D. L. et al., Immunity 7:97, 1997; Bertoni, R. et al., J. Clin. Invest. 100:503, 1997; Threlkeld, S. C. et al., J. Immunol. 159:1648, 1997; Diepolder, H. M. et al., J. Virol. 71:6011, 1997). Accordingly, recall responses are detected by culturing PBL from subjects that have been exposed to the antigen due to disease and thus have generated an immune response "naturally", or from patients who were vaccinated against the antigen. PBL from subjects are cultured in vitro for 1-2 weeks in the presence of test peptide plus antigen presenting cells (APC) to allow activation of "memory" T cells, as compared to "naive" T cells. At the end of the culture period, T cell activity is detected using assays including 51Cr release involving peptide-sensitized targets, T cell proliferation, or lymphokine release.

VI.) 58P1D12 Transgenic Animals

Nucleic acids that encode a 58P1D12-related protein can also be used to generate either transgenic animals or "knock out" animals that, in turn, are useful in the development and screening of therapeutically useful reagents. In accordance with established techniques, cDNA encoding 58P1D12 can be used to clone genomic DNA that encodes 58P1D12. The cloned genomic sequences can then be used to generate transgenic animals containing cells that express DNA that encode 58P1D12. Methods for generating transgenic animals, particularly animals such as mice or rats, have become conventional in the art and are described, for example, in U.S. Pat.

Nos. 4,736,866 issued 12 Apr. 1988, and 4,870,009 issued 26 Sep. 1989. Typically, particular cells would be targeted for 58P1D12 transgene incorporation with tissue-specific enhancers.

Transgenic animals that include a copy of a transgene encoding 58P1D12 can be used to examine the effect of increased expression of DNA that encodes 58P1D12. Such animals can be used as tester animals for reagents thought to confer protection from, for example, pathological conditions associated with its overexpression. In accordance with this aspect of the invention, an animal is treated with a reagent and a reduced incidence of a pathological condition, compared to untreated animals that bear the transgene, would indicate a potential therapeutic intervention for the pathological condition.

Alternatively, non-human homologues of 58P1D12 can be used to construct a 58P1D12 "knock out" animal that has a defective or altered gene encoding 58P1D12 as a result of homologous recombination between the endogenous gene encoding 58P1D12 and altered genomic DNA encoding 58P1D12 introduced into an embryonic cell of the animal. For example, cDNA that encodes 58P1D12 can be used to clone genomic DNA encoding 58P1D12 in accordance with established techniques. A portion of the genomic DNA encoding 58P1D12 can be deleted or replaced with another gene, such as a gene encoding a selectable marker that can be used to monitor integration. Typically, several kilobases of unaltered flanking DNA (both at the 5' and 3' ends) are included in the vector (see, e.g., Thomas and Capecchi, Cell, 51:503 (1987) for a description of homologous recombination vectors). The vector is introduced into an embryonic stem cell line (e.g., by electroporation) and cells in which the introduced DNA has homologously recombined with the endogenous DNA are selected (see, e.g., Li et al., Cell, 69:915 (1992)). The selected cells are then injected into a blastocyst of an animal (e.g., a mouse or rat) to form aggregation chimeras (see, e.g., Bradley, in Teratocarcinomas and Embryonic Stem Cells: A Practical Approach, E. J. Robertson, ed. (IRL, Oxford, 1987), pp. 113-152). A chimeric embryo can then be implanted into a suitable pseudopregnant female foster animal, and the embryo brought to term to create a "knock out" animal. Progeny harboring the homologously recombined DNA in their germ cells can be identified by standard techniques and used to breed animals in which all cells of the animal contain the homologously recombined DNA. Knock out animals can be characterized, for example, for their ability to defend against certain pathological conditions or for their development of pathological conditions due to absence of a 58P1D12 polypeptide.

VII.) Methods for the Detection of 58P1D12

Another aspect of the present invention relates to methods for detecting 58P1D12 polynucleotides and 58P1D12-related proteins, as well as methods for identifying a cell that expresses 58P1D12. The expression profile of 58P1D12 makes it a diagnostic marker for metastasized disease. Accordingly, the status of 58P1D12 gene products provides information useful for predicting a variety of factors including susceptibility to advanced stage disease, rate of progression, and/or tumor aggressiveness. As discussed in detail herein, the status of 58P1D12 gene products in patient samples can be analyzed by a variety protocols that are well known in the art including immunohistochemical analysis, the variety of Northern blotting techniques including in situ hybridization, RT-PCR analysis (for example on laser capture micro-dissected samples), Western blot analysis and tissue array analysis.

More particularly, the invention provides assays for the detection of 58P1D12 polynucleotides in a biological sample, such as serum, bone, prostate, and other tissues, urine, semen, cell preparations, and the like. Detectable 58P1D12 polynucleotides include, for example, a 58P1D12 gene or fragment thereof, 58P1D12 mRNA, alternative splice variant 58P1D12 mRNAs, and recombinant DNA or RNA molecules that contain a 58P1D12 polynucleotide. A number of methods for amplifying and/or detecting the presence of 58P1D12 polynucleotides are well known in the art and can be employed in the practice of this aspect of the invention.

In one embodiment, a method for detecting a 58P1D12 mRNA in a biological sample comp rises producing cDNA from the sample by reverse transcription using at least one primer; amplifying the cDNA so produced using a 58P1D12 polynucleotides as sense and antisense primers to amplify 58P1D12 cDNAs therein; and detecting the presence of the amplified 58P1D12 cDNA. Optionally, the sequence of the amplified 58P1D12 cDNA can be determined.

In another embodiment, a method of detecting a 58P1D12 gene in a biological sample comprises first isolating genomic DNA from the sample; amplifying the isolated genomic DNA using 58P1D12 polynucleotides as sense and antisense primers; and detecting the presence of the amplified 58P1D12 gene. Any number of appropriate sense and antisense probe combinations can be designed from a 58P1D12 nucleotide sequence (see, e.g., FIG. 2) and used for this purpose.

The invention also provides assays for detecting the presence of a 58P1D12 protein in a tissue or other biological sample such as serum, semen, bone, prostate, urine, cell preparations, and the like. Methods for detecting a 58P1D12-related protein are also well known and include, for example, immunoprecipitation, immunohistochemical analysis, Western blot analysis, molecular binding assays, ELISA, ELIFA and the like. For example, a method of detecting the presence of a 58P1D12-related protein in a biological sample comprises first contacting the sample with a 58P1D12 antibody, a 58P1D12-reactive fragment thereof, or a recombinant protein containing an antigen-binding region of a 58P1D12 antibody; and then detecting the binding of 58P1D12-related protein in the sample.

Methods for identifying a cell that expresses 58P1D12 are also within the scope of the invention. In one embodiment, an assay for identifying a cell that expresses a 58P1D12 gene comprises detecting the presence of 58P1D12 mRNA in the cell. Methods for the detection of particular mRNAs in cells are well known and include, for example, hybridization assays using complementary DNA probes (such as in situ hybridization using labeled 58P1D12 riboprobes, Northern blot and related techniques) and various nucleic acid amplification assays (such as RT-PCR using complementary primers specific for 58P1D12, and other amplification type detection methods, such as, for example, branched DNA, SISBA, TMA and the like). Alternatively, an assay for identifying a cell that expresses a 58P1D12 gene comprises detecting the presence of 58P1D12-related protein in the cell or secreted by the cell. Various methods for the detection of proteins are well known in the art and are employed for the detection of 58P1D12-related proteins and cells that express 58P1D12-related proteins.

58P1D12 expression analysis is also useful as a tool for identifying and evaluating agents that modulate 58P1D12 gene expression. For example, 58P1D12 expression is significantly upregulated in prostate cancer, and is expressed in cancers of the tissues listed in Table I. Identification of a molecule or biological agent that inhibits 58P1D12 expression or over-expression in cancer cells is of therapeutic value.

For example, such an agent can be identified by using a screen that quantifies 58P1D12 expression by RT-PCR, nucleic acid hybridization or antibody binding.

VIII.) Methods for Monitoring the Status of 58P1D12-Related Genes and Their Products

Oncogenesis is known to be a multistep process where cellular growth becomes progressively dysregulated and cells progress from a normal physiological state to precancerous and then cancerous states (see, e.g., Alers et al., Lab Invest. 77(5): 437-438 (1997) and Isaacs et al., Cancer Surv. 23: 19-32 (1995)). In this context, examining a biological sample for evidence of dysregulated cell growth (such as aberrant 58P1D12 expression in cancers) allows for early detection of such aberrant physiology, before a pathologic state such as cancer has progressed to a stage that therapeutic options are more limited and or the prognosis is worse. In such examinations, the status of 58P1D12 in a biological sample of interest can be compared, for example, to the status of 58P1D12 in a corresponding normal sample (e.g. a sample from that individual or alternatively another individual that is not affected by a pathology). An alteration in the status of 58P1D12 in the biological sample (as compared to the normal sample) provides evidence of dysregulated cellular growth. In addition to using a biological sample that is not affected by a pathology as a normal sample, one can also use a predetermined normative value such as a predetermined normal level of mRNA expression (see, e.g., Grever et al., J. Comp. Neurol. 1996 Dec. 9; 376(2): 306-14 and U.S. Pat. No. 5,837,501) to compare 58P1D12 status in a sample.

The term “status” in this context is used according to its art accepted meaning and refers to the condition or state of a gene and its products. Typically, skilled artisans use a number of parameters to evaluate the condition or state of a gene and its products. These include, but are not limited to the location of expressed gene products (including the location of 58P1D12 expressing cells) as well as the level, and biological activity of expressed gene products (such as 58P1D12 mRNA, polynucleotides and polypeptides). Typically, an alteration in the status of 58P1D12 comprises a change in the location of 58P1D12 and/or 58P1D12 expressing cells and/or an increase in 58P1D12 mRNA and/or protein expression.

58P1D12 status in a sample can be analyzed by a number of means well known in the art, including without limitation, immunohistochemical analysis, in situ hybridization, RT-PCR analysis on laser capture micro-dissected samples, Western blot analysis, and tissue array analysis. Typical protocols for evaluating the status of a 58P1D12 gene and gene products are found, for example in Ausubel et al. eds., 1995, Current Protocols In Molecular Biology, Units 2 (Northern Blotting), 4 (Southern Blotting), 15 (Immunoblotting) and 18 (PCR Analysis). Thus, the status of 58P1D12 in a biological sample is evaluated by various methods utilized by skilled artisans including, but not limited to genomic Southern analysis (to examine, for example perturbations in a 58P1D12 gene), Northern analysis and/or PCR analysis of 58P1D12 mRNA (to examine, for example alterations in the polynucleotide sequences or expression levels of 58P1D12 mRNAs), and, Western and/or immunohistochemical analysis (to examine, for example alterations in polypeptide sequences, alterations in polypeptide localization within a sample, alterations in expression levels of 58P1D12 proteins and/or associations of 58P1D12 proteins with polypeptide binding partners). Detectable 58P1D12 polynucleotides include, for example, a 58P1D12 gene or fragment thereof, 58P1D12 mRNA, alternative splice variants, 58P1D12 mRNAs, and recombinant DNA or RNA molecules containing a 58P1D12 polynucleotide.

The expression profile of 58P1D12 makes it a diagnostic marker for local and/or metastasized disease, and provides information on the growth or oncogenic potential of a biological sample. In particular, the status of 58P1D12 provides information useful for predicting susceptibility to particular disease stages, progression, and/or tumor aggressiveness. The invention provides methods and assays for determining 58P1D12 status and diagnosing cancers that express 58P1D12, such as cancers of the tissues listed in Table I. For example, because 58P1D12 mRNA is so highly expressed in prostate and other cancers relative to normal prostate tissue, assays that evaluate the levels of 58P1D12 mRNA transcripts or proteins in a biological sample can be used to diagnose a disease associated with 58P1D12 dysregulation, and can provide prognostic information useful in defining appropriate therapeutic options.

The expression status of 58P1D12 provides information including the presence, stage and location of dysplastic, precancerous and cancerous cells, predicting susceptibility to various stages of disease, and/or for gauging tumor aggressiveness. Moreover, the expression profile makes it useful as an imaging reagent for metastasized disease. Consequently, an aspect of the invention is directed to the various molecular prognostic and diagnostic methods for examining the status of 58P1D12 in biological samples such as those from individuals suffering from, or suspected of suffering from a pathology characterized by dysregulated cellular growth, such as cancer.

As described above, the status of 58P1D12 in a biological sample can be examined by a number of well-known procedures in the art. For example, the status of 58P1D12 in a biological sample taken from a specific location in the body can be examined by evaluating the sample for the presence or absence of 58P1D12 expressing cells (e.g. those that express 58P1D12 mRNAs or proteins). This examination can provide evidence of dysregulated cellular growth, for example, when 58P1D12-expressing cells are found in a biological sample that does not normally contain such cells (such as a lymph node), because such alterations in the status of 58P1D12 in a biological sample are often associated with dysregulated cellular growth. Specifically, one indicator of dysregulated cellular growth is the metastases of cancer cells from an organ of origin (such as the prostate) to a different area of the body (such as a lymph node). In this context, evidence of dysregulated cellular growth is important for example because occult lymph node metastases can be detected in a substantial proportion of patients with prostate cancer, and such metastases are associated with known predictors of disease progression (see, e.g., Murphy et al., Prostate 42(4): 315-317 (2000); Su et al., Semin. Surg. Oncol. 18(1): 17-28 (2000) and Freeman et al., J Urol 1995 August 154(2 Pt 1):474-8).

In one aspect, the invention provides methods for monitoring 58P1D12 gene products by determining the status of 58P1D12 gene products expressed by cells from an individual suspected of having a disease associated with dysregulated cell growth (such as hyperplasia or cancer) and then comparing the status so determined to the status of 58P1D12 gene products in a corresponding normal sample. The presence of aberrant 58P1D12 gene products in the test sample relative to the normal sample provides an indication of the presence of dysregulated cell growth within the cells of the individual.

In another aspect, the invention provides assays useful in determining the presence of cancer in an individual, comprising detecting a significant increase in 58P1D12 mRNA or protein expression in a test cell or tissue sample relative to expression levels in the corresponding normal cell or tissue. The presence of 58P1D12 mRNA can, for example, be evaluated in tissues including but not limited to those listed in Table 1. The presence of significant 58P1D12 expression in any of these tissues is useful to indicate the emergence, presence and/or severity of a cancer, since the corresponding normal tissues do not express 58P1D12 mRNA or express it at lower levels.

In a related embodiment, 58P1D12 status is determined at the protein level rather than at the nucleic acid level. For example, such a method comprises determining the level of 58P1D12 protein expressed by cells in a test tissue sample and comparing the level so determined to the level of 58P1D12 expressed in a corresponding normal sample. In one embodiment, the presence of 58P1D12 protein is evaluated, for example, using immunohistochemical methods. 58P1D12 antibodies or binding partners capable of detecting 58P1D12 protein expression are used in a variety of assay formats well known in the art for this purpose.

In a further embodiment, one can evaluate the status of 58P1D12 nucleotide and amino acid sequences in a biological sample in order to identify perturbations in the structure of these molecules. These perturbations can include insertions, deletions, substitutions and the like. Such evaluations are useful because perturbations in the nucleotide and amino acid sequences are observed in a large number of proteins associated with a growth dysregulated phenotype (see, e.g., Marrogi et al., 1999, J. Cutan. Pathol. 26(8):369-378). For example, a mutation in the sequence of 58P1D12 may be indicative of the presence or promotion of a tumor. Such assays therefore have diagnostic and predictive value where a mutation in 58P1D12 indicates a potential loss of function or increase in tumor growth.

A wide variety of assays for observing perturbations in nucleotide and amino acid sequences are well known in the art. For example, the size and structure of nucleic acid or amino acid sequences of 58P1D12 gene products are observed by the Northern, Southern, Western, PCR and DNA sequencing protocols discussed herein. In addition, other methods for observing perturbations in nucleotide and amino acid sequences such as single strand conformation polymorphism analysis are well known in the art (see, e.g., U.S. Pat. Nos. 5,382,510 issued 7 Sep. 1999, and 5,952,170 issued 17 Jan. 1995).

Additionally, one can examine the methylation status of a 58P1D12 gene in a biological sample. Aberrant demethylation and/or hypermethylation of CpG islands in gene 5' regulatory regions frequently occurs in immortalized and transformed cells, and can result in altered expression of various genes. For example, promoter hypermethylation of the pi-class glutathione S-transferase (a protein expressed in normal prostate but not expressed in >90% of prostate carcinomas) appears to permanently silence transcription of this gene and is the most frequently detected genomic alteration in prostate carcinomas (De Marzo et al., Am. J. Pathol. 155(6): 1985-1992 (1999)). In addition, this alteration is present in at least 70% of cases of high-grade prostatic intraepithelial neoplasia (PIN) (Brooks et al., Cancer Epidemiol. Biomarkers Prev., 1998, 7:531-536). In another example, expression of the LAGE-1 tumor specific gene (which is not expressed in normal prostate but is expressed in 25-50% of prostate cancers) is induced by deoxy-azacytidine in lymphoblastoid cells, suggesting that tumoral expression is due to demethylation (Lethe et al., Int. J. Cancer 76(6): 903-908 (1998)). A variety of assays for examining methylation status of a gene are well known in the art. For example, one can utilize, in Southern hybridization approaches, methylation-sensitive restriction enzymes that cannot cleave sequences that contain methylated CpG sites to assess the methylation status of CpG islands. In addition, MSP (methylation specific PCR) can rapidly profile the methylation status of all the CpG sites present in a CpG island of a given gene. This procedure involves initial modification of DNA by sodium bisulfite (which will convert all unmethylated cytosines to uracil) followed by amplification using primers specific for methylated versus unmethylated DNA. Protocols involving methylation interference can also be found for example in Current Protocols In Molecular Biology, Unit 12, Frederick M. Ausubel et al. eds., 1995.

Gene amplification is an additional method for assessing the status of 58P1D12. Gene amplification is measured in a sample directly, for example, by conventional Southern blotting or Northern blotting to quantitate the transcription of mRNA (Thomas, 1980, Proc. Natl. Acad. Sci. USA, 77:5201-5205), dot blotting (DNA analysis), or in situ hybridization, using an appropriately labeled probe, based on the sequences provided herein. Alternatively, antibodies are employed that recognize specific duplexes, including DNA duplexes, RNA duplexes, and DNA-RNA hybrid duplexes or DNA-protein duplexes. The antibodies in turn are labeled and the assay carried out where the duplex is bound to a surface, so that upon the formation of duplex on the surface, the presence of antibody bound to the duplex can be detected.

Biopsied tissue or peripheral blood can be conveniently assayed for the presence of cancer cells using for example, Northern, dot blot or RT-PCR analysis to detect 58P1D12 expression. The presence of RT-PCR amplifiable 58P1D12 mRNA provides an indication of the presence of cancer. RT-PCR assays are well known in the art. RT-PCR detection assays for tumor cells in peripheral blood are currently being evaluated for use in the diagnosis and management of a number of human solid tumors. In the prostate cancer field, these include RT-PCR assays for the detection of cells expressing PSA and PSM (Verkaik et al., 1997, Urol. Res. 25:373-384; Ghossein et al., 1995, J. Clin. Oncol. 13:1195-2000; Heston et al., 1995, Clin. Chem. 41:1687-1688).

A further aspect of the invention is an assessment of the susceptibility that an individual has for developing cancer. In one embodiment, a method for predicting susceptibility to cancer comprises detecting 58P1D12 mRNA or 58P1D12 protein in a tissue sample, its presence indicating susceptibility to cancer, wherein the degree of 58P1D12 mRNA expression correlates to the degree of susceptibility. In a specific embodiment, the presence of 58P1D12 in prostate or other tissue is examined, with the presence of 58P1D12 in the sample providing an indication of prostate cancer susceptibility (or the emergence or existence of a prostate tumor). Similarly, one can evaluate the integrity 58P1D12 nucleotide and amino acid sequences in a biological sample, in order to identify perturbations in the structure of these molecules such as insertions, deletions, substitutions and the like. The presence of one or more perturbations in 58P1D12 gene products in the sample is an indication of cancer susceptibility (or the emergence or existence of a tumor).

The invention also comprises methods for gauging tumor aggressiveness. In one embodiment, a method for gauging aggressiveness of a tumor comprises determining the level of 58P1D12 mRNA or 58P1D12 protein expressed by tumor cells, comparing the level so determined to the level of 58P1D12 mRNA or 58P1D12 protein expressed in a corresponding normal tissue taken from the same individual or a normal tissue reference sample, wherein the degree of 58P1D12 mRNA or 58P1D12 protein expression in the tumor sample relative to the normal sample indicates the degree of aggressiveness. In a specific embodiment, aggressiveness of a tumor is evaluated by determining the extent to which 58P1D12 is expressed in the tumor cells, with higher expression levels indicating more aggressive tumors. Another embodiment is the evaluation of the integrity of 58P1D12 nucleotide and amino acid sequences in a biological sample, in order to identify perturbations in the structure of these molecules such as insertions, deletions, substitutions and the like. The presence of one or more perturbations indicates more aggressive tumors.

Another embodiment of the invention is directed to methods for observing the progression of a malignancy in an individual over time. In one embodiment, methods for observing the progression of a malignancy in an individual over time comprise determining the level of 58P1D12 mRNA or 58P1D12 protein expressed by cells in a sample of the tumor, comparing the level so determined to the level of 58P1D12 mRNA or 58P1D12 protein expressed in an equivalent tissue sample taken from the same individual at a different time, wherein the degree of 58P1D12 mRNA or 58P1D12 protein expression in the tumor sample over time provides information on the progression of the cancer. In a specific embodiment, the progression of a cancer is evaluated by determining 58P1D12 expression in the tumor cells over time, where increased expression overtime indicates a progression of the cancer. Also, one can evaluate the integrity 58P1D12 nucleotide and amino acid sequences in a biological sample in order to identify perturbations in the structure of these molecules such as insertions, deletions, substitutions and the like, where the presence of one or more perturbations indicates a progression of the cancer.

The above diagnostic approaches can be combined with any one of a wide variety of prognostic and diagnostic protocols known in the art. For example, another embodiment of the invention is directed to methods for observing a coincidence between the expression of 58P1D12 gene and 58P1D12 gene products (or perturbations in 58P1D12 gene and 58P1D12 gene products) and a factor that is associated with malignancy, as a means for diagnosing and prognosticating the status of a tissue sample. A wide variety of factors associated with malignancy can be utilized, such as the expression of genes associated with malignancy (e.g. PSA, PSCA and PSM expression for prostate cancer etc.) as well as gross cytological observations (see, e.g., Bocking et al., 1984, Anal. Quant. Cytol. 6(2):74-88; Epstein, 1995, Hum. Pathol. 26(2):223-9; Thorson et al., 1998, Mod. Pathol. 11(6):543-51; Baisden et al., 1999, Am. J. Surg. Pathol. 23(8):918-24). Methods for observing a coincidence between the expression of 58P1D12 gene and 58P1D12 gene products (or perturbations in 58P1D12 gene and 58P1D12 gene products) and another factor that is associated with malignancy are useful, for example, because the presence of a set of specific factors that coincide with disease provides information crucial for diagnosing and prognosticating the status of a tissue sample.

In one embodiment, methods for observing a coincidence between the expression of 58P1D12 gene and 58P1D12 gene products (or perturbations in 58P1D12 gene and 58P1D12 gene products) and another factor associated with malignancy entails detecting the overexpression of 58P1D12 mRNA or protein in a tissue sample, detecting the overexpression of PSA mRNA or protein in a tissue sample (or PSCA or PSM expression), and observing a coincidence of 58P1D12 mRNA or protein and PSA mRNA or protein overexpression (or PSCA or PSM expression). In a specific embodiment, the expression of 58P1D12 and PSA mRNA in prostate tissue is examined, where the coincidence of 58P1D12 and PSA mRNA overexpression in the sample indicates the existence of prostate cancer, prostate cancer susceptibility or the emergence or status of a prostate tumor.

Methods for detecting and quantifying the expression of 58P1D12 mRNA or protein are described herein, and standard nucleic acid and protein detection and quantification technologies are well known in the art. Standard methods for the detection and quantification of 58P1D12 mRNA include in situ hybridization using labeled 58P1D12 riboprobes, Northern blot and related techniques using 58P1D12 polynucleotide probes, RT-PCR analysis using primers specific for 58P1D12, and other amplification type detection methods, such as, for example, branched DNA, SISBA, TMA and the like. In a specific embodiment, semi-quantitative RT-PCR is used to detect and quantify 58P1D12 mRNA expression. Any number of primers capable of amplifying 58P1D12 can be used for this purpose, including but not limited to the various primer sets specifically described herein. In a specific embodiment, polyclonal or monoclonal antibodies specifically reactive with the wild-type 58P1D12 protein can be used in an immunohistochemical assay of biopsied tissue.

IX.) Identification of Molecules that Interact with 58P1D12

The 58P1D12 protein and nucleic acid sequences disclosed herein allow a skilled artisan to identify proteins, small molecules and other agents that interact with 58P1D12, as well as pathways activated by 58P1D12 via any one of a variety of art accepted protocols. For example, one can utilize one of the so-called interaction trap systems (also referred to as the "two-hybrid assay"). In such systems, molecules interact and reconstitute a transcription factor which directs expression of a reporter gene, whereupon the expression of the reporter gene is assayed. Other systems identify protein-protein interactions in vivo through reconstitution of a eukaryotic transcriptional activator, see, e.g., U.S. Pat. Nos. 5,955,280 issued 21 Sep. 1999, 5,925,523 issued 20 Jul. 1999, 5,846,722 issued 8 Dec. 1998 and 6,004,746 issued 21 Dec. 1999. Algorithms are also available in the art for genome-based predictions of protein function (see, e.g., Marcotte, et al., Nature 402: 4 Nov. 1999, 83-86).

Alternatively one can screen peptide libraries to identify molecules that interact with 58P1D12 protein sequences. In such methods, peptides that bind to 58P1D12 are identified by screening libraries that encode a random or controlled collection of amino acids. Peptides encoded by the libraries are expressed as fusion proteins of bacteriophage coat proteins, the bacteriophage particles are then screened against the 58P1D12 protein(s).

Accordingly, peptides having a wide variety of uses, such as therapeutic, prognostic or diagnostic reagents, are thus identified without any prior information on the structure of the expected ligand or receptor molecule. Typical peptide libraries and screening methods that can be used to identify molecules that interact with 58P1D12 protein sequences are disclosed for example in U.S. Pat. Nos. 5,723,286 issued 3 Mar. 1998 and 5,733,731 issued 31 Mar. 1998.

Alternatively, cell lines that express 58P1D12 are used to identify protein-protein interactions mediated by 58P1D12. Such interactions can be examined using immunoprecipitation techniques (see, e.g., Hamilton B. J., et al. Biochem. Biophys. Res. Commun. 1999, 261:646-51). 58P1D12 protein can be immunoprecipitated from 58P1D12-expressing cell lines using anti-58P1D12 antibodies. Alternatively, antibodies against His-tag can be used in a cell line engineered to express fusions of 58P1D12 and a His-tag (vectors mentioned above). The immunoprecipitated complex can be examined for protein association by procedures such as Western blotting, 35S-methionine labeling of proteins, protein microsequencing, silver staining and two-dimensional gel electrophoresis.

Small molecules and ligands that interact with 58P1D12 can be identified through related embodiments of such screening assays. For example, small molecules can be identified that interfere with protein function, including molecules that interfere with 58P1D12's ability to mediate phosphorylation and de-phosphorylation, interaction with DNA or RNA molecules as an indication of regulation of cell cycles, second messenger signaling or tumorigenesis. Similarly, small molecules that modulate 58P1D12-related ion channel, protein pump, or cell communication functions are identified and used to treat patients that have a cancer that expresses 58P1D12 (see, e.g., Hille, B., Ionic Channels of Excitable Membranes 2nd Ed., Sinauer Assoc., Sunderland, Mass., 1992). Moreover, ligands that regulate 58P1D12 function can be identified based on their ability to bind 58P1D12 and activate a reporter construct. Typical methods are discussed for example in U.S. Pat. No. 5,928,868 issued 27 Jul. 1999, and include methods for forming hybrid ligands in which at least one ligand is a small molecule. In an illustrative embodiment, cells engineered to express a fusion protein of 58P1D12 and a DNA-binding protein are used to co-express a fusion protein of a hybrid ligand/small molecule and a cDNA library transcriptional activator protein. The cells further contain a reporter gene, the expression of which is conditioned on the proximity of the first and second fusion proteins to each other, an event that occurs only if the hybrid ligand binds to target sites on both hybrid proteins. Those cells that express the reporter gene are selected and the unknown small molecule or the unknown ligand is identified. This method provides a means of identifying modulators, which activate or inhibit 58P1D12.

An embodiment of this invention comprises a method of screening for a molecule that interacts with a 58P1D12 amino acid sequence shown in FIG. **2** or FIG. **3**, comprising the steps of contacting a population of molecules with a 58P1D12 amino acid sequence, allowing the population of molecules and the 58P1D12 amino acid sequence to interact under conditions that facilitate an interaction, determining the presence of a molecule that interacts with the 58P1D12 amino acid sequence, and then separating molecules that do not interact with the 58P1D12 amino acid sequence from molecules that do. In a specific embodiment, the method further comprises purifying, characterizing and identifying a molecule that interacts with the 58P1D12 amino acid sequence. The identified molecule can be used to modulate a function performed by 58P1D12. In a preferred embodiment, the 58P1D12 amino acid sequence is contacted with a library of peptides.

X.) Therapeutic Methods and Compositions

The identification of 58P1D12 as a protein that is normally expressed in a restricted set of tissues, but which is also expressed in cancers such as those listed in Table I, opens a number of therapeutic approaches to the treatment of such cancers.

Of note, targeted antitumor therapies have been useful even when the targeted protein is expressed on normal tissues, even vital normal organ tissues. A vital organ is one that is necessary to sustain life, such as the heart or colon. A non-vital organ is one that can be removed whereupon the individual is still able to survive. Examples of non-vital organs are ovary, breast, and prostate.

For example, Herceptin® is an FDA approved pharmaceutical that has as its active ingredient an antibody which is immunoreactive with the protein variously known as HER2, HER2/neu, and erb-b-2. It is marketed by Genentech and has been a commercially successful antitumor agent. Herceptin sales reached almost $400 million in 2002. Herceptin is a treatment for HER2 positive metastatic breast cancer. However, the expression of HER2 is not limited to such tumors. The same protein is expressed in a number of normal tissues. In particular, it is known that HER2/neu is present in normal kidney and heart, thus these tissues are present in all human recipients of Herceptin. The presence of HER2/neu in normal kidney is also confirmed by Latif, Z., et al., B.J.U. International (2002) 89:5-9. As shown in this article (which evaluated whether renal cell carcinoma should be a preferred indication for anti-HER2 antibodies such as Herceptin) both protein and mRNA are produced in benign renal tissues. Notably, HER2/neu protein was strongly overexpressed in benign renal tissue.

Despite the fact that HER2/neu is expressed in such vital tissues as heart and kidney, Herceptin is a very useful, FDA approved, and commercially successful drug. The effect of Herceptin on cardiac tissue, i.e., "cardiotoxicity," has merely been a side effect to treatment. When patients were treated with Herceptin alone, significant cardiotoxicity occurred in a very low percentage of patients.

Of particular note, although kidney tissue is indicated to exhibit normal expression, possibly even higher expression than cardiac tissue, kidney has no appreciable Herceptin side effect whatsoever. Moreover, of the diverse array of normal tissues in which HER2 is expressed, there is very little occurrence of any side effect. Only cardiac tissue has manifested any appreciable side effect at all. A tissue such as kidney, where HER2/neu expression is especially notable, has not been the basis for any side effect.

Furthermore, favorable therapeutic effects have been found for antitumor therapies that target epidermal growth factor receptor (EGFR). EGFR is also expressed in numerous normal tissues. There have been very limited side effects in normal tissues following use of anti-EGFR therapeutics.

Thus, expression of a target protein in normal tissue, even vital normal tissue, does not defeat the utility of a targeting agent for the protein as a therapeutic for certain tumors in which the protein is also overexpressed.

Accordingly, therapeutic approaches that inhibit the activity of a 58P1D12 protein are useful for patients suffering from a cancer that expresses 58P1D12. These therapeutic approaches generally fall into two classes. One class comprises various methods for inhibiting the binding or association of a 58P1D12 protein with its binding partner or with other proteins. Another class comprises a variety of methods for inhibiting the transcription of a 58P1D12 gene or translation of 58P1D12 mRNA.

X.A.) Anti-Cancer Vaccines

The invention provides cancer vaccines comprising a 58P1D12-related protein or 58P1D12-related nucleic acid. In view of the expression of 58P1D12, cancer vaccines prevent and/or treat 58P1D12-expressing cancers with minimal or no effects on non-target tissues. The use of a tumor antigen in a vaccine that generates humoral and/or cell-mediated immune responses as anti-cancer therapy is well known in the art and has been employed in prostate cancer using human PSMA and rodent PAP immunogens (Hodge et al., 1995, Int. J. Cancer 63:231-237; Fong et al., 1997, J. Immunol. 159:3113-3117).

Such methods can be readily practiced by employing a 58P1D12-related protein, or a 58P1D12-encoding nucleic acid molecule and recombinant vectors capable of expressing and presenting the 58P1D12 immunogen (which typically comprises a number of antibody or T cell epitopes). Skilled artisans understand that a wide variety of vaccine systems for delivery of immunoreactive epitopes are known in the art (see, e.g., Heryln et al., Ann Med 1999 February 31(1):66-78; Maruyama et al., Cancer lmmunol lmmunother 2000 June 49(3):123-32) Briefly, such methods of generating an immune response (e.g. humoral and/or cell-mediated) in a mammal, comprise the steps of: exposing the mammal's immune system to an immunoreactive epitope (e.g. an epitope present in a 58P1D12 protein shown in or analog or homolog thereof) so that the mammal generates an immune response that is specific for that epitope (e.g. generates antibodies that specifically recognize that epitope). ln a preferred method, a 58P1D12 immunogen contains a biological motif, see e.g., Tables Vlll-XXl and XXll-XLlX, or a peptide of a size range from 58P1D12 indicated in FlG. **5**, FlG. **6**, FlG. **7**, FlG. **8**, and FlG. **9**.

The entire 58P1D12 protein, immunogenic regions or epitopes thereof can be combined and delivered by various means. Such vaccine compositions can include, for example, lipopeptides (e.g., Vitiello, A. et al., J. Clin. lnvest. 95:341, 1995), peptide compositions encapsulated in poly(DL-lactide-co-glycolide) ("PLG") microspheres (see, e.g., Eldridge, et al., Molec. lmmunol. 28:287-294, 1991: Alonso et al., Vaccine 12:299-306, 1994; Jones et al., Vaccine 13:675-681, 1995), peptide compositions contained in immune stimulating complexes (lSCOMS) (see, e.g., Takahashi et al., Nature 344:873-875, 1990; Hu et al., Clin Exp lmmunol. 113:235-243, 1998), multiple antigen peptide systems (MAPs) (see e.g., Tam, J. P., Proc. Natl. Acad. Sci. U.S.A. 85:5409-5413, 1988; Tam, J. P., J. lmmunol. Methods 196:17-32, 1996), peptides formulated as multivalent peptides; peptides for use in ballistic delivery systems, typically crystallized peptides, viral delivery vectors (Perkus, M. E. et al., ln: Concepts in vaccine development, Kaufmann, S. H. E., ed., p. 379, 1996; Chakrabarti, S. et al., Nature 320:535, 1986; Hu, S. L. et al., Nature 320:537, 1986; Kieny, M.-P. et al., AlDS Bio/Technology 4:790, 1986; Top, F. H. et al., J. lnfect. Dis. 124:148, 1971; Chanda, P. K. et al., Virology 175:535, 1990), particles of viral or synthetic origin (e.g., Kofler, N. et al., J. lmmunol. Methods. 192:25, 1996; Eldridge, J. H. et al., Sem. Hematol. 30:16, 1993; Falo, L. D., Jr. et al., Nature Med. 7:649, 1995), adjuvants (Warren, H. S., Vogel, F. R., and Chedid, L. A. Annu. Rev. lmmunol. 4:369, 1986; Gupta, R. K. et al., Vaccine 11:293, 1993), liposomes (Reddy, R. et al., J. lmmunol. 148:1585, 1992; Rock, K. L., lmmunol. Today 17:131, 1996), or, naked or particle absorbed cDNA (Ulmer, J. B. et al., Science 259:1745, 1993; Robinson, H. L., Hunt, L. A., and Webster, R. G., Vaccine 11:957, 1993; Shiver, J. W. et al., ln: Concepts in vaccine development, Kaufmann, S. H. E., ed., p. 423, 1996; Cease, K. B., and Berzofsky, J. A., Annu. Rev. lmmunol. 12:923, 1994 and Eldridge, J. H. et al., Sem. Hematol. 30:16, 1993). Toxin-targeted delivery technologies, also known as receptor mediated targeting, such as those of Avant lmmunotherapeutics, lnc. (Needham, Mass.) may also be used.

ln patients with 58P1D12-associated cancer, the vaccine compositions of the invention can also be used in conjunction with other treatments used for cancer, e.g., surgery, chemotherapy, drug therapies, radiation therapies, etc. including use in combination with immune adjuvants such as lL-2, lL-12, GM-CSF, and the like.

Cellular Vaccines:

CTL epitopes can be determined using specific algorithms to identify peptides within 58P1D12 protein that bind corresponding HLA alleles (see e.g., Table lV; Epimer™ and Epimatrix™, Brown University (URL brown.edu/Research/TB-HlV_Lab/epimatrix/epimatrix.html); and, BlMAS, (URL bimas.dcrt.nih.gov/; SYFPElTHl at URL syfpeithi.bmi-heidelberg.com/). ln a preferred embodiment, a 58P1D12 immunogen contains one or more amino acid sequences identified using techniques well known in the art, such as the sequences shown in Tables Vlll-XXl and XXll-XLlX or a peptide of 8, 9, 10 or 11 amino acids specified by an HLA Class l motif/supermotif (e.g., Table lV (A), Table lV (D), or Table lV (E)) and/or a peptide of at least 9 amino acids that comprises an HLA Class ll motif/supermotif (e.g., Table lV (B) or Table lV (C)). As is appreciated in the art, the HLA Class l binding groove is essentially closed ended so that peptides of only a particular size range can fit into the groove and be bound, generally HLA Class l epitopes are 8, 9, 10, or 11 amino acids long. ln contrast, the HLA Class ll binding groove is essentially open ended; therefore a peptide of about 9 or more amino acids can be bound by an HLA Class ll molecule. Due to the binding groove differences between HLA Class l and ll, HLA Class l motifs are length specific, i.e., position two of a Class l motif is the second amino acid in an amino to carboxyl direction of the peptide. The amino acid positions in a Class ll motif are relative only to each other, not the overall peptide, i.e., additional amino acids can be attached to the amino and/or carboxyl termini of a motif-bearing sequence. HLA Class ll epitopes are often 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, or 25 amino acids long, or longer than 25 amino acids.

Antibody-Based Vaccines

A wide variety of methods for generating an immune response in a mammal are known in the art (for example as the first step in the generation of hybridomas). Methods of generating an immune response in a mammal comprise exposing the mammal's immune system to an immunogenic epitope on a protein (e.g. a 58P1D12 protein) so that an immune response is generated. A typical embodiment consists of a method for generating an immune response to 58P1D12 in a host, by contacting the host with a sufficient amount of at least one 58P1D12 B cell or cytotoxic T-cell epitope or analog thereof; and at least one periodic interval thereafter re-contacting the host with the 58P1D12 B cell or cytotoxic T-cell epitope or analog thereof. A specific embodiment consists of a method of generating an immune response against a 58P1D12-related protein or a man-made multiepitopic peptide comprising: administering 58P1D12 immunogen (e.g. a 58P1D12 protein or a peptide fragment thereof, a 58P1D12 fusion protein or analog etc.) in a vaccine preparation to a human or another mammal. Typically, such vaccine preparations further contain a suitable adjuvant (see, e.g., U.S. Pat. No. 6,146,635) or a universal helper epitope such as a PADRE™ peptide (Epimmune lnc., San Diego, Calif.; see, e.g., Alexander et al., J. lmmunol. 2000 164(3); 164(3): 1625-1633; Alexander et al., lmmunity 1994 1(9): 751-761 and Alexander et al., lmmunol. Res. 1998 18(2): 79-92). An alternative method comprises generating an immune response in an individual against a 58P1D12 immunogen by: administering in vivo to muscle or skin of the individual's body a DNA molecule that comprises a DNA sequence that encodes a 58P1D12 immunogen, the DNA sequence operatively linked to regulatory sequences which control the expression of the DNA sequence; wherein the DNA molecule is taken up by cells, the DNA sequence is expressed in the cells and an immune response is generated against the immunogen (see, e.g., U.S. Pat. No. 5,962,428). Optionally a genetic vaccine facilitator such as anionic lipids; saponins; lectins; estrogenic compounds; hydroxylated lower alkyls; dimethyl sulfoxide; and urea is also administered. ln addition, an antiidiotypic antibody can be administered that mimics 58P1D12, in order to generate a response to the target antigen.

Nucleic Acid Vaccines:

Vaccine compositions of the invention include nucleic acid-mediated modalities. DNA or RNA that encode protein(s) of the invention can be administered to a patient. Genetic immunization methods can be employed to generate prophylactic or therapeutic humoral and cellular immune responses directed against cancer cells expressing 58P1D12. Constructs comprising DNA encoding a 58P1D12-related protein/immunogen and appropriate regulatory sequences can be injected directly into muscle or skin of an individual, such that the cells of the muscle or skin take-up the construct and express the encoded 58P1D12 protein/immunogen. Alternatively, a vaccine comprises a 58P1D12-related protein. Expression of the 58P1D12-related protein immunogen results in the generation of prophylactic or therapeutic humoral and cellular immunity against cells that bear a 58P1D12 protein. Various prophylactic and therapeutic genetic immunization techniques known in the art can be used (for review, see information and references published at Internet address genweb.com). Nucleic acid-based delivery is described, for instance, in Wolff et. al., Science 247:1465 (1990) as well as U.S. Pat. Nos. 5,580,859; 5,589,466; 5,804,566; 5,739,118; 5,736,524; 5,679,647; WO 98/04720. Examples of DNA-based delivery technologies include "naked DNA", facilitated (bupivicaine, polymers, peptide-mediated) delivery, cationic lipid complexes, and particle-mediated ("gene gun") or pressure-mediated delivery (see, e.g., U.S. Pat. No. 5,922,687).

For therapeutic or prophylactic immunization purposes, proteins of the invention can be expressed via viral or bacterial vectors. Various viral gene delivery systems that can be used in the practice of the invention include, but are not limited to, vaccinia, fowlpox, canarypox, adenovirus, influenza, poliovirus, adeno-associated virus, lentivirus, and sindbis virus (see, e.g., Restifo, 1996, Curr. Opin. Immunol. 8:658-663; Tsang et al. J. Natl. Cancer Inst. 87:982-990 (1995)). Non-viral delivery systems can also be employed by introducing naked DNA encoding a 58P1D12-related protein into the patent (e.g., intramuscularly or intradermally) to induce an anti-tumor response.

Vaccinia virus is used, for example, as a vector to express nucleotide sequences that encode the peptides of the invention. Upon introduction into a host, the recombinant vaccinia virus expresses the protein immunogenic peptide, and thereby elicits a host immune response. Vaccinia vectors and methods useful in immunization protocols are described in, e.g., U.S. Pat. No. 4,722,848. Another vector is BCG (Bacille Calmette Guerin). BCG vectors are described in Stover et al., Nature 351:456-460 (1991). A wide variety of other vectors useful for therapeutic administration or immunization of the peptides of the invention, e.g. adeno and adeno-associated virus vectors, retroviral vectors, *Salmonella typhi* vectors, detoxified anthrax toxin vectors, and the like, will be apparent to those skilled in the art from the description herein.

Thus, gene delivery systems are used to deliver a 58P1D12-related nucleic acid molecule. In one embodiment, the full-length human 58P1D12 cDNA is employed. In another embodiment, 58P1D12 nucleic acid molecules encoding specific cytotoxic T lymphocyte (CTL) and/or antibody epitopes are employed.

Ex Vivo Vaccines

Various ex vivo strategies can also be employed to generate an immune response. One approach involves the use of antigen presenting cells (APCs) such as dendritic cells (DC) to present 58P1D12 antigen to a patient's immune system. Dendritic cells express MHC class I and II molecules, B7 co-stimulator, and IL-12, and are thus highly specialized antigen presenting cells. In prostate cancer, autologous dendritic cells pulsed with peptides of the prostate-specific membrane antigen (PSMA) are being used in a Phase I clinical trial to stimulate prostate cancer patients' immune systems (Tjoa et al., 1996, Prostate 28:65-69; Murphy et al., 1996, Prostate 29:371-380). Thus, dendritic cells can be used to present 58P1D12 peptides to T cells in the context of MHC class I or II molecules. In one embodiment, autologous dendritic cells are pulsed with 58P1D12 peptides capable of binding to MHC class I and/or class II molecules. In another embodiment, dendritic cells are pulsed with the complete 58P1D12 protein. Yet another embodiment involves engineering the overexpression of a 58P1D12 gene in dendritic cells using various implementing vectors known in the art, such as adenovirus (Arthur et al., 1997, Cancer Gene Ther. 4:17-25), retrovirus (Henderson et al., 1996, Cancer Res. 56:3763-3770), lentivirus, adeno-associated virus, DNA transfection (Ribas et al., 1997, Cancer Res. 57:2865-2869), or tumor-derived RNA transfection (Ashley et al., 1997, J. Exp. Med. 186: 1177-1182). Cells that express 58P1D12 can also be engineered to express immune modulators, such as G M-CSF, and used as immunizing agents.

X.B.) 58P1D12 as a Target for Antibody-Based Therapy

58P1D12 is an attractive target for antibody-based therapeutic strategies. A number of antibody strategies are known in the art for targeting both extracellular and intracellular molecules (see, e.g., complement and ADCC mediated killing as well as the use of intrabodies). Because 58P1D12 is expressed by cancer cells of various lineages relative to corresponding normal cells, systemic administration of 58P1D12-immunoreactive compositions are prepared that exhibit excellent sensitivity without toxic, non-specific and/or non-target effects caused by binding of the immunoreactive composition to non-target organs and tissues. Antibodies specifically reactive with domains of 58P1D12 are useful to treat 58P1D12-expressing cancers systemically, either as conjugates with a toxin or therapeutic agent, or as naked antibodies capable of inhibiting cell proliferation or function.

58P1D12 antibodies can be introduced into a patient such that the antibody binds to 58P1D12 and modulates a function, such as an interaction with a binding partner, and consequently mediates destruction of the tumor cells and/or inhibits the growth of the tumor cells. Mechanisms by which such antibodies exert a therapeutic effect can include complement-mediated cytolysis, antibody-dependent cellular cytotoxicity, modulation of the physiological function of 58P1D12, inhibition of ligand binding or signal transduction pathways, modulation of tumor cell differentiation, alteration of tumor angiogenesis factor profiles, and/or apoptosis.

Those skilled in the art understand that antibodies can be used to specifically target and bind immunogenic molecules such as an immunogenic region of a 58P1D12 sequence shown in FIG. **2** or FIG. **3**. In addition, skilled artisans understand that it is routine to conjugate antibodies to cytotoxic agents (see, e.g., Slevers et al. Blood 93:11 3678-3684 (Jun. 1, 1999)). When cytotoxic and/or therapeutic agents are delivered directly to cells, such as by conjugating them to antibodies specific for a molecule expressed by that cell (e.g. 58P1D12), the cytotoxic agent will exert its known biological effect (i.e. cytotoxicity) on those cells.

A wide variety of compositions and methods for using antibody-cytotoxic agent conjugates to kill cells are known in the art. In the context of cancers, typical methods entail administering to an animal having a tumor a biologically effective a mount of a conjugate comprising a selected cytotoxic and/or therapeutic agent linked to a targeting agent (e.g. an anti-58P1D12 antibody) that binds to a marker (e.g.

58P1D12) expressed, accessible to binding or localized on the cell surfaces. A typical embodiment is a method of delivering a cytotoxic and/or therapeutic agent to a cell expressing 58P1D12, comprising conjugating the cytotoxic agent to an antibody that immunospecifically binds to a 58P1D12 epitope, and, exposing the cell to the antibody-agent conjugate. Another illustrative embodiment is a method of treating an individual suspected of suffering from metastasized cancer, comprising a step of administering parenterally to said individual a pharmaceutical composition comprising a therapeutically effective amount of an antibody conjugated to a cytotoxic and/or therapeutic agent.

Cancer immunotherapy using anti-58P1D12 antibodies can be done in accordance with various approaches that have been successfully employed in the treatment of other types of cancer, including but not limited to colon cancer (Arlen et al., 1998, Crit. Rev. Immunol. 18:133-138), multiple myeloma (Ozaki et al., 1997, Blood 90:3179-3186, Tsunenari et al., 1997, Blood 90:2437-2444), gastric cancer (Kasprzyk et al., 1992, Cancer Res. 52:2771-2776), B-cell lymphoma (Funakoshi et al., 1996, J. Immunother. Emphasis Tumor Immunol. 19:93-101), leukemia (Zhong et al., 1996, Leuk. Res. 20:581-589), colorectal cancer (Moun et al., 1994, Cancer Res. 54:6160-6166; Velders et al., 1995, Cancer Res. 55:4398-4403), and breast cancer (Shepard et al., 1991, J. Clin. Immunol. 11: 117-127). Some therapeutic approaches involve conjugation of naked antibody to a toxin or radioisotope, such as the conjugation of Y91 or I131 to anti-CD20 antibodies (e.g., Zevalin™, IDEC Pharmaceuticals Corp. or Bexxar™, Coulter Pharmaceuticals), while others involve co-administration of antibodies and other therapeutic agents, such as Herceptin™ (trastuzumab) with paclitaxel (Genentech, Inc.). The antibodies can be conjugated to a therapeutic agent. To treat prostate cancer, for example, 58P1D12 antibodies can be administered in conjunction with radiation, chemotherapy or hormone ablation. Also, antibodies can be conjugated to a toxin such as calicheamicin (e.g., Mylotarg™, Wyeth-Ayerst, Madison, N.J., a recombinant humanized IgG4 kappa antibody conjugated to antitumor antibiotic calicheamicin) or a maytansinoid (e.g., taxane-based Tumor-Activated Prodrug, TAP, platform, ImmunoGen, Cambridge, Mass., also see e.g., U.S. Pat. No. 5,416,064).

Although 58P1D12 antibody therapy is useful for all stages of cancer, antibody therapy can be particularly appropriate in advanced or metastatic cancers. Treatment with the antibody therapy of the invention is indicated for patients who have received one or more rounds of chemotherapy. Alternatively, antibody therapy of the invention is combined with a chemotherapeutic or radiation regimen for patients who have not received chemotherapeutic treatment. Additionally, antibody therapy can enable the use of reduced dosages of concomitant chemotherapy, particularly for patients who do not tolerate the toxicity of the chemotherapeutic agent very well. Fan et al. (Cancer Res. 53:4637-4642, 1993), Prewett et al. (International J. of Onco. 9:217-224, 1996), and Hancock et al. (Cancer Res. 51:4575-4580, 1991) describe the use of various antibodies together with chemotherapeutic agents.

Although 58P1D12 antibody therapy is useful for all stages of cancer, antibody therapy can be particularly appropriate in advanced or metastatic cancers. Treatment with the antibody therapy of the invention is indicated for patients who have received one or more rounds of chemotherapy. Alternatively, antibody therapy of the invention is combined with a chemotherapeutic or radiation regimen for patients who have not received chemotherapeutic treatment. Additionally, antibody therapy can enable the use of reduced dosages of concomitant chemotherapy, particularly for patients who do not tolerate the toxicity of the chemotherapeutic agent very well.

Cancer patients can be evaluated for the presence and level of 58P1D12 expression, preferably using immunohistochemical assessments of tumor tissue, quantitative 58P1D12 imaging, or other techniques that reliably indicate the presence and degree of 58P1D12 expression. Immunohistochemical analysis of tumor biopsies or surgical specimens is preferred for this purpose. Methods for immunohistochemical analysis of tumor tissues are well known in the art.

Anti-58P1D12 monoclonal antibodies that treat prostate and other cancers include those that initiate a potent immune response against the tumor or those that are directly cytotoxic. In this regard, anti-58P1D12 monoclonal antibodies (mAbs) can elicit tumor cell lysis by either complement-mediated or antibody-dependent cell cytotoxicity (ADCC) mechanisms, both of which require an intact Fc portion of the immunoglobulin molecule for interaction with effector cell Fc receptor sites on complement proteins. In addition, anti-58P1D12 mAbs that exert a direct biological effect on tumor growth are useful to treat cancers that express 58P1D12. Mechanisms by which directly cytotoxic mAbs act include: inhibition of cell growth, modulation of cellular differentiation, modulation of tumor angiogenesis factor profiles, and the induction of apoptosis. The mechanism(s) by which a particular anti-58P1D12 mAb exerts an anti-tumor effect is evaluated using any number of in vitro assays that evaluate cell death such as ADCC, ADMMC, complement-mediated cell lysis, and so forth, as is generally known in the art.

In some patients, the use of murine or other non-human monoclonal antibodies, or human/mouse chimeric mAbs can induce moderate to strong immune responses against the non-human antibody. This can result in clearance of the antibody from circulation and reduced efficacy. In the most severe cases, such an immune response can lead to the extensive formation of immune complexes which, potentially, can cause renal failure. Accordingly, preferred monoclonal antibodies used in the therapeutic methods of the invention are those that are either fully human or humanized and that bind specifically to the target 58P1D12 antigen with high affinity but exhibit low or no antigenicity in the patient.

Therapeutic methods of the invention contemplate the administration of single anti-58P1D12 mAbs as well as combinations, or cocktails, of different mAbs. Such mAb cocktails can have certain advantages inasmuch as they contain mAbs that target different epitopes, exploit different effector mechanisms or combine directly cytotoxic mAbs with mAbs that rely on immune effector functionality. Such mAbs in combination can exhibit synergistic therapeutic effects. In addition, anti-58P1D12 mAbs can be administered concomitantly with other therapeutic modalities, including but not limited to various chemotherapeutic agents, androgen-blockers, immune modulators (e.g., IL-2, GM-CSF), surgery or radiation. The anti-58P1D12 mAbs are administered in their "naked" or unconjugated form, or can have a therapeutic agent(s) conjugated to them.

Anti-58P1D12 antibody formulations are administered via any route capable of delivering the antibodies to a tumor cell. Routes of administration include, but are not limited to, intravenous, intraperitoneal, intramuscular, intratumor, intradermal, and the like. Treatment generally involves repeated administration of the anti-58P1D12 antibody preparation, via an acceptable route of administration such as intravenous injection (IV), typically at a dose in the range of about 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9., 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, or 25 mg/kg body weight. In general, doses in the range of 10-1000 mg mAb per week are effective and well tolerated.

Based on clinical experience with the Herceptin™ mAb in the treatment of metastatic breast cancer, an initial loading dose of approximately 4 mg/kg patient body weight IV, followed by weekly doses of about 2 mg/kg IV of the anti-58P1D12 mAb preparation represents an acceptable dosing regimen. Preferably, the initial loading dose is administered as a 90-minute or longer infusion. The periodic maintenance dose is administered as a 30 minute or longer infusion, provided the initial dose was well tolerated. As appreciated by those of skill in the art, various factors can influence the ideal dose regimen in a particular case. Such factors include, for example, the binding affinity and half life of the Ab or mAbs used, the degree of 58P1D12 expression in the patient, the extent of circulating shed 58P1D12 antigen, the desired steady-state antibody concentration level, frequency of treatment, and the influence of chemotherapeutic or other agents used in combination with the treatment method of the invention, as well as the health status of a particular patient.

Optionally, patients should be evaluated for the levels of 58P1D12 in a given sample (e.g. the levels of circulating 58P1D12 antigen and/or 58P1D12 expressing cells) in order to assist in the determination of the most effective dosing regimen, etc. Such evaluations are also used for monitoring purposes throughout therapy, and are useful to gauge therapeutic success in combination with the evaluation of other parameters (for example, urine cytology and/or ImmunoCyt levels in bladder cancer therapy, or by analogy, serum PSA levels in prostate cancer therapy).

Anti-idiotypic anti-58P1D12 antibodies can also be used in anti-cancer therapy as a vaccine for inducing an immune response to cells expressing a 58P1D12-related protein. In particular, the generation of anti-idiotypic antibodies is well known in the art; this methodology can readily be adapted to generate anti-idiotypic anti-58P1D12 antibodies that mimic an epitope on a 58P1D12-related protein (see, for example, Wagner et al., 1997, Hybridoma 16: 33-40; Foon et al., 1995, J. Clin. Invest. 96:334-342; Herlyn et al., 1996, Cancer Immunol. Immunother. 43:65-76). Such an anti-idiotypic antibody can be used in cancer vaccine strategies.

X.C.) 58P1D12 as a Target for Cellular Immune Responses

Vaccines and methods of preparing vaccines that contain an immunogenically effective amount of one or more HLA-binding peptides as described herein are further embodiments of the invention. Furthermore, vaccines in accordance with the invention encompass compositions of one or more of the claimed peptides. A peptide can be present in a vaccine individually. Alternatively, the peptide can exist as a homopolymer comprising multiple copies of the same peptide, or as a heteropolymer of various peptides. Polymers have the advantage of increased immunological reaction and, where different peptide epitopes are used to make up the polymer, the additional ability to induce antibodies and/or CTLs that react with different antigenic determinants of the pathogenic organism or tumor-related peptide targeted for an immune response. The composition can be a naturally occurring region of an antigen or can be prepared, e.g., recombinantly or by chemical synthesis.

Carriers that can be used with vaccines of the invention are well known in the art, and include, e.g., thyroglobulin, albumins such as human serum albumin, tetanus toxoid, polyamino acids such as poly l-lysine, poly l-glutamic acid, influenza, hepatitis B virus core protein, and the like. The vaccines can contain a physiologically tolerable (i.e., acceptable) diluent such as water, or saline, preferably phosphate buffered saline. The vaccines also typically include an adjuvant. Adjuvants such as incomplete Freund's adjuvant, aluminum phosphate, aluminum hydroxide, or alum are examples of materials well known in the art. Additionally, as disclosed herein, CTL responses can be primed by conjugating peptides of the invention to lipids, such as tripalmitoyl-S-glycerylcysteinlyseryl-serine (P3CSS). Moreover, an adjuvant such as a synthetic cytosine-phosphorothiolated-guanine-containing (CpG) oligonucleotides has been found to increase CTL responses 10- to 100-fold. (see, e.g. Davila and Celis, J. Immunol. 165:539-547 (2000)).

Upon immunization with a peptide composition in accordance with the invention, via injection, aerosol, oral, transdermal, transmucosal, intrapleural, intrathecal, or other suitable routes, the immune system of the host responds to the vaccine by producing large amounts of CTLs and/or HTLs specific for the desired antigen. Consequently, the host becomes at least partially immune to later development of cells that express or overexpress 58P1D12 antigen, or derives at least some therapeutic benefit when the antigen was tumor-associated.

In some embodiments, it may be desirable to combine the class I peptide components with components that induce or facilitate neutralizing antibody and or helper T cell responses directed to the target antigen. A preferred embodiment of such a composition comprises class I and class II epitopes in accordance with the invention. An alternative embodiment of such a composition comprises a class I and/or class II epitope in accordance with the invention, along with a cross reactive HTL epitope such as PADRE™ (Epimmune, San Diego, Calif.) molecule (described e.g., in U.S. Pat. No. 5,736,142).

A vaccine of the invention can also include antigen-presenting cells (APC), such as dendritic cells (DC), as a vehicle to present peptides of the invention. Vaccine compositions can be created in vitro, following dendritic cell mobilization and harvesting, whereby loading of dendritic cells occurs in vitro. For example, dendritic cells are transfected, e.g., with a minigene in accordance with the invention, or are pulsed with peptides. The dendritic cell can then be administered to a patient to elicit immune responses in vivo. Vaccine compositions, either DNA- or peptide-based, can also be administered in vivo in combination with dendritic cell mobilization whereby loading of dendritic cells occurs in vivo.

Preferably, the following principles are utilized when selecting an array of epitopes for inclusion in a polyepitopic composition for use in a vaccine, or for selecting discrete epitopes to be included in a vaccine and/or to be encoded by nucleic acids such as a minigene. It is preferred that each of the following principles be balanced in order to make the selection. The multiple epitopes to be incorporated in a given vaccine composition may be, but need not be, contiguous in sequence in the native antigen from which the epitopes are derived.

1.) Epitopes are selected which, upon administration, mimic immune responses that have been observed to be correlated with tumor clearance. For HLA Class I this includes 3-4 epitopes that come from at least one tumor associated antigen (TAA). For HLA Class II a similar rationale is employed; again 3-4 epitopes are selected from at least one TAA (see, e.g., Rosenberg et al., Science 278:1447-1450). Epitopes from one TAA may be used in combination with epitopes from one or more additional TAAs to produce a vaccine that targets tumors with varying expression patterns of frequently-expressed TAAs.

2.) Epitopes are selected that have the requisite binding affinity established to be correlated with immunogenicity: for

HLA Class I an IC50 of 500 nM or less, often 200 nM or less; and for Class II an IC50 of 1000 nM or less.

3.) Sufficient supermotif bearing-peptides, or a sufficient array of allele-specific motif-bearing peptides, are selected to give broad population coverage. For example, it is preferable to have at least 80% population coverage. A Monte Carlo analysis, a statistical evaluation known in the art, can be employed to assess the breadth, or redundancy of, population coverage.

4.) When selecting epitopes from cancer-related antigens it is often useful to select analogs because the patient may have developed tolerance to the native epitope.

5.) Of particular relevance are epitopes referred to as "nested epitopes." Nested epitopes occur where at least two epitopes overlap in a given peptide sequence. A nested peptide sequence can comprise B cell, HLA class I and/or HLA class II epitopes. When providing nested epitopes, a general objective is to provide the greatest number of epitopes per sequence. Thus, an aspect is to avoid providing a peptide that is any longer than the amino terminus of the amino terminal epitope and the carboxyl terminus of the carboxyl terminal epitope in the peptide. When providing a multi-epitopic sequence, such as a sequence comprising nested epitopes, it is generally important to screen the sequence in order to insure that it does not have pathological or other deleterious biological properties.

6.) If a polyepitopic protein is created, or when creating a minigene, an objective is to generate the smallest peptide that encompasses the epitopes of interest. This principle is similar, if not the same as that employed when selecting a peptide comprising nested epitopes. However, with an artificial polyepitopic peptide, the size minimization objective is balanced against the need to integrate any spacer sequences between epitopes in the polyepitopic protein. Spacer amino acid residues can, for example, be introduced to avoid junctional epitopes (an epitope recognized by the immune system, not present in the target antigen, and only created by the man-made juxtaposition of epitopes), or to facilitate cleavage between epitopes and thereby enhance epitope presentation. Junctional epitopes are generally to be avoided because the recipient may generate an immune response to that non-native epitope. Of particular concern is a junctional epitope that is a "dominant epitope." A dominant epitope may lead to such a zealous response that immune responses to other epitopes are diminished or suppressed.

7.) Where the sequences of multiple variants of the same target protein are present, potential peptide epitopes can also be selected on the basis of their conservancy. For example, a criterion for conservancy may define that the entire sequence of an HLA class I binding peptide or the entire 9-mer core of a class II binding peptide be conserved in a designated percentage of the sequences evaluated for a specific protein antigen.

X.C.1. Minigene Vaccines

A number of different approaches are available which allow simultaneous delivery of multiple epitopes. Nucleic acids encoding the peptides of the invention are a particularly useful embodiment of the invention. Epitopes for inclusion in a minigene are preferably selected according to the guidelines set forth in the previous section. A preferred means of administering nucleic acids encoding the peptides of the invention uses minigene constructs encoding a peptide comprising one or multiple epitopes of the invention.

The use of multi-epitope minigenes is described below and in, Ishioka et al., J. Immunol. 162:3915-3925, 1999; An, L. and Whitton, J. L., J. Virol. 71:2292, 1997; Thomson, S. A. et al., J. Immunol. 157:822, 1996; Whitton, J. L. et al., J. Virol. 67:348, 1993; Hanke, R. et al., Vaccine 16:426, 1998. For example, a multi-epitope DNA plasmid encoding supermotif- and/or motif-bearing epitopes derived 58P1D12, the PADRE® universal helper T cell epitope or multiple HTL epitopes from 58P1D12 (see e.g., Tables VIII-XXI and XXII to XLIX), and an endoplasmic reticulum-translocating signal sequence can be engineered. A vaccine may also comprise epitopes that are derived from other TAAs.

The immunogenicity of a multi-epitopic minigene can be confirmed in transgenic mice to evaluate the magnitude of CTL induction responses against the epitopes tested. Further, the immunogenicity of DNA-encoded epitopes in vivo can be correlated with the in vitro responses of specific CTL lines against target cells transfected with the DNA plasmid. Thus, these experiments can show that the minigene serves to both: 1.) generate a CTL response and 2.) that the induced CTLs recognized cells expressing the encoded epitopes.

For example, to create a DNA sequence encoding the selected epitopes (minigene) for expression in human cells, the amino acid sequences of the epitopes may be reverse translated. A human codon usage table can be used to guide the codon choice for each amino acid. These epitope-encoding DNA sequences may be directly adjoined, so that when translated, a continuous polypeptide sequence is created. To optimize expression and/or immunogenicity, additional elements can be incorporated into the minigene design. Examples of amino acid sequences that can be reverse translated and included in the minigene sequence include: HLA class I epitopes, HLA class II epitopes, antibody epitopes, a ubiquitination signal sequence, and/or an endoplasmic reticulum targeting signal. In addition, HLA presentation of CTL and HTL epitopes may be improved by including synthetic (e.g. poly-alanine) or naturally-occurring flanking sequences adjacent to the CTL or HTL epitopes; these larger peptides comprising the epitope(s) are within the scope of the invention.

The minigene sequence may be converted to DNA by assembling oligonucleotides that encode the plus and minus strands of the minigene. Overlapping oligonucleotides (30-100 bases long) may be synthesized, phosphorylated, purified and annealed under appropriate conditions using well known techniques. The ends of the oligonucleotides can be joined, for example, using T4 DNA ligase. This synthetic minigene, encoding the epitope polypeptide, can then be cloned into a desired expression vector.

Standard regulatory sequences well known to those of skill in the art are preferably included in the vector to ensure expression in the target cells. Several vector elements are desirable: a promoter with a down-stream cloning site for minigene insertion; a polyadenylation signal for efficient transcription termination; an *E. coli* origin of replication; and an *E. coli* selectable marker (e.g. ampicillin or kanamycin resistance). Numerous promoters can be used for this purpose, e.g., the human cytomegalovirus (hCMV) promoter. See, e.g., U.S. Pat. Nos. 5,580,859 and 5,589,466 for other suitable promoter sequences.

Additional vector modifications may be desired to optimize minigene expression and immunogenicity. In some cases, introns are required for efficient gene expression, and one or more synthetic or naturally-occurring introns could be incorporated into the transcribed region of the minigene. The inclusion of mRNA stabilization sequences and sequences for replication in mammalian cells may also be considered for increasing minigene expression.

Once an expression vector is selected, the minigene is cloned into the polylinker region downstream of the promoter. This plasmid is transformed into an appropriate *E. coli* strain, and DNA is prepared using standard techniques. The orientation and DNA sequence of the minigene, as well as all other elements included in the vector, are confirmed using restriction mapping and DNA sequence analysis. Bacterial cells harboring the correct plasmid can be stored as a master cell bank and a working cell bank.

In addition, immunostimulatory sequences (ISSs or CpGs) appear to play a role in the immunogenicity of DNA vaccines. These sequences may be included in the vector, outside the minigene coding sequence, if desired to enhance immunogenicity.

In some embodiments, a bi-cistronic expression vector which allows production of both the minigene-encoded epitopes and a second protein (included to enhance or decrease immunogenicity) can be used. Examples of proteins or polypeptides that could beneficially enhance the immune response if co-expressed include cytokines (e.g., IL-2, IL-12, GM-CSF), cytokine-inducing molecules (e.g., LeIF), costimulatory molecules, or for HTL responses, pan-DR binding proteins (PADRE™, Epimmune, San Diego, Calif.). Helper (HTL) epitopes can be joined to intracellular targeting signals and expressed separately from expressed CTL epitopes; this allows direction of the HTL epitopes to a cell compartment different than that of the CTL epitopes. If required, this could facilitate more efficient entry of HTL epitopes into the HLA class II pathway, thereby improving HTL induction. In contrast to HTL or CTL induction, specifically decreasing the immune response by co-expression of immunosuppressive molecules (e.g. TGF-β) may be beneficial in certain diseases.

Therapeutic quantities of plasmid DNA can be produced for example, by fermentation in *E. coli*, followed by purification. Aliquots from the working cell bank are used to inoculate growth medium, and grown to saturation in shaker flasks or a bioreactor according to well-known techniques. Plasmid DNA can be purified using standard bioseparation technologies such as solid phase anion-exchange resins supplied by QIAGEN, Inc. (Valencia, Calif.). If required, supercoiled DNA can be isolated from the open circular and linear forms using gel electrophoresis or other methods.

Purified plasmid DNA can be prepared for injection using a variety of formulations. The simplest of these is reconstitution of lyophilized DNA in sterile phosphate-buffer saline (PBS). This approach, known as "naked DNA," is currently being used for intramuscular (IM) administration in clinical trials. To maximize the immunotherapeutic effects of minigene DNA vaccines, an alternative method for formulating purified plasmid DNA may be desirable. A variety of methods have been described, and new techniques may become available. Cationic lipids, glycolipids, and fusogenic liposomes can also be used in the formulation (see, e.g., as described by WO 93/24640; Mannino & Gould-Fogerite, BioTechniques 6(7): 682 (1988); U.S. Pat. No. 5,279,833; WO 91/06309; and Felgner, et al., Proc. Nat'l Acad. Sci. USA 84:7413 (1987). In addition, peptides and compounds referred to collectively as protective, interactive, non-condensing compounds (PINC) could also be complexed to purified plasmid DNA to influence variables such as stability, intramuscular dispersion, or trafficking to specific organs or cell types.

Target cell sensitization can be used as a functional assay for expression and HLA class I presentation of minigene-encoded CTL epitopes. For example, the plasmid DNA is introduced into a mammalian cell line that is suitable as a target for standard CTL chromium release assays. The transfection method used will be dependent on the final formulation. Electroporation can be used for "naked" DNA, whereas cationic lipids allow direct in vitro transfection. A plasmid expressing green fluorescent protein (GFP) can be co-transfected to allow enrichment of transfected cells using fluorescence activated cell sorting (FACS). These cells are then chromium-51 (51Cr) labeled and used as target cells for epitope-specific CTL lines; cytolysis, detected by 51Cr release, indicates both production of, and HLA presentation of, minigene-encoded CTL epitopes. Expression of HTL epitopes may be evaluated in an analogous manner using assays to assess HTL activity.

In vivo immunogenicity is a second approach for functional testing of minigene DNA formulations. Transgenic mice expressing appropriate human HLA proteins are immunized with the DNA product. The dose and route of administration are formulation dependent (e.g., IM for DNA in PBS, intraperitoneal (i.p.) for lipid-complexed DNA). Twenty-one days after immunization, splenocytes are harvested and restimulated for one week in the presence of peptides encoding each epitope being tested. Thereafter, for CTL effector cells, assays are conducted for cytolysis of peptide-loaded, 51Cr-labeled target cells using standard techniques. Lysis of target cells that were sensitized by HLA loaded with peptide epitopes, corresponding to minigene-encoded epitopes, demonstrates DNA vaccine function for in vivo induction of CTLs. Immunogenicity of HTL epitopes is confirmed in transgenic mice in an analogous manner.

Alternatively, the nucleic acids can be administered using ballistic delivery as described, for instance, in U.S. Pat. No. 5,204,253. Using this technique, particles comprised solely of DNA are administered. In a further alternative embodiment, DNA can be adhered to particles, such as gold particles.

Minigenes can also be delivered using other bacterial or viral delivery systems well known in the art, e.g., an expression construct encoding epitopes of the invention can be incorporated into a viral vector such as vaccinia.

X.C.2. Combinations of CTL Peptides with Helper Peptides

Vaccine compositions comprising CTL peptides of the invention can be modified, e.g., analoged, to provide desired attributes, such as improved serum half life, broadened population coverage or enhanced immunogenicity.

For instance, the ability of a peptide to induce CTL activity can be enhanced by linking the peptide to a sequence which contains at least one epitope that is capable of inducing a T helper cell response. Although a CTL peptide can be directly linked to a T helper peptide, often CTL epitope/HTL epitope conjugates are linked by a spacer molecule. The spacer is typically comprised of relatively small, neutral molecules, such as amino acids or amino acid mimetics, which are substantially uncharged under physiological conditions. The spacers are typically selected from, e.g., Ala, Gly, or other neutral spacers of nonpolar amino acids or neutral polar amino acids. It will be understood that the optionally present spacer need not be comprised of the same residues and thus may be a hetero- or homo-oligomer. When present, the spacer will usually be at least one or two residues, more usually three to six residues and sometimes 10 or more residues. The CTL peptide epitope can be linked to the T helper peptide epitope either directly or via a spacer either at the amino or carboxy terminus of the CTL peptide. The amino terminus of either the immunogenic peptide or the T helper peptide may be acylated.

In certain embodiments, the T helper peptide is one that is recognized by T helper cells present in a majority of a genetically diverse population. This can be accomplished by selecting peptides that bind to many, most, or all of the HLA class II molecules. Examples of such amino acid bind many HLA Class II molecules include sequences from antigens such as tetanus toxoid at positions 830-843 QYIKANSKFIGITE; (SEQ ID NO: 18), *Plasmodium falciparum* circumsporozoite (CS) protein at positions 378-398 DIEKKIAKMEKASSVFNVVNS; (SEQ ID NO: 19), and *Streptococcus* 18 kD protein at positions 116-131 GAVDSILGGVATYGAA; (SEQ ID NO: 20). Other examples include peptides bearing a DR 1-4-7 supermotif, or either of the DR3 motifs.

Alternatively, it is possible to prepare synthetic peptides capable of stimulating T helper lymphocytes, in a loosely HLA-restricted fashion, using amino acid sequences not found in nature (see, e.g., PCT publication WO 95/07707). These synthetic compounds called Pan-DR-binding epitopes (e.g., PADRE™, Epimmune, Inc., San Diego, Calif.) are designed, most preferably, to bind most HLA-DR (human HLA class II) molecules. For instance, a pan-DR-binding epitope peptide having the formula: xKXVAAWTLKAAx (SEQ ID NO: 21), where "X" is either cyclohexylalanine, phenylalanine, or tyrosine, and a is either d-alanine or l-alanine, has been found to bind to most HLA-DR alleles, and to stimulate the response of T helper lymphocytes from most individuals, regardless of their HLA type. An alternative of a pan-DR binding epitope comprises all "L" natural amino acids and can be provided in the form of nucleic acids that encode the epitope.

HTL peptide epitopes can also be modified to alter their biological properties. For example, they can be modified to include d-amino acids to increase their resistance to proteases and thus extend their serum half life, or they can be conjugated to other molecules such as lipids, proteins, carbohydrates, and the like to increase their biological activity. For example, a T helper peptide can be conjugated to one or more palmitic acid chains at either the amino or carboxyl termini.

X.C.3. Combinations of CTL Peptides with T Cell Priming Agents

In some embodiments it may be desirable to include in the pharmaceutical compositions of the invention at least one component which primes B lymphocytes or T lymphocytes. Lipids have been identified as agents capable of priming CTL in vivo. For example, palmitic acid residues can be attached to the ϵ- and α-amino groups of a lysine residue and then linked, e.g., via one or more linking residues such as Gly, Gly-Gly-, Ser, Ser-Ser, or the like, to an immunogenic peptide. The lipidated peptide can then be administered either directly in a micelle or particle, incorporated into a liposome, or emulsified in an adjuvant, e.g., incomplete Freund's adjuvant. In a preferred embodiment, a particularly effective immunogenic composition comprises palmitic acid attached to ϵ- and α-amino groups of Lys, which is attached via linkage, e.g., Ser-Ser, to the amino terminus of the immunogenic peptide.

As another example of lipid priming of CTL responses, *E. coli* lipoproteins, such as tripalmitoyl-S-glycerylcysteinlyseryl-serine (P3CSS) can be used to prime virus specific CTL when covalently attached to an appropriate peptide (see, e.g., Deres, et al., Nature 342:561, 1989). Peptides of the invention can be coupled to P3CSS, for example, and the lipopeptide administered to an individual to prime specifically an immune response to the target antigen. Moreover, because the induction of neutralizing antibodies can also be primed with P3CSS-conjugated epitopes, two such compositions can be combined to more effectively elicit both humoral and cell-mediated responses.

X.C.4. Vaccine Compositions Comprising DC Pulsed with CTL and/or HTL Peptides

An embodiment of a vaccine composition in accordance with the invention comprises ex vivo administration of a cocktail of epitope-bearing peptides to PBMC, or isolated DC therefrom, from the patient's blood. A pharmaceutical to facilitate harvesting of DC can be used, such as Progenipoietin™ (Pharmacia-Monsanto, St. Louis, Mo.) or GM-CSF/IL-4. After pulsing the DC with peptides and prior to reinfusion into patients, the DC are washed to remove unbound peptides. In this embodiment, a vaccine comprises peptide-pulsed DCs which present the pulsed peptide epitopes complexed with HLA molecules on their surfaces.

The DC can be pulsed ex vivo with a cocktail of peptides, some of which stimulate CTL responses to 58P1D12. Optionally, a helper T cell (HTL) peptide, such as a natural or artificial loosely restricted HLA Class II peptide, can be included to facilitate the CTL response. Thus, a vaccine in accordance with the invention is used to treat a cancer which expresses or overexpresses 58P1D12.

X.D. Adoptive Immunotherapy

Antigenic 58P1D12-related peptides are used to elicit a CTL and/or HTL response ex vivo, as well. The resulting CTL or HTL cells, can be used to treat tumors in patients that do not respond to other conventional forms of therapy, or will not respond to a therapeutic vaccine peptide or nucleic acid in accordance with the invention. Ex vivo CTL or HTL responses to a particular antigen are induced by incubating in tissue culture the patent's, or genetically compatible, CTL or HTL precursor cells together with a source of antigen-presenting cells (APC), such as dendritic cells, and the appropriate immunogenic peptide. After an appropriate incubation time (typically about 7-28 days), in which the precursor cells are activated and expanded into effector cells, the cells are infused back into the patient, where they will destroy (CTL) or facilitate destruction (HTL) of their specific target cell (e.g., a tumor cell). Transfected dendritic cells may also be used as antigen presenting cells.

X.E. Administration of Vaccines for Therapeutic or Prophylactic Purposes

Pharmaceutical and vaccine compositions of the invention are typically used to treat and/or prevent a cancer that expresses or overexpresses 58P1D12. In therapeutic applications, peptide and/or nucleic acid compositions are administered to a patient in an amount sufficient to elicit an effective B cell, CTL and/or HTL response to the antigen and to cure or at least partially arrest or slow symptoms and/or complications. An amount adequate to accomplish this is defined as "therapeutically effective dose." Amounts effective for this use will depend on, e.g., the particular composition administered, the manner of administration, the stage and severity of the disease being treated, the weight and general state of health of the patient, and the judgment of the prescribing physician.

For pharmaceutical compositions, the immunogenic peptides of the invention, or DNA encoding them, are generally administered to an individual already bearing a tumor that expresses 58P1D12. The peptides or DNA encoding them can be administered individually or as fusions of one or more peptide sequences. Patients can be treated with the immunogenic peptides separately or in conjunction with other treatments, such as surgery, as appropriate.

For therapeutic use, administration should generally begin at the first diagnosis of 58P1D12-associated cancer. This is followed by boosting doses until at least symptoms are substantially abated and for a period thereafter. The embodiment of the vaccine composition (i.e., including, but not limited to embodiments such as peptide cocktails, polyepitopic polypeptides, minigenes, or TAA-specific CTLs or pulsed dendritic cells) delivered to the patient may vary according to the stage of the disease or the patient's health status. For example, in a patient with a tumor that expresses 58P1D12, a vaccine comprising 58P1D12-specific CTL may be more efficacious in killing tumor cells in patient with advanced disease than alternative embodiments.

It is generally important to provide an amount of the peptide epitope delivered by a mode of administration sufficient to stimulate effectively a cytotoxic T cell response; compositions which stimulate helper T cell responses can also be given in accordance with this embodiment of the invention.

The dosage for an initial therapeutic immunization generally occurs in a unit dosage range where the lower value is about 1, 5, 50, 500, or 1,000 μg and the higher value is about 10,000; 20,000; 30,000; or 50,000 μg. Dosage values for a human typically range from about 500 μg to about 50,000 μg per 70 kilogram patient. Boosting dosages of between about 1.0 μg to about 50,000 μg of peptide pursuant to a boosting regimen over weeks to months may be administered depending upon the patient's response and condition as determined by measuring the specific activity of CTL and HTL obtained from the patient's blood. Administration should continue until at least clinical symptoms or laboratory tests indicate that the neoplasia, has been eliminated or reduced and for a period thereafter. The dosages, routes of administration, and dose schedules are adjusted in accordance with methodologies known in the art.

In certain embodiments, the peptides and compositions of the present invention are employed in serious disease states, that is, life-threatening or potentially life threatening situations. In such cases, as a result of the minimal amounts of extraneous substances and the relative nontoxic nature of the peptides in preferred compositions of the invention, it is possible and may be felt desirable by the treating physician to administer substantial excesses of these peptide compositions relative to these stated dosage amounts.

The vaccine compositions of the invention can also be used purely as prophylactic agents. Generally the dosage for an initial prophylactic immunization generally occurs in a unit dosage range where the lower value is about 1, 5, 50, 500, or 1000 μg and the higher value is about 10,000; 20,000; 30,000; or 50,000 μg. Dosage values for a human typically range from about 500 μg to about 50,000 μg per 70 kilogram patient. This is followed by boosting dosages of between about 1.0 μg to about 50,000 μg of peptide administered at defined intervals from about four weeks to six months after the initial administration of vaccine. The immunogenicity of the vaccine can be assessed by measuring the specific activity of CTL and HTL obtained from a sample of the patient's blood.

The pharmaceutical compositions for therapeutic treatment are intended for parenteral, topical, oral, nasal, intrathecal, or local (e.g. as a cream or topical ointment) administration. Preferably, the pharmaceutical compositions are administered parentally, e.g., intravenously, subcutaneously, intradermally, or intramuscularly. Thus, the invention provides compositions for parenteral administration which comprise a solution of the immunogenic peptides dissolved or suspended in an acceptable carrier, preferably an aqueous carrier.

A variety of aqueous carriers may be used, e.g., water, buffered water, 0.8% saline, 0.3% glycine, hyaluronic acid and the like. These compositions may be sterilized by conventional, well-known sterilization techniques, or may be sterile filtered. The resulting aqueous solutions may be packaged for use as is, or lyophilized, the lyophilized preparation being combined with a sterile solution prior to administration.

The compositions may contain pharmaceutically acceptable auxiliary substances as required to approximate physiological conditions, such as pH-adjusting and buffering agents, tonicity adjusting agents, wetting agents, preservatives, and the like, for example, sodium acetate, sodium lactate, sodium chloride, potassium chloride, calcium chloride, sorbitan monolaurate, triethanolamine oleate, etc.

The concentration of peptides of the invention in the pharmaceutical formulations can vary widely, i.e., from less than about 0.1%, usually at or at least about 2% to as much as 20% to 50% or more by weight, and will be selected primarily by fluid volumes, viscosities, etc., in accordance with the particular mode of administration selected.

A human unit dose form of a composition is typically included in a pharmaceutical composition that comprises a human unit dose of an acceptable carrier, in one embodiment an aqueous carrier, and is administered in a volume/quantity that is known by those of skill in the art to be used for administration of such compositions to humans (see, e.g., Remington's Pharmaceutical Sciences, 17th Edition, A. Gennaro, Editor, Mack Publishing Co., Easton, Pa., 1985). For example a peptide dose for initial immunization can be from about 1 to about 50,000 μg, generally 100-5,000 μg, for a 70 kg patient. For example, for nucleic acids an initial immunization may be performed using an expression vector in the form of naked nucleic acid administered IM (or SC or ID) in the amounts of 0.5-5 mg at multiple sites. The nucleic acid (0.1 to 1000 μg) can also be administered using a gene gun. Following an incubation period of 3-4 weeks, a booster dose is then administered. The booster can be recombinant fowlpox virus administered at a dose of 5-107 to 5×109 pfu.

For antibodies, a treatment generally involves repeated administration of the anti-58P1D12 antibody preparation, via an acceptable route of administration such as intravenous injection (IV), typically at a dose in the range of about 0.1 to about 10 mg/kg body weight. In general, doses in the range of 10-500 mg mAb per week are effective and well tolerated. Moreover, an initial loading dose of approximately 4 mg/kg patient body weight IV, followed by weekly doses of about 2 mg/kg IV of the anti-58P1D12 mAb preparation represents an acceptable dosing regimen. As appreciated by those of skill in the art, various factors can influence the ideal dose in a particular case. Such factors include, for example, half life of a composition, the binding affinity of an Ab, the immunogenicity of a substance, the degree of 58P1D12 expression in the patient, the extent of circulating shed 58P1D12 antigen, the desired steady-state concentration level, frequency of treatment, and the influence of chemotherapeutic or other agents used in combination with the treatment method of the invention, as well as the health status of a particular patient. Non-limiting preferred human unit doses are, for example, 500 μg-1 mg, 1 mg-50 mg, 50 mg-100 mg, 100 mg-200 mg, 200 mg-300 mg, 400 mg-500 mg, 500 mg-600 mg, 600 mg-700 mg, 700 mg-800 mg, 800 mg-900 mg, 900 mg-1 g, or 1 mg-700 mg. In certain embodiments, the dose is in a range of 2-5 mg/kg bodyweight, e.g., with follow on weekly doses of 1-3 mg/kg; 0.5 mg, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 mg/kg body weight followed, e.g., in two, three or four weeks by weekly doses; 0.5-10 mg/kg body weight, e.g., followed in two, three or four weeks by weekly doses; 225, 250, 275, 300, 325, 350, 375, 400 mg m2 of body area weekly; 1-600 mg m2 of body area weekly; 225-400 mg m2 of body area weekly; these does can be followed by weekly doses for 2, 3, 4, 5, 6, 7, 8, 9, 19, 11, 12 or more weeks.

In one embodiment, human unit dose forms of polynucleotides comprise a suitable dosage range or effective amount that provides any therapeutic effect. As appreciated by one of ordinary skill in the art a therapeutic effect depends on a number of factors, including the sequence of the polynucleotide, molecular weight of the polynucleotide and route of administration. Dosages are generally selected by the physician or other health care professional in accordance with a variety of parameters known in the art, such as severity of symptoms, history of the patient and the like. Generally, for a polynucleotide of about 20 bases, a dosage range may be selected from, for example, an independently selected lower limit such as about 0.1, 0.25, 0.5, 1, 2, 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400 or 500 mg/kg up to an independently selected upper limit, greater than the lower limit, of about 60, 80, 100, 200, 300, 400, 500, 750, 1000, 1500, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000 or 10,000 mg/kg. For example, a dose may be about any of the following: 0.1 to 100 mg/kg, 0.1 to 50 mg/kg, 0.1 to 25 mg/kg, 0.1 to 10 mg/kg, 1 to 500 mg/kg, 100 to 400 mg/kg, 200 to 300 mg/kg, 1 to 100 mg/kg, 100 to 200 mg/kg, 300 to 400 mg/kg, 400 to 500 mg/kg, 500 to 1000 mg/kg, 500 to 5000 mg/kg, or 500 to 10,000 mg/kg. Generally, parenteral routes of administration may require higher doses of polynucleotide compared to more direct application to the nucleotide to diseased tissue, as do polynucleotides of increasing length.

In one embodiment, human unit dose forms of T-cells comprise a suitable dosage range or effective amount that provides any therapeutic effect. As appreciated by one of ordinary skill in the art, a therapeutic effect depends on a number of factors. Dosages are generally selected by the physician or other health care professional in accordance with a variety of parameters known in the art, such as severity of symptoms, history of the patient and the like. A dose may be about 104 cells to about 106 cells, about 106 cells to about 108 cells, about 108 to about 1011 cells, or about 108 to about 5×1010 cells. A dose may also about 106 cells/m2 to about 1010 cells/m2, or about 106 cells/m2 to about 108 cells/m2.

Proteins(s) of the invention, and/or nucleic acids encoding the protein(s), can also be administered via liposomes, which may also serve to: 1) target the proteins(s) to a particular tissue, such as lymphoid tissue; 2) to target selectively to diseases cells; or, 3) to increase the half-life of the peptide composition. Liposomes include emulsions, foams, micelles, insoluble monolayers, liquid crystals, phospholipid dispersions, lamellar layers and the like. In these preparations, the peptide to be delivered is incorporated as part of a liposome, alone or in conjunction with a molecule which binds to a receptor prevalent among lymphoid cells, such as monoclonal antibodies which bind to the CD45 antigen, or with other therapeutic or immunogenic compositions. Thus, liposomes either filled or decorated with a desired peptide of the invention can be directed to the site of lymphoid cells, where the liposomes then deliver the peptide compositions. Liposomes for use in accordance with the invention are formed from standard vesicle-forming lipids, which generally include neutral and negatively charged phospholipids and a sterol, such as cholesterol. The selection of lipids is generally guided by consideration of, e.g., liposome size, acid lability and stability of the liposomes in the blood stream. A variety of methods are available for preparing liposomes, as described in, e.g., Szoka, et al., Ann. Rev. Biophys. Bioeng. 9:467 (1980), and U.S. Pat. Nos. 4,235,871, 4,501,728, 4,837,028, and 5,019,369.

For targeting cells of the immune system, a ligand to be incorporated into the liposome can include, e.g., antibodies or fragments thereof specific for cell surface determinants of the desired immune system cells. A liposome suspension containing a peptide may be administered intravenously, locally, topically, etc. in a dose which varies according to, inter alia, the manner of administration, the peptide being delivered, and the stage of the disease being treated.

For solid compositions, conventional nontoxic solid carriers may be used which include, for example, pharmaceutical grades of mannitol, lactose, starch, magnesium stearate, sodium saccharin, talcum, cellulose, glucose, sucrose, magnesium carbonate, and the like. For oral administration, a pharmaceutically acceptable nontoxic composition is formed by incorporating any of the normally employed excipients, such as those carriers previously listed, and generally 10-95% of active ingredient, that is, one or more peptides of the invention, and more preferably at a concentration of 25%-75%.

For aerosol administration, immunogenic peptides are preferably supplied in finely divided form along with a surfactant and propellant. Typical percentages of peptides are about 0.01%-20% by weight, preferably about 1%-10%. The surfactant must, of course, be nontoxic, and preferably soluble in the propellant. Representative of such agents are the esters or partial esters of fatty acids containing from about 6 to 22 carbon atoms, such as caproic, octanoic, lauric, palmitic, stearic, linoleic, linolenic, olesteric and oleic acids with an aliphatic polyhydric alcohol or its cyclic anhydride. Mixed esters, such as mixed or natural glycerides may be employed. The surfactant may constitute about 0.1%-20% by weight of the composition, preferably about 0.25-5%. The balance of the composition is ordinarily propellant. A carrier can also be included, as desired, as with, e.g., lecithin for intranasal delivery.

XI.) Diagnostic and Prognostic Embodiments of 58P1D12.

As disclosed herein, 58P1D12 polynucleotides, polypeptides, reactive cytotoxic T cells (CTL), reactive helper T cells (HTL) and anti-polypeptide antibodies are used in well known diagnostic, prognostic and therapeutic assays that examine conditions associated with dysregulated cell growth such as cancer, in particular the cancers listed in Table I (see, e.g., both its specific pattern of tissue expression as well as its overexpression in certain cancers as described for example in the Example entitled "Expression analysis of 58P1D12 in normal tissues, and patient specimens").

58P1D12 can be analogized to a prostate associated antigen PSA, the archetypal marker that has been used by medical practitioners for years to identify and monitor the presence of prostate cancer (see, e.g., Merrill et al., J. Urol. 163(2): 503-5120 (2000); Polascik et al., J. Urol. August; 162(2):293-306 (1999) and Fortier et al., J. Nat. Cancer Inst. 91(19): 1635-1640(1999)). A variety of other diagnostic markers are also used in similar contexts including p53 and K-ras (see, e.g., Tulchinsky et al., Int J Mol Med 1999 July 4(1):99-102 and Minimoto et al., Cancer Detect Prev 2000; 24(1):1-12). Therefore, this disclosure of 58P1D12 polynucleotides and polypeptides (as well as 58P1D12 polynucleotide probes and anti-58P1D12 antibodies used to identify the presence of these molecules) and their properties allows skilled artisans to utilize these molecules in methods that are analogous to those used, for example, in a variety of diagnostic assays directed to examining conditions associated with cancer.

Typical embodiments of diagnostic methods which utilize the 58P1D12 polynucleotides, polypeptides, reactive T cells and antibodies are analogous to those methods from well-established diagnostic assays, which employ, e.g., PSA polynucleotides, polypeptides, reactive T cells and antibodies. For example, just as PSA polynucleotides are used as probes (for example in Northern analysis, see, e.g., Sharief et al., Biochem. Mol. Biol. Int. 33(3):567-74(1994)) and primers (for example in PCR analysis, see, e.g., Okegawa et al., J. Urol. 163(4): 1189-1190 (2000)) to observe the presence and/or the level of PSA mRNAs in methods of monitoring PSA overexpression or the metastasis of prostate cancers, the 58P1D12 polynucleotides described herein can be utilized in the same way to detect 58P1D12 overexpression or the metastasis of prostate and other cancers expressing this gene. Alternatively, just as PSA polypeptides are used to generate antibodies specific for PSA which can then be used to observe the presence and/or the level of PSA proteins in methods to monitor PSA protein overexpression (see, e.g., Stephan et al., Urology 55(4):560-3 (2000)) or the metastasis of prostate cells (see, e.g., Alanen et al., Pathol. Res. Pract. 192(3):233-7 (1996)), the 58P1D12 polypeptides described herein can be utilized to generate antibodies for use in detecting 58P1D12 overexpression or the metastasis of prostate cells and cells of other cancers expressing this gene.

Specifically, because metastases involves the movement of cancer cells from an organ of origin (such as the lung or prostate gland etc.) to a different area of the body (such as a lymph node), assays which examine a biological sample for the presence of cells expressing 58P1D12 polynucleotides and/or polypeptides can be used to provide evidence of metastasis. For example, when a biological sample from tissue that does not normally contain 58P1D12-expressing cells (lymph node) is found to contain 58P1D12-expressing cells such as the 58P1D12 expression seen in LAPC4 and LAPC9, xenografts isolated from lymph node and bone metastasis, respectively, this finding is indicative of metastasis.

Alternatively 58P1D12 polynucleotides and/or polypeptides can be used to provide evidence of cancer, for example, when cells in a biological sample that do not normally express 58P1D12 or express 58P1D12 at a different level are found to express 58P1D12 or have an increased expression of 58P1D12 (see, e.g., the 58P1D12 expression in the cancers listed in Table 1 and in patient samples etc. shown in the accompanying Figures). In such assays, artisans may further wish to generate supplementary evidence of metastasis by testing the biological sample for the presence of a second tissue restricted marker (in addition to 58P1D12) such as PSA, PSCA etc. (see, e.g., Alanen et al., Pathol. Res. Pract. 192(3): 233-237 (1996)).

The use of immunohistochemistry to identify the presence of a 58P1D12 polypeptide within a tissue section can indicate an altered state of certain cells within that tissue. It is well understood in the art that the ability of an antibody to localize to a polypeptide that is expressed in cancer cells is a way of diagnosing presence of disease, disease stage, progression and/or tumor aggressiveness. Such an antibody can also detect an altered distribution of the polypeptide within the cancer cells, as compared to corresponding non-malignant tissue.

The 58P1D12 polypeptide and immunogenic compositions are also useful in view of the phenomena of altered subcellular protein localization in disease states. Alteration of cells from normal to diseased state causes changes in cellular morphology and is often associated with changes in subcellular protein localization/distribution. For example, cell membrane proteins that are expressed in a polarized manner in normal cells can be altered in disease, resulting in distribution of the protein in a non-polar manner over the whole cell surface.

The phenomenon of altered subcellular protein localization in a disease state has been demonstrated with MUC1 and Her2 protein expression by use of immunohistochemical means. Normal epithelial cells have a typical apical distribution of MUC1, in addition to some supranuclear localization of the glycoprotein, whereas malignant lesions often demonstrate an apolar staining pattern (Diaz et al, The Breast Journal, 7; 40-45 (2001); Zhang et al, Clinical Cancer Research, 4; 2669-2676 (1998): Cao, et al, The Journal of Histochemistry and Cytochemistry, 45: 1547-1557 (1997)). In addition, normal breast epithelium is either negative for Her2 protein or exhibits only a basolateral distribution whereas malignant cells can express the protein over the whole cell surface (De Potter, et al, International Journal of Cancer, 44; 969-974 (1989): McCormick, et al, 117; 935-943 (2002)). Alternatively, distribution of the protein may be altered from a surface only localization to include diffuse cytoplasmic expression in the diseased state. Such an example can be seen with MUC1 (Diaz, et al, The Breast Journal, 7: 40-45 (2001)).

Alteration in the localization/distribution of a protein in the cell, as detected by immunohistochemical methods, can also provide valuable information concerning the favorability of certain treatment modalities. This last point is illustrated by a situation where a protein may be intracellular in normal tissue, but cell surface in malignant cells; the cell surface location makes the cells favorably amenable to antibody-based diagnostic and treatment regimens. When such an alteration of protein localization occurs for 58P1D12, the 58P1D12 protein and immune responses related thereto are very useful. Accordingly, the ability to determine whether alteration of subcellular protein localization occurred for 24P4C12 make the 58P1D12 protein and immune responses related thereto very useful. Use of the 58P1D12 compositions allows those skilled in the art to make important diagnostic and therapeutic decisions.

Immunohistochemical reagents specific to 58P1D12 are also useful to detect metastases of tumors expressing 58P1D12 when the polypeptide appears in tissues where 58P1D12 is not normally produced.

Thus, 58P1D12 polypeptides and antibodies resulting from immune responses thereto are useful in a variety of important contexts such as diagnostic, prognostic, preventative and/or therapeutic purposes known to those skilled in the art.

Just as PSA polynucleotide fragments and polynucleotide variants are employed by skilled artisans for use in methods of monitoring PSA, 58P1D12 polynucleotide fragments and polynucleotide variants are used in an analogous manner. In particular, typical PSA polynucleotides used in methods of monitoring PSA are probes or primers which consist of fragments of the PSA cDNA sequence. Illustrating this, primers used to PCR amplify a PSA polynucleotide must include less than the whole PSA sequence to function in the polymerase chain reaction. In the context of such PCR reactions, skilled artisans generally create a variety of different polynucleotide fragments that can be used as primers in order to amplify different portions of a polynucleotide of interest or to optimize amplification reactions (see, e.g., Caetano-Anolles, G. Biotechniques 25(3): 472-476, 478-480 (1998); Robertson et al., Methods Mol. Biol. 98:121-154 (1998)). An additional illustration of the use of such fragments is provided in the Example entitled "Expression analysis of 58P1D12 in normal tissues, and patient specimens," where a 58P1D12 polynucleotide fragment is used as a probe to show the expression of 58P1D12 RNAs in cancer cells. In addition, variant polynucleotide sequences are typically used as primers and probes for the corresponding mRNAs in PCR and Northern analyses (see, e.g., Sawai et al., Fetal Diagn. Ther. 1996 November-December 11(6):407-13 and Current Protocols In Molecular Biology, Volume 2, Unit 2, Frederick M. Ausubel et al. eds., 1995)). Polynucleotide fragments and variants are useful in this context where they are capable of binding to a target polynucleotide sequence (e.g., a 58P1D12 polynucleotide shown in FIG. **2** or variant thereof) under conditions of high stringency.

Furthermore, PSA polypeptides which contain an epitope that can be recognized by an antibody or T cell that specifically binds to that epitope are used in methods of monitoring

PSA. 58P1D12 polypeptide fragments and polypeptide analogs or variants can also be used in an analogous manner. This practice of using polypeptide fragments or polypeptide variants to generate antibodies (such as anti-PSA antibodies or T cells) is typical in the art with a wide variety of systems such as fusion proteins being used by practitioners (see, e.g., Current Protocols In Molecular Biology, Volume 2, Unit 16, Frederick M. Ausubel et al. eds., 1995). In this context, each epitope(s) functions to provide the architecture with which an antibody or T cell is reactive. Typically, skilled artisans create a variety of different polypeptide fragments that can be used in order to generate immune responses specific for different portions of a polypeptide of interest (see, e.g., U.S. Pat. No. 5,840,501 and U.S. Pat. No. 5,939,533). For example it may be preferable to utilize a polypeptide comprising one of the 58P1D12 biological motifs discussed herein or a motif-bearing subsequence which is readily identified by one of skill in the art based on motifs available in the art. Polypeptide fragments, variants or analogs are typically useful in this context as long as they comprise an epitope capable of generating an antibody or T cell specific for a target polypeptide sequence (e.g. a 58P1D12 polypeptide shown in FIG. **3**).

As shown herein, the 58P1D12 polynucleotides and polypeptides (as well as the 58P1D12 polynucleotide probes and anti-58P1D12 antibodies or T cells used to identify the presence of these molecules) exhibit specific properties that make them useful in diagnosing cancers such as those listed in Table 1. Diagnostic assays that measure the presence of 58P1D12 gene products, in order to evaluate the presence or onset of a disease condition described herein, such as prostate cancer, are used to identify patients for preventive measures or further monitoring, as has been done so successfully with PSA. Moreover, these materials satisfy a need in the art for molecules having similar or complementary characteristics to PSA in situations where, for example, a definite diagnosis of metastasis of prostatic origin cannot be made on the basis of a test for PSA alone (see, e.g., Alanen et al., Pathol. Res. Pract. 192(3): 233-237 (1996)), and consequently, materials such as 58P1D12 polynucleotides and polypeptides (as well as the 58P1D12 polynucleotide probes and anti-58P1D12 antibodies used to identify the presence of these molecules) need to be employed to confirm a metastases of prostatic origin.

Finally, in addition to their use in diagnostic assays, the 58P1D12 polynucleotides disclosed herein have a number of other utilities such as their use in the identification of oncogenetic associated chromosomal abnormalities in the chromosomal region to which the 58P1D12 gene maps (see the Example entitled "Chromosomal Mapping of 58P1D12" below). Moreover, in addition to their use in diagnostic assays, the 58P1D12-related proteins and polynucleotides disclosed herein have other utilities such as their use in the forensic analysis of tissues of unknown origin (see, e.g., Takahama K Forensic Sci Int 1996 Jun. 28; 80(1-2): 63-9).

Additionally, 58P1D12-related proteins or polynucleotides of the invention can be used to treat a pathologic condition characterized by the over-expression of 58P1D12. For example, the amino acid or nucleic acid sequence of FIG. **2** or FIG. **3**, or fragments of either, can be used to generate an immune response to a 58P1D12 antigen. Antibodies or other molecules that react with 58P1D12 can be used to modulate the function of this molecule, and thereby provide a therapeutic benefit.

XII.) Inhibition of 58P1D12 Protein Function

The invention includes various methods and compositions for inhibiting the binding of 58P1D12 to its binding partner or its association with other protein(s) as well as methods for inhibiting 58P1D12 function.

XII.A.) Inhibition of 58P1D12 with Intracellular Antibodies

In one approach, a recombinant vector that encodes single chain antibodies that specifically bind to 58P1D12 are introduced into 58P1D12 expressing cells via gene transfer technologies. Accordingly, the encoded single chain anti-58P1D12 antibody is expressed intracellularly, binds to 58P1D12 protein, and thereby inhibits its function. Methods for engineering such intracellular single chain antibodies are well known. Such intracellular antibodies, also known as "intrabodies", are specifically targeted to a particular compartment within the cell, providing control over where the inhibitory activity of the treatment is focused. This technology has been successfully applied in the art (for review, see Richardson and Marasco, 1995, TIBTECH vol. 13). Intrabodies have been shown to virtually eliminate the expression of otherwise abundant cell surface receptors (see, e.g., Richardson et al., 1995, Proc. Natl. Acad. Sci. USA 92: 3137-3141; Beerli et al., 1994, J. Biol. Chem. 289: 23931-23936; Deshane et al., 1994, Gene Ther. 1: 332-337).

Single chain antibodies comprise the variable domains of the heavy and light chain joined by a flexible linker polypeptide, and are expressed as a single polypeptide. Optionally, single chain antibodies are expressed as a single chain variable region fragment joined to the light chain constant region. Well-known intracellular trafficking signals are engineered into recombinant polynucleotide vectors encoding such single chain antibodies in order to target precisely the intrabody to the desired intracellular compartment. For example, intrabodies targeted to the endoplasmic reticulum (ER) are engineered to incorporate a leader peptide and, optionally, a C-terminal ER retention signal, such as the KDEL amino acid motif. Intrabodies intended to exert activity in the nucleus are engineered to include a nuclear localization signal. Lipid moieties are joined to intrabodies in order to tether the intrabody to the cytosolic side of the plasma membrane. Intrabodies can also be targeted to exert function in the cytosol. For example, cytosolic intrabodies are used to sequester factors within the cytosol, thereby preventing them from being transported to their natural cellular destination.

In one embodiment, intrabodies are used to capture 58P1D12 in the nucleus, thereby preventing its activity within the nucleus. Nuclear targeting signals are engineered into such 58P1D12 intrabodies in order to achieve the desired targeting. Such 58P1D12 intrabodies are designed to bind specifically to a particular 58P1D12 domain. In another embodiment, cytosolic intrabodies that specifically bind to a 58P1D12 protein are used to prevent 58P1D12 from gaining access to the nucleus, thereby preventing it from exerting any biological activity within the nucleus (e.g., preventing 58P1D12 from forming transcription complexes with other factors).

In order to specifically direct the expression of such intrabodies to particular cells, the transcription of the intrabody is placed under the regulatory control of an appropriate tumor-specific promoter and/or enhancer. In order to target intrabody expression specifically to prostate, for example, the PSA promoter and/or promoter/enhancer can be utilized (See, for example, U.S. Pat. No. 5,919,652 issued 6 Jul. 1999).

XII.B.) Inhibition of 58P1D12 with Recombinant Proteins

In another approach, recombinant molecules bind to 58P1D12 and thereby inhibit 58P1D12 function. For example, these recombinant molecules prevent or inhibit 58P1D12 from accessing/binding to its binding partner(s) or associating with other protein(s). Such recombinant molecules can, for example, contain the reactive part(s) of a 58P1D12 specific antibody molecule. In a particular embodiment, the 58P1D12 binding domain of a 58P1D12 binding partner is engineered into a dimeric fusion protein, whereby the fusion protein comprises two 58P1D12 ligand binding domains linked to the Fc portion of a human IgG, such as human IgG1. Such IgG portion can contain, for example, the CH2 and CH3 domains and the hinge region, but not the CH1 domain. Such dimeric fusion proteins are administered in soluble form to patients suffering from a cancer associated with the expression of 58P1D12, whereby the dimeric fusion protein specifically binds to 58P1D12 and blocks 58P1D12 interaction with a binding partner. Such dimeric fusion proteins are further combined into multimeric proteins using known antibody linking technologies.

XII.C.) Inhibition of 58P1D12 Transcription or Translation

The present invention also comprises various methods and compositions for inhibiting the transcription of the 58P1D12 gene. Similarly, the invention also provides methods and compositions for inhibiting the translation of 58P1D12 mRNA into protein.

In one approach, a method of inhibiting the transcription of the 58P1D12 gene comprises contacting the 58P1D12 gene with a 58P1D12 antisense polynucleotide. In another approach, a method of inhibiting 58P1D12 mRNA translation comprises contacting a 58P1D12 mRNA with an antisense polynucleotide. In another approach, a 58P1D12 specific ribozyme is used to cleave a 58P1D12 message, thereby inhibiting translation. Such antisense and ribozyme based methods can also be directed to the regulatory regions of the 58P1D12 gene, such as 58P1D12 promoter and/or enhancer elements. Similarly, proteins capable of inhibiting a 58P1D12 gene transcription factor are used to inhibit 58P1D12 mRNA transcription. The various polynucleotides and compositions useful in the aforementioned methods have been described above. The use of antisense and ribozyme molecules to inhibit transcription and translation is well known in the art.

Other factors that inhibit the transcription of 58P1D12 by interfering with 58P1D12 transcriptional activation are also useful to treat cancers expressing 58P1D12. Similarly, factors that interfere with 58P1D12 processing are useful to treat cancers that express 58P1D12. Cancer treatment methods utilizing such factors are also within the scope of the invention.

XII.D.) General Considerations for Therapeutic Strategies

Gene transfer and gene therapy technologies can be used to deliver therapeutic polynucleotide molecules to tumor cells synthesizing 58P1D12 (i.e., antisense, ribozyme, polynucleotides encoding intrabodies and other 58P1D12 inhibitory molecules). A number of gene therapy approaches are known in the art. Recombinant vectors encoding 58P1D12 antisense polynucleotides, ribozymes, factors capable of interfering with 58P1D12 transcription, and so forth, can be delivered to target tumor cells using such gene therapy approaches.

The above therapeutic approaches can be combined with any one of a wide variety of surgical, chemotherapy or radiation therapy regimens. The therapeutic approaches of the invention can enable the use of reduced dosages of chemotherapy (or other therapies) and/or less frequent administration, an advantage for all patients and particularly for those that do not tolerate the toxicity of the chemotherapeutic agent well.

The anti-tumor activity of a particular composition (e.g., antisense, ribozyme, intrabody), or a combination of such compositions, can be evaluated using various in vitro and in vivo assay systems. In vitro assays that evaluate therapeutic activity include cell growth assays, soft agar assays and other assays indicative of tumor promoting activity, binding assays capable of determining the extent to which a therapeutic composition will inhibit the binding of 58P1D12 to a binding partner, etc.

In vivo, the effect of a 58P1D12 therapeutic composition can be evaluated in a suitable animal model. For example, xenogenic prostate cancer models can be used, wherein human prostate cancer explants or passaged xenograft tissues are introduced into immune compromised animals, such as nude or SCID mice (Klein et al., 1997, Nature Medicine 3: 402-408). For example, PCT Patent Application WO98/16628 and U.S. Pat. No. 6,107,540 describe various xenograft models of human prostate cancer capable of recapitulating the development of primary tumors, micrometastasis, and the formation of osteoblastic metastases characteristic of late stage disease. Efficacy can be predicted using assays that measure inhibition of tumor formation, tumor regression or metastasis, and the like.

In vivo assays that evaluate the promotion of apoptosis are useful in evaluating therapeutic compositions. In one embodiment, xenografts from tumor bearing mice treated with the therapeutic composition can be examined for the presence of apoptotic foci and compared to untreated control xenograft-bearing mice. The extent to which apoptotic foci are found in the tumors of the treated mice provides an indication of the therapeutic efficacy of the composition.

The therapeutic compositions used in the practice of the foregoing methods can be formulated into pharmaceutical compositions comprising a carrier suitable for the desired delivery method. Suitable carriers include any material that when combined with the therapeutic composition retains the anti-tumor function of the therapeutic composition and is generally non-reactive with the patient's immune system. Examples include, but are not limited to, any of a number of standard pharmaceutical carriers such as sterile phosphate buffered saline solutions, bacteriostatic water, and the like (see, generally, Remington's Pharmaceutical Sciences 16th Edition, A. Osal., Ed., 1980).

Therapeutic formulations can be solubilized and administered via any route capable of delivering the therapeutic composition to the tumor site. Potentially effective routes of administration include, but are not limited to, intravenous, parenteral, intraperitoneal, intramuscular, intratumor, intradermal, intraorgan, orthotopic, and the like. A preferred formulation for intravenous injection comprises the therapeutic composition in a solution of preserved bacteriostatic water, sterile unpreserved water, and/or diluted in polyvinylchloride or polyethylene bags containing 0.9% sterile Sodium Chloride for Injection, USP. Therapeutic protein preparations can be lyophilized and stored as sterile powders, preferably under vacuum, and then reconstituted in bacteriostatic water (containing for example, benzyl alcohol preservative) or in sterile water prior to injection.

Dosages and administration protocols for the treatment of cancers using the foregoing methods will vary with the method and the target cancer, and will generally depend on a number of other factors appreciated in the art.

XIII.) Identification, Characterization and Use of Modulators of 58P1D12

Methods to Identify and Use Modulators

In one embodiment, screening is performed to identify modulators that induce or suppress a particular expression profile, suppress or induce specific pathways, preferably generating the associated phenotype thereby. In another embodiment, having identified differentially expressed genes important in a particular state; screens are performed to identify modulators that alter expression of individual genes, either increase or decrease. In another embodiment, screening is performed to identify modulators that alter a biological function of the expression product of a differentially expressed gene. Again, having identified the importance of a gene in a particular state, screens are performed to identify agents that bind and/or modulate the biological activity of the gene product.

In addition, screens are done for genes that are induced in response to a candidate agent. After identifying a modulator (one that suppresses a cancer expression pattern leading to a normal expression pattern, or a modulator of a cancer gene that leads to expression of the gene as in normal tissue) a screen is performed to identify genes that are specifically modulated in response to the agent. Comparing expression profiles between normal tissue and agent-treated cancer tissue reveals genes that are not expressed in normal tissue or cancer tissue, but are expressed in agent treated tissue, and vice versa. These agent-specific sequences are identified and used by methods described herein for cancer genes or proteins. In particular these sequences and the proteins they encode are used in marking or identifying agent-treated cells. In addition, antibodies are raised against the agent-induced proteins and used to target novel therapeutics to the treated cancer tissue sample.

Modulator-Related Identification and Screening Assays:

Gene Expression-Related Assays

Proteins, nucleic acids, and antibodies of the invention are used in screening assays. The cancer-associated proteins, antibodies, nucleic acids, modified proteins and cells containing these sequences are used in screening assays, such as evaluating the effect of drug candidates on a "gene expression profile," expression profile of polypeptides or alteration of biological function. In one embodiment, the expression profiles are used, preferably in conjunction with high throughput screening techniques to allow monitoring for expression profile genes after treatment with a candidate agent (e.g., Davis, G F, et al, J Biol Screen 7:69 (2002); Zlokarnik, et al., Science 279:84-8 (1998); Heid, Genome Res 6:986-94, 1996).

The cancer proteins, antibodies, nucleic acids, modified proteins and cells containing the native or modified cancer proteins or genes are used in screening assays. That is, the present invention comprises methods for screening for compositions which modulate the cancer phenotype or a physiological function of a cancer protein of the invention. This is done on a gene itself or by evaluating the effect of drug candidates on a "gene expression profile" or biological function. In one embodiment, expression profiles are used, preferably in conjunction with high throughput screening techniques to allow monitoring after treatment with a candidate agent, see Zlokarnik, supra.

A variety of assays are executed directed to the genes and proteins of the invention. Assays are run on an individual nucleic acid or protein level. That is, having identified a particular gene as up regulated in cancer, test compounds are screened for the ability to modulate gene expression or for binding to the cancer protein of the invention. "Modulation" in this context includes an increase or a decrease in gene expression. The preferred amount of modulation will depend on the original change of the gene expression in normal versus tissue undergoing cancer, with changes of at least 10%, preferably 50%, more preferably 100-300%, and in some embodiments 300-1000% or greater. Thus, if a gene exhibits a 4-fold increase in cancer tissue compared to normal tissue, a decrease of about four-fold is often desired; similarly, a 10-fold decrease in cancer tissue compared to normal tissue a target value of a 10-fold increase in expression by the test compound is often desired. Modulators that exacerbate the type of gene expression seen in cancer are also useful, e.g., as an upregulated target in further analyses.

The amount of gene expression is monitored using nucleic acid probes and the quantification of gene expression levels, or, alternatively, a gene product itself is monitored, e.g., through the use of antibodies to the cancer protein and standard immunoassays. Proteomics and separation techniques also allow for quantification of expression.

Expression Monitoring to Identify Compounds that Modify Gene Expression

In one embodiment, gene expression monitoring, i.e., an expression profile, is monitored simultaneously for a number of entities. Such profiles will typically involve one or more of the genes of FIG. **2**. In this embodiment, e.g., cancer nucleic acid probes are attached to biochips to detect and quantify cancer sequences in a particular cell. Alternatively, PCR can be used. Thus, a series, e.g., wells of a microtiter plate, can be used with dispensed primers in desired wells. A PCR reaction can then be performed and analyzed for each well.

Expression monitoring is performed to identify compounds that modify the expression of one or more cancer-associated sequences, e.g., a polynucleotide sequence set out in FIG. **2**. Generally, a test modulator is added to the cells prior to analysis. Moreover, screens are also provided to identify agents that modulate cancer, modulate cancer proteins of the invention, bind to a cancer protein of the invention, or interfere with the binding of a cancer protein of the invention and an antibody or other binding partner.

In one embodiment, high throughput screening methods involve providing a library containing a large number of potential therapeutic compounds (candidate compounds). Such "combinatorial chemical libraries" are then screened in one or more assays to identify those library members (particular chemical species or subclasses) that display a desired characteristic activity. The compounds thus identified can serve as conventional "lead compounds," as compounds for screening, or as therapeutics.

In certain embodiments, combinatorial libraries of potential modulators are screened for an ability to bind to a cancer polypeptide or to modulate activity. Conventionally, new chemical entities with useful properties are generated by identifying a chemical compound (called a "lead compound") with some desirable property or activity, e.g., inhibiting activity, creating variants of the lead compound, and evaluating the property and activity of those variant compounds. Often, high throughput screening (HTS) methods are employed for such an analysis.

As noted above, gene expression monitoring is conveniently used to test candidate modulators (e.g., protein, nucleic acid or small molecule). After the candidate agent has been added and the cells allowed to incubate for a period, the sample containing a target sequence to be analyzed is, e.g., added to a biochip.

If required, the target sequence is prepared using known techniques. For example, a sample is treated to lyse the cells, using known lysis buffers, electroporation, etc., with purification and/or amplification such as PCR performed as appropriate. For example, an in vitro transcription with labels covalently attached to the nucleotides is performed. Generally, the nucleic acids are labeled with biotin-FITC or PE, or with cy3 or cy5.

The target sequence can be labeled with, e.g., a fluorescent, a chemiluminescent, a chemical, or a radioactive signal, to provide a means of detecting the target sequence's specific binding to a probe. The label also can be an enzyme, such as alkaline phosphatase or horseradish peroxidase, which when provided with an appropriate substrate produces a product that is detected. Alternatively, the label is a labeled compound or small molecule, such as an enzyme inhibitor, that binds but is not catalyzed or altered by the enzyme. The label also can be a moiety or compound, such as, an epitope tag or biotin which specifically binds to streptavidin. For the example of biotin, the streptavidin is labeled as described above, thereby, providing a detectable signal for the bound target sequence. Unbound labeled streptavidin is typically removed prior to analysis.

As will be appreciated by those in the art, these assays can be direct hybridization assays or can comprise "sandwich assays", which include the use of multiple probes, as is generally outlined in U.S. Pat. Nos. 5,681,702; 5,597,909; 5,545,730; 5,594,117; 5,591,584; 5,571,670; 5,580,731; 5,571,670; 5,591,584; 5,624,802; 5,635,352; 5,594,118; 5,359,100; 5,124,246; and 5,681,697. In this embodiment, in general, the target nucleic acid is prepared as outlined above, and then added to the biochip comprising a plurality of nucleic acid probes, under conditions that allow the formation of a hybridization complex.

A variety of hybridization conditions are used in the present invention, including high, moderate and low stringency conditions as outlined above. The assays are generally run under stringency conditions which allow formation of the label probe hybridization complex only in the presence of target. Stringency can be controlled by altering a step parameter that is a thermodynamic variable, including, but not limited to, temperature, formamide concentration, salt concentration, chaotropic salt concentration pH, organic solvent concentration, etc. These parameters may also be used to control non-specific binding, as is generally outlined in U.S. Pat. No. 5,681,697. Thus, it can be desirable to perform certain steps at higher stringency conditions to reduce non-specific binding.

The reactions outlined herein can be accomplished in a variety of ways. Components of the reaction can be added simultaneously, or sequentially, in different orders, with preferred embodiments outlined below. In addition, the reaction may include a variety of other reagents. These include salts, buffers, neutral proteins, e.g. albumin, detergents, etc. which can be used to facilitate optimal hybridization and detection, and/or reduce nonspecific or background interactions. Reagents that otherwise improve the efficiency of the assay, such as protease inhibitors, nuclease inhibitors, anti-microbial agents, etc., may also be used as appropriate, depending on the sample preparation methods and purity of the target. The assay data are analyzed to determine the expression levels of individual genes, and changes in expression levels as between states, forming a gene expression profile.

Biological Activity-Related Assays

The invention provides methods identify or screen for a compound that modulates the activity of a cancer-related gene or protein of the invention. The methods comprise adding a test compound, as defined above, to a cell comprising a cancer protein of the invention. The cells contain a recombinant nucleic acid that encodes a cancer protein of the invention. In another embodiment, a library of candidate agents is tested on a plurality of cells.

In one aspect, the assays are evaluated in the presence or absence or previous or subsequent exposure of physiological signals, e.g. hormones, antibodies, peptides, antigens, cytokines, growth factors, action potentials, pharmacological agents including chemotherapeutics, radiation, carcinogenics, or other cells (i.e., cell-cell contacts). In another example, the determinations are made at different stages of the cell cycle process. In this way, compounds that modulate genes or proteins of the invention are identified. Compounds with pharmacological activity are able to enhance or interfere with the activity of the cancer protein of the invention. Once identified, similar structures are evaluated to identify critical structural features of the compound.

In one embodiment, a method of modulating (e.g., inhibiting) cancer cell division is provided; the method comprises administration of a cancer modulator. In another embodiment, a method of modulating (e.g., inhibiting) cancer is provided; the method comprises administration of a cancer modulator. In a further embodiment, methods of treating cells or individuals with cancer are provided; the method comprises administration of a cancer modulator.

In one embodiment, a method for modulating the status of a cell that expresses a gene of the invention is provided. As used herein status comprises such art-accepted parameters such as growth, proliferation, survival, function, apoptosis, senescence, location, enzymatic activity, signal transduction, etc. of a cell. In one embodiment, a cancer inhibitor is an antibody as discussed above. In another embodiment, the cancer inhibitor is an antisense molecule. A variety of cell growth, proliferation, and metastasis assays are known to those of skill in the art, as described herein.

High Throughput Screening to Identify Modulators

The assays to identify suitable modulators are amenable to high throughput screening. Preferred assays thus detect enhancement or inhibition of cancer gene transcription, inhibition or enhancement of polypeptide expression, and inhibition or enhancement of polypeptide activity.

In one embodiment, modulators evaluated in high throughput screening methods are proteins, often naturally occurring proteins or fragments of naturally occurring proteins. Thus, e.g., cellular extracts containing proteins, or random or directed digests of proteinaceous cellular extracts, are used. In this way, libraries of proteins are made for screening in the methods of the invention. Particularly preferred in this embodiment are libraries of bacterial, fungal, viral, and mammalian proteins, with the latter being preferred, and human proteins being especially preferred. Particularly useful test compound will be directed to the class of proteins to which the target belongs, e.g., substrates for enzymes, or ligands and receptors.

Use of Soft Agar Growth and Colony Formation to Identify and Characterize Modulators

Normal cells require a solid substrate to attach and grow. When cells are transformed, they lose this phenotype and grow detached from the substrate. For example, transformed cells can grow in stirred suspension culture or suspended in semi-solid media, such as semi-solid or soft agar. The transformed cells, when transfected with tumor suppressor genes, can regenerate normal phenotype and once again require a solid substrate to attach to and grow. Soft agar growth or colony formation in assays are used to identify modulators of cancer sequences, which when expressed in host cells, inhibit abnormal cellular proliferation and transformation. A modulator reduces or eliminates the host cells' ability to grow suspended in solid or semisolid media, such as agar.

Techniques for soft agar growth or colony formation in suspension assays are described in Freshney, Culture of Animal Cells a Manual of Basic Technique (3rd ed., 1994). See also, the methods section of Garkavtsev et al. (1996), supra.

Evaluation of Contact lnhibition and Growth Density Limitation to ldentify and Characterize Modulators

Normal cells typically grow in a flat and organized pattern in cell culture until they touch other cells. When the cells touch one another, they are contact inhibited and stop growing. Transformed cells, however, are not contact inhibited and continue to grow to high densities in disorganized foci. Thus, transformed cells grow to a higher saturation density than corresponding normal cells. This is detected morphologically by the formation of a disoriented monolayer of cells or cells in foci. Alternatively, labeling index with (3H)-thymidine at saturation density is used to measure density limitation of growth, similarly an MTT or Alamar blue assay will reveal proliferation capacity of cells and the the ability of modulators to affect same. See Freshney (1994), supra. Transformed cells, when transfected with tumor suppressor genes, can regenerate a normal phenotype and become contact inhibited and would grow to a lower density.

ln this assay, labeling index with 3H)-thymidine at saturation density is a preferred method of measuring density limitation of growth. Transformed host cells are transfected with a cancer-associated sequence and are grown for 24 hours at saturation density in non-limiting medium conditions. The percentage of cells labeling with (3H)-thymidine is determined by incorporated cpm.

Contact independent growth is used to identify modulators of cancer sequences, which had led to abnormal cellular proliferation and transformation. A modulator reduces or eliminates contact independent growth, and returns the cells to a normal phenotype.

Evaluation of Growth Factor or Serum Dependence to ldentify and Characterize Modulators

Transformed cells have lower serum dependence than their normal counterparts (see, e.g., Temin, J. Natl. Cancer lnst. 37:167-175 (1966); Eagle et al., J. Exp. Med 131:836-879 (1970)); Freshney, supra. This is in part due to release of various growth factors by the transformed cells. The degree of growth factor or serum dependence of transformed host cells can be compared with that of control. For example, growth factor or serum dependence of a cell is monitored in methods to identify and characterize compounds that modulate cancer-associated sequences of the invention.

Use of Tumor-Specific Marker Levels to ldentify and Characterize Modulators

Tumor cells release an increased amount of certain factors (hereinafter "tumor specific markers") than their normal counterparts. For example, plasminogen activator (PA) is released from human glioma at a higher level than from normal brain cells (see, e.g., Gullino, Angiogenesis, Tumor Vascularization, and Potential lnterference with Tumor Growth, in Biological Responses in Cancer, pp. 178-184 (Mihich (ed.) 1985)). Similarly, Tumor Angiogenesis Factor (TAF) is released at a higher level in tumor cells than their normal counterparts. See, e.g., Folkman, Angiogenesis and Cancer, Sem Cancer Biol. (1992)), while bFGF is released from endothelial tumors (Ensoli, B et al).

Various techniques which measure the release of these factors are described in Freshney (1994), supra. Also, see, Unkless et al., J. Biol. Chem. 249:4295-4305 (1974); Strickland & Beers, J. Biol. Chem. 251:5694-5702 (1976); Whur et al., Br. J. Cancer 42:305 312 (1980); Gullino, Angiogenesis, Tumor Vascularization, and Potential lnterference with Tumor Growth, in Biological Responses in Cancer, pp. 178-184 (Mihich (ed.) 1985); Freshney, Anticancer Res. 5:111-130 (1985). For example, tumor specific marker levels are monitored in methods to identify and characterize compounds that modulate cancer-associated sequences of the invention.

lnvasiveness into Matrigel to ldentify and Characterize Modulators

The degree of invasiveness into Matrigel or an extracellular matrix constituent can be used as an assay to identify and characterize compounds that modulate cancer associated sequences. Tumor cells exhibit a positive correlation between malignancy and invasiveness of cells into Matrigel or some other extracellular matrix constituent. ln this assay, tumorigenic cells are typically used as host cells. Expression of a tumor suppressor gene in these host cells would decrease invasiveness of the host cells. Techniques described in Cancer Res. 1999; 59:6010; Freshney (1994), supra, can be used. Briefly, the level of invasion of host cells is measured by using filters coated with Matrigel or some other extracellular matrix constituent. Penetration into the gel, or through to the distal side of the filter, is rated as invasiveness, and rated histologically by number of cells and distance moved, or by prelabeling the cells with 1251 and counting the radioactivity on the distal side of the filter or bottom of the dish. See, e.g., Freshney (1984), supra.

Evaluation of Tumor Growth ln Vivo to ldentify and Characterize Modulators

Effects of cancer-associated sequences on cell growth are tested in transgenic or immune-suppressed organisms. Transgenic organisms are prepared in a variety of art-accepted ways. For example, knock-out transgenic organisms, e.g., mammals such as mice, are made, in which a cancer gene is disrupted or in which a cancer gene is inserted. Knock-out transgenic mice are made by insertion of a marker gene or other heterologous gene into the endogenous cancer gene site in the mouse genome via homologous recombination. Such mice can also be made by substituting the endogenous cancer gene with a mutated version of the cancer gene, or by mutating the endogenous cancer gene, e.g., by exposure to carcinogens.

To prepare transgenic chimeric animals, e.g., mice, a DNA construct is introduced into the nuclei of embryonic stem cells. Cells containing the newly engineered genetic lesion are injected into a host mouse embryo, which is re-implanted into a recipient female. Some of these embryos develop into chimeric mice that possess germ cells some of which are derived from the mutant cell line. Therefore, by breeding the chimeric mice it is possible to obtain a new line of mice containing the introduced genetic lesion (see, e.g., Capecchi et al., Science 244:1288 (1989)). Chimeric mice can be derived according to U.S. Pat. No. 6,365,797, issued 2 Apr. 2002; U.S. Pat. No. 6,107,540 issued 22 Aug. 2000; Hogan et al., Manipulating the Mouse Embryo: A laboratory Manual, Cold Spring Harbor Laboratory (1988) and Teratocarcinomas and Embryonic Stem Cells: A Practical Approach, Robertson, ed., lRL Press, Washington, D.C., (1987).

Alternatively, various immune-suppressed or immune-deficient host animals can be used. For example, a genetically athymic "nude" mouse (see, e.g., Giovanella et al., J. Natl. Cancer lnst. 52:921 (1974)), a SClD mouse, a thymectomized mouse, or an irradiated mouse (see, e.g., Bradley et al., Br. J. Cancer 38:263 (1978); Selby et al., Br. J. Cancer 41:52 (1980)) can be used as a host. Transplantable tumor cells (typically about 106 cells) injected into isogenic hosts produce invasive tumors in a high proportion of cases, while normal cells of similar origin will not. In hosts which developed invasive tumors, cells expressing cancer-associated sequences are injected subcutaneously or orthotopically. Mice are then separated into groups, including control groups and treated experimental groups) e.g. treated with a modulator). After a suitable length of time, preferably 4-8 weeks, tumor growth is measured (e.g., by volume or by its two largest dimensions, or weight) and compared to the control. Tumors that have statistically significant reduction (using, e.g., Student's T test) are said to have inhibited growth.

In Vitro Assays to Identify and Characterize Modulators

Assays to identify compounds with modulating activity can be performed in vitro. For example, a cancer polypeptide is first contacted with a potential modulator and incubated for a suitable amount of time, e.g., from 0.5 to 48 hours. In one embodiment, the cancer polypeptide levels are determined in vitro by measuring the level of protein or mRNA. The level of protein is measured using immunoassays such as Western blotting, ELISA and the like with an antibody that selectively binds to the cancer polypeptide or a fragment thereof. For measurement of mRNA, amplification, e.g., using PCR, LCR, or hybridization assays, e.g., Northern hybridization, RNAse protection, dot blotting, are preferred. The level of protein or mRNA is detected using directly or indirectly labeled detection agents, e.g., fluorescently or radioactively labeled nucleic acids, radioactively or enzymatically labeled antibodies, and the like, as described herein.

Alternatively, a reporter gene system can be devised using a cancer protein promoter operably linked to a reporter gene such as luciferase, green fluorescent protein, CAT, or P-gal. The reporter construct is typically transfected into a cell. After treatment with a potential modulator, the amount of reporter gene transcription, translation, or activity is measured according to standard techniques known to those of skill in the art (Davis G F, supra; Gonzalez, J. & Negulescu, P. Curr. Opin. Biotechnol. 1998: 9:624).

As outlined above, in vitro screens are done on individual genes and gene products. That is, having identified a particular differentially expressed gene as important in a particular state, screening of modulators of the expression of the gene or the gene product itself is performed.

In one embodiment, screening for modulators of expression of specific gene(s) is performed. Typically, the expression of only one or a few genes is evaluated. In another embodiment, screens are designed to first find compounds that bind to differentially expressed proteins. These compounds are then evaluated for the ability to modulate differentially expressed activity. Moreover, once initial candidate compounds are identified, variants can be further screened to better evaluate structure activity relationships.

Binding Assays to Identify and Characterize Modulators

In binding assays in accordance with the invention, a purified or isolated gene product of the invention is generally used. For example, antibodies are generated to a protein of the invention, and immunoassays are run to determine the amount and/or location of protein. Alternatively, cells comprising the cancer proteins are used in the assays.

Thus, the methods comprise combining a cancer protein of the invention and a candidate compound such as a ligand, and determining the binding of the compound to the cancer protein of the invention. Preferred embodiments utilize the human cancer protein; animal models of human disease of can also be developed and used. Also, other analogous mammalian proteins also can be used as appreciated by those of skill in the art. Moreover, in some embodiments variant or derivative cancer proteins are used.

Generally, the cancer protein of the invention, or the ligand, is non-diffusibly bound to an insoluble support. The support can, e.g., be one having isolated sample receiving areas (a microtiter plate, an array, etc.). The insoluble supports can be made of any composition to which the compositions can be bound, is readily separated from soluble material, and is otherwise compatible with the overall method of screening. The surface of such supports can be solid or porous and of any convenient shape.

Examples of suitable insoluble supports include microtiter plates, arrays, membranes and beads. These are typically made of glass, plastic (e.g., polystyrene), polysaccharide, nylon, nitrocellulose, or Teflon™, etc. Microtiter plates and arrays are especially convenient because a large number of assays can be carried out simultaneously, using small amounts of reagents and samples. The particular manner of binding of the composition to the support is not crucial so long as it is compatible with the reagents and overall methods of the invention, maintains the activity of the composition and is nondiffusable. Preferred methods of binding include the use of antibodies which do not sterically block either the ligand binding site or activation sequence when attaching the protein to the support, direct binding to "sticky" or ionic supports, chemical crosslinking, the synthesis of the protein or agent on the surface, etc. Following binding of the protein or lig and/binding agent to the support, excess unbound material is removed by washing. The sample receiving areas may then be blocked through incubation with bovine serum albumin (BSA), casein or other innocuous protein or other moiety.

Once a cancer protein of the invention is bound to the support, and a test compound is added to the assay. Alternatively, the candidate binding agent is bound to the support and the cancer protein of the invention is then added. Binding agents include specific antibodies, non-natural binding agents identified in screens of chemical libraries, peptide analogs, etc.

Of particular interest are assays to identify agents that have a low toxicity for human cells. A wide variety of assays can be used for this purpose, including proliferation assays, cAMP assays, labeled in vitro protein-protein binding assays, electrophoretic mobility shift assays, immunoassays for protein binding, functional assays (phosphorylation assays, etc.) and the like.

A determination of binding of the test compound (ligand, binding agent, modulator, etc.) to a cancer protein of the invention can be done in a number of ways. The test compound can be labeled, and binding determined directly, e.g., by attaching all or a portion of the cancer protein of the invention to a solid support, adding a labeled candidate compound (e.g., a fluorescent label), washing off excess reagent, and determining whether the label is present on the solid support. Various blocking and washing steps can be utilized as appropriate.

In certain embodiments, only one of the components is labeled, e.g., a protein of the invention or ligands labeled. Alternatively, more than one component is labeled with different labels, e.g., I125, for the proteins and a fluorophoro for the compound. Proximity reagents, e.g., quenching or energy transfer reagents are also useful.

Competitive Binding to Identify and Characterize Modulators

In one embodiment, the binding of the "test compounds" is determined by competitive binding assay with a "competitor." The competitor is a binding moiety that binds to the target molecule (e.g., a cancer protein of the invention). Competitors include compounds such as antibodies, peptides, binding partners, ligands, etc. Under certain circumstances, the competitive binding between the test compound and the competitor displaces the test compound. In one embodiment, the test compound is labeled. Either the test compound, the competitor, or both, is added to the protein for a time sufficient to allow binding. Incubations are performed at a temperature that facilitates optimal activity, typically between four and 40° C. Incubation periods are typically optimized, e.g., to facilitate rapid high throughput screening; typically between zero and one hour will be sufficient. Excess reagent is generally removed or washed away. The second component is then added, and the presence or absence of the labeled component is followed, to indicate binding.

In one embodiment, the competitor is added first, followed by the test compound. Displacement of the competitor is an indication that the test compound is binding to the cancer protein and thus is capable of binding to, and potentially modulating, the activity of the cancer protein. In this embodiment, either component can be labeled. Thus, e.g., if the competitor is labeled, the presence of label in the post-test compound wash solution indicates displacement by the test compound. Alternatively, if the test compound is labeled, the presence of the label on the support indicates displacement.

In an alternative embodiment, the test compound is added first, with incubation and washing, followed by the competitor. The absence of binding by the competitor indicates that the test compound binds to the cancer protein with higher affinity than the competitor. Thus, if the test compound is labeled, the presence of the label on the support, coupled with a lack of competitor binding, indicates that the test compound binds to and thus potentially modulates the cancer protein of the invention.

Accordingly, the competitive binding methods comprise differential screening to identity agents that are capable of modulating the activity of the cancer proteins of the invention. In this embodiment, the methods comprise combining a cancer protein and a competitor in a first sample. A second sample comprises a test compound, the cancer protein, and a competitor. The binding of the competitor is determined for both samples, and a change, or difference in binding between the two samples indicates the presence of an agent capable of binding to the cancer protein and potentially modulating its activity. That is, if the binding of the competitor is different in the second sample relative to the first sample, the agent is capable of binding to the cancer protein.

Alternatively, differential screening is used to identify drug candidates that bind to the native cancer protein, but cannot bind to modified cancer proteins. For example the structure of the cancer protein is modeled and used in rational drug design to synthesize agents that interact with that site, agents which generally do not bind to site-modified proteins. Moreover, such drug candidates that affect the activity of a native cancer protein are also identified by screening drugs for the ability to either enhance or reduce the activity of such proteins.

Positive controls and negative controls can be used in the assays. Preferably control and test samples are performed in at least triplicate to obtain statistically significant results. Incubation of all samples occurs for a time sufficient to allow for the binding of the agent to the protein. Following incubation, samples are washed free of non-specifically bound material and the amount of bound, generally labeled agent determined. For example, where a radiolabel is employed, the samples can be counted in a scintillation counter to determine the amount of bound compound.

A variety of other reagents can be included in the screening assays. These include reagents like salts, neutral proteins, e.g. albumin, detergents, etc. which are used to facilitate optimal protein-protein binding and/or reduce non-specific or background interactions. Also reagents that otherwise improve the efficiency of the assay, such as protease inhibitors, nuclease inhibitors, anti-microbial agents, etc., can be used. The mixture of components is added in an order that provides for the requisite binding.

Use of Polynucleotides to Down-Regulate or Inhibit a Protein of the Invention.

Polynucleotide modulators of cancer can be introduced into a cell containing the target nucleotide sequence by formation of a conjugate with a ligand-binding molecule, as described in WO 91/04753. Suitable ligand-binding molecules include, but are not limited to, cell surface receptors, growth factors, other cytokines, or other ligands that bind to cell surface receptors. Preferably, conjugation of the ligand binding molecule does not substantially interfere with the ability of the ligand binding molecule to bind to its corresponding molecule or receptor, or block entry of the sense or antisense oligonucleotide or its conjugated version into the cell. Alternatively, a polynucleotide modulator of cancer can be introduced into a cell containing the target nucleic acid sequence, e.g., by formation of a polynucleotide-lipid complex, as described in WO 90/10448. It is understood that the use of antisense molecules or knock out and knock in models may also be used in screening assays as discussed above, in addition to methods of treatment.

Inhibitory and Antisense Nucleotides

In certain embodiments, the activity of a cancer-associated protein is down-regulated, or entirely inhibited, by the use of antisense polynucleotide or inhibitory small nuclear RNA (snRNA), i.e., a nucleic acid complementary to, and which can preferably hybridize specifically to, a coding mRNA nucleic acid sequence, e.g., a cancer protein of the invention, mRNA, or a subsequence thereof. Binding of the antisense polynucleotide to the mRNA reduces the translation and/or stability of the mRNA.

In the context of this invention, antisense polynucleotides can comprise naturally occurring nucleotides, or synthetic species formed from naturally occurring subunits or their close homologs. Antisense polynucleotides may also have altered sugar moieties or inter-sugar linkages. Exemplary among these are the phosphorothioate and other sulfur containing species which are known for use in the art. Analogs are comprised by this invention so long as they function effectively to hybridize with nucleotides of the invention. See, e.g., Isis Pharmaceuticals, Carlsbad, Calif.; Sequitor, Inc., Natick, Mass.

Such antisense polynucleotides can readily be synthesized using recombinant means, or can be synthesized in vitro. Equipment for such synthesis is sold by several vendors, including Applied Biosystems. The preparation of other oligonucleotides such as phosphorothioates and alkylated derivatives is also well known to those of skill in the art.

Antisense molecules as used herein include antisense or sense oligonucleotides. Sense oligonucleotides can, e.g., be employed to block transcription by binding to the anti-sense strand. The antisense and sense oligonucleotide comprise a single stranded nucleic acid sequence (either RNA or DNA) capable of binding to target mRNA (sense) or DNA (antisense) sequences for cancer molecules. Antisense or sense oligonucleotides, according to the present invention, comprise a fragment generally at least about 12 nucleotides, preferably from about 12 to 30 nucleotides. The ability to derive an antisense or a sense oligonucleotide, based upon a cDNA sequence encoding a given protein is described in, e.g., Stein & Cohen (Cancer Res. 48:2659 (1988 and van der Krol et al. (BioTechniques 6:958 (1988)).

Ribozymes

In addition to antisense polynucleotides, ribozymes can be used to target and inhibit transcription of cancer-associated nucleotide sequences. A ribozyme is an RNA molecule that catalytically cleaves other RNA molecules. Different kinds of ribozymes have been described, including group I ribozymes, hammerhead ribozymes, hairpin ribozymes, RNase P, and axhead ribozymes (see, e.g., Castanotto et al., Adv. in Pharmacology 25: 289-317 (1994) for a general review of the properties of different ribozymes).

The general features of hairpin ribozymes are described, e.g., in Hampel et al., Nucl. Acids Res. 18:299-304 (1990); European Patent Publication No. 0360257; U.S. Pat. No. 5,254,678. Methods of preparing are well known to those of skill in the art (see, e.g., WO 94/26877; Ojwang et al., Proc. Natl. Acad. Sci. USA 90:6340-6344 (1993); Yamada et al., Human Gene Therapy 1:39-45 (1994); Leavitt et al., Proc. Natl. Acad. Sci. USA 92:699-703 (1995); Leavitt et al., Human Gene Therapy 5: 1151-120 (1994); and Yamada et al., Virology 205:121-126 (1994)).

Use of Modulators in Phenotypic Screening

In one embodiment, a test compound is administered to a population of cancer cells, which have an associated cancer expression profile. By "administration" or "contacting" herein is meant that the modulator is added to the cells in such a manner as to allow the modulator to act upon the cell, whether by uptake and intracellular action, or by action at the cell surface. In some embodiments, a nucleic acid encoding a proteinaceous agent (i.e., a peptide) is put into a viral construct such as an adenoviral or retroviral construct, and added to the cell, such that expression of the peptide agent is accomplished, e.g., PCT US97/01019. Regulatable gene therapy systems can also be used. Once the modulator has been administered to the cells, the cells are washed if desired and are allowed to incubate under preferably physiological conditions for some period. The cells are then harvested and a new gene expression profile is generated. Thus, e.g., cancer tissue is screened for agents that modulate, e.g., induce or suppress, the cancer phenotype. A change in at least one gene, preferably many, of the expression profile indicates that the agent has an effect on cancer activity. Similarly, altering a biological function or a signaling pathway is indicative of modulator activity. By defining such a signature for the cancer phenotype, screens for new drugs that alter the phenotype are devised. With this approach, the drug target need not be known and need not be represented in the original gene/protein expression screening platform, nor does the level of transcript for the target protein need to change. The modulator inhibiting function will serve as a surrogate marker.

As outlined above, screens are done to assess genes or gene products. That is, having identified a particular differentially expressed gene as important in a particular state, screening of modulators of either the expression of the gene or the gene product itself is performed.

Use of Modulators to Affect Peptides of the Invention

Measurements of cancer polypeptide activity, or of the cancer phenotype are performed using a variety of assays. For example, the effects of modulators upon the function of a cancer polypeptide(s) are measured by examining parameters described above. A physiological change that affects activity is used to assess the influence of a test compound on the polypeptides of this invention. When the functional outcomes are determined using intact cells or animals, a variety of effects can be assesses such as, in the case of a cancer associated with solid tumors, tumor growth, tumor metastasis, neovascularization, hormone release, transcriptional changes to both known and uncharacterized genetic markers (e.g., by Northern blots), changes in cell metabolism such as cell growth or pH changes, and changes in intracellular second messengers such as cGNIP.

Methods of Identifying Characterizing Cancer-Associated Sequences

Expression of various gene sequences is correlated with cancer. Accordingly, disorders based on mutant or variant cancer genes are determined. In one embodiment, the invention provides methods for identifying cells containing variant cancer genes, e.g., determining the presence of, all or part, the sequence of at least one endogenous cancer gene in a cell. This is accomplished using any number of sequencing techniques. The invention comprises methods of identifying the cancer genotype of an individual, e.g., determining all or part of the sequence of at least one gene of the invention in the individual. This is generally done in at least one tissue of the individual, e.g., a tissue set forth in Table I, and may include the evaluation of a number of tissues or different samples of the same tissue. The method may include comparing the sequence of the sequenced gene to a known cancer gene, i.e., a wild-type gene to determine the presence of family members, homologies, mutations or variants. The sequence of all or part of the gene can then be compared to the sequence of a known cancer gene to determine if any differences exist. This is done using any number of known homology programs, such as BLAST, Bestfit, etc. The presence of a difference in the sequence between the cancer gene of the patient and the known cancer gene correlates with a disease state or a propensity for a disease state, as outlined herein.

In a preferred embodiment, the cancer genes are used as probes to determine the number of copies of the cancer gene in the genome. The cancer genes are used as probes to determine the chromosomal localization of the cancer genes. Information such as chromosomal localization finds use in providing a diagnosis or prognosis in particular when chromosomal abnormalities such as translocations, and the like are identified in the cancer gene locus.

XIV.) Kits/Articles of Manufacture

For use in the laboratory, prognostic, prophylactic, diagnostic and therapeutic applications described herein, kits are within the scope of the invention. Such kits can comprise a carrier, package, or container that is compartmentalized to receive one or more containers such as vials, tubes, and the like, each of the container(s) comprising one of the separate elements to be used in the method, along with a label or insert comprising instructions for use, such as a use described herein. For example, the container(s) can comprise a probe that is or can be detectably labeled. Such probe can be an antibody or polynucleotide specific for a protein or a gene or message of the invention, respectively. Where the method utilizes nucleic acid hybridization to detect the target nucleic acid, the kit can also have containers containing nucleotide(s) for amplification of the target nucleic acid sequence. Kits can comprise a container comprising a reporter, such as a biotin-binding protein, such as avidin or streptavidin, bound to a reporter molecule, such as an enzymatic, fluorescent, or radioisotope label; such a reporter can be used with, e.g., a nucleic acid or antibody. The kit can include all or part of the amino acid sequences in FIG. **2** or FIG. **3** or analogs thereof, or a nucleic acid molecule that encodes such amino acid sequences.

The kit of the invention will typically comprise the container described above and one or more other containers associated therewith that comprise materials desirable from a commercial and user standpoint, including buffers, diluents, filters, needles, syringes; carrier, package, container, vial and/

or tube labels listing contents and/or instructions for use, and package inserts with instructions for use.

A label can be present on or with the container to indicate that the composition is used for a specific therapy or non-therapeutic application, such as a prognostic, prophylactic, diagnostic or laboratory application, and can also indicate directions for either in vivo or in vitro use, such as those described herein. Directions and or other information can also be included on an insert(s) or label(s) which is included with or on the kit. The label can be on or associated with the container. A label a can be on a container when letters, numbers or other characters forming the label are molded or etched into the container itself; a label can be associated with a container when it is present within a receptacle or carrier that also holds the container, e.g., as a package insert. The label can indicate that the composition is used for diagnosing, treating, prophylaxing or prognosing a condition, such as a neoplasia of a tissue set forth in Table I.

The terms "kit" and "article of manufacture" can be used as synonyms.

In another embodiment of the invention, an article(s) of manufacture containing compositions, such as amino acid sequence(s), small molecule(s), nucleic acid sequence(s), and/or antibody(s), e.g., materials useful for the diagnosis, prognosis, prophylaxis and/or treatment of neoplasias of tissues such as those set forth in Table I is provided. The article of manufacture typically comprises at least one container and at least one label. Suitable containers include, for example, bottles, vials, syringes, and test tubes. The containers can be formed from a variety of materials such as glass, metal or plastic. The container can hold amino acid sequence(s), small molecule(s), nucleic acid sequence(s), cell population(s) and/or antibody(s). In one embodiment, the container holds a polynucleotide for use in examining the mRNA expression profile of a cell, together with reagents used for this purpose. In another embodiment a container comprises an antibody, binding fragment thereof or specific binding protein for use in evaluating protein expression of 58P1D12 in cells and tissues, or for relevant laboratory, prognostic, diagnostic, prophylactic and therapeutic purposes; indications and/or directions for such uses can be included on or with such container, as can reagents and other compositions or tools used for these purposes. In another embodiment, a container comprises materials for eliciting a cellular or humoral immune response, together with associated indications and/or directions. In another embodiment, a container comprises materials for adoptive immunotherapy, such as cytotoxic T cells (CTL) or helper T cells (HTL), together with associated indications and/or directions; reagents and other compositions or tools used for such purpose can also be included.

The container can alternatively hold a composition that is effective for treating, diagnosis, prognosing or prophylaxing a condition and can have a sterile access port (for example the container can be an intravenous solution bag or a vial having a stopper pierceable by a hypodermic injection needle). The active agents in the composition can be an antibody capable of specifically binding 58P1D12 and modulating the function of 58P1D12.

The article of manufacture can further comprise a second container comprising a pharmaceutically-acceptable buffer, such as phosphate-buffered saline, Ringers solution and/or dextrose solution. It can further include other materials desirable from a commercial and user standpoint, including other buffers, diluents, filters, stirrers, needles, syringes, and/or package inserts with indications and/or instructions for use.

EXAMPLES

Various aspects of the invention are further described and illustrated by way of the several examples that follow, none of which is intended to limit the scope of the invention.

Example 1

SSH-Generated Isolation of a cDNA Fragment of the 58P1D12 Gene

To isolate genes that are over-expressed in prostate cancer we used the Suppression Subtractive Hybridization (SSH) procedure using cDNA derived from prostate cancer tissues. The 58P1D12 SSH cDNA sequence was identified from an SSH experiment consisting of LAPC-9AD cDNA minus LAPC-9AI cDNA.

Materials and Methods:

Human Tissues:

The patient cancer and normal tissues were purchased from different sources such as the NDRI (Philadelphia, Pa.). mRNA for some normal tissues were purchased from different companies such as Clontech, Palo Alto, Calif.

RNA Isolation:

Tissues were homogenized in Trizol reagent (Life Technologies, Gibco BRL) using 10 ml/g tissue to isolate total RNA. Poly A RNA was purified from total RNA using Qiagen's Oligotex mRNA Mini and Midi kits. Total and mRNA were quantified by spectrophotometric analysis (O.D. 260/280 nm) and analyzed by gel electrophoresis.

Oligonucleotides:

The following HPLC purified oligonucleotides were used:

DPNCDN (cDNA synthesis primer):
(SEQ ID NO: 22)
5'TTTTGATCAAGCTT$_{30}$3'

Adaptor 1:
(SEQ ID NO: 23)
5'CTAATACGACTCACTATAGGGCTCGAGCGGCCGCCCGGGCAG3'

(SEQ ID NO: 24)
3'GGCCCGTCCTAG5'

Adaptor 2:
(SEQ ID NO: 25)
5'GTAATACGACTCACTATAGGGCAGCGTGGTCGCGGCCGAG3'

(SEQ ID NO: 26)
3'CGGCTCCTAG5'

PCR primer 1:
(SEQ ID NO: 27)
5'CTAATACGACTCACTATAGGGC3'

Nested primer (NP)1:
(SEQ ID NO: 28)
5'TCGAGCGGCCGCCCGGGCAGGA3'

Nested primer (NP)2:
(SEQ ID NO: 29)
5'AGCGTGGTCGCGGCCGAGGA3'

Suppression Subtractive Hybridization:

Suppression Subtractive Hybridization (SSH) was used to identify cDNAs corresponding to genes that may be differentially expressed in prostate cancer. The SSH reaction utilized cDNA from LAPC-9AD and LAPC-9AI prostate cancer xenograft tissues.

The gene 58P1D12 sequence was derived from the androgen dependent xenograft LAPC-9AD minus the androgen independent xenograft LAPC-9AI cDNA subtraction. The SSH DNA sequence 58P1D12 (FIG. **1**) was identified.

The cDNA derived from LAPC-9AI was used as the source of the "driver" cDNA, while the cDNA from LAPC-9AD was used as the source of the "tester" cDNA. Double stranded cDNAs corresponding to tester and driver cDNAs were synthesized from 2 μg of poly(A)$^+$ RNA isolated from the relevant xenograft tissue, as described above, using CLONTECH's PCR-Select cDNA Subtraction Kit and 1 ng of oligonucleotide DPNCDN as primer. First- and second-strand synthesis were carried out as described in the Kit's user manual protocol (CLONTECH Protocol No. PT1117-1, Catalog No. K1804-1). The resulting cDNA was digested with Dpn II for 3 hrs at 37° C. Digested cDNA was extracted with phenol/chloroform (1:1) and ethanol precipitated.

Tester cDNA was generated by diluting 1 μl of Dpn II digested cDNA from the relevant tissue source (see above) (400 ng) in 5 μl of water. The diluted cDNA (2 μl, 160 ng) was then ligated to 2 μl of Adaptor 1 and Adaptor 2 (10 μM), in separate ligation reactions, in a total volume of 10 μl at 16° C. overnight, using 400 μl of T4 DNA ligase (CLONTECH). Ligation was terminated with 1 μl of 0.2 M EDTA and heating at 72° C. for 5 min.

The first hybridization was performed by adding 1.5 μl (600 ng) of driver cDNA to each of two tubes containing 1.5 μl (20 ng) Adaptor 1- and Adaptor 2-ligated tester cDNA. In a final volume of 4 μl, the samples were overlaid with mineral oil, denatured in an MJ Research thermal cycler at 98° C. for 1.5 minutes, and then were allowed to hybridize for 8 hrs at 68° C. The two hybridizations were then mixed together with an additional 1 μl of fresh denatured driver cDNA and were allowed to hybridize overnight at 68° C. The second hybridization was then diluted in 200 μl of 20 mM Hepes, pH 8.3, 50 mM NaCl, 0.2 mM EDTA, heated at 70° C. for 7 min. and stored at −20° C.

PCR Amplification, Cloning and Sequencing of Gene Fragments Generated from SSH: To amplify gene fragments resulting from SSH reactions, two PCR amplifications were performed. In the primary PCR reaction 1 μl of the diluted final hybridization mix was added to 1 μl of PCR primer 1 (10 μM), 0.5 μl dNTP mix (10 μM), 2.5 μl 10× reaction buffer (CLONTECH) and 0.5 μl 50× Advantage cDNA polymerase Mix (CLONTECH) in a final volume of 25 μl. PCR 1 was conducted using the following conditions: 75° C. for 5 min., 94° C. for 25 sec., then 27 cycles of 94° C. for 10 sec, 66° C. for 30 sec, 72° C. for 1.5 min. Five separate primary PCR reactions were performed for each experiment. The products were pooled and diluted 1:10 with water. For the secondary PCR reaction, 1 μl from the pooled and diluted primary PCR reaction was added to the same reaction mix as used for PCR 1, except that primers NP1 and NP2 (10 μM) were used instead of PCR primer 1. PCR 2 was performed using 10-12 cycles of 94° C. for 10 sec, 68° C. for 30 sec, and 72° C. for 1.5 minutes. The PCR products were analyzed using 2% agarose gel electrophoresis.

The PCR products were inserted into pCR2.1 using the T/A vector cloning kit (Invitrogen). Transformed *E. coli* were subjected to blue/white and ampicillin selection. White colonies were picked and arrayed into 96 well plates and were grown in liquid culture overnight. To identify inserts, PCR amplification was performed on 1 μl of bacterial culture using the conditions of PCR1 and NP1 and NP2 as primers. PCR products were analyzed using 2% agarose gel electrophoresis.

Bacterial clones were stored in 20% glycerol in a 96 well format. Plasmid DNA was prepared, sequenced, and subjected to nucleic acid homology searches of the GenBank, dBest, and NCI-CGAP databases.

RT-PCR Expression Analysis: First strand cDNAs can be generated from 1 μg of mRNA with oligo (dT)12-18 priming using the Gibco-BRL Superscript Preamplification system. The manufacturer's protocol was used which included an incubation for 50 min at 42° C. with reverse transcriptase followed by RNAse H treatment at 37° C. for 20 min. After completing the reaction, the volume can be increased to 200 μl with water prior to normalization. First strand cDNAs from 16 different normal human tissues can be obtained from Clontech.

Normalization of the first strand cDNAs from multiple tissues was performed by using the primers 5'atatcgccgcgctcgtcgtcgacaa3' (SEQ ID NO: 30) and 5'agccacacgcagctcattgtagaagg 3' (SEQ ID NO: 31) to amplify β-actin. First strand cDNA (5 μl) were amplified in a total volume of 50 μl containing 0.4 μM primers, 0.2 μM each dNTPs, 1×PCR buffer (Clontech, 10 mM Tris-HCL, 1.5 mM $MgCl_2$, 50 mM KCl, pH8.3) and 1×Klentaq DNA polymerase (Clontech). Five μl of the PCR reaction can be removed at 18, 20, and 22 cycles and used for agarose gel electrophoresis. PCR was performed using an MJ Research thermal cycler under the following conditions: Initial denaturation can be at 94° C. for 15 sec, followed by a 18, 20, and 22 cycles of 94° C. for 15, 65° C. for 2 min, 72° C. for 5 sec. A final extension at 72° C. was carried out for 2 min. After agarose gel electrophoresis, the band intensities of the 283 b.p. β-actin bands from multiple tissues were compared by visual inspection. Dilution factors for the first strand cDNAs were calculated to result in equal β-actin band intensities in all tissues after 22 cycles of PCR. Three rounds of normalization can be required to achieve equal ban d intensifies in all tissues after 22 cycles of PCR.

To determine expression levels of the 58P1D12 gene, 5 μl of normalized first strand cDNA were analyzed by PCR using 26, and 30 cycles of amplification. Semi-quantitative expression analysis can be achieved by comparing the PCR products at cycle numbers that give light band intensities. The primers used for RT-PCR were designed using the 58P1D12 SSH sequence and are listed below:

```
58P1D12.1
5'- CCTGCTTCAGTAACAACCACATTCT - 3' (SEQ ID NO: 32)

58P1D12.2
5'- CTTTACCAGTGGAATGATGACAGG - 3'  (SEQ ID NO: 33)
```

A typical RT-PCR expression analysis is shown in FIG. **14**.

Example 2

Full Length Cloning of 58P1D12

The 58P1D12 SSH sequence of 427 bp (FIG. **1**) was identified from LAPC xenograft SSH experiment. A full length cDNA clone for 58P1D12 was isolated from an LAPC-9 AD library. The cDNA (clone 2) is 2550 bp in length and encodes a 273 amino acid ORF (FIG. **2** and FIG. **3**). 58P1D12 v.1

exhibited 100% homology to human chondrolectin. Other variants of 58P1D12 were also identified and these are listed in FIG. **2** and FIG. **3**.

Example 3

Chromosomal Mapping of 58P1D12

Chromosomal localization can implicate genes in disease pathogenesis. Several chromosome mapping approaches are available including fluorescent in situ hybridization (FISH), human/hamster radiation hybrid (RH) panels (Walter et al., 1994; Nature Genetics 7:22; Research Genetics, Huntsville Ala.), human-rodent somatic cell hybrid panels such as is available from the Coriell Institute (Camden, N.J.), and genomic viewers utilizing BLAST homologies to sequenced and mapped genomic clones (NCBI, Bethesda, Md.).

58P1D12 maps to chromosome 21q11.2 using 58P1D12 sequence and the NCBI BLAST tool located on the World Wide Web at: (.ncbi.nlm.nih.gov/genome/seq/page.cgi?F=HsBlast.html&&ORG=Hs).

Example 4

Expression Analysis of 58P1D12 in Normal Tissues and Patient Specimens

58P1D12 expression in normal and patient cancer tissues was tested by PCR (FIG. **14**). First strand cDNA was prepared from normal tissues (bladder, brain, heart, kidney, liver, lung, prostate, spleen, skeletal muscle, tests, pancreas, colon and stomach), and from pools of patient cancer specimens (prostate cancer pool, bladder cancer pool, lung cancer pool, ovary cancer pool, prostate cancer xenograft pool, bone and melanoma cancer pool, and cervical cancer pool). Normalization was performed by PCR using primers to actin. Semi-quantitative PCR, using primers to 58P1D12, was performed at 30 cycles of amplification. Results expression of 58P1D12 predominantly in testis amongst the normal tissues tested. 58P1D12 is also expressed in prostate cancer, bladder cancer, lung cancer, ovary cancer, prostate cancer xenograft, bone/melanoma cancer pool, and cervical cancer.

FIG. **15** shows expression of 58P1D12 in normal tissues by Northern blot analysis. Two multiple tissue northern blots (Clontech) both with 2 μg of mRNA/lane were probed with the 58P1D12 sequence. Size standards in kilobases (kb) are indicated on the side. Results show predominant expression of 58P1D12 transcript in normal testis. A significantly weaker signal is detected in spleen, lung, kidney and pancreas.

In order to show expression of 58P1D12 in prostate cancer tissues, RNA was extracted from LAPC prostate cancer xenografts, LAPC4AD, LAPC-4AI, LAPC-9AD and LAPC-9AI, as well as from normal prostate. Northern blots with 10 μg of total RNA were probed with the 58P1D12 sequence. Size standards in kilobases are on the side. Results show strong expression of 58P1D12 in LAPC-9AD, and a weaker signal in LAPC-9AI, but not in the other tissues tested. (FIG. **16**).

FIG. **17** shows 58P1D12 expression in normal and patient cancer specimens. RNA was extracted from a pool of 3 ovary tumor patient specimens, normal prostate (NP), normal bladder (NB), normal kidney (NK), and normal colon (NC). Northern blots with 10 μg of total RNA were probed with the 58P1D12 sequence. Size standards in kilobases are on the side. Results show expression of 58P1D12 in ovary tumor patient specimens but not in the normal tissues tested.

Analysis of 58P1D12 expression in individual ovarian cancer patient specimens is shown in FIG. **18**. RNA was extracted from normal ovary, LAPC-9AD prostate cancer xenograft, and a panel of ovary tumor patient specimens (Ovary T). Northern blots with 10 μg of total RNA were probed with the 58P1D12 sequence. Size standards in kilobases are on the side. Results show expression of 58P1D12 in most of the patient specimens tested as well as in the prostate xenograft.

58P1D12 expression was tested in a large panel of ovarian cancer patent specimens (FIG. **19**). First strand cDNA was prepared from a panel of ovary patient cancer specimens as well as two different normal ovary tissues. Normalization was performed by PCR using primers to actin. Semi-quantitative PCR, using primers to 58P1D12, was performed at 26 and 30 cycles of amplification. Samples were run on an agarose gel, and PCR products were quantitated using the AlphaImager software. Expression was recorded as negative, weak, medium or strong. Results show expression of 58P1D12 in the majority of patient cancer specimens tested (75%), but not in normal ovary.

Example 5

Transcript Variants of 58P1D12

Transcript variants are variants of mature mRNA from the same gene which arise by alternative transcription or alternative splicing. Alternative transcripts are transcripts from the same gene but start transcription at different points. Splice variants are mRNA variants spliced differently from the same transcript. In eukaryotes, when a multi-exon gene is transcribed from genomic DNA, the initial RNA is spliced to produce functional mRNA, which has only exons and is used for translation into an amino acid sequence. Accordingly, a given gene can have zero to many alternative transcripts and each transcript can have zero to many splice variants. Each transcript variant has a unique exon makeup, and can have different coding and/or non-coding (5' or 3' end) portions, from the original transcript. Transcript variants can code for similar or different proteins with the same or a similar function or can encode proteins with different functions, and can be expressed in the same tissue at the same time, or in different tissues at the same time, or in the same tissue at different times, or in different tissues at different times. Proteins encoded by transcript variants can have similar or different cellular or extracellular localizations, e.g., secreted versus intracellular.

Transcript variants are identified by a variety of art-accepted methods. For example, alternative transcripts and splice variants are identified by full-length cloning experiment, or by use of full-length transcript and EST sequences. First, all human ESTs were grouped into clusters which show direct or indirect identity with each other. Second, ESTs in the same cluster were further grouped into sub-clusters and assembled into a consensus sequence. The original gene sequence is compared to the consensus sequence(s) or other full-length sequences. Each consensus sequence is a potential splice variant for that gene (see, e.g., URL www.doubletwist.com/products/c11_agentsOverview.jhtml). Even when a variant is identified that is not a full-length clone, that portion of the variant is very useful for antigen generation and for further cloning of the full-length splice variant, using techniques known in the art.

Moreover, computer programs are available in the art that identify transcript variants based on genomic sequences. Genomic-based transcript variant identification programs include FgenesH (A. Salamov and V. Solovyev, "Ab initio gene finding in *Drosophila* genomic DNA," Genome Research. 2000 April; 10(4):516-22); Grail (URL compbio.ornl.gov/Grail-bin/EmptyGrailForm) and GenScan (URL genes.mit.edu/GENSCAN.html). For a general discussion of splice variant identification protocols see., e.g., Southan, C., A genomic perspective on human proteases, FEBS Lett. 2001 Jun. 8; 498(2-3):214-8; de Souza, S. J., et al., Identification of human chromosome 22 transcribed sequences with ORF expressed sequence tags, Proc. Natl. Acad Sci USA. 2000 Nov. 7; 97(23):12690-3.

To further confirm the parameters of a transcript variant, a variety of techniques are available in the art, such as full-length cloning, proteomic validation, PCR-based validation, and 5' RACE validation, etc. (see e.g., Proteomic Validation: Brennan, S. O., et al., Albumin banks peninsula: a new termination variant characterized by electrospray mass spectrometry, Biochem Biophys Acta. 1999 Aug. 17; 1433(1-2): 321-6; Ferranti P, et al., Differential splicing of pre-messenger RNA produces multiple forms of mature caprine alpha(s1)-casein, Eur J. Biochem. 1997 Oct. 1; 249(1):1-7. For PCR-based Validation: Wellmann S, et al., Specific reverse transcription-PCR quantification of vascular endothelial growth factor (VEGF) splice variants by LightCycler technology, Clin Chem. 2001 April; 47(4):654-60; Jia, H. P., et al, Discovery of new human beta-defensins using a genomics-based approach, Gene. 2001 Jan. 24; 263(1-2):211-8. For PCR-based and 5' RACE Validation: Brigle, K. E., et al., Organization of the murine reduced folate carrier gene and identification of variant splice forms, Biochem Biophys Acta. 1997 Aug. 7; 1353(2): 191-8).

It is known in the art that genomic regions are modulated in cancers. When the genomic region to which a gene maps is modulated in a particular cancer, the alternative transcripts or splice variants of the gene are modulated as well. Disclosed herein is that 58P1D12 has a particular expression profile related to cancer. Alternative transcripts and splice variants of 58P1D12 may also be involved in cancers in the same or different tissues, thus serving as tumor-associated markers/antigens.

Using the full-length gene and EST sequences, four transcript variants were identified, designated as 58P1D12 v.2, v.3, v.4 and v.5. The boundaries of the exon in the original transcript, 58P1D12 v.1 were shown in Table LI. Exon compositions of the variants were shown in FIG. **10**. Theoretically, each different combination of exons in spatial order, e.g. exon 1 of v.2 and exons 3, 4, 5 and 6 of v.1, is a potential splice variant.

Tables LII(a)-(d) through LV(a)-(d) are set forth on a variant-by-variant bases. LII(a)-(d) shows nucleotide sequence of the transcript variant. Table LIII(a)-(d) shows the alignment of the transcript variant with nucleic acid sequence of 58P1D12 v.1. LIV(a)-(d) lays out amino acid translation of the transcript variant for the identified reading frame orientation. Table LV(a)-(d) displays alignments of the amino acid sequence encoded by the splice variant with that of 58P1D12 v.1.

Example 6

Single Nucleotide Polymorphisms of 58P1D12

A Single Nucleotide Polymorphism (SNP) is a single base pair variation in a nucleotide sequence at a specific location. At any given point of the genome, there are four possible nucleotide base pairs: A/T, C/G, G/C and T/A. Genotype refers to the specific base pair sequence of one or more locations in the genome of an individual. Haplotype refers to the base pair sequence of more than one location on the same DNA molecule (or the same chromosome in higher organisms), often in the context of one gene or in the context of several tightly linked genes. SNP that occurs on a cDNA is called cSNP. This cSNP may change amino acids of the protein encoded by the gene and thus change the functions of the protein. Some SNP cause inherited diseases; others contribute to quantitative variations in phenotype and reactions to environmental factors including diet and drugs among individuals. Therefore, SNP and/or combinations of alleles (called haplotypes) have many applications, including diagnosis of inherited diseases, determination of drug reactions and dosage, identification of genes responsible for diseases, and analysis of the genetic relationship between individuals (P. Nowotny, J. M. Kwon and A. M. Goate, "SNP analysis to dissect human traits," Curr. Opin. Neurobiol. 2001 October; 11(5):637-641; M. Pirmohamed and B. K. Park, "Genetic susceptibility to adverse drug reactions," Trends Pharmacol. Sci. 2001 June; 22(6):298-305; J. H. Riley, C. J. Allan, E. Lai and A. Roses, "The use of single nucleotide polymorphisms in the isolation of common disease genes," Pharmacogenomics. 2000 February; 1(1):3947; R. Judson, J. C. Stephens and A. Windemuth, "The predictive power of haplotypes in clinical response," Pharmacogenomics. 2000 February; 1(1):15-26).

SNP are identified by a variety of art-accepted methods (P. Bean, "The promising voyage of SNP target discovery," Am. Clin. Lab. 2001 October-November; 20(9):18-20; K. M. Weiss, "In search of human variation," Genome Res. 1998 July; 8(7):691-697; M. M. She, "Enabling large-scale pharmacogenetic studies by high-throughput mutation detection and genotyping technologies," Clin. Chem. 2001 February; 47(2):164-172). For example, SNP can be identified by sequencing DNA fragments that show polymorphism by gel-based methods such as restriction fragment length polymorphism (RFLP) and denaturing gradient gel electrophoresis (DGGE). They can also be discovered by direct sequencing of DNA samples pooled from different individuals or by comparing sequences from different DNA samples. With the rapid accumulation of sequence data in public and private databases, one can discover SNP by comparing sequences using computer programs (Z. Gu, L. Hillier and P. Y. Kwok, "Single nucleotide polymorphism hunting in cyberspace," Hum. Mutat. 1998; 12(4):221-225). SNP can be verified and genotype or haplotype of an individual can be determined by a variety of methods including direct sequencing and high throughput microarrays (P. Y. Kwok, "Methods for genotyping single nucleotide polymorphisms," Annu. Rev. Genomics Hum. Genet. 2001; 2:235-258; M. Kokoris, K. Dix, K. Moynihan, J. Mathis, B. Erwin, P. Grass, B. Hines and A. Duesterhoeft, "High-throughput SNP genotyping with the Masscode system," Mol. Diagn. 2000 December; 5(4):329-340).

Using the methods described above, ten SNP were identified in the original transcript, 58P1D12 v.1, at positions 1764 (A/C), 1987 (G/A), 2045 (A/G), 2066 (T/C), 2134 (T/A), 2350 (G/T), 2435 (G/T), 302 (G/T), 304 (G/T) and 1533 (C/T). The transcripts or proteins with alternative allele were designated as variant 58P1D12 v.6 through v.15, respectively. FIG. **12** shows the schematic alignment of the SNP variants. FIG. **11** shows the schematic alignment of protein variants, corresponding to nucleotide variants. Nucleotide variants that code for the same amino acid sequence as v.1 are not shown in FIG. **11**. These alleles of the SNP, though shown separately here, can occur in different combinations (haplotypes) and in any one of the transcript variants (such as 58P1D12 v.2) that contains the site of the SNP.

Example 7

Production of Recombinant 58P1D12 in Prokaryotic Systems

To express recombinant 58P1D12 and 58P1D12 variants in prokaryotic cells, the full or partial length 58P1D12 and 58P1D12 variant cDNA sequences are cloned into any one of a variety of expression vectors known in the art. One or more of the following regions of 58P1D12 variants are expressed: the full length sequence presented in FIGS. **2** and **3**, or any 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30 or more contiguous amino acids from 58P1D12, variants, or analogs thereof.

A. In Vitro Transcription and Translation Constructs:

pCRII: To generate 58P1D12 sense and anti-sense RNA probes for RNA in situ investigations, pCRII constructs (Invitrogen, Carlsbad Calif.) are generated encoding either all or fragments of the 58P1D12 cDNA. The pCRII vector has Sp6 and T7 promoters flanking the insert to drive the transcription of 58P1D12 RNA for use as probes in RNA in situ hybridization experiments. These probes are used to analyze the cell and tissue expression of 58P1D12 at the RNA level. Transcribed 58P1D12 RNA representing the cDNA amino acid coding region of the 58P1D12 gene is used in in vitro translation systems such as the TnT™ Coupled Reticulolysate System (Promega, Corp., Madison, Wis.) to synthesize 58P1D12 protein.

B. Bacterial Constructs:

pGEX Constructs: To generate recombinant 58P1D12 proteins in bacteria that are fused to the Glutathione S-transferase (GST) protein, all or parts of the 58P1D12 cDNA protein coding sequence are cloned into the pGEX family of GST-fusion vectors (Amersham Pharmacia Biotech, Piscataway, N.J.). These constructs allow controlled expression of recombinant 58P1D12 protein sequences with GST fused at the amino-terminus and a six histidine epitope (6×His) at the carboxyl-terminus. The GST and 6×His tags permit purification of the recombinant fusion protein from induced bacteria with the appropriate affinity matrix and allow recognition of the fusion protein with anti-GST and anti-His antibodies. The 6×His tag is generated by adding 6 histidine codons to the cloning primer at the 3' end, e.g., of the open reading frame (ORF). A proteolytic cleavage site, such as the PreScission™ recognition site in pGEX-6P-1, may be employed such that it permits cleavage of the GST tag from 58P1D12-related protein. The ampicillin resistance gene and pBR322 origin permits selection and maintenance of the pGEX plasmids in *E. coli.*

pMAL Constructs: To generate, in bacteria, recombinant 58P1D12 proteins that are fused to maltose-binding protein (MBP), all or parts of the 58P1D12 cDNA protein coding sequence are fused to the MBP gene by cloning into the pMAL-c2X and pMAL-p2X vectors (New England Biolabs, Beverly, Mass.). These constructs allow controlled expression of recombinant 58P1D12 protein sequences with MBP fused at the amino-terminus and a 6×His epitope tag at the carboxyl-terminus. The MBP and 6×His tags permit purification of the recombinant protein from induced bacteria with the appropriate affinity matrix and allow recognition of the fusion protein with anti-MBP and anti-His antibodies. The 6×His epitope tag is generated by adding 6 histidine codons to the 3' cloning primer. A Factor Xa recognition site permits cleavage of the pMAL tag from 58P1D12. The pMAL-c2X and pMAL-p2X vectors are optimized to express the recombinant protein in the cytoplasm or periplasm respectively. Periplasm expression enhances folding of proteins with disulfide bonds.

pET Constructs: To express 58P1D12 in bacterial cells, all or parts of the 58P1D12 cDNA protein coding sequence are cloned into the pET family of vectors (Novagen, Madison, Wis.). These vectors allow tightly controlled expression of recombinant 58P1D12 protein in bacteria with and without fusion to proteins that enhance solubility, such as NusA and thioredoxin (Trx), and epitope tags, such as 6×His and S-Tag™ that aid purification and detection of the recombinant protein. For example, constructs are made utilizing pET NusA fusion system 43.1 such that regions of the 58P1D12 protein are expressed as amino-terminal fusions to NusA. The cDNA encoding amino acids 22-213 of 58P1D12 was cloned into the pET-21b vector, expressed, and purified from bacteria. The recombinant protein was used to generate rabbit polyclonal antibodies.

C. Yeast Constructs:

pESC Constructs: To express 58P1D12 in the yeast species *Saccharomyces cerevisiae* for generation of recombinant protein and functional studies, all or parts of the 58P1D12 cDNA protein coding sequence are cloned into the pESC family of vectors each of which contain 1 of 4 selectable markers, HIS3, TRP1, LEU2, and URA3 (Stratagene, La Jolla, Calif.). These vectors allow controlled expression from the same plasmid of up to 2 different genes or cloned sequences containing either Flag™ or Myc epitope tags in the same yeast cell. This system is useful to confirm protein-protein interactions of 58P1D12. In addition, expression in yeast yields similar post-translational modifications, such as glycosylations and phosphorylations that are found when expressed in eukaryotic cells.

pESP Constructs: To express 58P1D12 in the yeast species *Saccharomyces pombe*, all or parts of the 58P1D12 cDNA protein coding sequence are cloned into the pESP family of vectors. These vectors allow controlled high level of expression of a 58P1D12 protein sequence that is fused at either the amino terminus or at the carboxyl terminus to GST which aids purification of the recombinant protein. A Flag™ epitope tag allows detection of the recombinant protein with anti-Flag™ antibody.

Example 8

Production of Recombinant 58P1D12 in Higher Eukaryotic Systems

A. Mammalian Constructs:

To express recombinant 58P1D12 in eukaryotic cells, the full or partial length 58P1D12 cDNA sequences were cloned into any one of a variety of expression vectors known in the art. One or more of the following regions of 58P1D12 were expressed in these constructs, amino acids 1 to 273 or any 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30 or more contiguous amino acids from 58P1D12 v.1; amino acids 1 to 232 of v.2; amino acids 1 to 236 of v.4; or any 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30 or more contiguous amino acids from 58P1D12 variants, or analogs thereof.

The constructs can be transfected into any one of a wide variety of mammalian cells such as 293T cells. Transfected

293T cell lysates can be probed with the anti-58P1D12 polyclonal serum, described herein.

pcDNA4/HisMax Constructs: To express 58P1D12 in mammalian cells, a 58P1D12 ORF, or portions thereof, of 58P1D12 are cloned into pcDNA4/HisMax Version A (Invitrogen, Carlsbad, Calif.). Protein expression is driven from the cytomegalovirus (CMV) promoter and the SP16 translational enhancer. The recombinant protein has Xpress™ and six histidine (6×His) epitopes fused to the amino-terminus. The pcDNA4/HisMax vector also contains the bovine growth hormone (BGH) polyadenylation signal and transcription termination sequence to enhance mRNA stability along with the SV40 origin for episomal replication and simple vector rescue in cell lines expressing the large T antigen. The Zeocin resistance gene allows for selection of mammalian cells expressing the protein and the ampicillin resistance gene and ColE1 origin permits selection and maintenance of the plasmid in *E. coli.*

pcDNA3.1/MycHis Constructs: To express 58P1D12 in mammalian cells, a 58P1D12 ORF, or portions thereof, of 58P1D12 with a consensus Kozak translation initiation site was cloned into pcDNA3.1/MycHis Version A (Invitrogen, Carlsbad, Calif.). Protein expression is driven from the cytomegalovirus (CMV) promoter. The recombinant proteins have the myc epitope and 6×His epitope fused to the carboxyl-terminus. The pcDNA3.1/MycHis vector also contains the bovine growth hormone (BGH) polyadenylation signal and transcription termination sequence to enhance mRNA stability, along with the SV40 origin for episomal replication and simple vector rescue in cell lines expressing the large T antigen. The Neomycin resistance gene can be used, as it allows for selection of mammalian cells expressing the protein and the ampicillin resistance gene and ColE1 origin permits selection and maintenance of the plasmid in *E. coli.*

The complete ORF of 58P1D12 v.1 was cloned into the pcDNA3.1/MycHis construct to generate 58P1D12.pcDNA3.1/MycHis.

pcDNA3.1/CT-GFP-TOPO Construct: To express 58P1D12 in mammalian cells and to allow detection of the recombinant proteins using fluorescence, a 58P1D12 ORF, or portions thereof, with a consensus Kozak translation initiation site are cloned into pcDNA3.1/CT-GFP-TOPO (Invitrogen, Calif.). Protein expression is driven from the cytomegalovirus (CMV) promoter. The recombinant proteins have the Green Fluorescent Protein (GFP) fused to the carboxyl-terminus facilitating non-invasive, in vivo detection and cell biology studies. The pcDNA3.1CT-GFP-TOPO vector also contains the bovine growth hormone (BGH) polyadenylation signal and transcription termination sequence to enhance mRNA stability along with the SV40 origin for episomal replication and simple vector rescue in cell lines expressing the large T antigen. The Neomycin resistance gene allows for selection of mammalian cells that express the protein and the ampicillin resistance gene and ColE1 origin permits selection and maintenance of the plasmid in *E. coli*. Additional constructs with an amino-terminal GFP fusion are made in pcDNA3.1/NT-GFP-TOPO spanning the entire length of a 58P1D12 protein.

PAPtag: A 58P1D12 ORF, or portions thereof, were cloned into pAPtag-5 (GenHunter Corp. Nashville, Tenn.). This construct generates an alkaline phosphatase fusion at the carboxyl-terminus of a 58P1D12 protein while fusing the IgGκ signal sequence to the amino-terminus. Constructs are also generated in which alkaline phosphatase with an amino-terminal IgGκ signal sequence is fused to the amino-terminus of a 58P1D12 protein. The resulting recombinant 58P1D12 proteins are optimized for secretion into the media of transfected mammalian cells and can be used to identify proteins such as ligands or receptors that interact with 58P1D12 proteins. Protein expression is driven from the CMV promoter and the recombinant proteins also contain myc and 6×His epitopes fused at the carboxyl-terminus that facilitates detection and purification. The Zeocin resistance gene present in the vector allows for selection of mammalian cells expressing the recombinant protein and the ampicillin resistance gene permits selection of the plasmid in *E. coli.*

pTag5: A 58P1D12 ORF, or portions thereof, were cloned into pTag-5. This vector is similar to pAPtag but without the alkaline phosphatase fusion. This construct generated 58P1D12 protein with an amino-terminal IgGκ signal sequence and myc and 6×His epitope tags at the carboxyl-terminus that facilitate detection and affinity purification. The resulting recombinant 58P1D12 protein was optimized for secretion into the media of transfected mammalian cells, and was used as immunogen or ligand to identify proteins such as ligands or receptors that interact with the 58P1D12 proteins. Protein expression is driven from the CMV promoter. The Zeocin resistance gene present in the vector allows for selection of mammalian cells expressing the protein, and the ampicillin resistance gene permits selection of the plasmid in *E. coli.*

PsecFc: A 58P1D12 ORF, or portions thereof, were cloned into psecFc. The psecFc vector was assembled by cloning the human immunoglobulin G1 (IgG) Fc (hinge, CH2, CH3 regions) into pSec Tag2 (Invitrogen, California). This construct generated an IgG1 Fc fusion at the carboxyl-terminus of the 58P1D12 proteins, while fusing the IgGK signal sequence to N-terminus. 58P1D12 fusions utilizing the murine IgG1 Fc region are also used. The resulting recombinant 58P1D12 proteins are optimized for secretion into the media of transfected mammalian cells, and can be used as immunogens or to identify proteins such as ligands or receptors that interact with 58P1D12 protein. Protein expression is driven from the CMV promoter. The hygromycin resistance gene present in the vector allows for selection of mammalian cells that express the recombinant protein, and the ampicillin resistance gene permits selection of the plasmid in *E. coli.*

pSRα Constructs: To generate mammalian cell lines that express 58P1D12 constitutively, 58P1D12 ORF, or portions thereof, were cloned into pSRα constructs. Amphotropic and ecotropic retroviruses were generated by transfection of pSRα constructs into the 293T-10A1 packaging line or co-transfection of pSRα and a helper plasmid (containing deleted packaging sequences) into the 293 cells, respectively. The retrovirus is used to infect a variety of mammalian cell lines, resulting in the integration of the cloned gene, 58P1D12, into the host cell-lines. Protein expression is driven from a long terminal repeat (LTR). The Neomycin resistance gene present in the vector allows for selection of mammalian cells that express the protein, and the ampicillin resistance gene and ColE1 origin permit selection and maintenance of the plasmid in *E. coli*. The retroviral vectors can thereafter be used for infection and generation of various cell lines using, for example, PC3, NIH 3T3, TsuPr1, 293 or rat-1 cells.

Additional pSRα constructs are made that fuse an epitope tag such as the FLAG™ tag to the carboxyl-terminus of 58P1D12 sequences to allow detection using anti-Flag antibodies. For example, the FLAG™ sequence 5' gat tac aag gat gac gac gat aag 3' (SEQ ID NO: 34) is added to cloning primer at the 3' end of the ORF. Additional pSRα constructs are made to produce both amino-terminal and carboxyl-terminal GFP and myc/6×His fusion proteins of the full-length 58P1D12 proteins.

Additional Viral Vectors: Additional constructs are made for viral-mediated delivery and expression of 58P1D12. High virus titer leading to high level expression of 58P1D12 is achieved in viral delivery systems such as adenoviral vectors and herpes amplicon vectors. A 58P1D12 coding sequence or fragments thereof are amplified by PCR and subcloned into the AdEasy shuttle vector (Stratagene). Recombination and virus packaging are performed according to the manufacturers instructions to generate adenoviral vectors. Alternatively, 58P1D12 coding sequences or fragments thereof are cloned into the HSV-1 vector (Imgenex) to generate herpes viral vectors. The viral vectors are thereafter used for infection of various cell lines such as PC3, NIH 3T3, 293 or rat-1 cells.

Regulated Expression Systems: To control expression of 58P1D12 in mammalian cells, coding sequences of 58P1D12, or portions thereof, are cloned into regulated mammalian expression systems such as the T-Rex System (Invitrogen), the GeneSwitch System (Invitrogen) and the tightly-regulated Ecdysone System (Sratagene). These systems allow the study of the temporal and concentration dependent effects of recombinant 58P1D12. These vectors are thereafter used to control expression of 58P1D12 in various cell lines such as PC3, NIH 3T3, 293 or rat-1 cells.

B. Baculovirus Expression Systems:

To generate recombinant 58P1D12 proteins in a baculovirus expression system, 58P1D12 ORF, or portions thereof, are cloned into the baculovirus transfer vector pBlueBac 4.5 (Invitrogen), which provides a His-tag at the N-terminus. Specifically, pBlueBac-58P1D12 is co-transfected with helper plasmid pBac-N-Blue (Invitrogen) into SF9 (*Spodoptera frugiperda*) insect cells to generate recombinant baculovirus (see Invitrogen instruction manual for details). Baculovirus is then collected from cell supernatant and purified by plaque assay.

Recombinant 58P1D12 protein is then generated by infection of HighFive insect cells (Invitrogen) with purified baculovirus. Recombinant 58P1D12 protein can be detected using anti-58P1D12 or anti-His-tag antibody. 58P1D12 protein can be purified and used in various cell-based assays or as immunogen to generate polyclonal and monoclonal antibodies specific for 58P1D12.

Example 9

Antigenicity Profiles and Secondary Structure

FIGS. **5**A-C, FIGS. **6**A-C, FIGS. **7**A-C, FIGS. **8**A-C, and FIGS. **9**A-C depict graphically five amino acid profiles of 58P1D12 variant 1, variant 2, and variant 3, A through C respectively, each assessment available by accessing the ProtScale website (URL www.expasy.ch/cgi-bin/protscale.pl) on the ExPasy molecular biology server.

These profiles: FIG. **5**, Hydrophilicity, (Hopp T. P., Woods K. R., 1981. Proc. Natl. Acad. Sci. U.S.A. 78:3824-3828); FIG. **6**, Hydropathicity, (Kyte J., Doolittle R. F., 1982. J. Mol. Biol. 157:105-132); FIG. **7**, Percentage Accessible Residues (Janin J., 1979 Nature 277:491-492); FIG. **8**, Average Flexibility, (Bhaskaran R., and Ponnuswamy P. K., 1988. Int. J. Pept. Protein Res. 32:242-255); FIG. **9**, Beta-turn (Deleage, G., Roux B. 1987 Protein Engineering 1:289-294); and optionally others available in the art, such as on the ProtScale website, were used to identify antigenic regions of each of the 58P1D12 variant proteins. Each of the above amino acid profiles of 58P1D12 variants were generated using the following ProtScale parameters for analysis: 1) A window size of 9; 2) 100% weight of the window edges compared to the window center; and, 3) amino acid profile values normalized to lie between 0 and 1.

Hydrophilicity (FIG. **5**), Hydropathicity (FIG. **6**) and Percentage Accessible Residues (FIG. **7**) profiles were used to determine stretches of hydrophilic amino acids (i.e., values greater than 0.5 on the Hydrophilicity and Percentage Accessible Residues profile, and values less than 0.5 on the Hydropathicity profile). Such regions are likely to be exposed to the aqueous environment, be present on the surface of the protein, and thus available for immune recognition, such as by antibodies.

Average Flexibility (FIG. **8**) and Beta-turn (FIG. **9**) profiles determine stretches of amino acids (i.e., values greater than 0.5 on the Beta-turn profile and the Average Flexibility profile) that are not constrained in secondary structures such as beta sheets and alpha helices. Such regions are also more likely to be exposed on the protein and thus accessible to immune recognition, such as by antibodies.

Antigenic sequences of the 58P1D12 variant proteins indicated, e.g., by the profiles set forth in FIG. **5**, FIG. **6**, FIG. **7**, FIG. **8**, and/or FIG. **9** are used to prepare immunogens, either peptides or nucleic acids that encode them, to generate therapeutic and diagnostic anti-58P1D12 antibodies. The immunogen can be any 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 35, 40, 45, 50 or more than 50 contiguous amino acids, or the corresponding nucleic acids that encode them, from the 58P1D12 protein variants listed in FIGS. **2** and **3**. In particular, peptide immunogens of the invention can comprise, a peptide region of at least 5 amino acids of FIGS. **2** and **3** in any whole number increment that includes an amino acid position having a value greater than 0.5 in the Hydrophilicity profiles of FIG. **5**; a peptide region of at least 5 amino acids of FIGS. **2** and **3** in any whole number increment that includes an amino acid position having a value less than 0.5 in the Hydropathicity profile of FIG. **6**; a peptide region of at least 5 amino acids of FIGS. **2** and **3** in any whole number increment that includes an amino acid position having a value greater than 0.5 in the Percent Accessible Residues profiles of FIG. **7**; a peptide region of at least 5 amino acids of FIGS. **2** and **3** in any whole number increment that includes an amino acid position having a value greater than 0.5 in the Average Flexibility profiles on FIG. **8**; and, a peptide region of at least 5 amino acids of FIGS. **2** and **3** in any whole number increment that includes an amino acid position having a value greater than 0.5 in the Beta-turn profile of FIG. **9**. Peptide immunogens of the invention can also comprise nucleic acids that encode any of the forgoing.

All immunogens of the invention, peptide or nucleic acid, can be embodied in human unit dose form, or comprised by a composition that includes a pharmaceutical excipient compatible with human physiology.

The secondary structures of 58P1D12 protein variants 1, 2, and 3, namely the predicted presence and location of alpha helices, extended strands, and random coils, are predicted from their primary amino acid sequences using the HNN—Hierarchical Neural Network method (NPS@: Network Protein Sequence Analysis TIBS 2000 March Vol. 25, No 3 [291]:147-150 Combet C., Blanchet C., Geourjon C. and Deléage G., http://pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_nn.html), accessed from the ExPasy molecular biology server (http://www.expasy.ch/tools/). The analysis indicates that 58P1D12 variant 1 is composed of 24.18% alpha helix, 18.68% extended strand, and 57.14% random coil (FIG. **13**A). 58P1D12 variant 2 is composed of 19.83% alpha helix, 18.97% extended strand, and 61.21% random coil (FIG. **13**B). 58P1D12 variant 3 is composed of 32.20% alpha helix, 15.25% extended strand, and 52.54% random coil (FIG. **13**C).

Analysis for the potential presence of transmembrane domains in the 58P1D12 variant proteins 1, 2 and 3, was carried out using a variety of transmembrane prediction algorithms accessed from the ExPasy molecular biology server (http://www.expasy.ch/tools/). Shown graphically in FIG. **13**D is the result of analysis of variant 1 using the TMpred program and in FIG. **13**E results using the TMHMM program, both of which predict the presence of 1 transmembrane domain. Shown graphically in FIG. **13**F is the result of analysis of variant 2 using the TMpred program and in FIG. **13**G results using the TMHMM program, both of which predict the presence of 1 transmembrane domain. Variant 3 is not predicted to contain transmembrane domains (FIGS. **13**H and **13**I). Analyses of the variants using other structural prediction programs are summarized in Table VI and Table L.

Example 10

Generation of 58P1D12 Polyclonal Antibodies

Polyclonal antibodies can be raised in a mammal, for example, by one or more injections of an immunizing agent and, if desired, an adjuvant. Typically, the immunizing agent and/or adjuvant will be injected in the mammal by multiple subcutaneous or intraperitoneal injections. In addition to immunizing with a full length 58P1D12 protein variant, computer algorithms are employed in design of immunogens that, based on amino acid sequence analysis contain characteristics of being antigenic and available for recognition by the immune system of the immunized host (see the Example entitled "Antigenicity Profiles and Secondary Structure"). Such regions would be predicted to be hydrophilic, flexible, in beta-turn conformations, and be exposed on the surface of the protein (see, e.g., FIG. **5**, FIG. **6**, FIG. **7**, FIG. **8**, or FIG. **9** for amino acid profiles that indicate such regions of 58P1D12 protein variant 1).

For example, recombinant bacterial fusion proteins or peptides containing hydrophilic, flexible, beta-turn regions of 58P1D12 protein variants are used as antigens to generate polyclonal antibodies in New Zealand White rabbits or monoclonal antibodies as described in Example 11. For example, in 58P1D12 variant 1, such regions include, but are not limited to, amino acids 19-30, amino acids 49-66, amino acids 70-82, amino acids 88-115, amino acids 131-145, amino acids 165-203, amino acids 243-255, and amino acids 262-273. It is useful to conjugate the immunizing agent to a protein known to be immunogenic in the mammal being immunized. Examples of such immunogenic proteins include, but are not limited to, keyhole limpet hemocyanin (KLH), serum albumin, bovine thyroglobulin, and soybean trypsin inhibitor. In one embodiment, a peptide encoding amino acids 102-115 of 58P1D12 variant 1 was conjugated to KLH and used to immunize a rabbit. Alternatively the immunizing agent may include all or portions of the 58P1D12 variant proteins, analogs or fusion proteins thereof. For example, the 58P1D12 variant 1 amino acid sequence can be fused using recombinant DNA techniques to any one of a variety of fusion protein partners that are well known in the art, such as glutathione-S-transferase (GST) and HIS tagged fusion proteins. In another embodiment, amino acids 22-213 of 58P1D12 variant 1 is fused to His using recombinant techniques and the pET21b expression vector. The protein was then expressed, purified, and used to immunize 2 rabbits. Such fusion proteins are purified from induced bacteria using the appropriate affinity matrix.

Other recombinant bacterial fusion proteins that may be employed include maltose binding protein, LacZ, thioredoxin, NusA, or an immunoglobulin constant region (see the section entitled "Production of 58P1D12 in Prokaryotic Systems" and Current Protocols In Molecular Biology, Volume 2, Unit 16, Frederick M. Ausubul et al. eds., 1995; Linsley, P. S., Brady, W., Umes, M., Grosmaire, L., Damle, N., and Ledbetter, J. (1991) J. Exp. Med. 174, 561-566).

In addition to bacterial derived fusion proteins, mammalian expressed protein antigens are also used. These antigens are expressed from mammalian expression vectors such as the Tag5 and Fc-fusion vectors (see the section entitled "Production of Recombinant 58P1D12 in Eukaryotic Systems"), and retain post-translational modifications such as glycosylations found in native protein. In one embodiment, amino acids 22-213 of 58P1D12 variant 1 was cloned into the Tag5 mammalian secretion vector, and expressed in 293T cells. The recombinant protein was purified by metal chelate chromatography from tissue culture supernatants of 293T cells stably expressing the recombinant vector. The purified Tag5 58P1D12 protein was then used as immunogen.

During the immunization protocol, it is useful to mix or emulsify the antigen in adjuvants that enhance the immune response of the host animal. Examples of adjuvants include, but are not limited to, complete Freund's adjuvant (CFA) and MPL-TDM adjuvant (monophosphoryl Lipid A, synthetic trehalose dicorynomycolate).

In a typical protocol, rabbits are initially immunized subcutaneously with up to 200 μg, typically 100-200 μg, of fusion protein or peptide conjugated to KLH mixed in complete Freund's adjuvant (CFA). Rabbits are then injected subcutaneously every two weeks with up to 200 μg, typically 100-200 μg, of the immunogen in incomplete Freund's adjuvant (IFA). Test bleeds are taken approximately 7-10 days following each immunization and used to monitor the titer of the antiserum by ELISA.

To test reactivity and specificity of immune serum, such as the rabbit serum derived from immunization with the His-fusion of 58P1D12 variant 1 protein, the full-length 58P1D12 variant 1 cDNA was cloned into pCDNA 3.1 myc-his expression vector (Invitrogen, see the Example entitled "Production of Recombinant 58P1D12 in Eukaryotic Systems"). After transfection of the constructs into 293T cells, cell lysates are probed with the anti-58P1D12 serum and with anti-His antibody (Santa Cruz Biotechnologies, Santa Cruz, Calif.) to determine specific reactivity to denatured 58P1D12 protein using the Western blot technique. In addition, the immune serum is tested by fluorescence microscopy, flow cytometry and immunoprecipitation against 293T and other recombinant 58P1D12-expressing cells to determine specific recognition of native protein. Western blot, immunoprecipitation, fluorescent microscopy, and flow cytometric techniques using cells that endogenously express 58P1D12 are also carried out to test reactivity and specificity.

Anti-serum from rabbits immunized with 58P1D12 variant fusion proteins, such as GST and MBP fusion proteins, are purified by depletion of antibodies reactive to the fusion partner sequence by passage over an affinity column containing the fusion partner either alone or in the context of an irrelevant fusion protein. For example, antiserum derived from a GST-58P1D12 variant 1 fusion protein is first purified by passage over a column of GST protein covalently coupled to AffiGel matrix (BioRad, Hercules, Calif.). The antiserum is then affinity purified by passage over a column composed of a MBP-58P1D12 fusion protein covalently coupled to Affigel matrix. The serum is then further purified by protein G affinity chromatography to isolate the IgG fraction. Sera from other

His-tagged antigens and peptide immunized rabbits as well as fusion partner depleted sera are affinity purified by passage over a column matrix composed of the original protein immunogen or free peptide.

Example 11

Generation of 58P1D12 Monoclonal Antibodies (mAbs)

In one embodiment, therapeutic mAbs to 58P1D12 variants comprise those that react with epitopes specific for each variant protein or specific to sequences in common between the variants that would disrupt or modulate the biological function of the 58P1D12 variants, for example those that would disrupt the interaction with ligands and binding partners. Immunogens for generation of such mAbs include those designed to encode or contain the entire 58P1D12 protein variant sequence, regions predicted to contain functional motifs, and regions of the 58P1D12 protein variants predicted to be antigenic from computer analysis of the amino acid sequence (see, e.g., FIG. **5**, FIG. **6**, FIG. **7**, FIG. **8**, or FIG. **9**, and the Example entitled "Antigenicity Profiles and Secondary Structure"). Immunogens include peptides, recombinant bacterial proteins, and mammalian expressed Tag 5 proteins and human and murine IgG FC fusion proteins. In addition, cells engineered to express high levels of a respective 58P1D12 variant, such as Rat1-58P1D12 variant 1 or 300.19-58P1D12 variant 1 murine Pre-B cells, were used to immunize mice.

To generate mAbs to a 58P1D12 variant, mice are first immunized intraperitoneally (IP) with, typically, 10-50 μg of protein immunogen or 10^7 58P1D12-expressing cells mixed in complete Freund's adjuvant. Mice are then subsequently immunized IP every 2-4 weeks with, typically, 10-50 μg of protein immunogen or 10^7 cells mixed in incomplete Freund's adjuvant. Alternatively, MPL-TDM adjuvant is used in immunizations. In addition to the above protein and cell-based immunization strategies, a DNA-based immunization protocol is employed in which a mammalian expression vector encoding a 58P1D12 variant sequence is used to immunize mice by direct injection of the plasmid DNA. For example, amino acids 22-213 of 58P1D12 of variant 1 is cloned into the Tag5 mammalian secretion vector and the recombinant vector will then be used as immunogen. In another example the same amino acids are cloned into an Fc-fusion secretion vector in which the 58P1D12 variant 1 sequence is fused at the amino-terminus to an IgK leader sequence and at the carboxyl-terminus to the coding sequence of the human or murine IgG Fc region. This recombinant vector is then used as immunogen. The plasmid immunization protocols are used in combination with purified proteins expressed from the same vector and with cells expressing the respective 58P1D12 variant.

During the immunization protocol, test bleeds are taken 7-10 days following an injection to monitor titer and specificity of the immune response. Once appropriate reactivity and specificity is obtained as determined by ELISA, Western blotting, immunoprecipitation, fluorescence microscopy, and flow cytometric analyses, fusion and hybridoma generation is then carried out with established procedures well known in the art (see, e.g., Harlow and Lane, 1988).

In one embodiment for generating 58P1D12 monoclonal antibodies, amino acids 22-213 of 58P1D12 variant 1 was cloned into the pTag5 vector, the protein was then expressed in 293T cells and purified by metal chelate chromatography. Balb C mice were immunized with 10 μg of recombinant variant 1 protein mixed in adjuvant. Mice were subsequently immunized over several weeks with the protein. ELISA using the antigen determined the titer of serum from the immunized mice. Reactivity and specificity of the serum to full length 58P1D12 variant 1 protein was monitored by flow cytometry using Rat-58P1D12 cells compared to Rat1-neo control cells (FIG. **20**). Other recombinant 58P1D12 variant 1-expressing cells or cells endogenously expressing 58P1D12 variant 1 are also used. Mice showing the strongest reactivity were rested and given a final injection of antigen and sacrificed for fusion. The lymph nodes of the sacrificed mice were harvested and fused to SP2/0 myeloma cells using standard procedures (Harlow and Lane, 1988). Supernatants from HAT selected growth wells were analyzed by flow cytometry to identify specific-58P1D12 surface binding MAbs (FIG. **21**). Supernatants are also be screened by ELISA, Western blot, immunoprecipitation, and fluorescent microscopy to identify 58P1D12 specific antibody-producing clones.

In another embodiment, 58P1D12 variant 3 specific MAbs are generated by employing immunogens that encode amino acid sequences unique to variant 3. For example, a GST-fusion protein is constructed to encode amino acids 171-236, expressed, purified, and used to immunize mice. Hybridomas resulting from fusion of the B-cells from the mice are screened on cells expressing 58P1D12 variant 3 protein and cross-screened on cells expressing variant 1 protein to identify variant 3-specific MAbs.

The binding affinity of 58P1D12 variant specific monoclonal antibodies are determined using standard technologies. Affinity measurements quantify the strength of antibody to epitope binding and are used to help define which 58P1D12 variant monoclonal antibodies preferred for diagnostic or therapeutic use, as appreciated by one of skill in the art. The BIAcore system (Uppsala, Sweden) is a preferred method for determining binding affinity. The BIAcore system uses surface plasmon resonance (SPR, Welford K. 1991, Opt. Quant. Elect. 23:1; Morton and Myszka, 1998, Methods in Enzymology 295: 268) to monitor biomolecular interactions in real time. BIAcore analysis conveniently generates association rate constants, dissociation rate constants, equilibrium dissociation constants, and affinity constants. In addition, equilibrium binding analysis of a dilution series of the MAb is also used to determine affinity defined by the dissociation constant (KD). The KD is determined by non-linear regression of the equilibrium binding data of the concentration series. The KD is defined as the concentration at which half-maximal binding of the MAb to the antigen is attained. Such analysis for MAb M8-143(5).10.1 is shown in FIG. **22**.

Example 12

HLA Class I and Class II Binding Assays

HLA class I and class II binding assays using purified HLA molecules are performed in accordance with disclosed protocols (e.g., PCT publications WO 94/20127 and WO 94/03205; Sidney et al., Current Protocols in Immunology 18.3.1 (1998); Sidney, et al., J. Immunol. 154:247 (1995); Sette, et al., Mol. Immunol. 31:813 (1994)). Briefly, purified MHC molecules (5 to 500 nM) are incubated with various unlabeled peptide inhibitors and 1-10 nM 125I-radiolabeled probe peptides as described. Following incubation, MHC-peptide complexes are separated from free peptide by gel filtration and the fraction of peptide bound is determined. Typically, in preliminary experiments, each MHC preparation is titered in the presence of fixed amounts of radiolabeled peptides to determine the concentration of HLA molecules necessary to bind 10-20% of the total radioactivity. All subsequent inhibition and direct binding assays are performed using these HLA concentrations.

Since under these conditions [label]<[HLA] and lC50≧[HLA], the measured lC50 values are reasonable approximations of the true KD values. Peptide inhibitors are typically tested at concentrations ranging from 120 μg/ml to 1.2 ng/ml, and are tested in two to four completely independent experiments. To allow comparison of the data obtained in different experiments, a relative binding figure is calculated for each peptide by dividing the lC50 of a positive control for inhibition by the lC50 for each tested peptide (typically unlabeled versions of the radiolabeled probe peptide). For database purposes, and inter-experiment comparisons, relative binding values are compiled. These values can subsequently be converted back into lC50 nM values by dividing the lC50 nM of the positive controls for inhibition by the relative binding of the peptide of interest. This method of data compilation is accurate and consistent for comparing peptides that have been tested on different days, or with different lots of purified MHC.

Binding assays as outlined above may be used to analyze HLA supermotif and/or HLA motif-bearing peptides (see Table lV).

Example 13

ldentification of HLA Supermotif- and Motif-Bearing CTL Candidate Epitopes

HLA vaccine compositions of the invention can include multiple epitopes. The multiple epitopes can comprise multiple HLA supermotifs or motifs to achieve broad population coverage. This example illustrates the identification and confirmation of supermotif- and motif-bearing epitopes for the inclusion in such a vaccine composition. Calculation of population coverage is performed using the strategy described below.

Computer Searches and Algorithms for ldentification of Supermotif and/or Motif-Bearing Epitopes

The searches performed to identify the motif-bearing peptide sequences in the Example entitled “Antigenicity Profiles” and Tables Vlll-XXl and XXll-XLlX employ the protein sequence data from the gene product of 58P1D12 set forth in FlGS. **2** and **3**, the specific search peptides used to generate the tables are listed in Table Vll.

Computer searches for epitopes bearing HLA Class l or Class ll supermotifs or motifs are performed as follows. All translated 58P1D12 protein sequences are analyzed using a text string search software program to identify potential peptide sequences containing appropriate HLA binding motifs; such programs are readily produced in accordance with information in the art in view of known motif/supermotif disclosures. Furthermore, such calculations can be made mentally.

ldentified A2-, A3-, and DR-supermotif sequences are scored using polynomial algorithms to predict their capacity to bind to specific HLA-Class l or Class ll molecules. These polynomial algorithms account for the impact of different amino acids at different positions, and are essentially based on the premise that the overall affinity (or AG) of peptide-HLA molecule interactions can be approximated as a linear polynomial function of the type:

$$“\Delta G”=a1i\times a2i\times a3i\ldots\times ani$$

where aji is a coefficient which represents the effect of the presence of a given amino acid (j) at a given position (i) along the sequence of a peptide of n amino acids. The crucial assumption of this method is that the effects at each position are essentially independent of each other (i.e., independent binding of individual side-chains). When residue j occurs at position i in the peptide, it is assumed to contribute a constant amount ji to the free energy of binding of the peptide irrespective of the sequence of the rest of the peptide.

The method of derivation of specific algorithm coefficients has been described in Gulukota et al., J. Mol. Biol. 267:1258-126, 1997; (see also Sidney et al., Human lmmunol. 45:79-93, 1996; and Southwood et al., J. lmmunol. 160:3363-3373, 1998). Briefly, for all i positions, anchor and non-anchor alike, the geometric mean of the average relative binding (ARB) of all peptides carrying j is calculated relative to the remainder of the group, and used as the estimate of ji. For Class ll peptides, if multiple alignments are possible, only the highest scoring alignment is utilized, following an iterative procedure. To calculate an algorithm score of a given peptide in a test set, the ARB values corresponding to the sequence of the peptide are multiplied. lf this product exceeds a chosen threshold, the peptide is predicted to bind. Appropriate thresholds are chosen as a function of the degree of stringency of prediction desired.

Selection of HLA-A2 Supertype Cross-Reactive Peptides

Protein sequences from 58P1D12 are scanned utilizing motif identification software, to identify 8-, 9- 10- and 11-mer sequences containing the HLA-A2-supermotif main anchor specificity. Typically, these sequences are then scored using the protocol described above and the peptides corresponding to the positive-scoring sequences are synthesized and tested for their capacity to bind purified HLA-A*0201 molecules in vitro (HLA-A*0201 is considered a prototype A2 supertype molecule).

These peptides are then tested for the capacity to bind to additional A2-supertype molecules (A*0202, A*0203, A*0206, and A*6802). Peptides that bind to at least three of the five A2-supertype alleles tested are typically deemed A2-supertype cross-reactive binders. Preferred peptides bind at an affinity equal to or less than 500 nM to three or more HLA-A2 supertype molecules.

Selection of HLA-A3 Supermotif-Bearing Epitopes

The 58P1D12 protein sequence(s) scanned above is also examined for the presence of peptides with the HLA-A3-supermotif primary anchors. Peptides corresponding to the HLA A3 supermotif-bearing sequences are then synthesized and tested for binding to HLA-A*0301 and HLA-A*1101 molecules, the molecules encoded by the two most prevalent A3-supertype alleles. The peptides that bind at least one of the two alleles with binding affinities of ≦500 nM, often ≦200 nM, are then tested for binding cross-reactivity to the other common A3-supertype alleles (e.g., A*3101, A*3301, and A*6801) to identify those that can bind at least three of the five HLA-A3-supertype molecules tested.

Selection of HLA-B7 Supermotif Bearing Epitopes

The 58P1D12 protein(s) scanned above is also analyzed for the presence of 8-, 9- 10-, or 11-mer peptides with the HLA-B7-supermotif. Corresponding peptides are synthesized and tested for binding to HLA-B*0702, the molecule encoded by the most common B7-supertype allele (i.e., the prototype B7 supertype allele). Peptides binding B*0702 with lC50 of ≦500 nM are identified using standard methods. These peptides are then tested for binding to other common B7-supertype molecules (e.g., B*3501, B*5101, B*5301, and B*5401). Peptides capable of binding to three or more of the five B7-supertype alleles tested are thereby identified.

Selection of A1 and A24 Motif-Bearing Epitopes

To further increase population coverage, HLA-A1 and -A24 epitopes can also be incorporated into vaccine compositions. An analysis of the 58P1D12 protein can also be performed to identify HLA-A1- and A24-motif-containing sequences.

High affinity and/or cross-reactive binding epitopes that bear other motif and/or supermotifs are identified using analogous methodology.

Example 14

Confirmation of Immunogenicity

Cross-reactive candidate CTL A2-supermotif-bearing peptides that are identified as described herein are selected to confirm in vitro immunogenicity. Confirmation is performed using the following methodology:

Target Cell Lines for Cellular Screening:

The 221A2.1 cell line, produced by transferring the HLA-A2.1 gene into the HLA-A, -B, -C null mutant human B-lymphoblastoid cell line 721.221, is used as the peptide-loaded target to measure activity of HLA-A2.1-restricted CTL. This cell line is grown in RPMI-1640 medium supplemented with antibiotics, sodium pyruvate, nonessential amino acids and 10% (v/v) heat inactivated FCS. Cells that express an antigen of interest, or transfectants comprising the gene encoding the antigen of interest, can be used as target cells to confirm the ability of peptide-specific CTLs to recognize endogenous antigen.

Primary CTL Induction Cultures:

Generation of Dendritic Cells (DC): PBMCs are thawed in RPMI with 30 μg/ml DNAse, washed twice and resuspended in complete medium (RPMI-1640 plus 5% AB human serum, non-essential amino acids, sodium pyruvate, L-glutamine and penicillin/streptomycin). The monocytes are purified by plating 10×106 PBMC/well in a 6-well plate. After 2 hours at 37° C., the non-adherent cells are removed by gently shaking the plates and aspirating the supernatants. The wells are washed a total of three times with 3 ml RPMI to remove most of the non-adherent and loosely adherent cells. Three ml of complete medium containing 50 ng/ml of GM-CSF and 1,000 U/ml of IL-4 are then added to each well. TNFα is added to the DCs on day 6 at 75 ng/ml and the cells are used for CTL induction cultures on day 7.

Induction of CTL with DC and Peptide: CD8+ T-cells are isolated by positive selection with Dynal immunomagnetic beads (Dynabeads® M-450) and the Detacha-Bead® reagent. Typically about 200-250×106 PBMC are processed to obtain 24×106 CD8+ T-cells (enough for a 48-well plate culture). Briefly, the PBMCs are thawed in RPMI with 30 μg/ml DNAse, washed once with PBS containing 1% human AB serum and resuspended in PBS/1% AB serum at a concentration of 20×106 cells/ml. The magnetic beads are washed 3 times with PBS/AB serum, added to the cells (140 μl beads/20×106 cells) and incubated for 1 hour at 4° C. with continuous mixing. The beads and cells are washed 4× with PBS/AB serum to remove the nonadherent cells and resuspended at 100×106 cells/ml (based on the original cell number) in PBS/AB serum containing 100 μl/ml Detacha-Bead® reagent and 30 μg/ml DNAse. The mixture is incubated for 1 hour at room temperature with continuous mixing. The beads are washed again with PBS/AB/DNAse to collect the CD8+ T-cells. The DC are collected and centrifuged at 1300 rpm for 5-7 minutes, washed once with PBS with 1% BSA, counted and pulsed with 40 μg/ml of peptide at a cell concentration of 1-2×106/ml in the presence of 3 μg/ml β2-microglobulin for 4 hours at 20° C. The DC are then irradiated (4,200 rads), washed 1 time with medium and counted again.

Setting up induction cultures: 0.25 ml cytokine generated DC (at 1×105 cells/ml) are co-cultured with 0.25 ml of CD8+ T-cells (at 2×106 cell/ml) in each well of a 48-well plate in the presence of 10 ng/ml of IL-7. Recombinant human IL-10 is added the next day at a final concentration of 10 ng/ml and rhuman IL-2 is added 48 hours later at 10 IU/ml.

Restimulation of the induction cultures with peptide-pulsed adherent cells: Seven and fourteen days after the primary induction, the cells are restimulated with peptide-pulsed adherent cells. The PBMCs are thawed and washed twice with RPMI and DNAse. The cells are resuspended at 5×106 cells/ml and irradiated at ~4200 rads. The PBMCs are plated at 2×106 in 0.5 ml complete medium per well and incubated for 2 hours at 37° C. The plates are washed twice with RPMI by tapping the plate gently to remove the nonadherent cells and the adherent cells pulsed with 10 μg/ml of peptide in the presence of 3 μg/ml β2 microglobulin in 0.25 ml RPMI/5% AB per well for 2 hours at 37° C. Peptide solution from each well is aspirated and the wells are washed once with RPMI. Most of the media is aspirated from the induction cultures (CD8+ cells) and brought to 0.5 ml with fresh media. The cells are then transferred to the wells containing the peptide-pulsed adherent cells. Twenty four hours later recombinant human IL-10 is added at a final concentration of 10 ng/ml and recombinant human IL2 is added the next day and again 2-3 days later at 50 IU/ml (Tsai et al., Critical Reviews in Immunology 18(1-2):65-75, 1998). Seven days later, the cultures are assayed for CTL activity in a 51Cr release assay. In some experiments the cultures are assayed for peptide-specific recognition in the in situ IFNγ ELISA at the time of the second restimulation followed by assay of endogenous recognition 7 days later. After expansion, activity is measured in both assays for a side-by-side comparison.

Measurement of CTL Lytic Activity by 51Cr Release.

Seven days after the second restimulation, cytotoxicity is determined in a standard (5 hr) 51Cr release assay by assaying individual wells at a single E:T. Peptide-pulsed targets are prepared by incubating the cells with 10 μg/ml peptide overnight at 37° C.

Adherent target cells are removed from culture flasks with trypsin-EDTA. Target cells are labeled with 200 μCi of 51Cr sodium chromate (Dupont, Wilmington, Del.) for 1 hour at 37° C. Labeled target cells are resuspended at 106 per ml and diluted 1:10 with K562 cells at a concentration of 3.3×106/ml (an NK-sensitive erythroblastoma cell line used to reduce non-specific lysis). Target cells (100 μl) and effectors (100 μl) are plated in 96 well round-bottom plates and incubated for 5 hours at 37° C. At that time, 100 μl of supernatant are collected from each well and percent lysis is determined according to the formula:

[(cpm of the test sample−cpm of the spontaneous 51Cr release sample)/(cpm of the maximal 51Cr release sample−cpm of the spontaneous 51Cr release sample)]×100.

Maximum and spontaneous release are determined by incubating the labeled targets with 1% Triton X-100 and media alone, respectively. A positive culture is defined as one in which the specific lysis (sample-background) is 10% or higher in the case of individual wells and is 15% or more at the two highest E:T ratios when expanded cultures are assayed.

In Situ Measurement of Human IFNγ Production as an Indicator of Peptide-Specific and Endogenous Recognition

Immulon 2 plates are coated with mouse anti-human IFNγ monoclonal antibody (4 μg/ml 0.1M NaHCO3, pH8.2) overnight at 4° C. The plates are washed with Ca2+, Mg2+-free PBS/0.05% Tween 20 and blocked with PBS/10% FCS for two hours, after which the CTLs (100 μl/well) and targets (100 μl/well) are added to each well, leaving empty wells for the standards and blanks (which received media only). The target cells, either peptide-pulsed or endogenous targets, are used at a concentration of 1×106 cells/ml. The plates are incubated for 48 hours at 37° C. with 5% CO2.

Recombinant human lFN-gamma is added to the standard wells starting at 400 pg or 1200 pg/100 microliter/well and the plate incubated for two hours at 37° C. The plates are washed and 100 μl of biotinylated mouse anti-human lFN-gamma monoclonal antibody (2 microgram/ml in PBS/3% FCS/0.05% Tween 20) are added and incubated for 2 hours at room temperature. After washing again, 100 microliter HRP-streptavidin (1:4000) are added and the plates incubated for one hour at room temperature. The plates are then washed 6× with wash buffer, 100 microliter/well developing solution (TMB 1:1) are added, and the plates allowed to develop for 5-15 minutes. The reaction is stopped with 50 microliter/well 1M H3PO4 and read at OD450. A culture is considered positive if it measured at least 50 pg of lFN-gamma/well above background and is twice the background level of expression.

CTL Expansion.

Those cultures that demonstrate specific lytic activity against peptide-pulsed targets and/or tumor targets are expanded over a two week period with anti-CD3. Briefly, 5×104 CD8+ cells are added to a T25 flask containing the following: 1×106 irradiated (4,200 rad) PBMC (autologous or allogeneic) per ml, 2×105 irradiated (8,000 rad) EBV-transformed cells per ml, and OKT3 (anti-CD3) at 30 ng per ml in RPMl-1640 containing 10% (v/v) human AB serum, non-essential amino acids, sodium pyruvate, 25 μM 2-mercaptoethanol, L-glutamine and penicillin/streptomycin. Recombinant human lL2 is added 24 hours later at a final concentration of 200 lU/ml and every three days thereafter with fresh media at 50 lU/ml. The cells are split if the cell concentration exceeds 1×106/ml and the cultures are assayed between days 13 and 15 at E:T ratios of 30, 10, 3 and 1:1 in the 51Cr release assay or at 1×106/ml in the in situ lFNγ assay using the same targets as before the expansion.

Cultures are expanded in the absence of anti-CD3+ as follows. Those cultures that demonstrate specific lytic activity against peptide and endogenous targets are selected and 5×104 CD8+ cells are added to a T25 flask containing the following: 1×106 autologous PBMC per ml which have been peptide-pulsed with 10 μg/ml peptide for two hours at 37° C. and irradiated (4,200 rad); 2×105 irradiated (8,000 rad) EBV-transformed cells per ml RPMl-1640 containing 10%(v/v) human AB serum, non-essential AA, sodium pyruvate, 25 mM 2-ME, L-glutamine and gentamicin.

lmmunogenicity of A2 Supermotif-Bearing Peptides

A2-supermotif cross-reactive binding peptides are tested in the cellular assay for the ability to induce peptide-specific CTL in normal individuals. ln this analysis, a peptide is typically considered to be an epitope if it induces peptide-specific CTLs in at least individuals, and preferably, also recognizes the endogenously expressed peptide.

lmmunogenicity can also be confirmed using PBMCs isolated from patients bearing a tumor that expresses 58P1D12. Briefly, PBMCs are isolated from patients, re-stimulated with peptide-pulsed monocytes and assayed for the ability to recognize peptide-pulsed target cells as well as transfected cells endogenously expressing the antigen.

Evaluation of A*03/A11 lmmunogenicity

HLA-A3 supermotif-bearing cross-reactive binding peptides are also evaluated for immunogenicity using methodology analogous for that used to evaluate the immunogenicity of the HLA-A2 supermotif peptides.

Evaluation of B7 lmmunogenicity

lmmunogenicity screening of the B7-supertype cross-reactive binding peptides identified as set forth herein are confirmed in a manner analogous to the confirmation of A2- and A3-supermotif-bearing peptides.

Peptides bearing other supermotifs/motifs, e.g., HLA-A1, HLA-A24 etc. are also confirmed using similar methodology.

Example 15

lmplementation of the Extended Supermotif to lmprove the Binding Capacity of Native Epitopes by Creating Analogs

HLA motifs and supermotifs (comprising primary and/or secondary residues) are useful in the identification and preparation of highly cross-reactive native peptides, as demonstrated herein. Moreover, the definition of HLA motifs and supermotifs also allows one to engineer highly cross-reactive epitopes by identifying residues within a native peptide sequence which can be analoged to confer upon the peptide certain characteristics, e.g. greater cross-reactivity within the group of HLA molecules that comprise a supertype, and/or greater binding affinity for some or all of those HLA molecules. Examples of analoging peptides to exhibit modulated binding affinity are set forth in this example.

Analoging at Primary Anchor Residues

Peptide engineering strategies are implemented to further increase the cross-reactivity of the epitopes. For example, the main anchors of A2-supermotif-bearing peptides are altered, for example, to introduce a preferred L, l, V, or M at position 2, and l or V at the C-terminus.

To analyze the cross-reactivity of the analog peptides, each engineered analog is initially tested for binding to the prototype A2 supertype allele A*0201, then, if A*0201 binding capacity is maintained, for A2-supertype cross-reactivity.

Alternatively, a peptide is confirmed as binding one or all supertype members and then analoged to modulate binding affinity to any one (or more) of the supertype members to add population coverage.

The selection of analogs for immunogenicity in a cellular screening analysis is typically further restricted by the capacity of the parent wild type (WT) peptide to bind at least weakly, i.e., bind at an lC50 of 5000 nM or less, to three of more A2 supertype alleles. The rationale for this requirement is that the WT peptides must be present endogenously in sufficient quantity to be biologically relevant. Analoged peptides have been shown to have increased immunogenicity and cross-reactivity by T cells specific for the parent epitope (see, e.g., Parkhurst et al., J. lmmunol. 157:2539, 1996; and Pogue et al., Proc. Natl. Acad. Sci. USA 92:8166, 1995).

ln the cellular screening of these peptide analogs, it is important to confirm that analog-specific CTLs are also able to recognize the wild-type peptide and, when possible, target cells that endogenously express the epitope.

Analoging of HLA-A3 and B7-Supermotif-Bearing Peptides

Analogs of HLA-A3 supermotif-bearing epitopes are generated using strategies similar to those employed in analoging HLA-A2 supermotif-bearing peptides. For example, peptides binding to 3/5 of the A3-supertype molecules are engineered at primary anchor residues to possess a preferred residue (V, S, M, or A) at position 2.

The analog peptides are then tested for the ability to bind A*03 and A*11 (prototype A3 supertype alleles). Those peptides that demonstrate ≦500 nM binding capacity are then confirmed as having A3-supertype cross-reactivity.

Similarly to the A2- and A3-motif bearing peptides, peptides binding 3 or more B7-supertype alleles can be improved, where possible, to achieve increased cross-reactive binding or greater binding affinity or binding half life. B7 supermotif-bearing peptides are, for example, engineered to possess a preferred residue (V, I, L, or F) at the C-terminal primary anchor position, as demonstrated by Sidney et al. (J. Immunol. 157:3480-3490, 1996).

Analoging at primary anchor residues of other motif and/or supermotif-bearing epitopes is performed in a like manner.

The analog peptides are then be confirmed for immunogenicity, typically in a cellular screening assay. Again, it is generally important to demonstrate that analog-specific CTLs are also able to recognize the wild-type peptide and, when possible, targets that endogenously express the epitope.

Analoging at Secondary Anchor Residues

Moreover, HLA supermotifs are of value in engineering highly cross-reactive peptides and/or peptides that bind HLA molecules with increased affinity by identifying particular residues at secondary anchor positions that are associated with such properties. For example, the binding capacity of a B7 supermotif-bearing peptide with an F residue at position 1 is analyzed. The peptide is then analoged to, for example, substitute L for F at position 1. The analoged peptide is evaluated for increased binding affinity, binding half life and/or increased cross-reactivity. Such a procedure identifies analoged peptides with enhanced properties.

Engineered analogs with sufficiently improved binding capacity or cross-reactivity can also be tested for immunogenicity in HLA-B7-transgenic mice, following for example, IFA immunization or lipopeptide immunization. Analoged peptides are additionally tested for the ability to stimulate a recall response using PBMC from patients with 58P1D12-expressing tumors.

Other Analoging Strategies

Another form of peptide analoging, unrelated to anchor positions, involves the substitution of a cysteine with α-amino butyric acid. Due to its chemical nature, cysteine has the propensity to form disulfide bridges and sufficiently alter the peptide structurally so as to reduce binding capacity. Substitution of α-amino butyric acid for cysteine not only alleviates this problem, but has been shown to improve binding and crossbinding capabilities in some instances (see, e.g., the review by Sette et al., In: Persistent Viral Infections, Eds. R. Ahmed and I. Chen, John Wiley & Sons, England, 1999).

Thus, by the use of single amino acid substitutions, the binding properties and/or cross-reactivity of peptide ligands for HLA supertype molecules can be modulated.

Example 16

Identification and Confirmation of 58P1D12-Derived Sequences with HLA-DR Binding Motifs

Peptide epitopes bearing an HLA class II supermotif or motif are identified and confirmed as outlined below using methodology similar to that described for HLA Class I peptides.

Selection of HLA-DR-Supermotif-Bearing Epitopes.

To identify 58P1D12-derived, HLA class II HTL epitopes, a 58P1D12 antigen is analyzed for the presence of sequences bearing an HLA-DR-motif or supermotif. Specifically, 15-mer sequences are selected comprising a DR-supermotif, comprising a 9-mer core, and three-residue N- and C-terminal flanking regions (15 amino acids total).

Protocols for predicting peptide binding to DR molecules have been developed (Southwood et al., J. Immunol. 160: 3363-3373, 1998). These protocols, specific for individual DR molecules, allow the scoring, and ranking, of 9-mer core regions. Each protocol not only scores peptide sequences for the presence of DR-supermotif primary anchors (i.e., at position 1 and position 6) within a 9-mer core, but additionally evaluates sequences for the presence of secondary anchors. Using allele-specific selection tables (see, e.g., Southwood et al., ibid.), it has been found that these protocols efficiently select peptide sequences with a high probability of binding a particular DR molecule. Additionally, it has been found that performing these protocols in tandem, specifically those for DR1, DR4w4, and DR7, can efficiently select DR cross-reactive peptides.

The 58P1D12-derived peptides identified above are tested for their binding capacity for various common HLA-DR molecules. All peptides are initially tested for binding to the DR molecules in the primary panel: DR1, DR4w4, and DR7. Peptides binding at least two of these three DR molecules are then tested for binding to DR2w2 β1, DR2w2 β2, DR6w19, and DR9 molecules in secondary assays. Finally, peptides binding at least two of the four secondary panel DR molecules, and thus cumulatively at least four of seven different DR molecules, are screened for binding to DR4w15, DR5w11, and DR8w2 molecules in tertiary assays. Peptides binding at least seven of the ten DR molecules comprising the primary, secondary, and tertiary screening assays are considered cross-reactive DR binders. 58P1D12-derived peptides found to bind common HLA-DR alleles are of particular interest.

Selection of DR3 Motif Peptides

Because HLA-DR3 is an allele that is prevalent in Caucasian, Black, and Hispanic populations, DR3 binding capacity is a relevant criterion in the selection of HTL epitopes. Thus, peptides shown to be candidates may also be assayed for their DR3 binding capacity. However, in view of the binding specificity of the DR3 motif, peptides binding only to DR3 can also be considered as candidates for inclusion in a vaccine formulation.

To efficiently identify peptides that bind DR3, target 58P1D12 antigens are analyzed for sequences carrying one of the two DR3-specific binding motifs reported by Geluk et al. (J. Immunol. 152:5742-5748, 1994). The corresponding peptides are then synthesized and confirmed as having the ability to bind DR3 with an affinity of 1 μM or better, i.e., less than 1 μM. Peptides are found that meet this binding criterion and qualify as HLA class II high affinity binders.

DR3 binding epitopes identified in this manner are included in vaccine compositions with DR supermotif-bearing peptide epitopes.

Similarly to the case of HLA class I motif-bearing peptides, the class II motif-bearing peptides are analoged to improve affinity or cross-reactivity. For example, aspartic acid at position 4 of the 9-mer core sequence is an optimal residue for DR3 binding, and substitution for that residue often improves DR 3 binding.

Example 17

Immunogenicity of 58P1D12-Derived HTL Epitopes

This example determines immunogenic DR supermotif- and DR3 motif-bearing epitopes among those identified using the methodology set forth herein.

Immunogenicity of HTL epitopes are confirmed in a manner analogous to the determination of immunogenicity of CTL epitopes, by assessing the ability to stimulate HTL responses and/or by using appropriate transgenic mouse models. Immunogenicity is determined by screening for: 1.) in vitro primary induction using normal PBMC or 2.) recall responses from patients who have 58P1D12-expressing tumors.

Example 18

Calculation of Phenotypic Frequencies of HLA-Supertypes in Various Ethnic Backgrounds to Determine Breadth of Population Coverage

This example illustrates the assessment of the breadth of population coverage of a vaccine composition comprised of multiple epitopes comprising multiple supermotifs and/or motifs.

In order to analyze population coverage, gene frequencies of HLA alleles are determined. Gene frequencies for each HLA allele are calculated from antigen or allele frequencies utilizing the binomial distribution formulae gf=1−(SQRT(1−af)) (see, e.g., Sidney et al., Human Immunol. 45:79-93, 1996). To obtain overall phenotypic frequencies, cumulative gene frequencies are calculated, and the cumulative antigen frequencies derived by the use of the inverse formula [af=1−(1−Cgf)2].

Where frequency data is not available at the level of DNA typing, correspondence to the serologically defined antigen frequencies is assumed. To obtain total potential supertype population coverage no linkage disequilibrium is assumed, and only alleles confirmed to belong to each of the supertypes are included (minimal estimates). Estimates of total potential coverage achieved by inter-loci combinations are made by adding to the A coverage the proportion of the non-A covered population that could be expected to be covered by the B alleles considered (e.g., total=A+B*(1-A)). Confirmed members of the A3-like supertype are A3, A11, A31, A*3301, and A*6801. Although the A3-like supertype may also include A34, A66, and A*7401, these alleles were not included in overall frequency calculations. Likewise, confirmed members of the A2-like supertype family are A*0201, A*0202, A*0203, A*0204, A*0205, A*0206, A*0207, A*6802, and A*6901. Finally, the B7-like supertype-confirmed alleles are: B7, B*3501-03, B51, B*5301, B*5401, B*5501-2, B*5601, B*6701, and B*7801 (potentially also B*1401, B*3504-06, B*4201, and B*5602).

Population coverage achieved by combining the A2-, A3- and B7-supertypes is approximately 86% in five major ethnic groups. Coverage may be extended by including peptides bearing the A1 and A24 motifs. On average, A1 is present in 12% and A24 in 29% of the population across five different major ethnic groups (Caucasian, North American Black, Chinese, Japanese, and Hispanic). Together, these alleles are represented with an average frequency of 39% in these same ethnic populations. The total coverage across the major ethnicities when A1 and A24 are combined with the coverage of the A2-, A3- and B7-supertype alleles is >95%, see, e.g., Table IV (G). An analogous approach can be used to estimate population coverage achieved with combinations of class II motif-bearing epitopes.

Immunogenicity studies in humans (e.g., Bertoni et al., J. Clin. Invest. 100:503, 1997; Doolan et al., Immunity 7:97, 1997; and Threlkeld et al., J. Immunol. 159:1648, 1997) have shown that highly cross-reactive binding peptides are almost always recognized as epitopes. The use of highly cross-reactive binding peptides is an important selection criterion in identifying candidate epitopes for inclusion in a vaccine that is immunogenic in a diverse population.

With a sufficient number of epitopes (as disclosed herein and from the art), an average population coverage is predicted to be greater than 95% in each of five major ethnic populations. The game theory Monte Carlo simulation analysis, which is known in the art (see e.g., Osborne, M. J. and Rubinstein, A. "A course in game theory" MIT Press, 1994), can be used to estimate what percentage of the individuals in a population comprised of the Caucasian, North American Black, Japanese, Chinese, and Hispanic ethnic groups would recognize the vaccine epitopes described herein. A preferred percentage is 90%. A more preferred percentage is 95%.

Example 19

CTL Recognition of Endogenously Processed Antigens After Priming

This example confirms that CTL induced by native or analoged peptide epitopes identified and selected as described herein recognize endogenously synthesized, i.e., native antigens.

Effector cells isolated from transgenic mice that are immunized with peptide epitopes, for example HLA-A2 supermotif-bearing epitopes, are re-stimulated in vitro using peptide-coated stimulator cells. Six days later, effector cells are assayed for cytotoxicity and the cell lines that contain peptide-specific cytotoxic activity are further re-stimulated. An additional six days later, these cell lines are tested for cytotoxic activity on ^{51}Cr labeled Jurkat-A2.1/K^{b} target cells in the absence or presence of peptide, and also tested on ^{51}Cr labeled target cells bearing the endogenously synthesized antigen, i.e. cells that are stably transfected with 58P1D12 expression vectors.

The results demonstrate that CTL lines obtained from animals primed with peptide epitope recognize endogenously synthesized 58P1D12 antigen. The choice of transgenic mouse model to be used for such an analysis depends upon the epitope(s) that are being evaluated. In addition to HLA-A*0201/K^{b} transgenic mice, several other transgenic mouse models including mice with human A11, which may also be used to evaluate A3 epitopes, and B7 alleles have been characterized and others (e.g., transgenic mice for HLA-A1 and A24) are being developed. HLA-DR1 and HLA-DR3 mouse models have also been developed, which may be used to evaluate HTL epitopes.

Example 20

Activity of CTL-HTL Conjugated Epitopes in Transgenic Mice

This example illustrates the induction of CTLs and HTLs in transgenic mice, by use of a 58P1D12-derived CTL and HTL peptide vaccine compositions. The vaccine composition used herein comprise peptides to be administered to a patient with a 58P1D12-expressing tumor. The peptide composition can comprise multiple CTL and/or HTL epitopes. The epitopes are identified using methodology as described herein. This example also illustrates that enhanced immunogenicity can be achieved by inclusion of one or more HTL epitopes in a CTL vaccine composition; such a peptide composition can comprise an HTL epitope conjugated to a CTL epitope. The CTL epitope can be one that binds to multiple HLA family members at an affinity of 500 nM or less, or analogs of that epitope. The peptides may be lipidated, if desired.

Immunization procedures: Immunization of transgenic mice is performed as described (Alexander et al., *J. Immunol.* 159:4753-4761, 1997). For example, A2/K^b mice, which are transgenic for the human HLA A2.1 allele and are used to confirm the immunogenicity of HLA-A*0201 motif- or HLA-A2 supermotif-bearing epitopes, and are primed subcutaneously (base of the tail) with a 0.1 ml of peptide in Incomplete Freund's Adjuvant, or if the peptide composition is a lipidated CTL/HTL conjugate, in DMSO/saline, or if the peptide composition is a polypeptide, in PBS or Incomplete Freund's Adjuvant. Seven days after priming, splenocytes obtained from these animals are restimulated with syngenic irradiated LPS-activated lymphoblasts coated with peptide.

Cell lines: Target cells for peptide-specific cytotoxicity assays are Jurkat cells transfected with the HLA-A2.1/K^b chimeric gene (e.g., Vitiello et al., *J. Exp. Med.* 173:1007, 1991).

In vitro CTL activation: One week after priming, spleen cells (30x10^6 cells/flask) are co-cultured at 37° C. with syngeneic, irradiated (3000 rads), peptide coated lymphoblasts (10x10^6 cells/flask) in 10 ml of culture medium/T25 flask. After six days, effector cells are harvested and assayed for cytotoxic activity.

Assay for cytotoxic activity: Target cells (1.0 to 1.5x10^6) are incubated at 37° C. in the presence of 200 μl of ^{51}Cr. After 60 minutes, cells are washed three times and resuspended in R10 medium. Peptide is added where required at a concentration of 1 μg/ml. For the assay, 10^4 ^{51}Cr-labeled target cells are added to different concentrations of effector cells (final volume of 200 μl) in U-bottom 96-well plates. After a six hour incubation period at 37° C., a 0.1 ml aliquot of supernatant is removed from each well and radioactivity is determined in a Micromedic automatic gamma counter. The percent specific lysis is determined by the formula: percent specific release=100×(experimental release−spontaneous release)/(maximum release−spontaneous release). To facilitate comparison between separate CTL assays run under the same conditions, % ^{51}Cr release data is expressed as lytic units/10^6 cells. One lytic unit is arbitrarily defined as the number of effector cells required to achieve 30% lysis of 10,000 target cells in a six hour ^{51}Cr release assay. To obtain specific lytic units/10^6, the lytic units/10^6 obtained in the absence of peptide is subtracted from the lytic units/10^6 obtained in the presence of peptide. For example, if 30% ^{51}Cr release is obtained at the effector (E): target (T) ratio of 50:1 (i.e., 5x10^5 effector cells for 10,000 targets) in the absence of peptide and 5:1 (i.e., 5x10^4 effector cells for 10,000 targets) in the presence of peptide, the specific lytic units would be: [(1/50,000)−(1/500,000)]x10^6=18 LU.

The results are analyzed to assess the magnitude of the CTL responses of animals injected with the immunogenic CTL/HTL conjugate vaccine preparation and are compared to the magnitude of the CTL response achieved using, for example, CTL epitopes as outlined above in the Example entitled "Confirmation of Immunogenicity." Analyses similar to this may be performed to confirm the immunogenicity of peptide conjugates containing multiple CTL epitopes and/or multiple HTL epitopes. In accordance with these procedures, it is found that a CTL response is induced, and concomitantly that an HTL response is induced upon administration of such compositions.

Example 21

Selection of CTL and HTL Epitopes for Inclusion in a 58P1D12-Specific Vaccine

This example illustrates a procedure for selecting peptide epitopes for vaccine compositions of the invention. The peptides in the composition can be in the form of a nucleic acid sequence, either single or one or more sequences (i.e., minigene) that encodes peptide(s), or can be single and/or polyepitopic peptides.

The following principles are utilized when selecting a plurality of epitopes for inclusion in a vaccine composition. Each of the following principles is balanced in order to make the selection.

Epitopes are selected which, upon administration, mimic immune responses that are correlated with 58P1D12 clearance. The number of epitopes used depends on observations of patients who spontaneously clear 58P1D12. For example, if it has been observed that patients who spontaneously clear 58P1D12-expressing cells generate an immune response to at least three (3) epitopes from 58P1D12 antigen, then at least three epitopes should be included for HLA class I. A similar rationale is used to determine HLA class II epitopes.

Epitopes are often selected that have a binding affinity of an IC$_{50}$ of 500 nM or less for an HLA class I molecule, or for class II, an IC$_{50}$ of 1000 nM or less; or HLA Class I peptides with high binding scores from the BIMAS web site, at URL bimas.dcrt.nih.gov/.

In order to achieve broad coverage of the vaccine through out a diverse population, sufficient supermotif bearing peptides, or a sufficient array of allele-specific motif bearing peptides, are selected to give broad population coverage. In one embodiment, epitopes are selected to provide at least 80% population coverage. A Monte Carlo analysis, a statistical evaluation known in the art, can be employed to assess breadth, or redundancy, of population coverage.

When creating polyepitopic compositions, or a minigene that encodes same, it is typically desirable to generate the smallest peptide possible that encompasses the epitopes of interest. The principles employed are similar, if not the same, as those employed when selecting a peptide comprising nested epitopes. For example, a protein sequence for the vaccine composition is selected because it has maximal number of epitopes contained within the sequence, i.e., it has a high concentration of epitopes. Epitopes may be nested or overlapping (i.e., frame shifted relative to one another). For example, with overlapping epitopes, two 9-mer epitopes and one 10-mer epitope can be present in a 10 amino acid peptide. Each epitope can be exposed and bound by an HLA molecule upon administration of such a peptide. A multi-epitopic, peptide can be generated synthetically, recombinantly, or via cleavage from the native source. Alternatively, an analog can be made of this native sequence, whereby one or more of the epitopes comprise substitutions that alter the cross-reactivity and/or binding affinity properties of the polyepitopic peptide. Such a vaccine composition is administered for therapeutic or prophylactic purposes. This embodiment provides for the possibility that an as yet undiscovered aspect of immune system processing will apply to the native nested sequence and thereby facilitate the production of therapeutic or prophylactic immune response-inducing vaccine compositions. Additionally such an embodiment provides for the possibility of motif-bearing epitopes for an HLA makeup that is presently unknown. Furthermore, this embodiment (absent the creating of any analogs) directs the immune response to multiple peptide sequences that are actually present in 58P1D12, thus avoiding the need to evaluate any junctional epitopes. Lastly, the embodiment provides an economy of scale when producing nucleic acid vaccine compositions. Related to this embodiment, computer programs can be derived in accordance with principles in the art, which identify in a target sequence, the greatest number of epitopes per sequence length.

A vaccine composition comprised of selected peptides, when administered, is safe, efficacious, and elicits an immune response similar in magnitude to an immune response that controls or clears cells that bear or overexpress 58P1D12.

Example 22

Construction of "Minigene" Multi-Epitope DNA Plasmids

This example discusses the construction of a minigene expression plasmid. Minigene plasmids may, of course, contain various configurations of B cell, CTL and/or HTL epitopes or epitope analogs as described herein.

A minigene expression plasmid typically includes multiple CTL and HTL peptide epitopes. In the present example, HLA-A2, -A3, -B7 supermotif-bearing peptide epitopes and HLA-A1 and -A24 motif-bearing peptide epitopes are used in conjunction with DR supermotif-bearing epitopes and/or DR3 epitopes. HLA class I supermotif or motif-bearing peptide epitopes derived 58P1D12, are selected such that multiple supermotifs/motifs are represented to ensure broad population coverage. Similarly, HLA class II epitopes are selected from 58P1D12 to provide broad population coverage, i.e. both HLA DR-1-4-7 supermotif-bearing epitopes and HLA DR-3 motif-bearing epitopes are selected for inclusion in the minigene construct. The selected CTL and HTL epitopes are then incorporated into a minigene for expression in an expression vector.

Such a construct may additionally include sequences that direct the HTL epitopes to the endoplasmic reticulum. For example, the Ii protein may be fused to one or more HTL epitopes as described in the art, wherein the CLIP sequence of the Ii protein is removed and replaced with an HLA class II epitope sequence so that HLA class II epitope is directed to the endoplasmic reticulum, where the epitope binds to an HLA class II molecules.

This example illustrates the methods to be used for construction of a minigene-bearing expression plasmid. Other expression vectors that may be used for minigene compositions are available and known to those of skill in the art.

The minigene DNA plasmid of this example contains a consensus Kozak sequence and a consensus murine kappa Ig-light chain signal sequence followed by CTL and/or HTL epitopes selected in accordance with principles disclosed herein. The sequence encodes an open reading frame fused to the Myc and His antibody epitope tag coded for by the pcDNA 3.1 Myc-His vector.

Overlapping oligonucleotides that can, for example, average about 70 nucleotides in length with 15 nucleotide overlaps, are synthesized and HPLC-purified. The oligonucleotides encode the selected peptide epitopes as well as appropriate linker nucleotides, Kozak sequence, and signal sequence. The final multiepitope minigene is assembled by extending the overlapping oligonucleotides in three sets of reactions using PCR. A Perkin/Elmer 9600 PCR machine is used and a total of 30 cycles are performed using the following conditions: 95° C. for 15 sec, annealing temperature (5° below the lowest calculated Tm of each primer pair) for 30 sec, and 72° C. for 1 min.

For example, a minigene is prepared as follows. For a first PCR reaction, 5 μg of each of two oligonucleotides are annealed and extended: In an example using eight oligonucleotides, i.e., four pairs of primers, oligonucleotides 1+2, 3+4, 5+6, and 7+8 are combined in 100 μl reactions containing Pfu polymerase buffer (1×=10 mM KCL, 10 mM $(NH4)_2SO_4$, 20 mM Tris-chloride, pH 8.75, 2 mM $MgSO_4$, 0.1% Triton X-100, 100 μg/ml BSA), 0.25 mM each dNTP, and 2.5 U of Pfu polymerase. The full-length dimer products are gel-purified, and two reactions containing the product of 1+2 and 3+4, and the product of 5+6 and 7+8 are mixed, annealed, and extended for 10 cycles. Half of the two reactions are then mixed, and 5 cycles of annealing and extension carried out before flanking primers are added to amplify the full length product. The full-length product is gel-purified and cloned into pCR-blunt (Invitrogen) and individual clones are screened by sequencing.

Example 23

The Plasmid Construct and the Degree to which it Induces Immunogenicity

The degree to which a plasmid construct, for example a plasmid constructed in accordance with the previous Example, is able to induce immunogenicity is confirmed in vitro by determining epitope presentation by APC following transduction or transfection of the APC with an epitope-expressing nucleic acid construct. Such a study determines "antigenicity" and allows the use of human APC. The assay determines the ability of the epitope to be presented by the APC in a context that is recognized by a T cell by quantifying the density of epitope-HLA class I complexes on the cell surface. Quantitation can be performed by directly measuring the amount of peptide eluted from the APC (see, e.g., Sijts et al., *J. Immunol.* 156:683-692, 1996; Demotz et al., *Nature* 342:682-684, 1989); or the number of peptide-HLA class I complexes can be estimated by measuring the amount of lysis or lymphokine release induced by diseased or transfected target cells, and then determining the concentration of peptide necessary to obtain equivalent levels of lysis or lymphokine release (see, e.g., Kageyama et al., *J. Immunol.* 154:567-576, 1995).

Alternatively, immunogenicity is confirmed through in vivo injections into mice and subsequent in vitro assessment of CTL and HTL activity, which are analyzed using cytotoxicity and proliferation assays, respectively, as detailed e.g., in Alexander et al., *Immunity* 1:751-761, 1994.

For example, to confirm the capacity of a DNA minigene construct containing at least one HLA-A2 supermotif peptide to induce CTLs in vivo, HLA-A2.1/K^b transgenic mice, for example, are immunized intramuscularly with 100 μg of naked cDNA. As a means of comparing the level of CTLs induced by cDNA immunization, a control group of animals is also immunized with an actual peptide composition that comprises multiple epitopes synthesized as a single polypeptide as they would be encoded by the minigene.

Splenocytes from immunized animals are stimulated twice with each of the respective compositions (peptide epitopes encoded in the minigene or the polyepitopic peptide), then assayed for peptide-specific cytotoxic activity in a ^{51}Cr release assay. The results indicate the magnitude of the CTL response directed against the A2-restricted epitope, thus indicating the in vivo immunogenicity of the minigene vaccine and polyepitopic vaccine.

It is, therefore, found that the minigene elicits immune responses directed toward the HLA-A2 supermotif peptide epitopes as does the polyepitopic peptide vaccine. A similar analysis is also performed using other HLA-A3 and HLA-B7 transgenic mouse models to assess CTL induction by HLA-A3 and HLA-B7 motif or supermotif epitopes, whereby it is also found that the minigene elicits appropriate immune responses directed toward the provided epitopes.

To confirm the capacity of a class II epitope-encoding minigene to induce HTLs in vivo, DR transgenic mice, or for those epitopes that cross react with the appropriate mouse MHC molecule, I-Ab-restricted mice, for example, are immunized intramuscularly with 100 μg of plasmid DNA. As a means of comparing the level of HTLs induced by DNA immunization, a group of control animals is also immunized with an actual peptide composition emulsified in complete Freund's adjuvant. CD4+ T cells, i.e. HTLs, are purified from splenocytes of immunized animals and stimulated with each of the respective compositions (peptides encoded in the minigene). The HTL response is measured using a ^{3}H-thymidine incorporation proliferation assay, (see, e.g., Alexander et al. Immunity 1:751-761, 1994). The results indicate the magnitude of the HTL response, thus demonstrating the in vivo immunogenicity of the minigene.

DNA minigenes, constructed as described in the previous Example, can also be confirmed as a vaccine in combination with a boosting agent using a prime boost protocol. The boosting agent can consist of recombinant protein (e.g., Barnett et at, *Aids Res. and Human Retroviruses* 14, Supplement 3:S299-S309, 1998) or recombinant vaccinia, for example, expressing a minigene or DNA encoding the complete protein of interest (see, e.g., Hanke et al., *Vaccine* 16:439-445, 1998; Sedegah et al., *Proc. Natl. Acad. Sci USA* 95:7648-53, 1998; Hanke and McMichael, *Immunol. Letters* 66:177-181, 1999; and Robinson et at., *Nature Med.* 5:526-34, 1999).

For example, the efficacy of the DNA minigene used in a prime boost protocol is initially evaluated in transgenic mice. In this example, A2.1/K^{b} transgenic mice are immunized IM with 100 μg of a DNA minigene encoding the immunogenic peptides including at least one HLA-A2 supermotif-bearing peptide. After an incubation period (ranging from 3-9 weeks), the mice are boosted IP with 10^{7} pfu/mouse of a recombinant vaccinia virus expressing the same sequence encoded by the DNA minigene. Control mice are immunized with 100 μg of DNA or recombinant vaccinia without the minigene sequence, or with DNA encoding the minigene, but without the vaccinia boost. After an additional incubation period of two weeks, splenocytes from the mice are immediately assayed for peptide-specific activity in an ELISPOT assay. Additionally, splenocytes are stimulated in vitro with the A2-restricted peptide epitopes encoded in the minigene and recombinant vaccinia, then assayed for peptide-specific activity in an alpha, beta and/or gamma IFN ELISA.

It is found that the minigene utilized in a prime-boost protocol elicits greater immune responses toward the HLA-A2 supermotif peptides than with DNA alone. Such an analysis can also be performed using HLA-A11 or HLA-B7 transgenic mouse models to assess CTL induction by HLA-A3 or HLA-B7 motif or supermotif epitopes. The use of prime boost protocols in humans is described below in the Example entitled "Induction of CTL Responses Using a Prime Boost Protocol."

Example 24

Peptide Compositions for Prophylactic Uses

Vaccine compositions of the present invention can be used to prevent 58P1D12 expression in persons who are at risk for tumors that bear this antigen. For example, a polyepitopic peptide epitope composition (or a nucleic acid comprising the same) containing multiple CTL and HTL epitopes such as those selected in the above Examples, which are also selected to target greater than 80% of the population, is administered to individuals at risk for a 58P1D12-associated tumor.

For example, a peptide-based composition is provided as a single polypeptide that encompasses multiple epitopes. The vaccine is typically administered in a physiological solution that comprises an adjuvant, such as Incomplete Freunds Adjuvant. The dose of peptide for the initial immunization is from about 1 to about 50,000 μg, generally 100-5,000 μg, for a 70 kg patient. The initial administration of vaccine is followed by booster dosages at 4 weeks followed by evaluation of the magnitude of the immune response in the patient, by techniques that determine the presence of epitope-specific CTL populations in a PBMC sample. Additional booster doses are administered as required. The composition is found to be both safe and efficacious as a prophylaxis against 58P1D12-associated disease.

Alternatively, a composition typically comprising transfecting agents is used for the administration of a nucleic acid-based vaccine in accordance with methodologies known in the art and disclosed herein.

Example 25

Polyepitopic Vaccine Compositions Derived from Native 58P1D12 Sequences

A native 58P1D12 polyprotein sequence is analyzed, preferably using computer algorithms defined for each class I and/or class II supermotif or motif, to identify "relatively short" regions of the polyprotein that comprise multiple epitopes. The "relatively short" regions are preferably less in length than an entire native antigen. This relatively short sequence that contains multiple distinct or overlapping, "nested" epitopes can be used to generate a minigene construct. The construct is engineered to express the peptide, which corresponds to the native protein sequence. The "relatively short" peptide is generally less than 250 amino acids in length, often less than 100 amino acids in length, preferably less than 75 amino acids in length, and more preferably less than 50 amino acids in length. The protein sequence of the vaccine composition is selected because it has maximal number of epitopes contained within the sequence, i.e., it has a high concentration of epitopes. As noted herein, epitope motifs may be nested or overlapping (i.e., frame shifted relative to one another). For example, with overlapping epitopes, two 9-mer epitopes and one 10-mer epitope can be present in a 10 amino acid peptide. Such a vaccine composition is administered for therapeutic or prophylactic purposes.

The vaccine composition will include, for example, multiple CTL epitopes from 58P1D12 antigen and at least one HTL epitope. This polyepitopic native sequence is administered either as a peptide or as a nucleic acid sequence which encodes the peptide. Alternatively, an analog can be made of this native sequence, whereby one or more of the epitopes comprise substitutions that alter the cross-reactivity and/or binding affinity properties of the polyepitopic peptide.

The embodiment of this example provides for the possibility that an as yet undiscovered aspect of immune system processing will apply to the native nested sequence and thereby facilitate the production of therapeutic or prophylactic immune response-inducing vaccine compositions. Additionally, such an embodiment provides for the possibility of motif-bearing epitopes for an HLA makeup(s) that is presently unknown. Furthermore, this embodiment (excluding an analoged embodiment) directs the immune response to multiple peptide sequences that are actually present in native 58P1D12, thus avoiding the need to evaluate any junctional epitopes. Lastly, the embodiment provides an economy of scale when producing peptide or nucleic acid vaccine compositions.

Related to this embodiment, computer programs are available in the art which can be used to identify in a target sequence, the greatest number of epitopes per sequence length.

Example 26

Polyepitopic Vaccine Compositions from Multiple Antigens

The 58P1D12 peptide epitopes of the present invention are used in conjunction with epitopes from other target tumor-associated antigens, to create a vaccine composition that is useful for the prevention or treatment of cancer that expresses 58P1D12 and such other antigens. For example, a vaccine composition can be provided as a single polypeptide that incorporates multiple epitopes from 58P1D12 as well as tumor-associated antigens that are often expressed with a target cancer associated with 58P1D12 expression, or can be administered as a composition comprising a cocktail of one or more discrete epitopes. Alternatively, the vaccine can be administered as a minigene construct or as dendritic cells which have been loaded with the peptide epitopes in vitro.

Example 27

Use of Peptides to Evaluate an Immune Response

Peptides of the invention may be used to analyze an immune response for the presence of specific antibodies, CTL or HTL directed to 58P1D12. Such an analysis can be performed in a manner described by Ogg et al., *Science* 279: 2103-2106, 1998. In this Example, peptides in accordance with the invention are used as a reagent for diagnostic or prognostic purposes, not as an immunogen.

In this example highly sensitive human leukocyte antigen tetrameric complexes ("tetramers") are used for a cross-sectional analysis of, for example, 58P1D12 HLA-A*0201-specific CTL frequencies from HLA A*0201-positive individuals at different stages of disease or following immunization comprising a 58P1D12 peptide containing an A*0201 motif. Tetrameric complexes are synthesized as described (Musey et al, *N. Engl. J. Med.* 337:1267, 1997). Briefly, purified HLA heavy chain (A*0201 in this example) and β2-microglobulin are synthesized by means of a prokaryotic expression system. The heavy chain is modified by deletion of the transmembrane-cytosolic tail and COOH-terminal addition of a sequence containing a BirA enzymatic biotinylation site. The heavy chain, β2-microglobulin, and peptide are refolded by dilution. The 45-kD refolded product is isolated by fast protein liquid chromatography and then biotinylated by BirA in the presence of biotin (Sigma, St. Louis, Mo.), adenosine 5' triphosphate and magnesium. Streptavidin-phycoerythrin conjugate is added in a 1:4 molar ratio, and the tetrameric product is concentrated to 1 mg/ml. The resulting product is referred to as tetramer-phycoerythrin.

For the analysis of patient blood samples, approximately one million PBMCs are centrifuged at 300 g for 5 minutes and resuspended in 50 μl of cold phosphate-buffered saline. Tri-color analysis is performed with the tetramer-phycoerythrin, along with anti-CD8-Tricolor, and ant-CD38. The PBMCs are incubated with tetramer and antibodies on ice for 30 to 60 min and then washed twice before formaldehyde fixation. Gates are applied to contain >99.98% of control samples. Controls for the tetramers include both A*0201-negative individuals and A*0201-positive non-diseased donors. The percentage of cells stained with the tetramer is then determined by flow cytometry. The results indicate the number of cells in the PBMC sample that contain epitope-restricted CTLs, thereby readily indicating the extent of immune response to the 58P1D12 epitope, and thus the status of exposure to 58P1D12, or exposure to a vaccine that elicits a protective or therapeutic response.

Example 28

Use of Peptide Epitopes to Evaluate Recall Responses

The peptide epitopes of the invention are used as reagents to evaluate T cell responses, such as acute or recall responses, in patients. Such an analysis may be performed on patients who have recovered from 58P1D12-associated disease or who have been vaccinated with a 58P1D12 vaccine.

For example, the class I restricted CTL response of persons who have been vaccinated may be analyzed. The vaccine may be any 58P1D12 vaccine. PBMC are collected from vaccinated individuals and HLA typed. Appropriate peptide epitopes of the invention that, optimally, bear supermotifs to provide cross-reactivity with multiple HLA supertype family members, are then used for analysis of samples derived from individuals who bear that HLA type.

PBMC from vaccinated individuals are separated on Ficoll-Histopaque density gradients (Sigma Chemical Co., St. Louis, Mo.), washed three times in HBSS (GIBCO Laboratories), resuspended in RPMI-1640 (GIBCO Laboratories) supplemented with L-glutamine (2 mM), penicillin (50 U/ml), streptomycin (50 μg/ml), and Hepes (10 mM) containing 10% heat-inactivated human AB serum (complete RPMI) and plated using microculture formats. A synthetic peptide comprising an epitope of the invention is added at 10 μg/ml to each well and HBV core 128-140 epitope is added at 1 μg/ml to each well as a source of T cell help during the first week of stimulation.

In the microculture format, 4×10^5 PBMC are stimulated with peptide in 8 replicate cultures in 96-well round bottom plate in 100 μl/well of complete RPMI. On days 3 and 10, 100 μl of complete RPMI and 20 U/ml final concentration of rIL-2 are added to each well. On day 7 the cultures are transferred into a 96-well flat-bottom plate and restimulated with peptide, rIL-2 and 10^5 irradiated (3,000 rad) autologous feeder cells. The cultures are tested for cytotoxic activity on day 14. A positive CTL response requires two or more of the eight replicate cultures to display greater than 10% specific ^{51}Cr release, based on comparison with non-diseased control subjects as previously described (Rehermann, et al., *Nature Med.* 2:1104, 1108, 1996; Rehermann et al., *J. Clin. Invest.* 97:1655-1665, 1996; and Rehermann et al. *J. Clin. Invest.* 98:1432-1440, 1996).

Target cell lines are autologous and allogeneic EBV-transformed B-LCL that are either purchased from the American Society for Histocompatibility and Immunogenetics (ASHI, Boston, Mass.) or established from the pool of patients as described (Guilhot, et al. *J. Virol.* 66:2670-2678, 1992).

Cytotoxicity assays are performed in the following manner. Target cells consist of either allogeneic HLA-matched or autologous EBV-transformed B lymphoblastoid cell line that are incubated overnight with the synthetic peptide epitope of the invention at 10 μM, and labeled with 100 μCi of ^{51}Cr (Amersham Corp., Arlington Heights, Ill.) for 1 hour after which they are washed four times with HBSS.

Cytolytic activity is determined in a standard 4-h, split well ^{51}Cr release assay using U-bottomed 96 well plates containing 3,000 targets/well. Stimulated PBMC are tested at effector/target (E/T) ratios of 20-50:1 on day 14. Percent cytotoxicity is determined from the formula: 100×[(experimental release-spontaneous release)/maximum release-spontaneous release)]. Maximum release is determined by lysis of targets by detergent (2% Triton X-100; Sigma Chemical Co., St. Louis, Mo.). Spontaneous release is <25% of maximum release for all experiments.

The results of such an analysis indicate the extent to which HLA-restricted CTL populations have been stimulated by previous exposure to 58P1D12 or a 58P1D12 vaccine.

Similarly, Class II restricted HTL responses may also be analyzed. Purified PBMC are cultured in a 96-well flat bottom plate at a density of 1.5×10^5 cells/well and are stimulated with 10 μg/ml synthetic peptide of the invention, whole 58P1D12 antigen, or PHA. Cells are routinely plated in replicates of 4-6 wells for each condition. After seven days of culture, the medium is removed and replaced with fresh medium containing 10 U/ml IL-2. Two days later, 1 μCi ^{3}H-thymidine is added to each well and incubation is continued for an additional 18 hours. Cellular DNA is then harvested on glass fiber mats and analyzed for ^{3}H-thymidine incorporation. Antigen-specific T cell proliferation is calculated as the ratio of ^{3}H-thymidine incorporation in the presence of antigen divided by the ^{3}H-thymidine incorporation in the absence of antigen.

Example 29

Induction of Specific CTL Response in Humans

A human clinical trial for an immunogenic composition comprising CTL and HTL epitopes of the invention is set up as an IND Phase I, dose escalation study and carried out as a randomized, double-blind, placebo-controlled trial. Such a trial is designed, for example, as follows:

A total of about 27 individuals are enrolled and divided into 3 groups:

Group I: 3 subjects are injected with placebo and 6 subjects are injected with 5 μg of peptide composition;

Group II: 3 subjects are injected with placebo and 6 subjects are injected with 50 μg peptide composition;

Group III: 3 subjects are injected with placebo and 6 subjects are injected with 500 μg of peptide composition.

After 4 weeks following the first injection, all subjects receive a booster inoculation at the same dosage.

The endpoints measured in this study relate to the safety and tolerability of the peptide composition as well as its immunogenicity. Cellular immune responses to the peptide composition are an index of the intrinsic activity of this the peptide composition, and can therefore be viewed as a measure of biological efficacy. The following summarize the clinical and laboratory data that relate to safety and efficacy endpoints.

Safety: The incidence of adverse events is monitored in the placebo and drug treatment group and assessed in terms of degree and reversibility.

Evaluation of Vaccine Efficacy: For evaluation of vaccine efficacy, subjects are bled before and after injection. Peripheral blood mononuclear cells are isolated from fresh heparinized blood by Ficoll-Hypaque density gradient centrifugation, aliquoted in freezing media and stored frozen. Samples are assayed for CTL and HTL activity.

The vaccine is found to be both safe and efficacious.

Example 30

Phase II Trials in Patients Expressing 58P1D12

Phase II trials are performed to study the effect of administering the CTL-HTL peptide compositions to patients having cancer that expresses 58P1D12. The main objectives of the trial are to determine an effective dose and regimen for inducing CTLs in cancer patients that express 58P1D12, to establish the safety of inducing a CTL and HTL response in these patients, and to see to what extent activation of CTLs improves the clinical picture of these patients, as manifested, e.g., by the reduction and/or shrinking of lesions. Such a study is designed, for example, as follows:

The studies are performed in multiple centers. The trial design is an open-label, uncontrolled, dose escalation protocol wherein the peptide composition is administered as a single dose followed six weeks later by a single booster shot of the same dose. The dosages are 50, 500 and 5,000 micrograms per injection. Drug-associated adverse effects (severity and reversibility) are recorded.

There are three patient groupings. The first group is injected with 50 micrograms of the peptide composition and the second and third groups with 500 and 5,000 micrograms of peptide composition, respectively. The patients within each group range in age from 21-65 and represent diverse ethnic backgrounds. All of them have a tumor that expresses 58P1D12.

Clinical manifestations or antigen-specific T-cell responses are monitored to assess the effects of administering the peptide compositions. The vaccine composition is found to be both safe and efficacious in the treatment of 58P1D12-associated disease.

Example 31

Induction of CTL Responses Using a Prime Boost Protocol

A prime boost protocol similar in its underlying principle to that used to confirm the efficacy of a DNA vaccine in transgenic mice, such as described above in the Example entitled "The Plasmid Construct and the Degree to Which It Induces Immunogenicity," can also be used for the administration of the vaccine to humans. Such a vaccine regimen can include an initial administration of, for example, naked DNA followed by a boost using recombinant virus encoding the vaccine, or recombinant protein/polypeptide or a peptide mixture administered in an adjuvant.

For example, the initial immunization may be performed using an expression vector, such as that constructed in the Example entitled "Construction of "Minigene" Multi-Epitope DNA Plasmids" in the form of naked nucleic acid administered IM (or SC or ID) in the amounts of 0.5-5 mg at multiple sites. The nucleic acid (0.1 to 1000 μg) can also be administered using a gene gun. Following an incubation period of 3-4 weeks, a booster dose is then administered. The booster can be recombinant fowlpox virus administered at a dose of 5-10^7 to 5×10^9 pfu. An alternative recombinant virus, such as an MVA, canarypox, adenovirus, or adeno-associated virus, can also be used for the booster, or the polyepitopic protein or a mixture of the peptides can be administered. For evaluation of vaccine efficacy, patient blood samples are obtained before immunization as well as at intervals following administration of the initial vaccine and booster doses of the vaccine. Peripheral blood mononuclear cells are isolated from fresh heparinized blood by Ficoll-Hypaque density gradient centrifugation, aliquoted in freezing media and stored frozen. Samples are assayed for CTL and HTL activity.

Analysis of the results indicates that a magnitude of response sufficient to achieve a therapeutic or protective immunity against 58P1D12 is generated.

Example 32

Administration of Vaccine Compositions Using Dendritic Cells (DC)

Vaccines comprising peptide epitopes of the invention can be administered using APCs, or "professional" APCs such as DC. In this example, peptide-pulsed DC are administered to a patient to stimulate a CTL response in vivo. In this method, dendritic cells are isolated, expanded, and pulsed with a vaccine comprising peptide CTL and HTL epitopes of the invention. The dendritic cells are infused back into the patient to elicit CTL and HTL responses in vivo. The induced CTL and HTL then destroy or facilitate destruction, respectively, of the target cells that bear the 58P1D12 protein from which the epitopes in the vaccine are derived.

For example, a cocktail of epitope-comprising peptides is administered ex vivo to PBMC, or isolated DC therefrom. A pharmaceutical to facilitate harvesting of DC can be used, such as Progenipoietin™ (Monsanto, St. Louis, Mo.) or GM-CSF/IL-4. After pulsing the DC with peptides, and prior to reinfusion into patients, the DC are washed to remove unbound peptides.

As appreciated clinically, and readily determined by one of skill based on clinical outcomes, the number of DC reinfused into the patient can vary (see, e.g., *Nature Med.* 4:328, 1998; *Nature Med.* 2:52, 1996 and *Prostate* 32:272, 1997). Although 2-50×10^6 DC per patient are typically administered, larger number of DC, such as 10^7 or 10^8 can also be provided. Such cell populations typically contain between 50-90% DC.

In some embodiments, peptide-loaded PBMC are injected into patients without purification of the DC. For example, PBMC generated after treatment with an agent such as Progenipoietin™ are injected into patients without purification of the DC. The total number of PBMC that are administered often ranges from 10^8 to 10^{10}. Generally, the cell doses injected into patents is based on the percentage of DC in the blood of each patient, as determined, for example, by immunofluorescence analysis with specific anti-DC antibodies. Thus, for example, if Progenipoietin™ mobilizes 2% DC in the peripheral blood of a given patient, and that patient is to receive 5×10^6 DC, then the patent will be injected with a total of 2.5×10^8 peptide-loaded PBMC. The percent DC mobilized by an agent such as Progenipoietin™ is typically estimated to be between 2-10%, but can vary as appreciated by one of skill in the art.

Ex Vivo Activation of CTL/HTL Responses

Alternatively, ex vivo CTL or HTL responses to 58P1D12 antigens can be induced by incubating, in tissue culture, the patient's, or genetically compatible, CTL or HTL precursor cells together with a source of APC, such as DC, and immunogenic peptides. After an appropriate incubation time (typically about 7-28 days), in which the precursor cells are activated and expanded into effector cells, the cells are infused into the patient, where they will destroy (CTL) or facilitate destruction (HTL) of their specific target cells, i.e., tumor cells.

Example 33

An Alternative Method of Identifying and Confirming Motif-Bearing Peptides

Another method of identifying and confirming motif-bearing peptides is to elute them from cells bearing defined MHC molecules. For example, EBV transformed B cell lines used for tissue typing have been extensively characterized to determine which HLA molecules they express. In certain cases these cells express only a single type of HLA molecule. These cells can be transfected with nucleic acids that express the antigen of interest, e.g. 58P1D12. Peptides produced by endogenous antigen processing of peptides produced as a result of transfection will then bind to HLA molecules within the cell and be transported and displayed on the cell's surface. Peptides are then eluted from the HLA molecules by exposure to mild acid conditions and their amino acid sequence determined, e.g., by mass spectral analysis (e.g., Kubo et al, *J. Immunol.* 152:3913, 1994). Because the majority of peptides that bind a particular HLA molecule are motif-bearing, this is an alternative modality for obtaining the motif-bearing peptides correlated with the particular HLA molecule expressed on the cell.

Alternatively, cell lines that do not express endogenous HLA molecules can be transfected with an expression construct encoding a single HLA allele. These cells can then be used as described, i.e., they can then be transfected with nucleic acids that encode 58P1D12 to isolate peptides corresponding to 58P1D12 that have been presented on the cell surface. Peptides obtained from such an analysis will bear motif(s) that correspond to binding to the single HLA allele that is expressed in the cell.

As appreciated by one in the art, one can perform a similar analysis on a cell bearing more than one HLA allele and subsequently determine peptides specific for each HLA allele expressed. Moreover, one of skill would also recognize that means other than transfection, such as loading with a protein antigen, can be used to provide a source of antigen to the cell.

Example 34

Complementary Polynucleotides

Sequences complementary to the 58P1D12-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring 58P1D12. Although use of oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are designed using, e.g., OLIGO 4.06 software (National Biosciences) and the coding sequence of 58P1D12. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding to a 58P1D12-encoding transcript.

Example 35

Purification of Naturally-Occurring or Recombinant 58P1D12 Using 58P1D12-Specific Antibodies

Naturally occurring or recombinant 58P1D12 is substantially purified by immunoaffinity chromatography using antibodies specific for 58P1D12. An immunoaffinity column is constructed by covalently coupling anti-58P1D12 antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturers instructions.

Media containing 58P1D12 are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of 58P1D12 (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/58P1D12 binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and GCR.P is collected.

Example 36

Identification of Molecules which Interact with 58P1D12

58P1D12, or biologically active fragments thereof, are labeled with 121 1 Bolton-Hunter reagent. (See, e.g., Bolton et al. (1973) Biochem. J. 133:529.) Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled 58P1D12, washed, and any wells with labeled 58P1D12 complex are assayed. Data obtained using different concentrations of 58P1D12 are used to calculate values for the number, affinity, and association of 58P1D12 with the candidate molecules.

Example 37

In Vivo Assay for 58P1D12 Tumor Growth Promotion

In Vivo Assay of 3T3 Cell Growth by Recombinant Expression of 58P1D12.

To address the potential for 58P1D12 to accelerate the growth of non-tumorigenic cells in an in vivo mouse model, non-transformed 3T3 cells were prepared by infection with either a virus containing an empty vector control (Neo gene alone) or with a vector containing the 58P1D12 full-length gene. 3T3 cells were selected for survival in G-418, and expression of 58P1D12 confirmed by Northern blot analysis. To assess the growth potential of these cells, 1×10^6 58P1D12 expressing 3T3 cells or 1×10^6 Neo control were mixed with Matrigel®, then injected intratibially in SCID mice and allowed to grow for 31 days. The growth of these cells was assessed on day 31 by visual inspection and by necropsy (FIG. **28**). Clearly, the 58P1D12 expressing 3T3 cells showed a potent effect in comparison to the 3T3-Neo cells, indicating that the 58P1D12 protein enhanced the growth of the cells in Matrigel®. These data indicate that 58P1D12 promotes the growth of non-tumorigenic cells and provides a growth advantage in vivo that mimics the role of this protein in human malignancies.

Example 38

58P1D12 Monoclonal Antibody-Mediated Inhibition of Bladder, Prostate, and Ovarian Tumors In Vivo

The significant expression of 58P1D12 in cancer tissues, together with its restrictive expression in normal tissues makes 58P1D12 a good target for antibody therapy. Similarly, 58P1D12 is a target for T cell-based immunotherapy. Thus, the therapeutic efficacy of anti-58P1D12 MAbs in human prostate cancer xenograft mouse models is evaluated by using recombinant cell lines such as PC3-58P1D12, and 3T3-58P1D12 (see, e.g., Kaighn, M. E., et al, Invest Urol, 1979. 17(1): p. 16-23), as well as human prostate xenograft models such as LAPC9 (Saffran et al, Proc Natl Acad Sci USA. 2001, 98:2658). Similarly, anti-58P1D12 MAbs are evaluated in human bladder cancer xenograft models using recombinant cell lines such as J82-58P1D12, kidney cancer xenograft models using 769-P cells, or ovarian cancer xenograft models using recombinant cell lines such as OVCAR-58P1D12.

Antibody efficacy on tumor growth and metastasis formation is studied, e.g., in a mouse orthotopic bladder cancer xenograft model, and a mouse prostate cancer xenograft model. The antibodies can be unconjugated, as discussed in this Example, or can be conjugated to a therapeutic modality (see below), as appreciated in the art. Anti-58P1D12 MAbs inhibit formation of prostate and bladder xenografts. Anti-58P1D12 MAbs also retard the growth of established orthotopic tumors and prolonged survival of tumor-bearing mice. These results indicate the utility of anti-58P1D12 MAbs in the treatment of local and advanced stages of prostate and bladder cancer (see, e.g., Saffran, D., 2001, et al., PNAS 10:1073-1078 or on the World Wide Web at www.pnas.org/cgi/doi/10.1073/pnas.051624698).

Administration of the anti-58P1D12 MAbs led to retardation of established orthotopic tumor growth and inhibition of metastasis to distant sites, resulting in a significant prolongation in the survival of tumor-bearing mice. These studies indicate that 58P1D12 is an attractive target for immunotherapy and demonstrate the therapeutic potential of anti-58P1D12 MAbs for the treatment of local and metastatic cancer. This example demonstrates that unconjugated 58P1D12 monoclonal antibodies are effective to inhibit the growth of human bladder and prostate tumor xenografts grown in SCID mice; accordingly, a combination of such efficacious MAbs is also effective.

MAb-toxin conjugates: Another embodiment of MAb therapy is through the use of toxin conjugation of MAbs for targeted delivery of cytotoxic agents to cells expressing the protein target. Major advances have been made in the clinical application of MAb toxin conjugates with the development of Mylotarg for acute myeloid leukemia (Bross, P. F., et al., 2001, Clin. Cancer Res. 7:1490-1496). Mylotarg is a humanized MAb directed to CD33 which is conjugated to a highly potent DNA-alkylating agent (calichemicin) via an acid labile hydrazone bond (Hamann, P. R., et al., 2002, Bioconjug. Chem. 13:40-46; ibid., 13:47-58). Additional toxins for MAb conjugation in development include maytansinoid, doxorubicin, taxoids and the potent synthetic dolastatin 10 analogs auristatin E and monomethylauristatin E (Doronina, S. O., et al., 2003, Nature Biotech. 21:778-784; Ross, S., et al., 2002, Cancer Res. 62:2546-2553; Francisco, J. A., et el., 2003, Blood 102:1458-1465; Mao, W., et al., 2004, Cancer Res. 64:781-788). Such applications have potential to deliver a cytotoxic agent to cells expressing the protein target of the MAb. Internalization of the target protein upon MAb binding is important for toxin delivery, and the mechanism spares the non-targeted tissues from the potentially harmful effects of the cytotoxic agent.

58P1D12 MAbs conjugated to toxins are used to induce cell killing in vitro using established protocols for cytotoxicity assays and clonogenic assays (Doronina, S. O., et al., 2003, Nature Biotech. 21:778-784; Mao, W., et al., 2004, Cancer Res. 64:781-788). Toxin conjugated anti-58P1D12 MAbs induce cytotoxicity of cells expressing endogenous 58P1D12 (769-P cells) and recombinant 58P1D12 (PC3-58P1D12, 3T3-58P1D12, Rat-1-58P1D12 and B300.19-

58P1D12). This methodology allows confirmation that the toxin conjugated MAb is functional against cells expressing the 58P1D12 protein on their surface versus those that do not express the target.

The MAb toxin conjugates are tested for their ability to inhibit tumor growth in vivo. Antibody efficacy on tumor growth and metastasis formation is studied, e.g., in a mouse orthotopic bladder cancer xenograft model, a mouse prostate cancer xenograft model, or mouse ovarian cancer xenograft model. Administration of the anti-58P1D12 MAbs led to retardation of established orthotopic tumor growth and inhibition of metastasis to distant sites, resulting in a significant prolongation in the survival of tumor-bearing mice. These studies indicate that 58P1D12 is an attractive target for immunotherapy and demonstrate the therapeutic potential of toxin-conjugated anti-58P1D12 MAbs for the treatment of local and metastatic cancer. This example demonstrates that toxin-conjugated 58P1D12 monoclonal antibodies are effective to inhibit the growth of human bladder, prostate and ovarian tumor xenografts grown in SCID mice; accordingly, a combination of such efficacious MAbs is also effective. The methodology allows the targeted delivery of a cytotoxin using a plasma stable linker in a MAb-toxin conjugate. Such a mechanism of action reduces the potential harmful effects of the toxin on non-targeted tissues.

Tumor Inhibition Using Multiple Unconjugated or Toxin-Conjugated 58P1D12 MAbs

Materials and Methods

58P1D12 Monoclonal Antibodies:

Monoclonal antibodies are raised against 58P1D12 as described in the Example entitled "Generation of 58P1D12 Monoclonal Antibodies (MAbs)." The antibodies are characterized by ELISA, Western blot, FACS, and immunoprecipitation for their capacity to bind 58P1D12. Epitope mapping data for the anti-58P1D12 MAbs, as determined by ELISA and Western analysis, recognize epitopes on the 58P1D12 protein. Immunohistochemical analysis of prostate cancer tissues and cells with these antibodies is performed.

The monoclonal antibodies are purified from ascites or hybridoma tissue culture supernatants by Protein-G Sepharose chromatography, dialyzed against PBS, filter sterilized, and stored at −20° C. Protein determinations are performed by a Bradford assay (Bio-Rad, Hercules, Calif.). A therapeutic monoclonal antibody or a cocktail comprising a mixture of individual monoclonal antibodies is prepared and used for the treatment of mice receiving subcutaneous or orthotopic injections of SCABER, J82, A498, 769-P, CaOv1, OVCAR or PA1 tumor xenografts.

The MAbs to 58P1D12 are conjugated to various different toxins (listed above) using any of a variety of methods described elsewhere in the art (Hamann, P. R., et al., 2002, Bioconjug. Chem. 13:40-46; ibid., 13:47-58; Doronina, S. O., et al., 2003, Nature Biotech. 21:778-784; Ojima, I., et al. 2002, J. Med. Chem. 45:5620-5623; Dubowchik, G. M., et al., 2002, Bioconjug. Chem. 13:855-869; King, H. D., 2002, J. Med. Chem 45:4336-4343; Ross, S., et al., 2002, Cancer Res. 62:2546-2553; Francisco, J. A., et el., 2003, Blood 102: 1458-1465; Mao, W., et al., 2004, Cancer Res. 64:781-788).

Cell Lines

The bladder and prostate carcinoma cell lines, J82 and PC3 as well as the fibroblast line NIH 3T3 (American Type Culture Collection) are maintained in media supplemented with L-glutamine and 10% FBS. J82-58P1D12, PC3-58P1D12 and 3T3-58P1D12 cell populations are generated by retroviral gene transfer as described in Hubert, R. S., et al., Proc. Natl. Acad. Sci. USA, 1999, 96(25):14523.

Xenograft Mouse Models

Subcutaneous (s.c.) tumors are generated by injection of 1×10^6 cancer cells mixed at a 1:1 dilution with Matrigel® (Collaborative Research) in the right flank of male SCID mice. To test antibody efficacy on tumor formation, i.p. antibody injections are started on the same day as tumor-cell injections. As a control, mice are injected with either purified mouse IgG (ICN) or PBS; or a purified monoclonal antibody that recognizes an irrelevant protein not expressed in human cells. Tumor sizes are determined by caliper measurements, and the tumor volume is calculated as: Length×Width×Height. Mice with s.c. tumors greater than 1.5 cm in diameter are sacrificed.

Orthotopic injections are performed under anesthesia by using ketamine/xylazine. For bladder orthotopic studies, an incision is made through the abdomen to expose the bladder, and tumor cells (5×10^5) mixed with Matrigel® are injected into the bladder wall in a 10-μl volume. To monitor tumor growth, mice are palpated and blood is collected on a weekly basis to measure BTA levels. For prostate orthopotic models, an incision is made through the abdominal muscles to expose the bladder and seminal vesicles, which then are delivered through the incision to expose the dorsal prostate. Tumor cells, e.g. LAPC-9 cells (5×10^5) mixed with Matrigel® are injected into the prostate in a 10-μl volume (Yoshida Y et al, Anticancer Res. 1998, 18:327; Ahn et al, Tumour Biol. 2001, 22:146). To monitor tumor growth, blood is collected on a weekly basis measuring PSA levels. The mice are segregated into groups for the appropriate treatments, with anti-58P1D12 or control MAbs being injected i.p.

Anti-58P1D12 MAbs Inhibit Growth of 58P1D12-Expressing Xenograft-Cancer Tumors

The effect of anti-58P1D12 MAbs on tumor formation is tested on the growth and progression of bladder, and prostate cancer xenografts using J82-58P1D12, and PC3-58P1D12 orthotopic models. As compared with the s.c. tumor model, the orthotopic model, which requires injection of tumor cells directly in the mouse bladder, and prostate, respectively, results in a local tumor growth, development of metastasis in distal sites, deterioration of mouse health, and subsequent death (Saffran, D., et al., PNAS supra; Fu, X., et al., Int J Cancer, 1992. 52(6): p. 987-90; Kubota, T., J Cell Biochem, 1994. 56(1): p. 4-8). The features make the orthotopic model more representative of human disease progression and allowed us to follow the therapeutic effect of MAbs on clinically relevant end points.

Accordingly, tumor cells are injected into the mouse bladder, or prostate, and 2 days later, the mice are segregated into two groups and treated with either: a) 200-500 μg, of anti-58P1D12 Ab, or b) PBS three times per week for two to five weeks.

A major advantage of the orthotopic cancer models is the ability to study the development of metastases. Formation of metastasis in mice bearing established orthotopic tumors is studies by IHC analysis on lung sections using an antibody against a tumor-specific cell-surface protein such as anti-CK20 for bladder cancer, anti-STEAP-1 for prostate cancer models (Lin S et al, Cancer Detect Prev. 2001; 25:202; Saffran, D., et al., PNAS supra).

Mice bearing established orthotopic tumors are administered 1000 μg injections of either anti-58P1D12 MAb or PBS over a 4-week period. Mice in both groups are allowed to establish a high tumor burden, to ensure a high frequency of metastasis formation in mouse lungs. Mice then are killed and their bladders, livers, bone and lungs are analyzed for the presence of tumor cells by IHC analysis.

These studies demonstrate a broad anti-tumor efficacy of anti-58P1D12 antibodies on initiation and progression of prostate, kidney, bladder and ovarian cancer in xenograft mouse models. Anti-58P1D12 antibodies inhibit formation of tumors as well as retarding the growth of already established tumors and prolong the survival of treated mice. Moreover, anti-58P1D12 MAbs demonstrate a dramatic inhibitory effect on the spread of local bladder and prostate tumor to distal sites, even in the presence of a large tumor burden. Thus, anti-58P1D12 MAbs are efficacious on major clinically relevant end points (tumor growth), prolongation of survival, and health.

Example 39

Therapeutic and Diagnostic Use of Anti-58P1D12 Antibodies in Humans

Anti-58P1D12 monoclonal antibodies are safely and effectively used for diagnostic, prophylactic, prognostic and/or therapeutic purposes in humans. Western blot and immunohistochemical analysis of cancer tissues and cancer xenografts with anti-58P1D12 mAb show strong extensive staining in carcinoma but significantly lower or undetectable levels in normal tissues. Detection of 58P1D12 in carcinoma and in metastatic disease demonstrates the usefulness of the mAb as a diagnostic and/or prognostic indicator. Anti-58P1D12 antibodies are therefore used in diagnostic applications such as immunohistochemistry of kidney biopsy specimens to detect cancer from suspect patients.

As determined by flow cytometry, anti-58P1D12 mAb specifically binds to carcinoma cells. Thus, anti-58P1D12 antibodies are used in diagnostic whole body imaging applications, such as radioimmunoscintigraphy and radioimmunotherapy, (see, e.g., Potamianos S., et. al. Anticancer Res 20(2A):925-948 (2000)) for the detection of localized and metastatic cancers that exhibit expression of 58P1D12. Shedding or release of an extracellular domain of 58P1D12 into the extracellular milieu, such as that seen for alkaline phosphodiesterase B10 (Meerson, N. R., Hepatology 27:563-568 (1998)), allows diagnostic detection of 58P1D12 by anti-58P1D12 antibodies in serum and/or urine samples from suspect patients.

Anti-58P1D12 antibodies that specifically bind 58P1D12 are used in therapeutic applications for the treatment of cancers that express 58P1D12. Anti-58P1D12 antibodies are used as an unconjugated modality and as conjugated form in which the antibodies are attached to one of various therapeutic or imaging modalities well known in the art, such as a prodrugs, enzymes or radioisotopes. In preclinical studies, unconjugated and conjugated anti-58P1D12 antibodies are tested for efficacy of tumor prevention and growth inhibition in the SCID mouse cancer xenograft models, e.g., kidney cancer models AGS-K3 and AGS-K6, (see, e.g., the Example entitled "58P1D12 Monoclonal Antibody-mediated Inhibition of Bladder and Lung Tumors In Vivo"). Either conjugated and unconjugated anti-58P1D12 antibodies are used as a therapeutic modality in human clinical trials either alone or in combination with other treatments as described in following Examples.

Example 40

Human Clinical Trials for the Treatment and Diagnosis of Human Carcinomas through Use of Human Anti-58P1D12 Antibodies In vivo

Antibodies are used in accordance with the present invention which recognize an epitope on 58P1D12, and are used in the treatment of certain tumors such as those listed in Table I. Based upon a number of factors, including 58P1D12 expression levels, tumors such as those listed in Table I are presently preferred indications. In connection with each of these indications, three clinical approaches are successfully pursued.

I.) Adjunctive therapy: In adjunctive therapy, patients are treated with anti-58P1D12 antibodies in combination with a chemotherapeutic or antineoplastic agent and/or radiation therapy. Primary cancer targets, such as those listed in Table I, are treated under standard protocols by the addition anti-58P1D12 antibodies to standard first and second line therapy. Protocol designs address effectiveness as assessed by reduction in tumor mass as well as the ability to reduce usual doses of standard chemotherapy. These dosage reductions allow additional and/or prolonged therapy by reducing dose-related toxicity of the chemotherapeutic agent. Anti-58P1D12 antibodies are utilized in several adjunctive clinical trials in combination with the chemotherapeutic or antineoplastic agents adriamycin (advanced prostrate carcinoma), cisplatin (advanced head and neck and lung carcinomas), taxol (breast cancer), and doxorubicin (preclinical).

II.) Monotherapy: In connection with the use of the anti-58P1D12 antibodies in monotherapy of tumors, the antibodies are administered to patients without a chemotherapeutic or antineoplastic agent. In one embodiment, monotherapy is conducted clinically in end stage cancer patients with extensive metastatic disease. Patients show some disease stabilization. Trials demonstrate an effect in refractory patients with cancerous tumors.

III.) Imaging Agent: Through binding a radionuclide (e.g., iodine or yttrium (I^{131}, Y^{90}) to anti-58P1D12 antibodies, the radiolabeled antibodies are utilized as a diagnostic and/or imaging agent. In such a role, the labeled antibodies localize to both solid tumors, as well as, metastatic lesions of cells expressing 58P1D12. In connection with the use of the anti-58P1D12 antibodies as imaging agents, the antibodies are used as an adjunct to surgical treatment of solid tumors, as both a pre-surgical screen as well as a post-operative follow-up to determine what tumor remains and/or returns. In one embodiment, a (^{111}In)-58P1D12 antibody is used as an imaging agent in a Phase I human clinical trial in patients having a carcinoma that expresses 58P1D12 (by analogy see, e.g., Divgi et al. *J. Natl. Cancer Inst.* 83:97-104 (1991)). Patients are followed with standard anterior and posterior gamma camera. The results indicate that primary lesions and metastatic lesions are identified.

Dose and Route of Administration

As appreciated by those of ordinary skill in the art, dosing considerations can be determined through comparison with the analogous products that are in the clinic. Thus, anti-58P1D12 antibodies can be administered with doses in the range of 5 to 400 mg/m^2, with the lower doses used, e.g., in connection with safety studies. The affinity of anti-58P1D12 antibodies relative to the affinity of a known antibody for its target is one parameter used by those of skill in the art for determining analogous dose regimens. Further, anti-58P1D12 antibodies that are fully human antibodies, as compared to the chimeric antibody, have slower clearance; accordingly, dosing in patients with such fully human anti-58P1D12 antibodies can be lower, perhaps in the range of 50 to 300 mg/m^2, and still remain efficacious. Dosing in mg/m^2, as opposed to the conventional measurement of dose in mg/kg, is a measurement based on surface area and is a convenient dosing measurement that is designed to include patients of all sizes from infants to adults.

Three distinct delivery approaches are useful for delivery of anti-58P1D12 antibodies. Conventional intravenous delivery is one standard delivery technique for many tumors. However, in connection with tumors in the peritoneal cavity, such as tumors of the ovaries, biliary duct, other ducts, and the like, intraperitoneal administration may prove favorable for obtaining high dose of antibody at the tumor and to also minimize antibody clearance. In a similar manner, certain solid tumors possess vasculature that is appropriate for regional perfusion. Regional perfusion allows for a high dose of antibody at the site of a tumor and minimizes short term clearance of the antibody.

Clinical Development Plan (CDP)

Overview: The CDP follows and develops treatments of anti-58P1D12 antibodies in connection with adjunctive therapy, monotherapy, and as an imaging agent. Trials initially demonstrate safety and thereafter confirm efficacy in repeat doses. Trails are open label comparing standard chemotherapy with standard therapy plus anti-58P1D12 antibodies. As will be appreciated, one criteria that can be utilized in connection with enrollment of patients is 58P1D12 expression levels in their tumors as determined by biopsy.

As with any protein or antibody infusion-based therapeutic, safety concerns are related primarily to (i) cytokine release syndrome, i.e., hypotension, fever, shaking, chills; (ii) the development of an immunogenic response to the material (i.e., development of human antibodies by the patient to the antibody therapeutic, or HAHA response); and, (iii) toxicity to normal cells that express 58P1D12. Standard tests and follow-up are utilized to monitor each of these safety concerns. Anti-58P1D12 antibodies are found to be safe upon human administration.

Example 41

Human Clinical Trial Adjunctive Therapy with Human Anti-58P1D12 Antibody and Chemotherapeutic Agent

A phase I human clinical trial is initiated to assess the safety of six intravenous doses of a human anti-58P1D12 antibody in connection with the treatment of a solid tumor, e.g., a cancer of a tissue listed in Table I. In the study, the safety of single doses of anti-58P1D12 antibodies when utilized as an adjunctive therapy to an antineoplastic or chemotherapeutic agent as defined herein, such as, without limitation: cisplatin, topotecan, doxorubicin, adriamycin, taxol, or the like, is assessed. The trial design includes delivery of six single doses of an anti-58P1D12 antibody with dosage of antibody escalating from approximately about 25 mg/m^2 to about 275 mg/m^2 over the course of the treatment in accordance with the following schedule:

	Day 0	Day 7	Day 14	Day 21	Day 28	Day 35
mAb Dose	25 mg/m^2	75 mg/m^2	125 mg/m^2	175 mg/m^2	225 mg/m^2	275 mg/m^2
Chemotherapy (standard dose)	+	+	+	+	+	+

Patients are closely followed for one-week following each administration of antibody and chemotherapy. In particular, patients are assessed for the safety concerns mentioned above: (i) cytokine release syndrome, i.e., hypotension, fever, shaking, chills; (ii) the development of an immunogenic response to the material (i.e., development of human antibodies by the patient to the human antibody therapeutic, or HAHA response); and, (iii) toxicity to normal cells that express 58P1D12. Standard tests and follow-up are utilized to monitor each of these safety concerns. Patients are also assessed for clinical outcome, and particularly reduction in tumor mass as evidenced by MRI or other imaging.

The anti-58P1D12 antibodies are demonstrated to be safe and efficacious, Phase II trials confirm the efficacy and refine optimum dosing.

Example 42

Human Clinical Trial: Monotherapy with Human Anti-58P1D12 Antibody

Anti-58P1D12 antibodies are safe in connection with the above-discussed adjunctive trial, a Phase II human clinical trial confirms the efficacy and optimum dosing for monotherapy. Such trial is accomplished, and entails the same safety and outcome analyses, to the above-described adjunctive trial with the exception being that patients do not receive chemotherapy concurrently with the receipt of doses of anti-58P1D12 antibodies.

Example 43

Human Clinical Trial: Diagnostic Imaging with Anti-58P1D12 Antibody

Once again, as the adjunctive therapy discussed above is safe within the safety criteria discussed above, a human clinical trial is conducted concerning the use of anti-58P1D12 antibodies as a diagnostic imaging agent. The protocol is designed in a substantially similar manner to those described in the art, such as in Divgi et al. *J. Natl. Cancer Inst.* 83:97-104 (1991). The antibodies are found to be both safe and efficacious when used as a diagnostic modality.

Example 44

I. Enhanced Angiogenesis (Tube Formation) of Human Umbilical Vein Endothelial Cells by Recombinant Extracellular Domain (ECD) of 58P1D12

Angiogenesis is a critical event in the maintenance of tumors for their growth, nutrition and oxygenation. Growth and migration of endothelial cells is a hallmark of angiogenesis, and the stimulation of such cells requires multiple serum growth factors including vascular endothelial growth factor (VEGF). To address whether 58P1D12 facilitates angiogenesis of primary endothelium, an assay was established to detect the effect of the extracellular domain (ECD) (amino acids x-y) of 58P1D12 on the formation of endothelial matrices (tubes) when plated on Matrigel®. Human umbilical vein endothelial cells (HUVEC) or human lung microvascular endothelial cells (HMVEC) were grown in 10% FBS in media to 70% confluency (passage 2-passage 6). Cells were detached in 1:1 trypsin:PBS or 10 mM EDTA, washed and resuspended in EBM-0.1% FBS. Recombinant 58P1D12 ECD (described in Example entitled "Production of Recombinant 58P1D12 in Eukaryotic Systems") was added to the cells and then incubated on 200 μl Matrigel® for 8-16 hours at 37° C. Tube formation was quantitated and captured by photography. FIG. **23** shows that the 58P1D12 ECD induced the formation of endothelial tubes similarly as generated by treatment with purified VEGF protein. FIG. **24** shows that similar results were obtained when treating mature lung-derived microvascular endothelial cells (HMVEC) with 58P1D12 ECD. Together, these results indicate that the 58P1D12 protein is involved in promoting angiogenesis that facilitates tumor growth, activation of endothelium for tumor vascularization or tumor cell invasion. FIG. **25** shows inhibition of endothelial tubes induced by 58P1D12 ECD with monoclonal antibody M8-143(5)10 at 30 μg/ml. An isotype control monoclonal antibody at the same concentration did not inhibit tube formation.

11. Enhanced Survival and Proliferation of Endothelial Cells by Recombinant Extracellular Domain (ECD) of 58P1D12

Enhanced proliferation, entry into S-phase and survival of endothelial cells is a hallmark of both angiogenesis and the cancer cell phenotype. To address the effect of the 58P1D12 protein on the survival and proliferation of normal endothelial cells, both HUVEC and HMVEC were treated with the 58P1D12 ECD and assayed in low serum conditions. The cells were grown overnight in 0.5% FBS and then compared to cells treated with VEGF or 58P1D12 ECD. The cells were evaluated for survival by the Alamar blue assay. The results in FIG. **26** indicate that HUVEC cell viability was increased by the endothelial growth factor VEGF (positive control) as well as increasing concentrations of recombinant 58P1D12 ECD. Further, the results in FIG. **27** show that proliferation of HMVEC in response to 58P1D12 ECD was enhanced at 72 hr post-treatment measured by a ^{3}H-thymidine incorporation assay. These results confirm those measured by Alamar blue assay, and indicate that cells that are treated with 58P1D12 protein have an enhanced proliferative capacity and survive in low serum conditions. Accordingly, 58P1D12 expressing cells induce increased potential for growth of endothelium in vivo.

Together, these data indicate that the 58P1D12 protein has growth promoting, survival potentiating and angiogenic activities for primary endothelium. Such activities are vital for the establishment and maintenance of tumors by stimulating capillary bed formation by endothelial cells. This increased angiogenesis provides a means of increased nutrition and oxygenation of growing tumors, and affords a therapeutic intervention strategy for targeting of 58P1D12 in malignancies. The 58P1D12 protein also plays a role in cell migration, particularly for cells that express the protein, and also serves as an attractant for endothelial cells and other cells expressing its ligand. Cell migration is important for the metastatic potential of tumor cells, and facilitates their homing to remote tissues. Expression of 58P1D12 protein in bone metastases indicates that the migratory phenotype of 58P1D12-expressing cells imparts a strong potential for such cells to metastasize.

Example 45

58P1D12 RNA Interference (RNAi)

RNA interference (RNAi) technology is implemented to a variety of cell assays relevant to oncology. RNAi is a post-transcriptional gene silencing mechanism activated by double-stranded RNA (dsRNA). RNAi induces specific mRNA degradation leading to changes in protein expression and subsequently in gene function. In mammalian cells, these dsRNAs called short interfering RNA (siRNA) have the correct composition to activate the RNAi pathway targeting for degradation, specifically some mRNAs. See, Elbashir S. M., et. al., *Duplexes of* 21-*nucleotide RNAs Mediate RNA interference in Cultured Mammalian Cells*, Nature 411 (6836): 494-8 (2001). Thus, RNAi technology is used successfully in mammalian cells to silence targeted genes.

Loss of cell proliferation control is a hallmark of cancerous cells; thus, assessing the role of 58P1D12 in cell survival/proliferation assays is relevant. Accordingly, RNAi was used to investigate the function of the 58P1D12 antigen. To generate siRNA for 58P1D12, algorithms were used that predict oligonucleotides that exhibit the critical molecular parameters (G:C content, melting temperature, etc.) and have the ability to significantly reduce the expression levels of the 58P1D12 protein when introduced into cells. In accordance with this Example, 58P1D12 siRNA compositions are used that comprise siRNA (double stranded, short interfering RNA) that correspond to the nucleic acid ORF sequence of the 58P1D12 protein or subsequences thereof. Thus, siRNA subsequences are used in this manner are generally 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35 or more than 35 contiguous RNA nucleotides in length. These siRNA sequences are complementary and non-complementary to at least a portion of the mRNA coding sequence. In a preferred embodiment, the subsequences are 19-25 nucleotides in length, most preferably 21-23 nucleotides in length. In preferred embodiments, these siRNA achieve knockdown of 58P1D12 antigen in cells expressing the protein and have functional effects as described below.

Mammalian siRNA transfections: The day before siRNA transfection, the different cell lines are plated in media (RPMI 1640 with 10% FBS w/o antibiotics) at 2×10^3 cells/well in 80μ (96 well plate format) for the proliferation assay. In parallel with the 58P1D12 specific siRNA oligo, the following sequences are included in every experiment as controls: a) Mock transfected cells with Lipofectamine 2000 (Invitrogen, Carlsbad, Calif.) and annealing buffer (no siRNA); b) Luciferase-4 specific siRNA (targeted sequence: 5'-AAGGGACGAAGACGAACACUUCTT-3') (SEQ ID NO: 35); and, c) Eg5 specific siRNA (targeted sequence: 5'-AACTGAAGACCTGAAGACAATAA-3') (SEQ ID NO: 36). SiRNAs are used at 10 nM and μg/ml Lipofectamine 2000 final concentration.

The procedure is as follows: The siRNAs are first diluted in OPTIMEM (serum-free transfection media, Invitrogen) at 0.1 μM (10-fold concentrated) and incubated 5-10 min RT. Lipofectamine 2000 was diluted at 10 μg/ml (10-fold concentrated) for the total number transfections and incubated 5-10 minutes at room temperature (RT). Appropriate amounts of diluted 10-fold concentrated Lipofectamine 2000 were mixed 1:1 with diluted 10-fold concentrated siRNA and incubated at

RT for 20-30" (5-fold concentrated transfection solution). 20 μls of the 5-fold concentrated transfection solutions were added to the respective samples and incubated at 37° C. for 96 hours before analysis.

^{3}H-Thymidine incorporation assay: The proliferation assay is a ^{3}H-thymidine incorporation method for determining the proliferation of viable cells by uptake and incorporation of label into DNA.

The procedure was as follows: Cells growing in log phase are trypsinized, washed, counted and plated in 96-well plates at 1000-4000 cells/well in 10% FBS. After 4-8 hrs, the media is replaced. The cells are incubated for 24-72 hrs, pulsed with ^{3}H-Thy at 1.5 μCi/ml for 14 hrs, harvested onto a filtermat and counted in scintillation cocktail on a Microbeta trilux or other counter.

These data indicate that 58P1D12 plays an important role in the proliferation of cancer cells and that the lack of 58P1D12 clearly decreases the survival potential of these cells. It is to be noted that 58P1D12 is constitutively expressed in many tumor cell lines. 58P1D12 serves a role in malignancy; its expression is a primary indicator of disease, where such disease is often characterized by high rates of uncontrolled cell proliferation and diminished apoptosis. Correlating cellular phenotype with gene knockdown following RNAi treatments is important, and allows one to draw valid conclusions and rule out toxicity or other non-specific effects of these reagents. To this end, assays to measure the levels of expression of both protein and mRNA for the target after RNAi treatments are important, including Western blotting, FACS staining with antibody, immunoprecipitation, Northern blotting or RT-PCR (Taqman or standard methods). Any phenotypic effect of the siRNAs in these assays should be correlated with the protein and/or mRNA knockdown levels in the same cell lines. 58P1D12 protein is reduced after treatment with siRNA oligos described above (e.g., 58P1D12.a, etc.).

A method to analyze 58P1D12 related cell proliferation is the measurement of DNA synthesis as a marker for proliferation. Labeled DNA precursors (i.e. ^{3}H-Thymidine) are used and their incorporation to DNA is quantified. Incorporation of the labeled precursor into DNA is directly proportional to the amount of cell division occurring in the culture. Another method used to measure cell proliferation is performing clonogenic assays. In these assays, a defined number of cells are plated onto the appropriate matrix and the number of colonies formed after a period of growth following siRNA treatment is counted.

In 58P1D12 cancer target validation, complementing the cell survival/proliferation analysis with apoptosis and cell cycle profiling studies are considered. The biochemical hallmark of the apoptotic process is genomic DNA fragmentation, an irreversible event that commits the cell to die. A method to observe fragmented DNA in cells is the immunological detection of histone-complexed DNA fragments by an immunoassay (i.e. cell death detection ELISA) which measures the enrichment of histone-complexed DNA fragments (mono- and oligo-nucleosomes) in the cytoplasm of apoptotic cells. This assay does not require pre-labeling of the cells and can detect DNA degradation in cells that do not proliferate in vitro (i.e. freshly isolated tumor cells).

The most important effector molecules for triggering apoptotic cell death are caspases. Caspases are proteases that when activated cleave numerous substrates at the carboxy-terminal site of an aspartate residue mediating very early stages of apoptosis upon activation. All caspases are synthesized as pro-enzymes and activation involves cleavage at aspartate residues. In particular, caspase 3 seems to play a central role in the initiation of cellular events of apoptosis. Assays for determination of caspase 3 activation detect early events of apoptosis. Following RNAi treatments, Western blot detection of active caspase 3 presence or proteolytic cleavage of products (i.e. PAR P) found in apoptotic cells further support an active induction of apoptosis. Because the cellular mechanisms that result in apoptosis are complex, each has its advantages and limitations. Consideration of other criteria/endpoints such as cellular morphology, chromatin condensation, membrane blebbing, apoptotic bodies help to further support cell death as apoptotic. Since not all the gene targets that regulate cell growth are anti-apoptotic, the DNA content of permeabilized cells is measured to obtain the profile of DNA content or cell cycle profile. Nuclei of apoptotic cells contain less DNA due to the leaking out to the cytoplasm (sub-G1 population). In addition, the use of DNA stains (i.e., propidium iodide) also differentiate between the different phases of the cell cycle in the cell population due to the presence of different quantities of DNA in G0/G1, S and G2/M. In these studies the subpopulations can be quantified.

For the 58P1D12 gene, RNAi studies facilitate the understanding of the contribution of the gene product in cancer pathways. Such active RNAi molecules have use in identifying assays to screen for mAbs that are active anti-tumor therapeutics. Further, siRNA are administered as therapeutics to cancer patients for reducing the malignant growth of several cancer types, including those listed in Table 1. When 58P1D12 plays a role in cell survival, cell proliferation, tumorigenesis, or apoptosis, it is used as a target for diagnostic, prognostic, preventative and/or therapeutic purposes.

Example 46

Homology Comparison of 58P1D12 to Known Sequences

The 58P1D12 protein has 273 amino acids with a calculated molecular weight of 30.47 kDa and a pI of 6.38. 58P1D12 is predicted to be a plasma membrane protein (PSORT http://psort.nibb.ac.jp/form.html). 58P1D12 contains a single transmembrane region from amino acids 218-240 with high probability that the amino-terminus resides outside, consistent with the topology of a Type 1 transmembrane protein (http://www.cbs.dtu.dk/services/TMHMM). Also visualized is a short hydrophobic stretch from amino acids 1-21, consistent with the existence of an amino-terminal signal peptide (http://www.cbs.dtu.dk/services/SignalP). Based on the TMpred algorithm of Hofmann and Stoffel which utilizes TMBASE (K. Hofmann, W. Stoffel, TMBASE—A database of membrane spanning protein segments Biol. Chem. Hoppe-Seyler 374:166, 1993), 58P1D12 contains a primary transmembrane region from amino acids 220-243 and a secondary transmembrane region from amino acids 1-21 (contiguous amino acids with values greater than 0 on the plot have high probability of being transmembrane regions) with an orientation in which the amino terminus resides inside and the carboxyl terminus outside. An alternative model is also predicted that 58P1D12 is a Type 1 transmembrane protein in which the amino-terminus resides outside and the protein contains a secondary transmembrane domain signal peptide from amino acids 1-21 and a primary transmembrane domain from aa214-aa246. The transmembrane prediction algorithms are accessed through the ExPasy molecular biology server (http://www.expasy.ch/tools/).

The C-type lectins are a family of both transmembrane and secreted proteins belonging to the lectin superfamily (8 different types of lectins) of carbohydrate recognition/adhesion molecules. They have related carbohydrate recognition domains (CRDs) in their extracellular domains, but for the transmembrane anchored proteins, they contain unique cytoplasmic domains. There are 14 identified families of C-type lectins, and they are all believed to have carbohydrate recognition properties, and have roles in various cell functions including cell adhesion, glycoprotein clearance and innate immunity. The adhesion function of C-type lectins is well-established for the selectin family. Additional functions of C-type lectins include signal transduction, and may provide a platform for protein associations. Some C-type lectins have been shown to facilitate glycoprotein clearance, elimination of foreign pathogens during innate immune responses and migration of cells. In mammals, C-type lectins are widely expressed, yet play roles in other species.

58P1D12 belongs to the type VIII family of C-type lectins, which includes one other member, layilin. Layilin has ~45% homology with 58P1D12, and interacts with the carbohydrate ligand hyaluronate. Hyaluronate, an extracellular matrix component consisting of repeating disaccharide units, is a ligand for several C-type lectins and is a putative ligand for 58P1D12, but was not detected in a binding reaction with 58P1D12 using the cetylpyridinium precipitation technique (Weng, L. et al, 2002, Genomics 80:62-70). Layilin localizes in membrane ruffles during cell migration, and interacts with the cytoskeletal protein talin (Borowsky, M. L. and Hynes, R. O., 1998, J. Cell Biol. 143:429-442). 58P1D12 also is observed in membrane ruffles and associates with talin through its cytoplasmic domain.

58P1D12 also interacts with other cytoplasmic signaling molecules, including serine/threonine and dual kinases that phosphorylate its intracellular serine and threonine residues. Such phosphorylation leads to the interaction of 58P1D12 with signaling molecules that are regulated by their interaction with the plasma membrane through association with 58P1D12. Such molecules induce signals that are critical to the function of 58P1D12 protein, which mediate the growth pathways that are stimulated in tumor cells expressing 58P1D12. Further, the adhesion properties of the 58P1D12 protein expressed on tumor cells enhances their ability to migrate during metastases, facilitates their homing to distal sites (lymph nodes), and promotes interactions with, and activation of, other cells such as endothelia during the formation of tumor masses. The adhesive and growth signals of 58P1D12 protein significantly promote unregulated growth of cancer cells, contributing to their growth advantage in vivo.

Similar to other C-type lectins, 58P1D12 is membrane associated. Functional analysis shows that a recombinant soluble form of 58P1D12 has both angiogenic and proliferative enhancing properties for endothelial cells (see Example 44). The membrane-associated form of the 58P1D12 protein or a secreted portion promotes establishment and maintenance of the malignant state of tumors expressing the 58P1D12 antigen. The formation of capillary beds surrounding tumors through the angiogenic function of 58P1D12 protein as well as the proliferative signals that are imparted by the protein promote survival of tumors which require nutrition and oxygenation.

Throughout this application, various website data content, publications, patent applications and patents are referenced. (Websites are referenced by their Uniform Resource Locator, or URL, addresses on the World Wide Web.) The disclosures of each of these references are hereby incorporated by reference herein in their entireties.

The present invention is not to be limited in scope by the embodiments disclosed herein, which are intended as single illustrations of individual aspects of the invention, and any that are functionally equivalent are within the scope of the invention. Various modifications to the models and methods of the invention, in addition to those described herein, will become apparent to those skilled in the art from the foregoing description and teachings, and are similarly intended to fall within the scope of the invention. Such modifications or other embodiments can be practiced without departing from the true scope and spirit of the invention.

SEQUENCE LISTING

```
<160> NUMBER OF SEQ ID NOS: 83

<210> SEQ ID NO 1
<211> LENGTH: 427
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 1

cagaattaca aaattctgga accaagtcaa tgagttatta tacttccatg ccactttctt     60

ttctggtact ctttgaaatc cacagtgtag actggtttgg actagttttt gttcttcctt    120

tacttttatg cagcatctgg aaacaacagg ttccaaaagc aaccagtatc agtaagagca    180

ggggtattgt tggtataaca acataaatta gattgggaat tatacctgct tcagtaacaa    240

ccacattctg atgggtgtct cctggttgat ttgtaagata aggcttttct acaggggctg    300

ttggattaat ctctggttca tacttgcaaa tataattgtg cttcatgtta cacctgtcat    360

cattccactg gtaaaggtag ggacccccaa ggccaggatt ggcagttggt tggtgatcat    420

cacaaca                                                               427


<210> SEQ ID NO 2
```

-continued

```
<211> LENGTH: 2550
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (380)...(1201)

<400> SEQUENCE: 2

atccaggacc agggcgcacc ggctcagcct ctcacttgtc agaggccggg gaagagaagc     60

aaagcgcaac ggtgtggtcc aagccggggc ttctgcttcg cctctaggac atacacggga    120

ccccctaact tcagtccccc aaacgcgcac cctcgaagtc ttgaactcca gccccgcaca    180

tccacgcgcg gcacaggcgc ggcaggcggc aggtcccggc cgaaggcgat gcgcgcaggg    240

ggtcgggcag ctgggctcgg gcggcgggag tagggcccgg cagggaggca gggaggctgc    300

agagtcagag tcgcgggctg cgccctgggc agaggccgcc ctcgctccac gcaacacctg    360

ctgctgccac cgcgccgcg atg agc cgc gtg gtc tcg ctg ctg ctg ggc gcc     412
                     Met Ser Arg Val Val Ser Leu Leu Leu Gly Ala
                      1               5                   10

gcg ctg ctc tgc ggc cac gga gcc ttc tgc cgc cgc gtg gtc agc ggc      460
Ala Leu Leu Cys Gly His Gly Ala Phe Cys Arg Arg Val Val Ser Gly
             15                  20                  25

caa aag gtg tgt ttt gct gac ttc aag cat ccc tgc tac aaa atg gcc      508
Gln Lys Val Cys Phe Ala Asp Phe Lys His Pro Cys Tyr Lys Met Ala
         30                  35                  40

tac ttc cat gaa ctg tcc agc cga gtg agc ttt cag gag gca cgc ctg      556
Tyr Phe His Glu Leu Ser Ser Arg Val Ser Phe Gln Glu Ala Arg Leu
     45                  50                  55

gct tgt gag agt gag gga gga gtc ctc ctc agc ctt gag aat gaa gca      604
Ala Cys Glu Ser Glu Gly Gly Val Leu Leu Ser Leu Glu Asn Glu Ala
 60                  65                  70                  75

gaa cag aag tta ata gag agc atg ttg caa aac ctg aca aaa ccc ggg      652
Glu Gln Lys Leu Ile Glu Ser Met Leu Gln Asn Leu Thr Lys Pro Gly
                 80                  85                  90

aca ggg att tct gat ggt gat ttc tgg ata ggg ctt tgg agg aat gga      700
Thr Gly Ile Ser Asp Gly Asp Phe Trp Ile Gly Leu Trp Arg Asn Gly
             95                 100                 105

gat ggg caa aca tct ggt gcc tgc cca gat ctc tac cag tgg tct gat      748
Asp Gly Gln Thr Ser Gly Ala Cys Pro Asp Leu Tyr Gln Trp Ser Asp
        110                 115                 120

gga agc aat tcc cag tac cga aac tgg tac aca gat gaa cct tcc tgc      796
Gly Ser Asn Ser Gln Tyr Arg Asn Trp Tyr Thr Asp Glu Pro Ser Cys
    125                 130                 135

gga agt gaa aag tgt gtt gtg atg tat cac caa cca act gcc aat cct      844
Gly Ser Glu Lys Cys Val Val Met Tyr His Gln Pro Thr Ala Asn Pro
140                 145                 150                 155

ggc ctt ggg ggt ccc tac ctt tac cag tgg aat gat gac agg tgt aac      892
Gly Leu Gly Gly Pro Tyr Leu Tyr Gln Trp Asn Asp Asp Arg Cys Asn
                160                 165                 170

atg aag cac aat tat att tgc aag tat gaa cca gag att aat cca aca      940
Met Lys His Asn Tyr Ile Cys Lys Tyr Glu Pro Glu Ile Asn Pro Thr
            175                 180                 185

gcc cct gta gaa aag cct tat ctt aca aat caa cca gga gac acc cat      988
Ala Pro Val Glu Lys Pro Tyr Leu Thr Asn Gln Pro Gly Asp Thr His
        190                 195                 200

cag aat gtg gtt gtt act gaa gca ggt ata att ccc aat cta att tat     1036
Gln Asn Val Val Val Thr Glu Ala Gly Ile Ile Pro Asn Leu Ile Tyr
    205                 210                 215

gtt gtt ata cca aca ata ccc ctg ctc tta ctg ata ctg gtt gct ttt     1084
Val Val Ile Pro Thr Ile Pro Leu Leu Leu Leu Ile Leu Val Ala Phe
```

-continued

```
220                 225                 230                 235

gga acc tgt tgt ttc cag atg ctg cat aaa agt aaa gga aga aca aaa     1132
Gly Thr Cys Cys Phe Gln Met Leu His Lys Ser Lys Gly Arg Thr Lys
                240                 245                 250

act agt cca aac cag tct aca ctg tgg att tca aag agt acc aga aaa     1180
Thr Ser Pro Asn Gln Ser Thr Leu Trp Ile Ser Lys Ser Thr Arg Lys
            255                 260                 265

gaa agt ggc atg gaa gta taa taactcattg acttggttcc agaattttgt        1231
Glu Ser Gly Met Glu Val  *
        270

aattctggat ctgtataagg aatggcatca gaacaatagc ttggaatggc ttgaaatcac   1291

aaaggatctg caagatgaac tgtaagctcc cccttgaggc aaatattaaa gtaattttta   1351

tatgtctatt atttcattta aagaatatgc tgtgctaata atggagtgag acatgcttat   1411

tttgctaaag gatgcaccca aacttcaaac ttcaagcaaa tgaaatggac aatgcagata   1471

aagttgttat caacacgtcg ggagtatgtg tgttagaagc aattcctttt atttctttca   1531

cctttcataa gttgttatct agtcaatgta atgtatattg tattgaaatt tacagtgtgc   1591

aaaagtattt tacctttgca taagtgtttg ataaaaatga actgttctaa tatttatttt   1651

tatggcatct cattttttcaa tacatgctct tttgattaaa gaaacttatt actgttgtca   1711

actgaattca cacacacaca aatatagtac catagaaaaa gtttgttttc tcgaaataat   1771

tcatctttca gcttctctgc ttttggtcaa tgtctaggaa atctcttcag aaataagaag   1831

ctatttcatt aagtgtgata taaacctcct caaacatttt acttagaggc aaggattgtc   1891

taatttcaat tgtgcaagac atgtgcctta taattatttt tagcttaaaa ttaaacagat   1951

tttgtaataa tgtaactttg ttaataggtg cataaacact aatgcagtca atttgaacaa   2011

aagaagtgac atacacaata taaatcatat gtcttcacac gttgcctata taatgagaag   2071

cagctctctg agggttctga aatcaatgtg gtccctctct tgcccactaa acaaagatgg   2131

ttgttcgggg tttgggattg acactggagg cagatagttg caaagttagt ctaaggtttc   2191

cctagctgta tttagcctct gactatatta gtatacaaag aggtcatgtg gttgagacca   2251

ggtgaatagt cactatcagt gtggagacaa gcacagcaca cagacatttt aggaaggaaa   2311

ggaactacga aatcgtgtga aaatgggttg gaacccatca gtgatcgcat attcattgat   2371

gagggtttgc ttgagataga aaatggtggc tcctttctgt cttatctcct agtttcttca   2431

atgcttacgc cttgttcttc tcaagagaaa gttgtaactc tctggtcttc atatgtccct   2491

gtgctccttt taaccaaata aagagttctt gtttctgaag aaaaaaaaaa aaaaaaaaa    2550


<210> SEQ ID NO 3
<211> LENGTH: 273
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 3

Met Ser Arg Val Val Ser Leu Leu Leu Gly Ala Ala Leu Leu Cys Gly
 1               5                  10                  15

His Gly Ala Phe Cys Arg Arg Val Val Ser Gly Gln Lys Val Cys Phe
            20                  25                  30

Ala Asp Phe Lys His Pro Cys Tyr Lys Met Ala Tyr Phe His Glu Leu
        35                  40                  45

Ser Ser Arg Val Ser Phe Gln Glu Ala Arg Leu Ala Cys Glu Ser Glu
    50                  55                  60

Gly Gly Val Leu Leu Ser Leu Glu Asn Glu Ala Glu Gln Lys Leu Ile
```

```
65                  70                  75                  80

Glu Ser Met Leu Gln Asn Leu Thr Lys Pro Gly Thr Gly Ile Ser Asp
                85                  90                  95

Gly Asp Phe Trp Ile Gly Leu Trp Arg Asn Gly Asp Gly Gln Thr Ser
            100                 105                 110

Gly Ala Cys Pro Asp Leu Tyr Gln Trp Ser Asp Gly Ser Asn Ser Gln
        115                 120                 125

Tyr Arg Asn Trp Tyr Thr Asp Glu Pro Ser Cys Gly Ser Glu Lys Cys
    130                 135                 140

Val Val Met Tyr His Gln Pro Thr Ala Asn Pro Gly Leu Gly Gly Pro
145                 150                 155                 160

Tyr Leu Tyr Gln Trp Asn Asp Asp Arg Cys Asn Met Lys His Asn Tyr
                165                 170                 175

Ile Cys Lys Tyr Glu Pro Glu Ile Asn Pro Thr Ala Pro Val Glu Lys
            180                 185                 190

Pro Tyr Leu Thr Asn Gln Pro Gly Asp Thr His Gln Asn Val Val Val
        195                 200                 205

Thr Glu Ala Gly Ile Ile Pro Asn Leu Ile Tyr Val Val Ile Pro Thr
    210                 215                 220

Ile Pro Leu Leu Leu Leu Ile Leu Val Ala Phe Gly Thr Cys Cys Phe
225                 230                 235                 240

Gln Met Leu His Lys Ser Lys Gly Arg Thr Lys Thr Ser Pro Asn Gln
                245                 250                 255

Ser Thr Leu Trp Ile Ser Lys Ser Thr Arg Lys Glu Ser Gly Met Glu
            260                 265                 270

Val


<210> SEQ ID NO 4
<211> LENGTH: 2418
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (388)...(1086)

<400> SEQUENCE: 4

ctgtggtgtt tttcccccgc tcctctggct gccttcctga tggatctctg tggtcccagg      60

caggaatggc ctgcttgggg acccagcgag ctcccaaggc ctttcctgct gcttcctcta     120

tccctgtgtt ttgcttggct ctctaaattg actcagctcc aggacatcag gaccccaggt     180

tctctggtct tgggactctg agacttgcac caggaatcct gcccaggctc tcaggccttt     240

ggactcagac tgagctactt cactggcttt cctggttctc cagcttgaag atggcagatc     300

gtgggacttc tcagcctcca taattgagtg agccaattcc ctggccaaaa ggtgtgtttt     360

gctgacttca agcatccctg ctacaaa atg gcc tac ttc cat gaa ctg tcc agc     414
                              Met Ala Tyr Phe His Glu Leu Ser Ser
                               1               5

cga gtg agc ttt cag gag gca cgc ctg gct tgt gag agt gag gga gga     462
Arg Val Ser Phe Gln Glu Ala Arg Leu Ala Cys Glu Ser Glu Gly Gly
 10                  15                  20                  25

gtc ctc ctc agc ctt gag aat gaa gca gaa cag aag tta ata gag agc     510
Val Leu Leu Ser Leu Glu Asn Glu Ala Glu Gln Lys Leu Ile Glu Ser
                 30                  35                  40

atg ttg caa aac ctg aca aaa ccc ggg aca ggg att tct gat ggt gat     558
Met Leu Gln Asn Leu Thr Lys Pro Gly Thr Gly Ile Ser Asp Gly Asp
             45                  50                  55
```

-continued

```
ttc tgg ata ggg ctt tgg agg aat gga gat ggg caa aca tct ggt gcc      606
Phe Trp Ile Gly Leu Trp Arg Asn Gly Asp Gly Gln Thr Ser Gly Ala
         60                 65                   70

tgc cca gat ctc tac cag tgg tct gat gga agc aat tcc cag tac cga      654
Cys Pro Asp Leu Tyr Gln Trp Ser Asp Gly Ser Asn Ser Gln Tyr Arg
     75                  80                  85

aac tgg tac aca gat gaa cct tcc tgc gga agt gaa aag tgt gtt gtg      702
Asn Trp Tyr Thr Asp Glu Pro Ser Cys Gly Ser Glu Lys Cys Val Val
 90                  95                 100                 105

atg tat cac caa cca act gcc aat cct ggc ctt ggg ggt ccc tac ctt      750
Met Tyr His Gln Pro Thr Ala Asn Pro Gly Leu Gly Gly Pro Tyr Leu
                110                 115                 120

tac cag tgg aat gat gac agg tgt aac atg aag cac aat tat att tgc      798
Tyr Gln Trp Asn Asp Asp Arg Cys Asn Met Lys His Asn Tyr Ile Cys
            125                 130                 135

aag tat gaa cca gag att aat cca aca gcc cct gta gaa aag cct tat      846
Lys Tyr Glu Pro Glu Ile Asn Pro Thr Ala Pro Val Glu Lys Pro Tyr
        140                 145                 150

ctt aca aat caa cca gga gac acc cat cag aat gtg gtt gtt act gaa      894
Leu Thr Asn Gln Pro Gly Asp Thr His Gln Asn Val Val Val Thr Glu
    155                 160                 165

gca ggt ata att ccc aat cta att tat gtt gtt ata cca aca ata ccc      942
Ala Gly Ile Ile Pro Asn Leu Ile Tyr Val Val Ile Pro Thr Ile Pro
170                 175                 180                 185

ctg ctc tta ctg ata ctg gtt gct ttt gga acc tgt tgt ttc cag atg      990
Leu Leu Leu Leu Ile Leu Val Ala Phe Gly Thr Cys Cys Phe Gln Met
                190                 195                 200

ctg cat aaa agt aaa gga aga aca aaa act agt cca aac cag tct aca     1038
Leu His Lys Ser Lys Gly Arg Thr Lys Thr Ser Pro Asn Gln Ser Thr
            205                 210                 215

ctg tgg att tca aag agt acc aga aaa gaa agt ggc atg gaa gta taa     1086
Leu Trp Ile Ser Lys Ser Thr Arg Lys Glu Ser Gly Met Glu Val  *
        220                 225                 230

taactcattg acttggttcc agaattttgt aattctggat ctgtataagg aatggcatca   1146

gaacaatagc ttggaatggc ttgaaatcac aaaggatctg caagatgaac tgtaagctcc   1206

cccttgaggc aaatattaaa gtaattttta tatgtctatt atttcattta aagaatatgc   1266

tgtgctaata atggagtgag acatgcttat tttgctaaag gatgcaccca aacttcaaac   1326

ttcaagcaaa tgaaatggac aatgcagata aagttgttat caacacgtcg ggagtatgtg   1386

tgttagaagc aattcctttt atttctttca cctttcataa gttgttatct agtcaatgta   1446

atgtatattg tattgaaatt tacagtgtgc aaaagtattt tacctttgca taagtgtttg   1506

ataaaaatga actgttctaa tatttatttt tatggcatct catttttcaa tacatgctct   1566

tttgattaaa gaaacttatt actgttgtca actgaattca cacacacaca aatatagtac   1626

catagaaaaa gtttgttttc tcgaaataat tcatctttca gcttctctgc ttttggtcaa   1686

tgtctaggaa atctcttcag aaataagaag ctatttcatt aagtgtgata taaacctcct   1746

caaacatttt acttagaggc aaggattgtc taatttcaat tgtgcaagac atgtgcctta   1806

taattatttt tagcttaaaa ttaaacagat tttgtaataa tgtaactttg ttaataggtg   1866

cataaacact aatgcagtca atttgaacaa aagaagtgac atacacaata taaatcatat   1926

gtcttcacac gttgcctata taatgagaag cagctctctg agggttctga aatcaatgtg   1986

gtccctctct tgcccactaa acaaagatgg ttgttcgggg tttgggattg acactggagg   2046

cagatagttg caaagttagt ctaaggtttc cctagctgta tttagcctct gactatatta   2106

gtatacaaag aggtcatgtg gttgagacca ggtgaatagt cactatcagt gtggagacaa   2166
```

-continued

```
gcacagcaca cagacatttt aggaaggaaa ggaactacga aatcgtgtga aaatgggttg   2226

gaacccatca gtgatcgcat attcattgat gagggtttgc ttgagataga aaatggtggc   2286

tcctttctgt cttatctcct agtttcttca atgcttacgc cttgttcttc tcaagagaaa   2346

gttgtaactc tctggtcttc atatgtccct gtgctccttt taaccaaata aagagttctt   2406

gtttctgaag aa                                                       2418


<210> SEQ ID NO 5
<211> LENGTH: 232
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 5

Met Ala Tyr Phe His Glu Leu Ser Ser Arg Val Ser Phe Gln Glu Ala
 1               5                   10                  15

Arg Leu Ala Cys Glu Ser Glu Gly Gly Val Leu Leu Ser Leu Glu Asn
            20                  25                  30

Glu Ala Glu Gln Lys Leu Ile Glu Ser Met Leu Gln Asn Leu Thr Lys
        35                  40                  45

Pro Gly Thr Gly Ile Ser Asp Gly Asp Phe Trp Ile Gly Leu Trp Arg
    50                  55                  60

Asn Gly Asp Gly Gln Thr Ser Gly Ala Cys Pro Asp Leu Tyr Gln Trp
65                  70                  75                  80

Ser Asp Gly Ser Asn Ser Gln Tyr Arg Asn Trp Tyr Thr Asp Glu Pro
                85                  90                  95

Ser Cys Gly Ser Glu Lys Cys Val Val Met Tyr His Gln Pro Thr Ala
            100                 105                 110

Asn Pro Gly Leu Gly Gly Pro Tyr Leu Tyr Gln Trp Asn Asp Asp Arg
        115                 120                 125

Cys Asn Met Lys His Asn Tyr Ile Cys Lys Tyr Glu Pro Glu Ile Asn
    130                 135                 140

Pro Thr Ala Pro Val Glu Lys Pro Tyr Leu Thr Asn Gln Pro Gly Asp
145                 150                 155                 160

Thr His Gln Asn Val Val Val Thr Glu Ala Gly Ile Ile Pro Asn Leu
                165                 170                 175

Ile Tyr Val Val Ile Pro Thr Ile Pro Leu Leu Leu Leu Ile Leu Val
            180                 185                 190

Ala Phe Gly Thr Cys Cys Phe Gln Met Leu His Lys Ser Lys Gly Arg
        195                 200                 205

Thr Lys Thr Ser Pro Asn Gln Ser Thr Leu Trp Ile Ser Lys Ser Thr
    210                 215                 220

Arg Lys Glu Ser Gly Met Glu Val
225                 230


<210> SEQ ID NO 6
<211> LENGTH: 2236
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (206)...(904)

<400> SEQUENCE: 6

gcttgaagat ggcagatcgt gggacttctc agcctccata attgagtgag ccaattccct      60

gaatacaaca agaagatggc catctgtaga ccaggaggtg gtccctcccc agaaactgga     120
```

-continued

```
tgggccagca cgttgattct gagcttctag cctccagaac tgccaaaagg tgtgttttgc    180

tgacttcaag catccctgct acaaa atg gcc tac ttc cat gaa ctg tcc agc      232
                           Met Ala Tyr Phe His Glu Leu Ser Ser
                            1               5

cga gtg agc ttt cag gag gca cgc ctg gct tgt gag agt gag gga gga      280
Arg Val Ser Phe Gln Glu Ala Arg Leu Ala Cys Glu Ser Glu Gly Gly
 10                  15                  20                  25

gtc ctc ctc agc ctt gag aat gaa gca gaa cag aag tta ata gag agc      328
Val Leu Leu Ser Leu Glu Asn Glu Ala Glu Gln Lys Leu Ile Glu Ser
                 30                  35                  40

atg ttg caa aac ctg aca aaa ccc ggg aca ggg att tct gat ggt gat      376
Met Leu Gln Asn Leu Thr Lys Pro Gly Thr Gly Ile Ser Asp Gly Asp
             45                  50                  55

ttc tgg ata ggg ctt tgg agg aat gga gat ggg caa aca tct ggt gcc      424
Phe Trp Ile Gly Leu Trp Arg Asn Gly Asp Gly Gln Thr Ser Gly Ala
         60                  65                  70

tgc cca gat ctc tac cag tgg tct gat gga agc aat tcc cag tac cga      472
Cys Pro Asp Leu Tyr Gln Trp Ser Asp Gly Ser Asn Ser Gln Tyr Arg
     75                  80                  85

aac tgg tac aca gat gaa cct tcc tgc gga agt gaa aag tgt gtt gtg      520
Asn Trp Tyr Thr Asp Glu Pro Ser Cys Gly Ser Glu Lys Cys Val Val
 90                  95                 100                 105

atg tat cac caa cca act gcc aat cct ggc ctt ggg ggt ccc tac ctt      568
Met Tyr His Gln Pro Thr Ala Asn Pro Gly Leu Gly Gly Pro Tyr Leu
                110                 115                 120

tac cag tgg aat gat gac agg tgt aac atg aag cac aat tat att tgc      616
Tyr Gln Trp Asn Asp Asp Arg Cys Asn Met Lys His Asn Tyr Ile Cys
            125                 130                 135

aag tat gaa cca gag att aat cca aca gcc cct gta gaa aag cct tat      664
Lys Tyr Glu Pro Glu Ile Asn Pro Thr Ala Pro Val Glu Lys Pro Tyr
        140                 145                 150

ctt aca aat caa cca gga gac acc cat cag aat gtg gtt gtt act gaa      712
Leu Thr Asn Gln Pro Gly Asp Thr His Gln Asn Val Val Val Thr Glu
    155                 160                 165

gca ggt ata att ccc aat cta att tat gtt gtt ata cca aca ata ccc      760
Ala Gly Ile Ile Pro Asn Leu Ile Tyr Val Val Ile Pro Thr Ile Pro
170                 175                 180                 185

ctg ctc tta ctg ata ctg gtt gct ttt gga acc tgt tgt ttc cag atg      808
Leu Leu Leu Leu Ile Leu Val Ala Phe Gly Thr Cys Cys Phe Gln Met
                190                 195                 200

ctg cat aaa agt aaa gga aga aca aaa act agt cca aac cag tct aca      856
Leu His Lys Ser Lys Gly Arg Thr Lys Thr Ser Pro Asn Gln Ser Thr
            205                 210                 215

ctg tgg att tca aag agt acc aga aaa gaa agt ggc atg gaa gta taa      904
Leu Trp Ile Ser Lys Ser Thr Arg Lys Glu Ser Gly Met Glu Val  *
        220                 225                 230

taactcattg acttggttcc agaattttgt aattctggat ctgtataagg aatggcatca    964

gaacaatagc ttggaatggc ttgaaatcac aaaggatctg caagatgaac tgtaagctcc   1024

cccttgaggc aaatattaaa gtaattttta tatgtctatt atttcattta aagaatatgc   1084

tgtgctaata atggagtgag acatgcttat tttgctaaag gatgcaccca aacttcaaac   1144

ttcaagcaaa tgaaatggac aatgcagata aagttgttat caacacgtcg ggagtatgtg   1204

tgttagaagc aattcctttt atttctttca cctttcataa gttgttatct agtcaatgta   1264

atgtatattg tattgaaatt tacagtgtgc aaaagtattt tacctttgca taagtgtttg   1324

ataaaaatga actgttctaa tatttatttt tatggcatct cattttcaa tacatgctct   1384

tttgattaaa gaaacttatt actgttgtca actgaattca cacacacaca aatatagtac   1444
```

-continued

```
catagaaaaa gtttgttttc tcgaaataat tcatctttca gcttctctgc ttttggtcaa   1504

tgtctaggaa atctcttcag aaataagaag ctatttcatt aagtgtgata taaacctcct   1564

caaacatttt acttagaggc aaggattgtc taatttcaat tgtgcaagac atgtgcctta   1624

taattatttt tagcttaaaa ttaaacagat tttgtaataa tgtaactttg ttaataggtg   1684

cataaacact aatgcagtca atttgaacaa aagaagtgac atacacaata taaatcatat   1744

gtcttcacac gttgcctata taatgagaag cagctctctg agggttctga aatcaatgtg   1804

gtccctctct tgcccactaa acaaagatgg ttgttcgggg tttgggattg acactggagg   1864

cagatagttg caaagttagt ctaaggtttc cctagctgta tttagcctct gactatatta   1924

gtatacaaag aggtcatgtg gttgagacca ggtgaatagt cactatcagt gtggagacaa   1984

gcacagcaca cagacatttt aggaaggaaa ggaactacga aatcgtgtga aaatgggttg   2044

gaacccatca gtgatcgcat attcattgat gagggtttgc ttgagataga aaatggtggc   2104

tcctttctgt cttatctcct agtttcttca atgcttacgc cttgttcttc tcaagagaaa   2164

gttgtaactc tctggtcttc atatgtccct gtgctccttt taaccaaata aagagttctt   2224

gtttctgaag aa                                                       2236


<210> SEQ ID NO 7
<211> LENGTH: 232
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 7

Met Ala Tyr Phe His Glu Leu Ser Ser Arg Val Ser Phe Gln Glu Ala
 1               5                  10                  15

Arg Leu Ala Cys Glu Ser Glu Gly Gly Val Leu Leu Ser Leu Glu Asn
            20                  25                  30

Glu Ala Glu Gln Lys Leu Ile Glu Ser Met Leu Gln Asn Leu Thr Lys
        35                  40                  45

Pro Gly Thr Gly Ile Ser Asp Gly Asp Phe Trp Ile Gly Leu Trp Arg
    50                  55                  60

Asn Gly Asp Gly Gln Thr Ser Gly Ala Cys Pro Asp Leu Tyr Gln Trp
65                  70                  75                  80

Ser Asp Gly Ser Asn Ser Gln Tyr Arg Asn Trp Tyr Thr Asp Glu Pro
                85                  90                  95

Ser Cys Gly Ser Glu Lys Cys Val Val Met Tyr His Gln Pro Thr Ala
            100                 105                 110

Asn Pro Gly Leu Gly Gly Pro Tyr Leu Tyr Gln Trp Asn Asp Asp Arg
        115                 120                 125

Cys Asn Met Lys His Asn Tyr Ile Cys Lys Tyr Glu Pro Glu Ile Asn
    130                 135                 140

Pro Thr Ala Pro Val Glu Lys Pro Tyr Leu Thr Asn Gln Pro Gly Asp
145                 150                 155                 160

Thr His Gln Asn Val Val Val Thr Glu Ala Gly Ile Ile Pro Asn Leu
                165                 170                 175

Ile Tyr Val Val Ile Pro Thr Ile Pro Leu Leu Leu Leu Ile Leu Val
            180                 185                 190

Ala Phe Gly Thr Cys Cys Phe Gln Met Leu His Lys Ser Lys Gly Arg
        195                 200                 205

Thr Lys Thr Ser Pro Asn Gln Ser Thr Leu Trp Ile Ser Lys Ser Thr
    210                 215                 220
```

-continued

```
Arg Lys Glu Ser Gly Met Glu Val
225                 230


<210> SEQ ID NO 8
<211> LENGTH: 2133
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (206)...(916)

<400> SEQUENCE: 8

gcttgaagat ggcagatcgt gggacttctc agcctccata attgagtgag ccaattccct     60

gaatacaaca agaagatggc catctgtaga ccaggaggtg gtccctcccc agaaactgga    120

tgggccagca cgttgattct gagcttctag cctccagaac tgccaaaagg tgtgttttgc    180

tgacttcaag catccctgct acaaa atg gcc tac ttc cat gaa ctg tcc agc      232
                            Met Ala Tyr Phe His Glu Leu Ser Ser
                             1               5

cga gtg agc ttt cag gag gca cgc ctg gct tgt gag agt gag gga gga      280
Arg Val Ser Phe Gln Glu Ala Arg Leu Ala Cys Glu Ser Glu Gly Gly
 10                  15                  20                  25

gtc ctc ctc agc ctt gag aat gaa gca gaa cag aag tta ata gag agc      328
Val Leu Leu Ser Leu Glu Asn Glu Ala Glu Gln Lys Leu Ile Glu Ser
                 30                  35                  40

atg ttg caa aac ctg aca aaa ccc ggg aca ggg att tct gat ggt gat      376
Met Leu Gln Asn Leu Thr Lys Pro Gly Thr Gly Ile Ser Asp Gly Asp
             45                  50                  55

ttc tgg ata ggg ctt tgg agg aat gga gat ggg caa aca tct ggt gcc      424
Phe Trp Ile Gly Leu Trp Arg Asn Gly Asp Gly Gln Thr Ser Gly Ala
         60                  65                  70

tgc cca gat ctc tac cag tgg tct gat gga agc aat tcc cag tac cga      472
Cys Pro Asp Leu Tyr Gln Trp Ser Asp Gly Ser Asn Ser Gln Tyr Arg
     75                  80                  85

aac tgg tac aca gat gaa cct tcc tgc gga agt gaa aag tgt gtt gtg      520
Asn Trp Tyr Thr Asp Glu Pro Ser Cys Gly Ser Glu Lys Cys Val Val
 90                  95                 100                 105

atg tat cac caa cca act gcc aat cct ggc ctt ggg ggt ccc tac ctt      568
Met Tyr His Gln Pro Thr Ala Asn Pro Gly Leu Gly Gly Pro Tyr Leu
                110                 115                 120

tac cag tgg aat gat gac agg tgt aac atg aag cac aat tat att tgc      616
Tyr Gln Trp Asn Asp Asp Arg Cys Asn Met Lys His Asn Tyr Ile Cys
            125                 130                 135

aag tat gaa cca gag att aat cca aca gcc cct gta gaa aag cct tat      664
Lys Tyr Glu Pro Glu Ile Asn Pro Thr Ala Pro Val Glu Lys Pro Tyr
        140                 145                 150

ctt aca aat caa cca gga gac acc cat cag aat gtg gtt gtt act gaa      712
Leu Thr Asn Gln Pro Gly Asp Thr His Gln Asn Val Val Val Thr Glu
    155                 160                 165

gca gta aag gaa gaa caa aaa cta gtc caa acc agt cta cac tgt gga      760
Ala Val Lys Glu Glu Gln Lys Leu Val Gln Thr Ser Leu His Cys Gly
170                 175                 180                 185

ttt caa aga gta cca gaa aag aaa gtg gca tgg aag tat aat aac tca      808
Phe Gln Arg Val Pro Glu Lys Lys Val Ala Trp Lys Tyr Asn Asn Ser
                190                 195                 200

ttg act tgg ttc cag aat ttt gta att ctg gat ctg tat aag gaa tgg      856
Leu Thr Trp Phe Gln Asn Phe Val Ile Leu Asp Leu Tyr Lys Glu Trp
            205                 210                 215

cat cag aac aat agc ttg gaa tgg ctt gaa atc aca aag gat ctg caa      904
His Gln Asn Asn Ser Leu Glu Trp Leu Glu Ile Thr Lys Asp Leu Gln
        220                 225                 230
```

-continued

```
gat gaa ctg taa gctccccctt gaggcaaata ttaaagtaat ttttatatgt          956
Asp Glu Leu  *
    235

ctattatttc atttaaagaa tatgctgtgc taataatgga gtgagacatg cttattttgc   1016

taaaggatgc acccaaactt caaacttcaa gcaaatgaaa tggacaatgc agataaagtt   1076

gttatcaaca cgtcgggagt atgtgtgtta gaagcaattc cttttatttc tttcaccttt   1136

cataagttgt tatctagtca atgtaatgta tattgtattg aaatttacag tgtgcaaaag   1196

tattttacct ttgcataagt gtttgataaa aatgaactgt tctaatattt atttttatgg   1256

catctcattt ttcaatacat gctcttttga ttaaagaaac ttattactgt tgtcaactga   1316

attcacacac acacaaatat agtaccatag aaaaagtttg ttttctcgaa ataattcatc   1376

tttcagcttc tctgcttttg gtcaatgtct aggaaatctc ttcagaaata agaagctatt   1436

tcattaagtg tgatataaac ctcctcaaac attttactta gaggcaagga ttgtctaatt   1496

tcaattgtgc aagacatgtg ccttataatt atttttagct taaaattaaa cagattttgt   1556

aataatgtaa ctttgttaat aggtgcataa acactaatgc agtcaatttg aacaaaagaa   1616

gtgacataca caatataaat catatgtctt cacacgttgc ctatataatg agaagcagct   1676

ctctgagggt tctgaaatca atgtggtccc tctcttgccc actaaacaaa gatggttgtt   1736

cggggtttgg gattgacact ggaggcagat agttgcaaag ttagtctaag gtttccctag   1796

ctgtatttag cctctgacta tattagtata caaagaggtc atgtggttga gaccaggtga   1856

atagtcacta tcagtgtgga gacaagcaca gcacacagac attttaggaa ggaaaggaac   1916

tacgaaatcg tgtgaaaatg ggttggaacc catcagtgat cgcatattca ttgatgaggg   1976

tttgcttgag atagaaaatg gtggctcctt tctgtcttat ctcctagttt cttcaatgct   2036

tacgccttgt tcttctcaag agaaagttgt aactctctgg tcttcatatg tccctgtgct   2096

ccttttaacc aaataaagag ttcttgtttc tgaagaa                             2133


<210> SEQ ID NO 9
<211> LENGTH: 236
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 9

Met Ala Tyr Phe His Glu Leu Ser Ser Arg Val Ser Phe Gln Glu Ala
 1               5                  10                  15

Arg Leu Ala Cys Glu Ser Glu Gly Gly Val Leu Leu Ser Leu Glu Asn
            20                  25                  30

Glu Ala Glu Gln Lys Leu Ile Glu Ser Met Leu Gln Asn Leu Thr Lys
        35                  40                  45

Pro Gly Thr Gly Ile Ser Asp Gly Asp Phe Trp Ile Gly Leu Trp Arg
    50                  55                  60

Asn Gly Asp Gly Gln Thr Ser Gly Ala Cys Pro Asp Leu Tyr Gln Trp
65                  70                  75                  80

Ser Asp Gly Ser Asn Ser Gln Tyr Arg Asn Trp Tyr Thr Asp Glu Pro
                85                  90                  95

Ser Cys Gly Ser Glu Lys Cys Val Val Met Tyr His Gln Pro Thr Ala
            100                 105                 110

Asn Pro Gly Leu Gly Gly Pro Tyr Leu Tyr Gln Trp Asn Asp Asp Arg
        115                 120                 125

Cys Asn Met Lys His Asn Tyr Ile Cys Lys Tyr Glu Pro Glu Ile Asn
    130                 135                 140
```

-continued

```
Pro Thr Ala Pro Val Glu Lys Pro Tyr Leu Thr Asn Gln Pro Gly Asp
145                 150                 155                 160

Thr His Gln Asn Val Val Val Thr Glu Ala Val Lys Glu Glu Gln Lys
                165                 170                 175

Leu Val Gln Thr Ser Leu His Cys Gly Phe Gln Arg Val Pro Glu Lys
            180                 185                 190

Lys Val Ala Trp Lys Tyr Asn Asn Ser Leu Thr Trp Phe Gln Asn Phe
        195                 200                 205

Val Ile Leu Asp Leu Tyr Lys Glu Trp His Gln Asn Asn Ser Leu Glu
    210                 215                 220

Trp Leu Glu Ile Thr Lys Asp Leu Gln Asp Glu Leu
225                 230                 235


<210> SEQ ID NO 10
<211> LENGTH: 2033
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (106)...(816)

<400> SEQUENCE: 10

gcttgaagat ggcagatcgt gggacttctc agcctccata attgagtgag ccaattccct     60

ggccaaaagg tgtgttttgc tgacttcaag catccctgct acaaa atg gcc tac ttc    117
                                                  Met Ala Tyr Phe
                                                   1

cat gaa ctg tcc agc cga gtg agc ttt cag gag gca cgc ctg gct tgt      165
His Glu Leu Ser Ser Arg Val Ser Phe Gln Glu Ala Arg Leu Ala Cys
 5                   10                  15                  20

gag agt gag gga gga gtc ctc ctc agc ctt gag aat gaa gca gaa cag      213
Glu Ser Glu Gly Gly Val Leu Leu Ser Leu Glu Asn Glu Ala Glu Gln
                 25                  30                  35

aag tta ata gag agc atg ttg caa aac ctg aca aaa ccc ggg aca ggg      261
Lys Leu Ile Glu Ser Met Leu Gln Asn Leu Thr Lys Pro Gly Thr Gly
             40                  45                  50

att tct gat ggt gat ttc tgg ata ggg ctt tgg agg aat gga gat ggg      309
Ile Ser Asp Gly Asp Phe Trp Ile Gly Leu Trp Arg Asn Gly Asp Gly
         55                  60                  65

caa aca tct ggt gcc tgc cca gat ctc tac cag tgg tct gat gga agc      357
Gln Thr Ser Gly Ala Cys Pro Asp Leu Tyr Gln Trp Ser Asp Gly Ser
     70                  75                  80

aat tcc cag tac cga aac tgg tac aca gat gaa cct tcc tgc gga agt      405
Asn Ser Gln Tyr Arg Asn Trp Tyr Thr Asp Glu Pro Ser Cys Gly Ser
 85                  90                  95                 100

gaa aag tgt gtt gtg atg tat cac caa cca act gcc aat cct ggc ctt      453
Glu Lys Cys Val Val Met Tyr His Gln Pro Thr Ala Asn Pro Gly Leu
                105                 110                 115

ggg ggt ccc tac ctt tac cag tgg aat gat gac agg tgt aac atg aag      501
Gly Gly Pro Tyr Leu Tyr Gln Trp Asn Asp Asp Arg Cys Asn Met Lys
            120                 125                 130

cac aat tat att tgc aag tat gaa cca gag att aat cca aca gcc cct      549
His Asn Tyr Ile Cys Lys Tyr Glu Pro Glu Ile Asn Pro Thr Ala Pro
        135                 140                 145

gta gaa aag cct tat ctt aca aat caa cca gga gac acc cat cag aat      597
Val Glu Lys Pro Tyr Leu Thr Asn Gln Pro Gly Asp Thr His Gln Asn
    150                 155                 160

gtg gtt gtt act gaa gca gta aag gaa gaa caa aaa cta gtc caa acc      645
Val Val Val Thr Glu Ala Val Lys Glu Glu Gln Lys Leu Val Gln Thr
165                 170                 175                 180
```

-continued

```
agt cta cac tgt gga ttt caa aga gta cca gaa aag aaa gtg gca tgg      693
Ser Leu His Cys Gly Phe Gln Arg Val Pro Glu Lys Lys Val Ala Trp
                185                 190                 195

aag tat aat aac tca ttg act tgg ttc cag aat ttt gta att ctg gat      741
Lys Tyr Asn Asn Ser Leu Thr Trp Phe Gln Asn Phe Val Ile Leu Asp
            200                 205                 210

ctg tat aag gaa tgg cat cag aac aat agc ttg gaa tgg ctt gaa atc      789
Leu Tyr Lys Glu Trp His Gln Asn Asn Ser Leu Glu Trp Leu Glu Ile
        215                 220                 225

aca aag gat ctg caa gat gaa ctg taa gctcccctt gaggcaaata             836
Thr Lys Asp Leu Gln Asp Glu Leu  *
    230                 235

ttaaagtaat ttttatatgt ctattatttc atttaaagaa tatgctgtgc taataatgga    896

gtgagacatg cttattttgc taaaggatgc acccaaactt caaacttcaa gcaaatgaaa    956

tggacaatgc agataaagtt gttatcaaca cgtcgggagt atgtgtgtta gaagcaattc   1016

cttttatttc tttcaccttt cataagttgt tatctagtca atgtaatgta tattgtattg   1076

aaatttacag tgtgcaaaag tattttacct ttgcataagt gtttgataaa aatgaactgt   1136

tctaatattt atttttatgg catctcattt ttcaatacat gctcttttga ttaaagaaac   1196

ttattactgt tgtcaactga attcacacac acacaaatat agtaccatag aaaaagtttg   1256

ttttctcgaa ataattcatc tttcagcttc tctgcttttg gtcaatgtct aggaaatctc   1316

ttcagaaata agaagctatt tcattaagtg tgatataaac ctcctcaaac attttactta   1376

gaggcaagga ttgtctaatt tcaattgtgc aagacatgtg ccttataatt atttttagct   1436

taaaattaaa cagattttgt aataatgtaa ctttgttaat aggtgcataa acactaatgc   1496

agtcaatttg aacaaaagaa gtgacataca caatataaat catatgtctt cacacgttgc   1556

ctatataatg agaagcagct ctctgagggt tctgaaatca atgtggtccc tctcttgccc   1616

actaaacaaa gatggttgtt cggggtttgg gattgacact ggaggcagat agttgcaaag   1676

ttagtctaag gtttccctag ctgtatttag cctctgacta tattagtata caaagaggtc   1736

atgtggttga gaccaggtga atagtcacta tcagtgtgga gacaagcaca gcacacagac   1796

attttaggaa ggaaaggaac tacgaaatcg tgtgaaaatg ggttggaacc catcagtgat   1856

cgcatattca ttgatgaggg tttgcttgag atagaaaatg gtggctcctt tctgtcttat   1916

ctcctagttt cttcaatgct tacgccttgt tcttctcaag agaaagttgt aactctctgg   1976

tcttcatatg tccctgtgct ccttttaacc aaataaagag ttcttgtttc tgaagaa      2033


<210> SEQ ID NO 11
<211> LENGTH: 236
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 11

Met Ala Tyr Phe His Glu Leu Ser Ser Arg Val Ser Phe Gln Glu Ala
 1               5                  10                  15

Arg Leu Ala Cys Glu Ser Glu Gly Gly Val Leu Leu Ser Leu Glu Asn
            20                  25                  30

Glu Ala Glu Gln Lys Leu Ile Glu Ser Met Leu Gln Asn Leu Thr Lys
        35                  40                  45

Pro Gly Thr Gly Ile Ser Asp Gly Asp Phe Trp Ile Gly Leu Trp Arg
    50                  55                  60

Asn Gly Asp Gly Gln Thr Ser Gly Ala Cys Pro Asp Leu Tyr Gln Trp
65                  70                  75                  80
```

-continued

```
Ser Asp Gly Ser Asn Ser Gln Tyr Arg Asn Trp Tyr Thr Asp Glu Pro
                85                  90                  95

Ser Cys Gly Ser Glu Lys Cys Val Val Met Tyr His Gln Pro Thr Ala
            100                 105                 110

Asn Pro Gly Leu Gly Gly Pro Tyr Leu Tyr Gln Trp Asn Asp Asp Arg
        115                 120                 125

Cys Asn Met Lys His Asn Tyr Ile Cys Lys Tyr Glu Pro Glu Ile Asn
    130                 135                 140

Pro Thr Ala Pro Val Glu Lys Pro Tyr Leu Thr Asn Gln Pro Gly Asp
145                 150                 155                 160

Thr His Gln Asn Val Val Val Thr Glu Ala Val Lys Glu Glu Gln Lys
                165                 170                 175

Leu Val Gln Thr Ser Leu His Cys Gly Phe Gln Arg Val Pro Glu Lys
            180                 185                 190

Lys Val Ala Trp Lys Tyr Asn Asn Ser Leu Thr Trp Phe Gln Asn Phe
        195                 200                 205

Val Ile Leu Asp Leu Tyr Lys Glu Trp His Gln Asn Asn Ser Leu Glu
    210                 215                 220

Trp Leu Glu Ile Thr Lys Asp Leu Gln Asp Glu Leu
225                 230                 235


<210> SEQ ID NO 12
<211> LENGTH: 273
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 12

Met Ser Arg Val Val Ser Leu Leu Leu Gly Ala Ala Leu Leu Cys Gly
 1               5                   10                  15

His Gly Ala Phe Cys Arg Arg Val Val Ser Gly Gln Lys Val Cys Phe
            20                  25                  30

Ala Asp Phe Lys His Pro Cys Tyr Lys Met Ala Tyr Phe His Glu Leu
        35                  40                  45

Ser Ser Arg Val Ser Phe Gln Glu Ala Arg Leu Ala Cys Glu Ser Glu
    50                  55                  60

Gly Gly Val Leu Leu Ser Leu Glu Asn Glu Ala Glu Gln Lys Leu Ile
65                  70                  75                  80

Glu Ser Met Leu Gln Asn Leu Thr Lys Pro Gly Thr Gly Ile Ser Asp
                85                  90                  95

Gly Asp Phe Trp Ile Gly Leu Trp Arg Asn Gly Asp Gly Gln Thr Ser
            100                 105                 110

Gly Ala Cys Pro Asp Leu Tyr Gln Trp Ser Asp Gly Ser Asn Ser Gln
        115                 120                 125

Tyr Arg Asn Trp Tyr Thr Asp Glu Pro Ser Cys Gly Ser Glu Lys Cys
    130                 135                 140

Val Val Met Tyr His Gln Pro Thr Ala Asn Pro Gly Leu Gly Gly Pro
145                 150                 155                 160

Tyr Leu Tyr Gln Trp Asn Asp Asp Arg Cys Asn Met Lys His Asn Tyr
                165                 170                 175

Ile Cys Lys Tyr Glu Pro Glu Ile Asn Pro Thr Ala Pro Val Glu Lys
            180                 185                 190

Pro Tyr Leu Thr Asn Gln Pro Gly Asp Thr His Gln Asn Val Val Val
        195                 200                 205

Thr Glu Ala Gly Ile Ile Pro Asn Leu Ile Tyr Val Val Ile Pro Thr
```

-continued

```
    210                 215                 220

Ile Pro Leu Leu Leu Leu Ile Leu Val Ala Phe Gly Thr Cys Cys Phe
225                 230                 235                 240

Gln Met Leu His Lys Ser Lys Gly Arg Thr Lys Thr Ser Pro Asn Gln
                245                 250                 255

Ser Thr Leu Trp Ile Ser Lys Ser Thr Arg Lys Glu Ser Gly Met Glu
            260                 265                 270

Val


<210> SEQ ID NO 13
<211> LENGTH: 232
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 13

Met Ala Tyr Phe His Glu Leu Ser Ser Arg Val Ser Phe Gln Glu Ala
 1               5                  10                  15

Arg Leu Ala Cys Glu Ser Glu Gly Gly Val Leu Leu Ser Leu Glu Asn
            20                  25                  30

Glu Ala Glu Gln Lys Leu Ile Glu Ser Met Leu Gln Asn Leu Thr Lys
        35                  40                  45

Pro Gly Thr Gly Ile Ser Asp Gly Asp Phe Trp Ile Gly Leu Trp Arg
    50                  55                  60

Asn Gly Asp Gly Gln Thr Ser Gly Ala Cys Pro Asp Leu Tyr Gln Trp
65                  70                  75                  80

Ser Asp Gly Ser Asn Ser Gln Tyr Arg Asn Trp Tyr Thr Asp Glu Pro
                85                  90                  95

Ser Cys Gly Ser Glu Lys Cys Val Val Met Tyr His Gln Pro Thr Ala
            100                 105                 110

Asn Pro Gly Leu Gly Gly Pro Tyr Leu Tyr Gln Trp Asn Asp Asp Arg
        115                 120                 125

Cys Asn Met Lys His Asn Tyr Ile Cys Lys Tyr Glu Pro Glu Ile Asn
    130                 135                 140

Pro Thr Ala Pro Val Glu Lys Pro Tyr Leu Thr Asn Gln Pro Gly Asp
145                 150                 155                 160

Thr His Gln Asn Val Val Val Thr Glu Ala Gly Ile Ile Pro Asn Leu
                165                 170                 175

Ile Tyr Val Val Ile Pro Thr Ile Pro Leu Leu Leu Leu Ile Leu Val
            180                 185                 190

Ala Phe Gly Thr Cys Cys Phe Gln Met Leu His Lys Ser Lys Gly Arg
        195                 200                 205

Thr Lys Thr Ser Pro Asn Gln Ser Thr Leu Trp Ile Ser Lys Ser Thr
    210                 215                 220

Arg Lys Glu Ser Gly Met Glu Val
225                 230


<210> SEQ ID NO 14
<211> LENGTH: 236
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 14

Met Ala Tyr Phe His Glu Leu Ser Ser Arg Val Ser Phe Gln Glu Ala
 1               5                  10                  15

Arg Leu Ala Cys Glu Ser Glu Gly Gly Val Leu Leu Ser Leu Glu Asn
            20                  25                  30
```

-continued

```
Glu Ala Glu Gln Lys Leu Ile Glu Ser Met Leu Gln Asn Leu Thr Lys
        35                  40                  45

Pro Gly Thr Gly Ile Ser Asp Gly Asp Phe Trp Ile Gly Leu Trp Arg
    50                  55                  60

Asn Gly Asp Gly Gln Thr Ser Gly Ala Cys Pro Asp Leu Tyr Gln Trp
65                  70                  75                  80

Ser Asp Gly Ser Asn Ser Gln Tyr Arg Asn Trp Tyr Thr Asp Glu Pro
                85                  90                  95

Ser Cys Gly Ser Glu Lys Cys Val Val Met Tyr His Gln Pro Thr Ala
            100                 105                 110

Asn Pro Gly Leu Gly Gly Pro Tyr Leu Tyr Gln Trp Asn Asp Asp Arg
        115                 120                 125

Cys Asn Met Lys His Asn Tyr Ile Cys Lys Tyr Glu Pro Glu Ile Asn
    130                 135                 140

Pro Thr Ala Pro Val Glu Lys Pro Tyr Leu Thr Asn Gln Pro Gly Asp
145                 150                 155                 160

Thr His Gln Asn Val Val Val Thr Glu Ala Val Lys Glu Glu Gln Lys
                165                 170                 175

Leu Val Gln Thr Ser Leu His Cys Gly Phe Gln Arg Val Pro Glu Lys
            180                 185                 190

Lys Val Ala Trp Lys Tyr Asn Asn Ser Leu Thr Trp Phe Gln Asn Phe
        195                 200                 205

Val Ile Leu Asp Leu Tyr Lys Glu Trp His Gln Asn Asn Ser Leu Glu
    210                 215                 220

Trp Leu Glu Ile Thr Lys Asp Leu Gln Asp Glu Leu
225                 230                 235


<210> SEQ ID NO 15
<211> LENGTH: 273
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 15

Met Ser Arg Val Val Ser Leu Leu Leu Gly Ala Ala Leu Leu Cys Gly
 1               5                  10                  15

His Gly Ala Phe Cys Arg Arg Val Val Ser Gly Gln Lys Val Cys Phe
            20                  25                  30

Ala Asp Phe Lys His Pro Cys Tyr Lys Met Ala Tyr Phe His Glu Leu
        35                  40                  45

Ser Ser Arg Val Ser Phe Gln Glu Ala Arg Leu Ala Cys Glu Ser Glu
    50                  55                  60

Gly Gly Val Leu Leu Ser Leu Glu Asn Glu Ala Glu Gln Lys Leu Ile
65                  70                  75                  80

Glu Ser Met Leu Gln Asn Leu Thr Lys Pro Gly Thr Gly Ile Ser Asp
                85                  90                  95

Gly Asp Phe Trp Ile Gly Leu Trp Arg Asn Gly Asp Gly Gln Thr Ser
            100                 105                 110

Gly Ala Cys Pro Asp Leu Tyr Gln Trp Ser Asp Gly Ser Asn Ser Gln
        115                 120                 125

Tyr Arg Asn Trp Tyr Thr Asp Glu Pro Ser Cys Gly Ser Glu Lys Cys
    130                 135                 140

Val Val Met Tyr His Gln Pro Thr Ala Asn Pro Gly Leu Gly Gly Pro
145                 150                 155                 160

Tyr Leu Tyr Gln Trp Asn Asp Asp Arg Cys Asn Met Lys His Asn Tyr
```

-continued

```
                165                 170                 175

Ile Cys Lys Tyr Glu Pro Glu Ile Asn Pro Thr Ala Pro Val Glu Lys
            180                 185                 190

Pro Tyr Leu Thr Asn Gln Pro Gly Asp Thr His Gln Asn Val Val Val
        195                 200                 205

Thr Glu Ala Gly Ile Ile Pro Asn Leu Ile Tyr Val Val Ile Pro Thr
    210                 215                 220

Ile Pro Leu Leu Leu Leu Ile Leu Val Ala Phe Gly Thr Cys Cys Phe
225                 230                 235                 240

Gln Met Leu His Lys Ser Lys Gly Arg Thr Lys Thr Ser Pro Asn Gln
                245                 250                 255

Ser Thr Leu Trp Ile Ser Lys Ser Thr Arg Lys Glu Ser Gly Met Glu
            260                 265                 270

Val


<210> SEQ ID NO 16
<211> LENGTH: 232
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 16

Met Ala Tyr Phe His Glu Leu Ser Ser Arg Val Ser Phe Gln Glu Ala
 1               5                  10                  15

Arg Leu Ala Cys Glu Ser Glu Gly Gly Val Leu Leu Ser Leu Glu Asn
            20                  25                  30

Glu Ala Glu Gln Lys Leu Ile Glu Ser Met Leu Gln Asn Leu Thr Lys
        35                  40                  45

Pro Gly Thr Gly Ile Ser Asp Gly Asp Phe Trp Ile Gly Leu Trp Arg
    50                  55                  60

Asn Gly Asp Gly Gln Thr Ser Gly Ala Cys Pro Asp Leu Tyr Gln Trp
65                  70                  75                  80

Ser Asp Gly Ser Asn Ser Gln Tyr Arg Asn Trp Tyr Thr Asp Glu Pro
                85                  90                  95

Ser Cys Gly Ser Glu Lys Cys Val Val Met Tyr His Gln Pro Thr Ala
            100                 105                 110

Asn Pro Gly Leu Gly Gly Pro Tyr Leu Tyr Gln Trp Asn Asp Asp Arg
        115                 120                 125

Cys Asn Met Lys His Asn Tyr Ile Cys Lys Tyr Glu Pro Glu Ile Asn
    130                 135                 140

Pro Thr Ala Pro Val Glu Lys Pro Tyr Leu Thr Asn Gln Pro Gly Asp
145                 150                 155                 160

Thr His Gln Asn Val Val Val Thr Glu Ala Gly Ile Ile Pro Asn Leu
                165                 170                 175

Ile Tyr Val Val Ile Pro Thr Ile Pro Leu Leu Leu Leu Ile Leu Val
            180                 185                 190

Ala Phe Gly Thr Cys Cys Phe Gln Met Leu His Lys Ser Lys Gly Arg
        195                 200                 205

Thr Lys Thr Ser Pro Asn Gln Ser Thr Leu Trp Ile Ser Lys Ser Thr
    210                 215                 220

Arg Lys Glu Ser Gly Met Glu Val
225                 230


<210> SEQ ID NO 17
<211> LENGTH: 236
<212> TYPE: PRT
```

```
-continued

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 17

Met Ala Tyr Phe His Glu Leu Ser Ser Arg Val Ser Phe Gln Glu Ala
 1               5                  10                  15

Arg Leu Ala Cys Glu Ser Glu Gly Gly Val Leu Leu Ser Leu Glu Asn
            20                  25                  30

Glu Ala Glu Gln Lys Leu Ile Glu Ser Met Leu Gln Asn Leu Thr Lys
        35                  40                  45

Pro Gly Thr Gly Ile Ser Asp Gly Asp Phe Trp Ile Gly Leu Trp Arg
    50                  55                  60

Asn Gly Asp Gly Gln Thr Ser Gly Ala Cys Pro Asp Leu Tyr Gln Trp
65                  70                  75                  80

Ser Asp Gly Ser Asn Ser Gln Tyr Arg Asn Trp Tyr Thr Asp Glu Pro
                85                  90                  95

Ser Cys Gly Ser Glu Lys Cys Val Val Met Tyr His Gln Pro Thr Ala
            100                 105                 110

Asn Pro Gly Leu Gly Gly Pro Tyr Leu Tyr Gln Trp Asn Asp Asp Arg
        115                 120                 125

Cys Asn Met Lys His Asn Tyr Ile Cys Lys Tyr Glu Pro Glu Ile Asn
    130                 135                 140

Pro Thr Ala Pro Val Glu Lys Pro Tyr Leu Thr Asn Gln Pro Gly Asp
145                 150                 155                 160

Thr His Gln Asn Val Val Val Thr Glu Ala Val Lys Glu Glu Gln Lys
                165                 170                 175

Leu Val Gln Thr Ser Leu His Cys Gly Phe Gln Arg Val Pro Glu Lys
            180                 185                 190

Lys Val Ala Trp Lys Tyr Asn Asn Ser Leu Thr Trp Phe Gln Asn Phe
        195                 200                 205

Val Ile Leu Asp Leu Tyr Lys Glu Trp His Gln Asn Asn Ser Leu Glu
    210                 215                 220

Trp Leu Glu Ile Thr Lys Asp Leu Gln Asp Glu Leu
225                 230                 235


<210> SEQ ID NO 18
<211> LENGTH: 14
<212> TYPE: PRT
<213> ORGANISM: Tetanus toxoid

<400> SEQUENCE: 18

Gln Tyr Ile Lys Ala Asn Ser Lys Phe Ile Gly Ile Thr Glu
 1               5                  10


<210> SEQ ID NO 19
<211> LENGTH: 21
<212> TYPE: PRT
<213> ORGANISM: Plasmodium falciparum

<400> SEQUENCE: 19

Asp Ile Glu Lys Lys Ile Ala Lys Met Glu Lys Ala Ser Ser Val Phe
 1               5                  10                  15

Asn Val Val Asn Ser
            20


<210> SEQ ID NO 20
<211> LENGTH: 16
<212> TYPE: PRT
<213> ORGANISM: Streptococcus
```

-continued

```
<400> SEQUENCE: 20

Gly Ala Val Asp Ser Ile Leu Gly Gly Val Ala Thr Tyr Gly Ala Ala
 1               5                  10                  15


<210> SEQ ID NO 21
<211> LENGTH: 13
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Pan-DR binding epitope
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: 3
<223> OTHER INFORMATION: Xaa = cyclohexyalanine, phenylalanine, or
      tyrosine
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: 1, 13
<223> OTHER INFORMATION: Xaa = d-alanine or l-alanine

<400> SEQUENCE: 21

Xaa Lys Xaa Val Ala Ala Trp Thr Leu Lys Ala Ala Xaa
 1               5                  10


<210> SEQ ID NO 22
<211> LENGTH: 14
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: primer

<400> SEQUENCE: 22

ttttgatcaa gctt                                                    14


<210> SEQ ID NO 23
<211> LENGTH: 42
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: primer

<400> SEQUENCE: 23

ctaatacgac tcactatagg gctcgagcgg ccgcccgggc ag                     42


<210> SEQ ID NO 24
<211> LENGTH: 12
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: primer

<400> SEQUENCE: 24

gatcctgccc gg                                                      12


<210> SEQ ID NO 25
<211> LENGTH: 40
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: primer

<400> SEQUENCE: 25

gtaatacgac tcactatagg gcagcgtggt cgcggccgag                        40


<210> SEQ ID NO 26
<211> LENGTH: 10
```

<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: primer

<400> SEQUENCE: 26

gatcctcggc 10

<210> SEQ ID NO 27
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: primer

<400> SEQUENCE: 27

ctaatacgac tcactatagg gc 22

<210> SEQ ID NO 28
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: primer

<400> SEQUENCE: 28

tcgagcggcc gcccgggcag ga 22

<210> SEQ ID NO 29
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: primer

<400> SEQUENCE: 29

agcgtggtcg cggccgagga 20

<210> SEQ ID NO 30
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: primer

<400> SEQUENCE: 30

atatcgccgc gctcgtcgtc gacaa 25

<210> SEQ ID NO 31
<211> LENGTH: 26
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: primer

<400> SEQUENCE: 31

agccacacgc agctcattgt agaagg 26

<210> SEQ ID NO 32
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: primer

<400> SEQUENCE: 32

-continued

```
cctgcttcag taacaaccac attct                                            25


<210> SEQ ID NO 33
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: primer

<400> SEQUENCE: 33

ctttaccagt ggaatgatga cagg                                             24


<210> SEQ ID NO 34
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: FLAG Tag

<400> SEQUENCE: 34

gattacaagg atgacgacga taag                                             24


<210> SEQ ID NO 35
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 35

aagggacgaa gacgaacacu uctt                                             24


<210> SEQ ID NO 36
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 36

aactgaagac ctgaagacaa taa                                              23


<210> SEQ ID NO 37
<211> LENGTH: 4
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 37

Asn Leu Thr Lys
 1


<210> SEQ ID NO 38
<211> LENGTH: 4
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 38

Asn Pro Thr Ala
 1


<210> SEQ ID NO 39
<211> LENGTH: 9
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 39

Ser Ser Arg Val Ser Phe Gln Glu Ala
 1               5
```

-continued

```
<210> SEQ ID NO 40
<211> LENGTH: 9
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 40

Gly Val Leu Leu Ser Leu Glu Asn Glu
 1               5


<210> SEQ ID NO 41
<211> LENGTH: 9
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 41

Gly Thr Gly Ile Ser Asp Gly Asp Phe
 1               5


<210> SEQ ID NO 42
<211> LENGTH: 9
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 42

Asp Gly Gln Thr Ser Gly Ala Cys Pro
 1               5


<210> SEQ ID NO 43
<211> LENGTH: 9
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 43

Asp Gly Ser Asn Ser Gln Tyr Arg Asn
 1               5


<210> SEQ ID NO 44
<211> LENGTH: 9
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 44

Pro Ser Cys Gly Ser Glu Lys Cys Val
 1               5


<210> SEQ ID NO 45
<211> LENGTH: 9
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 45

Met Leu His Lys Ser Lys Gly Arg Thr
 1               5


<210> SEQ ID NO 46
<211> LENGTH: 9
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 46

Arg Thr Lys Thr Ser Pro Asn Gln Ser
 1               5


<210> SEQ ID NO 47
```

-continued

```
<211> LENGTH: 9
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 47

Thr Arg Lys Glu Ser Gly Met Glu Val
 1               5


<210> SEQ ID NO 48
<211> LENGTH: 9
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 48

Ile Ser Lys Ser Thr Arg Lys Glu Ser
 1               5


<210> SEQ ID NO 49
<211> LENGTH: 9
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 49

Cys Pro Asp Leu Tyr Gln Trp Ser Asp
 1               5


<210> SEQ ID NO 50
<211> LENGTH: 9
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 50

Ser Asn Ser Gln Tyr Arg Asn Trp Tyr
 1               5


<210> SEQ ID NO 51
<211> LENGTH: 9
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 51

Tyr Arg Asn Trp Tyr Thr Asp Glu Pro
 1               5


<210> SEQ ID NO 52
<211> LENGTH: 9
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 52

Met Lys His Asn Tyr Ile Cys Lys Tyr
 1               5


<210> SEQ ID NO 53
<211> LENGTH: 9
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 53

Val Glu Lys Pro Tyr Leu Thr Asn Gln
 1               5


<210> SEQ ID NO 54
<211> LENGTH: 9
<212> TYPE: PRT
```

-continued

```
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 54

Pro Asn Leu Ile Tyr Val Val Ile Pro
 1               5


<210> SEQ ID NO 55
<211> LENGTH: 23
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 55

Ser Met Asx Asn Asp Asn Gly Ser Thr Val Cys Phe Ala Asp Phe Lys
 1               5                  10                  15

His Pro Cys Tyr Lys Met Ala
            20


<210> SEQ ID NO 56
<211> LENGTH: 273
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 56

Met Ser Arg Val Val Ser Leu Leu Leu Gly Ala Ala Leu Leu Cys Gly
 1               5                  10                  15

His Gly Ala Phe Cys Arg Arg Val Val Ser Gly Gln Lys Val Cys Phe
            20                  25                  30

Ala Asp Phe Lys His Pro Cys Tyr Lys Met Ala Tyr Phe His Glu Leu
        35                  40                  45

Ser Ser Arg Val Ser Phe Gln Glu Ala Arg Leu Ala Cys Glu Ser Glu
    50                  55                  60

Gly Gly Val Leu Leu Ser Leu Glu Asn Glu Ala Glu Gln Lys Leu Ile
65                  70                  75                  80

Glu Ser Met Leu Gln Asn Leu Thr Lys Pro Gly Thr Gly Ile Ser Asp
                85                  90                  95

Gly Asp Phe Trp Ile Gly Leu Trp Arg Asn Gly Asp Gly Gln Thr Ser
            100                 105                 110

Gly Ala Cys Pro Asp Leu Tyr Gln Trp Ser Asp Gly Ser Asn Ser Gln
        115                 120                 125

Tyr Arg Asn Trp Tyr Thr Asp Glu Pro Ser Cys Gly Ser Glu Lys Cys
    130                 135                 140

Val Val Met Tyr His Gln Pro Thr Ala Asn Pro Gly Leu Gly Gly Pro
145                 150                 155                 160

Tyr Leu Tyr Gln Trp Asn Asp Asp Arg Cys Asn Met Lys His Asn Tyr
                165                 170                 175

Ile Cys Lys Tyr Glu Pro Glu Ile Asn Pro Thr Ala Pro Val Glu Lys
            180                 185                 190

Pro Tyr Leu Thr Asn Gln Pro Gly Asp Thr His Gln Asn Val Val Val
        195                 200                 205

Thr Glu Ala Gly Ile Ile Pro Asn Leu Ile Tyr Val Val Ile Pro Thr
    210                 215                 220

Ile Pro Leu Leu Leu Leu Ile Leu Val Ala Phe Gly Thr Cys Cys Phe
225                 230                 235                 240

Gln Met Leu His Lys Ser Lys Gly Arg Thr Lys Thr Ser Pro Asn Gln
                245                 250                 255

Ser Thr Leu Trp Ile Ser Lys Ser Thr Arg Lys Glu Ser Gly Met Glu
            260                 265                 270
```

```
Val


<210> SEQ ID NO 57
<211> LENGTH: 74
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 57

Gln Asn Val Val Val Thr Glu Ala Val Lys Glu Glu Gln Lys Leu Val
 1              5                   10                  15

Gln Thr Ser Leu His Cys Gly Phe Gln Arg Val Pro Glu Lys Lys Val
            20                  25                  30

Ala Trp Lys Tyr Asn Asn Ser Leu Thr Trp Phe Gln Asn Phe Val Ile
        35                  40                  45

Leu Asp Leu Tyr Lys Glu Trp His Gln Asn Asn Ser Leu Glu Trp Leu
    50                  55                  60

Glu Ile Thr Lys Asp Leu Gln Asp Glu Leu
65                  70


<210> SEQ ID NO 58
<211> LENGTH: 75
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 58

His Gln Asn Val Val Val Thr Glu Ala Val Lys Glu Glu Gln Lys Leu
 1              5                   10                  15

Val Gln Thr Ser Leu His Cys Gly Phe Gln Arg Val Pro Glu Lys Lys
            20                  25                  30

Val Ala Trp Lys Tyr Asn Asn Ser Leu Thr Trp Phe Gln Asn Phe Val
        35                  40                  45

Ile Leu Asp Leu Tyr Lys Glu Trp His Gln Asn Asn Ser Leu Glu Trp
    50                  55                  60

Leu Glu Ile Thr Lys Asp Leu Gln Asp Glu Leu
65                  70                  75


<210> SEQ ID NO 59
<211> LENGTH: 80
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 59

Gln Pro Gly Asp Thr His Gln Asn Val Val Val Thr Glu Ala Val Lys
 1              5                   10                  15

Glu Glu Gln Lys Leu Val Gln Thr Ser Leu His Cys Gly Phe Gln Arg
            20                  25                  30

Val Pro Glu Lys Lys Val Ala Trp Lys Tyr Asn Asn Ser Leu Thr Trp
        35                  40                  45

Phe Gln Asn Phe Val Ile Leu Asp Leu Tyr Lys Glu Trp His Gln Asn
    50                  55                  60

Asn Ser Leu Glu Trp Leu Glu Ile Thr Lys Asp Leu Gln Asp Glu Leu
65                  70                  75                  80


<210> SEQ ID NO 60
<211> LENGTH: 2418
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 60
```

-continued

ctgtggtgtt tttcccccgc tcctctggct gccttcctga tggatctctg tggtcccagg 60

caggaatggc ctgcttgggg acccagcgag ctcccaaggc ctttcctgct gcttcctcta 120

tccctgtgtt ttgcttggct ctctaaattg actcagctcc aggacatcag gaccccaggt 180

tctctggtct tgggactctg agacttgcac caggaatcct gcccaggctc tcaggccttt 240

ggactcagac tgagctactt cactggcttt cctggttctc cagcttgaag atggcagatc 300

gtgggacttc tcagcctcca taattgagtg agccaattcc ctggccaaaa ggtgtgtttt 360

gctgacttca agcatccctg ctacaaaatg gcctacttcc atgaactgtc cagccgagtg 420

agctttcagg aggcacgcct ggcttgtgag agtgagggag gagtcctcct cagccttgag 480

aatgaagcag aacagaagtt aatagagagc atgttgcaaa acctgacaaa acccgggaca 540

gggatttctg atggtgattt ctggataggg ctttggagga atggagatgg gcaaacatct 600

ggtgcctgcc cagatctcta ccagtggtct gatggaagca attcccagta ccgaaactgg 660

tacacagatg aaccttcctg cggaagtgaa aagtgtgttg tgatgtatca ccaaccaact 720

gccaatcctg gccttggggg tccctacctt taccagtgga atgatgacag gtgtaacatg 780

aagcacaatt atatttgcaa gtatgaacca gagattaatc caacagcccc tgtagaaaag 840

ccttatctta caaatcaacc aggagacacc catcagaatg tggttgttac tgaagcaggt 900

ataattccca atctaattta tgttgttata ccaacaatac ccctgctctt actgatactg 960

gttgcttttg gaacctgttg tttccagatg ctgcataaaa gtaaaggaag aacaaaaact 1020

agtccaaacc agtctacact gtggatttca aagagtacca gaaaagaaag tggcatggaa 1080

gtataataac tcattgactt ggttccagaa ttttgtaatt ctggatctgt ataaggaatg 1140

gcatcagaac aatagcttgg aatggcttga aatcacaaag gatctgcaag atgaactgta 1200

agctccccct tgaggcaaat attaaagtaa tttttatatg tctattattt catttaaaga 1260

atatgctgtg ctaataatgg agtgagacat gcttattttg ctaaaggatg cacccaaact 1320

tcaaacttca agcaaatgaa atggacaatg cagataaagt tgttatcaac acgtcgggag 1380

tatgtgtgtt agaagcaatt ccttttattt ctttcacctt tcataagttg ttatctagtc 1440

aatgtaatgt atattgtatt gaaatttaca gtgtgcaaaa gtattttacc tttgcataag 1500

tgtttgataa aaatgaactg ttctaatatt tatttttatg gcatctcatt tttcaataca 1560

tgctcttttg attaaagaaa cttattactg ttgtcaactg aattcacaca cacacaaata 1620

tagtaccata gaaaaagttt gttttctcga aataattcat ctttcagctt ctctgctttt 1680

ggtcaatgtc taggaaatct cttcagaaat aagaagctat ttcattaagt gtgatataaa 1740

cctcctcaaa cattttactt agaggcaagg attgtctaat ttcaattgtg caagacatgt 1800

gccttataat tatttttagc ttaaaattaa acagattttg taataatgta actttgttaa 1860

taggtgcata aacactaatg cagtcaattt gaacaaaaga agtgacatac acaatataaa 1920

tcatatgtct tcacacgttg cctatataat gagaagcagc tctctgaggg ttctgaaatc 1980

aatgtggtcc ctctcttgcc cactaaacaa agatggttgt tcggggtttg ggattgacac 2040

tggaggcaga tagttgcaaa gttagtctaa ggtttcccta gctgtattta gcctctgact 2100

atattagtat acaaagaggt catgtggttg agaccaggtg aatagtcact atcagtgtgg 2160

agacaagcac agcacacaga cattttagga aggaaaggaa ctacgaaatc gtgtgaaaat 2220

gggttggaac ccatcagtga tcgcatattc attgatgagg gtttgcttga gatagaaaat 2280

ggtggctcct ttctgtctta tctcctagtt tcttcaatgc ttacgccttg ttcttctcaa 2340

-continued

gagaaagttg taactctctg gtcttcatat gtccctgtgc tccttttaac caaataaaga 2400

gttcttgttt ctgaagaa 2418

<210> SEQ ID NO 61
<211> LENGTH: 2076
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 61

ggccaaaagg tgtgttttgc tgacttcaag catccctgct acaaaatggc ctacttccat 60

gaactgtcca gccgagtgag ctttcaggag gcacgcctgg cttgtgagag tgagggagga 120

gtcctcctca gccttgagaa tgaagcagaa cagaagttaa tagagagcat gttgcaaaac 180

ctgacaaaac ccgggacagg gatttctgat ggtgatttct ggatagggct ttggaggaat 240

ggagatgggc aaacatctgg tgcctgccca gatctctacc agtggtctga tggaagcaat 300

tcccagtacc gaaactggta cacagatgaa ccttcctgcg gaagtgaaaa gtgtgttgtg 360

atgtatcacc aaccaactgc caatcctggc cttgggggtc cctaccttta ccagtggaat 420

gatgacaggt gtaacatgaa gcacaattat atttgcaagt atgaaccaga gattaatcca 480

acagcccctg tagaaaagcc ttatcttaca aatcaaccag gagacaccca tcagaatgtg 540

gttgttactg aagcaggtat aattcccaat ctaatttatg ttgttatacc aacaataccc 600

ctgctcttac tgatactggt tgcttttgga acctgttgtt tccagatgct gcataaaagt 660

aaaggaagaa caaaaactag tccaaaccag tctacactgt ggatttcaaa gagtaccaga 720

aaagaaagtg gcatggaagt ataataactc attgacttgg ttccagaatt ttgtaattct 780

ggatctgtat aaggaatggc atcagaacaa tagcttggaa tggcttgaaa tcacaaagga 840

tctgcaagat gaactgtaag ctccccсttg aggcaaatat taaagtaatt tttatatgtc 900

tattatttca tttaaagaat atgctgtgct aataatggag tgagacatgc ttattttgct 960

aaaggatgca cccaaacttc aaacttcaag caaatgaaat ggacaatgca gataaagttg 1020

ttatcaacac gtcgggagta tgtgtgttag aagcaattcc ttttatttct ttcacctttc 1080

ataagttgtt atctagtcaa tgtaatgtat attgtattga aatttacagt gtgcaaaagt 1140

attttacctt tgcataagtg tttgataaaa atgaactgtt ctaatattta tttttatggc 1200

atctcatttt tcaatacatg ctcttttgat taaagaaact tattactgtt gtcaactgaa 1260

ttcacacaca cacaaatata gtaccataga aaaagtttgt tttctcgaaa taattcatct 1320

ttcagcttct ctgcttttgg tcaatgtcta ggaaatctct tcagaaataa gaagctattt 1380

cattaagtgt gatataaacc tcctcaaaca ttttacttag aggcaaggat tgtctaattt 1440

caattgtgca agacatgtgc cttataatta tttttagctt aaaattaaac agattttgta 1500

ataatgtaac tttgttaata ggtgcataaa cactaatgca gtcaatttga acaaaagaag 1560

tgacatacac aatataaatc atatgtcttc acacgttgcc tatataatga gaagcagctc 1620

tctgagggtt ctgaaatcaa tgtggtccct ctcttgccca ctaaacaaag atggttgttc 1680

ggggtttggg attgacactg gaggcagata gttgcaaagt tagtctaagg tttccctagc 1740

tgtatttagc ctctgactat attagtatac aaagaggtca tgtggttgag accaggtgaa 1800

tagtcactat cagtgtggag acaagcacag cacacagaca ttttaggaag gaaaggaact 1860

acgaaatcgt gtgaaaatgg gttggaaccc atcagtgatc gcatattcat tgatgagggt 1920

ttgcttgaga tagaaaatgg tggctccttt ctgtcttatc tcctagtttc ttcaatgctt 1980

acgccttgtt cttctcaaga gaaagttgta actctctggt cttcatatgt ccctgtgctc 2040

-continued

cttttaacca aataaagagt tcttgtttct gaagaa 2076

<210> SEQ ID NO 62
<211> LENGTH: 2418
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 62

ctgtggtgtt tttcccccgc tcctctggct gccttcctga tggatctctg tggtcccagg 60

caggaatggc ctgcttgggg acccagcgag ctcccaaggc ctttcctgct gcttcctcta 120

tccctgtgtt ttgcttggct ctctaaattg actcagctcc aggacatcag gaccccaggt 180

tctctggtct tgggactctg agacttgcac caggaatcct gcccaggctc tcaggccttt 240

ggactcagac tgagctactt cactggcttt cctggttctc cagcttgaag atggcagatc 300

gtgggacttc tcagcctcca taattgagtg agccaattcc ctggccaaaa ggtgtgtttt 360

gctgacttca agcatccctg ctacaaaatg gcctacttcc atgaactgtc cagccgagtg 420

agctttcagg aggcacgcct ggcttgtgag agtgagggag gagtcctcct cagccttgag 480

aatgaagcag aacagaagtt aatagagagc atgttgcaaa acctgacaaa acccgggaca 540

gggatttctg atggtgattt ctggataggg ctttggagga atggagatgg gcaaacatct 600

ggtgcctgcc cagatctcta ccagtggtct gatggaagca attcccagta ccgaaactgg 660

tacacagatg aaccttcctg cggaagtgaa aagtgtgttg tgatgtatca ccaaccaact 720

gccaatcctg gccttggggg tccctacctt taccagtgga atgatgacag gtgtaacatg 780

aagcacaatt atatttgcaa gtatgaacca gagattaatc caacagcccc tgtagaaaag 840

ccttatctta caaatcaacc aggagacacc catcagaatg tggttgttac tgaagcaggt 900

ataattccca atctaattta tgttgttata ccaacaatac ccctgctctt actgatactg 960

gttgcttttg gaacctgttg tttccagatg ctgcataaaa gtaaaggaag aacaaaaact 1020

agtccaaacc agtctacact gtggatttca aagagtacca gaaaagaaag tggcatggaa 1080

gtataataac tcattgactt ggttccagaa ttttgtaatt ctggatctgt ataaggaatg 1140

gcatcagaac aatagcttgg aatggcttga aatcacaaag gatctgcaag atgaactgta 1200

agctccccct tgaggcaaat attaaagtaa tttttatatg tctattattt catttaaaga 1260

atatgctgtg ctaataatgg agtgagacat gcttattttg ctaaaggatg cacccaaact 1320

tcaaacttca agcaaatgaa atggacaatg cagataaagt tgttatcaac acgtcgggag 1380

tatgtgtgtt agaagcaatt ccttttattt ctttcacctt tcataagttg ttatctagtc 1440

aatgtaatgt atattgtatt gaaatttaca gtgtgcaaaa gtattttacc tttgcataag 1500

tgtttgataa aaatgaactg ttctaatatt tatttttatg gcatctcatt tttcaataca 1560

tgctcttttg attaaagaaa cttattactg ttgtcaactg aattcacaca cacacaaata 1620

tagtaccata gaaaaagttt gttttctcga aataattcat ctttcagctt ctctgctttt 1680

ggtcaatgtc taggaaatct cttcagaaat aagaagctat ttcattaagt gtgatataaa 1740

cctcctcaaa cattttactt agaggcaagg attgtctaat ttcaattgtg caagacatgt 1800

gccttataat tatttttagc ttaaaattaa acagattttg taataatgta actttgttaa 1860

taggtgcata aacactaatg cagtcaattt gaacaaaaga agtgacatac acaatataaa 1920

tcatatgtct tcacacgttg cctatataat gagaagcagc tctctgaggg ttctgaaatc 1980

aatgtggtcc ctctcttgcc cactaaacaa agatggttgt tcggggtttg ggattgacac 2040

-continued

```
tggaggcaga tagttgcaaa gttagtctaa ggtttcccta gctgtattta gcctctgact   2100

atattagtat acaaagaggt catgtggttg agaccaggtg aatagtcact atcagtgtgg   2160

agacaagcac agcacacaga cattttagga aggaaaggaa ctacgaaatc gtgtgaaaat   2220

gggttggaac ccatcagtga tcgcatattc attgatgagg gtttgcttga gatagaaaat   2280

ggtggctcct ttctgtctta tctcctagtt tcttcaatgc ttacgccttg ttcttctcaa   2340

gagaaagttg taactctctg gtcttcatat gtccctgtgc tccttttaac caaataaaga   2400

gttcttgttt ctgaagaa                                                 2418


<210> SEQ ID NO 63
<211> LENGTH: 232
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 63

Met Ala Tyr Phe His Glu Leu Ser Ser Arg Val Ser Phe Gln Glu Ala
 1               5                   10                  15

Arg Leu Ala Cys Glu Ser Glu Gly Gly Val Leu Leu Ser Leu Glu Asn
            20                  25                  30

Glu Ala Glu Gln Lys Leu Ile Glu Ser Met Leu Gln Asn Leu Thr Lys
        35                  40                  45

Pro Gly Thr Gly Ile Ser Asp Gly Asp Phe Trp Ile Gly Leu Trp Arg
    50                  55                  60

Asn Gly Asp Gly Gln Thr Ser Gly Ala Cys Pro Asp Leu Tyr Gln Trp
65                  70                  75                  80

Ser Asp Gly Ser Asn Ser Gln Tyr Arg Asn Trp Tyr Thr Asp Glu Pro
                85                  90                  95

Ser Cys Gly Ser Glu Lys Cys Val Val Met Tyr His Gln Pro Thr Ala
            100                 105                 110

Asn Pro Gly Leu Gly Gly Pro Tyr Leu Tyr Gln Trp Asn Asp Asp Arg
        115                 120                 125

Cys Asn Met Lys His Asn Tyr Ile Cys Lys Tyr Glu Pro Glu Ile Asn
    130                 135                 140

Pro Thr Ala Pro Val Glu Lys Pro Tyr Leu Thr Asn Gln Pro Gly Asp
145                 150                 155                 160

Thr His Gln Asn Val Val Val Thr Glu Ala Gly Ile Ile Pro Asn Leu
                165                 170                 175

Ile Tyr Val Val Ile Pro Thr Ile Pro Leu Leu Leu Leu Ile Leu Val
            180                 185                 190

Ala Phe Gly Thr Cys Cys Phe Gln Met Leu His Lys Ser Lys Gly Arg
        195                 200                 205

Thr Lys Thr Ser Pro Asn Gln Ser Thr Leu Trp Ile Ser Lys Ser Thr
    210                 215                 220

Arg Lys Glu Ser Gly Met Glu Val
225                 230


<210> SEQ ID NO 64
<211> LENGTH: 232
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 64

Met Ala Tyr Phe His Glu Leu Ser Ser Arg Val Ser Phe Gln Glu Ala
 1               5                   10                  15

Arg Leu Ala Cys Glu Ser Glu Gly Gly Val Leu Leu Ser Leu Glu Asn
```

```
            20                  25                  30

Glu Ala Glu Gln Lys Leu Ile Glu Ser Met Leu Gln Asn Leu Thr Lys
        35                  40                  45

Pro Gly Thr Gly Ile Ser Asp Gly Asp Phe Trp Ile Gly Leu Trp Arg
    50                  55                  60

Asn Gly Asp Gly Gln Thr Ser Gly Ala Cys Pro Asp Leu Tyr Gln Trp
65                  70                  75                  80

Ser Asp Gly Ser Asn Ser Gln Tyr Arg Asn Trp Tyr Thr Asp Glu Pro
                85                  90                  95

Ser Cys Gly Ser Glu Lys Cys Val Val Met Tyr His Gln Pro Thr Ala
            100                 105                 110

Asn Pro Gly Leu Gly Gly Pro Tyr Leu Tyr Gln Trp Asn Asp Asp Arg
        115                 120                 125

Cys Asn Met Lys His Asn Tyr Ile Cys Lys Tyr Glu Pro Glu Ile Asn
    130                 135                 140

Pro Thr Ala Pro Val Glu Lys Pro Tyr Leu Thr Asn Gln Pro Gly Asp
145                 150                 155                 160

Thr His Gln Asn Val Val Val Thr Glu Ala Gly Ile Ile Pro Asn Leu
                165                 170                 175

Ile Tyr Val Val Ile Pro Thr Ile Pro Leu Leu Leu Leu Ile Leu Val
            180                 185                 190

Ala Phe Gly Thr Cys Cys Phe Gln Met Leu His Lys Ser Lys Gly Arg
        195                 200                 205

Thr Lys Thr Ser Pro Asn Gln Ser Thr Leu Trp Ile Ser Lys Ser Thr
    210                 215                 220

Arg Lys Glu Ser Gly Met Glu Val
225                 230


<210> SEQ ID NO 65
<211> LENGTH: 232
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 65

Met Ala Tyr Phe His Glu Leu Ser Ser Arg Val Ser Phe Gln Glu Ala
 1               5                  10                  15

Arg Leu Ala Cys Glu Ser Glu Gly Gly Val Leu Leu Ser Leu Glu Asn
            20                  25                  30

Glu Ala Glu Gln Lys Leu Ile Glu Ser Met Leu Gln Asn Leu Thr Lys
        35                  40                  45

Pro Gly Thr Gly Ile Ser Asp Gly Asp Phe Trp Ile Gly Leu Trp Arg
    50                  55                  60

Asn Gly Asp Gly Gln Thr Ser Gly Ala Cys Pro Asp Leu Tyr Gln Trp
65                  70                  75                  80

Ser Asp Gly Ser Asn Ser Gln Tyr Arg Asn Trp Tyr Thr Asp Glu Pro
                85                  90                  95

Ser Cys Gly Ser Glu Lys Cys Val Val Met Tyr His Gln Pro Thr Ala
            100                 105                 110

Asn Pro Gly Leu Gly Gly Pro Tyr Leu Tyr Gln Trp Asn Asp Asp Arg
        115                 120                 125

Cys Asn Met Lys His Asn Tyr Ile Cys Lys Tyr Glu Pro Glu Ile Asn
    130                 135                 140

Pro Thr Ala Pro Val Glu Lys Pro Tyr Leu Thr Asn Gln Pro Gly Asp
145                 150                 155                 160
```

-continued

```
Thr His Gln Asn Val Val Val Thr Glu Ala Gly Ile Ile Pro Asn Leu
                165                 170                 175

Ile Tyr Val Val Ile Pro Thr Ile Pro Leu Leu Leu Leu Ile Leu Val
            180                 185                 190

Ala Phe Gly Thr Cys Cys Phe Gln Met Leu His Lys Ser Lys Gly Arg
        195                 200                 205

Thr Lys Thr Ser Pro Asn Gln Ser Thr Leu Trp Ile Ser Lys Ser Thr
    210                 215                 220

Arg Lys Glu Ser Gly Met Glu Val
225                 230


<210> SEQ ID NO 66
<211> LENGTH: 2236
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 66

gcttgaagat ggcagatcgt gggacttctc agcctccata attgagtgag ccaattccct      60

gaatacaaca agaagatggc catctgtaga ccaggaggtg gtccctcccc agaaactgga     120

tgggccagca cgttgattct gagcttctag cctccagaac tgccaaaagg tgtgttttgc     180

tgacttcaag catccctgct acaaaatggc ctacttccat gaactgtcca gccgagtgag     240

ctttcaggag gcacgcctgg cttgtgagag tgagggagga gtcctcctca gccttgagaa     300

tgaagcagaa cagaagttaa tagagagcat gttgcaaaac ctgacaaaac ccgggacagg     360

gatttctgat ggtgatttct ggatagggct ttggaggaat ggagatgggc aaacatctgg     420

tgcctgccca gatctctacc agtggtctga tggaagcaat tcccagtacc gaaactggta     480

cacagatgaa ccttcctgcg gaagtgaaaa gtgtgttgtg atgtatcacc aaccaactgc     540

caatcctggc cttgggggtc cctaccttta ccagtggaat gatgacaggt gtaacatgaa     600

gcacaattat atttgcaagt atgaaccaga gattaatcca acagcccctg tagaaaagcc     660

ttatcttaca aatcaaccag gagacaccca tcagaatgtg gttgttactg aagcaggtat     720

aattcccaat ctaatttatg ttgttatacc aacaataccc ctgctcttac tgatactggt     780

tgcttttgga acctgttgtt tccagatgct gcataaaagt aaaggaagaa caaaaactag     840

tccaaaccag tctacactgt ggatttcaaa gagtaccaga aaagaaagtg gcatggaagt     900

ataataactc attgacttgg ttccagaatt ttgtaattct ggatctgtat aaggaatggc     960

atcagaacaa tagcttggaa tggcttgaaa tcacaaagga tctgcaagat gaactgtaag    1020

ctcccccttg aggcaaatat taaagtaatt tttatatgtc tattatttca tttaaagaat    1080

atgctgtgct aataatggag tgagacatgc ttattttgct aaaggatgca cccaaacttc    1140

aaacttcaag caaatgaaat ggacaatgca gataaagttg ttatcaacac gtcgggagta    1200

tgtgtgttag aagcaattcc ttttatttct ttcacctttc ataagttgtt atctagtcaa    1260

tgtaatgtat attgtattga aatttacagt gtgcaaaagt attttacctt tgcataagtg    1320

tttgataaaa atgaactgtt ctaatattta tttttatggc atctcatttt tcaatacatg    1380

ctcttttgat taaagaaact tattactgtt gtcaactgaa ttcacacaca cacaaatata    1440

gtaccataga aaaagtttgt ttctcgaaa taattcatct ttcagcttct ctgcttttgg    1500

tcaatgtcta ggaaatctct tcagaaataa gaagctattt cattaagtgt gatataaacc    1560

tcctcaaaca ttttacttag aggcaaggat tgtctaattt caattgtgca agacatgtgc    1620

cttataatta tttttagctt aaaattaaac agattttgta ataatgtaac tttgttaata    1680
```

-continued

ggtgcataaa cactaatgca gtcaatttga acaaaagaag tgacatacac aatataaatc 1740

atatgtcttc acacgttgcc tatataatga gaagcagctc tctgagggtt ctgaaatcaa 1800

tgtggtccct ctcttgccca ctaaacaaag atggttgttc ggggtttggg attgacactg 1860

gaggcagata gttgcaaagt tagtctaagg tttccctagc tgtatttagc ctctgactat 1920

attagtatac aaagaggtca tgtggttgag accaggtgaa tagtcactat cagtgtggag 1980

acaagcacag cacacagaca ttttaggaag gaaaggaact acgaaatcgt gtgaaaatgg 2040

gttggaaccc atcagtgatc gcatattcat tgatgagggt ttgcttgaga tagaaaatgg 2100

tggctccttt ctgtcttatc tcctagtttc ttcaatgctt acgccttgtt cttctcaaga 2160

gaaagttgta actctctggt cttcatatgt ccctgtgctc cttttaacca aataaagagt 2220

tcttgtttct gaagaa 2236

<210> SEQ ID NO 67
<211> LENGTH: 2076
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 67

ggccaaaagg tgtgttttgc tgacttcaag catccctgct acaaaatggc ctacttccat 60

gaactgtcca gccgagtgag ctttcaggag gcacgcctgg cttgtgagag tgagggagga 120

gtcctcctca gccttgagaa tgaagcagaa cagaagttaa tagagagcat gttgcaaaac 180

ctgacaaaac ccgggacagg gatttctgat ggtgatttct ggatagggct ttggaggaat 240

ggagatgggc aaacatctgg tgcctgccca gatctctacc agtggtctga tggaagcaat 300

tcccagtacc gaaactggta cacagatgaa ccttcctgcg gaagtgaaaa gtgtgttgtg 360

atgtatcacc aaccaactgc caatcctggc cttggggtc cctaccttta ccagtggaat 420

gatgacaggt gtaacatgaa gcacaattat atttgcaagt atgaaccaga gattaatcca 480

acagcccctg tagaaaagcc ttatcttaca aatcaaccag gagacaccca tcagaatgtg 540

gttgttactg aagcaggtat aattcccaat ctaatttatg ttgttatacc aacaataccc 600

ctgctcttac tgatactggt tgcttttgga acctgttgtt tccagatgct gcataaaagt 660

aaaggaagaa caaaaactag tccaaaccag tctacactgt ggatttcaaa gagtaccaga 720

aaagaaagtg gcatggaagt ataataactc attgacttgg ttccagaatt ttgtaattct 780

ggatctgtat aaggaatggc atcagaacaa tagcttggaa tggcttgaaa tcacaaagga 840

tctgcaagat gaactgtaag ctccccctg aggcaaatat taaagtaatt tttatatgtc 900

tattatttca tttaaagaat atgctgtgct aataatggag tgagacatgc ttattttgct 960

aaaggatgca cccaaacttc aaacttcaag caaatgaaat ggacaatgca gataaagttg 1020

ttatcaacac gtcgggagta tgtgtgttag aagcaattcc ttttatttct ttcacctttc 1080

ataagttgtt atctagtcaa tgtaatgtat attgtattga aatttacagt gtgcaaaagt 1140

attttacctt tgcataagtg tttgataaaa atgaactgtt ctaatattta tttttatggc 1200

atctcatttt tcaatacatg ctcttttgat taaagaaact tattactgtt gtcaactgaa 1260

ttcacacaca cacaaatata gtaccataga aaaagtttgt tttctcgaaa taattcatct 1320

ttcagcttct ctgcttttgg tcaatgtcta ggaaatctct tcagaaataa gaagctattt 1380

cattaagtgt gatataaacc tcctcaaaca ttttacttag aggcaaggat tgtctaattt 1440

caattgtgca agacatgtgc cttataatta tttttagctt aaaattaaac agattttgta 1500

ataatgtaac tttgttaata ggtgcataaa cactaatgca gtcaatttga acaaaagaag 1560

-continued

tgacatacac aatataaatc atatgtcttc acacgttgcc tatataatga gaagcagctc 1620

tctgagggtt ctgaaatcaa tgtggtccct ctcttgccca ctaaacaaag atggttgttc 1680

ggggtttggg attgacactg gaggcagata gttgcaaagt tagtctaagg tttccctagc 1740

tgtatttagc ctctgactat attagtatac aaagaggtca tgtggttgag accaggtgaa 1800

tagtcactat cagtgtggag acaagcacag cacacagaca ttttaggaag gaaaggaact 1860

acgaaatcgt gtgaaaatgg gttggaaccc atcagtgatc gcatattcat tgatgagggt 1920

ttgcttgaga tagaaaatgg tggctccttt ctgtcttatc tcctagtttc ttcaatgctt 1980

acgccttgtt cttctcaaga gaaagttgta actctctggt cttcatatgt ccctgtgctc 2040

cttttaacca aataaagagt tcttgtttct gaagaa 2076

<210> SEQ ID NO 68
<211> LENGTH: 2236
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 68

gcttgaagat ggcagatcgt gggacttctc agcctccata attgagtgag ccaattccct 60

gaatacaaca agaagatggc catctgtaga ccaggaggtg gtccctcccc agaaactgga 120

tgggccagca cgttgattct gagcttctag cctccagaac tgccaaaagg tgtgttttgc 180

tgacttcaag catccctgct acaaaatggc ctacttccat gaactgtcca gccgagtgag 240

ctttcaggag gcacgcctgg cttgtgagag tgagggagga gtcctcctca gccttgagaa 300

tgaagcagaa cagaagttaa tagagagcat gttgcaaaac ctgacaaaac ccgggacagg 360

gatttctgat ggtgatttct ggatagggct ttggaggaat ggagatgggc aaacatctgg 420

tgcctgccca gatctctacc agtggtctga tggaagcaat tcccagtacc gaaactggta 480

cacagatgaa ccttcctgcg gaagtgaaaa gtgtgttgtg atgtatcacc aaccaactgc 540

caatcctggc cttgggggtc cctaccttta ccagtggaat gatgacaggt gtaacatgaa 600

gcacaattat atttgcaagt atgaaccaga gattaatcca acagcccctg tagaaaagcc 660

ttatcttaca aatcaaccag gagacaccca tcagaatgtg gttgttactg aagcaggtat 720

aattcccaat ctaatttatg ttgttatacc aacaataccc ctgctcttac tgatactggt 780

tgcttttgga acctgttgtt tccagatgct gcataaaagt aaaggaagaa caaaaactag 840

tccaaaccag tctacactgt ggatttcaaa gagtaccaga aaagaaagtg gcatggaagt 900

ataataactc attgacttgg ttccagaatt ttgtaattct ggatctgtat aaggaatggc 960

atcagaacaa tagcttggaa tggcttgaaa tcacaaagga tctgcaagat gaactgtaag 1020

ctcccccttg aggcaaatat taaagtaatt tttatatgtc tattatttca tttaaagaat 1080

atgctgtgct aataatggag tgagacatgc ttattttgct aaaggatgca cccaaacttc 1140

aaacttcaag caaatgaaat ggacaatgca gataaagttg ttatcaacac gtcgggagta 1200

tgtgtgttag aagcaattcc ttttatttct ttcacctttc ataagttgtt atctagtcaa 1260

tgtaatgtat attgtattga aatttacagt gtgcaaaagt attttacctt tgcataagtg 1320

tttgataaaa atgaactgtt ctaatattta tttttatggc atctcatttt tcaatacatg 1380

ctcttttgat taaagaaact tattactgtt gtcaactgaa ttcacacaca cacaaatata 1440

gtaccataga aaaagtttgt tttctcgaaa taattcatct ttcagcttct ctgcttttgg 1500

tcaatgtcta ggaaatctct tcagaaataa gaagctattt cattaagtgt gatataaacc 1560

-continued

```
tcctcaaaca ttttacttag aggcaaggat tgtctaattt caattgtgca agacatgtgc   1620

cttataatta tttttagctt aaaattaaac agattttgta ataatgtaac tttgttaata   1680

ggtgcataaa cactaatgca gtcaatttga acaaaagaag tgacatacac aatataaatc   1740

atatgtcttc acacgttgcc tatataatga gaagcagctc tctgagggtt ctgaaatcaa   1800

tgtggtccct ctcttgccca ctaaacaaag atggttgttc ggggtttggg attgacactg   1860

gaggcagata gttgcaaagt tagtctaagg tttccctagc tgtatttagc ctctgactat   1920

attagtatac aaagaggtca tgtggttgag accaggtgaa tagtcactat cagtgtggag   1980

acaagcacag cacacagaca ttttaggaag gaaaggaact acgaaatcgt gtgaaaatgg   2040

gttggaaccc atcagtgatc gcatattcat tgatgagggt ttgcttgaga tagaaaatgg   2100

tggctccttt ctgtcttatc tcctagtttc ttcaatgctt acgccttgtt cttctcaaga   2160

gaaagttgta actctctggt cttcatatgt ccctgtgctc cttttaacca aataaagagt   2220

tcttgtttct gaagaa                                                   2236


<210> SEQ ID NO 69
<211> LENGTH: 232
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 69

Met Ala Tyr Phe His Glu Leu Ser Ser Arg Val Ser Phe Gln Glu Ala
 1               5                  10                  15

Arg Leu Ala Cys Glu Ser Glu Gly Gly Val Leu Leu Ser Leu Glu Asn
            20                  25                  30

Glu Ala Glu Gln Lys Leu Ile Glu Ser Met Leu Gln Asn Leu Thr Lys
        35                  40                  45

Pro Gly Thr Gly Ile Ser Asp Gly Asp Phe Trp Ile Gly Leu Trp Arg
    50                  55                  60

Asn Gly Asp Gly Gln Thr Ser Gly Ala Cys Pro Asp Leu Tyr Gln Trp
65                  70                  75                  80

Ser Asp Gly Ser Asn Ser Gln Tyr Arg Asn Trp Tyr Thr Asp Glu Pro
                85                  90                  95

Ser Cys Gly Ser Glu Lys Cys Val Val Met Tyr His Gln Pro Thr Ala
            100                 105                 110

Asn Pro Gly Leu Gly Gly Pro Tyr Leu Tyr Gln Trp Asn Asp Asp Arg
        115                 120                 125

Cys Asn Met Lys His Asn Tyr Ile Cys Lys Tyr Glu Pro Glu Ile Asn
    130                 135                 140

Pro Thr Ala Pro Val Glu Lys Pro Tyr Leu Thr Asn Gln Pro Gly Asp
145                 150                 155                 160

Thr His Gln Asn Val Val Val Thr Glu Ala Gly Ile Ile Pro Asn Leu
                165                 170                 175

Ile Tyr Val Val Ile Pro Thr Ile Pro Leu Leu Leu Leu Ile Leu Val
            180                 185                 190

Ala Phe Gly Thr Cys Cys Phe Gln Met Leu His Lys Ser Lys Gly Arg
        195                 200                 205

Thr Lys Thr Ser Pro Asn Gln Ser Thr Leu Trp Ile Ser Lys Ser Thr
    210                 215                 220

Arg Lys Glu Ser Gly Met Glu Val
225                 230


<210> SEQ ID NO 70
```

```
<211> LENGTH: 232
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 70

Met Ala Tyr Phe His Glu Leu Ser Ser Arg Val Ser Phe Gln Glu Ala
 1               5                   10                  15

Arg Leu Ala Cys Glu Ser Glu Gly Gly Val Leu Leu Ser Leu Glu Asn
            20                  25                  30

Glu Ala Glu Gln Lys Leu Ile Glu Ser Met Leu Gln Asn Leu Thr Lys
        35                  40                  45

Pro Gly Thr Gly Ile Ser Asp Gly Asp Phe Trp Ile Gly Leu Trp Arg
    50                  55                  60

Asn Gly Asp Gly Gln Thr Ser Gly Ala Cys Pro Asp Leu Tyr Gln Trp
65                  70                  75                  80

Ser Asp Gly Ser Asn Ser Gln Tyr Arg Asn Trp Tyr Thr Asp Glu Pro
                85                  90                  95

Ser Cys Gly Ser Glu Lys Cys Val Val Met Tyr His Gln Pro Thr Ala
            100                 105                 110

Asn Pro Gly Leu Gly Gly Pro Tyr Leu Tyr Gln Trp Asn Asp Asp Arg
        115                 120                 125

Cys Asn Met Lys His Asn Tyr Ile Cys Lys Tyr Glu Pro Glu Ile Asn
    130                 135                 140

Pro Thr Ala Pro Val Glu Lys Pro Tyr Leu Thr Asn Gln Pro Gly Asp
145                 150                 155                 160

Thr His Gln Asn Val Val Val Thr Glu Ala Gly Ile Ile Pro Asn Leu
                165                 170                 175

Ile Tyr Val Val Ile Pro Thr Ile Pro Leu Leu Leu Leu Ile Leu Val
            180                 185                 190

Ala Phe Gly Thr Cys Cys Phe Gln Met Leu His Lys Ser Lys Gly Arg
        195                 200                 205

Thr Lys Thr Ser Pro Asn Gln Ser Thr Leu Trp Ile Ser Lys Ser Thr
    210                 215                 220

Arg Lys Glu Ser Gly Met Glu Val
225                 230


<210> SEQ ID NO 71
<211> LENGTH: 232
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 71

Met Ala Tyr Phe His Glu Leu Ser Ser Arg Val Ser Phe Gln Glu Ala
 1               5                   10                  15

Arg Leu Ala Cys Glu Ser Glu Gly Gly Val Leu Leu Ser Leu Glu Asn
            20                  25                  30

Glu Ala Glu Gln Lys Leu Ile Glu Ser Met Leu Gln Asn Leu Thr Lys
        35                  40                  45

Pro Gly Thr Gly Ile Ser Asp Gly Asp Phe Trp Ile Gly Leu Trp Arg
    50                  55                  60

Asn Gly Asp Gly Gln Thr Ser Gly Ala Cys Pro Asp Leu Tyr Gln Trp
65                  70                  75                  80

Ser Asp Gly Ser Asn Ser Gln Tyr Arg Asn Trp Tyr Thr Asp Glu Pro
                85                  90                  95

Ser Cys Gly Ser Glu Lys Cys Val Val Met Tyr His Gln Pro Thr Ala
            100                 105                 110
```

-continued

```
Asn Pro Gly Leu Gly Gly Pro Tyr Leu Tyr Gln Trp Asn Asp Asp Arg
        115                 120                 125

Cys Asn Met Lys His Asn Tyr Ile Cys Lys Tyr Glu Pro Glu Ile Asn
    130                 135                 140

Pro Thr Ala Pro Val Glu Lys Pro Tyr Leu Thr Asn Gln Pro Gly Asp
145                 150                 155                 160

Thr His Gln Asn Val Val Val Thr Glu Ala Gly Ile Ile Pro Asn Leu
                165                 170                 175

Ile Tyr Val Val Ile Pro Thr Ile Pro Leu Leu Leu Leu Ile Leu Val
            180                 185                 190

Ala Phe Gly Thr Cys Cys Phe Gln Met Leu His Lys Ser Lys Gly Arg
        195                 200                 205

Thr Lys Thr Ser Pro Asn Gln Ser Thr Leu Trp Ile Ser Lys Ser Thr
    210                 215                 220

Arg Lys Glu Ser Gly Met Glu Val
225                 230


<210> SEQ ID NO 72
<211> LENGTH: 2133
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 72

gcttgaagat ggcagatcgt gggacttctc agcctccata attgagtgag ccaattccct     60

gaatacaaca agaagatggc catctgtaga ccaggaggtg gtccctcccc agaaactgga    120

tgggccagca cgttgattct gagcttctag cctccagaac tgccaaaagg tgtgttttgc    180

tgacttcaag catccctgct acaaaatggc ctacttccat gaactgtcca gccgagtgag    240

ctttcaggag gcacgcctgg cttgtgagag tgagggagga gtcctcctca gccttgagaa    300

tgaagcagaa cagaagttaa tagagagcat gttgcaaaac ctgacaaaac ccgggacagg    360

gatttctgat ggtgatttct ggatagggct ttggaggaat ggagatgggc aaacatctgg    420

tgcctgccca gatctctacc agtggtctga tggaagcaat tcccagtacc gaaactggta    480

cacagatgaa ccttcctgcg gaagtgaaaa gtgtgttgtg atgtatcacc aaccaactgc    540

caatcctggc cttgggggtc cctaccttta ccagtggaat gatgacaggt gtaacatgaa    600

gcacaattat atttgcaagt atgaaccaga gattaatcca acagcccctg tagaaaagcc    660

ttatcttaca aatcaaccag gagacaccca tcagaatgtg gttgttactg aagcagtaaa    720

ggaagaacaa aaactagtcc aaaccagtct acactgtgga tttcaaagag taccagaaaa    780

gaaagtggca tggaagtata ataactcatt gacttggttc cagaattttg taattctgga    840

tctgtataag gaatggcatc agaacaatag cttggaatgg cttgaaatca caaaggatct    900

gcaagatgaa ctgtaagctc ccccttgagg caaatattaa agtaattttt atatgtctat    960

tatttcattt aaagaatatg ctgtgctaat aatggagtga gacatgctta ttttgctaaa   1020

ggatgcaccc aaacttcaaa cttcaagcaa atgaaatgga caatgcagat aaagttgtta   1080

tcaacacgtc gggagtatgt gtgttagaag caattccttt tatttctttc acctttcata   1140

agttgttatc tagtcaatgt aatgtatatt gtattgaaat ttacagtgtg caaaagtatt   1200

ttacctttgc ataagtgttt gataaaaatg aactgttcta atatttattt ttatggcatc   1260

tcatttttca atacatgctc ttttgattaa agaaacttat tactgttgtc aactgaattc   1320

acacacacac aaatatagta ccatagaaaa agtttgtttt ctcgaaataa ttcatctttc   1380
```

-continued

agcttctctg cttttggtca atgtctagga aatctcttca gaaataagaa gctatttcat 1440

taagtgtgat ataaacctcc tcaaacattt tacttagagg caaggattgt ctaatttcaa 1500

ttgtgcaaga catgtgcctt ataattattt ttagcttaaa attaaacaga ttttgtaata 1560

atgtaacttt gttaataggt gcataaacac taatgcagtc aatttgaaca aaagaagtga 1620

catacacaat ataaatcata tgtcttcaca cgttgcctat ataatgagaa gcagctctct 1680

gagggttctg aaatcaatgt ggtccctctc ttgcccacta aacaaagatg gttgttcggg 1740

gtttgggatt gacactggag gcagatagtt gcaaagttag tctaaggttt ccctagctgt 1800

atttagcctc tgactatatt agtatacaaa gaggtcatgt ggttgagacc aggtgaatag 1860

tcactatcag tgtggagaca agcacagcac acagacattt taggaaggaa aggaactacg 1920

aaatcgtgtg aaaatgggtt ggaacccatc agtgatcgca tattcattga tgagggtttg 1980

cttgagatag aaaatggtgg ctcctttctg tcttatctcc tagtttcttc aatgcttacg 2040

ccttgttctt ctcaagagaa agttgtaact ctctggtctt catatgtccc tgtgctcctt 2100

ttaaccaaat aaagagttct tgtttctgaa gaa 2133

<210> SEQ ID NO 73
<211> LENGTH: 2076
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 73

ggccaaaagg tgtgttttgc tgacttcaag catccctgct acaaaatggc ctacttccat 60

gaactgtcca gccgagtgag ctttcaggag gcacgcctgg cttgtgagag tgagggagga 120

gtcctcctca gccttgagaa tgaagcagaa cagaagttaa tagagagcat gttgcaaaac 180

ctgacaaaac ccgggacagg gatttctgat ggtgatttct ggatagggct ttggaggaat 240

ggagatgggc aaacatctgg tgcctgccca gatctctacc agtggtctga tggaagcaat 300

tcccagtacc gaaactggta cacagatgaa ccttcctgcg gaagtgaaaa gtgtgttgtg 360

atgtatcacc aaccaactgc caatcctggc cttgggggtc cctaccttta ccagtggaat 420

gatgacaggt gtaacatgaa gcacaattat atttgcaagt atgaaccaga gattaatcca 480

acagcccctg tagaaaagcc ttatcttaca aatcaaccag gagacaccca tcagaatgtg 540

gttgttactg aagcaggtat aattcccaat ctaatttatg ttgttatacc aacaataccc 600

ctgctcttac tgatactggt tgcttttgga acctgttgtt tccagatgct gcataaaagt 660

aaaggaagaa caaaaactag tccaaaccag tctacactgt ggatttcaaa gagtaccaga 720

aaagaaagtg gcatggaagt ataataactc attgacttgg ttccagaatt ttgtaattct 780

ggatctgtat aaggaatggc atcagaacaa tagcttggaa tggcttgaaa tcacaaagga 840

tctgcaagat gaactgtaag ctcccccttg aggcaaatat taaagtaatt tttatatgtc 900

tattatttca tttaaagaat atgctgtgct aataatggag tgagacatgc ttattttgct 960

aaaggatgca cccaaacttc aaacttcaag caaatgaaat ggacaatgca gataaagttg 1020

ttatcaacac gtcgggagta tgtgtgttag aagcaattcc ttttatttct ttcacctttc 1080

ataagttgtt atctagtcaa tgtaatgtat attgtattga aatttacagt gtgcaaaagt 1140

attttacctt tgcataagtg tttgataaaa atgaactgtt ctaatattta tttttatggc 1200

atctcatttt tcaatacatg ctcttttgat taaagaaact tattactgtt gtcaactgaa 1260

ttcacacaca cacaaatata gtaccataga aaaagtttgt tttctcgaaa taattcatct 1320

ttcagcttct ctgcttttgg tcaatgtcta ggaaatctct tcagaaataa gaagctattt 1380

-continued

cattaagtgt gatataaacc tcctcaaaca ttttacttag aggcaaggat tgtctaattt 1440

caattgtgca agacatgtgc cttataatta tttttagctt aaaattaaac agattttgta 1500

ataatgtaac tttgttaata ggtgcataaa cactaatgca gtcaatttga acaaaagaag 1560

tgacatacac aatataaatc atatgtcttc acacgttgcc tatataatga gaagcagctc 1620

tctgagggtt ctgaaatcaa tgtggtccct ctcttgccca ctaaacaaag atggttgttc 1680

ggggtttggg attgacactg gaggcagata gttgcaaagt tagtctaagg tttccctagc 1740

tgtatttagc ctctgactat attagtatac aaagaggtca tgtggttgag accaggtgaa 1800

tagtcactat cagtgtggag acaagcacag cacacagaca ttttaggaag gaaaggaact 1860

acgaaatcgt gtgaaaatgg gttggaaccc atcagtgatc gcatattcat tgatgagggt 1920

ttgcttgaga tagaaaatgg tggctccttt ctgtcttatc tcctagtttc ttcaatgctt 1980

acgccttgtt cttctcaaga gaaagttgta actctctggt cttcatatgt ccctgtgctc 2040

cttttaacca aataaagagt tcttgtttct gaagaa 2076

<210> SEQ ID NO 74
<211> LENGTH: 2133
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 74

gcttgaagat ggcagatcgt gggacttctc agcctccata attgagtgag ccaattccct 60

gaatacaaca agaagatggc catctgtaga ccaggaggtg gtccctcccc agaaactgga 120

tgggccagca cgttgattct gagcttctag cctccagaac tgccaaaagg tgtgttttgc 180

tgacttcaag catccctgct acaaaatggc ctacttccat gaactgtcca gccgagtgag 240

ctttcaggag gcacgcctgg cttgtgagag tgagggagga gtcctcctca gccttgagaa 300

tgaagcagaa cagaagttaa tagagagcat gttgcaaaac ctgacaaaac ccgggacagg 360

gatttctgat ggtgatttct ggatagggct ttggaggaat ggagatgggc aaacatctgg 420

tgcctgccca gatctctacc agtggtctga tggaagcaat tcccagtacc gaaactggta 480

cacagatgaa ccttcctgcg gaagtgaaaa gtgtgttgtg atgtatcacc aaccaactgc 540

caatcctggc cttgggggtc cctaccttta ccagtggaat gatgacaggt gtaacatgaa 600

gcacaattat atttgcaagt atgaaccaga gattaatcca acagcccctg tagaaaagcc 660

ttatcttaca aatcaaccag gagacaccca tcagaatgtg gttgttactg aagcagtaaa 720

ggaagaacaa aaactagtcc aaaccagtct acactgtgga tttcaaagag taccagaaaa 780

gaaagtggca tggaagtata ataactcatt gacttggttc cagaattttg taattctgga 840

tctgtataag gaatggcatc agaacaatag cttggaatgg cttgaaatca caaaggatct 900

gcaagatgaa ctgtaagctc ccccttgagg caaatattaa agtaattttt atatgtctat 960

tatttcattt aaagaatatg ctgtgctaat aatggagtga gacatgctta ttttgctaaa 1020

ggatgcaccc aaacttcaaa cttcaagcaa atgaaatgga caatgcagat aaagttgtta 1080

tcaacacgtc gggagtatgt gtgttagaag caattccttt tatttctttc acctttcata 1140

agttgttatc tagtcaatgt aatgtatatt gtattgaaat ttacagtgtg caaaagtatt 1200

ttacctttgc ataagtgttt gataaaaatg aactgttcta atatttattt ttatggcatc 1260

tcatttttca atacatgctc ttttgattaa agaaacttat tactgttgtc aactgaattc 1320

acacacacac aaatatagta ccatagaaaa agtttgtttt ctcgaaataa ttcatctttc 1380

-continued

```
agcttctctg cttttggtca atgtctagga aatctcttca gaaataagaa gctatttcat   1440

taagtgtgat ataaacctcc tcaaacattt tacttagagg caaggattgt ctaatttcaa   1500

ttgtgcaaga catgtgcctt ataattattt ttagcttaaa attaaacaga ttttgtaata   1560

atgtaacttt gttaataggt gcataaacac taatgcagtc aatttgaaca aaagaagtga   1620

catacacaat ataaatcata tgtcttcaca cgttgcctat ataatgagaa gcagctctct   1680

gagggttctg aaatcaatgt ggtccctctc ttgcccacta aacaaagatg gttgttcggg   1740

gtttgggatt gacactggag gcagatagtt gcaaagttag tctaaggttt ccctagctgt   1800

atttagcctc tgactatatt agtatacaaa gaggtcatgt ggttgagacc aggtgaatag   1860

tcactatcag tgtggagaca agcacagcac acagacattt taggaaggaa aggaactacg   1920

aaatcgtgtg aaaatgggtt ggaacccatc agtgatcgca tattcattga tgagggtttg   1980

cttgagatag aaaatggtgg ctcctttctg tcttatctcc tagtttcttc aatgcttacg   2040

ccttgttctt ctcaagagaa agttgtaact ctctggtctt catatgtccc tgtgctcctt   2100

ttaaccaaat aaagagttct tgtttctgaa gaa                                2133


<210> SEQ ID NO 75
<211> LENGTH: 236
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 75

Met Ala Tyr Phe His Glu Leu Ser Ser Arg Val Ser Phe Gln Glu Ala
 1               5                  10                  15

Arg Leu Ala Cys Glu Ser Glu Gly Gly Val Leu Leu Ser Leu Glu Asn
            20                  25                  30

Glu Ala Glu Gln Lys Leu Ile Glu Ser Met Leu Gln Asn Leu Thr Lys
        35                  40                  45

Pro Gly Thr Gly Ile Ser Asp Gly Asp Phe Trp Ile Gly Leu Trp Arg
    50                  55                  60

Asn Gly Asp Gly Gln Thr Ser Gly Ala Cys Pro Asp Leu Tyr Gln Trp
65                  70                  75                  80

Ser Asp Gly Ser Asn Ser Gln Tyr Arg Asn Trp Tyr Thr Asp Glu Pro
                85                  90                  95

Ser Cys Gly Ser Glu Lys Cys Val Val Met Tyr His Gln Pro Thr Ala
            100                 105                 110

Asn Pro Gly Leu Gly Gly Pro Tyr Leu Tyr Gln Trp Asn Asp Asp Arg
        115                 120                 125

Cys Asn Met Lys His Asn Tyr Ile Cys Lys Tyr Glu Pro Glu Ile Asn
    130                 135                 140

Pro Thr Ala Pro Val Glu Lys Pro Tyr Leu Thr Asn Gln Pro Gly Asp
145                 150                 155                 160

Thr His Gln Asn Val Val Val Thr Glu Ala Val Lys Glu Glu Gln Lys
                165                 170                 175

Leu Val Gln Thr Ser Leu His Cys Gly Phe Gln Arg Val Pro Glu Lys
            180                 185                 190

Lys Val Ala Trp Lys Tyr Asn Asn Ser Leu Thr Trp Phe Gln Asn Phe
        195                 200                 205

Val Ile Leu Asp Leu Tyr Lys Glu Trp His Gln Asn Asn Ser Leu Glu
    210                 215                 220

Trp Leu Glu Ile Thr Lys Asp Leu Gln Asp Glu Leu
225                 230                 235
```

```
<210> SEQ ID NO 76
<211> LENGTH: 232
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 76

Met Ala Tyr Phe His Glu Leu Ser Ser Arg Val Ser Phe Gln Glu Ala
 1               5                  10                  15

Arg Leu Ala Cys Glu Ser Glu Gly Gly Val Leu Leu Ser Leu Glu Asn
            20                  25                  30

Glu Ala Glu Gln Lys Leu Ile Glu Ser Met Leu Gln Asn Leu Thr Lys
        35                  40                  45

Pro Gly Thr Gly Ile Ser Asp Gly Asp Phe Trp Ile Gly Leu Trp Arg
    50                  55                  60

Asn Gly Asp Gly Gln Thr Ser Gly Ala Cys Pro Asp Leu Tyr Gln Trp
65                  70                  75                  80

Ser Asp Gly Ser Asn Ser Gln Tyr Arg Asn Trp Tyr Thr Asp Glu Pro
                85                  90                  95

Ser Cys Gly Ser Glu Lys Cys Val Val Met Tyr His Gln Pro Thr Ala
            100                 105                 110

Asn Pro Gly Leu Gly Gly Pro Tyr Leu Tyr Gln Trp Asn Asp Asp Arg
        115                 120                 125

Cys Asn Met Lys His Asn Tyr Ile Cys Lys Tyr Glu Pro Glu Ile Asn
    130                 135                 140

Pro Thr Ala Pro Val Glu Lys Pro Tyr Leu Thr Asn Gln Pro Gly Asp
145                 150                 155                 160

Thr His Gln Asn Val Val Val Thr Glu Ala Gly Ile Ile Pro Asn Leu
                165                 170                 175

Ile Tyr Val Val Ile Pro Thr Ile Pro Leu Leu Leu Leu Ile Leu Val
            180                 185                 190

Ala Phe Gly Thr Cys Cys Phe Gln Met Leu His Lys Ser Lys Gly Arg
        195                 200                 205

Thr Lys Thr Ser Pro Asn Gln Ser Thr Leu Trp Ile Ser Lys Ser Thr
    210                 215                 220

Arg Lys Glu Ser Gly Met Glu Val
225                 230


<210> SEQ ID NO 77
<211> LENGTH: 236
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 77

Met Ala Tyr Phe His Glu Leu Ser Ser Arg Val Ser Phe Gln Glu Ala
 1               5                  10                  15

Arg Leu Ala Cys Glu Ser Glu Gly Gly Val Leu Leu Ser Leu Glu Asn
            20                  25                  30

Glu Ala Glu Gln Lys Leu Ile Glu Ser Met Leu Gln Asn Leu Thr Lys
        35                  40                  45

Pro Gly Thr Gly Ile Ser Asp Gly Asp Phe Trp Ile Gly Leu Trp Arg
    50                  55                  60

Asn Gly Asp Gly Gln Thr Ser Gly Ala Cys Pro Asp Leu Tyr Gln Trp
65                  70                  75                  80

Ser Asp Gly Ser Asn Ser Gln Tyr Arg Asn Trp Tyr Thr Asp Glu Pro
                85                  90                  95
```

-continued

```
Ser Cys Gly Ser Glu Lys Cys Val Val Met Tyr His Gln Pro Thr Ala
            100                 105                 110

Asn Pro Gly Leu Gly Gly Pro Tyr Leu Tyr Gln Trp Asn Asp Asp Arg
        115                 120                 125

Cys Asn Met Lys His Asn Tyr Ile Cys Lys Tyr Glu Pro Glu Ile Asn
    130                 135                 140

Pro Thr Ala Pro Val Glu Lys Pro Tyr Leu Thr Asn Gln Pro Gly Asp
145                 150                 155                 160

Thr His Gln Asn Val Val Val Thr Glu Ala Val Lys Glu Glu Gln Lys
                165                 170                 175

Leu Val Gln Thr Ser Leu His Cys Gly Phe Gln Arg Val Pro Glu Lys
            180                 185                 190

Lys Val Ala Trp Lys Tyr Asn Asn Ser Leu Thr Trp Phe Gln Asn Phe
        195                 200                 205

Val Ile Leu Asp Leu Tyr Lys Glu Trp His Gln Asn Asn Ser Leu Glu
    210                 215                 220

Trp Leu Glu Ile Thr Lys Asp Leu Gln Asp Glu Leu
225                 230                 235


<210> SEQ ID NO 78
<211> LENGTH: 2033
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 78

gcttgaagat ggcagatcgt gggacttctc agcctccata attgagtgag ccaattccct     60

ggccaaaagg tgtgttttgc tgacttcaag catccctgct acaaaatggc ctacttccat    120

gaactgtcca gccgagtgag ctttcaggag gcacgcctgg cttgtgagag tgagggagga    180

gtcctcctca gccttgagaa tgaagcagaa cagaagttaa tagagagcat gttgcaaaac    240

ctgacaaaac ccgggacagg gatttctgat ggtgatttct ggatagggct ttggaggaat    300

ggagatgggc aaacatctgg tgcctgccca gatctctacc agtggtctga tggaagcaat    360

tcccagtacc gaaactggta cacagatgaa ccttcctgcg gaagtgaaaa gtgtgttgtg    420

atgtatcacc aaccaactgc caatcctggc cttgggggtc cctaccttta ccagtggaat    480

gatgacaggt gtaacatgaa gcacaattat atttgcaagt atgaaccaga gattaatcca    540

acagcccctg tagaaaagcc ttatcttaca aatcaaccag gagacaccca tcagaatgtg    600

gttgttactg aagcagtaaa ggaagaacaa aaactagtcc aaaccagtct acactgtgga    660

tttcaaagag taccagaaaa gaaagtggca tggaagtata ataactcatt gacttggttc    720

cagaattttg taattctgga tctgtataag gaatggcatc agaacaatag cttggaatgg    780

cttgaaatca caaaggatct gcaagatgaa ctgtaagctc ccccttgagg caaatattaa    840

agtaattttt atatgtctat tatttcattt aaagaatatg ctgtgctaat aatggagtga    900

gacatgctta ttttgctaaa ggatgcaccc aaacttcaaa cttcaagcaa atgaaatgga    960

caatgcagat aaagttgtta tcaacacgtc gggagtatgt gtgttagaag caattccttt   1020

tatttctttc acctttcata agttgttatc tagtcaatgt aatgtatatt gtattgaaat   1080

ttacagtgtg caaaagtatt ttacctttgc ataagtgttt gataaaaatg aactgttcta   1140

atatttattt ttatggcatc tcatttttca atacatgctc ttttgattaa agaaacttat   1200

tactgttgtc aactgaattc acacacacac aaatatagta ccatagaaaa agtttgtttt   1260

ctcgaaataa ttcatctttc agcttctctg cttttggtca atgtctagga aatctcttca   1320
```

-continued

gaaataagaa gctatttcat taagtgtgat ataaacctcc tcaaacattt tacttagagg 1380

caaggattgt ctaatttcaa ttgtgcaaga catgtgcctt ataattattt ttagcttaaa 1440

attaaacaga ttttgtaata atgtaacttt gttaataggt gcataaacac taatgcagtc 1500

aatttgaaca aaagaagtga catacacaat ataaatcata tgtcttcaca cgttgcctat 1560

ataatgagaa gcagctctct gagggttctg aaatcaatgt ggtccctctc ttgcccacta 1620

aacaaagatg gttgttcggg gtttgggatt gacactggag gcagatagtt gcaaagttag 1680

tctaaggttt ccctagctgt atttagcctc tgactatatt agtatacaaa gaggtcatgt 1740

ggttgagacc aggtgaatag tcactatcag tgtggagaca agcacagcac acagacattt 1800

taggaaggaa aggaactacg aaatcgtgtg aaaatgggtt ggaacccatc agtgatcgca 1860

tattcattga tgagggtttg cttgagatag aaaatggtgg ctcctttctg tcttatctcc 1920

tagtttcttc aatgcttacg ccttgttctt ctcaagagaa agttgtaact ctctggtctt 1980

catatgtccc tgtgctcctt ttaaccaaat aaagagttct tgtttctgaa gaa 2033

<210> SEQ ID NO 79
<211> LENGTH: 2076
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 79

ggccaaaagg tgtgttttgc tgacttcaag catccctgct acaaaatggc ctacttccat 60

gaactgtcca gccgagtgag ctttcaggag gcacgcctgg cttgtgagag tgagggagga 120

gtcctcctca gccttgagaa tgaagcagaa cagaagttaa tagagagcat gttgcaaaac 180

ctgacaaaac ccgggacagg gatttctgat ggtgatttct ggatagggct ttggaggaat 240

ggagatgggc aaacatctgg tgcctgccca gatctctacc agtggtctga tggaagcaat 300

tcccagtacc gaaactggta cacagatgaa ccttcctgcg gaagtgaaaa gtgtgttgtg 360

atgtatcacc aaccaactgc caatcctggc cttgggggtc cctaccttta ccagtggaat 420

gatgacaggt gtaacatgaa gcacaattat atttgcaagt atgaaccaga gattaatcca 480

acagcccctg tagaaaagcc ttatcttaca aatcaaccag gagacaccca tcagaatgtg 540

gttgttactg aagcaggtat aattcccaat ctaatttatg ttgttatacc aacaataccc 600

ctgctcttac tgatactggt tgcttttgga acctgttgtt tccagatgct gcataaaagt 660

aaaggaagaa caaaaactag tccaaaccag tctacactgt ggatttcaaa gagtaccaga 720

aaagaaagtg gcatggaagt ataataactc attgacttgg ttccagaatt ttgtaattct 780

ggatctgtat aaggaatggc atcagaacaa tagcttggaa tggcttgaaa tcacaaagga 840

tctgcaagat gaactgtaag ctcccccttg aggcaaatat taaagtaatt tttatatgtc 900

tattatttca tttaaagaat atgctgtgct aataatggag tgagacatgc ttattttgct 960

aaaggatgca cccaaacttc aaacttcaag caaatgaaat ggacaatgca gataaagttg 1020

ttatcaacac gtcgggagta tgtgtgttag aagcaattcc ttttatttct ttcacctttc 1080

ataagttgtt atctagtcaa tgtaatgtat attgtattga aatttacagt gtgcaaaagt 1140

attttacctt tgcataagtg tttgataaaa atgaactgtt ctaatattta tttttatggc 1200

atctcatttt tcaatacatg ctcttttgat taaagaaact tattactgtt gtcaactgaa 1260

ttcacacaca cacaaatata gtaccataga aaaagtttgt tttctcgaaa taattcatct 1320

ttcagcttct ctgcttttgg tcaatgtcta ggaaatctct tcagaaataa gaagctattt 1380

cattaagtgt gatataaacc tcctcaaaca ttttacttag aggcaaggat tgtctaattt 1440

-continued

caattgtgca agacatgtgc cttataatta tttttagctt aaaattaaac agattttgta 1500

ataatgtaac tttgttaata ggtgcataaa cactaatgca gtcaatttga acaaaagaag 1560

tgacatacac aatataaatc atatgtcttc acacgttgcc tatataatga gaagcagctc 1620

tctgagggtt ctgaaatcaa tgtggtccct ctcttgccca ctaaacaaag atggttgttc 1680

ggggtttggg attgacactg gaggcagata gttgcaaagt tagtctaagg tttccctagc 1740

tgtatttagc ctctgactat attagtatac aaagaggtca tgtggttgag accaggtgaa 1800

tagtcactat cagtgtggag acaagcacag cacacagaca ttttaggaag gaaaggaact 1860

acgaaatcgt gtgaaaatgg gttggaaccc atcagtgatc gcatattcat tgatgagggt 1920

ttgcttgaga tagaaaatgg tggctccttt ctgtcttatc tcctagtttc ttcaatgctt 1980

acgccttgtt cttctcaaga gaaagttgta actctctggt cttcatatgt ccctgtgctc 2040

cttttaacca aataaagagt tcttgtttct gaagaa 2076

<210> SEQ ID NO 80
<211> LENGTH: 2033
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 80

gcttgaagat ggcagatcgt gggacttctc agcctccata attgagtgag ccaattccct 60

ggccaaaagg tgtgttttgc tgacttcaag catccctgct acaaaatggc ctacttccat 120

gaactgtcca gccgagtgag ctttcaggag gcacgcctgg cttgtgagag tgagggagga 180

gtcctcctca gccttgagaa tgaagcagaa cagaagttaa tagagagcat gttgcaaaac 240

ctgacaaaac ccgggacagg gatttctgat ggtgatttct ggatagggct ttggaggaat 300

ggagatgggc aaacatctgg tgcctgccca gatctctacc agtggtctga tggaagcaat 360

tcccagtacc gaaactggta cacagatgaa ccttcctgcg gaagtgaaaa gtgtgttgtg 420

atgtatcacc aaccaactgc caatcctggc cttgggggtc cctaccttta ccagtggaat 480

gatgacaggt gtaacatgaa gcacaattat atttgcaagt atgaaccaga gattaatcca 540

acagcccctg tagaaaagcc ttatcttaca aatcaaccag gagacaccca tcagaatgtg 600

gttgttactg aagcagtaaa ggaagaacaa aaactagtcc aaaccagtct acactgtgga 660

tttcaaagag taccagaaaa gaaagtggca tggaagtata ataactcatt gacttggttc 720

cagaattttg taattctgga tctgtataag gaatggcatc agaacaatag cttggaatgg 780

cttgaaatca caaaggatct gcaagatgaa ctgtaagctc cccttgagg caaatattaa 840

agtaattttt atatgtctat tatttcattt aaagaatatg ctgtgctaat aatggagtga 900

gacatgctta ttttgctaaa ggatgcaccc aaacttcaaa cttcaagcaa atgaaatgga 960

caatgcagat aaagttgtta tcaacacgtc gggagtatgt gtgttagaag caattccttt 1020

tatttctttc acctttcata agttgttatc tagtcaatgt aatgtatatt gtattgaaat 1080

ttacagtgtg caaaagtatt ttacctttgc ataagtgttt gataaaaatg aactgttcta 1140

atatttattt ttatggcatc tcatttttca atacatgctc ttttgattaa agaaacttat 1200

tactgttgtc aactgaattc acacacacac aaatatagta ccatagaaaa agtttgtttt 1260

ctcgaaataa ttcatctttc agcttctctg cttttggtca atgtctagga aatctcttca 1320

gaaataagaa gctatttcat taagtgtgat ataaacctcc tcaaacattt tacttagagg 1380

caaggattgt ctaatttcaa ttgtgcaaga catgtgcctt ataattattt ttagcttaaa 1440

-continued

```
attaaacaga ttttgtaata atgtaacttt gttaataggt gcataaacac taatgcagtc   1500

aatttgaaca aaagaagtga catacacaat ataaatcata tgtcttcaca cgttgcctat   1560

ataatgagaa gcagctctct gagggttctg aaatcaatgt ggtccctctc ttgcccacta   1620

aacaaagatg gttgttcggg gtttgggatt gacactggag gcagatagtt gcaaagttag   1680

tctaaggttt ccctagctgt atttagcctc tgactatatt agtatacaaa gaggtcatgt   1740

ggttgagacc aggtgaatag tcactatcag tgtggagaca agcacagcac acagacattt   1800

taggaaggaa aggaactacg aaatcgtgtg aaaatgggtt ggaacccatc agtgatcgca   1860

tattcattga tgagggtttg cttgagatag aaaatggtgg ctcctttctg tcttatctcc   1920

tagtttcttc aatgcttacg ccttgttctt ctcaagagaa agttgtaact ctctggtctt   1980

catatgtccc tgtgctcctt ttaaccaaat aaagagttct tgtttctgaa gaa           2033


<210> SEQ ID NO 81
<211> LENGTH: 236
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 81

Met Ala Tyr Phe His Glu Leu Ser Ser Arg Val Ser Phe Gln Glu Ala
 1               5                  10                  15

Arg Leu Ala Cys Glu Ser Glu Gly Gly Val Leu Leu Ser Leu Glu Asn
            20                  25                  30

Glu Ala Glu Gln Lys Leu Ile Glu Ser Met Leu Gln Asn Leu Thr Lys
        35                  40                  45

Pro Gly Thr Gly Ile Ser Asp Gly Asp Phe Trp Ile Gly Leu Trp Arg
    50                  55                  60

Asn Gly Asp Gly Gln Thr Ser Gly Ala Cys Pro Asp Leu Tyr Gln Trp
65                  70                  75                  80

Ser Asp Gly Ser Asn Ser Gln Tyr Arg Asn Trp Tyr Thr Asp Glu Pro
                85                  90                  95

Ser Cys Gly Ser Glu Lys Cys Val Val Met Tyr His Gln Pro Thr Ala
            100                 105                 110

Asn Pro Gly Leu Gly Gly Pro Tyr Leu Tyr Gln Trp Asn Asp Asp Arg
        115                 120                 125

Cys Asn Met Lys His Asn Tyr Ile Cys Lys Tyr Glu Pro Glu Ile Asn
    130                 135                 140

Pro Thr Ala Pro Val Glu Lys Pro Tyr Leu Thr Asn Gln Pro Gly Asp
145                 150                 155                 160

Thr His Gln Asn Val Val Val Thr Glu Ala Val Lys Glu Glu Gln Lys
                165                 170                 175

Leu Val Gln Thr Ser Leu His Cys Gly Phe Gln Arg Val Pro Glu Lys
            180                 185                 190

Lys Val Ala Trp Lys Tyr Asn Asn Ser Leu Thr Trp Phe Gln Asn Phe
        195                 200                 205

Val Ile Leu Asp Leu Tyr Lys Glu Trp His Gln Asn Asn Ser Leu Glu
    210                 215                 220

Trp Leu Glu Ile Thr Lys Asp Leu Gln Asp Glu Leu
225                 230                 235


<210> SEQ ID NO 82
<211> LENGTH: 170
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
```

```
<400> SEQUENCE: 82

Met Ala Tyr Phe His Glu Leu Ser Ser Arg Val Ser Phe Gln Glu Ala
 1               5                  10                  15

Arg Leu Ala Cys Glu Ser Glu Gly Gly Val Leu Leu Ser Leu Glu Asn
            20                  25                  30

Glu Ala Glu Gln Lys Leu Ile Glu Ser Met Leu Gln Asn Leu Thr Lys
        35                  40                  45

Pro Gly Thr Gly Ile Ser Asp Gly Asp Phe Trp Ile Gly Leu Trp Arg
    50                  55                  60

Asn Gly Asp Gly Gln Thr Ser Gly Ala Cys Pro Asp Leu Tyr Gln Trp
65                  70                  75                  80

Ser Asp Gly Ser Asn Ser Gln Tyr Arg Asn Trp Tyr Thr Asp Glu Pro
                85                  90                  95

Ser Cys Gly Ser Glu Lys Cys Val Val Met Tyr His Gln Pro Thr Ala
            100                 105                 110

Asn Pro Gly Leu Gly Gly Pro Tyr Leu Tyr Gln Trp Asn Asp Asp Arg
        115                 120                 125

Cys Asn Met Lys His Asn Tyr Ile Cys Lys Tyr Glu Pro Glu Ile Asn
    130                 135                 140

Pro Thr Ala Pro Val Glu Lys Pro Tyr Leu Thr Asn Gln Pro Gly Asp
145                 150                 155                 160

Thr His Gln Asn Val Val Val Thr Glu Ala
                165                 170


<210> SEQ ID NO 83
<211> LENGTH: 236
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 83

Met Ala Tyr Phe His Glu Leu Ser Ser Arg Val Ser Phe Gln Glu Ala
 1               5                  10                  15

Arg Leu Ala Cys Glu Ser Glu Gly Gly Val Leu Leu Ser Leu Glu Asn
            20                  25                  30

Glu Ala Glu Gln Lys Leu Ile Glu Ser Met Leu Gln Asn Leu Thr Lys
        35                  40                  45

Pro Gly Thr Gly Ile Ser Asp Gly Asp Phe Trp Ile Gly Leu Trp Arg
    50                  55                  60

Asn Gly Asp Gly Gln Thr Ser Gly Ala Cys Pro Asp Leu Tyr Gln Trp
65                  70                  75                  80

Ser Asp Gly Ser Asn Ser Gln Tyr Arg Asn Trp Tyr Thr Asp Glu Pro
                85                  90                  95

Ser Cys Gly Ser Glu Lys Cys Val Val Met Tyr His Gln Pro Thr Ala
            100                 105                 110

Asn Pro Gly Leu Gly Gly Pro Tyr Leu Tyr Gln Trp Asn Asp Asp Arg
        115                 120                 125

Cys Asn Met Lys His Asn Tyr Ile Cys Lys Tyr Glu Pro Glu Ile Asn
    130                 135                 140

Pro Thr Ala Pro Val Glu Lys Pro Tyr Leu Thr Asn Gln Pro Gly Asp
145                 150                 155                 160

Thr His Gln Asn Val Val Val Thr Glu Ala Val Lys Glu Glu Gln Lys
                165                 170                 175

Leu Val Gln Thr Ser Leu His Cys Gly Phe Gln Arg Val Pro Glu Lys
            180                 185                 190
```

-continued

```
Lys Val Ala Trp Lys Tyr Asn Asn Ser Leu Thr Trp Phe Gln Asn Phe
        195                 200                 205

Val Ile Leu Asp Leu Tyr Lys Glu Trp His Gln Asn Asn Ser Leu Glu
    210                 215                 220

Trp Leu Glu Ile Thr Lys Asp Leu Gln Asp Glu Leu
225                 230                 235
```

The invention claimed is:

1. A method for detecting expression levels of a 58P1D12 protein consisting of the sequence of SEQ ID NO: 13 or SEQ ID NO: 14 in a biological sample and normal sample, comprising:

determining expression levels of the 58P1D12 protein in the biological sample and the normal sample; and

comparing the expression levels of the 58P1D12 protein detected in the biological sample and the normal sample.

2. The method of claim **1**, wherein the biological sample is from a patient who has or is suspected of having a cancer selected from the group consisting of prostate cancer, bladder cancer, lung cancer, ovarian cancer, bone cancer, and cervical cancer.

3. The method of claim **1**, wherein the biological sample is from serum, blood or urine, or tissue selected from the group consisting of prostate, bladder, lung, ovary, bone, and cervix.

4. The method of claim **1**, wherein the presence of elevated 58P1D12 protein in the biological sample relative to the normal sample provides an indication of the presence of cancer.

5. The method of claim **1**, wherein the 58P1D12 protein consists of the sequence of SEQ ID NO: 13.

6. The method of claim **1**, wherein the 58P1D12 protein consists of the sequence of SEQ ID NO: 14.

7. A method for detecting expression levels of a 58P1D12 transcription variant polynucleotide encoding a protein consisting of the sequence of SEQ ID NO: 13 or SEQ ID NO: 14 in a biological sample and a corresponding normal sample, comprising:

determining expression levels of the 58P1D12 transcription variant polynucleotide in the biological sample and the corresponding normal sample; and

comparing the expression levels of the 58P1D12 transcription variant polynucleotide detected in the biological sample and the corresponding normal sample, wherein the biological sample is from a patient who has or is suspected of having a cancer selected from the group consisting of prostate cancer, bladder cancer, lung cancer, ovarian cancer, bone cancer, and cervical cancer.

8. The method of claim **7**, wherein the biological sample is from serum, blood or urine, or tissue selected from the group consisting of prostate, bladder, lung, ovary, bone, and cervix.

9. The method of claim **7**, wherein the presence of elevated 58P1D12 transcription variant polynucleotide in the biological sample relative to the corresponding normal sample provides an indication of the presence of cancer.

10. The method of claim **7**, wherein the 58P1D12 transcription variant polynucleotide comprises the sequence of SEQ ID NO: 4.

11. The method of claim **7**, wherein the 58P1D12 transcription variant polynucleotide comprises the sequence of SEQ ID NO: 8.

* * * * *

专利名称(译)	名为58P1D12的核酸和相应蛋白质可用于治疗和检测癌症		
公开(公告)号	US7842458	公开(公告)日	2010-11-30
申请号	US10/919654	申请日	2004-08-16
[标]申请(专利权)人(译)	坎纳史蒂芬乙 JAKOBOVITS AYA CHALLITA EID PIA中号 GE WANGMAO RAITANO ARTHUR乙		
申请(专利权)人(译)	坎纳史蒂芬B. JAKOBOVITS AYA challita PIA的M -开斋节. GE WANGMAO RAITANO ARTHUR B.		
当前申请(专利权)人(译)	AGENSYS INC.		
[标]发明人	KANNER STEVEN B JAKOBOVITS AYA CHALLITA EID PIA M GE WANGMAO RAITANO ARTHUR B		
发明人	KANNER, STEVEN B. JAKOBOVITS, AYA CHALLITA-EID, PIA M. GE, WANGMAO RAITANO, ARTHUR B.		
IPC分类号	C12Q1/68 G01N33/53 G01N33/574 C07H21/04 C07K14/435 A61K38/00 A61K39/00 C07K14/705		
CPC分类号	C07K14/7056 C07K16/30 C12Q1/6881 G01N33/574 G01N33/57434 A01K2217/05 A61K38/00 A61K39/00 A61K2039/505 C07K2317/73 C12N2799/026 C12N2799/027 G01N2500/04 A61P35/00 C12Q2600/136 C12Q2600/156 C12Q2600/158		
代理机构(译)	美富律师事务所		
优先权	60/148935 1999-08-12 US		
其他公开文献	US20050136435A1		
外部链接	Espacenet USPTO		

摘要(译)

描述了新基因58P1D12及其编码的蛋白质及其变体，其中58P1D12在正常成人组织中表现出组织特异性表达，并且在表I中列出的癌症中异常表达。因此，58P1D12提供诊断，预后，预防和/或或癌症的治疗靶点。 58P1D12基因或其片段，或其编码的蛋白质，或其变体，或其片段，可用于引发体液或细胞免疫应答;与58P1D12反应的抗体或T细胞可用于主动或被动免疫。

FIGURE 1: 58P1D12 SSH SEQUENCE OF 427 NUCLEOTIDES. (SEQ ID NO: 1).

```
  1 CAGAATTACA AAATTCTGGA ACCAAGTCAA TGAGTTATTA TACTTCCATG CCACTTTCTT
 61 TTCTGGTACT CTTTGAAATC CACAGTGTAG ACTGGTTTGG ACTAGTTTTT GTTCTTCCTT
121 TACTTTTATG CAGCATCTGG AAACAACAGG TTCCAAAAGC AACCAGTATC AGTAAGAGCA
181 GGGGTATTGT TGGTATAACA ACATAAATTA GATTGGGAAT TATACCTGCT TCAGTAACAA
241 CCACATTCTG ATGGGTGTCT CCTGGTTGAT TTGTAAGATA AGGCTTTTCT ACAGGGGCTG
301 TTGGATTAAT CTCTGGTTCA TACTTGCAAA TATAATTGTG CTTCATGTTA CACCTGTCAT
361 CATTCCACTG GTAAAGGTAG GGACCCCCAA GGCCAGGATT GGCAGTTGGT TGGTGATCAT
421 CACAACA
```