



US 20030143632A1

(19) **United States**(12) **Patent Application Publication**
Holtzman et al.(10) **Pub. No.: US 2003/0143632 A1**(43) **Pub. Date: Jul. 31, 2003**(54) **NOVEL GENES ENCODING PROTEINS
HAVING PROGNOSTIC, DIAGNOSTIC,
PREVENTIVE, THERAPEUTIC AND OTHER
USES**(75) Inventors: **Douglas A. Holtzman**, Jamaica Plain,
MA (US); **Andrew D.J. Goodearl**,
Natick, MA (US); **Sean A. McCarthy**,
San Diego, CA (US)

Correspondence Address:

MILLENNIUM PHARMACEUTICALS, INC.
75 Sidney Street
Cambridge, MA 02139 (US)(73) Assignee: **Millennium Pharmaceuticals, Inc.**(21) Appl. No.: **10/365,227**(22) Filed: **Feb. 12, 2003****Related U.S. Application Data**

- (60) Continuation of application No. 09/802,582, filed on Mar. 8, 2001, now abandoned.
Continuation-in-part of application No. 09/128,709, filed on Aug. 4, 1998, now abandoned.
Said application No. 09/802,582 is a continuation-in-part of application No. 09/130,491, filed on Aug. 6, 1998, now Pat. No. 6,416,974.
Said application No. 09/802,582 is a continuation-in-part of application No. 09/388,280, filed on Sep. 1, 1999, and which is a division of application No. 09/130,491, filed on Aug. 6, 1998, now Pat. No. 6,416,974.
Said application No. 09/802,582 is a continuation-in-part of application No. 09/388,279, filed on Sep. 2, 1999, now abandoned, and which is a division of application No. 09/130,491, filed on Aug. 6, 1998, now Pat. No. 6,416,974.

- (60) Provisional application No. 60/054,645, filed on Aug. 4, 1997. Provisional application No. 60/054,966, filed on Aug. 6, 1997. Provisional application No. 60/058,108, filed on Sep. 5, 1997. Provisional application No. 60/054,966, filed on Aug. 6, 1997. Provisional application No. 60/058,108, filed on Sep. 5, 1997. Provisional application No. 60/054,966, filed on Aug. 6, 1997. Provisional application No. 60/058,108, filed on Sep. 5, 1997.

Publication Classification

- (51) **Int. Cl.⁷** **G01N 33/53**; C07H 21/04;
C12P 21/02; C12N 5/06; C07K 14/705
(52) **U.S. Cl.** **435/7.1**; 435/69.1; 435/320.1;
435/325; 530/350; 536/23.5

(57) **ABSTRACT**

The invention relates to Tango-73, Tango-74, Tango-76, Tango-78, and Tango-83 polypeptides, nucleic acid molecules encoding Tango-73, Tango-74, Tango-76, Tango-78, and Tango-83, and uses thereof. The invention provides isolated nucleic acids encoding a variety of proteins having diagnostic, preventive, therapeutic, and other uses. These nucleic acids and proteins are useful for diagnosis, prevention, and therapy of a number of human and other animal disorders. The invention also provides antisense nucleic acid molecules, expression vectors containing the nucleic acid molecules of the invention, host cells into which the expression vectors have been introduced, and non-human transgenic animals in which a nucleic acid molecule of the invention has been introduced or disrupted. The invention still further provides isolated polypeptides, fusion polypeptides, antigenic peptides and antibodies. Diagnostic, screening, and therapeutic methods using compositions of the invention are also provided. The nucleic acids and polypeptides of the present invention are useful as modulating agents in regulating a variety of cellular processes.

FIG. 1A

R	Q	L	G	G	S	T	L	L	W	E	A	E	S	S	W	R	A	Q	E	91
AGG	CAG	CTG	GGT	GGG	TCC	ACC	TTG	CTG	TGG	GAA	GCC	GAG	AGC	TCC	TGG	CGG	GCC	CAG	GAG	273
G	Q	L	S	W	E	W	G	K	R	H	R	R	H	H	L	P	D	R	S	111
GGA	CAG	CTG	TCC	TGG	GAG	TGG	GGC	AAG	AGG	CAC	CGT	CGA	CAT	CAC	TTG	CCA	GAC	AGA	AGT	333
Q	L	C	R	K	V	K	F	Q	V	D	F	N	L	I	G	W	G	S	W	131
CAA	CTG	TGT	CGG	AAG	GTC	AAC	TTC	CAG	GTG	GAC	TTC	AAC	CTG	ATC	GGA	TGG	GGC	TCC	TGG	393
I	I	Y	P	K	Q	Y	N	A	Y	R	C	E	G	E	C	P	N	P	V	151
ATC	ATC	TAC	CCC	AAG	CAG	TAC	AAC	GCC	TAT	CGC	TGT	GAG	GGC	GAG	TGT	CCT	AAT	CCT	GTT	453
G	E	E	F	H	P	T	N	H	A	Y	I	Q	V	G	C	Q	A			169
GGG	GAG	GAG	TTT	CAT	CCG	ACC	AAC	CAT	GCA	TAC	ATC	CAG	GTG	GGA	TGC	CAG	GCG			507

T

FIG. 1B

T78	1	MDLFTVTLSQVTFSLG	16
		: . .	
nodal	101	PMDIPTEGPLTIDIFEQAKGDPERDPADCLERIWMETFTVIPSQVTFASG	150
			
T78	17	SMVLEVTRPLSKWLKRPGALEKQMSRVAGECWPRPPTPP...ATNVLLML	63
		: : . : :	
nodal	151	STVLEVTKPLSKWLKDPRALEKQVSSRAEKCWHQPYTPPPVPVASTNVLM	200
			
T78	64	YSNLSQEQRQLGGSTLLWEAESSWRAQEGQLSWE...WGKRHRRHHLDPDR	110
		 : :	
nodal	201	YSNRPQEQRQLGGATLLWEAESSWRAQEGQLSVERGGWGRRRHHLDPDR	250
			
T78	111	SQLCRKVKFQVDFNLIGWSWIIYPKQYNAYRCEGECPNPVGEEFHPNTNH	160
		:	
nodal	251	SQLCRRVKFQVDFNLIGWSWIIYPKQYNAYRCEGECPNPVGEEFHPNTNH	300
			
T78	161	AYIQV...GCQA.....	300
		: :	
nodal	301	AYIQSLLKRYQPHRVPSTCCAPVKTPLSMLYVDNGRVLLEHHKDMIVEE	350

GTCGACCCACGGCTCCGGGAGCAACCGCAGCTCTAGTATCCAGACTCCAGCGCCCGCCCGCGCGGAGCCCAACCC 79

CGACCCAGAGCTCTCCAGCGCGCGCCAGCGAGCGGCTCCCGCCCTTACTTCTCCCGCGCGCCAGCCACCTTC 158

GGGAGTCCGGGTGCCCCACCTGCAAACTCTCCGCTTCTGCACCTGCCACCCCTGAGCCAGCGCGCGCGCCAGCGGAG 237

M A N A G L Q L L G F I L A F L G W I 19
TC ATG GCC AAC GCG GGG CTG CAG CTG TTG GGC TTC ATT CTC GCC TTC CTG GGA TGG ATC 296

G A I V S T A L P Q W R I Y S Y A G D N 39
GGC GCC ATC GTC AGC ACT GCC CTG CCC CAG TGG AGG ATT TAC TCC TAT GCC GCC GAC AAC 356

I V T A Q A M Y E G L W M S C V S Q S T 59
ATC GTG ACC GCC CAG GCC ATG TAC GAG GGG CTG TGG ATG TCC TCC GTG TCG CAG AGC ACC 416

G Q I Q C K V F D S L L N L S S T L Q A 79
GGG CAG ATC CAG TGC AAA GTC TTT GAC TCC TTG CTG AAT CTG AGC AGC ACA TTG CAA GCA 476

T R A L M V V G I L L G V I A I F V A T 99
ACC CGT GCC TTG ATG GTG GTT GGC ATC CTC CTG GGA GTG ATA GCA ATC TTT GTG GCC ACC 536

V G M K C M K C L E D D E V Q K M R M A 119
GTT GGC ATG AAG TGT ATG AAG TGC TTG GAA GAC GAT GAG GTG CAG AAG ATG AGG ATG GCT 596

FIG. 3A

V I G G A I F L L A G L A I L V A T A W 139
GTC ATT GGG GGT GCG ATA TTT CTT CTT CCA GGT CTG GCT ATT TTA GTT GCC ACA GCA TGG 656

Y G N R I V Q E F Y D P M T P V N A R Y 159
TAT GGC AAT AGA ATC GTT CAA GAA TTC TAT GAC CCT ATG ACC CCA GTC AAT GCC AGG TAC 716

E F G Q A L F T G W A A A S L C L L G G 179
GAA TTT GGT CAG GCT CTC TTC ACT GGC TGG GCT GCT GCT TCT CTC TGC CTT CTG GGA GGT 776

A L L C C S C P R K T T S Y P T P R P Y 199
GCC CTA CTT TGC TGT TCC TCC CCC CCA AAA ACA ACC TCT TAC CCA ACA CCA AGG CCC TAT 836

P K P A P S S G K D Y V * 212
CCA AAA CCT GCA CCT TCC ACC GGG AAA GAC TAC GTG TGA 875

CACAGAGGCAAAAGGAGAAATCATGTTGAAACAAACCGAAATGGACATTGAGATACATCATTAACATTAGGACCTT 954

AGAAATTTGGGTATGTAACTCTGAAGTATGGTATTACAAACAAACAAACAAACAAACCCATGTGTFAAAATACTC 1033

AGTGCTAAACATGGCTTAACTTATTTTATCTTCTTTCCTTCAATATAGGAGGAGATTTTTCCTATTTGTATTACTGCT 1112

TCCCATTTGAGTAATCATCTCACTGGGGGAGGGGTCCTTAATATATATAGATATGTATATATACATGTTTTTC 1191

TATTAATAATAGACAGTAAATACTATTCTTCATTAAGTTGATACTAGCATACTTAAATAATCTCTRAAATAGGTAAATG 1270

TATTTAATTCATATTGATGAAGAAGTTTATTGGTATATTTTCTTTTTCGTCCTATATATACATAATGTAACAGTCAATA 1349

FIG. 3B

[illegible]

FIG. 3C

2376 CAAACCTACGCACATACCTTCATGTGGCTCAGTGGCTTCCTCTCTCTACAGTCTATTTCCATTCTTTCAGCTGTGTC
2455 TGACATGTTTGTGCTCTGTTCOCATTTTAAACAACTGCTCTTACCTTTCAGTCTGTACAGAAATGCTATTTTCACCTTGAGCA
2534 AGATGATGTAATGGAAGGGTGTGGCAATGGTGTCTCTGGAGACCTGGATTTTCAGTCTTGGTGTATCAATCACCGTCTG
2613 TGTTTGAGCAAGGCATTTGGCTGCTGTAAAGCTTATTTGCTTCACTGTAAAGCGGTGTTTGTAAATTCCTGATCTTTCACC
2692 CTCACAGTGAATGTGTGGGATCCAGTGAATAGAAATACATGTAAAGTGTGTTTGTAAATTTTAAAGTGTCTATACTAA
2771 GGGAAAGAAATTCAGGAATTAACCTGCATACGTTTTCAGTGTGTGCTTTTCAAAATGTTTTCGAAACAAAAAATGTTAAGAAA
2850 TGGGTTTCTTGCCTTAACCAAGTCTCTCAAGTGAATGAGACAGTGAAGTAAATTTAGTGCCTAAACAAATTAAGATTCCTG
2929 AGGAAGTCTTAATCTTCTGCAAGTGAATGCCCCGATGCTTTCCTGCTTAAACAGATGTAATGGGAAGAAATAAAAGCC
3008 TACGTGTTGGTAAATCCAAACAGCAAGGAGATTTTTCGAATCAATAAATCACTCAATAGGTGCTATCTGTTTCAGTCAATGCC
3087 TCAGAGCTCTTGTGTTAGCTGGCAGCTGACGGCTGCTAGGATAGTTAGTTTGGAAATGGTACTTTCATAATAAACTACAC
3166 AAGGAAGTCAGCCACTGTGTCTTATGAGGAATTTGGACCTAATAAATTTTAGTGTGCCCTTCCAAACCTTGAGAAATATATG
3245 CTTTTGGAAGTTAAATTTAAATGGCTTTTTCGCCACATACATAGATCTTCAATGATGTGTGAGTGTAAATTCATGTGGATA
3324 TCAGTTACCAACATTAACAAAAATTTTATGGCCCCAAATGACCAACGAATGTGTACAAATAGAAATTTTATCCAAATTTT
3403 GATCTTTTATATTTCTTACCACACCTCTGGAACAGACCAATAGACATTTTGGGGTTTTATATAGGAATTTGTATATAA
3482 GCATTAAGCTTTTCAATAAATTTGTTTTTTTAAATTTTAAAAAAGGAAAAAATAAAAAAAGGGCGGCGG

FIG. 3D

ccagactcca ccaccgcta ccgggaccag aagccaggag cctcgccccg cagctgcaca 60
gagagcaagg gtataggcac taacttgttt gcagagaccc catcaccttc gggagctcag 120
gtgcgcacct tgcaaaactcc actttctgca tctgccactg agcccgggg agcctcggaa 180
agagcc atg gcc aac gcg ggg ctg cag ctg ggt ttc atc ctg gct 228
Met Ala Asn Ala Gly Leu Gln Leu Gly Phe Ile Leu Ala
1 5 10
tct ctg gga tgg atc ggc tcc atc gtc agc act gcc ctg ccc cag tgg 276
Ser Leu Gly Trp Ile Gly Ser Ile Val Ser Thr Ala Leu Pro Gln Trp 30
15 20 25 30
aag att tac tcc tat gct ggt ggg gac aac atc gtg acc gct cag gcc atc 324
Lys Ile Tyr Ser Tyr Ala Gly Asp Asn Ile Val Thr Ala Gln Ala Ile 45
35 40 45
tac gag gga ctg tgg atg tcc tgc gtt tcg caa agc acc ggg cag ata 372
Tyr Glu Gly Leu Trp Met Ser Cys Val Ser Gln Ser Thr Gly Gln Ile 60
50 55 60
cag tgc aaa gtc ttc gac tcc ttg ctg aat ctg aac agt act ttg cag 420
Gln Cys Lys Val Phe Asp Ser Leu Leu Asn Leu Ser Thr Leu Gln 75
65 70 75

FIG. 4A

gca acc cga gcc ttg atg gta att ggc atc ctg ggg ctg gca 468
Ala Thr Arg Ala Leu Met Val Ile Gly Ile Leu Leu Gly Leu Ile Ala
80 85 90

atc ttt gtg tcc acc att ggc atg aag tgc atg agg tgc ctg gaa gat 516
Ile Phe Val Ser Thr Ile Gly Met Lys Cys Met Arg Cys Leu Glu Asp
95 100 105 110

gat gag gtg cag aag atg ttg atg gct gtc att ggg ggc ata ata ttt 564
Asp Glu Val Gln Lys Met Trp Met Ala Val Ile Gly Gly Ile Ile Phe
115 120 125

tta att tca ggt ctg gcg aca tta gtg gcc aca gca tgg tat gga aac 612
Leu Ile Ser Gly Leu Ala Thr Leu Val Ala Thr Ala Trp Tyr Gly Asn
130 135 140

aga att gtt caa gaa ttc tat gac ccc ttg acc ccc atc aat gcc agg 660
Arg Ile Val Gln Glu Phe Tyr Asp Pro Leu Thr Pro Ile Asn Ala Arg
145 150 155

tat gaa ttt ggc cag gcc ctc ttt act ggc ttg ggc gct gcc tcc ctc 708
Tyr Glu Phe Gly Gln Ala Leu Phe Thr Gly Trp Ala Ala Ser Leu
160 165 170

FIG. 4B

```
tgc ctt ctg gga ggt gtc cta ctt tcc tgc tcc tgt ccc cgg aaa aca 756
Cys Leu Leu Gly Gly Val Leu Ser Cys Ser Cys Pro Arg Lys Thr 190
175 180 185

acc tct tac cca aca cga cgt tat ccc aag cca aca cct tct agt 804
Thr Ser Tyr Pro Thr Pro Arg Pro Tyr Pro Lys Pro Thr Pro Ser Ser 205
195 200

ggg aaa gac tat gtg tgacagaggc aaaggaagag atcttcctgg agcaaataca 859
Gly Lys Asp Tyr Val
210

aaatggacat tgaacctagg attgacatta acgccttaga ctgttgatga tggttatcgg 919
aactgtggta gaacagaagg aagcatatctt ttatacatcc ccattggctat gcaggccttg 979
gctgtacctt accatcttcc ctgagcacag gaggaaggc ttttgcctgt gaactgctgc 1039
ttccctctga gaaatcacac tcaaacgggg ataaggctgt ccttgcattgt gtatagatat 1099
gtacagatac atagtttcta ttaaaaatag aaccccttat tctcctcata 1159
ctgtaccagc acactttaaa tgactctaca atatatacaa ttatgttttg attaaaaaa 1219
aaaaaaaaa aaaaa 1234
```

FIG. 4C

GTCGACCCACGGCTCCGGCTCCGAGAACCTTTCACGGCCACAACTACGGGACGATTTCTGATTGATTTTGGCGCT 79
TTTCGATCCACCCCTCCTCCCTTCTC ATG GGA CTT TCG GGA CAA AGC GTC CCG ACC GCC TOG AGC 13
M G L W G Q S V P T A S S 142
A R A G R Y P G A R T A S G T R P W L L 33
GCT CGA GCA GGG CGC TAT CCA GGA GCC AGG ACA GCG TCG GGA ACC AGA CCA TGG CTC CTG 202
D S K I L K F V V F I V A V L L P V R V 53
GAC TCC AAG ATC CTT AAG TTC GTC GTC TTC ATC GTC CCG GTT CTG CTG CCG GTC CGG GTT 262
D S A T I P R Q D E V P Q Q T V A P Q Q 73
GAC TCT GCC ACC ATC CCC CGG CAG GAC GAA GTT CCC CAG CAG ACA GTG GCC CCA CAG CAA 322
Q R R S L K E E C P A G S H R S E Y T 93
CAG AGG CGC AGC CTC AAG GAG GAG GAG TGT CCA GCA GGA TCT CAT AGA TCA GAA TAT ACT 382
G A C N P C T E G V D Y T I A S N N L P 113
GGA GCC TGT AAC CCG TGC ACA GAG GGT GTG GAT TAC ACC ATT GCT TCC AAC AAT TTG CCT 442
S C L L C T V C K S G Q T N K S S C T T 133
TCT TCC CTG CTA TGT ACA GTT TGT AAA TCA GGT CAA ACA AAT AAA AGT TCC TGT ACC ACG 502

FIG. 5A

T R D T V C Q C E K G S F Q D K N S P E 153
ACC AGA GAC ACC GTG TGT CAG TGT GAA AAA GGA AGC TTC CAG GAT AAA AAC TCC CCT GAG 562

M C R T C R T G C P R G M V K V S N C T 173
ATG TGC CGG ACG TGT AGA ACA GGG TGT CCC AGA GGG ATG GTC AAG GTC AGT AAT TGT ACG 622

P R S D I K C K N E S A A S S T G K T P 193
CCC CGG AGT GAC ATC AAG TGC AAA AAT GAA TCA GCT GCC AGT TCC ACT GGG AAA ACC CCA 682

A A E E T V T T I L G M L A S P Y H Y L 213
GCA GCG GAG GAG ACA GTG ACC ACC ATC CTG GGG ATG CTT GCC TCT CCC TAT CAC TAC CTT 742

I I I V V L V I I L A V V V G F S C R 233
ATC ATC ATA GTG GTT TTA GTC ATC ATT TTA GCT GTG GTT GTG GTT GGC TTT TCA TGT CCG 802

K K F I S Y L K G I C S G G G G P E R 253
AAG AAA TTC ATT TCT TAC CTC AAA GGC ATC TGC TCA GGT GGA GGA GGT CCC GAA CGT 862

V H R V L F R R R S C P S R V P G A E D 273
GTG CAC AGA GTC CTT TTC CGG CGG CGT TCA TGT CCT TCA CGA GTT CCT GCG GCG GAG GAC 922

N A R N E T L S N R Y L Q P T Q V S E Q 293
AAT GCC CGC AAC GAG ACC CTG AGT AAC AGA TAC TTG CAG CCC ACC CAG GTC TCT GAG CAG 982

FIG. 5B

E	I	Q	G	Q	E	L	A	E	L	T	G	V	T	V	E	S	P	E	E	313	
GAA	ATC	CAA	GGT	CAG	GAG	CTG	GCA	GAG	CTA	ACA	GGT	GTG	ACT	GTA	GAG	TCG	CCA	GAG	GAG	1042	
P	Q	R	L	L	E	Q	A	E	A	E	G	C	Q	R	R	R	L	L	V	333	
CCA	CAG	CGT	CTG	CTG	GAA	CAG	GCA	GAA	GCT	GAA	GGG	TGT	CAG	AGG	AGG	CTG	CTG	GTT	1102		
P	V	N	D	A	D	S	A	D	I	S	T	L	L	D	A	S	A	T	L	353	
CCA	GTG	AAT	GAC	GCT	GAC	TCC	GCT	GAC	ATC	AGC	ACC	TTG	CTG	GAT	GCC	TCG	GCA	ACA	CTG	1162	
E	E	G	H	A	K	E	T	I	Q	D	Q	L	V	G	S	E	K	L	F	373	
GAA	GAA	GGA	CAT	GCA	AAG	GAA	ACA	ATT	CAG	GAC	CAA	CTG	GTG	GGC	TCC	GAA	AAG	CTC	TTT	1222	
Y	E	E	D	E	A	G	S	A	T	S	C	L	.							387	
TAT	GAA	GAA	GAT	GAA	GCA	GGC	TCT	GC	CG	TCC	TCC	CTG	TGA							1264	
A	A	G	A	T	C	T	C	T	C	T	C	T	C	T	A	A	G	G	A	A	1343
AGT	A	C	T	G	A	C	C	C	A	A	C	T	C	T	A	T	G	G	G	C	1422
ACT	T	G	G	A	G	T	A	T	G	A	A	C	G	G	A	A	T	T	A	T	1501
AT	A	C	G	A	T	T	A	G	G	T	C	G	T	C	A	T	G	C	C	A	1580
GC	A	G	G	T	G	A	T	T	G	A	G	A	C	C	A	G	T	G	A	A	1659

FIG. 5C

AAAAGTGGTTTAGGATGTCATTCCTTTCAGTTCTTTCATCATGAGACAAGTCTTTTCTCTGCTTCTTATATTGCAAGC 1738
TCCATCTCTACTGGTGTGTCATTTAATGACATCTAACTACAGATGCCGCCACAGCCACAATGCTTTGCTTATATAATTTT 1817
TTAACTTTAGAACGGGATTATCTTGTATTACCTGTATTTTTCAGTTTTCGGATATTTTTCGACTTAAATGATGAGATTATCA 1896
AGACGTAGCCCTATGCTAAGTCATGAGCATATGGAAGTTTACGAGGGTTCGACTTACGAGTTTTCGAGCTTTAAGATACGATT 1975
ATTGGGGCTTACCCCACTTAATTAGAGAAACATTTATATTCCTTACTACTGTAGGCTGTACATCTCTTTTCCGATTT 2054
TTGTATAATGATGTAAACATGGAAAACCTTTAGGAAATGCCACTTATTAGGCTGTATTACATGGGTTCCTGGATACAAAT 2133
CAGCAGTCAAAAATGACTAAAATATACTAGTACGGAGGAGAAATCCTCCTCTGTGCGAGGCACCTTACTGCAATTC 2212
CAGTTCTCCTCCTGCGCCCTGAGACTGACAGGGTTTGATGGCTGGCAGCTTCTCAAGGGCAGCTTGTCTTACTTTC 2291
TTAATTTTAGAGGTATATAGCCATATTTTATTTATAAATAATATTTTATTTTATAAGTAGATGTTTACATATG 2370
CCCAGGATTTTGAAGAGCCTGGTATCTTTTGGGAAGCCATGTGTCTGTTTGTCTGCTGCGGACAGTCAITGGGACTGCA 2449
CTTCCGACTTGTCCACAGCAGATGAGGACAGTGAGAAATTAAGTTAGATCCCGAGACTGCGAAGAGCTTCTCTTTCAAGCG 2528
CCATTACAGTTGAACGTTAGTGAATCTTTGAGCCTCATTTTGGGCTCAGGGCAGAGCAGGTGTTTATCTGCCCGGCATCT 2607
GCCATGGCATCAAGAGGGAAGAGTGGACGGTCTTGGGAATGGTGTGAAATGTTTCCGACTCAGGCCATGGATGGGCCC 2686
CTCTCGCTTCTGGTGTCTGTGAACCTGAGTCCCTGGGATGCCCTTTTAGGGCAGAGATTCTGTAGCTGCGTTTATGGGTA 2765

FIG. 5D

2844 CAGATTCCCTGTTTGAGGAGCTTGGCCCTCTGTAAAGCATCTGACTCATCTCAGAGATATCAATTCTTAAACACTGTGA
2923 CAACAGGATCTAAATGGCTGACACATTTGTCCTMGTCACGTTCCATTATTTTATTTAAACCTCAGTAATCGTTT
3002 TAGCTTCTTTCCAGCAAACTCTTCTCCACAGTAGCCAGTCGCTGGTAGGATAAATTACGGATATAGTCATTCTAGGGGT
3081 TTCAGTCTTTMCCATCTCAAGGCATTGTGTGTTTGTTCGGGACTGGTTTGGCTGGGACAAAGTTAGAACTGCCTGAA
3160 GTTCGCACATTCAGATTTGTGTGTGTCATGGAGTTTATAGGAGGGGATGGCCTTTCGGGTCTTCGCACCTTCCATCTCTTCC
3239 CACTTCCATCTGCGGTCCTCCACACCTTGTTCCTCCCTGCACCTTCTGTGGATGACACAGGGTGCTGCTCCCTCTAGTCTTTGCT
3318 TTGCTCGGCCTTCTGTGCAGGAGACTTGGTCTCAAGCTCAGAGAGAGCCAGTCCGGTCCAGCTCCTTTGTGCTCCCTTCC
3397 TCAGAGGCCTTCCCTTGAAGATGCATCTAGACTACAGCCTTATCAGTGTTTTAAGCTTATTCCTTTAACATAAGCTTCTCT
3476 GACAACATGAATTTGTTGGGTTTMMGGCGTTGGTTGAATTGTTTAGGTTTGGCTTTATACCGGGCCAAATAGCACA
3555 TAACACCTGTTATATATGAATACTCATATGTTTATGACCACAAATAAATATGAACCTCATATTAATAAAAAAAAAA
3569 AAAAGGCGGCGCG

FIG. 5E

GOCITTTAGGGCAGAGATTCTGAGCTGOGTTTTAGGGTACAGATTTOOCT
 GTTTGAGGAGCTTGGOOOCTCTGTAAAGCATCTGACTCATCTCAGAGATAT
 CAATTCCTTAAACACTGTGACAACAGGATCTAAATGGCTGACACATTTGT
 OCTTGTGTCAOGTTOCATTTATTTTAAAAOCTCAGTAAATOGTTTTTA
 GCTTCTTTTOCAGCAAACTCTTCTOCACAGTAGOOCAGTOGTGTAGGATA
 AATTACGGATATAGTCATTCTAGGGGTTTCAGTCTTTTOCATCTCAAGGC
 ATTGTGTTTTTGTTOGGGACTGGTTTTGGCTGGGACAAAGTTAGAACTG
 OCTGAAGTTGGCACATTCAGATTGTTGTGTGTOCATGGAGTTTTAGGAGGGG
 ATGGOCTTTDOGGTCTTOGCACCTTOCATOCTCTOOCACCTTOCATCTGGOG
 TOOCACAOCTTGTDOOCTGCACCTTCTGGATGACACAGGGTGTCTGCTOCT
 OCTAGTCTTTTGOCTTTGCTGGGOCTTCTGTGCAGGAGACTTGGTCTCAA
 GCTCAGAGAGAGOCAGTDOGGTDOOCAGCTOCTTTGTDOOCTTOCTCAGAGG
 OCTTOCTTTGAAGATGCATCTAGACTAOCAGOCTTATCAGTGTTTAAGCTT
 ATTOCTTTAACATAAGCTTTOCTGACAACATGAATTTGTGGGGTTTTTG
 GOGTTGGTTGATTTGTTTAGGTTTTGCTTTATACOOGGGOCAAATAGCAC
 ATAACAOCTGGTTATATATGAATAACTCATATGTTTATGAOCCAAATAAA
 TATGAAAOCTCATATTAAAAAAATAAAAGGGGGG

FIG. 6

56.2% identity in 176 aa overlap

```
30      40      50      60      70      80
TRAIL AAATPSKVGSSAGRIEPRGGGRGALPTSMGQHGPSA-RARAGRAPGPRPAREASPRLLKV
      :: :: :: :: :: :: :: :: :: :: :: ::
T74    MGLWCQSVPTASSARAGRYPGARTASGTRFWLLD
      10      20      30

90      100     110     120     130     140
TRAIL HRTFKFVW--VGVLLQVVPSSAATIK---LHDQSIGTQQWEHSPGLGELCPFGSHRSEIRPG
      :: :: :: :: :: :: :: :: :: :: :: ::
T74    SKILKFVVFIVAVLLFVRVDSATIPRQDEVPPQIVAPQQRRSLKEECPAGSHRSEITTG
      40      50      60      70      80      90

150     160     170     180     190     200
TRAIL ACNRCTEGVGYTNASNNLFACLPCCTACKSDEEERSPCCTTTRNTACQCPGTFERNDSAEM
      :: :: :: :: :: :: :: :: :: :: :: ::
T74    ACNPCTEGVDYTLASNNLPSCLLCTVCKSGQTNKSSCTTTRDTVCQCEKGSFQDRKNSPEM
      100     110     120     130     140     150

210     220     230     240     250     260
TRAIL CRKCSTGCPRGVKKVRCCTPWSIDECVHGESGNGHNIVVILVVTLVVPLLVAVLIVCCC
      :: :: :: :: :: :: :: :: :: :: :: ::
T74    CRTCTGCPRGVKKVSNCTPRSDIKCKNESAAASSTGKTPAAEETVTTILGMLASPHYLI
      160     170     180     190     200     210
```

FIG. 7

V R N K T L P W S P C S A V Y L T E L 19
GC GTC CCG AAC AAG ACG CTG CCC TGG TCT CCC TGC AGT GCT CTC TAC CTC ACG GAG CTC 59

L D D G H G D C L L D D G H S T L Y E L 39
CTG GAT GAT GGT CAC GGA GAC TGC CTC CTG GAT GAT GGC CAC AGC ACC CTC TAT GAG CTG 119

D Q Q C K Q I F G P D F R H C P N T S V 59
GAC CAG CAG TGC AAG CAG ATC TTT GGG CCT GAT TTC CGA CAC TGC CCC AAC ACC TCT GTG 179

E D I C V Q L W C R H R D S D E P I C H 79
GAG GAC ATC TGT GTC CAG CTC TGG TGC CGT CAT CCG GAT AGT GAT GAG CCE ATT TGC CAC 239

T K N G S L L W A D G T P C G P G H L C 99
ACA AAG AAT GGC AGC TTG CTC TGG GCT GAT GGT ACG CCC TGT GGC CCT GCG CAC CTG TGC 299

L D G S C V L R E E V E N P K A V V D G 119
CTG GAT GGT AGC TGT GTG CTC CCG GAG GAA GTA GAG AAT CCC AAG GCT GTG GTA GAT GGA 359

D W G P W G P W G Q C S R T C G G I Q 139
GAC TGG GGT CCC TGG GGA CCC TGG GGA CAA TGT TCT CGC ACC TGT GGT GGA GGG ATA CAG 419

F S N R E C D N P A P Q N G G R F C L G 159
TTT TCG AAC CGT GAG TGT GAT AAT CCA GCA CCT CAG AAT GGA GGA AGA TTT TCC CTG GGA 479

FIG. 8A

E R V K Y Q S C K T E E C P P N G K S F 179
GAG AGA GTC AAG TAC CAA TCT TGC AAG ACA GAG GAA TGT CCA CCA AAT GGA AAA AGC TTC 539

R E Q Q C E K Y N A Y N H T D L D G N F 199
AGG GAG CAG CAG TGT GAA AAA TAT AAT GCC TAC AAC CAC ACG GAC CTG GAT GGG AAT TTC 599

L Q W V P K Y S G V S P R D R C K L F C 219
CTT CAG TGG GTC CCC AAA TAC TCA GGA GTG TCC CCC CGA GAC CGA TGC AAA CTG TTT TGC 659

R A R G R S E F K V F E T K V I D G T L 239
AGA GCC CGT GGG AGG AGT CAG TTC AAA GTG TTT GAA ACT AAG GTG ATC GAT GGC ACT CTG 719

C G P D T L A I C V R G Q C V K A G C D 259
TGC CGA CCG GAT ACT CTG GCC ATC TGT GTG CCG GGA CAG TGC GTT AAG GGT GGC TGT GAC 779

H V V N S P K K L D K C G V C G K G T 279
CAT GTG GTG AAC TCA CCT AAG AAG CTG GAC AAG TGT GGG GTG TGT GGG GGC AAA GGC ACT 839

A C R K V S G S F T P F S Y G Y N D I V 299
GCC TGT AGG AAG GTC TCA GGT TCT TTC ACC CCC TTC AGT TAT GGC TAC AAT GAC ATT GTC 899

T I P A G A T N I D V K Q R S H P G V Q 319
ACC ATC CCA GCT GGT GCC ACA AAT ATT GAT GTG AAA CAA CCG AGC CAC CCA GGG GTC CAG 959

FIG. 8B

N D G S Y L A L K T A N G Q Y L L N G N 339
AAT GAC GGC AGC TAC CTG GCA CTG AAG ACA GCC AAT GGG CAG TAC CTG CTC AAT GGT AAC 1019

L A I S A I E Q D I L M K G T I L K Y S 359
CTA GCC ATC TCT GCC ATA GAG CAA GAC ATC TTG ATG AAG GGG ACC ATC CTA AAG TAC AGT 1079

G S M A T L E R L Q S F Q A L P E P L T 379
GGT TCC ATG GCC ACC CTG GAG CGG CTG CAG AGC TTC CAA GCC CTC CCT GAG CCT CTT ACA 1139

V Q L L T V S G E V F P P K V K Y T F F 399
GTA CAG CTC CTG ACT GTG TCT GGT GAG GTC TTC CCT CCA AAA GTC AAA TAT ACC TTC TTC 1199

V P N D T D F N V Q S S K E R A S T N I 419
GTC CCC AAT GAC ACG GAC TTC AAC CTG CAG AGT AGC AAA GAA AGA GCA AGC ACC AAC ATC 1259

I Q S L P Y A E W V L G D W S E C P S T 439
ATT CAG TCC TTG CCC TAT GCA GAG TGG GTG CTG GGG GAC TGG TCT GAA TGT CCA AGC ACA 1319

C G G G W Q R R T V E C R D P S G Q A S 459
TGT GGA GGT GGC TGG CAG CGG CGG ACT GTG GAA TGC AGG GAC CCC TCA GGT CAG GCC TCT 1379

D T C D E A L K P E D A K P C G S Q P C 479
GAC ACC TGT GAT GAG GCT CTG AAA CCT GAG GAT GCC AAG CCC TGT GGA AGC CAG CCA TGT 1439

L L * 482
CTC CTC TGA 1448

FIG. 8C

TCCCTTGGTGACATGTCTAAGGCTTATGGATTMYGGGCTACTGGGCTACAGACAAAGGTCTCCTCTGAGGTGACACTA 1527
CATATCAAGATGGCATGGCCCTTCCAGGCCCTTCTATTACTACAACCCCTTTCGGGTACCACCTAATTCTATAAGGAAGAGAG 1606
AAGAGGATGTAAAGGGTAACAGACTGTAAAGTTGACTGTCTAGTGGACTGGACCTTGTATTATGACCAAGAAGATGGGATA 1685
CGTTAAAGGTAAAGTGTGCTTATTGATCCAAAGGTGAGATTTCAGAACCCCTCTTTCGCAAGGACTAGAAAGGTT 1764
AAATGAGAAAGAGAATTMTTCTCTTTCGGTMTCTCCATAATAATCAATCTACCTCAGCGGGAGGAACCTTGGTGAT 1843
AAGGCCAGGTGTTAGTGGTGAGTGCCCAAGGCACCTCTCCATAGATATCTTCGAGCCATCTTCAGAAATGCCCATGGCTGT 1922
TTTCAGTATTAAACTCTGTCTCTCAAAGGTGGTGTGCCATCAGGGTTATAGAAAGCCACCTTGTCTCAGGCT 2001
GCCTCCTGCTGGGGCGGACCCCTTTTCAAGTATTATTATGCAAAATATGTTTCTGAACTAAAGTGTGATCTTTACACCAAAAAA 2080
AAAAAAAAAAAAAAAAAAAAAGGCGGCGCC 2114

FIG. 8D

>fthxe29c9F, 1448 bases, 3233 checksum.

GTCGACCCACGCGTCCGGGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAA
 AATGGGGCAGACTGTGACAGTGTGACTGGAAAGTGCACCTGTGCCCCAGG
 ATTCAAAGGAATTGACTGCTCTACCCCATGCCCTCTGGGAACCTATGGGA
 TAAACTGTTCCCTCTCGCTGTGGCTGTAAAAATGATGCAGTCTGCTCTCCT
 GTGGACGGGTCTTGTACTTGCAAGGCAGGCTGGCACGGGGTGGACTGCTC
 CATCAGATGTCCCAGTGGCACATGGGGCTTTGGCTGTAACCTAACATGCC
 AGTGCCCTCAACGGGGGAGCCTGCAACACCCCTGGACGGGACCTGCACGTGT
 GCACCTGGATGGCGGGGGAGAAATGCCAACTTCCCTGCCAGGATGGCAC
 GTAAGGGCTGAACTGTGCTGAGCGCTGCGACTGCAGCCACGCAGATGGCT
 GCCACCCCTACCAAGGGCCATTGCCGCTGCCCTCCCCGGATGGTCAGGTGTC
 CACTGTGACAGCGTGTGTGCTGAGGGACGCTGGGGCCCCAACTGCTCCCT
 GCCCTGCTACTGTAAAAATGGGGCTTCATGCTCCCTGATGATGGCATCT
 GCGAGTGTGCACCAGGCTTCGGAGGGCAACACTTGTGAGAGGATCTGCTCC
 CCTGGTTTTTATGGGCATCGCTGCAGCCAGACATGCCACAGTGGGTTCA
 CAGCAGCGGGCCCTGCCACCACATCACCGGCTGTGTGACTGCTTGCTG
 GCTTCACAGGGCGCCCTCTGCAATGAAGTGTGTCCCAGTGGCAGATTTGGG
 AAAAAGTGTGCAGGAATTTGTACCTGCACCAACAACGGAACTGTAAACC
 CATTGACAGATCTTGTGAGTGTTAACCCGGTTGGATTGGCAGTGAAGTCT
 CTCAACCATGTCCACCTGCCCACTGGGGGCCCCAACTGCATCCACAGTGC
 AACTGCCATAATGGAGCTTTCTGCAGCGCCTACGATGGGGGAATGTAAATG
 CACTCCTGGCTGGACAGGGCTCTACTGCACTCAGAGATGTCTCTAGGGT
 TTTATGGAAAAGATTGTGCACTGATATGCCAATGTCAAAACGGAGCTGAC
 TGCGACCACATTTCTGGGCAGTGTACTTGCCGCACTGGATTCATGGGACG
 GCACTGTGAGCAGAAGTGCCCTTCAGGAACATATGGCTATGGCTGTGCCC
 AGATATGTGATTGTCTGAACAACCTCCACCTGCGACCACATCACTGGGACC
 TGTTACTGCAGCCCCGGATGGAAGGGAGCGAGATGTGATCAAGCTGGTGT
 TATCATAGTTGGAAATCTGAACAGCTTAAGCCGAACCAGTACTGCTCTCC
 CTGCTGATTCCCTACCAAATCGGGGCCATTGCAGGCATCATCATTCTTGTC
 CTAGTTGTTCTCTTCCCTACTGGCATTGTTTATTATAGACACAGC

FIG. 9

>fthxe29c9R, 1578 bases, 5059 checksum.

NAGCCCAACAGGAATGTTCTATGAAAGTGAACCTAACAGTGAGTGTTGTT
 CCCAAGGAGTATTCAGCAATAATGGGCGTCTNTCCCAAGGATCCATATGA
 CCTCCCAAAGAACAGTCACATCCCTTGTCATTATGACCTGCTGOCAGTCC
 GAGACAGTTTCATCCTCCCTAAGCAAGAGGACAGTGGAGGTAGCAGCAGC
 AACAGCAGCAGCAGCAGTGAATGACACCAAAGGACCGCTTGGTAGCCACT
 GGAACCCCTTCCAGAACTGCTGTTTGGTTCTTCTCCATCCTCAATTTTGC
 CACTTTTCATGTGAATGTTAGTCAATTCGGTGGGCAATTTTGGACATGAA
 CCAGAAAGCTGAAAGCTGAGGCTGACACGGACTGTAGGTGCTTTTTGTTT
 AGGTGGATTGGAAGGAGTTAGAGATGTGATTTGCCATTGCTGTTAGTTTT
 AGAACTATACCGTGAAGCATGACTTATTGTAAGATGTTGGCTGAAAGCA
 TGAACCTTGCAAGAACTCCCTCGGAGACGCAGGTTGCAGTGGACATTGGGAT
 TGTTGCTTGAAAAATTAAATTTGAATATTTTCTCTCTCATTTGCATCAT
 ACAGCTCTACCTAGGATTGTACAGTTTACCATAAAATTTACTTCATGAAA
 GTGGGAATCACTGAACATGTAGAAGACAAGGAACATATTGTTAACTCCTG
 ATTCTTAACTTTATTCAACTGGACTCAGAATTGTAGGGATAATATGAATG
 CAGGAGGAAACATTCTGTCAGGCGGTATGACTGGACAGACTTTGAATATA
 CTCTAAAAGTGGACAGAAAATTTACGAAAATCTTAGATTTTGTTTAGAAT
 GAGAAAATATACAATTAGAATTATTTTAGAAATAGTAGGAAGTATTGCAG
 AAGTCAATACACAAATGTGCCAGGCAGAGGTGGTTTTCTCTGTTTGACTC
 TCAACCAACTTCAGATCTATGACATTATTCTGATCACTGGCTCCATCATA
 CATATTCAOCACTTGAGATTTCATAACATATCAATAGTTATTTTCATAAATA
 TAGAAATGAAATAATTTTATTTTGGACAGACTGGATGGAATGAGTGTGTA
 ATGATTGATAAAGGTTGTAAATTTTAAATGCAAGATGACGCTTACGTTCT
 GTAAACCATTAGTAATACATGCTGTAATATAGAATTAGTGGAACATTTTG
 ATTAATCTTCCCTAGAAGTGAAGTGAATATTTTGTGCATATTTGAGAA
 AGGGAACCTTTCCTTTTATTAATTGTCAATTTAGAGAACTATGCTTAAGC
 TGGTCTTTTGCATTGCTAATGTGACATGTACCCAACCTTTTCATTAATTTG
 TATTTCCATTTTAAATTGCATATTCTATGTTTTGTAGTGTGGATTGT
 TAATGAAAAAATATTATATGTTTCGTTATTCCTTGTATTATTGCCACTTAT
 CTTTGTCTTGATAAAAAATGCGTTGTTCTTTTTTCTTTTGGAGGGACAAGA
 TGAAAAATATATAATTTGAATTGATTAAAATTGGTCGTTACTAAAATAGTA
 TAGTAAAAAAAAAAAAAAAAAGGGCGGCGG

FIG. 10

ggg gaa gct tgc cag cag atc tgc agc tgc caa aat ggg gca gac tgt	48
Gly Glu Ala Cys Gln Gln Ile Cys Ser Cys Gln Asn Gly Ala Asp Cys	15
1	5
gac agt gtg act gga aag tgc acc tgc ggc cca gga ttc aaa gga att	96
Asp Ser Val Thr Thr Gly Lys Cys Thr Cys Ala Pro Gly Phe Lys Gly Ile	30
20	25
gac tgc tct acc cca tgc cct ctg gga acc tat ggg ata aac tgt tcc	144
Asp Cys Ser Thr Thr Pro Cys Pro Leu Gly Thr Tyr Gly Ile Asn Cys Ser	45
35	40
tct cgc tgt ggc tgt aaa aat gat gca gtc tgc tct cct gtg gac ggg	192
Ser Arg Cys Gly Cys Lys Asn Asp Ala Val Cys Ser Pro Val Asp Gly	60
50	55

FIG. 11A

tct	tgt	act	tgc	aag	gca	ggc	tgg	cac	ggg	gtg	gac	tgc	tcc	atc	aga	240
Ser	Cys	Thr	Cys	Lys	Ala	Gly	Trp	His	Gly	Val	Asp	Cys	Ser	Ile	Arg	
65					70					75					80	
tgt	ccc	agt	ggc	aca	tgg	ggc	ttt	ggc	tgt	aac	tta	aca	tgc	cag	tgc	288
Cys	Pro	Ser	Gly	Thr	Trp	Gly	Phe	Gly	Cys	Asn	Leu	Thr	Cys	Gln	Cys	
										90					95	
ctc	aac	ggg	gga	gcc	tgc	aac	acc	ctg	gac	ggg	acc	tgc	acg	tgt	gca	336
Leu	Asn	Gly	Gly	Ala	Cys	Asn	Thr	Leu	Asp	Gly	Thr	Cys	Thr	Cys	Ala	
cct	gga	tgg	cgc	ggg	gag	aaa	tgc	gaa	ctt	ccc	tgc	cag	gat	ggc	acg	384
Pro	Gly	Trp	Arg	Gly	Glu	Lys	Cys	Glu	Leu	Pro	Cys	Gln	Asp	Gly	Thr	
tac	ggg	ctg	aac	tgt	gct	gag	cgc	tgc	gac	tgc	agc	cac	gca	gat	ggc	432
Tyr	Gly	Leu	Asn	Cys	Ala	Glu	Arg	Cys	Asp	Cys	Ser	His	Ala	Asp	Gly	

FIG. 11B

tgc cac cct acc acg ggc cat tgc cgc tgc ctc ccc gga tgg tca ggt	480
Cys His Pro Thr Thr Gly His Cys Arg Cys Leu Pro Gly Trp Ser Gly	160
145	
gtc cac tgt gac agc gtg tgt gct gag gga cgc tgg ggc ccc aac tgc	528
Val His Cys Asp Ser Val Cys Ala Glu Gly Arg Trp Gly Pro Asn Cys	175
165	
170	
tcc ctg ccc tgc tac tgt aaa aat ggg gct tca tgc tcc cct gat gat	576
Ser Leu Pro Cys Tyr Cys Lys Asn Gly Ala Ser Cys Ser Pro Asp Asp	190
180	
185	
ggc atc tgc gag tgt gca cca ggc ttc cga ggc acc act tgt cag agg	624
Gly Ile Cys Glu Cys Ala Pro Gly Phe Arg Gly Thr Thr Cys Gln Arg	205
195	
200	
205	
atc tgc tcc cct ggt ttt tat ggg cat cgc tgc agc cag aca tgc cca	672
Ile Cys Ser Pro Gly Phe Tyr Gly His Arg Cys Ser Gln Thr Cys Pro	220
210	
215	

FIG. 11C

cag tgc gtt cac agc agc ggc ccc tgc cac cac atc acc ggc ctg tgt	720
Gln Cys Val His Ser Ser Gly Pro Cys His His Ile Thr Gly Leu Cys	240
225	
gac tgc ttg cct ggc ttc aca ggc gcc ctc tgc aat gaa gtg tgt ccc	768
Asp Cys Leu Pro Gly Phe Thr Gly Ala Leu Cys Asn Glu Val Cys Pro	255
245	
agt ggc aga ttt ggg aaa aac tgt gca gga att tgt acc tgc acc aac	816
Ser Gly Arg Phe Gly Lys Asn Cys Ala Gly Ile Cys Thr Cys Thr Asn	270
260	
aac gga acc tgt aac ccc att gac aga tct tgt cag tgt tac ccc ggt	864
Asn Gly Thr Cys Asn Pro Ile Asp Arg Ser Cys Gln Cys Tyr Pro Gly	285
275	
tgg att ggc agt gac tgc tct caa cca tgt cca cct gcc cac tgg ggc	912
Trp Ile Gly Ser Asp Cys Ser Gln Pro Cys Pro Ala His Trp Gly	300
290	

FIG. 11D

cca aac tgc atc cac acg tgc aac tgc cat aat gga gct ttc tgc agc	960
Pro Asn Cys Ile His Thr Cys Asn Cys His Asn Gly Ala Phe Cys Ser	320
305	310
315	
gcc tac gat ggg gaa tgt aaa tgc act cct ggc tgg aca ggg ctc tac	1008
Ala Tyr Asp Gly Glu Cys Lys Cys Thr Pro Gly Trp Thr Gly Leu Tyr	335
325	330
tgc act cag aga tgt cct cta ggg ttt tat gga aaa gat tgt gca ctg	1056
Cys Thr Gln Arg Cys Pro Leu Gly Phe Tyr Gly Lys Asp Cys Ala Leu	350
340	345
ata tgc caa tgt caa aac gga gct gac tgc gac cac att tct ggg cag	1104
Ile Cys Gln Cys Gln Asn Gly Ala Asp Cys Asp His Ile Ser Gly Gln	365
355	360
tgt act tgc cgc act gga ttc atg gga cgc cac tgt gag cag aag tgc	1152
Cys Thr Cys Arg Thr Gly Phe Met Gly Arg His Cys Glu Gln Lys Cys	380
370	375

FIG. 11E

cct	tca	gga	aca	tat	ggc	tat	ggc	tgt	cgc	cag	ata	tgt	gat	tgt	ctg	1200
Pro	Ser	Gly	Thr	Tyr	Gly	Tyr	Gly	Cys	Arg	Gln	Ile	Cys	Asp	Cys	Leu	400
385					390					395						
aac	aac	tcc	acc	tgc	gac	cac	atc	act	ggg	acc	tgt	tac	tgc	agc	ccc	1248
Asn	Asn	Ser	Thr	Cys	Asp	His	Ile	Thr	Gly	Thr	Cys	Tyr	Cys	Ser	Pro	
				405					410					415		
gga	tgg	aag	gga	gcg	aga	tgt	gat	caa	gct	ggg	ggt	gtt	ata	gtt	gga	1296
Gly	Trp	Lys	Gly	Ala	Arg	Cys	Asp	Gln	Ala	Gly	Val	Ile	Ile	Val	Gly	
			420					425					430			
aat	ctg	aac	agc	tta	agc	cga	acc	agt	act	gct	ctc	cct	gct	gat	tcc	1344
Asn	Leu	Asn	Ser	Leu	Ser	Arg	Thr	Ser	Thr	Ala	Leu	Pro	Ala	Asp	Ser	
							440					445				
tac	cac	atc	ggg	gcc	att	gca	ggc	atc	atc	att	ctt	gtc	cta	gtt	gtt	1392
Tyr	His	Ile	Gly	Ala	Ile	Ala	Gly	Ile	Ile	Ile	Leu	Val	Leu	Val	Val	
						455					460					

FIG. 11F

ctc ttc cta ctg gca ttg ttc att att att tat aga cac aag cag aag gga	1440
Leu Phe Leu Leu Ala Leu Phe Ile Ile Tyr Arg His Lys Gln Lys Gly	480
465	
aag gaa tca agc atg cca gca gtt acc tac acc cct gct atg agg gtc	1488
Lys Glu Ser Ser Met Pro Ala Val Thr Tyr Thr Pro Ala Met Arg Val	495
485	
gtc aat gca gat tat acc att tca gga acc ctt cct cac agc aat ggt	1536
Val Asn Ala Asp Tyr Thr Ile Ser Gly Thr Leu Pro His Ser Asn Gly	510
500	
gga aac gct aat agc cac tac ttc acc acc aat ccc agt tac cac acg ctc	1584
Gly Asn Ala Asn Ser His Tyr Phe Thr Asn Pro Ser Tyr His Thr Leu	525
515	
acc cag tgt gcc aca tcc cct cac gtc aac aac agg gac agg atg act	1632
Thr Gln Cys Ala Thr Ser Pro His Val Asn Asn Arg Asp Arg Met Thr	540
530	

FIG. 11G

gtc acg aag tca aaa aac aat caa ctg ttt gtg aat ctt aaa aat gtg 1680
Val Thr Lys Ser Lys Asn Asn Gln Leu Phe Val Asn Leu Lys Asn Val 560
545 550 555

aac cct ggg aag aga ggc cct gtg ggg gac tgc atg gga cat tgc cgg 1728
Asn Pro Gly Lys Arg Gly Pro Val Gly Asp Cys Met Gly His Cys Arg 575
565 570

ctg act gga aac atg gcg gct acc tca acg agc tcg gtg ctt ttg gac 1776
Leu Thr Gly Asn Met Ala Ala Thr Ser Thr Ser Ser Val Leu Leu Asp 590
580 585

ttg aca gaa gct ata tgg gaa aat cct taaaagacct gggaagaat 1823
Leu Thr Glu Ala Ile Trp Glu Asn Pro 600
595 600

tctgaatata attcaagtaa ctgctcccta agcagttctg agaaccata tgccactatt 1883
aaagacccac ctgtacttat ccgaaaagc tcagagtgtg gttatgtgga gatgaaatcg 1943
ccggcacgaa gagattcccc atatgcagag atcaataact caacttcagc caacaggaat 2003
gtctatgaag ttgaacctac agtgagtgtt gtccaaggag tattcagcaa taatggggcgt 2063
ctctcccagg atccatatga cctcccaaag aacagtcaca tcccttgta ttatgacctg 2123

FIG. 11H

ctgccagtcc	gagacagttc	atcctcccct	aagcaagagg	acagtggagg	tagcagcagc	2183
aacagcagca	gcagcagtga	atgacaccaa	aggaccgctt	ggtagccact	ggaacccttt	2243
ccagaactgc	tggttggttc	ttctccatcc	tcaattttgc	cactttcatg	tgaatgttag	2303
tcaattcggc	gggcaatttt	tggacatgaa	ccagaaagct	gaaagctgag	gctgacacgg	2363
actgtaggtg	ctttttgttc	aggtaggattc	gaaggagtta	gagatgtgat	ttgccattgc	2423
tgtagtattt	agaactatac	ccgtgaagca	tgacttattg	taagatgttg	gctgaaagca	2483
tgaacttgca	gaactccctc	ggagacgcag	gttgcagtg	acattgggat	tgtagcttga	2543
aaaattaaaa	tttgaatatt	ttctctctca	tttgcattcat	acagctctac	ctaggattgt	2603
acagtttacc	ataaaattta	cttcatgaaa	gtgggaatca	ctgaacatgt	agaagacaag	2663
gaacatatgt	ttaactcctg	attcttaact	ttattcaact	ggactccagaa	ttgtagggat	2723
aatatgaatg	caggaggaaa	cattctgtca	ggcggtatga	ctggacagac	tttgaatata	2783
ctctaaaagt	ggacagaaaa	tttacgaaaa	tcttagattt	tgtttagaat	gagaaaatat	2843
acaattagaa	ttattttaga	aatagtagga	agtattgcag	aagtcaatac	acaaaatgtgc	2903
caggcagagg	tggtttttctc	tgtttgactc	tcaaccaact	tcagatctat	gacattatct	2963
tgatccactgg	ctccatcata	cataattcacc	acttgagatt	cataacatat	caatagttat	3023
ttcataaata	tagaaatgaa	ataattttat	ttttgacaga	ctggatggaa	tgagtgtgta	3083
atgattgata	aaggttgtaa	attttaaattg	caagatgacg	cttacgttct	gtaaacatt	3143
agtaatacat	gctgtaatat	agaattagtg	gaacattttg	attaatcttt	ccctagaagt	3203
gactgaaata	tttttggtgca	tatttgagaa	agggaacttt	ccttttatta	attgtcaatt	3263
tagagaaact	atgcttaagc	tggtcttttg	cattgctaata	gtgacatgta	cccaactttt	3323
cattaaatttg	tatttccatt	tttaaaattgc	atatctctatg	ttttgtagtg	tttggattgt	3383
taatgaaaaa	atatatatg	ttcgttattc	cttgatttat	tgccacttat	cttttgcttg	3443
ataaaaaatgc	gttgttcttt	tttcttttgg	agggacaaga	tgaaaaatata	taatttgaat	3503
tgattaaaaat	tggtcgttac	taaaatagta	tagtaaaaaa	aaaaaaaaaa		3552

FIG. 11

Comparison of TANGO 73 and RVP1

Percent Similarity: 68.750 Percent Identity: 48.077

t73pro x RVP1.pep

July 29, 1997 17:38

```

1 MANAGLQLLGFILAFILGWIGAIIVSTALPQWRIYSYAGDNIVTAQAMYEG 50
  . :|:: | ||.||||:..|||:|||| ||: .: |..|:|||| :|||
1 .MSMSLEITGTSLAVLGWLCTIVCCALPMWRVSAFIGSSIITAQITWEG 49
51 WMSCVSQSTGQIQCKVFDSLNLSSSTLQATRALMVVGILLGVIAIFVATV 100
  ||.|| |||||:||||:|||||. . |||.|||:||||:||||:..:::| |
50 WMNCV.QSTGQMCKMYDSLALPQDLQAARALIVVSILLAFAFGLLVALV 98
101 GMKCMKCLEDDEVQKMRMAVIGGAIFLLAGLAILVATAWYGNRIVQEFYD 150
  | .| .|: :||..| ::::|:|:||||:| .||:..| :| |::|:|:
99 GAQCTNCV.QDETAKAKITIVAGVLFLLAAVLTLPVSVSANTIIRDFYN 147
151 PMTPVNARYEFGQALFTGWAAASLCLLGGALLCCSCPRKTTSYPTPRYP 200
  |:| ..: |:| :|:|||||. | ||||| |||||:.. .. ||. |.
148 PLVPEAQKREMGTYGVGWAAALQLLGGALLCCSCPPREKYAPTKILYS 197
201 KP...APSSGKDYV..... 211
  | :|:|:|: .
198 APRSTGPGTGTGTAYDRKTTSERPGARTPHHHHYQPSMYPTRPACSLASE 247

```

FIG. 12

COMPARISON OF TANGO 73 AND TMVCF
(SIROTKIN ET AL, GENOMICS 1997, 42:245-251)

PERCENT SIMILARITY: 67.619 PERCENT IDENTITY: 46.190

TMVCF.PEP X T73PRO

JULY 29, 1997 16:32 ..

```

1 MGSAALEILGLVLCLVWGGLILACGLPMWQVTAFLDHNIIVTAQT TWKGL 50
|.:|:|::||:|.::|| | |:..:| | :. : .:|||||. :.||
1 MANAGLQLLGFI LAFLGWIGAIVSTALPQWRIYSYAGDNIIVTAQAMY EGL 50

51 WMSCVVQSTGHMQCKVYDSVLALSTE VQAARALT VSAVLLAFVALFVT LA 100
|||| | |||:| |||:| |:|. | ..:| |-|| | ::||:|:|:| |. .
51 WMSCVSQSTGQIQCKVFDSLNL SSTLQATRALMVVGILLGVIAIFVATV 100

101 GAQCTTCVAPG.PAKARVALTGGVLYLFCGL LALVPLCW FANIVVREFYD 149
| . | .:..: ..| |:|:|. |:~:|:~:| | | :. |:~:| :|:| |||
101 GMKCMKCLEDDEVQKM RMAVIGGAIFLLAGLA I LVATAWYGNRIVQE FYD 150

150 PSVPVSQKYELGAALYIGWAATALLMVG GCLCCGAWVCTGRPDLSFPVK 199
| . |||..:| |:|. |:|. |||||..| :~:| | ||| :..| .. |:|..
151 PMTPVNARYEFGQALFTGWAAASLC LLGGALLCC.... SCPRKTTSYPT P 196

200 YSAPRRPTATGDYDKKNYV 218
. |:~:~:~| |:~|
197 RPYPKPAPSSG.... KDYV 211

```

FIG. 13

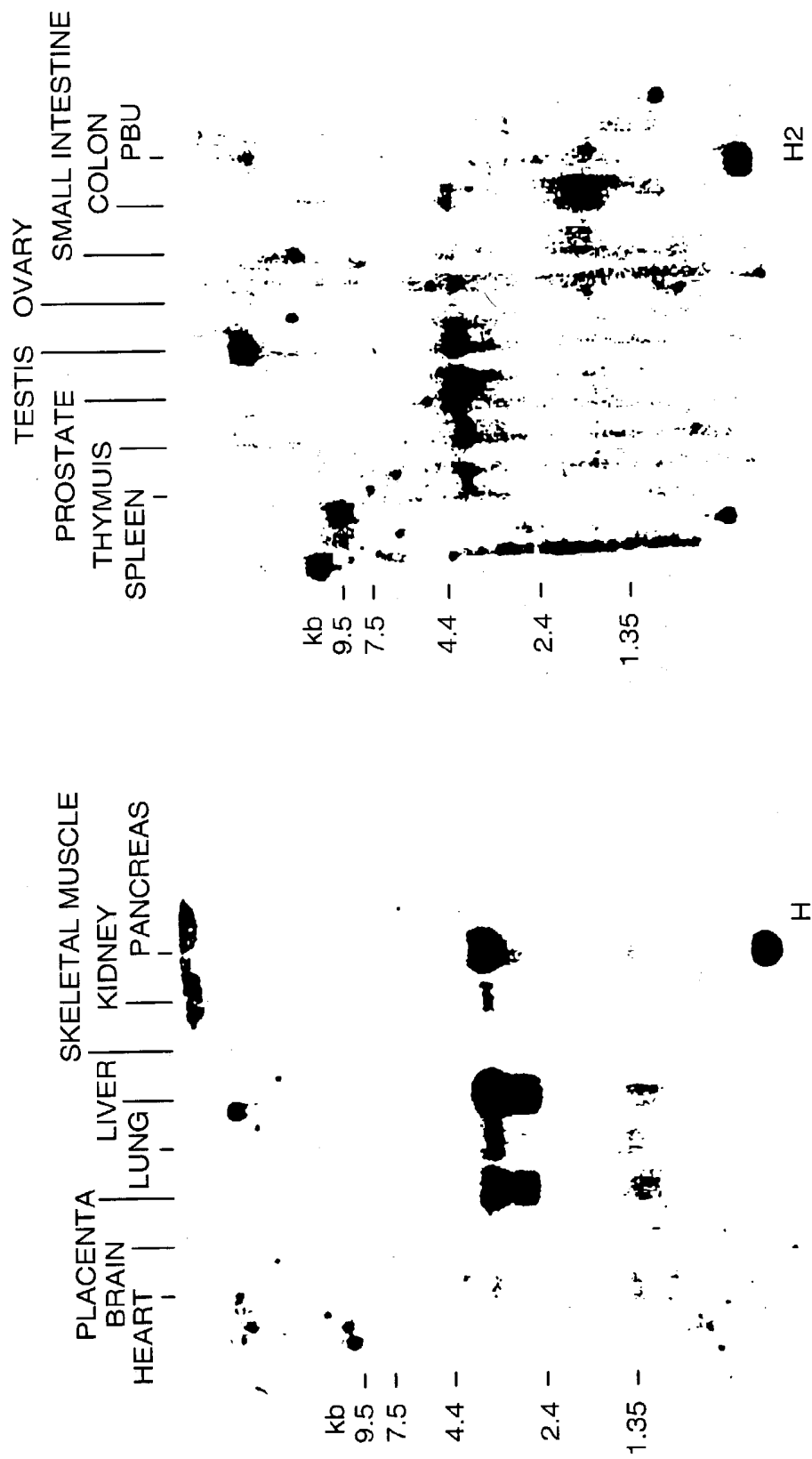


FIG. 14A

FIG. 14B

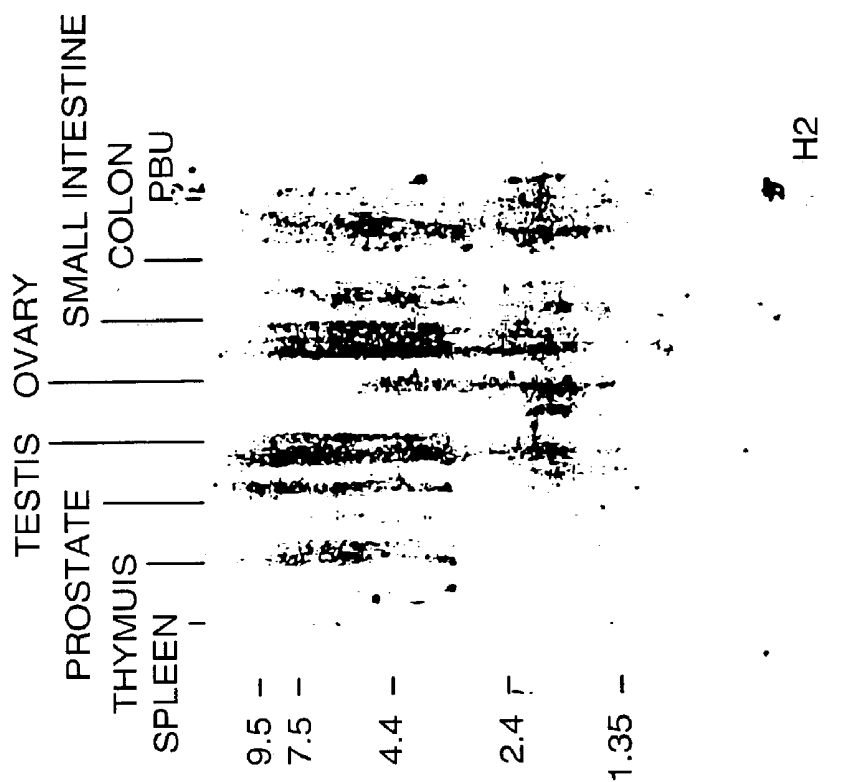


FIG. 15B

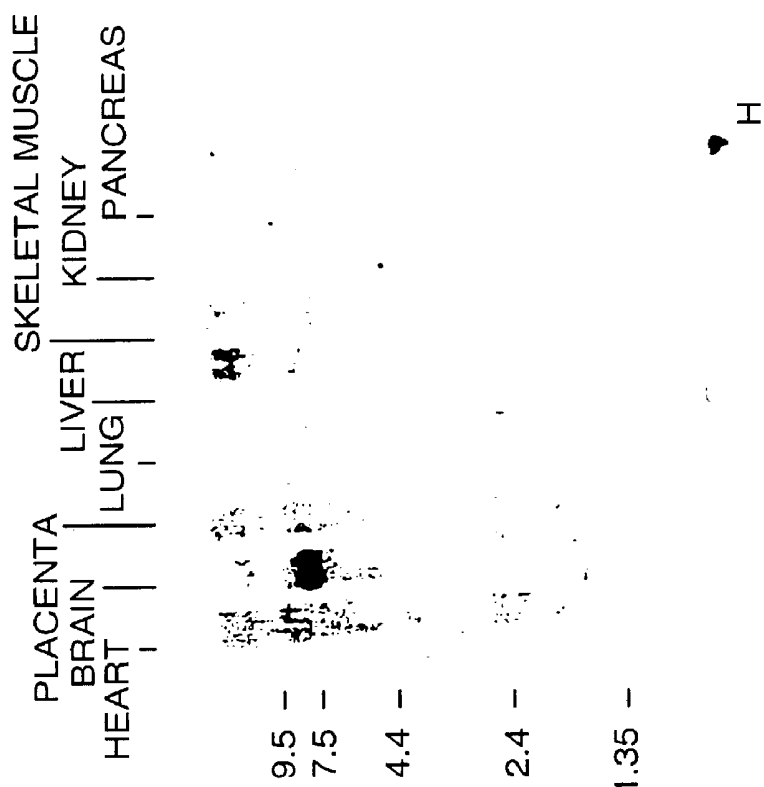


FIG. 15A

```

1 CASLNGVTGDSHLMASMLSSLDHSQPWSPCSAYMVTSLFDNKGHECLMDK 50
      . . . . . ||||| . | |||.||:|:|
1 .....VRNKTLPWSPCSAVYLTELDDGHGDCLLD. 30

51 PQNP IKLP S D L P G T LYDANRQCQTFGEESKHCPDAA..STCTTLWCTGT 98
      | |||: ..||. || : :|||. . | |||
31 .....DGHSTLYELDQQCKQIFGPDFRHCPNTSVEDICVQLWCRHR 71

99 SGGLLV C QTKH..FPWADGTSCGEGKWCVSGKC VNKTDMKH F ATPVHG SW 146
      :| ||. ||||| || | |. | || : :... | | |
72 DSDEPICHTKNGSLLWADGTPCGPGHLCLDGSCVLREEVENPKAVVDGDW 121

1147 GPWGPWGDCSR TC GG V Q Y TMRECDNPVPKNGGKYCEGKRVRYS CNIED 196
      ||||| ||||| |:|: ||||| |. |||:|.||:|:
1122 GPWGPWGQCSR TC GG GI Q FS N RECDNPAPQN GGRFCLGERVKYQS CKTEE 171

1197 CPDN NG K T FR EE Q CE A H NE FSKA SF G NE PTVE W TP KYAGVS PKDRCK LTC 246
      || |||.|||: ||| :|: |||.|||: ||||| |
1172 CPP.NGKSFR E QQCEKYNAYNH TD LDGN.FLOWVPKYSGVSPRDRCKLFC 219

```

FIG. 16A

```

2247 EAKGIGYFFVLQPKVVDGTPCSPDSTSVCVGQCVKAGCDRIIDSKKKFD 296
      | : | | : ||: ||| | | | . : | | | | | | | | : . | | | |
2220 RARGRSEFKVFETKVIDGTLGCPDTLAICVRGQCVKAGCDHVVNSPKKLD 269
      . . . . . . . . . . . . . . . . . . . . . . . . . . . .
2297 KCGVCGGNGSTCKKMSGIVTSTRPGYHDI VTIPAGATNIEVKHRNQGRSR 346
      | | | | | | | . | : | . | | | | | | | | | | | | | | | |
2270 KCGVCGGKGTA CRKVSGSFTPFSGYNDI VTIPAGATNIDVKORSHPGVQ 319
      . . . . . . . . . . . . . . . . . . . . . . . . . . . .
3347 NNGSFLAIRAADGTYYILNGNFTLSTLEQDLTYKGTVLRYSGSSAALERIR 396
      | . | : | : | : | : | : | : | : | : | : | : | : | : | :
3320 NDGSYLALKTANGQYLLNGNLAISAEQDILMKGTILKYSGSMATLERLQ 369
      . . . . . . . . . . . . . . . . . . . . . . . . . . . .
3397 SFSPLKEPLTIQVLMV.GHALRPKIKFTYFMKKKTE.....SFNA 435
      | | | | | | | : | | | | | : | : | : | : | : | | | |
3370 SFQALPEPLTVQLLTVSGEVFPKVKYTF FVPNDTDFNVQSSKERASTNI 419
      . . . . . . . . . . . . . . . . . . . . . . . . . . . .
436 IPT..FSEWVIEEWGECSTCGSGWQRRVVQCRDINGHPASECAKEVKPA 483
      | . : . | | : | | | | | | | | | | | : | | . | . . | |
420 IQSLPYAEWVLGDWSECPSTCGGGWQRRRTVECRDPGQASDTCDEALKPE 469
      . . . . . . . . . . . . . . . . . . . . . . . . . . . .
4484 STRPCADLPCPHWQVDWSPCSKTCGKGKRTLKCVSHDGGVLSNESCD 533
      : | | | |
470 DAKPCGSQPCLL..... 481

```

FIG. 16B

NOVEL GENES ENCODING PROTEINS HAVING PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC AND OTHER USES

CROSS REFERENCE TO RELATED APPLICATIONS

[0001] This application is a continuation-in-part (and claims the benefit of priority under 35 USC 120) of the following applications:

[0002] 1. U.S. application Ser. No. 09/128,709 (filed Aug. 4, 1998), which application claims priority from U.S. Ser. No. 60/054,645 (filed Aug. 4, 1997).

[0003] 2. U.S. application Ser. No. 09/130,491 (filed Aug. 6, 1998), which application claims priority from U.S. Ser. No. 60/054,966 (filed Aug. 6, 1997) and U.S. Ser. No. 60/058,108 (filed Sep. 5, 1997).

[0004] 3. U.S. application Ser. No. 09/388,280 (filed Sep. 1, 1999), a divisional of U.S. application Ser. No. 09/130,491 (filed Aug. 6, 1998), which application claims priority from U.S. Ser. No. 60/054,966 (filed Aug. 6, 1997) and U.S. Ser. No. 60/058,108 (filed Sep. 5, 1997).

[0005] 4. U.S. application Ser. No. 09/388,279 (filed Sep. 1, 1999), a divisional of U.S. application Ser. No. 09/130,491 (filed Aug. 6, 1998), which application claims priority from U.S. Ser. No. 60/054,966 (filed Aug. 6, 1997) and U.S. Ser. No. 60/058,108 (filed Sep. 5, 1997).

TECHNICAL FIELD OF THE INVENTION

[0006] This invention relates to polypeptides and the genes encoding them.

SUMMARY OF THE INVENTION

[0007] The invention relates to the discovery and characterization of the genes encoding Tango-73, Tango-74, Tango-76, Tango-78, and Tango-83. Tango-73 cDNA (SEQ ID NO: 3; **FIG. 3**) encodes a human protein (SEQ ID NO: 4; **FIG. 3**) that is 48% identical to rat RVP.1 (Briehl et al., *Mol. Endocrinol.* 5:1381, 1991). Murine Tango-73 cDNA (SEQ ID NO: 17; **FIG. 4**) encodes a murine protein (SEQ ID NO: 18; **FIG. 4**). Tango-74 cDNA (SEQ ID NO: 5; **FIG. 5**) encodes a human protein (SEQ ID NO: 6; **FIG. 5**) with homology to TRAIL receptor (Pan et al., *Science* 276:111, 1997). Tango-76 cDNA (SEQ ID NO: 7; **FIG. 7**) is a rat protein (SEQ ID NO: 8; **FIG. 8**) that is 62.1% identical to murine ADAMTS-1 (**FIG. 16**). Tango-78 cDNA (SEQ ID NO: 1; **FIG. 1**) was isolated from a human bone marrow cDNA library (Clontech; Palo Alto, Calif.) and encodes a 169 amino acid portion of Tango-78, a novel protein (SEQ ID NO: 2; **FIG. 1**) that is highly homologous to the murine nodal protein (Collignon et al., *Nature* 381:155, 1996). Tango-83 cDNA (SEQ ID NO: 9, SEQ ID NO: 10, and SEQ ID NO: 19; **FIG. 10**, **FIG. 12**, and **FIG. 11**, respectively) encodes a protein (SEQ ID NO: 20; **FIG. 11**) that is expressed by stimulated human astrocytes.

[0008] The invention features isolated nucleic acid molecules encoding Tango-73, Tango-74, Tango-76, Tango-78, or Tango-83 polypeptides, isolated nucleic acid molecules that encode polypeptides that are substantially identical to

the Tango-73, Tango-74, Tango-76, Tango-78, or Tango-83 protein sequences described herein (SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 18, or SEQ ID NO: 20) and isolated nucleic acid molecules which hybridize under stringent conditions to the protein coding portions of the Tango-73, Tango-74, Tango-76, Tango-78, or Tango-83 nucleic acid sequences described herein (SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 17, or SEQ ID NO: 19).

[0009] The invention also features a host cell that includes an isolated nucleic acid molecule encoding a polypeptide of the invention and a nucleic acid vector (e.g., an expression vector; a vector which includes a regulatory element; a vector that is a virus; a vector that is a retrovirus) containing an isolated nucleic acid molecule encoding a polypeptide of the invention.

[0010] In one embodiment, the invention features a substantially pure polypeptide of the invention (e.g., a polypeptide of the invention that is soluble under physiological conditions); a polypeptide of the invention that includes a signal sequence; a Tango-73 polypeptide that is at least 85%, 90%, 95%, or 100% identical to the amino acid sequence of SEQ ID NO: 4 or SEQ ID NO: 18; a Tango-74 polypeptide that is at least 85%, 90%, 95%, or 100% identical to the amino acid sequence of SEQ ID NO: 6; a Tango-76 polypeptide that is at least 85%, 90%, 95%, or 100% identical to the amino acid sequence of SEQ ID NO: 8; a Tango-78 polypeptide that is at least 85%, 90%, 95%, or 100% identical to the amino acid sequence of SEQ ID NO: 2; a Tango-83 polypeptide that is at least 85%, 90%, 95%, or 100% identical to the amino acid sequence of SEQ ID NO: 20.

[0011] In other embodiments the invention also features substantially pure polypeptides which include a first portion and a second portion, the first portion including a polypeptide of the invention and the second portion including a detectable marker.

[0012] The invention also features antibodies, e.g., monoclonal, polyclonal, and engineered antibodies, which specifically bind polypeptides of the invention. By "specifically binds" is meant an antibody that recognizes and binds to a particular antigen, e.g., a Tango-73, Tango-74, Tango-76, Tango-78, or Tango-83 polypeptide of the invention, but which does not substantially recognize or bind to other molecules in a sample, e.g., a biological sample, which includes the polypeptide (e.g., Tango-73, Tango-74, Tango-76, Tango-78, or Tango-83).

[0013] The invention also features a pharmaceutical composition that includes a polypeptide of the invention.

[0014] The invention includes methods for diagnosing a disorder associated with aberrant expression of a protein of the invention (i.e., Tango-73, Tango-74, Tango-76, Tango-78, or Tango-83), the method including obtaining a biological sample from a patient and measuring the expression of the protein in the biological sample, wherein increased or decreased expression of the protein in the biological sample compared to a control indicates that the patient suffers from a disorder associated with aberrant expression of the protein.

[0015] The invention encompasses isolated nucleic acid molecules encoding a polypeptide of the invention or a polypeptide fragment thereof; vectors containing these

nucleic acid molecules; cells harboring recombinant DNA encoding a polypeptide of the invention; fusion proteins which include all or a portion of a polypeptide of the invention; transgenic animals which express a polypeptide of the invention; and recombinant knock-out animals which fail to express a polypeptide of the invention.

[0016] The nucleic acid molecules of the invention can be inserted into vectors, as described below, which will facilitate expression of the insert. The nucleic acid molecules and the polypeptides they encode can be used directly as diagnostic or therapeutic agents, or (in the case of a polypeptide) can be used to generate antibodies that, in turn, are therapeutically useful. Accordingly, expression vectors containing the nucleic acid molecules of the invention, cells transfected with these vectors, the polypeptides expressed, and antibodies generated (against either the entire polypeptide or an antigenic fragment thereof) are among the preferred embodiments.

[0017] A transformed cell is any cell into which (or into an ancestor of which) has been introduced, by means of recombinant DNA techniques, a nucleic acid encoding a polypeptide of the invention (e.g., a Tango-73, Tango-74, Tango-76, Tango-78, or Tango-83 polypeptide).

[0018] The invention also encompasses nucleic acid molecules that hybridize, preferably under stringent conditions, to a nucleic acid molecule encoding a polypeptide of the invention (e.g., the polypeptide encoding portions of a nucleic acid molecule having the sequence shown in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 17, or SEQ ID NO: 19). Preferably the hybridizing nucleic acid molecule consists of 400, more preferably 200 nucleotides. Preferred hybridizing nucleic acid molecules have a biological activity possessed by a nucleic acid of the invention.

[0019] The invention also features substantially pure or isolated polypeptides of the invention, including those that correspond to various functional domains of polypeptides of the invention, or fragments thereof.

[0020] The polypeptides of the invention can be prepared by recombinant gene expression, chemically synthesized, or purified from tissues in which they are naturally expressed using standard biochemical methods of purification.

[0021] Also included in the invention are functional polypeptides, which possess one or more of the biological functions or activities of Tango-73, Tango-74, Tango-76, or Tango-78, or Tango-83. These functions include the ability to bind some or all of the proteins that normally bind to polypeptides of the invention. A functional polypeptide is also considered within the scope of the invention if it serves as an antigen for production of antibodies that specifically bind to a polypeptide of the invention. In many cases, functional polypeptides retain one or more domains present in the naturally-occurring form of the polypeptide.

[0022] The functional polypeptides may contain a primary amino acid sequence that has been modified from those disclosed herein. Preferably these modifications consist of conservative amino acid substitutions, as described herein.

[0023] Another aspect of this invention features isolated or recombinant proteins and polypeptides of the invention, or modulators thereof. Preferred proteins and polypeptides

possess at least one biological activity possessed by the corresponding naturally-occurring human polypeptide. An activity, a biological activity, and a functional activity of a polypeptide of the invention refers to an activity exerted by a protein or polypeptide of the invention on a responsive cell as determined *in vivo*, or *in vitro*, according to standard techniques. Such activities can be a direct activity, such as an association with or an enzymatic activity on a second protein or an indirect activity, such as a cellular signaling activity mediated by interaction of the protein with a second protein. Thus, such activities include, e.g., (1) the ability to form protein-protein interactions with proteins in the signaling pathway of the naturally-occurring polypeptide; (2) the ability to bind a ligand of the naturally-occurring polypeptide; (3) the ability to bind to an intracellular target of the naturally-occurring polypeptide.

[0024] Further activities of polypeptides of the invention include the ability to modulate (this term, as used herein, includes, but is not limited to, stabilize, promote, inhibit or disrupt, protein-protein interactions (e.g., homophilic and/or heterophilic)), protein-ligand interactions, e.g., in receptor-ligand recognition, development, differentiation, maturation, proliferation and/or activity of cells function, survival, morphology, proliferation and/or differentiation of cells of tissues in which it is expressed. Additional activities include but are not limited to: (1) the ability to modulate cell surface recognition; (2) the ability to transduce an extracellular signal (e.g., by interacting with a ligand and/or a cell-surface receptor); (3) the ability to modulate a signal transduction pathway; and (4) the ability to modulate intracellular signaling cascades (e.g., signal transduction cascades).

[0025] The invention also features antagonists and agonists of Tango-73, Tango-74, Tango-76, Tango-78, or Tango-83 that can inhibit or enhance, respectively, one or more of the biological activities of nucleic acids or polypeptides of the invention. Suitable antagonists can include: small molecules (i.e., molecules with a molecular weight below about 500); large molecules (i.e., molecules with a molecular weight above about 500); antibodies that bind and "neutralize" polypeptides of the invention (as described below); polypeptides that compete with a native form of a polypeptide of the invention for binding to a functional binding partner of the native protein of the invention; and nucleic acid molecules that interfere with transcription of nucleic acids of the invention (for example, antisense nucleic acid molecules and ribozymes). Agonists of nucleic acids or polypeptides of the invention also include small and large molecules, and antibodies other than neutralizing antibodies.

[0026] In addition, the invention features substantially pure polypeptides that functionally interact with polypeptides of the invention and the nucleic acid molecules that encode them.

[0027] The invention encompasses methods for treating disorders associated with aberrant expression or activity of a protein of the invention (i.e., Tango-73, Tango-74, Tango-76, Tango-78, or Tango-83). Thus, the invention includes methods for treating disorders associated with excessive expression or activity of a protein of the invention. Such methods entail administering a compound that decreases the expression or activity of the protein. The invention also includes methods for treating disorders associated with insufficient expression or activity of a protein of the inven-

tion. These methods entail administering a compound that increases the expression or activity of the protein.

[0028] The invention also features methods for detecting a protein of the invention (i.e., Tango-73, Tango-74, Tango-76, Tango-78, or Tango-83). Such methods include: obtaining a biological sample; contacting the sample with an antibody that specifically binds to the protein under conditions that permit specific binding; and detecting any antibody-protein complexes formed.

[0029] In addition, the present invention encompasses methods and compositions for the diagnostic evaluation, typing, and prognosis of disorders associated with inappropriate expression or activity of nucleic acids or polypeptides of the invention. For example, the nucleic acid molecules of the invention can be used as diagnostic hybridization probes to detect, for example, inappropriate expression of nucleic acids or polypeptides of the invention or mutations in the genes of the invention. Such methods may be used to classify cells by the level of expression of nucleic acids or polypeptides of the invention.

[0030] Thus, the invention features a method for diagnosing a disorder associated with aberrant activity of a protein of the invention, the method including obtaining a biological sample from a patient and measuring the activity of the protein in the biological sample, wherein increased or decreased activity in the biological sample compared to a control indicates that the patient suffers from a disorder associated with aberrant activity of the protein.

[0031] The nucleic acid molecules of the invention can be used as primers for diagnostic PCR analysis for the identification of gene mutations, allelic variations, and regulatory defects in a gene of the invention. The present invention further provides for diagnostic kits for the practice of such methods.

[0032] The invention features methods of identifying compounds that modulate the expression or activity of a protein of the invention by assessing the expression or activity of the protein in the presence and absence of a selected compound. A difference in the level of expression or activity of the protein in the presence and absence of the selected compound indicates that the selected compound is capable of modulating expression or activity of the protein. Expression can be assessed either at the level of gene expression (e.g., by measuring mRNA) or protein expression by techniques that are well known to skilled artisans. The activity of nucleic acids or polypeptides of the invention can be assessed functionally.

[0033] The preferred methods and materials are described below in examples that are meant to illustrate, not limit, the invention. Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, the preferred methods and materials are described herein. All publications, patent applications, patents, and other references mentioned herein are incorporated by reference in their entirety. In the case of conflict, the present specification, including definitions, will control. In addition, the materials, methods, and examples are illustrative only and are not intended to be limiting.

[0034] The details of one or more embodiments of the invention are set forth in the accompanying drawings and the description below. Other features, objects, and advantages of the invention will be apparent from the description and drawings, and from the claims.

DESCRIPTION OF DRAWINGS

[0035] **FIG. 1** depicts the nucleic acid sequence (SEQ ID NO: 1; open reading frame from nucleotide 588-1094) and deduced amino acid sequence (SEQ ID NO: 2) of Tango-78. The open reading frame extends from nucleotide

[0036] **FIG. 2** depicts an alignment of the amino acid sequence of Tango-78 (top sequence; SEQ ID NO: 2) and the amino acid sequence of murine nodal protein (bottom sequence; SEQ ID NO: 12). The sequences show 79.4% identity.

[0037] **FIG. 3** depicts the nucleotide acid sequence (SEQ ID NO: 3; open reading frame from nucleotide 240-875) and deduced amino acid sequence (SEQ ID NO: 4) of Tango-73.

[0038] **FIG. 4** depicts the nucleotide acid sequence (SEQ ID NO: 17; open reading frame from nucleotide 187-819) and deduced amino acid sequence (SEQ ID NO: 18) of murine Tango-73.

[0039] **FIG. 5** depicts the nucleotide acid sequence (SEQ ID NO: 5; open reading frame from nucleotide 143-1264) and deduced amino acid sequence (SEQ ID NO: 6) of Tango-74.

[0040] **FIG. 6** depicts the nucleotide acid sequence of a 3' non-coding portion of Tango-74 (SEQ ID NO: 11).

[0041] **FIG. 7** depicts an alignment of a portion of the amino acid sequence of Tango-74 (bottom sequence; SEQ ID NO: 6) and the amino acid sequence of TRAIL (top sequence; SEQ ID NO: 13). The sequences show 56.2% identity.

[0042] **FIG. 8** depicts the partial nucleotide sequence (SEQ ID NO: 7; open reading frame from nucleotide 3-1448) and deduced amino acid sequence (SEQ ID NO: 8) of Tango-76.

[0043] **FIG. 9** depicts the nucleotide sequence of a 5' portion of Tango-83 (SEQ ID NO: 9).

[0044] **FIG. 10** depicts the nucleotide sequence of a 3' portion of Tango-83 (SEQ ID NO: 10).

[0045] **FIG. 11** depicts the nucleotide sequence (SEQ ID NO: 19; open reading frame from nucleotide 1-1803) and deduced amino acid sequence (SEQ ID NO: 20) of Tango-83.

[0046] **FIG. 12** depicts an alignment of the amino acid sequence of Tango-73 (top sequence; SEQ ID NO: 4) and the amino acid sequence of RVPI (bottom sequence; SEQ ID NO: 14). The sequences show 48.077% identity.

[0047] **FIG. 13** depicts an alignment of the amino acid sequence of Tango-73 (bottom sequence; SEQ ID NO: 4) and TMVCF (top sequence; SEQ ID NO: 15). The sequences show 46.190% identity.

[0048] **FIG. 14** depicts a Northern blot analysis of Tango-73 mRNA.

[0049] **FIG. 15** depicts a Northern blot analysis of Tango-83 mRNA.

[0050] **FIG. 16** depicts an alignment of amino acid sequence of Tango-76 (top sequence; SEQ ID NO: 8) and ADAMTS-1 (bottom sequence; SEQ ID NO: 16). The sequences show 62.1% identity.

DETAILED DESCRIPTION OF THE INVENTION

[0051] The present invention is based, at least in part, on the discovery of a variety of cDNA molecules which encode proteins which are herein designated Tango-73, Tango-74, Tango-76, Tango-78, and Tango-83. These proteins exhibit a variety of physiological activities, and are included in a single application for the sake of convenience. It is understood that the allowability or non-allowability of claims directed to one of these proteins has no bearing on the allowability of claims directed to the others. The characteristics of each of these proteins and the cDNAs encoding them are described separately in the ensuing sections. In addition to the full length mature and immature human proteins described in the following sections, the invention includes fragments, derivatives, and variants of these proteins, as described herein. These proteins, fragments, derivatives, and variants are collectively referred to herein as polypeptides of the invention or proteins of the invention.

[0052] Tango-73 cDNA (**FIG. 3**; SEQ ID NO: 3) was isolated from human prostate epithelial cells as follows.

[0053] Human prostate epithelial cells (Clonetics) were expanded in culture with Prostate Epithelial Growth Medium (PEGM) (Clonetics). When cells reached confluence cells were grown in Prostate Basal Media (Clonetics) for 24 hours. They were stimulated with PEGM (prostate epithelial growth medium; Clonetics) and 40 ug/ml cycloheximide for 3 hours.

[0054] Total RNA was isolated using the RNeasy Midi Kit (Qiagen). Poly (A)+ was isolated using the Oligotex beads (Qiagen). Next, cDNA was constructed using the Superscript cDNA Synthesis Kit (Gibco BRL). The cDNA was cloned into the expression vector pMET7 using the SalI and NotI sites in the polylinker. Transformants were picked and sequenced.

[0055] Northern blot analysis of Tango-73 expression was carried out as described above. This analysis revealed the presence of 4.0 kb and 3.0 kb transcripts in the placenta and liver. A 4.0 kb transcript was present in lung, kidney, thymus, prostate, spleen, testes, and colon, with the highest expression in lung, pancreas, prostate, and testes.

[0056] The amino acid sequence of Tango-73 is 48% identical to rat RVP.1 (Briehl et al., *Mol. Endocrinol.* 5:1381, 1991) and 46.1% identical to TMVCF (Sirotkin et al., *Genomics* 42:245, 1997).

[0057] RVP.1 ("Rat Ventral Prostate 1") is predicted to encode a 280 amino acid protein that lacks significant homology to known protein functional domains (Briehl et al., supra). RVP.1 is up-regulated during apoptosis (Briehl et al., supra).

[0058] TMVCF ("transmembrane protein deleted in VCFS"), a 219 amino acid protein with two putative membrane-spanning domains, is deleted in velo-cardio-facial

syndrome (Sirotkin et al., supra). TMVCF shows high expression in human adult lung, heart, and skeletal muscle, and its transcripts have been detected as early as Day 9 of mouse development (Sirotkin et al., supra).

[0059] The invention also includes murine Tango-73 (**FIG. 4**).

[0060] Tango-83 (**FIG. 9**, **FIG. 10**, and **FIG. 11**) and Tango-74 cDNAs (**FIG. 5**) were isolated from human astrocytes as follows.

[0061] Human astrocytes (Clonetics) were expanded in culture with Astrocyte Growth Media (AGM; Clonetics) according to the recommendations of the supplier. When the cells reached ~80-90% confluence, they were stimulated with 200 units/ml Interleukin 1-Beta (Boehringer Mannheim) and cycloheximide (CHI: 40 micrograms/ml) for 4 hours. Total RNA was isolated using the RNeasy Midi Kit (Qiagen), and the poly A+ fraction was further purified using Oligotex beads (Qiagen).

[0062] Three micrograms of poly A+ RNA were used to synthesize a cDNA library using the Superscript cDNA Synthesis kit (Gibco BRL). Complementary DNA was directionally cloned into the expression plasmid pMET7 using the SalI and NotI sites in the polylinker to construct a plasmid library. Transformants were picked and grown up for single-pass sequencing. Additionally, astrocyte cDNA was ligated into the SalI/NotI sites of the ZipLox vector (Gibco BRL) for construction of a lambda phage cDNA library.

[0063] Northern blot analysis of Tango-83 expression, performed as described above, revealed that Tango-83 is expressed as an approximately 9.0 kb transcript in brain (**FIG. 15**).

[0064] Northern blot analysis, performed as described above, revealed that Tango-74 is expressed as an approximately 4.0 kb transcript in heart, brain, lung, liver, kidney, pancreas, spleen, prostate, testes, ovary, small intestine, colon and peripheral blood lymphocytes. Higher expression was seen in lung, liver, skeletal muscle, spleen, testes, colon and peripheral blood lymphocytes.

[0065] The amino acid sequence of Tango-74 is homologous to the amino acid sequence of the TRAIL receptor (Pan et al., *Science* 276:111, 1997) (**FIG. 7**). Human TRAIL receptor is a member of the TNF-receptor family (Pan et al., *Science* 276:111, 1997). It contains a cytoplasmic "death domain" and is involved in regulating cell suicide and tissue homeostasis (Pan et al., *Science* 276:111, 1997).

[0066] Tango-76 cDNA (SEQ ID NO: 7) was isolated from an adult rat frontal cortex library. The amino acid sequence of Tango-76 is homologous to the amino acid sequence of ADAMTS-1 (**FIG. 16**). ADAMTS ("A Disintegrin-like And Metalloprotease domain (reprolysin-type) with Thrombospondin type I motifs") proteases have been disproportionately linked to disease (See http://www.lerner.ccf.org/bme/staff/apte/adamts/biological_role.html).

ADAMTS-1 has shown involvement in inflammation and cancer cachexia (Kuno et al., *J Biol Chem.* 1997 Jan 3;272(1): 556-62) as well as ureteral and adrenal anomalies and decreased growth (Shindo et al., *J Clin Invest.* 2000 May 15;105(10):1345-1352). ADAMTS-1 protein is composed of the following domains (listed in order from N-terminal to

C-terminal end): signal sequence, pro-domain, metalloprotease domain, disintegrin-like domain, thrombospondin module, cysteine-rich domain, spacer domain, and two C-terminal thrombospondin modules (See http://www.lerner.ccf.org/bme/staff/apte/adamts/domain_organization.html).

[0067] Northern blot analysis of human mRNA probed with a Tango-76 probe revealed a 4.2 kb band in lung. Analysis of rat mRNA revealed a weak 3.8 kb transcript in heart, brain, spleen, liver, skeletal muscle, and kidney and a weak 1.8 kb transcript in spleen and liver.

[0068] Tango-78 cDNA (SEQ ID NO: 1; **FIG. 1**) was isolated from a human bone marrow cDNA library (Clontech; Palo Alto, Calif.). This Tango-78 cDNA encodes a 169 amino acid portion of Tango-78, a novel protein (SEQ ID NO: 2; **FIG. 1**) that is highly homologous to the murine nodal protein (Collignon et al., *Nature* 381:155, 1996).

[0069] Murine nodal is a growth factor related to TGF-beta (Collignon et al., *Nature* 381:155, 1996). Murine nodal is important to development; for example, it is required for the formation of the primitive streak during mouse gastrulation (Collignon et al., *Nature* 381:155, 1996).

[0070] The Tango-78 cDNA (SEQ ID NO: 1; **FIG. 1**) described herein was isolated using the method described in U.S. Ser. No. 08/752,307 (filed Nov. 19, 1996), hereby incorporated by reference. Tango-78 protein (SEQ ID NO: 2; **FIG. 1**) is highly homologous to murine nodal protein (Collignon et al., supra; **FIG. 2**).

[0071] An "isolated nucleic acid molecule" is a nucleic acid molecule that is separated from the 5' and 3' coding sequences with which it is immediately contiguous in the naturally occurring genome of an organism. Isolated nucleic acid molecules include nucleic acid molecules that are not naturally occurring, e.g., nucleic acid molecules created by recombinant DNA techniques. Nucleic acid molecules include both RNA and DNA, including cDNA, genomic DNA, and synthetic (e.g., chemically synthesized) DNA. Where single-stranded, the nucleic acid molecule may be a sense strand or an antisense strand.

[0072] As used herein, a "signal sequence" includes a peptide of at least about 15 or 20 amino acid residues in length which occurs at the N-terminus of secretory and membrane-bound proteins and which contains at least about 70% hydrophobic amino acid residues such as alanine, leucine, isoleucine, phenylalanine, proline, tyrosine, tryptophan, or valine. In a preferred embodiment, a signal sequence contains at least about 10 to 40 amino acid residues, preferably about 19-34 amino acid residues, and has at least about 60-80%, more preferably at least about 65-75%, and more preferably at least about 70% hydrophobic residues. A signal sequence serves to direct a protein containing such a sequence to a lipid bilayer. A signal sequence is usually cleaved during processing of the mature protein.

[0073] The term "purified" as used herein refers to a nucleic acid or peptide that is substantially free of cellular material, viral material, or culture medium when produced by recombinant DNA techniques, or chemical precursors or other chemicals when chemically synthesized.

[0074] Polypeptides or other compounds of interest are said to be "substantially pure" when they are within prepa-

rations that are at least 60% by weight (dry weight) the compound of interest. Preferably, the preparation is at least 75%, more preferably at least 90%, and most preferably at least 99%, by weight the compound of interest. Purity can be measured by any appropriate standard method, for example, by column chromatography, polyacrylamide gel electrophoresis, or HPLC analysis.

[0075] Where a particular polypeptide or nucleic acid molecule is said to have a specific percent identity to a reference polypeptide or nucleic acid molecule of a defined length, the percent identity is relative to the reference polypeptide or nucleic acid molecule. Thus, a peptide that is 50% identical to a reference polypeptide that is 100 amino acids long can be a 50 amino acid polypeptide that is completely identical to a 50 amino acid long portion of the reference polypeptide. It might also be a 100 amino acid long polypeptide that is 50% identical to the reference polypeptide over its entire length. Of course, many other polypeptides will meet the same criteria. The same rule applies for nucleic acid molecules.

[0076] For polypeptides, the length of the reference polypeptide sequence will generally be at least 16 amino acids, preferably at least 20 amino acids, more preferably at least 25 amino acids, and most preferably 35 amino acids, 50 amino acids, or 100 amino acids. For nucleic acids, the length of the reference nucleic acid sequence will generally be at least 50 nucleotides, preferably at least 60 nucleotides, more preferably at least 75 nucleotides, and most preferably 100 nucleotides or 300 nucleotides.

[0077] In the case of polypeptide sequences which are less than 100% identical to a reference sequence, the non-identical positions are preferably, but not necessarily, conservative substitutions for the reference sequence. Conservative substitutions typically include substitutions within the following groups: glycine and alanine; valine, isoleucine, and leucine; aspartic acid and glutamic acid; asparagine and glutamine; serine and threonine; lysine and arginine; and phenylalanine and tyrosine.

[0078] To determine the percent identity of two amino acid sequences or of two nucleic acids, the sequences are aligned for optimal comparison purposes (e.g., gaps can be introduced in the sequence of a first amino acid or nucleic acid sequence for optimal alignment with a second amino or nucleic acid sequence). The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are identical at that position. The percent identity between the two sequences is a function of the number of identical positions shared by the sequences (i.e., % identity = # of identical positions/total # of positions (e.g., overlapping positions) × 100). Preferably, the two sequences are the same length.

[0079] The determination of percent homology between two sequences can be accomplished using a mathematical algorithm. A preferred, non-limiting example of a mathematical algorithm utilized for the comparison of two sequences is the algorithm of Karlin and Altschul (1990) *Proc. Natl. Acad. Sci. USA* 87:2264-2268, modified as in Karlin and Altschul (1993) *Proc. Natl. Acad. Sci. USA* 90:5873-5877. Such an algorithm is incorporated into the

NBLAST and XBLAST programs of Altschul, et al. (1990) *J. Mol. Biol.* 215:403-410. BLAST nucleotide searches can be performed with the NBLAST program, score=100, wordlength=12 to obtain nucleotide sequences homologous to Tango-73, Tango-74, Tango-76, Tango-78, or Tango-83 nucleic acid molecules of the invention. BLAST protein searches can be performed with the XBLAST program, score=50, wordlength=3 to obtain amino acid sequences homologous to Tango-73, Tango-74, Tango-76, Tango-78, or Tango-83 protein molecules of the invention. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul et al. (1997) *Nucleic Acids Res.* 25:3389-3402. Alternatively, PSI-Blast can be used to perform an iterated search that detects distant relationships between molecules. Id. When utilizing BLAST, Gapped BLAST, and PSI-Blast programs, the default parameters of the respective programs (e.g., XBLAST and NBLAST) can be used. See <http://www.ncbi.nlm.nih.gov>. Another preferred, non-limiting example of a mathematical algorithm utilized for the comparison of sequences is the algorithm of Myers and Miller, (1988) *CABIOS* 4:11-17. Such an algorithm is incorporated into the ALIGN program (version 2.0) which is part of the GCG sequence alignment software package. When utilizing the ALIGN program for comparing amino acid sequences, a PAM120 weight residue table, a gap length penalty of 12, and a gap penalty of 4 can be used.

[0080] Another preferred, non-limiting example of a mathematical algorithm utilized for the comparison of sequences is the local homology algorithm of Smith and Waterman (*Advances in Applied Mathematics* 2: 482-489 (1981)). Such an algorithm is incorporated into the BestFit program, which is part of the Wisconsin™ package, and is used to find the best segment of similarity between two sequences. BestFit reads a scoring matrix that contains values for every possible GCG symbol match. The program uses these values to construct a path matrix that represents the entire surface of comparison with a score at every position for the best possible alignment to that point. The quality score for the best alignment to any point is equal to the sum of the scoring matrix values of the matches in that alignment, less the gap creation penalty multiplied by the number of gaps in that alignment, less the gap extension penalty multiplied by the total length of all gaps in that alignment. The gap creation and gap extension penalties are set by the user. If the best path to any point has a negative value, a zero is put in that position.

[0081] After the path matrix is complete, the highest value on the surface of comparison represents the end of the best region of similarity between the sequences. The best path from this highest value backwards to the point where the values revert to zero is the alignment shown by BestFit. This alignment is the best segment of similarity between the two sequences. Further documentation can be found at <http://ir.ucdavis.edu/GCGhelp/bestfit.html#algorithm>.

[0082] Additional algorithms for sequence analysis are known in the art and include ADVANCE and ADAM as described in Torellis and Robotti (1994) *Comput. Appl. Biosci.*, 10:3-5; and FASTA described in Pearson and Lipman (1988) *Proc. Natl. Acad. Sci.* 85:2444-8. Within FASTA, ktup is a control option that sets the sensitivity and speed of the search. If ktup=2, similar regions in the two sequences being compared are found by looking at pairs of

aligned residues; if ktup=1, single aligned amino acids are examined. ktup can be set to 2 or 1 for protein sequences, or from 1 to 6 for DNA sequences. The default if ktup is not specified is 2 for proteins and 6 for DNA. For a further description of FASTA parameters, see <http://bioweb.pasteur.fr/docs/man/man/fasta.1.html#sect2>, the contents of which are incorporated herein by reference.

[0083] The percent identity between two sequences can be determined using techniques similar to those described above, with or without allowing gaps. In calculating percent identity, typically exact matches are counted.

[0084] As used herein, the phrase "allelic variant" refers to a nucleotide sequence that occurs at a given locus or to a polypeptide encoded by the nucleotide sequence. Allelic variants of any of these genes can be identified by sequencing the corresponding chromosomal portion at the indication location in multiple individuals.

Tango-73, Tango-74, Tango-76, Tango-78, and Tango-83 Nucleic Acid Molecules

[0085] The invention encompasses nucleic acids that have a sequence that is substantially identical to the nucleic acid sequence of Tango-73, Tango-74, Tango-76, Tango-78, or Tango-83. A nucleic acid sequence which is substantially identical to a given reference nucleic acid; sequence is hereby defined as a nucleic acid having a sequence that has at least 85%, preferably 90%, and more preferably 95%, 98%, 99% or more identity to the sequence of the given reference nucleic acid sequence, e.g., the nucleic acid sequence of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 17, or SEQ ID NO: 19.

[0086] The Tango-73, Tango-74, Tango-76, Tango-78, or Tango-83 nucleic acid molecules of the invention can be cDNA, genomic DNA, synthetic DNA, or RNA, and can be double-stranded or single-stranded (i.e., either a sense or an antisense strand). Fragments of these molecules are also considered within the scope of the invention, and can be produced, for example, by the polymerase chain reaction (PCR) or generated by treatment with one or more restriction endonucleases. A ribonucleic acid (RNA) molecule can be produced by in vitro transcription.

[0087] The nucleic acid molecules of the invention can contain naturally occurring sequences, or sequences that differ from those that occur naturally, but, due to the degeneracy of the genetic code, encode the same polypeptide. In addition, these nucleic acid molecules are not limited to sequences that only encode polypeptides, and thus, can include some or all of the non-coding sequences that lie upstream or downstream from a coding sequence.

[0088] The nucleic acid molecules of the invention can be synthesized (for example, by phosphoramidite-based synthesis) or obtained from a biological cell, such as the cell of a mammal. Thus, the nucleic acids can be those of a human, mouse, rat, guinea pig, cow, sheep, horse, pig, rabbit, monkey, dog, or cat. Combinations or modifications of the nucleotides within these types of nucleic acids are also encompassed.

[0089] In addition, the isolated nucleic acid molecules of the invention encompass fragments that are not found as such in the natural state. Thus, the invention encompasses

recombinant molecules, such as those in which a nucleic acid molecule (for example, an isolated nucleic acid molecule encoding Tango-73, Tango-74, Tango-76, Tango-78, or Tango-83) is incorporated into a vector (for example, a plasmid or viral vector) or into the genome of a heterologous cell (or the genome of a homologous cell, at a position other than the natural chromosomal location). Recombinant nucleic acid molecules and uses therefor are discussed further below.

[0090] In the event the nucleic acid molecules of the invention encode or act as antisense molecules, they can be used for example, to regulate translation of mRNA of the invention. Techniques associated with detection or regulation of expression of nucleic acids or polypeptides of the invention are well known to skilled artisans and can be used to diagnose and/or treat disorders associated with aberrant expression of nucleic acids or polypeptides of the invention.

[0091] The invention also encompasses nucleic acid molecules that hybridize under stringent conditions to a nucleic acid molecule encoding a polypeptide of the invention (e.g., nucleic acid molecules having the sequence of the protein encoding portion of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 17, or SEQ ID NO: 19). The cDNA sequences described herein can be used to identify these hybridizing nucleic acids, which include, for example, nucleic acids that encode homologous polypeptides in other species, and splice variants of the genes of the invention in humans or other mammals. Accordingly, the invention features methods of detecting and isolating these nucleic acid molecules. Using these methods, a sample (for example, a nucleic acid library, such as a cDNA or genomic library) is contacted (or "screened") with a probe specific to a nucleotide of the invention (for example, a fragment of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 17, or SEQ ID NO: 19 that is at least 25 or 50 or 100 nucleotides long). The probe will selectively hybridize to nucleic acids encoding related polypeptides (or to complementary sequences thereof). The probe, which can contain at least 25 (for example, 25, 50, 100, or 200 nucleotides) can be produced using any of several standard methods (see, for example, Ausubel et al., "Current Protocols in Molecular Biology, Vol. I," Green Publishing Associates, Inc., and John Wiley & Sons, Inc., NY, 1989). For example, the probe can be generated using PCR amplification methods in which oligonucleotide primers are used to amplify a nucleic acid sequence specific to a nucleic acid of the invention that can be used as a probe to screen a nucleic acid library and thereby detect nucleic acid molecules (within the library) that hybridize to the probe.

[0092] One single-stranded nucleic acid is said to hybridize to another if a duplex forms between them. This occurs when one nucleic acid contains a sequence that is the reverse and complement of the other (this same arrangement gives rise to the natural interaction between the sense and antisense strands of DNA in the genome and underlies the configuration of the "double helix"). Complete complementarity between the hybridizing regions is not required in order for a duplex to form; it is only necessary that the number of paired bases is sufficient to maintain the duplex under the hybridization conditions used.

[0093] Typically, hybridization conditions are of low to moderate stringency. These conditions favor specific inter-

actions between completely complementary sequences, but allow some non-specific interaction between less than perfectly matched sequences to occur as well. After hybridization, the nucleic acids can be "washed" under moderate or high conditions of stringency to dissociate duplexes that are bound together by some non-specific interaction (the nucleic acids that form these duplexes are thus not completely complementary).

[0094] As is known in the art, the optimal conditions for washing are determined empirically, often by gradually increasing the stringency. The parameters that can be changed to affect stringency include, primarily, temperature and salt concentration. In general, the lower the salt concentration and the higher the temperature, the higher the stringency. Washing can be initiated at a low temperature (for example, room temperature) using a solution containing a salt concentration that is equivalent to or lower than that of the hybridization solution. Subsequent washing can be carried out using progressively warmer solutions having the same salt concentration. As alternatives, the salt concentration can be lowered and the temperature maintained in the washing step, or the salt concentration can be lowered and the temperature increased. Additional parameters can also be altered. For example, use of a destabilizing agent, such as formamide, alters the stringency conditions.

[0095] In reactions where nucleic acids are hybridized, the conditions used to achieve a given level of stringency will vary. There is not one set of conditions, for example, that will allow duplexes to form between all nucleic acids that are 85% identical to one another; hybridization also depends on unique features of each nucleic acid. The length of the sequence, the composition of the sequence (for example, the content of purine-like nucleotides versus the content of pyrimidine-like nucleotides) and the type of nucleic acid (for example, DNA or RNA) affect hybridization. An additional consideration is whether one of the nucleic acids is immobilized (for example, on a filter).

[0096] An example of a progression from lower to higher stringency conditions is the following, where the salt content is given as the relative abundance of SSC (a salt solution containing sodium chloride and sodium citrate; 2×SSC is 10-fold more concentrated than 0.2×SSC). Nucleic acids are hybridized at 42EC in 2×SSC/0.1% SDS (sodium dodecyl-sulfate; a detergent) and then washed in 0.2×SSC/0.1% SDS at room temperature (for conditions of low stringency); 0.2×SSC/0.1% SDS at 42EC (for conditions of moderate stringency); and 0.1×SSC at 68EC (for conditions of high stringency). Washing can be carried out using only one of the conditions given, or each of the conditions can be used (for example, washing for 10-15 minutes each in the order listed above). Any or all of the washes can be repeated. As mentioned above, optimal conditions will vary and can be determined empirically.

[0097] A second set of conditions that are considered "stringent conditions" are those in which hybridization is carried out at 50EC in Church buffer (7% SDS, 0.5% NaHPO₄, 1 M EDTA, 1% BSA) and washing is carried out at 50EC in 2×SSC.

[0098] Once detected, the nucleic acid molecules can be isolated by any of a number of standard techniques (see, for example, Sambrook et al., "Molecular Cloning, A Laboratory Manual," 2nd Ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989).

[0099] The invention also encompasses: (a) expression vectors that contain any of the foregoing coding sequences (related to a polypeptide of the invention) and/or their complements (that is, "antisense" sequence); (b) expression vectors that contain any of the foregoing coding sequences (related to a polypeptide of the invention) operatively associated with a regulatory element (examples of which are given below) that directs the expression of the coding sequences; (c) expression vectors containing, in addition to sequences encoding a polypeptide of the invention, nucleic acid sequences that are unrelated to nucleic acid sequences encoding a polypeptide of the invention, such as molecules encoding a reporter or marker; and (d) genetically engineered host cells that contain any of the foregoing expression vectors and thereby express the nucleic acid molecules of the invention in the host cell.

[0100] Recombinant nucleic acid molecules can contain a sequence encoding a soluble polypeptide of the invention; mature polypeptide of the invention; or polypeptide of the invention having an added or endogenous signal sequence. A full length polypeptide of the invention; a domain of a polypeptide of the invention; or a fragment thereof may be fused to additional polypeptides, as described below. Similarly, the nucleic acid molecules of the invention can encode the mature form of a polypeptide of the invention or a form that encodes a polypeptide that facilitates secretion. In the latter instance, the polypeptide is typically referred to as a proprotein, which can be converted into an active form by removal of the signal sequence, for example, within the host cell. Proproteins can be converted into the active form of the protein by removal of the inactivating sequence.

[0101] The regulatory elements referred to above include, but are not limited to, inducible and non-inducible promoters, enhancers, operators and other elements, which are known to those skilled in the art, and which drive or otherwise regulate gene expression. Such regulatory elements include but are not limited to the cytomegalovirus hCMV immediate early gene, the early or late promoters of SV40 adenovirus, the lac system, the trp system, the TAC system, the TRC system, the major operator and promoter regions of phage A, the control regions of fd coat protein, the promoter for 3-phosphoglycerate kinase, the promoters of acid phosphatase, and the promoters of the yeast α -mating factors.

[0102] Similarly, the nucleic acid can form part of a hybrid gene encoding additional polypeptide sequences, for example, sequences that function as a marker or reporter. Examples of marker or reporter genes include β -lactamase, chloramphenicol acetyltransferase (CAT), adenosine deaminase (ADA), aminoglycoside phosphotransferase (neo^r, G418^r), dihydrofolate reductase (DHFR), hygromycin-B-phosphotransferase (HPH), thymidine kinase (TK), lacZ (encoding β -galactosidase), and xanthine guanine phosphoribosyltransferase (XGPRT). As with many of the standard procedures associated with the practice of the invention, skilled artisans will be aware of additional useful reagents, for example, of additional sequences that can serve the function of a marker or reporter. Generally, the hybrid polypeptide will include a first portion and a second portion; the first portion being a polypeptide of the invention and the second portion being, for example, the reporter described above or an immunoglobulin constant region.

[0103] The expression systems that may be used for purposes of the invention include, but are not limited to, microorganisms such as bacteria (for example, *E. coli* and *B. subtilis*) transformed with recombinant bacteriophage DNA, plasmid DNA, or cosmid DNA expression vectors containing the nucleic acid molecules of the invention; yeast (for example, *Saccharomyces* and *Pichia*) transformed with recombinant yeast expression vectors containing the nucleic acid molecules of the invention (preferably containing the nucleic acid sequence encoding a polypeptide of the invention); insect cell systems infected with recombinant virus expression vectors (for example, baculovirus) containing the nucleic acid molecules of the invention; plant cell systems infected with recombinant virus expression vectors (for example, cauliflower mosaic virus (CaMV) and tobacco mosaic virus (TMV)) or transformed with recombinant plasmid expression vectors (for example, Ti plasmid) containing nucleotide sequences of nucleic acids of the invention; or mammalian cell systems (for example, COS, CHO, BHK, 293, VERO, HeLa, MDCK, WI38, and NIH 3T3 cells) harboring recombinant expression constructs containing promoters derived from the genome of mammalian cells (for example, the metallothionein promoter) or from mammalian viruses (for example, the adenovirus late promoter and the vaccinia virus 7.5K promoter).

[0104] In bacterial systems, a number of expression vectors may be advantageously selected depending upon the use intended for the gene product being expressed. For example, when a large quantity of such a protein is to be produced, for the generation of pharmaceutical compositions containing polypeptides of the invention or for raising antibodies to those polypeptides, vectors that are capable of directing the expression of high levels of fusion protein products that are readily purified may be desirable. Such vectors include, but are not limited to, the *E. coli* expression vector pUR278 (Ruther et al., *EMBO J.* 2:1791, 1983), in which the coding sequence of the insert may be ligated individually into the vector in frame with the lacZ coding region so that a fusion protein is produced; pIN vectors (Inouye and Inouye, *Nucleic Acids Res.* 13:3101-3109, 1985; Van Heeke and Schuster, *J. Biol. Chem.* 264:5503-5509, 1989); and the like. pGEX vectors may also be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to glutathione-agarose beads followed by elution in the presence of free glutathione. The pGEX vectors are designed to include thrombin or factor Xa protease cleavage sites so that the cloned target gene product can be released from the GST moiety.

[0105] In an insect system, *Autographa californica* nuclear polyhidrosis virus (AcNPV) can be used as a vector to express foreign genes. The virus grows in *Spodoptera frugiperda* cells. The coding sequence of the insert may be cloned individually into non-essential regions (for example the polyhedrin gene) of the virus and placed under control of an AcNPV promoter (for example the polyhedrin promoter). Successful insertion of the coding sequence will result in inactivation of the polyhedrin gene and production of non-occluded recombinant virus (i.e., virus lacking the proteinaceous coat coded for by the polyhedrin gene). These recombinant viruses are then used to infect *Spodoptera frugiperda*

cells in which the inserted gene is expressed. (for example, see Smith et al., *J. Virol.* 46:584, 1983; Smith, U.S. Pat. No. 4,215,051).

[0106] In mammalian host cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, the nucleic acid molecule of the invention may be ligated to an adenovirus transcription/translation control complex, for example, the late promoter and tripartite leader sequence. This chimeric gene may then be inserted in the adenovirus genome by in vitro or in vivo recombination. Insertion in a non-essential region of the viral genome (for example, region E1 or E3) will result in a recombinant virus that is viable and capable of expressing a gene product of the invention in infected hosts (for example, see Logan and Shenk, *Proc. Natl. Acad. Sci. USA* 81:3655-3659, 1984). Specific initiation signals may also be required for efficient translation of inserted nucleic acid molecules. These signals include the ATG initiation codon and adjacent sequences. In cases where an entire gene or cDNA, including its own initiation codon and adjacent sequences, is inserted into the appropriate expression vector, no additional translational control signals may be needed. However, in cases where only a portion of the coding sequence is inserted, exogenous translational control signals, including, perhaps, the ATG initiation codon, must be provided. Furthermore, the initiation codon must be in phase with the reading frame of the desired coding sequence to ensure translation of the entire insert. These exogenous translational control signals and initiation codons can be of a variety of origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of appropriate transcription enhancer elements, transcription terminators, etc. (see Bittner et al., *Methods in Enzymol.* 153:516-544, 1987).

[0107] In addition, a host cell strain may be chosen which modulates the expression of the inserted sequences, or modifies and processes the gene product in the specific fashion desired. Such modifications (for example, glycosylation) and processing (for example, cleavage) of protein products may be important for the function of the protein. Different host cells have characteristic and specific mechanisms for the post-translational processing and modification of proteins and gene products. Appropriate cell lines or host systems can be chosen to ensure the correct modification and processing of the foreign protein expressed. To this end, eukaryotic host cells that possess the cellular machinery for proper processing of the primary transcript, glycosylation, and phosphorylation of the gene product may be used. The mammalian cell types listed above are among those that could serve as suitable host cells.

[0108] For long-term, high-yield production of recombinant proteins, stable expression is preferred. For example, cell lines that stably express the sequences of nucleic acids or polypeptides of the invention described above may be engineered. Rather than using expression vectors that contain viral origins of replication, host cells can be transformed with DNA controlled by appropriate expression control elements (for example, promoter, enhancer sequences, transcription terminators, polyadenylation sites, etc.), and a selectable marker. Following the introduction of the foreign DNA, engineered cells may be allowed to grow for 1-2 days in an enriched media, and then switched to a selective media. The selectable marker in the recombinant plasmid confers

resistance to the selection and allows cells to stably integrate the plasmid into their chromosomes and grow to form foci that in turn can be cloned and expanded into cell lines. This method can advantageously be used to engineer cell lines that express nucleic acids or polypeptides of the invention. Such engineered cell lines may be particularly useful in screening and evaluation of compounds that affect the endogenous activity of the gene product.

[0109] A number of selection systems can be used. For example, the herpes simplex virus thymidine kinase (Wigler, et al., *Cell* 11:223, 1977), hypoxanthine-guanine phosphoribosyltransferase (Szybalska and Szybalski, *Proc. Natl. Acad. Sci. USA* 48:2026, 1962), and adenine phosphoribosyltransferase (Lowy, et al., *Cell* 22:817, 1980) genes can be employed in tk⁻, hgp⁺ or ap⁺ cells, respectively. Also, anti-metabolite resistance can be used as the basis of selection for the following genes: dhfr, which confers resistance to methotrexate (Wigler et al., *Proc. Natl. Acad. Sci. USA* 77:3567, 1980; O'Hare et al., *Proc. Natl. Acad. Sci. USA* 78:1527, 1981); gpt, which confers resistance to mycophenolic acid (Mulligan and Berg, *Proc. Natl. Acad. Sci. USA* 78:2072, 1981); neo, which confers resistance to the aminoglycoside G-418 (Colberre-Garapin et al., *J. Mol. Biol.* 150:1, 1981); and hyg⁺, which confers resistance to hygromycin (Santerre et al., *Gene* 30:147, 1984).

[0110] The nucleic acid molecules of the invention are useful for diagnosis of disorders associated with aberrant expression of nucleic acid molecules of the invention are also useful in genetic mapping and chromosome identification.

Tango-73, Tango-74, Tango-76, Tango-78, and Tango-83 Polypeptides

[0111] The invention also includes polypeptides that have a sequence that is substantially identical to the amino acid sequence of Tango-73, Tango-74, Tango-76, Tango-78, or Tango-83 (e.g., polypeptides that are substantially identical to the polypeptide of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 18, or SEQ ID NO: 20). A polypeptide which is "substantially identical" to a given reference polypeptide is a polypeptide having a sequence that has at least 85%, preferably 90%, and more preferably 95%, 98%, 99% or more identity to the sequence of the given reference polypeptide sequence (e.g., the amino sequence of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 18, or SEQ ID NO: 20).

[0112] The terms "protein" and "polypeptide" are used herein interchangeably to describe any chain of amino acids, regardless of length or post-translational modification (for example, glycosylation or phosphorylation). Thus, the term "Tango-73, Tango-74, Tango-76, Tango-78, or Tango-83 polypeptide" includes: full-length, naturally occurring protein of the invention; recombinantly or synthetically produced polypeptide that corresponds to a full-length naturally occurring protein of the invention; or particular domains or portions of the naturally occurring protein. The term also encompasses mature a polypeptide of the invention that has an added amino-terminal methionine (useful for expression in prokaryotic cells).

[0113] The polypeptides of the invention described herein are those encoded by any of the nucleic acid molecules described above and include fragments, mutants, truncated

forms, and fusion proteins of polypeptides of the invention. These polypeptides can be prepared for a variety of uses, including but not limited to the generation of antibodies, as reagents in diagnostic assays, for the identification of other cellular gene products or compounds that can modulate the activity or expression of nucleic acids or polypeptides of the invention, and as pharmaceutical reagents useful for the treatment of disorders associated with aberrant expression or activity of nucleic acids or polypeptides of the invention.

[0114] Preferred polypeptides are substantially pure polypeptides of the invention, including those that correspond to the polypeptide with an intact signal sequence, and the secreted form of the polypeptide. Especially preferred are polypeptides that are soluble under normal physiological conditions.

[0115] The invention also encompasses polypeptides that are functionally equivalent to polypeptides of the invention. These polypeptides are equivalent to polypeptides of the invention in that they are capable of carrying out one or more of the functions of polypeptides of the invention in a biological system. Preferred polypeptides of the invention have 20%, 40%, 50%, 75%, 80%, or even 90% of one or more of the biological activities of the full-length, mature human form of polypeptides of the invention. Such comparisons are generally based on an assay of biological activity in which equal concentrations of the polypeptides are used and compared. The comparison can also be based on the amount of the polypeptide required to reach 50% of the maximal stimulation obtainable.

[0116] Functionally equivalent proteins can be those, for example, that contain additional or substituted amino acid residues. Substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues involved. Amino acids that are typically considered to provide a conservative substitution for one another are specified in the summary of the invention.

[0117] Polypeptides that are functionally equivalent to polypeptides of the invention can be made using random mutagenesis techniques well known to those skilled in the art. It is more likely, however, that such polypeptides will be generated by site-directed mutagenesis (again using techniques well known to those skilled in the art). These polypeptides may have increased functionality or decreased functionality.

[0118] To design functionally equivalent polypeptides, it is useful to distinguish between conserved positions and variable positions. This can be done by aligning the amino acid sequence of a protein of the invention from one species with its homolog from another species. Skilled artisans will recognize that conserved amino acid residues are more likely to be necessary for preservation of function. Thus, it is preferable that conserved residues are not altered.

[0119] Mutations within the coding sequence of nucleic acid molecules of the invention can be made to generate variant genes that are better suited for expression in a selected host cell. For example, N-linked glycosylation sites can be altered or eliminated to achieve, for example, expression of a homogeneous product that is more easily recovered and purified from yeast hosts which are known to hyperglycosylate N-linked sites. To this end, a variety of amino

acid substitutions at one or both of the first or third amino acid positions of any one or more of the glycosylation recognition sequences which occur, and/or an amino acid deletion at the second position of any one or more of such recognition sequences, will prevent glycosylation at the modified tripeptide sequence (see, for example, Miyajima et al., *EMBO J.* 5:1193, 1986).

[0120] The polypeptides of the invention can be expressed fused to another polypeptide, for example, a marker polypeptide or fusion partner. For example, the polypeptide can be fused to a hexa-histidine tag to facilitate purification of bacterially expressed protein or a hemagglutinin tag to facilitate purification of protein expressed in eukaryotic cells.

[0121] A fusion protein may be readily purified by utilizing an antibody specific for the fusion protein being expressed. For example, a system described by Janknecht et al. allows for the ready purification of non-denatured fusion proteins expressed in human cell lines (*Proc. Natl. Acad. Sci. USA* 88: 8972-8976, 1991). In this system, the gene of interest is subcloned into a vaccinia recombination plasmid such that the gene's open reading frame is translationally fused to an amino-terminal tag consisting of six histidine residues. Extracts from cells infected with recombinant vaccinia virus are loaded onto Ni²⁺ nitriloacetic acid-agarose columns and histidine-tagged proteins are selectively eluted with imidazole-containing buffers.

[0122] The polypeptides of the invention can be chemically synthesized (for example, see Creighton, "Proteins: Structures and Molecular Principles," W.H. Freeman & Co., NY, 1983), or, perhaps more advantageously, produced by recombinant DNA technology as described herein. For additional guidance, skilled artisans may consult Ausubel et al. (supra), Sambrook et al. ("Molecular Cloning, A Laboratory Manual," Cold Spring Harbor Press, Cold Spring Harbor, N.Y., 1989), and, particularly for examples of chemical synthesis Gait, M. J. Ed. ("Oligonucleotide Synthesis," IRL Press, Oxford, 1984).

[0123] The invention also features polypeptides that interact with nucleic acids or polypeptides of the invention (and the genes that encode them) and thereby alter the function of nucleic acids or polypeptides of the invention. Interacting polypeptides can be identified using methods known to those skilled in the art. One suitable method is the "two-hybrid system," which detects protein interactions in vivo (Chien et al., *Proc. Natl. Acad. Sci. USA*, 88:9578, 1991). A kit for practicing this method is available from Clontech (Palo Alto, Calif.).

Transgenic Animals

[0124] Polypeptides of the invention can also be expressed in transgenic animals. These animals represent a model system for the study of disorders that are caused by or exacerbated by over expression or under expression of nucleic acids or polypeptides of the invention, and for the development of therapeutic agents that modulate the expression or activity of nucleic acids or polypeptides of the invention.

[0125] Transgenic animals can be farm animals (pigs, goats, sheep, cows, horses, rabbits, and the like), rodents (such as rats, guinea pigs, and mice), non-human primates

(for example, baboons, monkeys, and chimpanzees), and domestic animals (for example, dogs and cats). Transgenic mice are especially preferred.

[0126] Any technique known in the art can be used to introduce a Tango-73, Tango-74, Tango-76, Tango-78, or Tango-83 transgene into animals to produce the founder lines of transgenic animals. Such techniques include, but are not limited to, pronuclear microinjection (U.S. Pat. No. 4,873,191); retrovirus mediated gene transfer into germ lines (Van der Putten et al., *Proc. Natl. Acad. Sci. USA* 82:6148, 1985); gene targeting into embryonic stem cells (Thompson et al., *Cell* 56:313, 1989); and electroporation of embryos (Lo, *Mol. Cell. Biol.* 3:1803, 1983).

[0127] The present invention provides for transgenic animals that carry a transgene of the invention in all their cells, as well as animals that carry a transgene in some, but not all of their cells. That is, the invention provides for mosaic animals. The transgene can be integrated as a single transgene or in concatamers, e.g., head-to-head tandems or head-to-tail tandems. The transgene can also be selectively introduced into and activated in a particular cell type (Lasko et al., *Proc. Natl. Acad. Sci. USA* 89:6232, 1992). The regulatory sequences required for such a cell-type specific activation will depend upon the particular cell type of interest, and will be apparent to those of skill in the art.

[0128] When it is desired that the transgene of the invention be integrated into the chromosomal site of the endogenous gene, gene targeting is preferred. Briefly, when such a technique is to be used, vectors containing some nucleotide sequences homologous to an endogenous gene of the invention are designed for the purpose of integrating, via homologous recombination with chromosomal sequences, into and disrupting the function of the nucleotide sequence of the endogenous gene. The transgene also can be selectively introduced into a particular cell type, thus inactivating the endogenous gene of the invention in only that cell type (Gu et al., *Science* 265:103, 1984). The regulatory sequences required for such a cell-type specific inactivation will depend upon the particular cell type of interest, and will be apparent to those of skill in the art. These techniques are useful for preparing "knock outs" lacking a functional gene.

[0129] Once transgenic animals have been generated, the expression of the recombinant gene of the invention can be assayed utilizing standard techniques. Initial screening may be accomplished by Southern blot analysis or PCR techniques to determine whether integration of the transgene has taken place. The level of mRNA expression of the transgene in the tissues of the transgenic animals may also be assessed using techniques which include, but are not limited to, Northern blot analysis of tissue samples obtained from the animal, in situ hybridization analysis, and RT-PCR. Biological samples can also be evaluated immunocytochemically using antibodies specific for the product of the transgene of the invention. Samples of tissue expressing the gene of the invention can also be evaluated immunocytochemically using antibodies specific for the product of the transgene of the invention.

[0130] For a review of techniques that can be used to generate and assess transgenic animals, skilled artisans can consult Gordon (*Intl. Rev. Cytol.* 115:171-229, 1989), and may obtain additional guidance from, for example: Hogan et al. "Manipulating the Mouse Embryo" (Cold Spring Harbor

Press, Cold Spring Harbor, N.Y., 1986; Krimpenfort et al., *Bio/Technology* 9:86, 1991; Palmiter et al., *Cell* 41:343, 1985; Kraemer et al., "Genetic Manipulation of the Early Mammalian Embryo," Cold Spring Harbor Press, Cold Spring Harbor, N.Y., 1985; Hammer et al., *Nature* 315:680, 1985; Purcel et al., *Science*, 244:1281, 1986; Wagner et al., U.S. Pat. No. 5,175,385; and Krimpenfort et al., U.S. Pat. No. 5,175,384 (the latter two publications are hereby incorporated by reference).

Anti-Tango-73, Tango-74, Tango-76, Tango-78, or Tango-83 Antibodies

[0131] Human polypeptides of the invention (or immunogenic fragments or analogs) can be used to raise antibodies useful in the invention; such polypeptides can be produced by recombinant techniques or synthesized (see, for example, "Solid Phase Peptide Synthesis," supra; Ausubel et al., supra). In general, the peptides can be coupled to a carrier protein, such as KLH, as described in Ausubel et al., supra, mixed with an adjuvant, and injected into a host mammal. Antibodies can be purified by peptide antigen affinity chromatography.

[0132] In particular, various host animals can be immunized by injection with a polypeptide of the invention. Host animals include rabbits, mice, guinea pigs, and rats. Various adjuvants that can be used to increase the immunological response depend on the host species and include Freund's adjuvant (complete and incomplete), mineral gels such as aluminum hydroxide, surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanin, and dinitrophenol. Potentially useful human adjuvants include BCG (bacille Calmette-Guerin) and *Corynebacterium parvum*. Polyclonal antibodies are heterogeneous populations of antibody molecules that are contained in the sera of the immunized animals.

[0133] Antibodies within the invention therefore include polyclonal antibodies and, in addition, monoclonal antibodies, humanized or chimeric antibodies, single chain antibodies, Fab fragments, F(ab')₂ fragments, and molecules produced using a Fab expression library.

[0134] Monoclonal antibodies, which are homogeneous populations of antibodies to a particular antigen, can be prepared using the polypeptides of the invention described above and standard hybridoma technology (see, for example, Kohler et al., *Nature* 256:495, 1975; Kohler et al., *Eur. J. Immunol.* 6:511, 1976; Kohler et al., *Eur. J. Immunol.* 6:292, 1976; Hammerling et al., "Monoclonal Antibodies and T Cell Hybridomas," Elsevier, N.Y., 1981; Ausubel et al., supra).

[0135] In particular, monoclonal antibodies can be obtained by any technique that provides for the production of antibody molecules by continuous cell lines in culture such as described in Kohler et al., *Nature* 256:495, 1975, and U.S. Pat. No. 4,376,110; the human B-cell hybridoma technique (Kosbor et al., *Immunology Today* 4:72, 1983; Cole et al., *Proc. Natl. Acad. Sci. USA* 80:2026, 1983), and the EBV-hybridoma technique (Cole et al., "Monoclonal Antibodies and Cancer Therapy," Alan R. Liss, Inc., pp. 77-96, 1983). Such antibodies can be of any immunoglobulin class including IgG, IgM, IgE, IgA, IgD and any subclass thereof. The hybridoma producing the mAb of this invention may be

cultivated in vitro or in vivo. The ability to produce high titers of mAbs in vivo makes this a particularly useful method of production.

[0136] Once produced, polyclonal or monoclonal antibodies are tested for specific recognition of polypeptides of the invention by Western blot or immunoprecipitation analysis by standard methods, e.g., as described in Ausubel et al., *supra*. Antibodies that specifically recognize and bind to polypeptides of the invention are useful in the invention. For example, such antibodies can be used in an immunoassay to monitor the level of a polypeptide of the invention produced by a mammal (for example, to determine the amount or subcellular location of a polypeptide of the invention).

[0137] Preferably, antibodies of the invention are produced using fragments of the protein of the invention that lie outside highly conserved regions and appear likely to be antigenic, by criteria, such as high frequency of charged residues. In one specific example, such fragments are generated by standard techniques of PCR, and are then cloned into the pGEX expression vector (Ausubel et al., *supra*). Fusion proteins are expressed in *E. coli* and purified using a glutathione agarose affinity matrix as described in Ausubel, et al., *supra*.

[0138] In some cases it may be desirable to minimize the potential problems of low affinity or specificity of antisera. In such circumstances, two or three fusions can be generated for each protein, and each fusion can be injected into at least two rabbits. Antisera can be raised by injections in a series, preferably including at least three booster injections.

[0139] Antisera may also be checked for its ability to immunoprecipitate recombinant proteins of the invention or control proteins, such as glucocorticoid receptor, CAT, or luciferase.

[0140] The antibodies can be used, for example, in the detection of the polypeptide of the invention in a biological sample as part of a diagnostic assay. Antibodies also can be used in a screening assay to measure the effect of a candidate compound on expression or localization of a polypeptide of the invention. Additionally, such antibodies can be used in conjunction with the gene therapy techniques described to, for example, evaluate normal and/or genetically engineered cells that express nucleic acids or polypeptides of the invention prior to their introduction into the patient. Such antibodies additionally can be used in a method for inhibiting abnormal activity of nucleic acids or polypeptides of the invention.

[0141] In addition, techniques developed for the production of "chimeric antibodies" (Morrison et al., *Proc. Natl. Acad. Sci. USA*, 81:6851, 1984; Neuberger et al., *Nature*, 312:604, 1984; Takeda et al., *Nature*, 314:452, 1984) by splicing the genes from a mouse antibody molecule of appropriate antigen specificity together with genes from a human antibody molecule of appropriate biological activity can be used. A chimeric antibody is a molecule in which different portions are derived from different animal species, such as those having a variable region derived from a murine mAb and a human immunoglobulin constant region.

[0142] Generally, partially human antibodies and fully human antibodies have a longer half-life within the human body than other antibodies. Accordingly, lower dosages and less frequent administration are often possible. Modifica-

tions such as lipidation can be used to stabilize antibodies and to enhance uptake and tissue penetration (e.g., into the brain). A method for lipidation of antibodies is described by Cruikshank et al. ((1997) *J. Acquired Immune Deficiency Syndromes and Human Retrovirology* 14:193).

[0143] Alternatively, techniques described for the production of single chain antibodies (U.S. Pat. Nos. 4,946,778, 4,946,778, and 4,704,692) can be adapted to produce single chain antibodies against polypeptides of the invention. Single chain antibodies are formed by linking the heavy and light chain fragments of the Fv region via an amino acid bridge, resulting in a single chain polypeptide.

[0144] Antibody fragments that recognize and bind to specific epitopes can be generated by known techniques. For example, such fragments include but are not limited to F(ab')₂ fragments that can be produced by pepsin digestion of the antibody molecule, and Fab fragments that can be generated by reducing the disulfide bridges of F(ab')₂ fragments. Alternatively, Fab expression libraries can be constructed (Huse et al., *Science*, 246:1275, 1989) to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity.

[0145] Antibodies to polypeptides of the invention can, in turn, be used to generate anti-idiotypic antibodies that resemble a portion of the protein of the invention using techniques well known to those skilled in the art (see, e.g., Greenspan et al., *FASEB J.* 7:437, 1993; Nissinoff, *J. Immunol.* 147:2429, 1991). For example, antibodies that bind to the protein of the invention and competitively inhibit the binding of a binding partner of the protein can be used to generate anti-idiotypes that resemble a binding partner binding domain of the protein and, therefore, bind and neutralize a binding partner of the protein. Such neutralizing anti-idiotypic antibodies or Fab fragments of such anti-idiotypic antibodies can be used in therapeutic regimens.

[0146] Antibodies can be humanized by methods known in the art. For example, monoclonal antibodies with a desired binding specificity can be commercially humanized (Scotgene, Scotland; Oxford Molecular, Palo Alto, Calif.). Fully human antibodies, such as those expressed in transgenic animals are also features of the invention (Green et al., *Nature Genetics* 7:13-21, 1994; see also U.S. Pat. Nos. 5,545,806 and 5,569,825, both of which are hereby incorporated by reference).

[0147] The methods described herein in which anti-polypeptide-of-the-invention antibodies are employed may be performed, for example, by utilizing pre-packaged diagnostic kits comprising at least one specific polypeptide-of-the-invention antibody reagent described herein, which may be conveniently used, for example, in clinical settings, to diagnose patients exhibiting symptoms of disorders associated with aberrant expression of nucleic acids or polypeptides of the invention.

[0148] An antibody (or fragment thereof) can be conjugated to a therapeutic moiety such as a cytotoxin, a therapeutic agent, or a radioactive agent (e.g., a radioactive metal ion). Cytotoxins and cytotoxic agents include any agent that is detrimental to cells. Examples of such agents include taxol, cytochalasin B, gramicidin D, ethidium bromide, emetine, mitomycin, etoposide, teniposide, vincristine, vinblastine, colchicin, doxorubicin, daunorubicin, dihydroxy

anthracin dione, mitoxantrone, mithramycin, actinomycin D, 1-dehydrotosterone, glucocorticoids, procaine, tetracaine, lidocaine, propranolol, and puromycin and analogs or homologs thereof. Therapeutic agents include, but are not limited to, antimetabolites (e.g., methotrexate, 6-mercaptopurine, 6-thioguanine, cytarabine, and 5-fluorouracil decarbazine), alkylating agents (e.g., mechlorethamine, thioepa chlorambucil, melphalan, carmustine (BSNU) and lomustine (CCNU), cyclophosphamide, busulfan, dibromomannitol, streptozotocin, mitomycin C, and cis-dichlorodiamine platinum (II) (DDP) cisplatin), anthracyclines (e.g., daunorubicin {formerly designated daunomycin} and doxorubicin), antibiotics (e.g., dactinomycin {formerly designated actinomycin}, bleomycin, mithramycin, and anthramycin), and anti-mitotic agents (e.g., vincristine and vinblastine).

[0149] Conjugated antibodies of the invention can be used for modifying a given biological response, the drug moiety not being limited to classical chemical therapeutic agents. For example, the drug moiety can be a protein or polypeptide possessing a desired biological activity. Such proteins include, for example, toxins such as abrin, ricin A, *Pseudomonas* exotoxin, or diphtheria toxin; proteins such as tumor necrosis factor, alpha-interferon, beta-interferon, nerve growth factor, platelet derived growth factor, tissue plasminogen activator; and biological response modifiers such as lymphokines, interleukin-1, interleukin-2, interleukin-6, granulocyte macrophage colony stimulating factor, granulocyte colony stimulating factor, or other growth factors.

[0150] Techniques for conjugating a therapeutic moiety to an antibody are well known (see, e.g., Arnon et al., 1985, "Monoclonal Antibodies For Immunotargeting Of Drugs In Cancer Therapy", in *Monoclonal Antibodies And Cancer Therapy*, Reisfeld et al., Eds., Alan R. Liss, Inc. pp. 243-256; Hellstrom et al., 1987, "Antibodies For Drug Delivery", in *Controlled Drug Delivery*, 2nd ed., Robinson et al., Eds., Marcel Dekker, Inc., pp. 623-653; Thorpe, 1985, "Antibody Carriers Of Cytotoxic Agents In Cancer Therapy: A Review", in *Monoclonal Antibodies '84: Biological And Clinical Applications*, Pinchera et al., Eds., pp. 475-506; "Analysis, Results, And Future Prospective Of The Therapeutic Use Of Radiolabeled Antibody In Cancer Therapy", in *Monoclonal Antibodies For Cancer Detection And Therapy*, Baldwin et al., Eds., Academic Press, pp. 303-316, 1985; and Thorpe et al., 1982, *Immunol. Rev.*, 62:119-158). Alternatively, an antibody can be conjugated to a second antibody to form an antibody heteroconjugate as described by Segal in U.S. Pat. No. 4,676,980.

Antisense Nucleic Acids

[0151] Treatment regimes based on an "antisense" approach involve the design of oligonucleotides (either DNA or RNA) that are complementary to mRNA of the invention. These oligonucleotides bind to the complementary mRNA transcripts of the invention and prevent translation. Absolute complementarity, although preferred, is not required. A sequence "complementary" to a portion of an RNA, as referred to herein, means a sequence having sufficient complementarity to be able to hybridize with the RNA, forming a stable duplex; in the case of double-stranded antisense nucleic acids, a single strand of the duplex DNA may be tested, or triplex formation may be

assayed. The ability to hybridize will depend on both the degree of complementarity and the length of the antisense nucleic acid. Generally, the longer the hybridizing nucleic acid, the more base mismatches with an RNA it may contain and still form a stable duplex (or triplex, as the case may be). One skilled in the art can ascertain a tolerable degree of mismatch by use of standard procedures to determine the melting point of the hybridized complex.

[0152] Oligonucleotides that are complementary to the 5' end of the message, e.g., the 5' untranslated sequence up to and including the AUG initiation codon, should work most efficiently at inhibiting translation. However, sequences complementary to the 3' untranslated sequences of mRNAs recently have been shown to be effective at inhibiting translation of mRNAs as well (Wagner, *Nature* 372:333, 1984). Thus, oligonucleotides complementary to either the 5' or 3' non-translated, non-coding regions of the gene or mRNA could be used in an antisense approach to inhibit translation of endogenous mRNA. Oligonucleotides complementary to the 5' untranslated region of the mRNA should include the complement of the AUG start codon.

[0153] Antisense oligonucleotides complementary to mRNA coding regions are less efficient inhibitors of translation but could be used in accordance with the invention. Whether designed to hybridize to the 5', 3', or coding region of an mRNA, antisense nucleic acids should be at least six nucleotides in length, and are preferably oligonucleotides ranging from 6 to about 50 nucleotides in length. In specific aspects the oligonucleotide is at least 10 nucleotides, at least 17 nucleotides, at least 25 nucleotides, or at least 50 nucleotides.

[0154] Regardless of the choice of target sequence, it is preferred that in vitro studies are first performed to quantitate the ability of the antisense oligonucleotide to inhibit gene expression. It is preferred that these studies utilize controls that distinguish between antisense gene inhibition and nonspecific biological effects of oligonucleotides. It is also preferred that these studies compare levels of the target RNA or protein with that of an internal control RNA or protein. Additionally, it is envisioned that results obtained using the antisense oligonucleotide are compared with those obtained using a control oligonucleotide. It is preferred that the control oligonucleotide is of approximately the same length as the test oligonucleotide and that the nucleotide sequence of the oligonucleotide differs from the antisense sequence no more than is necessary to prevent specific hybridization to the target sequence.

[0155] The oligonucleotides can be DNA or RNA or chimeric mixtures or derivatives or modified versions thereof, single-stranded or double-stranded. The oligonucleotide can be modified at the base moiety, sugar moiety, or phosphate backbone, for example, to improve stability of the molecule, hybridization, etc. The oligonucleotide may include other appended groups such as peptides (e.g., for targeting host cell receptors in vivo), or agents facilitating transport across the cell membrane (as described, e.g., in Letsinger et al., *Proc. Natl. Acad. Sci. USA* 86:6553, 1989; Lemaire et al., *Proc. Natl. Acad. Sci. USA* 84:648, 1987; PCT Publication No. WO 88/09810) or the blood-brain barrier (see, for example, PCT Publication No. WO 89/10134), or hybridization-triggered cleavage agents (see, for example, Krol et al., *BioTechniques* 6:958, 1988), or

intercalating agents (see, for example, Zon, *Pharm. Res.* 5:539, 1988). To this end, the oligonucleotide can be conjugated to another molecule, e.g., a peptide, hybridization triggered cross-linking agent, transport agent, or hybridization-triggered cleavage agent.

[0156] The antisense oligonucleotide may comprise at least one modified base moiety which is selected from the group including, but not limited to, 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxymethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethyl-aminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methyl ester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 2-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine.

[0157] The antisense oligonucleotide may also comprise at least one modified sugar moiety selected from the group including, but not limited to, arabinose, 2-fluoroarabinose, xylulose, and hexose.

[0158] In yet another embodiment, the antisense oligonucleotide comprises at least one modified phosphate backbone selected from the group consisting of a phosphorothioate, a phosphorodithioate, a phosphoramidothioate, a phosphoramidate, a phosphordiamidate, a methylphosphonate, an alkyl phosphotriester, and a formacetal, or an analog of any of these backbones.

[0159] In yet another embodiment, the antisense oligonucleotide is an α -anomeric oligonucleotide. An α -anomeric oligonucleotide forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual β -units, the strands run parallel to each other (Gautier et al., *Nucl. Acids Res.* 15:6625, 1987). The oligonucleotide is a 2'-O-methylribonucleotide (Inoue et al., *Nucl. Acids Res.* 15:6131, 1987), or a chimeric RNA-DNA analog (Inoue et al., *FEBS Lett.* 215:327, 1987).

[0160] Antisense oligonucleotides of the invention can be synthesized by standard methods known in the art, e.g., by use of an automated DNA synthesizer (such as are commercially available from Biosearch, Applied Biosystems, etc.). As examples, phosphorothioate oligonucleotides can be synthesized by the method of Stein et al. (*Nucl. Acids Res.* 16:3209, 1988), and methylphosphonate oligonucleotides can be prepared by use of controlled pore glass polymer supports (Sarin et al., *Proc. Natl. Acad. Sci. USA* 85:7448, 1988).

[0161] The antisense molecules should be delivered to cells that express nucleic acids or polypeptides of the invention in vivo. A number of methods have been developed for delivering antisense DNA or RNA to cells; e.g., antisense molecules can be injected directly into the tissue site, or modified antisense molecules, designed to target the

desired cells (e.g., antisense linked to peptides or antibodies that specifically bind receptors or antigens expressed on the target cell surface) can be administered systemically.

[0162] However, it is often difficult to achieve intracellular concentrations of the antisense molecule sufficient to suppress translation of endogenous mRNAs. Therefore, a preferred approach uses a recombinant DNA construct in which the antisense oligonucleotide is placed under the control of a strong pol III or pol II promoter. The use of such a construct to transfect target cells in the patient will result in the transcription of sufficient amounts of single stranded RNAs that will form complementary base pairs with the endogenous transcripts of nucleic acids of the invention and thereby prevent translation of the endogenous mRNA. For example, a vector can be introduced in vivo such that it is taken up by a cell and directs the transcription of an antisense RNA. Such a vector can remain episomal or become chromosomally integrated, as long as it can be transcribed to produce the desired antisense RNA.

[0163] Such vectors can be constructed by recombinant DNA technology methods standard in the art. Vectors can be plasmid, viral, or others known in the art, used for replication and expression in mammalian cells. Expression of the sequence encoding the antisense RNA can be by any promoter known in the art to act in mammalian, preferably human cells. Such promoters can be inducible or constitutive. Such promoters include, but are not limited to: the SV40 early promoter region (Bernoist et al., *Nature* 290:304, 1981); the promoter contained in the 3' long terminal repeat of Rous sarcoma virus (Yamamoto et al., *Cell* 22:787-797, 1988); the herpes thymidine kinase promoter (Wagner et al., *Proc. Natl. Acad. Sci. USA* 78:1441, 1981); or the regulatory sequences of the metallothionein gene (Brinster et al., *Nature* 296:39, 1988).

Ribozymes

[0164] Ribozyme molecules designed to catalytically cleave mRNA transcripts of nucleic acids of the invention can be used to prevent translation and expression of mRNA of the invention. (see, e.g., PCT Publication WO 90/11364; Saraver et al., *Science* 247:1222, 1990). While various ribozymes that cleave mRNA at site-specific recognition sequences can be used to destroy mRNAs of the invention, the use of hammerhead ribozymes is preferred. Hammerhead ribozymes cleave mRNAs at locations dictated by flanking regions that form complementary base pairs with the target mRNA. The sole requirement is that the target mRNA have the following sequence of two bases: 5'-UG-3'. The construction and production of hammerhead ribozymes is well known in the art (Haseloff et al., *Nature* 334:585, 1988). There are numerous examples of potential hammerhead ribozyme cleavage sites within the nucleotide sequence of human cDNA of the invention. Preferably, the ribozyme is engineered so that the cleavage recognition site is located near the 5' end of the mRNA, i.e., to increase efficiency and minimize the intracellular accumulation of non-functional mRNA transcripts.

[0165] The ribozymes of the present invention also include RNA endoribonucleases (hereinafter "Cech-type ribozymes"), such as the one that occurs naturally in *Tetrahymena thermophila* (known as the IVS or L-19 IVS RNA), and which has been extensively described by Cech

and his collaborators (Zaug et al., *Science* 224:574, 1984; Zaug et al., *Science*, 231:470, 1986; Zug et al., *Nature* 324:429, 1986; PCT Application No. WO 88/04300; and Been et al., *Cell* 47:207, 1986). The Cech-type ribozymes have an eight base-pair sequence that hybridizes to a target RNA sequence, whereafter cleavage of the target RNA takes place. The invention encompasses those Cech-type ribozymes that target eight base-pair active site sequences present in nucleic acids of the invention.

[0166] As in the antisense approach, the ribozymes can be composed of modified oligonucleotides (e.g., for improved stability, targeting, etc.), and should be delivered to cells which express nucleic acids or polypeptides of the invention in vivo. A preferred method of delivery involves using a DNA construct "encoding" the ribozyme under the control of a strong constitutive pol III or pol II promoter, so that transfected cells will produce sufficient quantities of the ribozyme to destroy endogenous messages and inhibit translation. Because ribozymes, unlike antisense molecules, are catalytic, a lower intracellular concentration is required for efficiency.

Other Methods for Modulating Tango-73,
Tango-74, Tango-76, Tango-78, and Tango-83
Expression

[0167] Endogenous expression of a gene of the invention can also be modulated by inactivating the endogenous gene or its promoter using targeted homologous recombination (see, e.g., U.S. Pat. No. 5,464,764). For example, a mutant, non-functional gene of the invention (or a completely unrelated DNA sequence) flanked by DNA homologous to the endogenous gene of the invention (either the coding regions or regulatory regions of the gene of the invention) can be used, with or without a selectable marker and/or a negative selectable marker, to transfect cells that express the endogenous gene of the invention in vivo. Insertion of the DNA construct, via targeted homologous recombination, results in inactivation of the gene of the invention. Such approaches are particularly suited for use in the agricultural field where modifications to ES (embryonic stem) cells can be used to generate animal offspring with an inactive gene of the invention. However, this approach can be adapted for use in humans, provided the recombinant DNA constructs are directly administered or targeted to the required site in vivo using appropriate viral vectors.

[0168] Alternatively, endogenous expression of a gene of the invention can be modulated by targeting deoxyribonucleotide sequences complementary to the regulatory region of the gene of the invention (i.e., the promoter and/or enhancers of a gene of the invention) to form triple helical structures that prevent transcription of the gene of the invention in target cells in the body (Helene, *Anticancer Drug Res.* 6:569, 1981; Helene et al., *Ann. N.Y. Acad. Sci.* 660:27, 1992; and Maher, *Bioassays* 14:807, 1992).

[0169] The invention includes methods for preparing pharmaceutical compositions for modulating the expression or activity of a polypeptide or nucleic acid of the invention. Such methods comprise formulating a pharmaceutically acceptable carrier with an agent which modulates expression or activity of a polypeptide or nucleic acid of the invention. Such compositions can further include additional active agents. Thus, the invention further includes methods for

preparing a pharmaceutical composition by formulating a pharmaceutically acceptable carrier with an agent that modulates expression or activity of a polypeptide or nucleic acid of the invention and one or more additional active compounds.

[0170] The agent that modulates expression or activity can, for example, be a small molecule. For example, such small molecules include peptides, peptidomimetics, amino acids, amino acid analogs, polynucleotides, polynucleotide analogs, nucleotides, nucleotide analogs, organic or inorganic compounds (i.e., including heteroorganic and organometallic compounds) having a molecular weight less than about 10,000 grams per mole, organic or inorganic compounds having a molecular weight less than about 5,000 grams per mole, organic or inorganic compounds having a molecular weight less than about 1,000 grams per mole, organic or inorganic compounds having a molecular weight less than about 500 grams per mole, and salts, esters, and other pharmaceutically acceptable forms of such compounds.

[0171] It is understood that appropriate doses of small molecule agents and protein or polypeptide agents depends upon a number of factors within the ken of the ordinarily skilled physician, veterinarian, or researcher. The dose(s) of these agents will vary, for example, depending upon the identity, size, and condition of the subject or sample being treated, further depending upon the route by which the composition is to be administered, if applicable, and the effect which the practitioner desires the agent to have upon the nucleic acid or polypeptide of the invention. Examples of doses of a small molecule include milligram or microgram amounts per kilogram of subject or sample weight (e.g., about 1 microgram per kilogram to about 500 milligrams per kilogram, about 100 micrograms per kilogram to about 5 milligrams per kilogram, or about 1 microgram per kilogram to about 50 micrograms per kilogram). Examples of doses of a protein or polypeptide include gram, milligram or microgram amounts per kilogram of subject or sample weight (e.g., about 1 microgram per kilogram to about 5 grams per kilogram, about 100 micrograms per kilogram to about 500 milligrams per kilogram, or about 1 milligram per kilogram to about 50 milligrams per kilogram). For antibodies, examples of dosages are from about 0.1 milligram per kilogram to 100 milligrams per kilogram of body weight (generally 10 milligrams per kilogram to 20 milligrams per kilogram). If the antibody is to act in the brain, a dosage of 50 milligrams per kilogram to 100 milligrams per kilogram is usually appropriate. It is furthermore understood that appropriate doses of one of these agents depend upon the potency of the agent with respect to the expression or activity to be modulated. Such appropriate doses can be determined using the assays described herein. When one or more of these agents is to be administered to an animal (e.g., a human) in order to modulate expression or activity of a polypeptide or nucleic acid of the invention, a physician, veterinarian, or researcher can, for example, prescribe a relatively low dose at first, subsequently increasing the dose until an appropriate response is obtained. In addition, it is understood that the specific dose level for any particular animal subject will depend upon a variety of factors including the activity of the specific agent employed, the age, body weight, general health, gender, and diet of the subject, the time of administration, the route of administration, the rate

of excretion, any drug combination, and the degree of expression or activity to be modulated.

[0172] As an alternative to making determinations based on the absolute expression level of selected genes, determinations may be based on the normalized expression levels of these genes. Expression levels are normalized by correcting the absolute expression level of a gene encoding a polypeptide of the invention by comparing its expression to the expression of a different gene, e.g., a housekeeping gene that is constitutively expressed. Suitable genes for normalization include housekeeping genes such as the actin gene. This normalization allows the comparison of the expression level in one sample (e.g., a patient sample), to another sample, or between samples from different sources.

[0173] Alternatively, the expression level can be provided as a relative expression level. To determine a relative expression level of a gene, the level of expression of the gene is determined for 10 or more samples of different endothelial (e.g. intestinal endothelium, airway endothelium, or other mucosal epithelium) cell isolates, preferably 50 or more samples, prior to the determination of the expression level for the sample in question. The mean expression level of each of the genes assayed in the larger number of samples is determined and this is used as a baseline expression level for the gene(s) in question. The expression level of the gene determined for the test sample (absolute level of expression) is then divided by the mean expression value obtained for that gene. This provides a relative expression level and aids in identifying extreme cases of disorders associated with aberrant expression of a gene encoding a polypeptide of the invention protein or with aberrant expression of a ligand thereof.

[0174] Preferably, the samples used in the baseline determination will be from either or both of cells which aberrantly express a gene encoding a polypeptide of the invention or a ligand thereof (i.e. 'diseased cells') and cells which express a gene encoding a polypeptide of the invention at a normal level or a ligand thereof (i.e. 'normal' cells). The choice of the cell source is dependent on the use of the relative expression level. Using expression found in normal tissues as a mean expression score aids in validating whether aberrance in expression of a gene encoding a polypeptide of the invention occurs specifically in diseased cells. Such a use is particularly important in identifying whether a gene encoding a polypeptide of the invention can serve as a target gene. In addition, as more data is accumulated, the mean expression value can be revised, providing improved relative expression values based on accumulated data. Expression data from endothelial cells (e.g. mucosal-endothelial cells) provides a means for grading the severity of the disorder.

Detecting Proteins Associated with Tango-73, Tango-74, Tango-76, Tango-78, or Tango-83

[0175] The invention also features polypeptides that interact with Tango-73, Tango-74, Tango-76, Tango-78, or Tango-83. Any method suitable for detecting protein-protein interactions may be employed for identifying transmembrane proteins, intracellular, or extracellular proteins that interact with polypeptides of the invention. Among the traditional methods which may be employed are co-immunoprecipitation, cross-linking and co-purification through

gradients or chromatographic columns of cell lysates or proteins obtained from cell lysates and the use of polypeptides of the invention to identify proteins in the lysate that interact with polypeptides of the invention. For these assays, the polypeptide of the invention can be full length polypeptide of the invention, a soluble extracellular domain of a polypeptide of the invention, or some other suitable polypeptide of the invention. Once isolated, such an interacting protein can be identified and cloned and then used, in conjunction with standard techniques, to identify proteins with which it interacts. For example, at least a portion of the amino acid sequence of a protein which interacts with the polypeptide of the invention can be ascertained using techniques well known to those of skill in the art, such as via the Edman degradation technique. The amino acid sequence obtained may be used as a guide for the generation of oligonucleotide mixtures that can be used to screen for gene sequences encoding the interacting protein. Screening may be accomplished, for example, by standard hybridization or PCR techniques. Techniques for the generation of oligonucleotide mixtures and the screening are well-known. (Ausubel, supra; and "PCR Protocols: A Guide to Methods and Applications," Innis et al., eds. Academic Press, Inc., NY, 1990).

[0176] Additionally, methods may be employed which result directly in the identification of genes which encode proteins which interact with polypeptides of the invention. These methods include, for example, screening expression libraries, in a manner similar to the well known technique of antibody probing of λ gt11 libraries, using labeled polypeptide of the invention or a fusion protein of the invention, e.g., a polypeptide of the invention or domain thereof fused to a marker such as an enzyme, fluorescent dye, a luminescent protein, or to an IgFc domain.

[0177] There are also methods capable of detecting protein interaction. A method that detects protein interactions in vivo is the two-hybrid system (Chien et al., *Proc. Natl. Acad. Sci. USA*, 88:9578, 1991). A kit for practicing this method is available from Clontech (Palo Alto, Calif.).

Identification of Tango-73, Tango-74, Tango-76, Tango-78, or Tango-83 Receptors

[0178] Receptors of polypeptides of the invention can be identified as follows. First cells or tissues that bind polypeptides of the invention are identified. An expression library is prepared using mRNA isolated from cells that bind to polypeptides of the invention. The expression library is used to transfect; eukaryotic cells, e.g., CHO cells. Detectably labeled polypeptides of the invention are used to identify clones that bind polypeptides of the invention. These clones are isolated and purified. The expression plasmid is then isolated from polypeptides-of-the-invention-binding clones. These expression plasmids will encode putative receptors of polypeptides of the invention.

[0179] Cells or tissues bearing a receptor of a polypeptide of the invention can be identified by exposing detectably labeled polypeptide of the invention to various cells lines and tissues. Alternatively a microphysiometer can be used to determine whether a selected cells responds to the presence of a cell receptor ligand (McConnel et al., *Science* 257:1906, 1992).

Compounds that bind Tango-73, Tango-74, Tango-76, Tango-78, or Tango-83

[0180] Compounds that bind nucleic acids or polypeptides of the invention can be identified using any standard binding assay. For example, candidate compounds can be bound to a solid support. A nucleic acid or polypeptide of the invention is then exposed to the immobilized compound and binding is measured (European Patent Application 84/03564).

[0181] In one embodiment, the invention provides assays for screening candidate or test compounds that bind with or modulate the activity of the membrane-bound form of a polypeptide of the invention or biologically active portion thereof. The test compounds of the present invention can be obtained using any of the numerous approaches in combinatorial library methods known in the art, including: biological libraries; spatially addressable parallel solid phase or solution phase libraries; synthetic library methods requiring deconvolution; the "one-bead one-compound" library method; and synthetic library methods using affinity chromatography selection. The biological library approach is limited to peptide libraries, while the other four approaches are applicable to peptide, non-peptide oligomer, or small molecule libraries of compounds (Lam (1997) *Anticancer Drug Des.* 12:145).

[0182] Examples of methods useful for the synthesis of molecular libraries can be found in the art, for example in: DeWitt et al. (1993) *Proc. Natl. Acad. Sci. USA* 90:6909; Erb et al. (1994) *Proc. Natl. Acad. Sci. USA* 91:11422; Zuckermann et al. (1994) *J. Med. Chem.* 37:2678; Cho et al. (1993) *Science* 261:1303; Carrell et al. (1994) *Angew. Chem. Int. Ed. Engl.* 33:2059; Carell et al. (1994) *Angew. Chem. Int. Ed. Engl.* 33:2061; and Gallop et al. (1994) *J. Med. Chem.* 37:1233.

[0183] Libraries of compounds can be presented in solution (e.g., Houghten (1992) *Bio/Techniques* 13:412-421), or on beads (Lam (1991) *Nature* 354:82-84), chips (Fodor (1993) *Nature* 364:555-556), bacteria (U.S. Pat. No. 5,223,409), spores (U.S. Pat. Nos. 5,571,698; 5,403,484; and 5,223,409), plasmids (Cull et al. (1992) *Proc. Natl. Acad. Sci. USA* 89:1865-1869) or phage (Scott and Smith (1990) *Science* 249:386-390; Devlin (1990) *Science* 249:404-406; Cwirla et al. (1990) *Proc. Natl. Acad. Sci. USA* 87:6378-6382; and Felici (1991) *J. Mol. Biol.* 222:301-310).

[0184] In one embodiment, an assay is a cell-based assay in which a cell that expresses a membrane-bound form of a polypeptide of the invention, or a biologically active portion thereof, on the cell surface is contacted with a test compound and the ability of the test compound to bind with the polypeptide is determined. The cell, for example, can be a yeast cell or a cell of mammalian origin. Determining the ability of the test compound to bind with the polypeptide can be accomplished, for example, by coupling the test compound with a radioisotope or enzymatic label such that binding of the test compound to the polypeptide or biologically active portion thereof can be determined by detecting the labeled compound in a complex. For example, test compounds can be labeled with ^{125}I , ^{35}S , ^{14}C , or ^3H , either directly or indirectly, and the radioisotope detected by direct counting of radio-emission or by scintillation counting. Alternatively, test compounds can be enzymatically labeled with, for example, horseradish peroxidase, alkaline phos-

phatase, or luciferase, and the enzymatic label detected by determination of conversion of an appropriate substrate to product. In one embodiment, the assay comprises contacting a cell which expresses a membrane-bound form of a polypeptide of the invention, or a biologically active portion thereof, on the cell surface with a known compound that binds the polypeptide to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with the polypeptide, wherein determining the ability of the test compound to interact with the polypeptide comprises determining the ability of the test compound to preferentially bind with the polypeptide or a biologically active portion thereof as compared to the known compound.

[0185] In another embodiment, the assay involves assessment of an activity characteristic of the polypeptide, wherein binding of the test compound with the polypeptide or a biologically active portion thereof alters (i.e., increases or decreases) the activity of the polypeptide.

Uses and Methods of the Invention

[0186] The nucleic acid molecules, proteins, protein homologs, and antibodies described herein can be used in one or more of the following methods: a) screening assays; b) detection assays (e.g., chromosomal mapping, tissue typing, forensic biology); c) predictive medicine (e.g., diagnostic assays, prognostic assays, monitoring clinical trials, and pharmacogenomics); and d) methods of treatment (e.g., therapeutic and prophylactic). For example, polypeptides of the invention can be used to (i) modulate cellular proliferation; (ii) modulate cellular differentiation; and/or (iii) modulate cellular adhesion. The isolated nucleic acid molecules of the invention can be used to express proteins (e.g., via a recombinant expression vector in a host cell in gene therapy applications), to detect mRNA (e.g., in a biological sample) or a genetic lesion, and to modulate activity of a polypeptide of the invention. In addition, the polypeptides of the invention can be used to screen drugs or compounds which modulate activity or expression of a polypeptide of the invention as well as to treat disorders characterized by insufficient or excessive production of a protein of the invention or production of a form of a protein of the invention which has decreased or aberrant activity compared to the wild type protein. In addition, the antibodies of the invention can be used to detect and isolate a protein of the invention and modulate activity of a protein of the invention.

[0187] This invention further pertains to novel agents identified by the above-described screening assays and uses thereof for treatments as described herein.

A. Screening Assays

[0188] The invention provides a method (also referred to herein as a Screening assay@) for identifying modulators, i.e., candidate or test compounds or agents (e.g., peptides, peptidomimetics, small molecules or other drugs) which bind to polypeptide of the invention or have a stimulatory or inhibitory effect on, for example, expression or activity of a polypeptide of the invention.

[0189] In one embodiment, the invention provides assays for screening candidate or test compounds that bind to or modulate the activity of the membrane-bound form of a polypeptide of the invention or biologically active portion

thereof. The test compounds of the present invention can be obtained using any of the numerous approaches in combinatorial library methods known in the art, including: biological libraries; spatially addressable parallel solid phase or solution phase libraries; synthetic library methods requiring deconvolution; the Aone-bead one-compound@ library method; and synthetic library methods using affinity chromatography selection. The biological library approach is limited to peptide libraries, while the other four approaches are applicable to peptide, non-peptide oligomer or small molecule libraries of compounds (Lam (1997) *Anticancer Drug Des.* 12:145).

[0190] Examples of methods for the synthesis of molecular libraries can be found in the art, for example in: DeWitt et al. (1993) *Proc. Natl. Acad. Sci. USA* 90:6909; Erb et al. (1994) *Proc. Natl. Acad. Sci. USA* 91:11422; Zuckermann et al. (1994) *J. Med. Chem.* 37:2678; Cho et al. (1993) *Science* 261:1303; Carrell et al. (1994) *Angew. Chem. Int. Ed. Engl.* 33:2059; Carrell et al. (1994) *Angew. Chem. Int. Ed. Engl.* 33:2061; and Gallop et al. (1994) *J. Med. Chem.* 37:1233.

[0191] Libraries of compounds may be presented in solution (e.g., Houghten (1992) *Bio/Techniques* 13:412-421), or on beads (Lam (1991) *Nature* 354:82-84), chips (Fodor (1993) *Nature* 364:555-556), bacteria (U.S. Pat. No. 5,223,409), spores (U.S. Pat. Nos. 5,571,698; 5,403,484; and 5,223,409), plasmids (Cull et al. (1992) *Proc. Natl. Acad. Sci. USA* 89:1865-1869) or phage (Scott and Smith (1990) *Science* 249:386-390; Devlin (1990) *Science* 249:404-406; Cwirla et al. (1990) *Proc. Natl. Acad. Sci. USA* 87:6378-6382; and Felici (1991) *J. Mol. Biol.* 222:301-310).

[0192] In one embodiment, an assay is a cell-based assay in which a cell that expresses a membrane-bound form of a polypeptide of the invention, or a biologically active portion thereof, on the cell surface is contacted with a test compound and the ability of the test compound to bind to the polypeptide determined. The cell, for example, can be a yeast cell or a cell of mammalian origin. Determining the ability of the test compound to bind to the polypeptide can be accomplished, for example, by coupling the test compound with a radioisotope or enzymatic label such that binding of the test compound to the polypeptide or biologically active portion thereof can be determined by detecting the labeled compound in a complex. For example, test compounds can be labeled with ^{125}I , ^{35}S , ^{14}C , or ^3H , either directly or indirectly, and the radioisotope detected by direct counting of radioemission or by scintillation counting. Alternatively, test compounds can be enzymatically labeled with, for example, horseradish peroxidase, alkaline phosphatase, or luciferase, and the enzymatic label detected by determination of conversion of an appropriate substrate to product. In a preferred embodiment, the assay comprises contacting a cell which expresses a membrane-bound form of a polypeptide of the invention, or a biologically active portion thereof, on the cell surface with a known compound which binds the polypeptide to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with the polypeptide, wherein determining the ability of the test compound to interact with the polypeptide comprises determining the ability of the test compound to preferentially bind to the polypeptide or a biologically active portion thereof as compared to the known compound.

[0193] In another embodiment, an assay is a cell-based assay comprising contacting a cell expressing a membrane-bound form of a polypeptide of the invention, or a biologically active portion thereof, on the cell surface with a test compound and determining the ability of the test compound to modulate (e.g., stimulate or inhibit) the activity of the polypeptide or biologically active portion thereof. Determining the ability of the test compound to modulate the activity of the polypeptide or a biologically active portion thereof can be accomplished, for example, by determining the ability of the polypeptide protein to bind to or interact with a target molecule.

[0194] Determining the ability of a polypeptide of the invention to bind to or interact with a target molecule can be accomplished by one of the methods described above for determining direct binding. As used herein, a "target molecule" is a molecule with which a selected polypeptide (e.g., a polypeptide of the invention) binds or interacts with in nature, for example, a molecule on the surface of a cell which expresses the selected protein, a molecule on the surface of a second cell, a molecule in the extracellular milieu, a molecule associated with the internal surface of a cell membrane or a cytoplasmic molecule. A target molecule can be a polypeptide of the invention or some other polypeptide or protein. For example, a target molecule can be a component of a signal transduction pathway which facilitates transduction of an extracellular signal (e.g., a signal generated by binding of a compound to a polypeptide of the invention) through the cell membrane and into the cell or a second intercellular protein which has catalytic activity or a protein which facilitates the association of downstream signaling molecules with a polypeptide of the invention. Determining the ability of a polypeptide of the invention to bind to or interact with a target molecule can be accomplished by determining the activity of the target molecule. For example, the activity of the target molecule can be determined by detecting induction of a cellular second messenger of the target (e.g., intracellular Ca^{2+} , diacylglycerol, IP3, etc.), detecting catalytic/enzymatic activity of the target on an appropriate substrate, detecting the induction of a reporter gene (e.g., a regulatory element that is responsive to a polypeptide of the invention operably linked to a nucleic acid encoding a detectable marker, e.g., luciferase), or detecting a cellular response, for example, cellular differentiation, or cell proliferation.

[0195] In yet another embodiment, an assay of the present invention is a cell-free assay comprising contacting a polypeptide of the invention or biologically active portion thereof with a test compound and determining the ability of the test compound to bind to the polypeptide or biologically active portion thereof. Binding of the test compound to the polypeptide can be determined either directly or indirectly as described above. In a preferred embodiment, the assay includes contacting the polypeptide of the invention or biologically active portion thereof with a known compound which binds the polypeptide to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with the polypeptide, wherein determining the ability of the test compound to interact with the polypeptide comprises determining the ability of the test compound to preferentially bind to the polypeptide or biologically active portion thereof as compared to the known compound.

[0196] In another embodiment, an assay is a cell-free assay comprising contacting a polypeptide of the invention or biologically active portion thereof with a test compound and determining the ability of the test compound to modulate (e.g., stimulate or inhibit) the activity of the polypeptide or biologically active portion thereof. Determining the ability of the test compound to modulate the activity of the polypeptide can be accomplished, for example, by determining the ability of the polypeptide to bind to a target molecule by one of the methods described above for determining direct binding. In an alternative embodiment, determining the ability of the test compound to modulate the activity of the polypeptide can be accomplished by determining the ability of the polypeptide of the invention to further modulate the target molecule. For example, the catalytic/enzymatic activity of the target molecule on an appropriate substrate can be determined as previously described.

[0197] In yet another embodiment, the cell-free assay comprises contacting a polypeptide of the invention or biologically active portion thereof with a known compound which binds the polypeptide to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with the polypeptide, wherein determining the ability of the test compound to interact with the polypeptide comprises determining the ability of the polypeptide to preferentially bind to or modulate the activity of a target molecule.

[0198] The cell-free assays of the present invention are amenable to use of both a soluble form or the membrane-bound form of a polypeptide of the invention. In the case of cell-free assays comprising the membrane-bound form of the polypeptide, it may be desirable to utilize a solubilizing agent such that the membrane-bound form of the polypeptide is maintained in solution. Examples of such solubilizing agents include non-ionic detergents such as n-octylglucoside, n-dodecylglucoside, n-octylmaltoside, octanoyl-N-methylglucamide, decanoyl-N-methylglucamide, Triton X-100, Triton X-114, Thesit, Isotridecylpoly(ethylene glycol ether)n, 3-[(3-cholamidopropyl)dimethylamminio]-1-propane sulfonate (CHAPS), 3-[(3-cholamidopropyl)dimethylamminio]-2-hydroxy-1-propane sulfonate (CHAPSO), or N-dodecyl=N,N-dimethyl-3-ammonio-1-propane sulfonate.

[0199] In more than one embodiment of the above assay methods of the present invention, it may be desirable to immobilize either the polypeptide of the invention or its target molecule to facilitate separation of complexed from uncomplexed forms of one or both of the proteins, as well as to accommodate automation of the assay. Binding of a test compound to the polypeptide, or interaction of the polypeptide with a target molecule in the presence and absence of a candidate compound, can be accomplished in any vessel suitable for containing the reactants. Examples of such vessels include microtitre plates, test tubes, and micro-centrifuge tubes. In one embodiment, a fusion protein can be provided which adds a domain that allows one or both of the proteins to be bound to a matrix. For example, glutathione-S-transferase fusion proteins or glutathione-S-transferase fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical; St. Louis, Mo.) or glutathione derivatized microtitre plates, which are then combined with the test compound or the test compound and either the non-adsorbed target protein or A polypeptide of the inven-

tion, and the mixture incubated under conditions conducive to complex formation (e.g., at physiological conditions for salt and pH). Following incubation, the beads or microtitre plate wells are washed to remove any unbound components and complex formation is measured either directly or indirectly, for example, as described above. Alternatively, the complexes can be dissociated from the matrix, and the level of binding or activity of the polypeptide of the invention can be determined using standard techniques.

[0200] Other techniques for immobilizing proteins on matrices can also be used in the screening assays of the invention. For example, either the polypeptide of the invention or its target molecule can be immobilized utilizing conjugation of biotin and streptavidin. Biotinylated polypeptide of the invention or target molecules can be prepared from biotin-NHS (N-hydroxy-succinimide) using techniques well known in the art (e.g., biotinylation kit, Pierce Chemicals; Rockford, Ill.), and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). Alternatively, antibodies reactive with the polypeptide of the invention or target molecules but which do not interfere with binding of the polypeptide of the invention to its target molecule can be derivatized to the wells of the plate, and unbound target or polypeptide of the invention trapped in the wells by antibody conjugation. Methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies reactive with the polypeptide of the invention or target molecule, as well as enzyme-linked assays which rely on detecting an enzymatic activity associated with the polypeptide of the invention or target molecule.

[0201] In another embodiment, modulators of expression of a polypeptide of the invention are identified in a method in which a cell is contacted with a candidate compound and the expression of the selected mRNA or protein (i.e., the mRNA or protein corresponding to a polypeptide or nucleic acid of the invention) in the cell is determined. The level of expression of the selected mRNA or protein in the presence of the candidate compound is compared to the level of expression of the selected mRNA or protein in the absence of the candidate compound. The candidate compound can then be identified as a modulator of expression of the polypeptide of the invention based on this comparison. For example, when expression of the selected mRNA or protein is greater (statistically significantly greater) in the presence of the candidate compound than in its absence, the candidate compound is identified as a stimulator of the selected mRNA or protein expression. Alternatively, when expression of the selected mRNA or protein is less (statistically significantly less) in the presence of the candidate compound than in its absence, the candidate compound is identified as an inhibitor of the selected mRNA or protein expression. The level of the selected mRNA or protein expression in the cells can be determined by methods described herein.

[0202] In yet another aspect of the invention, a polypeptide of the inventions can be used as Abait proteins@ in a two-hybrid assay or three hybrid assay (see, e.g., U.S. Pat. No. 5,283,317; Zervos et al. (1993) *Cell* 72:223-232; Madura et al. (1993) *J. Biol. Chem.* 268:12046-12054; Bartel et al. (1993) *Bio/Techniques* 14:920-924; Twabuchi et al. (1993) *Oncogene* 8:1693-1696; and PCT Publication No. WO 94/10300) to identify other proteins that bind to or

interact with the polypeptide of the invention and modulate activity of the polypeptide of the invention. Such binding proteins are also likely to be involved in the propagation of signals by the polypeptide of the inventions as, for example, upstream or downstream elements of a signaling pathway involving the polypeptide of the invention.

[0203] This invention further pertains to novel agents identified by the above-described screening assays and uses thereof for treatments as described herein.

B. Detection Assays

[0204] Portions or fragments of the cDNA sequences identified herein (and the corresponding complete gene sequences) can be used in numerous ways as polynucleotide reagents. For example, these sequences can be used to: (i) map their respective genes on a chromosome and, thus, locate gene regions associated with genetic disease; (ii) identify an individual from a minute biological sample (tissue typing); and (iii) aid in forensic identification of a biological sample. These applications are described in the subsections below.

[0205] 1. Chromosome Mapping

[0206] Once the sequence (or a portion of the sequence) of a gene has been isolated, this sequence can be used to map the location of the gene on a chromosome. Accordingly, nucleic acid molecules described herein or fragments thereof, can be used to map the location of the corresponding genes on a chromosome. The mapping of the sequences to chromosomes is an important first step in correlating these sequences with genes associated with disease.

[0207] Briefly, genes can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp in length) from the sequence of a gene of the invention. Computer analysis of the sequence of a gene of the invention can be used to rapidly select primers that do not span more than one exon in the genomic DNA, thus complicating the amplification process. These primers can then be used for PCR screening of somatic cell hybrids containing individual human chromosomes. Only those hybrids containing the human gene corresponding to the gene sequences will yield an amplified fragment. For a review of this technique, see D'Eustachio et al. ((1983) *Science* 220:919-924).

[0208] PCR mapping of somatic cell hybrids is a rapid procedure for assigning a particular sequence to a particular chromosome. Three or more sequences can be assigned per day using a single thermal cycler. Using the nucleic acid sequences of the invention to design oligonucleotide primers, sublocalization can be achieved with panels of fragments from specific chromosomes. Other mapping strategies which can similarly be used to map a gene to its chromosome include in situ hybridization (described in Fan et al. (1990) *Proc. Natl. Acad. Sci. USA* 87:6223-27), pre-screening with labeled flow-sorted chromosomes (CITE), and pre-selection by hybridization to chromosome specific cDNA libraries. Fluorescence in situ hybridization (FISH) of a DNA sequence to a metaphase chromosomal spread can further be used to provide a precise chromosomal location in one step. For a review of this technique, see Verma et al., (Human Chromosomes: A Manual of Basic Techniques (Pergamon Press, New York, 1988)).

[0209] Reagents for chromosome mapping can be used individually to mark a single chromosome or a single site on

that chromosome, or panels of reagents can be used for marking multiple sites and/or multiple chromosomes. Reagents corresponding to noncoding regions of the genes actually are preferred for mapping purposes. Coding sequences are more likely to be conserved within gene families, thus increasing the chance of cross hybridizations during chromosomal mapping.

[0210] Once a sequence has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. (Such data are found, for example, in V. McKusick, Mendelian Inheritance in Man, available on-line through Johns Hopkins University Welch Medical Library). The relationship between genes and disease, mapped to the same chromosomal region, can then be identified through linkage analysis (co-inheritance of physically adjacent genes), described in, e.g., Egeland et al. (1987) *Nature* 325:783-787.

[0211] Moreover, differences in the DNA sequences between individuals affected and unaffected with a disease associated with a gene of the invention can be determined. If a mutation is observed in some or all of the affected individuals but not in any unaffected individuals, then the mutation is likely to be the causative agent of the particular disease. Comparison of affected and unaffected individuals generally involves first looking for structural alterations in the chromosomes such as deletions or translocations that are visible from chromosome spreads or detectable using PCR based on that DNA sequence. Ultimately, complete sequencing of genes from several individuals can be performed to confirm the presence of a mutation and to distinguish mutations from polymorphisms.

[0212] Furthermore, the nucleic acid sequences disclosed herein can be used to perform searches against "mapping databases", e.g., BLAST-type search, such that the chromosome position of the gene is identified by sequence homology or identity with known sequence fragments which have been mapped to chromosomes.

[0213] A polypeptide and fragments and sequences thereof and antibodies specific thereto can be used to map the location of the gene encoding the polypeptide on a chromosome. This mapping can be carried out by specifically detecting the presence of the polypeptide in members of a panel of somatic cell hybrids between cells of a first species of animal from which the protein originates and cells from a second species of animal and then determining which somatic cell hybrid(s) expresses the polypeptide and noting the chromosome(s) from the first species of animal that it contains. For examples of this technique, see Pajunen et al. (1988) *Cytogenet. Cell Genet.* 47:37-41 and Van Keuren et al. (1986) *Hum. Genet.* 74:34-40. Alternatively, the presence of the polypeptide in the somatic cell hybrids can be determined by assaying an activity or property of the polypeptide, for example, enzymatic activity, as described in Bordelon-Riser et al. (1979) *Somatic Cell Genetics* 5:597-613 and Owerbach et al. (1978) *Proc. Natl. Acad. Sci. USA* 75:5640-5644.

[0214] 2. Tissue Typing

[0215] The nucleic acid sequences of the present invention can also be used to identify individuals from minute biological samples. The United States military, for example, is considering the use of restriction fragment length polymor-

phism (RFLP) for identification of its personnel. In this technique, an individual's genomic DNA is digested with one or more restriction enzymes, and probed on a Southern blot to yield unique bands for identification. This method does not suffer from the current limitations of ADog Tags@ which can be lost, switched, or stolen, making positive identification difficult. The sequences of the present invention are useful as additional DNA markers for RFLP (described in U.S. Pat. No. 5,272,057).

[0216] Furthermore, the sequences of the present invention can be used to provide an alternative technique that determines the actual base-by-base DNA sequence of selected portions of an individual's genome. Thus, the nucleic acid sequences described herein can be used to prepare two PCR primers from the 5' and 3' ends of the sequences. These primers can then be used to amplify an individual's DNA and subsequently sequence it.

[0217] Panels of corresponding DNA sequences from individuals, prepared in this manner, can provide unique individual identifications, as each individual will have a unique set of such DNA sequences due to allelic differences. The sequences of the present invention can be used to obtain such identification sequences from individuals and from tissue. The nucleic acid sequences of the invention uniquely represent portions of the human genome. Allelic variation occurs to some degree in the coding regions of these sequences, and to a greater degree in the noncoding regions. It is estimated that allelic variation between individual humans occurs with a frequency at about once per each 500 bases. Each of the sequences described herein can, to some degree, be used as a standard against which DNA from an individual can be compared for identification purposes. Because greater numbers of polymorphisms occur in the noncoding regions, fewer sequences are necessary to differentiate individuals. The noncoding sequences of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 17, or SEQ ID NO: 19 can comfortably provide positive individual identification with a panel of perhaps 10 to 1,000 primers which each yield a noncoding amplified sequence of 100 bases. If predicted coding sequences, such as those in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 17, or SEQ ID NO: 19 are used, a more appropriate number of primers for positive individual identification would be 500 to 2,000.

[0218] If a panel of reagents from the nucleic acid sequences described herein is used to generate a unique identification database for an individual, those same reagents can later be used to identify tissue from that individual. Using the unique identification database, positive identification of the individual, living or dead, can be made from extremely small tissue samples.

[0219] 3. Use of Partial Gene Sequences in Forensic Biology

[0220] DNA-based identification techniques can also be used in forensic biology. Forensic biology is a scientific field employing genetic typing of biological evidence found at a crime scene as a means for positively identifying, for example, a perpetrator of a crime. To make such an identification, PCR technology can be used to amplify DNA sequences taken from very small biological samples such as tissues, e.g., hair or skin, or body fluids, e.g., blood, saliva,

or semen found at a crime scene. The amplified sequence can then be compared to a standard, thereby allowing identification of the origin of the biological sample.

[0221] The sequences of the present invention can be used to provide polynucleotide reagents, e.g., PCR primers, targeted to specific loci in the human genome, which can enhance the reliability of DNA-based forensic identifications by, for example, providing another identification marker@ (i.e. another DNA sequence that is unique to a particular individual). As mentioned above, actual base sequence information can be used for identification as an accurate alternative to patterns formed by restriction enzyme generated fragments. Sequences targeted to noncoding regions are particularly appropriate for this use as greater numbers of polymorphisms occur in the noncoding regions, making it easier to differentiate individuals using this technique. Examples of polynucleotide reagents include the nucleic acid sequences of the invention or portions thereof, e.g., fragments derived from noncoding regions having a length of at least 20 or 30 bases.

[0222] The nucleic acid sequences described herein can further be used to provide polynucleotide reagents, e.g., labeled or labelable probes which can be used in, for example, an in situ hybridization technique, to identify a specific tissue, e.g., brain tissue. This can be very useful in cases where a forensic pathologist is presented with a tissue of unknown origin. Panels of such probes can be used to identify tissue by species and/or by organ type.

C. Predictive Medicine

[0223] The present invention also pertains to the field of predictive medicine in which diagnostic assays, prognostic assays, pharmacogenomics, and monitoring clinical trials are used for prognostic (predictive) purposes to thereby treat an individual prophylactically. Accordingly, one aspect of the present invention relates to diagnostic assays for determining expression of a polypeptide or nucleic acid of the invention and/or activity of a polypeptide of the invention, in the context of a biological sample (e.g., blood, serum, cells, tissue) to thereby determine whether an individual is afflicted with a disease or disorder, or is at risk of developing a disorder, associated with aberrant expression or activity of a polypeptide of the invention, such as a proliferative disorder, e.g., psoriasis or cancer, or an angiogenic disorder. The invention also provides for prognostic (or predictive) assays for determining whether an individual is at risk of developing a disorder associated with aberrant expression or activity of a polypeptide of the invention. For example, mutations in a gene of the invention can be assayed in a biological sample. Such assays can be used for prognostic or predictive purpose to thereby prophylactically treat an individual prior to the onset of a disorder characterized by or associated with aberrant expression or activity of a polypeptide of the invention.

[0224] Another aspect of the invention provides methods for expression of a nucleic acid or polypeptide of the invention or activity of a polypeptide of the invention in an individual to thereby select appropriate therapeutic or prophylactic agents for that individual (referred to herein as "pharmacogenomics"). Pharmacogenomics allows for the selection of agents (e.g., drugs) for therapeutic or prophylactic treatment of an individual based on the genotype of the

individual (e.g., the genotype of the individual examined to determine the ability of the individual to respond to a particular agent).

[0225] Yet another aspect of the invention pertains to monitoring the influence of agents (e.g., drugs or other compounds) on the expression or activity of a polypeptide of the invention in clinical trials. These and other agents are described in further detail in the following sections.

[0226] 1. Diagnostic Assays

[0227] An exemplary method for detecting the presence or absence of a polypeptide or nucleic acid of the invention in a biological sample involves obtaining a biological sample from a test subject and contacting the biological sample with a compound or an agent capable of detecting a polypeptide or nucleic acid (e.g., mRNA, genomic DNA) of the invention such that the presence of a polypeptide or nucleic acid of the invention is detected in the biological sample. A preferred agent for detecting mRNA or genomic DNA encoding a polypeptide of the invention is a labeled nucleic acid probe capable of hybridizing to mRNA or genomic DNA encoding a polypeptide of the invention. The nucleic acid probe can be, for example, a full-length cDNA, such as the nucleic acid of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 17, or SEQ ID NO: 19 or a portion thereof, such as an oligonucleotide of at least 15, 30, 50, 100, 250 or 500 contiguous nucleotides in length and sufficient to specifically hybridize under stringent conditions to a mRNA or genomic DNA encoding a polypeptide of the invention. Other suitable probes for use in the diagnostic assays of the invention are described herein.

[0228] A preferred agent for detecting a polypeptide of the invention is an antibody capable of binding to a polypeptide of the invention, preferably an antibody with a detectable label. Antibodies can be polyclonal, or more preferably, monoclonal. An intact antibody, or a fragment thereof (e.g., Fab or F(ab')₂) can be used. The term Alabeled@, with regard to the probe or antibody, is intended to encompass direct labeling of the probe or antibody by coupling (i.e., physically linking) a detectable substance to the probe or antibody, as well as indirect labeling of the probe or antibody by reactivity with another reagent that is directly labeled. Examples of indirect labeling include detection of a primary antibody using a fluorescently labeled secondary antibody and end-labeling of a DNA probe with biotin such that it can be detected with fluorescently labeled streptavidin. The term Abiological sample@ is intended to include tissues, cells and biological fluids isolated from a subject, as well as tissues, cells and fluids present within a subject. That is, the detection method of the invention can be used to detect mRNA, protein, or genomic DNA in a biological sample in vitro as well as in vivo. For example, in vitro techniques for detection of mRNA include Northern hybridizations and in situ hybridizations. In vitro techniques for detection of a polypeptide of the invention include enzyme linked immunosorbent assays (ELISAs), Western blots, immunoprecipitations and immunofluorescence. In vitro techniques for detection of genomic DNA include Southern hybridizations. Furthermore, in vivo techniques for detection of a polypeptide of the invention include introducing into a subject a labeled antibody directed against the polypeptide. For example, the antibody can be labeled with

a radioactive marker whose presence and location in a subject can be detected by standard imaging techniques.

[0229] In one embodiment, the biological sample contains protein molecules from the test subject. Alternatively, the biological sample can contain mRNA molecules from the test subject or genomic DNA molecules from the test subject. A preferred biological sample is a peripheral blood leukocyte sample isolated by conventional means from a subject.

[0230] In another embodiment, the methods further involve obtaining a control biological sample from a control subject, contacting the control sample with a compound or agent capable of detecting a polypeptide of the invention or mRNA or genomic DNA encoding a polypeptide of the invention, such that the presence of the polypeptide or mRNA or genomic DNA encoding the polypeptide is detected in the biological sample, and comparing the presence of the polypeptide or mRNA or genomic DNA encoding the polypeptide in the control sample with the presence of the polypeptide or mRNA or genomic DNA encoding the polypeptide in the test sample.

[0231] The invention also encompasses kits for detecting the presence of a polypeptide or nucleic acid of the invention in a biological sample (a test sample). Such kits can be used to determine if a subject is suffering from or is at increased risk of developing a disorder associated with aberrant expression of a Tango-73, Tango-74, Tango-76, Tango-78, or Tango-83 gene as discussed, for example, in sections above relating to uses of the sequences of the invention.

[0232] In another example, kits can be used to determine if a subject is suffering from or is at risk for disorders involving Tango-73, Tango-74, Tango-76, Tango-78, or Tango-83.

[0233] In another example, kits can be used to determine if a subject is suffering from or is at risk for which are associated with aberrant Tango-73, Tango-74, Tango-76, Tango-78, or Tango-83 family member activity and/or expression.

[0234] The kit, for example, can comprise a labeled compound or agent capable of detecting the polypeptide or mRNA encoding the polypeptide in a biological sample and means for determining the amount of the polypeptide or mRNA in the sample (e.g., an antibody which binds the polypeptide or an oligonucleotide probe which binds to DNA or mRNA encoding the polypeptide). Kits can also include instructions for observing that the tested subject is suffering from or is at risk of developing a disorder associated with aberrant expression of the polypeptide if the amount of the polypeptide or mRNA encoding the polypeptide is above or below a normal level.

[0235] For antibody-based kits, the kit can comprise, for example: (1) a first antibody (e.g., attached to a solid support) which binds to a polypeptide of the invention; and, optionally, (2) a second, different antibody which binds to either the polypeptide or the first antibody and is conjugated to a detectable agent.

[0236] For oligonucleotide-based kits, the kit can comprise, for example: (1) an oligonucleotide, e.g., a detectably labeled oligonucleotide, which hybridizes to a nucleic acid sequence encoding a polypeptide of the invention or (2) a

pair of primers useful for amplifying a nucleic acid molecule encoding a polypeptide of the invention. The kit can also comprise, e.g., a buffering agent, a preservative, or a protein stabilizing agent. The kit can also comprise components necessary for detecting the detectable agent (e.g., an enzyme or a substrate). The kit can also contain a control sample or a series of control samples which can be assayed and compared to the test sample contained. Each component of the kit is usually enclosed within an individual container and all of the various containers are within a single package along with instructions for observing whether the tested subject is suffering from or is at risk of developing a disorder associated with aberrant expression of the polypeptide.

[0237] 2. Prognostic Assays

[0238] The methods described herein can furthermore be utilized as diagnostic or prognostic assays to identify subjects having or at risk of developing a disease or disorder associated with aberrant expression or activity of a polypeptide of the invention. For example, the assays described herein, such as the preceding diagnostic assays or the following assays, can be utilized to identify a subject having or at risk of developing a disorder associated with aberrant expression or activity of a polypeptide of the invention, e.g., an immunologic disorder, or embryonic disorders. Alternatively, the prognostic assays can be utilized to identify a subject having or at risk for developing such a disease or disorder. Thus, the present invention provides a method in which a test sample is obtained from a subject and a polypeptide or nucleic acid (e.g., mRNA, genomic DNA) of the invention is detected, wherein the presence of the polypeptide or nucleic acid is diagnostic for a subject having or at risk of developing a disease or disorder associated with aberrant expression or activity of the polypeptide. As used herein, a "test sample" refers to a biological sample obtained from a subject of interest. For example, a test sample can be a biological fluid (e.g., serum), cell sample, or tissue.

[0239] The prognostic assays described herein, for example, can be used to identify a subject having or at risk of developing disorders such as disorders discussed, for example, in sections above relating to uses of the sequences of the invention. For example, prognostic assays described herein can be used to identify a subject having or at risk of developing immunological disorders, e.g., autoimmune disorders (e.g., arthritis, graft rejection (e.g., allograft rejection), T cell disorders (e.g., AIDS)), inflammatory disorders (e.g., bacterial infection, psoriasis, septicemia, cerebral malaria, inflammatory bowel disease, arthritis (e.g., rheumatoid arthritis, osteoarthritis)), and allergic inflammatory disorders (e.g., asthma, psoriasis), which are associated with aberrant Tango-73, Tango-74, Tango-76, Tango-78, or Tango-83 activity and/or expression.

[0240] In another example, prognostic assays described herein can be used to identify a subject having or at risk of developing brain-related disorders, inflammations (e.g., bacterial and viral meningitis, encephalitis, and cerebral toxoplasmosis), and tumors (e.g., astrocytoma), and to treat injury or trauma to the brain, which are associated with aberrant Tango-73, Tango-74, Tango-76, Tango-78, or Tango-83 family member activity and/or expression.

[0241] Furthermore, the prognostic assays described herein can be used to determine whether a subject can be administered an agent (e.g., an agonist, antagonist, peptido-

mimetic, protein, peptide, nucleic acid, small molecule, or other drug candidate) to treat a disease or disorder associated with aberrant expression or activity of a polypeptide of the invention. For example, such methods can be used to determine whether a subject can be effectively treated with a specific agent or class of agents (e.g., agents of a type which decrease activity of the polypeptide). Thus, the present invention provides methods for determining whether a subject can be effectively treated with an agent for a disorder associated with aberrant expression or activity of a polypeptide of the invention in which a test sample is obtained and the polypeptide or nucleic acid encoding the polypeptide is detected (e.g., wherein the presence of the polypeptide or nucleic acid is diagnostic for a subject that can be administered the agent to treat a disorder associated with aberrant expression or activity of the polypeptide).

[0242] The methods of the invention can also be used to detect genetic lesions or mutations in a gene of the invention, thereby determining if a subject with the lesioned gene is at risk for a disorder characterized aberrant expression or activity of a polypeptide of the invention. In preferred embodiments, the methods include detecting, in a sample of cells from the subject, the presence or absence of a genetic lesion or mutation characterized by at least one of an alteration affecting the integrity of a gene encoding the polypeptide of the invention, or the mis-expression of the gene encoding the polypeptide of the invention. For example, such genetic lesions or mutations can be detected by ascertaining the existence of at least one of: 1) a deletion of one or more nucleotides from the gene; 2) an addition of one or more nucleotides to the gene; 3) a substitution of one or more nucleotides of the gene; 4) a chromosomal rearrangement of the gene; 5) an alteration in the level of a messenger RNA transcript of the gene; 6) an aberrant modification of the gene, such as of the methylation pattern of the genomic DNA; 7) the presence of a non-wild type splicing pattern of a messenger RNA transcript of the gene; 8) a non-wild type level of a the protein encoded by the gene; 9) an allelic loss of the gene; and 10) an inappropriate post-translational modification of the protein encoded by the gene. As described herein, there are a large number of assay techniques known in the art that can be used for detecting lesions in a gene.

[0243] In certain embodiments, detection of the lesion involves the use of a probe/primer in a polymerase chain reaction (PCR) (see, e.g., U.S. Pat. Nos. 4,683,195 and 4,683,202), such as anchor PCR or RACE PCR, or, alternatively, in a ligation chain reaction (LCR) (see, e.g., Landegran et al. (1988) *Science* 241:1077-1080; and Nakazawa et al. (1994) *Proc. Natl. Acad. Sci. USA* 91:360-364), the latter of which can be particularly useful for detecting point mutations in a gene (see, e.g., Abravaya et al. (1995) *Nucleic Acids Res.* 23:675-682). This method can include the steps of collecting a sample of cells from a patient, isolating nucleic acid (e.g., genomic, mRNA or both) from the cells of the sample, contacting the nucleic acid sample with one or more primers which specifically hybridize to the selected gene under conditions such that hybridization and amplification of the gene (if present) occurs, and detecting the presence or absence of an amplification product, or detecting the size of the amplification product and comparing the length to a control sample. It is anticipated that PCR and/or LCR may be desirable to use as a preliminary

amplification step in conjunction with any of the techniques used for detecting mutations described herein.

[0244] Alternative amplification methods include: self sustained sequence replication (Guatelli et al. (1990) *Proc. Natl. Acad. Sci. USA* 87:1874-1878), transcriptional amplification system (Kwoh, et al. (1989) *Proc. Natl. Acad. Sci. USA* 86:1173-1177), Q-Beta Replicase (Lizardi et al. (1988) *Bio/Technology* 6:1197), or any other nucleic acid amplification method, followed by the detection of the amplified molecules using techniques well known to those of skill in the art. These detection schemes are especially useful for the detection of nucleic acid molecules if such molecules are present in very low numbers.

[0245] In an alternative embodiment, mutations in a selected gene from a sample cell can be identified by alterations in restriction enzyme cleavage patterns. For example, sample and control DNA is isolated, amplified (optionally), digested with one or more restriction endonucleases, and fragment length sizes are determined by gel electrophoresis and compared. Differences in fragment length sizes between sample and control DNA indicates mutations in the sample DNA. Moreover, the use of sequence specific ribozymes (see, e.g., U.S. Pat. No. 5,498,531) can be used to score for the presence of specific mutations by development or loss of a ribozyme cleavage site.

[0246] In other embodiments genetic mutations can be identified by hybridizing a sample and control nucleic acids, e.g., DNA or RNA, to high density arrays containing hundreds or thousands of oligonucleotide probes (Cronin et al. (1996) *Human Mutation* 7:244-255; Kozal et al. (1996) *Nature Medicine* 2:753-759). For example, genetic mutations can be identified in two-dimensional arrays containing light-generated DNA probes as described in Cronin et al., supra. Briefly, a first hybridization array of probes can be used to scan through long stretches of DNA in a sample and control to identify base changes between the sequences by making linear arrays of sequential overlapping probes. This step allows the identification of point mutations. This step is followed by a second hybridization array that allows the characterization of specific mutations by using smaller, specialized probe arrays complementary to all variants or mutations detected. Each mutation array is composed of parallel probe sets, one complementary to the wild-type gene and the other complementary to the mutant gene.

[0247] In yet another embodiment, any of a variety of sequencing reactions known in the art can be used to directly sequence the selected gene and detect mutations by comparing the sequence of the sample nucleic acids with the corresponding wild-type (control) sequence. Examples of sequencing reactions include those based on techniques developed by Maxim and Gilbert ((1977) *Proc. Natl. Acad. Sci. USA* 74:560) or Sanger ((1977) *Proc. Natl. Acad. Sci. USA* 74:5463). It is also contemplated that any of a variety of automated sequencing procedures can be utilized when performing the diagnostic assays ((1995) *Bio/Techniques* 19:448), including sequencing by mass spectrometry (see, e.g., PCT Publication No. WO 94/16101; Cohen et al. (1996) *Adv. Chromatogr.* 36:127-162; and Griffin et al. (1993) *Appl. Biochem. Biotechnol.* 38:147-159).

[0248] Other methods for detecting mutations in a selected gene include methods in which protection from cleavage

agents is used to detect mismatched bases in RNA/RNA or RNA/DNA heteroduplexes (Myers et al. (1985) *Science* 230:1242). In general, the technique of Amismatch cleavage entails providing heteroduplexes formed by hybridizing (labeled) RNA or DNA containing the wild-type sequence with potentially mutant RNA or DNA obtained from a tissue sample. The double-stranded duplexes are treated with an agent which cleaves single-stranded regions of the duplex such as which will exist due to basepair mismatches between the control and sample strands. RNA/DNA duplexes can be treated with RNase to digest mismatched regions, and DNA/DNA hybrids can be treated with S1 nuclease to digest mismatched regions.

[0249] In other embodiments, either DNA/DNA or RNA/DNA duplexes can be treated with hydroxylamine or osmium tetroxide and with piperidine in order to digest mismatched regions. After digestion of the mismatched regions, the resulting material is then separated by size on denaturing polyacrylamide gels to determine the site of mutation. See, e.g., Cotton et al. (1988) *Proc. Natl. Acad. Sci. USA* 85:4397; Saleeba et al. (1992) *Methods Enzymol.* 217:286-295. In a preferred embodiment, the control DNA or RNA can be labeled for detection.

[0250] In still another embodiment, the mismatch cleavage reaction employs one or more proteins that recognize mismatched base pairs in double-stranded DNA (so called ADNA mismatch repair enzymes) in defined systems for detecting and mapping point mutations in cDNAs obtained from samples of cells. For example, the mutY enzyme of *E. coli* cleaves A at G/A mismatches and the thymidine DNA glycosylase from HeLa cells cleaves T at G/T mismatches (Hsu et al. (1994) *Carcinogenesis* 15:1657-1662). According to an exemplary embodiment, a probe based on a selected sequence, e.g., a wild-type sequence, is hybridized to a cDNA or other DNA product from a test cell(s). The duplex is treated with a DNA mismatch repair enzyme, and the cleavage products, if any, can be detected from electrophoresis protocols or the like. See, e.g., U.S. Pat. No. 5,459,039.

[0251] In other embodiments, alterations in electrophoretic mobility will be used to identify mutations in genes. For example, single strand conformation polymorphism (SSCP) may be used to detect differences in electrophoretic mobility between mutant and wild type nucleic acids (Orita et al. (1989) *Proc. Natl. Acad. Sci. USA* 86:2766; see also Cotton (1993) *Mutat. Res.* 285:125-144; Hayashi (1992) *Genet. Anal. Tech. Appl.* 9:73-79). Single-stranded DNA fragments of sample and control nucleic acids will be denatured and allowed to renature. The secondary structure of single-stranded nucleic acids varies according to sequence, and the resulting alteration in electrophoretic mobility enables the detection of even a single base change. The DNA fragments may be labeled or detected with labeled probes. The sensitivity of the assay may be enhanced by using RNA (rather than DNA), in which the secondary structure is more sensitive to a change in sequence. In a preferred embodiment, the subject method utilizes heteroduplex analysis to separate double stranded heteroduplex molecules on the basis of changes in electrophoretic mobility (Keen et al. (1991) *Trends Genet.* 7:5).

[0252] In yet another embodiment, the movement of mutant or wild-type fragments in polyacrylamide gels con-

taining a gradient of denaturant is assayed using denaturing gradient gel electrophoresis (DGGE) (Myers et al. (1985) *Nature* 313:495). When DGGE is used as the method of analysis, DNA will be modified to insure that it does not completely denature, for example by adding a 'GC clamp of approximately 40 bp of high-melting GC-rich DNA by PCR. In a further embodiment, a temperature gradient is used in place of a denaturing gradient to identify differences in the mobility of control and sample DNA (Rosenbaum and Reissner (1987) *Biophys. Chem.* 265:12753).

[0253] Examples of other techniques for detecting point mutations include, but are not limited to, selective oligonucleotide hybridization, selective amplification, or selective primer extension. For example, oligonucleotide primers may be prepared in which the known mutation is placed centrally and then hybridized to target DNA under conditions which permit hybridization only if a perfect match is found (Saiki et al. (1986) *Nature* 324:163); Saiki et al. (1989) *Proc. Natl. Acad. Sci. USA* 86:6230). Such allele specific oligonucleotides are hybridized to PCR amplified target DNA or a number of different mutations when the oligonucleotides are attached to the hybridizing membrane and hybridized with labeled target DNA.

[0254] Alternatively, allele specific amplification technology that depends on selective PCR amplification may be used in conjunction with the instant invention. Oligonucleotides used as primers for specific amplification may carry the mutation of interest in the center of the molecule (so that amplification depends on differential hybridization) (Gibbs et al. (1989) *Nucleic Acids Res.* 17:2437-2448) or at the extreme 3' end of one primer where, under appropriate conditions, mismatch can prevent or reduce polymerase extension (Prossner (1993) *Tibtech* 11:238). In addition, it may be desirable to introduce a novel restriction site in the region of the mutation to create cleavage-based detection (Gasparini et al. (1992) *Mol. Cell Probes* 6:1). It is anticipated that in certain embodiments amplification may also be performed using Taq ligase for amplification (Barany (1991) *Proc. Natl. Acad. Sci. USA* 88:189). In such cases, ligation will occur only if there is a perfect match at the 3' end of the 5' sequence making it possible to detect the presence of a known mutation at a specific site by looking for the presence or absence of amplification.

[0255] The methods described herein may be performed, for example, by utilizing pre-packaged diagnostic kits comprising at least one probe nucleic acid or antibody reagent described herein, which may be conveniently used, e.g., in clinical settings to diagnose patients exhibiting symptoms or family history of a disease or illness involving a gene encoding a polypeptide of the invention. Furthermore, any cell type or tissue, e.g., preferably peripheral blood leukocytes, in which the polypeptide of the invention is expressed may be utilized in the prognostic assays described herein.

[0256] 3. Pharmacogenomics

[0257] Agents or modulators that have a stimulatory or inhibitory effect on activity or expression of a polypeptide of the invention as identified by a screening assay described herein can be administered to individuals to treat (prophylactically or therapeutically) disorders associated with aberrant activity of the polypeptide. In conjunction with such treatment, the pharmacogenomics (i.e., the study of the relationship between an individual's genotype and that indi-

vidual's response to a foreign compound or drug) of the individual may be considered. Differences in metabolism of therapeutics can lead to severe toxicity or therapeutic failure by altering the relation between dose and blood concentration of the pharmacologically active drug. Thus, the pharmacogenomics of the individual permits the selection of effective agents (e.g., drugs) for prophylactic or therapeutic treatments based on a consideration of the individual's genotype. Such pharmacogenomics can further be used to determine appropriate dosages and therapeutic regimens. Accordingly, the activity of a polypeptide of the invention, expression of a nucleic acid of the invention, or mutation content of a gene of the invention in an individual can be determined to thereby select appropriate agent(s) for therapeutic or prophylactic treatment of the individual.

[0258] Pharmacogenomics deals with clinically significant hereditary variations in the response to drugs due to altered drug disposition and abnormal action in affected persons. See, e.g., Linder (1997) *Clin. Chem.* 43(2)-254-266. In general, two types of pharmacogenetic conditions can be differentiated. Genetic conditions transmitted as a single factor altering the way drugs act on the body are referred to as Altered drug action. Genetic conditions transmitted as single factors altering the way the body acts on drugs are referred to as Altered drug metabolism. These pharmacogenetic conditions can occur either as rare defects or as polymorphisms. For example, glucose-6-phosphate dehydrogenase deficiency (G6PD) is a common inherited enzymopathy in which the main clinical complication is haemolysis after ingestion of oxidant drugs (anti-malarials, sulfonamides, analgesics, nitrofurans) and consumption of fava beans.

[0259] As an illustrative embodiment, the activity of drug metabolizing enzymes is a major determinant of both the intensity and duration of drug action. The discovery of genetic polymorphisms of drug metabolizing enzymes (e.g., N-acetyltransferase 2 (NAT 2) and cytochrome P450 enzymes CYP2D6 and CYP2C19) has provided an explanation as to why some patients do not obtain the expected drug effects or show exaggerated drug response and serious toxicity after taking the standard and safe dose of a drug. These polymorphisms are expressed in two phenotypes in the population, the extensive metabolizer (EM) and poor metabolizer (PM). The prevalence of PM is different among different populations. For example, the gene coding for CYP2D6 is highly polymorphic and several mutations have been identified in PM, which all lead to the absence of functional CYP2D6. Poor metabolizers of CYP2D6 and CYP2C19 quite frequently experience exaggerated drug response and side effects when they receive standard doses. If a metabolite is the active therapeutic moiety, a PM will show no therapeutic response, as demonstrated for the analgesic effect of codeine mediated by its CYP2D6-formed metabolite morphine. The other extreme are the so called ultra-rapid metabolizers who do not respond to standard doses. Recently, the molecular basis of ultra-rapid metabolism has been identified to be due to CYP2D6 gene amplification.

[0260] Thus, the activity of a polypeptide of the invention, expression of a nucleic acid encoding the polypeptide, or mutation content of a gene encoding the polypeptide in an individual can be determined to thereby select appropriate agent(s) for therapeutic or prophylactic treatment of the

individual. In addition, pharmacogenetic studies can be used to apply genotyping of polymorphic alleles encoding drug-metabolizing enzymes to the identification of an individual's drug responsiveness phenotype. This knowledge, when applied to dosing or drug selection, can avoid adverse reactions or therapeutic failure and thus enhance therapeutic or prophylactic efficiency when treating a subject with a modulator of activity or expression of the polypeptide, such as a modulator identified by one of the exemplary screening assays described herein.

[0261] 4. Monitoring of Effects During Clinical Trials

[0262] Monitoring the influence of agents (e.g., drugs, compounds) on the expression or activity of a polypeptide of the invention (e.g., the ability to modulate aberrant cell proliferation, chemotaxis, and/or differentiation) can be applied not only in basic drug screening, but also in clinical trials. For example, the effectiveness of an agent, as determined by a screening assay as described herein, to increase gene expression, protein levels or protein activity, can be monitored in clinical trials of subjects exhibiting decreased gene expression, protein levels, or protein activity. Alternatively, the effectiveness of an agent, as determined by a screening assay, to decrease gene expression, protein levels or protein activity, can be monitored in clinical trials of subjects exhibiting increased gene expression, protein levels, or protein activity. In such clinical trials, expression or activity of a polypeptide of the invention and preferably, that of other polypeptide that have been implicated in for example, a cellular proliferation disorder, can be used as a marker of the immune responsiveness of a particular cell.

[0263] For example, and not by way of limitation, genes, including those of the invention, that are modulated in cells by treatment with an agent (e.g., compound, drug or small molecule) that modulates activity or expression of a polypeptide of the invention (e.g., as identified in a screening assay described herein) can be identified. Thus, to study the effect of agents on cellular proliferation disorders, for example, in a clinical trial, cells can be isolated and RNA prepared and analyzed for the levels of expression of a gene of the invention and other genes implicated in the disorder. The levels of gene expression (i.e., a gene expression pattern) can be quantified by Northern blot analysis or RT-PCR, as described herein, or alternatively by measuring the amount of protein produced, by one of the methods as described herein, or by measuring the levels of activity of a gene of the invention or other genes. In this way, the gene expression pattern can serve as a marker, indicative of the physiological response of the cells to the agent. Accordingly, this response state may be determined before, and at various points during, treatment of the individual with the agent.

[0264] In a preferred embodiment, the present invention provides a method for monitoring the effectiveness of treatment of a subject with an agent (e.g., an agonist, antagonist, peptidomimetic, protein, peptide, nucleic acid, small molecule, or other drug candidate identified by the screening assays described herein) comprising the steps of (i) obtaining a pre-administration sample from a subject prior to administration of the agent; (ii) detecting the level of the polypeptide or nucleic acid of the invention in the pre-administration sample; (iii) obtaining one or more post-administration samples from the subject; (iv) detecting the level of the polypeptide or nucleic acid of the invention in the

post-administration samples; (v) comparing the level of the polypeptide or nucleic acid of the invention in the pre-administration sample with the level of the polypeptide or nucleic acid of the invention in the post-administration sample or samples; and (vi) altering the administration of the agent to the subject accordingly. For example, increased administration of the agent may be desirable to increase the expression or activity of the polypeptide to higher levels than detected, i.e., to increase the effectiveness of the agent. Alternatively, decreased administration of the agent may be desirable to decrease expression or activity of the polypeptide to lower levels than detected, i.e., to decrease the effectiveness of the agent.

C. Methods of Treatment

[0265] Tango-73, Tango-74, Tango-76, Tango-78, and Tango-83 polypeptides, nucleic acids, and modulators thereof can be used to modulate the function, morphology, proliferation and/or differentiation of cells in the tissues in which it is expressed. Such molecules can be used to treat disorders associated with abnormal or aberrant metabolism or function of cells in the tissues in which it is expressed. Tissues in which nucleic acids and polypeptides of the invention are expressed include, for example, pancreas, kidney, testis, heart, brain, liver, placenta, lung, skeletal muscle, or small intestine.

[0266] As revealed by Northern blot analysis, Tango-74, Tango-76, and Tango-83 are expressed in the brain. Consequently, Tango-74, Tango-76, and Tango-83 polypeptides, nucleic acids, and modulators thereof can be used to treat disorders of the brain, such as cerebral edema, hydrocephalus, brain herniations, iatrogenic disease (due to, e.g., infection, toxins, or drugs), inflammations (e.g., bacterial and viral meningitis, encephalitis, and cerebral toxoplasmosis), cerebrovascular diseases (e.g., hypoxia, ischemia, and infarction, intracranial hemorrhage and vascular malformations, and hypertensive encephalopathy), and tumors (e.g., neuroglial tumors, neuronal tumors, tumors of pineal cells, meningeal tumors, primary and secondary lymphomas, intracranial tumors, and medulloblastoma), and to treat injury or trauma to the brain.

[0267] As revealed by Northern blot analysis, Tango-74 and Tango-76 are expressed in the cardiovascular system. Consequently, Tango-74 and Tango-76 polypeptides, nucleic acids, and modulators thereof can be used to treat cardiovascular disorders, such as ischemic heart disease (e.g., angina pectoris, myocardial infarction, and chronic ischemic heart disease), hypertensive heart disease, pulmonary heart disease, valvular heart disease (e.g., rheumatic fever and rheumatic heart disease, endocarditis, mitral valve prolapse, and aortic valve stenosis), congenital heart disease (e.g., valvular and vascular obstructive lesions, atrial or ventricular septal defect, and patent ductus arteriosus), or myocardial disease (e.g., myocarditis, congestive cardiomyopathy, and hypertrophic cardiomyopathy).

[0268] As revealed by Northern blot analysis, Tango-74 and Tango-76 are expressed in the heart. Consequently, Tango-74 and Tango-76 nucleic acids, proteins, and modulators thereof can be used to treat heart disorders, e.g., ischemic heart disease, atherosclerosis, hypertension, angina pectoris, Hypertrophic Cardiomyopathy, and congenital heart disease.

[0269] As revealed by Northern blot analysis, Tango-74 shows leukocyte expression. Consequently, Tango-74 polypeptides, nucleic acids, and modulators thereof can be used to treat leukocytic disorders, such as leukopenias (e.g., neutropenia, monocytopenia, lymphopenia, and granulocytopenia), leukocytosis (e.g., granulocytosis, lymphocytosis, eosinophilia, monocytosis, acute and chronic lymphadenitis), malignant lymphomas (e.g., Non-Hodgkin's lymphomas, Hodgkin's lymphomas, leukemias, agnogenic myeloid metaplasia, multiple myeloma, plasmacytoma, Waldenstrom's macroglobulinemia, heavy-chain disease, monoclonal gammopathy, histiocytoses, eosinophilic granuloma, and angioimmunoblastic lymphadenopathy).

[0270] As human Tango-78 was found in a bone marrow cDNA library, Tango-78 nucleic acids, proteins, and modulators thereof can be used to modulate the proliferation, differentiation, and/or function of cells that appear in the bone marrow, e.g., stem cells (e.g., hematopoietic stem cells), and blood cells, e.g., erythrocytes, platelets, and leukocytes. Thus Tango-78 nucleic acids, proteins, and modulators thereof can be used to treat bone marrow, blood, and hematopoietic associated diseases and disorders, e.g., acute myeloid leukemia, hemophilia, leukemia, anemia (e.g., sickle cell anemia), and thalassemia.

[0271] As revealed by Northern blot analysis, Tango-73, Tango-74, and Tango-76 are expressed in the spleen. Consequently, Tango-73, Tango-74, and Tango-76 nucleic acids, proteins, and modulators thereof can be used to modulate the proliferation, differentiation, and/or function of cells that form the spleen, e.g., cells of the splenic connective tissue, e.g., splenic smooth muscle cells and/or endothelial cells of the splenic blood vessels. Tango-73, Tango-74, and Tango-76 nucleic acids, proteins, and modulators thereof can also be used to modulate the proliferation, differentiation, and/or function of cells that are processed, e.g., regenerated or phagocytized within the spleen, e.g., erythrocytes and/or B and T lymphocytes and macrophages. Thus Tango-73, Tango-74, and Tango-76 nucleic acids, proteins, and modulators thereof can be used to treat spleen, e.g., the fetal spleen, associated diseases and disorders. Examples of splenic diseases and disorders include e.g., splenic lymphoma and/or splenomegaly, and/or phagocytotic disorders, e.g., those inhibiting macrophage engulfment of bacteria and viruses in the bloodstream.

[0272] As revealed by Northern blot analysis, Tango-73, Tango-74, and Tango-76 are expressed in the lung. Consequently, Tango-73, Tango-74, and Tango-76 polypeptides, nucleic acids, and modulators thereof can be used to treat pulmonary (lung) disorders, such as atelectasis, cystic fibrosis, rheumatoid lung disease, pulmonary congestion or edema, chronic obstructive airway disease (e.g., emphysema, chronic bronchitis, bronchial asthma, and bronchiectasis), diffuse interstitial diseases (e.g., sarcoidosis, pneumoconiosis, hypersensitivity pneumonitis, bronchiolitis, Goodpasture's syndrome, idiopathic pulmonary fibrosis, idiopathic pulmonary hemosiderosis, pulmonary alveolar proteinosis, desquamative interstitial pneumonitis, chronic interstitial pneumonia, fibrosing alveolitis, hamman-rich syndrome, pulmonary eosinophilia, diffuse interstitial fibrosis, Wegener's granulomatosis, lymphomatoid granulomatosis, and lipid pneumonia), or tumors (e.g., bronchogenic carcinoma, bronchioloalveolar carcinoma, bronchial carcinoid, hamartoma, and mesenchymal tumors).

[0273] As revealed by Northern blot analysis, Tango-73 and Tango-74 are expressed in the pancreas. Consequently, Tango-73 and Tango-74 polypeptides, nucleic acids, and modulators thereof can be used to treat pancreatic disorders, such as pancreatitis (e.g., acute hemorrhagic pancreatitis and chronic pancreatitis), pancreatic cysts (e.g., congenital cysts, pseudocysts, and benign or malignant neoplastic cysts), pancreatic tumors (e.g., pancreatic carcinoma and adenoma), diabetes mellitus (e.g., insulin- and non-insulin-dependent types, impaired glucose tolerance, and gestational diabetes), or islet cell tumors (e.g., insulinomas, adenomas, Zollinger-Ellison syndrome, glucagonomas, and somatostatinoma).

[0274] As revealed by Northern blot analysis, Tango-73, Tango-74, and Tango-76 are expressed in the liver. Consequently, Tango-73, Tango-74, and Tango-76 polypeptides, nucleic acids, and modulators thereof can be used to treat hepatic (liver) disorders, such as jaundice, hepatic failure, hereditary hyperbilirubinemias (e.g., Gilbert's syndrome, Crigler-Najjar syndromes and Dubin-Johnson and Rotor's syndromes), hepatic circulatory disorders (e.g., hepatic vein thrombosis and portal vein obstruction and thrombosis), hepatitis (e.g., chronic active hepatitis, acute viral hepatitis, and toxic and drug-induced hepatitis), cirrhosis (e.g., alcoholic cirrhosis, biliary cirrhosis, and hemochromatosis), or malignant tumors (e.g., primary carcinoma, hepatoma, hepatoblastoma, liver cysts, and angiosarcoma).

[0275] As revealed by Northern blot analysis, Tango-73, Tango-74, and Tango-76 are expressed in the kidney. Consequently, Tango-73, Tango-74, and Tango-76 polypeptides, nucleic acids, and modulators thereof can be used to treat renal (kidney) disorders, such as glomerular diseases (e.g., acute and chronic glomerulonephritis, rapidly progressive glomerulonephritis, nephrotic syndrome, focal proliferative glomerulonephritis, glomerular lesions associated with systemic disease, such as systemic lupus erythematosus, Goodpasture's syndrome, multiple myeloma, diabetes, polycystic kidney disease, neoplasia, sickle cell disease, and chronic inflammatory diseases), tubular diseases (e.g., acute tubular necrosis and acute renal failure, polycystic renal disease, medullary sponge kidney, medullary cystic disease, nephrogenic diabetes, and renal tubular acidosis), tubulointerstitial diseases (e.g., pyelonephritis, drug and toxin induced tubulointerstitial nephritis, hypercalcemic nephropathy, and hypokalemic nephropathy) acute and rapidly progressive renal failure, chronic renal failure, nephrolithiasis, gout, vascular diseases (e.g., hypertension and nephrosclerosis, microangiopathic hemolytic anemia, atheroembolic renal disease, diffuse cortical necrosis, and renal infarcts), or tumors (e.g., renal cell carcinoma and nephroblastoma).

[0276] As revealed by Northern blot analysis, Tango-73 and Tango-74 are expressed in the reproductive system. Consequently, Tango-73 or Tango-74 can be used to treat reproductive disorders, including ovulation disorder, blockage of the fallopian tubes (e.g., due to pelvic inflammatory disease or endometriosis), disorders due to infections (e.g., toxic shock syndrome, chlamydia infection, Herpes infection, human papillomavirus infection), and ovarian disorders (e.g., ovarian cyst, ovarian fibroma, ovarian endometriosis, ovarian teratoma).

[0277] As revealed by Northern blot analysis, Tango-74 is expressed in the ovaries. Consequently, Tango-74 polypep-

tides, nucleic acids, and modulators thereof can be used to treat ovarian disorders, such as ovarian endometriosis, non-neoplastic cysts (e.g., follicular and luteal cysts and polycystic ovaries) and tumors (e.g., tumors of surface epithelium, germ cell tumors, ovarian fibroma, sex cord-stromal tumors, and ovarian cancers (e.g., metastatic carcinomas, and ovarian teratoma).

[0278] As revealed by Northern blot analysis, Tango-73 is expressed in the placenta. Consequently, Tango-73 polypeptides, nucleic acids, and modulators thereof can be used to treat placental disorders, such as toxemia of pregnancy (e.g., preeclampsia and eclampsia), placentitis, or spontaneous abortion.

[0279] As revealed by Northern blot analysis, Tango-73 and Tango-74 are expressed in the testes. Consequently, Tango-73 and Tango-74 polypeptides, nucleic acids, and modulators thereof can be used to treat testicular disorders, such as unilateral testicular enlargement (e.g., nontuberculous, granulomatous orchitis); inflammatory diseases resulting in testicular dysfunction (e.g., gonorrhea and mumps); cryptorchidism; sperm cell disorders (e.g., immotile cilia syndrome and germinal cell aplasia); acquired testicular defects (e.g., viral orchitis); and tumors (e.g., germ cell tumors, interstitial cell tumors, and roblastoma, testicular lymphoma and adenomatoid tumors).

[0280] As revealed by Northern blot analysis, Tango-73 and Tango-74 are expressed in the prostate. Consequently, Tango-73 and Tango-74 polypeptides, nucleic acids, and modulators thereof can be used to treat prostate disorders, such as inflammatory diseases (e.g., acute and chronic prostatitis and granulomatous prostatitis), hyperplasia (e.g., benign prostatic hypertrophy or hyperplasia), or tumors (e.g., carcinomas).

[0281] As revealed by Northern blot analysis, Tango-73 and Tango-74 are expressed in the intestines. Consequently, Tango-73 and Tango-74 polypeptides, nucleic acids, and modulators thereof can be used to treat intestinal disorders, such as ischemic bowel disease, infective enterocolitis, Crohn's disease, benign tumors, malignant tumors (e.g., argentaffinomas, lymphomas, adenocarcinomas, and sarcomas), malabsorption syndromes (e.g., celiac disease, tropical sprue, Whipple's disease, and abetalipoproteinemia), obstructive lesions, hernias, intestinal adhesions, intussusception, or volvulus.

[0282] As revealed by Northern blot analysis, Tango-73 and Tango-74 are expressed in the colon. Consequently, Tango-73 and Tango-74 polypeptides, nucleic acids, and modulators thereof can be used to treat colonic disorders, such as congenital anomalies (e.g., megacolon and imperforate anus), idiopathic disorders (e.g., diverticular disease and melanosis coli), vascular lesions (e.g., ischemic colitis, hemorrhoids, angiodysplasia), inflammatory diseases (e.g., colitis (e.g., idiopathic ulcerative colitis, pseudomembranous colitis), and lymphopathia venereum), and tumors (e.g., hyperplastic polyps, adenomatous polyps, bronchogenic cancer, colonic carcinoma, squamous cell carcinoma, adenocarcinomas, sarcomas, lymphomas, argentaffinomas, carcinomas, and melanocarcinomas).

[0283] As revealed by Northern blot analysis, Tango-74 and Tango-76 are expressed in skeletal muscle tissue. Consequently, Tango-74 and Tango-76 polypeptides, nucleic

acids, and modulators thereof can be used to treat disorders of skeletal muscle, such as muscular dystrophy (e.g., Duchenne Muscular Dystrophy, Becker Muscular Dystrophy, Emery-Dreifuss Muscular Dystrophy, Limb-Girdle Muscular Dystrophy, Facioscapulohumeral Muscular Dystrophy, Myotonic Dystrophy, Oculopharyngeal Muscular Dystrophy, Distal Muscular Dystrophy, and Congenital Muscular Dystrophy), motor neuron diseases (e.g., Amyotrophic Lateral Sclerosis, Infantile Progressive Spinal Muscular Atrophy, Intermediate Spinal Muscular Atrophy, Spinal Bulbar Muscular Atrophy, and Adult Spinal Muscular Atrophy), myopathies (e.g., inflammatory myopathies (e.g., Dermatomyositis and Polymyositis), Myotonia Congenita, Paramyotonia Congenita, Central Core Disease, Nemaline Myopathy, Myotubular Myopathy, and Periodic Paralysis), and metabolic diseases of muscle (e.g., Phosphorylase Deficiency, Acid Maltase Deficiency, Phosphofructokinase Deficiency, Debrancher Enzyme Deficiency, Mitochondrial Myopathy, Carnitine Deficiency, Carnitine Palmitoyl Transferase Deficiency, Phosphoglycerate Kinase Deficiency, Phosphoglycerate Mutase Deficiency, Lactate Dehydrogenase Deficiency, and Myoadenylate Deaminase Deficiency).

[0284] The nucleic acids or polypeptides of the invention can be used to treat proliferative disorders, e.g., neoplasms or tumors (e.g., a carcinoma, a sarcoma, adenoma, or myeloid leukemia).

[0285] Disorders associated with abnormal Tango-73, Tango-74, Tango-76, Tango-78, or Tango-83 activity or expression may include proliferative disorders (e.g., carcinoma, lymphoma, e.g., follicular lymphoma).

[0286] Disorders associated with abnormal Tango-73 activity also include apoptotic disorders (e.g., rheumatoid arthritis, systemic lupus erythematosus, insulin-dependent diabetes mellitus), cytotoxic disorders, septic shock, cachexia, and proliferative disorders (e.g., B cell cancers stimulated by TNF).

[0287] 1. Prophylactic Methods

[0288] In one aspect, the invention provides a method for preventing in a subject, a disease or condition associated with an aberrant expression or activity of a polypeptide of the invention, by administering to the subject an agent that modulates expression or at least one activity of the polypeptide. Subjects at risk for a disease that is caused or contributed to by aberrant expression or activity of a polypeptide of the invention can be identified by, for example, any or a combination of diagnostic or prognostic assays as described herein. Administration of a prophylactic agent can occur prior to the manifestation of symptoms characteristic of the aberrancy, such that a disease or disorder is prevented or, alternatively, delayed in its progression. Depending on the type of aberrancy, for example, an agonist or antagonist agent can be used for treating the subject. The prophylactic agents described herein, for example, can be used to treat a subject at risk of developing disorders such as disorders discussed for example, in Sections above relative to the uses of the sequences of the invention. For example, an antagonist of a Tango-73, Tango-74, Tango-76, Tango-78, or Tango-83 protein may be used to modulate or treat an immunological disorder. The appropriate agent can be determined based on screening assays described herein.

[0289] 2. Therapeutic Methods

[0290] Another aspect of the invention pertains to methods of modulating expression or activity of a polypeptide of the invention for therapeutic purposes. The modulatory method of the invention involves contacting a cell with an agent that modulates one or more of the activities of the polypeptide. An agent that modulates activity can be an agent as described herein, such as a nucleic acid or a protein, a naturally-occurring cognate ligand of the polypeptide, a peptide, a peptidomimetic, or other small molecule. In one embodiment, the agent stimulates one or more of the biological activities of the polypeptide. Examples of such stimulatory agents include the active polypeptide of the invention and a nucleic acid molecule encoding the polypeptide of the invention that has been introduced into the cell. In another embodiment, the agent inhibits one or more of the biological activities of the polypeptide of the invention. Examples of such inhibitory agents include antisense nucleic acid molecules and antibodies. These modulatory methods can be performed *in vitro* (e.g., by culturing the cell with the agent) or, alternatively, *in vivo* (e.g., by administering the agent to a subject). As such, the present invention provides methods of treating an individual afflicted with a disease or disorder characterized by aberrant expression or activity of a polypeptide of the invention. In one embodiment, the method involves administering an agent (e.g., an agent identified by a screening assay described herein), or combination of agents that modulates (e.g., upregulates or down-regulates) expression or activity. In another embodiment, the method involves administering a polypeptide of the invention or a nucleic acid molecule of the invention as therapy to compensate for reduced or aberrant expression or activity of the polypeptide.

[0291] Stimulation of activity is desirable in situations in which activity or expression is abnormally low or down-regulated and/or in which increased activity is likely to have a beneficial effect. Conversely, inhibition of activity is desirable in situations in which activity or expression is abnormally high or upregulated and/or in which decreased activity is likely to have a beneficial effect.

[0292] This invention is further illustrated by the following examples which should not be construed as limiting. The contents of all references, patents and published patent applications cited throughout this application are hereby incorporated by reference.

Effective Dose

[0293] Toxicity and therapeutic efficacy of the polypeptides of the invention and the compounds that modulate their expression or activity can be determined by standard pharmaceutical procedures, using either cells in culture or experimental animals to determine the LD₅₀ (the dose lethal to 50% of the population) and the ED₅₀ (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio LD₅₀/ED₅₀. Polypeptides or other compounds that exhibit large therapeutic indices are preferred. While compounds that exhibit toxic side effects may be used, care should be taken to design a delivery system that targets such compounds to the site of affected tissue in order to minimize potential damage to uninfected cells and, thereby, reduce side effects.

[0294] The data obtained from the cell culture assays and animal studies can be used in formulating a range of dosage for use in humans. The dosage of such compounds lies

preferably within a range of circulating concentrations that include the ED₅₀ with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from cell culture assays. A dose may be formulated in animal models to achieve a circulating plasma concentration range that includes the IC₅₀ (that is, the concentration of the test compound which achieves a half-maximal inhibition of symptoms) as determined in cell culture. Such information can be used to more accurately determine useful doses in humans. Levels in plasma may be measured, for example, by high performance liquid chromatography.

Formulations and Use

[0295] Pharmaceutical compositions for use in accordance with the present invention may be formulated in conventional manner using one or more physiologically acceptable carriers or excipients.

[0296] Thus, the compounds and their physiologically acceptable salts and solvates may be formulated for administration by inhalation or insufflation (either through the mouth or the nose) or oral, buccal, parenteral or rectal administration.

[0297] For oral administration, the pharmaceutical compositions may take the form of, for example, tablets or capsules prepared by conventional means with pharmaceutically acceptable excipients such as binding agents (for example, pregelatinised maize starch, polyvinylpyrrolidone or hydroxypropyl methylcellulose); fillers (for example, lactose, microcrystalline cellulose or calcium hydrogen phosphate); lubricants (for example, magnesium stearate, talc or silica); disintegrants (for example, potato starch or sodium starch glycolate); or wetting agents (for example, sodium lauryl sulphate). The tablets may be coated by methods well known in the art. Liquid preparations for oral administration may take the form of, for example, solutions, syrups or suspensions, or they may be presented as a dry product for constitution with water or other suitable vehicle before use. Such liquid preparations may be prepared by conventional means with pharmaceutically acceptable additives such as suspending agents (for example, sorbitol syrup, cellulose derivatives or hydrogenated edible fats); emulsifying agents (for example, lecithin or acacia); non-aqueous vehicles (for example, almond oil, oily esters, ethyl alcohol or fractionated vegetable oils); and preservatives (for example, methyl or propyl-p-hydroxybenzoates or sorbic acid). The preparations may also contain buffer salts, flavoring, coloring and sweetening agents as appropriate. Preparations for oral administration may be suitably formulated to give controlled release of the active compound.

[0298] For buccal administration the compositions may take the form of tablets or lozenges formulated in conventional manner.

[0299] For administration by inhalation, the compounds for use according to the present invention are conveniently delivered in the form of an aerosol spray presentation from pressurized packs or a nebulizer, with the use of a suitable propellant, for example, dichlorodifluoromethane, trichlorofluoromethane, dichlorotetrafluoroethane, carbon dioxide

or other suitable gas. In the case of a pressurized aerosol the dosage unit may be determined by providing a valve to deliver a metered amount. Capsules and cartridges of, for example, gelatin for use in an inhaler or insufflator may be formulated containing a powder mix of the compound and a suitable powder base such as lactose or starch.

[0300] The compounds may be formulated for parenteral administration by injection, for example, by bolus injection or continuous infusion. Formulations for injection may be presented in unit dosage form, for example, in ampoules or in multi-dose containers, with an added preservative. The compositions may take such forms as suspensions, solutions or emulsions in oily or aqueous vehicles, and may contain formulatory agents such as suspending, stabilizing and/or dispersing agents. Alternatively, the active ingredient may be in powder form for constitution with a suitable vehicle, for example, sterile pyrogen-free water, before use.

[0301] The compounds may also be formulated in rectal compositions such as suppositories or retention enemas, for example, containing conventional suppository bases such as cocoa butter or other glycerides.

[0302] In addition to the formulations described previously, the compounds may also be formulated as a depot preparation. Such long acting formulations may be administered by implantation (for example subcutaneously or intramuscularly) or by intramuscular injection. Thus, for example, the compounds may be formulated with suitable polymeric or hydrophobic materials (for example as an emulsion in an acceptable oil) or ion exchange resins, or as sparingly soluble derivatives, for example, as a sparingly soluble salt.

[0303] The compositions may, if desired, be presented in a pack or dispenser device that may contain one or more unit dosage forms containing the active ingredient. The pack may for example comprise metal or plastic foil, such as a blister pack. The pack or dispenser device may be accompanied by instructions for administration.

[0304] The therapeutic compositions of the invention can also contain a carrier or excipient, many of which are known to skilled artisans. Excipients that can be used include buffers (for example, citrate buffer, phosphate buffer, acetate buffer, and bicarbonate buffer), amino acids, urea, alcohols, ascorbic acid, phospholipids, proteins (for example, serum

albumin), EDTA, sodium chloride, liposomes, mannitol, sorbitol, and glycerol. The nucleic acids, polypeptides, antibodies, or modulatory compounds of the invention can be administered by any standard route of administration. For example, administration can be parenteral, intravenous, subcutaneous, intramuscular, intracranial, intraorbital, ophthalmic, intraventricular, intracapsular, intraspinal, intracisternal, intraperitoneal, transmucosal, or oral. The modulatory compound can be formulated in various ways, according to the corresponding route of administration. For example, liquid solutions can be made for ingestion or injection; gels or powders can be made for ingestion, inhalation, or topical application. Methods for making such formulations are well known and can be found in, for example, "Remington's Pharmaceutical Sciences." It is expected that the preferred route of administration will be intravenous.

[0305] It is recognized that the pharmaceutical compositions and methods described herein can be used independently or in combination with one another. That is, subjects can be administered one or more of the pharmaceutical compositions, e.g., pharmaceutical compositions comprising a nucleic acid molecule or protein of the invention or a modulator thereof, subjected to one or more of the therapeutic methods described herein, or both, in temporally overlapping or non-overlapping regimens. When therapies overlap temporally, the therapies may generally occur in any order and can be simultaneous (e.g., administered simultaneously together in a composite composition or simultaneously but as separate compositions) or interspersed. By way of example, a subject afflicted with a disorder described herein can be simultaneously or sequentially administered both a cytotoxic agent which selectively kills aberrant cells and an antibody (e.g., an antibody of the invention) which can, in one embodiment, be conjugated or linked with a therapeutic agent, a cytotoxic agent, an imaging agent, or the like.

Equivalents

[0306] Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following claims.

SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 20

<210> SEQ ID NO 1

<211> LENGTH: 1095

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

<220> FEATURE:

<221> NAME/KEY: CDS

<222> LOCATION: (588)...(1094)

<400> SEQUENCE: 1

cctcgagggg agggggagat gcaactcatc acatttttac tgactgtcct ctggctgtgc 60

aagtattcct ggaaggggga ctggaagaac agtaattcgg agtctgggct tggcagttgg 120

-continued

```

gcaaatccag gtttactcct ggctctgcc ccttccaaga atgacacctt ggtcagatct 180
ttaaaccaca ctgagcctca gttttcctca tctctaaaag ggactcgaaa atcttaccaa 240
ctcatagagt tgggggtgaga attogaaggt aattctatat aaggttaaggc ctccagcaag 300
agctatggtg gttgtgacac tgactgaggc tgggggaggc cctcactcac cctccttcct 360
tcttggtttt ctctaccaca gatgtggcag tggatgggca gaactggacg tttgcttttg 420
acttctcctt cctgagccaa caagaggatc tggcatgggc tgagctccgg ctgcagctgt 480
ccagccctgt ggacctcccc actgagggt cacttgccat tgagattttc caccagccaa 540
agcccagacac agagcaggct tcagacagct gcttagagcg gtttcag atg gac cta 596
                                Met Asp Leu
                                1

ttc act gtc act ttg tcc cag gtc acc ttt tcc ttg ggc agc atg gtt 644
Phe Thr Val Thr Leu Ser Gln Val Thr Phe Ser Leu Gly Ser Met Val
    5              10              15

ttg gag gtg acc agg cct ctc tcc aag tgg ctg aag cgc cct ggg gcc 692
Leu Glu Val Thr Arg Pro Leu Ser Lys Trp Leu Lys Arg Pro Gly Ala
    20              25              30              35

ctg gag aag cag atg tcc agg gta gct gga gag tgc tgg ccg cgg ccc 740
Leu Glu Lys Gln Met Ser Arg Val Ala Gly Glu Cys Trp Pro Arg Pro
                40              45              50

ccc aca ccg cct gcc acc aat gtg ctc ctt atg ctc tac tcc aac ctc 788
Pro Thr Pro Pro Ala Thr Asn Val Leu Leu Met Leu Tyr Ser Asn Leu
                55              60              65

tcg cag gag cag agg cag ctg ggt ggg tcc acc ttg ctg tgg gaa gcc 836
Ser Gln Glu Gln Arg Gln Leu Gly Gly Ser Thr Leu Leu Trp Glu Ala
    70              75              80

gag agc tcc tgg ccg gcc cag gag gga cag ctg tcc tgg gag tgg ggc 884
Glu Ser Ser Trp Arg Ala Gln Glu Gly Gln Leu Ser Trp Glu Trp Gly
    85              90              95

aag agg cac cgt cga cat cac ttg cca gac aga agt caa ctg tgt cgg 932
Lys Arg His Arg Arg His His Leu Pro Asp Arg Ser Gln Leu Cys Arg
   100              105              110              115

aag gtc aag ttc cag gtg gac ttc aac ctg atc gga tgg ggc tcc tgg 980
Lys Val Lys Phe Gln Val Asp Phe Asn Leu Ile Gly Trp Gly Ser Trp
   120              125              130

atc atc tac ccc aag cag tac aac gcc tat cgc tgt gag ggc gag tgt 1028
Ile Ile Tyr Pro Lys Gln Tyr Asn Ala Tyr Arg Cys Glu Gly Glu Cys
   135              140              145

cct aat cct gtt ggg gag gag ttt cat ccg acc aac cat gca tac atc 1076
Pro Asn Pro Val Gly Glu Glu Phe His Pro Thr Asn His Ala Tyr Ile
   150              155              160

cag gtg gga tgc cag gcg t 1095
Gln Val Gly Cys Gln Ala
    165

```

<210> SEQ ID NO 2

<211> LENGTH: 169

<212> TYPE: PRT

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 2

```

Met Asp Leu Phe Thr Val Thr Leu Ser Gln Val Thr Phe Ser Leu Gly
 1          5          10          15

Ser Met Val Leu Glu Val Thr Arg Pro Leu Ser Lys Trp Leu Lys Arg
 20          25          30

```

-continued

Pro Gly Ala Leu Glu Lys Gln Met Ser Arg Val Ala Gly Glu Cys Trp
 35 40 45

Pro Arg Pro Pro Thr Pro Pro Ala Thr Asn Val Leu Leu Met Leu Tyr
 50 55 60

Ser Asn Leu Ser Gln Glu Gln Arg Gln Leu Gly Gly Ser Thr Leu Leu
 65 70 75 80

Trp Glu Ala Glu Ser Ser Trp Arg Ala Gln Glu Gly Gln Leu Ser Trp
 85 90 95

Glu Trp Gly Lys Arg His Arg Arg His His Leu Pro Asp Arg Ser Gln
 100 105 110

Leu Cys Arg Lys Val Lys Phe Gln Val Asp Phe Asn Leu Ile Gly Trp
 115 120 125

Gly Ser Trp Ile Ile Tyr Pro Lys Gln Tyr Asn Ala Tyr Arg Cys Glu
 130 135 140

Gly Glu Cys Pro Asn Pro Val Gly Glu Glu Phe His Pro Thr Asn His
 145 150 155 160

Ala Tyr Ile Gln Val Gly Cys Gln Ala
 165

<210> SEQ ID NO 3
 <211> LENGTH: 3483
 <212> TYPE: DNA
 <213> ORGANISM: Homo sapiens
 <220> FEATURE:
 <221> NAME/KEY: CDS
 <222> LOCATION: (240)...(872)

<400> SEQUENCE: 3

gtcgacccac gcgtccgggg agcaaccgca gcttctagta tccagactcc agcgccgccc 60

cgggcgcgga cccaacccc gaccagagc ttctccagcg gcggcgagc gagcagggt 120

ccccgcctta acttctccg cggggcccag ccaccttcgg gagtccgggt tgcccactg 180

caaactctcc gccttctgca cctgccacc ctgagccagc gcgggcgccc gagcgagtc 239

atg gcc aac gcg ggg ctg cag ctg ttg ggc ttc att ctc gcc ttc ctg 287
 Met Ala Asn Ala Gly Leu Gln Leu Leu Gly Phe Ile Leu Ala Phe Leu
 1 5 10 15

gga tgg atc ggc gcc atc gtc agc act gcc ctg ccc cag tgg agg att 335
 Gly Trp Ile Gly Ala Ile Val Ser Thr Ala Leu Pro Gln Trp Arg Ile
 20 25 30

tac tcc tat gcc ggc gac aac atc gtg acc gcc cag gcc atg tac gag 383
 Tyr Ser Tyr Ala Gly Asp Asn Ile Val Thr Ala Gln Ala Met Tyr Glu
 35 40 45

ggg ctg tgg atg tcc tgc gtg tgc cag agc acc ggg cag atc cag tgc 431
 Gly Leu Trp Met Ser Cys Val Ser Gln Ser Thr Gly Gln Ile Gln Cys
 50 55 60

aaa gtc ttt gac tcc ttg ctg aat ctg agc agc aca ttg caa gca acc 479
 Lys Val Phe Asp Ser Leu Leu Asn Leu Ser Ser Thr Leu Gln Ala Thr
 65 70 75 80

cgt gcc ttg atg gtg gtt ggc atc ctc ctg gga gtg ata gca atc ttt 527
 Arg Ala Leu Met Val Val Gly Ile Leu Leu Gly Val Ile Ala Ile Phe
 85 90 95

gtg gcc acc gtt ggc atg aag tgt atg aag tgc ttg gaa gac gat gag 575
 Val Ala Thr Val Gly Met Lys Cys Met Lys Cys Leu Glu Asp Asp Glu
 100 105 110

gtg cag aag atg agg atg gct gtc att ggg ggt gcg ata ttt ctt ctt 623

Val	Gln	Lys	Met	Arg	Met	Ala	Val	Ile	Gly	Gly	Ala	Ile	Phe	Leu	Leu	
		115					120				125					
gca	ggt	ctg	gct	att	tta	gtt	gcc	aca	gca	tgg	tat	ggc	aat	aga	atc	671
Ala	Gly	Leu	Ala	Ile	Leu	Val	Ala	Thr	Ala	Trp	Tyr	Gly	Asn	Arg	Ile	
		130				135					140					
gtt	caa	gaa	ttc	tat	gac	cct	atg	acc	cca	gtc	aat	gcc	agg	tac	gaa	719
Val	Gln	Glu	Phe	Tyr	Asp	Pro	Met	Thr	Pro	Val	Asn	Ala	Arg	Tyr	Glu	
		145			150					155					160	
ttt	ggt	cag	gct	ctc	ttc	act	ggc	tgg	gct	gct	gct	tct	ctc	tgc	ctt	767
Phe	Gly	Gln	Ala	Leu	Phe	Thr	Gly	Trp	Ala	Ala	Ala	Ser	Leu	Cys	Leu	
			165					170						175		
ctg	gga	ggt	gcc	cta	ctt	tgc	tgt	tcc	tgt	ccc	cga	aaa	aca	acc	tct	815
Leu	Gly	Gly	Ala	Leu	Leu	Cys	Cys	Ser	Cys	Pro	Arg	Lys	Thr	Thr	Ser	
			180					185					190			
tac	cca	aca	cca	agg	ccc	tat	cca	aaa	cct	gca	cct	tcc	agc	ggg	aaa	863
Tyr	Pro	Thr	Pro	Arg	Pro	Tyr	Pro	Lys	Pro	Ala	Pro	Ser	Ser	Gly	Lys	
		195				200					205					
gac	tac	gtg	tgacacagag	gcaaaaggag	aaaatcatgt	tgaaacaaac										912
Asp	Tyr	Val														
		210														
cgaaaatgga	cattgagata	ctatcattaa	cattaggacc	ttagaatttt	gggtattgta											972
atctgaagta	tggtattaca	aaacaaacaa	acaaacaaaa	aacccatgtg	ttaaaatact											1032
cagtgcataa	catggcctaa	tcttatttta	tcttcttttc	tcaatatagg	aggggaagatt											1092
tttccatttg	tattactgct	tcccattgag	taatcatact	caactggggg	aaggggtgct											1152
ccttaaatat	atatagatat	gtatatatac	atgtttttct	attaaaaata	gacagtaaaa											1212
tactattctc	attatgttga	tactagcata	cttaaaaatat	ctctaaaata	ggtaaatgta											1272
tttaattcca	tattgatgaa	gatgtttatt	ggtatatatt	ctttttcgtc	tatatataca											1332
tatgtaacag	tcaaatatca	tttactcttc	ttcattagct	ttgggtgcct	ttgccacaag											1392
accttagccta	atttaccaag	gatgaattct	ttcaattctt	catgcgtgcc	cttttcatat											1452
acttatttta	ttttttacca	taattottata	gcacttgcct	cgttattaag	cccttatttg											1512
ttttgtgttt	catgtgtctc	tatctcctga	atctaacaca	tttcatagcc	tacatttttag											1572
tttctaaagc	caagaagaat	tattatacaa	tcagaacttt	ggaggcaaat	ctttctgcat											1632
gaccaaagtg	ataaattcct	gttgaccttc	ccacacaatc	cctgtactct	gacccatagc											1692
actcttggtt	gctttgaaaa	tatttgtcca	attgagtagc	tgcatgctgt	tccccagggt											1752
gttgtaacac	aactttattg	attgaatttt	taagctactt	attcatagtt	ttatatcccc											1812
ctaaactacc	tttttgttcc	ccattcctta	attgtattgt	tttcccaagt	gtaattatca											1872
tgcgttttat	atcttcctaa	taagggtggg	tctgtttgtc	tgaacaaagt	gctagacttt											1932
ctggagtgat	aatctgggtga	caaattattct	ctctgtagct	gtaagcaagt	cacttaatct											1992
ttctacctct	tttttctatc	tgccaaattg	agataatgat	acttaaccag	ttagaagagg											2052
tagtgtgaat	attaattagt	ttatattact	ctcattcttt	gaacatgaac	tatgcctatg											2112
tagtgtcttt	atttgctcag	ctggctgaga	cactgaagaa	gtcactgaac	aaaacctaca											2172
cacgtacott	catgtgattc	actgccttcc	tctctctacc	agtcctatttc	cactgaacaa											2232
aacctacaca	cataccttca	tgtgggttcag	tgcttctctc	tctctaccag	tctattttcca											2292
ctgaacaaaa	cctacgcaca	taccttcatg	tggctcagtg	ccttctctct	tctaccagtc											2352
tattttccatt	cttttcagctg	tgtctgacat	gtttgtgctc	tgttccattt	taacaactgc											2412

-continued

```

tcttactttt ccagtcctgta cagaatgcta tttcacttga gcaagatgat gtaatggaaa 2472
gggtgttggtc attggtgtct ggagacctgg atttgagtct tggtgctatc aatcaccgtc 2532
tgtgtttgag caaggcattt ggctgctgta agcttattgc ttcactctgta agcgggtggtt 2592
tgtaattcct gatcttccca cctcacagtg atgttggtgg gatccagtga gatagaatac 2652
atgtaagtgt ggttttgtta tttaaaaagt gctatactaa gggaaagaat tgaggaatta 2712
actgcatacg ttttggtgtt gcttttcaaa tgtttgaaaa caaaaaaat gttaagaaat 2772
gggtttcttg cottaaccag tctctcaagt gatgagacag tgaagtaaaa ttgagtgcac 2832
taaacaaata agattctgag gaagtcttat cttctgcagt gagtatggcc cgatgctttc 2892
tgtggcctaa cagatgtaat gggaagaaat aaaagcctac gtgttggtta atccaacagc 2952
aaggagatt tttgaatcat aataactcat aaggtgctat ctgttcagtgt atgcctcag 3012
agctcttgct gtagctggc agctgacgct gctaggatag ttagtttga aatggtactt 3072
cataataaac tacacaagga aagtcagcca ctgtgtctta tgaggaattg gacctataa 3132
attttagtgt gccttccaaa cctgagaata tatgcttttg gaagttaaaa tttaaatggc 3192
ttttgccaca tacatagatc ttcattgatgt gtgagtgtta ttccatgtgg atatcagtta 3252
ccaaacatta caaaaaaatt ttatggccca aaatgaccaa cgaaattgtt acaatagaat 3312
ttatccaatt ttgatctttt tatattcttc taccacacct ggaacagac caatagacat 3372
tttgggggtt tataatagga atttgtataa agcattactc tttttcaata aattgttttt 3432
taatttaaaa aaaggaaaaa aaaaaaaaaa aaaaaaaaaa agggcgccg c 3483

```

<210> SEQ ID NO 4

<211> LENGTH: 211

<212> TYPE: PRT

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 4

```

Met Ala Asn Ala Gly Leu Gln Leu Leu Gly Phe Ile Leu Ala Phe Leu
 1             5             10            15
Gly Trp Ile Gly Ala Ile Val Ser Thr Ala Leu Pro Gln Trp Arg Ile
          20            25            30
Tyr Ser Tyr Ala Gly Asp Asn Ile Val Thr Ala Gln Ala Met Tyr Glu
          35            40            45
Gly Leu Trp Met Ser Cys Val Ser Gln Ser Thr Gly Gln Ile Gln Cys
          50            55            60
Lys Val Phe Asp Ser Leu Leu Asn Leu Ser Ser Thr Leu Gln Ala Thr
          65            70            75            80
Arg Ala Leu Met Val Val Gly Ile Leu Leu Gly Val Ile Ala Ile Phe
          85            90            95
Val Ala Thr Val Gly Met Lys Cys Met Lys Cys Leu Glu Asp Asp Glu
          100           105           110
Val Gln Lys Met Arg Met Ala Val Ile Gly Gly Ala Ile Phe Leu Leu
          115           120           125
Ala Gly Leu Ala Ile Leu Val Ala Thr Ala Trp Tyr Gly Asn Arg Ile
          130           135           140
Val Gln Glu Phe Tyr Asp Pro Met Thr Pro Val Asn Ala Arg Tyr Glu
          145           150           155           160
Phe Gly Gln Ala Leu Phe Thr Gly Trp Ala Ala Ala Ser Leu Cys Leu

```

165										170					175					
Leu	Gly	Gly	Ala	Leu	Leu	Cys	Cys	Ser	Cys	Pro	Arg	Lys	Thr	Thr	Ser					
			180				185						190							
Tyr	Pro	Thr	Pro	Arg	Pro	Tyr	Pro	Lys	Pro	Ala	Pro	Ser	Ser	Gly	Lys					
			195				200						205							
Asp Tyr Val																				
210																				
<210> SEQ ID NO 5																				
<211> LENGTH: 3569																				
<212> TYPE: DNA																				
<213> ORGANISM: Homo sapiens																				
<220> FEATURE:																				
<221> NAME/KEY: CDS																				
<222> LOCATION: (104)...(1261)																				
<400> SEQUENCE: 5																				
gtcgcacccac gcgtccggct gcgagaacct ttgcacgcgc acaaactacg gggacgattt 60																				
ctgattgatt tttggcgctt tcgatccacc ctctctccct ctc atg gga ctt tgg 115																				
Met Gly Leu Trp																				
1																				
gga caa agc gtc ccg acc gcc tcg agc gct cga gca ggg cgc tat cca 163																				
Gly Gln Ser Val Pro Thr Ala Ser Ser Ala Arg Ala Gly Arg Tyr Pro			10			15			20											
gga gcc agg aca gcg tcg gga acc aga cca tgg ctc ctg gac tcc aag 211																				
Gly Ala Arg Thr Ala Ser Gly Thr Arg Pro Trp Leu Leu Asp Ser Lys			25			30			35											
atc ctt aag ttc gtc gtc ttc atc gtc gcg gtt ctg ctg ccg gtc ccg 259																				
Ile Leu Lys Phe Val Val Phe Ile Val Ala Val Leu Leu Pro Val Arg			40			45			50											
gtt gac tct gcc acc atc ccc ccg cag gac gaa gtt ccc cag cag aca 307																				
Val Asp Ser Ala Thr Ile Pro Arg Gln Asp Glu Val Pro Gln Gln Thr			55			60			65											
gtg gcc cca cag caa cag agg ccg agc ctc aag gag gag gag tgt cca 355																				
Val Ala Pro Gln Gln Gln Arg Ser Leu Lys Glu Glu Glu Cys Pro			70			75			80											
gca gga tct cat aga tca gaa tat act gga gcc tgt aac ccg tgc aca 403																				
Ala Gly Ser His Arg Ser Glu Tyr Thr Gly Ala Cys Asn Pro Cys Thr			85			90			95			100								
gag ggt gtg gat tac acc att gct tcc aac aat ttg cct tct tgc ctg 451																				
Glu Gly Val Asp Tyr Thr Ile Ala Ser Asn Asn Leu Pro Ser Cys Leu			105			110			115											
cta tgt aca gtt tgt aaa tca ggt caa aca aat aaa agt tcc tgt acc 499																				
Leu Cys Thr Val Cys Lys Ser Gly Gln Thr Asn Lys Ser Ser Cys Thr			120			125			130											
acg acc aga gac acc gtg tgt cag tgt gaa aaa gga agc ttc cag gat 547																				
Thr Thr Arg Asp Thr Val Cys Gln Cys Glu Lys Gly Ser Phe Gln Asp			135			140			145											
aaa aac tcc cct gag atg tgc ccg acg tgt aga aca ggg tgt ccc aga 595																				
Lys Asn Ser Pro Glu Met Cys Arg Thr Cys Arg Thr Gly Cys Pro Arg			150			155			160											
ggg atg gtc aag gtc agt aat tgt acg ccc ccg agt gac atc aag tgc 643																				
Gly Met Val Lys Val Ser Asn Cys Thr Pro Arg Ser Asp Ile Lys Cys			165			170			175			180								
aaa aat gaa tca gct gcc agt tcc act ggg aaa acc cca gca gcg gag 691																				
Lys Asn Glu Ser Ala Ala Ser Ser Thr Gly Lys Thr Pro Ala Ala Glu			185			190			195											

-continued

gag aca gtg acc acc atc ctg ggg atg ctt gcc tct ccc tat cac tac Glu Thr Val Thr Thr Ile Leu Gly Met Leu Ala Ser Pro Tyr His Tyr 200 205 210	739
ctt atc atc ata gtg gtt tta gtc atc att tta gct gtg gtt gtg gtt Leu Ile Ile Ile Val Val Leu Val Ile Ile Leu Ala Val Val Val Val 215 220 225	787
ggc ttt tca tgt cgg aag aaa ttc att tct tac ctc aaa ggc atc tgc Gly Phe Ser Cys Arg Lys Lys Phe Ile Ser Tyr Leu Lys Gly Ile Cys 230 235 240	835
tca ggt ggt gga gga ggt ccc gaa cgt gtg cac aga gtc ctt ttc cgg Ser Gly Gly Gly Gly Gly Pro Glu Arg Val His Arg Val Leu Phe Arg 245 250 255 260	883
cgg cgt tca tgt cct tca cga gtt cct ggg gcg gag gac aat gcc cgc Arg Arg Ser Cys Pro Ser Arg Val Pro Gly Ala Glu Asp Asn Ala Arg 265 270 275	931
aac gag acc ctg agt aac aga tac ttg cag ccc acc cag gtc tct gag Asn Glu Thr Leu Ser Asn Arg Tyr Leu Gln Pro Thr Gln Val Ser Glu 280 285 290	979
cag gaa atc caa ggt cag gag ctg gca gag cta aca ggt gtg act gta Gln Glu Ile Gln Gly Gln Glu Leu Ala Glu Leu Thr Gly Val Thr Val 295 300 305	1027
gag tcg cca gag gag cca cag cgt ctg ctg gaa cag gca gaa gct gaa Glu Ser Pro Glu Glu Pro Gln Arg Leu Leu Glu Gln Ala Glu Ala Glu 310 315 320	1075
ggg tgt cag agg agg agg ctg ctg gtt cca gtg aat gac gct gac tcc Gly Cys Gln Arg Arg Arg Leu Leu Val Pro Val Asn Asp Ala Asp Ser 325 330 335 340	1123
gct gac atc agc acc ttg ctg gat gcc tcg gca aca ctg gaa gaa gga Ala Asp Ile Ser Thr Leu Leu Asp Ala Ser Ala Thr Leu Glu Glu Gly 345 350 355	1171
cat gca aag gaa aca att cag gac caa ctg gtg ggc tcc gaa aag ctc His Ala Lys Glu Thr Ile Gln Asp Gln Leu Val Gly Ser Glu Lys Leu 360 365 370	1219
ttt tat gaa gaa gat gaa gca ggc tct gct acg tcc tgc ctg Phe Tyr Glu Glu Asp Glu Ala Gly Ser Ala Thr Ser Cys Leu 375 380 385	1261
tgaaagaatc tcttcaggaa accagagcct ccttcattta ccttttctcc tacaaaggga	1321
agcagcctgg aagaacagct ccagtacttg acccatgccc caacaaactc tactatccaa	1381
tatggggcag cttaccaatg gtcctagaac tttgttaacg cacttggagt aatttttatg	1441
aaatactgcg tgtgataagc aaacgggaga aatttatatc agattcttggt ctgcatagtt	1501
atacgattgt gtattaaggg tcgttttagg ccacatgcgg tggctcatgc ctgtaatccc	1561
agcactttga taggctgagg caggtggatt gottgagctc gggagtttga gaccagcctc	1621
atcaacacag tgaaactcca tctcaattta aaaagaaaaa aagtggtttt aggatgtcat	1681
tctttgcagt tottcatcat gagacaagtc tttttttctg cttottatat tgcaagctcc	1741
atctctactg gtgtgtgcat ttaatgacat ctaactacag atgccgcaca gccacaatgc	1801
tttgcccttat aatttttttaa ctttagaacg ggattatctt gttattacct gtattttcag	1861
tttcggatat ttttgactta atgatgagat tatcaagacg tagccctatg ctaagtcatg	1921
agcatatgga cttacgaggg ttcgacttag agttttgagc tttaagatac gattattggg	1981
gcttaccccc accttaatta gagaaacatt tatattgctt actactgtag gctgtacatc	2041
tcttttccga tttttgtata atgatgtaaa catggaaaaa ctttaggaaa tgcacttatt	2101

-continued

```

aggctgttta catgggttgc ctggatacaa atcagcagtc aaaaatgact aaaaatataa 2161
ctagtgcagg agggagaaat cctccctctg tgggaggcac ttactgcatt ccagttctcc 2221
ctcctgcgcc ctgagactgg accagggttt gatggctggc agcttctcaa ggggcagctt 2281
gtcttacttg ttaattttag aggtatatag ccatatttat ttataaataa atatttattt 2341
atttatttat aagtagatgt ttacatatgc ccaggatttt gaagagcctg gtatctttgg 2401
gaagccatgt gtctggtttg tcgtgctggg acagtcatgg gactgcatct tccgacttgt 2461
ccacagcaga tgaggacagt gagaattaag ttagatccga gactgcgaag agcttctctt 2521
tcaagcgcca ttacagttga acgttagtga atcttgagcc tcatttgggc tcagggcaga 2581
gcaggtgttt atctgccccg gcactcgcca tggcatcaag agggaagagt ggacggtgct 2641
tgggaatggt gtgaaatggt tgccgactca ggcattgatg gggccctctc gcttctggtg 2701
gtctgtgaac tgagtccctg ggatgccttt tagggcagag attcctgagc tgcgttttag 2761
ggtacagatt ccctgtttga ggagcttggc ccctctgtaa gcacttgact catctcagag 2821
atatcaatto ttaaacactg tgacaacagg atctaaaatg gctgacacat ttgtccttgt 2881
gtcacgttcc attattttat ttaaaaacct cagtaatcgt tttagcttct ttccagcaaa 2941
ctcttctcca cagtagccca gtcgtggtag gataaattac ggatatagtc attctagggg 3001
tttcagtctt ttccatctca aggcattgtg tgtttgttc cgggactggt ttggctggga 3061
caaagttaga actgcctgaa gttgcacat tcagattggt gtgtccatgg agtttttagga 3121
ggggatggcc ttccggtct tcgcacttcc atcctctccc acttccatct ggcgtccac 3181
acottgtccc ctgcacttct ggatgacaca ggggtctgct gcctcctagt ctttgccttt 3241
gctgggcctt ctgtgcagga gacttggctt caaagctcag agagagccag tccggtccca 3301
gtccctttgt cccttctca gaggccttcc ttgaagatgc atctagacta ccagccttat 3361
cagtgtttaa gcttattcct ttaacataag cttcctgaca acatgaaatt gttggggttt 3421
tttggcgttg gttgaattgt ttaggttttg ctttataccc gggccaaata gcacataaca 3481
cctggttata tatgaaatac tcatatgttt atgaccaaaa taaatatgaa acctcatatt 3541
aaaaaaaaa aaaaaaaagg gcggccgc 3569

```

<210> SEQ ID NO 6

<211> LENGTH: 386

<212> TYPE: PRT

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 6

```

Met Gly Leu Trp Gly Gln Ser Val Pro Thr Ala Ser Ser Ala Arg Ala
 1             5             10            15

```

```

Gly Arg Tyr Pro Gly Ala Arg Thr Ala Ser Gly Thr Arg Pro Trp Leu
 20             25            30

```

```

Leu Asp Ser Lys Ile Leu Lys Phe Val Val Phe Ile Val Ala Val Leu
 35             40            45

```

```

Leu Pro Val Arg Val Asp Ser Ala Thr Ile Pro Arg Gln Asp Glu Val
 50             55            60

```

```

Pro Gln Gln Thr Val Ala Pro Gln Gln Gln Arg Arg Ser Leu Lys Glu
 65             70            75            80

```

```

Glu Glu Cys Pro Ala Gly Ser His Arg Ser Glu Tyr Thr Gly Ala Cys
 85             90            95

```

-continued

Asn	Pro	Cys	Thr	Glu	Gly	Val	Asp	Tyr	Thr	Ile	Ala	Ser	Asn	Asn	Leu
		100						105					110		
Pro	Ser	Cys	Leu	Leu	Cys	Thr	Val	Cys	Lys	Ser	Gly	Gln	Thr	Asn	Lys
		115					120					125			
Ser	Ser	Cys	Thr	Thr	Thr	Arg	Asp	Thr	Val	Cys	Gln	Cys	Glu	Lys	Gly
		130				135					140				
Ser	Phe	Gln	Asp	Lys	Asn	Ser	Pro	Glu	Met	Cys	Arg	Thr	Cys	Arg	Thr
	145				150					155					160
Gly	Cys	Pro	Arg	Gly	Met	Val	Lys	Val	Ser	Asn	Cys	Thr	Pro	Arg	Ser
			165						170					175	
Asp	Ile	Lys	Cys	Lys	Asn	Glu	Ser	Ala	Ala	Ser	Ser	Thr	Gly	Lys	Thr
		180						185					190		
Pro	Ala	Ala	Glu	Glu	Thr	Val	Thr	Thr	Ile	Leu	Gly	Met	Leu	Ala	Ser
	195						200					205			
Pro	Tyr	His	Tyr	Leu	Ile	Ile	Ile	Val	Val	Leu	Val	Ile	Ile	Leu	Ala
	210				215						220				
Val	Val	Val	Val	Gly	Phe	Ser	Cys	Arg	Lys	Lys	Phe	Ile	Ser	Tyr	Leu
	225			230						235					240
Lys	Gly	Ile	Cys	Ser	Gly	Gly	Gly	Gly	Gly	Pro	Glu	Arg	Val	His	Arg
			245					250						255	
Val	Leu	Phe	Arg	Arg	Arg	Ser	Cys	Pro	Ser	Arg	Val	Pro	Gly	Ala	Glu
		260						265					270		
Asp	Asn	Ala	Arg	Asn	Glu	Thr	Leu	Ser	Asn	Arg	Tyr	Leu	Gln	Pro	Thr
		275					280					285			
Gln	Val	Ser	Glu	Gln	Glu	Ile	Gln	Gly	Gln	Glu	Leu	Ala	Glu	Leu	Thr
	290					295					300				
Gly	Val	Thr	Val	Glu	Ser	Pro	Glu	Glu	Pro	Gln	Arg	Leu	Leu	Glu	Gln
	305			310						315					320
Ala	Glu	Ala	Glu	Gly	Cys	Gln	Arg	Arg	Arg	Leu	Leu	Val	Pro	Val	Asn
			325						330					335	
Asp	Ala	Asp	Ser	Ala	Asp	Ile	Ser	Thr	Leu	Leu	Asp	Ala	Ser	Ala	Thr
		340						345					350		
Leu	Glu	Glu	Gly	His	Ala	Lys	Glu	Thr	Ile	Gln	Asp	Gln	Leu	Val	Gly
	355					360						365			
Ser	Glu	Lys	Leu	Phe	Tyr	Glu	Glu	Asp	Glu	Ala	Gly	Ser	Ala	Thr	Ser
	370					375					380				
Cys	Leu														
	385														

<210> SEQ ID NO 7
 <211> LENGTH: 2114
 <212> TYPE: DNA
 <213> ORGANISM: Rattus norvegicus
 <220> FEATURE:
 <221> NAME/KEY: CDS
 <222> LOCATION: (3)...(1445)

<400> SEQUENCE: 7

gc	gtc	cgg	aac	aag	acg	ctg	ccc	tgg	tct	ccc	tgc	agt	gct	gtc	tac	47
Val	Arg	Asn	Lys	Thr	Leu	Pro	Trp	Ser	Pro	Cys	Ser	Ala	Val	Tyr		
1					5					10				15		
ctc	acg	gag	ctc	ctg	gat	gat	ggg	cac	gga	gac	tgc	ctc	ctg	gat	gat	95
Leu	Thr	Glu	Leu	Leu	Asp	Asp	Gly	His	Gly	Asp	Cys	Leu	Leu	Asp	Asp	
			20					25						30		

-continued

ggc cac agc acc ctc tat gag ctg gac cag cag tgc aag cag atc ttt Gly His Ser Thr Leu Tyr Glu Leu Asp Gln Gln Cys Lys Gln Ile Phe 35 40 45	143
ggg cct gat ttc cga cac tgc ccc aac acc tct gtg gag gac atc tgt Gly Pro Asp Phe Arg His Cys Pro Asn Thr Ser Val Glu Asp Ile Cys 50 55 60	191
gtc cag ctc tgg tgc cgt cat cgg gat agt gat gag ccc att tgc cac Val Gln Leu Trp Cys Arg His Arg Asp Ser Asp Glu Pro Ile Cys His 65 70 75	239
aca aag aat gcc agc ttg ctc tgg gct gat ggt acg ccc tgt ggc cct Thr Lys Asn Ala Ser Leu Leu Trp Ala Asp Gly Thr Pro Cys Gly Pro 80 85 90 95	287
ggg cac ctg tgc ctg gat ggt agc tgt gtg ctc cgg gag gaa gta gag Gly His Leu Cys Leu Asp Gly Ser Cys Val Leu Arg Glu Glu Val Glu 100 105 110	335
aat ccc aag gct gtg gta gat gga gac tgg ggt ccc tgg gga ccc tgg Asn Pro Lys Ala Val Val Asp Gly Asp Trp Gly Pro Trp Gly Pro Trp 115 120 125	383
gga caa tgt tct cgc acc tgt ggt gga ggg ata cag ttt tcg aac cgt Gly Gln Cys Ser Arg Thr Cys Gly Gly Ile Gln Phe Ser Asn Arg 130 135 140	431
gag tgt gat aat cca gca cct cag aat gga gga aga ttt tgc ctg gga Glu Cys Asp Asn Pro Ala Pro Gln Asn Gly Gly Arg Phe Cys Leu Gly 145 150 155	479
gag aga gtc aag tac caa tct tgc aag aca gag gaa tgt cca cca aat Glu Arg Val Lys Tyr Gln Ser Cys Lys Thr Glu Glu Cys Pro Pro Asn 160 165 170 175	527
gga aaa agc ttc agg gag cag cag tgt gaa aaa tat aat gcc tac aac Gly Lys Ser Phe Arg Glu Gln Gln Cys Glu Lys Tyr Asn Ala Tyr Asn 180 185 190	575
cac acg gac ctg gat ggg aat ttc ctt cag tgg gtc ccc aaa tac tca His Thr Asp Leu Asp Gly Asn Phe Leu Gln Trp Val Pro Lys Tyr Ser 195 200 205	623
gga gtg tcc ccc cga gac cga tgc aaa ctg ttt tgc aga gcc cgt ggg Gly Val Ser Pro Arg Asp Arg Cys Lys Leu Phe Cys Arg Ala Arg Gly 210 215 220	671
agg agt gag ttc aaa gtg ttt gaa act aag gtg atc gat ggc act ctg Arg Ser Glu Phe Lys Val Phe Glu Thr Lys Val Ile Asp Gly Thr Leu 225 230 235	719
tgc gga ccg gat act ctg gcc atc tgt gtg cgg gga cag tgc gtt aag Cys Gly Pro Asp Thr Leu Ala Ile Cys Val Arg Gly Gln Cys Val Lys 240 245 250 255	767
gct ggc tgt gac cat gtg gtg aac tca cct aag aag ctg gac aag tgt Ala Gly Cys Asp His Val Val Asn Ser Pro Lys Lys Leu Asp Lys Cys 260 265 270	815
ggg gtg tgt ggg ggc aaa ggc act gcc tgt agg aag gtc tca ggt tct Gly Val Cys Gly Gly Lys Gly Thr Ala Cys Arg Lys Val Ser Gly Ser 275 280 285	863
ttc acc ccc ttc agt tat ggc tac aat gac att gtc acc atc cca gct Phe Thr Pro Phe Ser Tyr Gly Tyr Asn Asp Ile Val Thr Ile Pro Ala 290 295 300	911
ggt gcc aca aat att gat gtg aaa caa cgg agc cac cca ggg gtc cag Gly Ala Thr Asn Ile Asp Val Lys Gln Arg Ser His Pro Gly Val Gln 305 310 315	959
aat gac ggc agc tac ctg gca ctg aag aca gcc aat ggg cag tac ctg Asn Asp Gly Ser Tyr Leu Ala Leu Lys Thr Ala Asn Gly Gln Tyr Leu 320 325 330 335	1007

-continued

ctc aat ggt aac cta gcc atc tct gcc ata gag caa gac atc ttg atg	1055
Leu Asn Gly Asn Leu Ala Ile Ser Ala Ile Glu Gln Asp Ile Leu Met	
340 345 350	
aag ggg acc atc cta aag tac agt ggt tcc atg gcc acc ctg gag cgg	1103
Lys Gly Thr Ile Leu Lys Tyr Ser Gly Ser Met Ala Thr Leu Glu Arg	
355 360 365	
ctg cag agc ttc caa gcc ctc cct gag cct ctt aca gta cag ctc ctg	1151
Leu Gln Ser Phe Gln Ala Leu Pro Glu Pro Leu Thr Val Gln Leu Leu	
370 375 380	
act gtg tct ggt gag gtc ttc cct cca aaa gtc aaa tat acc ttc ttc	1199
Thr Val Ser Gly Glu Val Phe Pro Pro Lys Val Lys Tyr Thr Phe Phe	
385 390 395	
gtc ccc aat gac acg gac ttc aac gtg cag agt agc aaa gaa aga gca	1247
Val Pro Asn Asp Thr Asp Phe Asn Val Gln Ser Ser Lys Glu Arg Ala	
400 405 410 415	
agc acc aac atc att cag tcc ttg ccc tat gca gag tgg gtg ctg ggg	1295
Ser Thr Asn Ile Ile Gln Ser Leu Pro Tyr Ala Glu Trp Val Leu Gly	
420 425 430	
gac tgg tct gaa tgt cca agc aca tgt gga ggt ggc tgg cag cgg cgg	1343
Asp Trp Ser Glu Cys Pro Ser Thr Cys Gly Gly Gly Trp Gln Arg Arg	
435 440 445	
act gtg gaa tgc agg gac ccc tca ggt cag gcc tct gac acc tgt gat	1391
Thr Val Glu Cys Arg Asp Pro Ser Gly Gln Ala Ser Asp Thr Cys Asp	
450 455 460	
gag gct ctg aaa cct gag gat gcc aag ccc tgt gga agc cag cca tgt	1439
Glu Ala Leu Lys Pro Glu Asp Ala Lys Pro Cys Gly Ser Gln Pro Cys	
465 470 475	
ctc ctc tgatcccctt ggtggacatg tctaaggctt atggatttgg gctactggcg	1495
Leu Leu	
480	
tacagacaaa ggtctcctct gaggtgacac tacatatcaa gatggcatgg cccttccagg	1555
ccttctatta ctacaacctt ttgggtacca cctaattcat aaggaagaga gaagaggatg	1615
taagggtaac agactgtaaa gttgactgtc tagtggactg gaccttgttt atgaccaaga	1675
agatgggata ggttaaaagg taaaagtgtg cttattgac caaaggtgag atttcagaac	1735
cagcctcttt gcaaaggact agaaaggtta aatgagaaag aagaattttt tttctctttg	1795
gtttctccaa taatcaatct acctcacagc gggaggaact tgggtgataa ggccaggtgt	1855
tagtgggtgag tgccaaggca ctctccatag atatcttcga gccatcttca gaaatggcca	1915
tggtgtgttt cagtattaaa actctgttgt ctcaaaagggt ggtgggtgtcc atcacagggt	1975
tatagaaagc cacttgttct caggctgcct cctgctgggg cggaacctt tcaagtattt	2035
atgcaaatat gtttctgaac taaagtgtga tottacacca aaaaaaaaaa aaaaaaaaaa	2095
aaaaaaaaa ggcgccgc	2114

<210> SEQ ID NO 8

<211> LENGTH: 481

<212> TYPE: PRT

<213> ORGANISM: Rattus norvegicus

<400> SEQUENCE: 8

Val Arg Asn Lys Thr Leu Pro Trp Ser Pro Cys Ser Ala Val Tyr Leu
1 5 10 15
Thr Glu Leu Leu Asp Asp Gly His Gly Asp Cys Leu Leu Asp Asp Gly
20 25 30

-continued

His	Ser	Thr	Leu	Tyr	Glu	Leu	Asp	Gln	Gln	Cys	Lys	Gln	Ile	Phe	Gly	35	40	45	
Pro	Asp	Phe	Arg	His	Cys	Pro	Asn	Thr	Ser	Val	Glu	Asp	Ile	Cys	Val	50	55	60	
Gln	Leu	Trp	Cys	Arg	His	Arg	Asp	Ser	Asp	Glu	Pro	Ile	Cys	His	Thr	65	70	75	80
Lys	Asn	Ala	Ser	Leu	Leu	Trp	Ala	Asp	Gly	Thr	Pro	Cys	Gly	Pro	Gly	85	90	95	
His	Leu	Cys	Leu	Asp	Gly	Ser	Cys	Val	Leu	Arg	Glu	Glu	Val	Glu	Asn	100	105	110	
Pro	Lys	Ala	Val	Val	Asp	Gly	Asp	Trp	Gly	Pro	Trp	Gly	Pro	Trp	Gly	115	120	125	
Gln	Cys	Ser	Arg	Thr	Cys	Gly	Gly	Gly	Ile	Gln	Phe	Ser	Asn	Arg	Glu	130	135	140	
Cys	Asp	Asn	Pro	Ala	Pro	Gln	Asn	Gly	Gly	Arg	Phe	Cys	Leu	Gly	Glu	145	150	155	160
Arg	Val	Lys	Tyr	Gln	Ser	Cys	Lys	Thr	Glu	Glu	Cys	Pro	Pro	Asn	Gly	165	170	175	
Lys	Ser	Phe	Arg	Glu	Gln	Gln	Cys	Glu	Lys	Tyr	Asn	Ala	Tyr	Asn	His	180	185	190	
Thr	Asp	Leu	Asp	Gly	Asn	Phe	Leu	Gln	Trp	Val	Pro	Lys	Tyr	Ser	Gly	195	200	205	
Val	Ser	Pro	Arg	Asp	Arg	Cys	Lys	Leu	Phe	Cys	Arg	Ala	Arg	Gly	Arg	210	215	220	
Ser	Glu	Phe	Lys	Val	Phe	Glu	Thr	Lys	Val	Ile	Asp	Gly	Thr	Leu	Cys	225	230	235	240
Gly	Pro	Asp	Thr	Leu	Ala	Ile	Cys	Val	Arg	Gly	Gln	Cys	Val	Lys	Ala	245	250	255	
Gly	Cys	Asp	His	Val	Val	Asn	Ser	Pro	Lys	Lys	Leu	Asp	Lys	Cys	Gly	260	265	270	
Val	Cys	Gly	Gly	Lys	Gly	Thr	Ala	Cys	Arg	Lys	Val	Ser	Gly	Ser	Phe	275	280	285	
Thr	Pro	Phe	Ser	Tyr	Gly	Tyr	Asn	Asp	Ile	Val	Thr	Ile	Pro	Ala	Gly	290	295	300	
Ala	Thr	Asn	Ile	Asp	Val	Lys	Gln	Arg	Ser	His	Pro	Gly	Val	Gln	Asn	305	310	315	320
Asp	Gly	Ser	Tyr	Leu	Ala	Leu	Lys	Thr	Ala	Asn	Gly	Gln	Tyr	Leu	Leu	325	330	335	
Asn	Gly	Asn	Leu	Ala	Ile	Ser	Ala	Ile	Glu	Gln	Asp	Ile	Leu	Met	Lys	340	345	350	
Gly	Thr	Ile	Leu	Lys	Tyr	Ser	Gly	Ser	Met	Ala	Thr	Leu	Glu	Arg	Leu	355	360	365	
Gln	Ser	Phe	Gln	Ala	Leu	Pro	Glu	Pro	Leu	Thr	Val	Gln	Leu	Leu	Thr	370	375	380	
Val	Ser	Gly	Glu	Val	Phe	Pro	Pro	Lys	Val	Lys	Tyr	Thr	Phe	Phe	Val	385	390	395	400
Pro	Asn	Asp	Thr	Asp	Phe	Asn	Val	Gln	Ser	Ser	Lys	Glu	Arg	Ala	Ser	405	410	415	
Thr	Asn	Ile	Ile	Gln	Ser	Leu	Pro	Tyr	Ala	Glu	Trp	Val	Leu	Gly	Asp	420	425	430	

-continued

Trp	Ser	Glu	Cys	Pro	Ser	Thr	Cys	Gly	Gly	Gly	Trp	Gln	Arg	Arg	Thr
		435					440					445			
Val	Glu	Cys	Arg	Asp	Pro	Ser	Gly	Gln	Ala	Ser	Asp	Thr	Cys	Asp	Glu
	450					455					460				
Ala	Leu	Lys	Pro	Glu	Asp	Ala	Lys	Pro	Cys	Gly	Ser	Gln	Pro	Cys	Leu
465					470					475					480

Leu

<210> SEQ ID NO 9
 <211> LENGTH: 1448
 <212> TYPE: DNA
 <213> ORGANISM: Homo sapiens

<400> SEQUENCE: 9

```

gtcgacccac gcgtccgggg gaagcttgcc agcagatctg cagctgccaa aatggggcag      60
actgtgacag tgtgactgga aagtgcacct gtgccccagg attcaaagga attgactgct      120
ctaccccatg cctcttgga acctatggga taaactgttc ctctcgctgt ggctgtaaaa      180
atgatgcagt ctgctctcct gtggacgggt cttgtacttg caaggcaggc tggcacgggg      240
tggactgctc catcagatgt cccagtggca catggggctt tggtgtaac ttaacatgcc      300
agtgcctcaa cgggggagcc tgcaacaccc tggacgggac ctgcacgtgt gcacctggat      360
ggcgcgggga gaaatgcgaa cttccctgcc aggatggcac gtacgggctg aactgtgctg      420
agcgctgcga ctgcagccac gcagatggct gccaccctac cacgggccat tgccgctgcc      480
tccccggatg gtcagggtgc cactgtgaca gcgtgtgtgc tgagggacgc tggggcccca      540
actgctccct gccctgctac tgtaaaaatg gggcttcctg ctcccctgat gatggcatct      600
gcgagtgtgc accaggcttc cgaggcacca cttgtcagag gatctgctcc cctggttttt      660
atgggcatcg ctgcagccag acatgcccac agtgcgttca cagcagcggg ccctgccacc      720
acatcaccgg cctgtgtgac tgcttgctg gcttcacagg cgccctctgc aatgaagtgt      780
gtcccagtgg cagatttggg aaaaactgtg caggaatttg tacctgcacc aacaacggaa      840
cctgtaaccc cattgacaga tcttgtcagt gttaccccg tggtattggc agtgactgct      900
ctcaaccatg tccacctgcc cactggggcc caaactgcat ccacacgtgc aactgccata      960
atggagcttt ctgcagcgcc tacgatggg aatgtaaatg cactcctggc tggacagggc      020
tctactgcac tcagagatgt cctctagggt tttatggaaa agattgtgca ctgatatgcc      080
aatgtcaaaa cggagctgac tgcgaccaca tttctgggca gtgtacttgc cgcactggat     1140
tcattgggacg gcactgtgag cagaagtgcc cttcaggaac atatggctat ggctgtcgcc     1200
agatatgtga ttgtctgaac aactccacct gcgaccacat cactgggacc tgttactgca     1260
gccccggatg gaagggagcg agatgtgatc aagctggtgt tatcatagtt ggaaatctga     1320
acagcttaag ccgaaccagt actgctctcc ctgctgattc ctacaaaatc ggggccattg     1380
caggcatcat cattcttgtc ctagtgttcc tcttcctact ggcattgttc attatttata     1440
gacacagc                                         1448

```

<210> SEQ ID NO 10
 <211> LENGTH: 1578
 <212> TYPE: DNA
 <213> ORGANISM: Homo sapiens
 <220> FEATURE:
 <221> NAME/KEY: misc_feature

-continued

<222> LOCATION: (1)...(1578)

<223> OTHER INFORMATION: n = A,T,C or G

<400> SEQUENCE: 10

```

nagcccaaca ggaatgttct atgaaagtga acctaacagt gagtgttggt cccaaggagt    60
attcagcaat aatgggcgctc tntcccaagg atccatatga cctcccaaag aacagtcaca    120
tcccttgatc ttatgacctg ctgccagtcg gagacagttc atcctcccct aagcaagagg    180
acagtggagg tagcagcagc aacagcagca gcagcagtgat atgacaccaa aggaccgctt    240
ggtagccact ggaacccttt ccagaactgc tgtttggttc ttctccatcc tcaattttgc    300
cactttcatg tgaatgttag tcaattcggg gggcaatttt tggacatgaa ccagaaagct    360
gaaagctgag gctgacacgg actgtagggt cttttgttgc aggtggattc gaaggagtta    420
gagatgtgat ttgccattgc tgttagtttt agaactatac cgtgaagca tgacttattg    480
taagatgttg gctgaaagca tgaacttgca gaactccctc ggagacgcag gttgcagtgg    540
acattgggat tgttgcttga aaaattaaaa tttgaatatt ttctctctca ttgcatcat    600
agagctctac ctaggattgt acagtttacc ataaaattta cttcatgaaa gtgggaatca    660
ctgaacatgt agaagacaag gaacatattg ttaactcctg attottaact ttattcaact    720
ggactcagaa ttgtagggat aatatgaatg caggaggaaa cattctgtca gccggtatga    780
ctggacagac tttgaatata ctctaaaagt ggacagaaaa ttacgaaaa tcttagattt    840
tgttttagat gagaaaatat acaattagaa ttattttaga aatagtagga agtattgcag    900
aagtcaatac acaaatgtgc caggcagagg tggttttctc tgtttgactc tcaaccaact    960
tcagatctat gacattattc tgatcactgg ctccatcata catattcacc acttgagatt   1020
cataacatat caatagttat ttcataaata tagaaatgaa ataattttat ttttgacaga   1080
ctggatggaa tgagtgtgta atgattgata aaggttgtaa attttaaatg caagatgacg   1140
cttacgttct gtaaaccatt agtaatacat gctgtaatat agaattagtg gaacattttg   1200
attaatcttt ccttagaagt gactgaaata tttttgtgca tatttgagaa agggaaacttt   1260
ccttttatta attgtcaatt tagagaaact atgcttaagc tggctttttg cattgctaatt   1320
gtgacatgta cccaactttt cattaatttg tatttccatt tttaaattgc atattctatg   1380
ttttgtagtg tttggattgt taatgaaaaa atattatatg ttcgttattc cttgtattat   1440
tgccacttat cttttgcttg ataaaaatgc gttgttcttt tttcttttgg agggacaaga   1500
tgaaaatata taatttgaat tgattaaaat tggtcgttac taaaatagta tagtaaaaaa   1560
aaaaaaaaag ggcggccg                                     1578

```

<210> SEQ ID NO 11

<211> LENGTH: 843

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 11

```

gccttttagg gcagagattc ctgagctgcg ttttagggta cagattccct gtttgaggag    60
cttgccccct ctgtaagcat ctgactcacc tcagagatat caattcttaa acaactgtgac    120
aacaggatct aaaatggctg acacatttgt ccttgtgtca cgttccatta ttttatttaa    180
aaacgtcagt aatcgtttta gcttctttcc agcaaaactct tctccacagt agcccagtcg    240
tggtaggata aattacggat atagtcattc taggggtttc agtccttttc atctcaaggc    300

```

-continued

```

attgtgtgtt ttgttccggg actggtttgg ctgggacaaa gttagaactg cctgaagtgc 360
gcacattcag attgttgtgt ccatggagtt ttaggagggg atggcctttc cggctctcgg 420
acttccatcc tctcccaactt ccatctggcg tcccacacct tgtcccctgc acttctggat 480
gacacagggg gctgctgcct cctagtcttt gcctttgctg ggccttctgt gcaggagact 540
tggctcctcaa gctcagagag agccagtcg gtcccagctc cttgtccct tcctcagagg 600
ccttccttga agatgcacat agactaccag ccttatcagt gtttaagctt attcctttaa 660
cataagcttc ctgacaacat gaaattgttg gggttttttg gcgttggttg atttgtttag 720
gttttgcttt ataccgggc caaatagcac ataacacctg gttatatatg aaatactcat 780
atgtttatga ccaaaataaa tatgaaacct catattaaaa aaaaaaaaaa aaaagggcgg 840
ccg 843

```

```

<210> SEQ ID NO 12
<211> LENGTH: 248
<212> TYPE: PRT
<213> ORGANISM: Mus musculus

```

```

<400> SEQUENCE: 12

```

```

Pro Met Asp Ile Pro Thr Glu Gly Pro Leu Thr Ile Asp Ile Phe His
 1             5             10            15
Gln Ala Lys Gly Asp Pro Glu Arg Asp Pro Ala Asp Cys Leu Glu Arg
      20             25            30
Ile Trp Met Glu Thr Phe Thr Val Ile Pro Ser Gln Val Thr Phe Ala
      35             40            45
Ser Gly Ser Thr Val Leu Glu Val Thr Lys Pro Leu Ser Lys Trp Lys
      50             55            60
Asp Pro Arg Ala Leu Glu Lys Gln Val Ser Ser Arg Ala Glu Lys Cys
      65             70            75            80
Trp His Gln Pro Tyr Thr Pro Pro Val Pro Val Ala Ser Thr Asn Val
      85             90            95
Leu Met Leu Tyr Ser Asn Arg Pro Gln Glu Gln Arg Gln Leu Gly Gly
      100            105           110
Ala Thr Leu Leu Trp Glu Ala Glu Ser Ser Trp Arg Ala Gln Gly Gln
      115            120           125
Leu Ser Val Glu Arg Gly Gly Trp Gly Arg Arg Gln Arg Arg His His
      130            135           140
Leu Pro Asp Arg Ser Gln Leu Cys Arg Arg Val Lys Phe Gln Val Asp
      145            150           155           160
Phe Asn Leu Ile Gly Trp Gly Ser Trp Ile Ile Tyr Pro Lys Gln Tyr
      165            170           175
Asn Ala Tyr Arg Cys Glu Gly Glu Cys Pro Asn Pro Val Gly Glu Glu
      180            185           190
Phe His Pro Thr Asn His Ala Tyr Ile Gln Ser Leu Leu Lys Arg Tyr
      195            200           205
Gln Pro His Arg Val Pro Ser Thr Cys Cys Ala Pro Val Lys Thr Lys
      210            215           220
Pro Leu Ser Met Leu Tyr Val Asp Asn Gly Arg Val Leu Leu Glu His
      225            230           235           240
His Lys Asp Met Ile Val Glu Glu
      245

```

-continued

<210> SEQ ID NO 13
 <211> LENGTH: 234
 <212> TYPE: PRT
 <213> ORGANISM: Homo sapiens

<400> SEQUENCE: 13

```

Ala Ala Ala Thr Pro Ser Lys Val Trp Gly Ser Ser Ala Gly Arg Ile
 1             5             10             15
Glu Pro Arg Gly Gly Gly Arg Gly Ala Leu Pro Thr Ser Met Gly Gln
             20             25             30
His Gly Pro Ser Ala Arg Ala Arg Ala Gly Arg Ala Pro Gly Pro Arg
             35             40             45
Pro Ala Arg Glu Ala Ser Pro Arg Leu Arg Val His Lys Thr Phe Lys
             50             55             60
Phe Val Val Val Gly Val Leu Leu Gln Val Val Pro Ser Ser Ala Ala
65             70             75             80
Thr Ile Lys Leu His Asp Gln Ser Ile Gly Thr Gln Gln Trp Glu His
             85             90             95
Ser Pro Leu Gly Glu Leu Cys Pro Pro Gly Ser His Arg Ser Glu Arg
             100            105            110
Pro Gly Ala Cys Asn Arg Cys Thr Glu Gly Val Gly Tyr Thr Asn Ala
             115            120            125
Ser Asn Asn Leu Phe Ala Cys Leu Pro Cys Thr Ala Cys Lys Ser Asp
             130            135            140
Glu Glu Glu Arg Ser Pro Cys Thr Thr Thr Arg Asn Thr Ala Cys Gln
145            150            155            160
Cys Lys Pro Gly Thr Phe Arg Asn Asp Asn Ser Ala Glu Met Cys Arg
             165            170            175
Lys Cys Ser Thr Gly Cys Pro Arg Gly Met Val Lys Val Lys Asp Cys
             180            185            190
Thr Pro Trp Ser Asp Ile Glu Cys Val His Lys Glu Ser Gly Asn Gly
             195            200            205
His Asn Ile Trp Val Ile Leu Val Val Thr Leu Val Val Pro Leu Leu
             210            215            220
Leu Val Ala Val Leu Ile Val Cys Cys Cys
225            230

```

<210> SEQ ID NO 14
 <211> LENGTH: 247
 <212> TYPE: PRT
 <213> ORGANISM: Rattus norvegicus

<400> SEQUENCE: 14

```

Met Ser Met Ser Leu Glu Ile Thr Gly Thr Ser Leu Ala Val Leu Gly
 1             5             10             15
Trp Leu Cys Thr Ile Val Cys Cys Ala Leu Pro Met Trp Arg Val Ser
             20             25             30
Ala Phe Ile Gly Ser Ser Ile Ile Thr Ala Gln Ile Thr Trp Glu Gly
             35             40             45
Leu Trp Met Asn Cys Val Gln Ser Thr Gly Gln Met Gln Cys Lys Met
             50             55             60
Tyr Asp Ser Leu Leu Ala Leu Pro Gln Asp Leu Gln Ala Ala Arg Ala
65             70             75             80

```

-continued

Leu Ile Val Val Ser Ile Leu Leu Ala Ala Phe Gly Leu Leu Val Ala
 85 90 95
 Leu Val Gly Ala Gln Cys Thr Asn Cys Val Gln Asp Glu Thr Ala Lys
 100 105 110
 Ala Lys Ile Thr Ile Val Ala Gly Val Leu Phe Leu Leu Ala Ala Val
 115 120 125
 Leu Thr Leu Val Pro Val Ser Trp Ser Ala Asn Thr Ile Ile Arg Asp
 130 135 140
 Phe Tyr Asn Pro Leu Val Pro Glu Ala Gln Lys Arg Glu Met Gly Thr
 145 150 155 160
 Gly Leu Tyr Val Gly Trp Ala Ala Ala Ala Leu Gln Leu Leu Gly Gly
 165 170 175
 Ala Leu Leu Cys Cys Ser Cys Pro Pro Arg Glu Lys Tyr Ala Pro Thr
 180 185 190
 Lys Ile Leu Tyr Ser Ala Pro Arg Ser Thr Gly Pro Gly Thr Gly Thr
 195 200 205
 Gly Thr Ala Tyr Asp Arg Lys Thr Thr Ser Glu Arg Pro Gly Ala Arg
 210 215 220
 Thr Pro His His His His Tyr Gln Pro Ser Met Tyr Pro Thr Arg Pro
 225 230 235 240
 Ala Cys Ser Leu Ala Ser Glu
 245

<210> SEQ ID NO 15

<211> LENGTH: 218

<212> TYPE: PRT

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 15

Met Gly Ser Ala Ala Leu Glu Ile Leu Gly Leu Val Leu Cys Leu Val
 1 5 10 15
 Gly Trp Gly Gly Leu Ile Leu Ala Cys Gly Leu Pro Met Trp Gln Val
 20 25 30
 Thr Ala Phe Leu Asp His Asn Ile Val Thr Ala Gln Thr Thr Trp Lys
 35 40 45
 Gly Leu Trp Met Ser Cys Val Val Gln Ser Thr Gly His Met Gln Cys
 50 55 60
 Lys Val Tyr Asp Ser Val Leu Ala Leu Ser Thr Glu Val Gln Ala Ala
 65 70 75 80
 Arg Ala Leu Thr Val Ser Ala Val Leu Leu Ala Phe Val Ala Leu Phe
 85 90 95
 Val Thr Leu Ala Gly Ala Gln Cys Thr Thr Cys Val Ala Pro Gly Pro
 100 105 110
 Ala Lys Ala Arg Val Ala Leu Thr Gly Gly Val Leu Tyr Leu Phe Cys
 115 120 125
 Gly Leu Leu Ala Leu Val Pro Leu Cys Trp Phe Ala Asn Ile Val Val
 130 135 140
 Arg Glu Phe Tyr Asp Pro Ser Val Pro Val Ser Gln Lys Tyr Glu Leu
 145 150 155 160
 Gly Ala Ala Leu Tyr Ile Gly Trp Ala Ala Thr Ala Leu Leu Met Val
 165 170 175
 Gly Gly Cys Leu Leu Cys Cys Gly Ala Trp Val Cys Thr Gly Arg Pro
 180 185 190

-continued

Asp Leu Ser Phe Pro Val Lys Tyr Ser Ala Pro Arg Arg Pro Thr Ala
195 200 205

Thr Gly Asp Tyr Asp Lys Lys Asn Tyr Val
210 215

<210> SEQ ID NO 16
<211> LENGTH: 551
<212> TYPE: PRT
<213> ORGANISM: Mus musculus

<400> SEQUENCE: 16

Cys Ala Ser Leu Asn Gly Val Ser Gly Asp Ser His Leu Met Ala Ser
1 5 10 15
Met Leu Ser Ser Leu Asp His Ser Gln Pro Trp Ser Pro Cys Ser Ala
20 25 30
Tyr Met Val Thr Ser Phe Leu Asp Asn Gly His Gly Glu Cys Leu Met
35 40 45
Asp Lys Pro Gln Asn Pro Ile Lys Leu Pro Ser Asp Leu Pro Gly Thr
50 55 60
Leu Tyr Asp Ala Asn Arg Gln Cys Gln Phe Thr Phe Gly Glu Glu Ser
65 70 75 80
Lys His Cys Pro Asp Ala Ala Ser Thr Cys Thr Thr Leu Trp Cys Thr
85 90 95
Gly Thr Ser Gly Gly Leu Leu Val Cys Gln Thr Lys His Phe Pro Trp
100 105 110
Ala Asp Gly Thr Ser Cys Gly Glu Gly Lys Trp Cys Val Ser Gly Lys
115 120 125
Cys Val Asn Lys Thr Asp Met Lys His Phe Ala Thr Pro Val His Gly
130 135 140
Ser Trp Gly Pro Trp Gly Pro Trp Gly Asp Cys Ser Arg Thr Cys Gly
145 150 155 160
Gly Gly Val Gln Tyr Thr Met Arg Glu Cys Asp Asn Pro Val Pro Lys
165 170 175
Asn Gly Gly Lys Tyr Cys Glu Gly Lys Arg Val Arg Tyr Arg Ser Cys
180 185 190
Asn Ile Glu Asp Cys Pro Asp Asn Asn Gly Lys Thr Phe Arg Glu Glu
195 200 205
Gln Cys Glu Ala His Asn Glu Phe Ser Lys Ala Ser Phe Gly Asn Glu
210 215 220
Pro Thr Val Glu Trp Thr Pro Lys Tyr Ala Gly Val Ser Pro Lys Asp
225 230 235 240
Arg Cys Lys Leu Thr Cys Glu Ala Lys Gly Ile Gly Tyr Phe Phe Val
245 250 255
Leu Gln Pro Lys Val Val Asp Gly Thr Pro Cys Ser Pro Asp Ser Thr
260 265 270
Ser Val Cys Val Gln Gly Gln Cys Val Lys Ala Gly Cys Asp Arg Ile
275 280 285
Ile Asp Ser Lys Lys Lys Phe Asp Lys Cys Gly Val Cys Gly Gly Asn
290 295 300
Gly Ser Thr Cys Lys Lys Met Ser Gly Ile Val Thr Ser Thr Arg Pro
305 310 315 320
Gly Tyr His Asp Ile Val Thr Ile Pro Ala Gly Ala Thr Asn Ile Glu
325 330 335
Val Lys His Arg Asn Gln Arg Gly Ser Arg Asn Asn Gly Ser Phe Leu
340 345 350
Ala Ile Arg Ala Ala Asp Gly Thr Tyr Ile Leu Asn Gly Asn Phe Thr
355 360 365
Leu Ser Thr Leu Glu Gln Asp Leu Thr Tyr Lys Gly Thr Val Leu Arg
370 375 380
Tyr Ser Gly Ser Ser Ala Ala Leu Glu Arg Ile Arg Ser Phe Ser Pro
385 390 395 400
Leu Lys Glu Pro Leu Thr Ile Gln Val Leu Met Val Gly His Ala Leu
405 410 415
Arg Pro Lys Ile Lys Phe Thr Tyr Phe Met Lys Lys Lys Thr Glu Ser
420 425 430
Phe Asn Ala Ile Pro Thr Phe Ser Glu Trp Val Ile Glu Glu Trp Gly
435 440 445
Glu Cys Ser Lys Thr Cys Gly Ser Gly Trp Gln Arg Arg Val Val Gln
450 455 460
Cys Arg Asp Ile Asn Gly His Pro Ala Ser Glu Cys Ala Lys Glu Val
465 470 475 480
Lys Pro Ala Ser Thr Arg Pro Cys Ala Asp Leu Pro Cys Pro His Trp

-continued

Gln	Val	Gly	Asp	Trp	Ser	Pro	Cys	Ser	Lys	Thr	Cys	Gly	Lys	Gly	Tyr	485	490	495
			500					505					510					
Lys	Lys	Arg	Thr	Leu	Lys	Cys	Val	Ser	His	Asp	Gly	Gly	Val	Leu	Ser			
		515				520						525						
Asn	Glu	Ser	Cys	Asp	Pro	Leu	Lys	Lys	Pro	Lys	His	Tyr	Ile	Asp	Phe			
	530					535					540							
Cys	Thr	Leu	Thr	Gln	Cys	Ser												
545					550													

<210> SEQ ID NO 17
 <211> LENGTH: 1234
 <212> TYPE: DNA
 <213> ORGANISM: Mus musculus
 <220> FEATURE:
 <221> NAME/KEY: CDS
 <222> LOCATION: (187)...(819)

<400> SEQUENCE: 17

ccagactcca ccaccgccta cccggaccag aagccaggag cctcgccccg cagctgcaca	60
gagagcaagg gtataggcac taacttgttt gcagagaccc catcaccttc gggagctcag	120
gtgcgcacct tgcaaaactcc actttctgca tctgccactg agcccgcggg agcctcggaa	180
agagcc atg gcc aac gcg ggg ctg cag ctg ctg ggt ttc atc ctg gct	228
Met Ala Asn Ala Gly Leu Gln Leu Leu Gly Phe Ile Leu Ala	
1 5 10	
tct ctg gga tgg atc ggc tcc atc gtc agc act gcc ctg ccc cag tgg	276
Ser Leu Gly Trp Ile Gly Ser Ile Val Ser Thr Ala Leu Pro Gln Trp	
15 20 25 30	
aag att tac tcc tat gct ggg gac aac atc gtg acc gct cag gcc atc	324
Lys Ile Tyr Ser Tyr Ala Gly Asp Asn Ile Val Thr Ala Gln Ala Ile	
35 40 45	
tac gag gga ctg tgg atg tcc tgc gtt tgc caa agc acc ggg cag ata	372
Tyr Glu Gly Leu Trp Met Ser Cys Val Ser Gln Ser Thr Gly Gln Ile	
50 55 60	
cag tgc aaa gtc ttc gac tcc ttg ctg aat ctg aac agt act ttg cag	420
Gln Cys Lys Val Phe Asp Ser Leu Leu Asn Leu Asn Ser Thr Leu Gln	
65 70 75	
gca acc cga gcc ttg atg gta att ggc atc ctg ctg ggg ctg atc gca	468
Ala Thr Arg Ala Leu Met Val Ile Gly Ile Leu Leu Gly Leu Ile Ala	
80 85 90	
atc ttt gtg tcc acc att ggc atg aag tgc atg agg tgc ctg gaa gat	516
Ile Phe Val Ser Thr Ile Gly Met Lys Cys Met Arg Cys Leu Glu Asp	
95 100 105 110	
gat gag gtg cag aag atg tgg atg gct gtc att ggg ggc ata ata ttt	564
Asp Glu Val Gln Lys Met Trp Met Ala Val Ile Gly Gly Ile Ile Phe	
115 120 125	
tta att tca ggt ctg gcg aca tta gtg gcc aca gca tgg tat gga aac	612
Leu Ile Ser Gly Leu Ala Thr Leu Val Ala Thr Ala Trp Tyr Gly Asn	
130 135 140	
aga att gtt caa gaa ttc tat gac ccc ttg acc ccc atc aat gcc agg	660
Arg Ile Val Gln Glu Phe Tyr Asp Pro Leu Thr Pro Ile Asn Ala Arg	
145 150 155	
tat gaa ttt ggc cag gcc ctc ttt act ggc tgg gcc gct gcc tcc ctc	708
Tyr Glu Phe Gly Gln Ala Leu Phe Thr Gly Trp Ala Ala Ala Ser Leu	
160 165 170	
tgc ctt ctg gga ggt gtc cta ctt tcc tgc tcc tgt ccc cgg aaa aca	756
Cys Leu Leu Gly Gly Val Leu Leu Ser Cys Ser Cys Pro Arg Lys Thr	
175 180 185 190	
acc tct tac cca aca cca cgg cct tat ccc aag cca aca cct tct agt	804

-continued

Thr	Ser	Tyr	Pro	Thr	Pro	Arg	Pro	Tyr	Pro	Lys	Pro	Thr	Pro	Ser	Ser	
				195					200					205		
ggg	aaa	gac	tat	gtg	tgacagaggc	aaaggaagag	atcttctctgg	agcaaataca								859
Gly	Lys	Asp	Tyr	Val												
				210												
aaatggacat	tgaacctagg	attgacatta	acgccttaga	ctgttgatga	tggttatcgg											919
aactgtggta	gaacagaagg	aagcatattt	ttatacatcc	ccatggctat	gcaggccttg											979
gctgtacctt	accatcttcc	ctgagcacag	gaggaagagc	ttttgcctgt	gaactgctgc											1039
ttccctctga	gaaatcacac	tcaaacgggg	ataaggtgct	ccttgcatgt	gtatagatat											1099
gtacagatac	atagtttcta	ttaaaaatag	acaagttaca	aatcccttat	tctcctcata											1159
ctgtaccagc	acacttttaa	tgactctaca	atatatacaa	ttatgttttg	attaaaaaaa											1219
aaaaaaaaaa	aaaaa															1234

<210> SEQ ID NO 18

<211> LENGTH: 211

<212> TYPE: PRT

<213> ORGANISM: Mus musculus

<400> SEQUENCE: 18

Met	Ala	Asn	Ala	Gly	Leu	Gln	Leu	Leu	Gly	Phe	Ile	Leu	Ala	Ser	Leu	
1				5					10					15		
Gly	Trp	Ile	Gly	Ser	Ile	Val	Ser	Thr	Ala	Leu	Pro	Gln	Trp	Lys	Ile	
			20					25					30			
Tyr	Ser	Tyr	Ala	Gly	Asp	Asn	Ile	Val	Thr	Ala	Gln	Ala	Ile	Tyr	Glu	
		35					40					45				
Gly	Leu	Trp	Met	Ser	Cys	Val	Ser	Gln	Ser	Thr	Gly	Gln	Ile	Gln	Cys	
	50					55					60					
Lys	Val	Phe	Asp	Ser	Leu	Leu	Asn	Leu	Asn	Ser	Thr	Leu	Gln	Ala	Thr	
65					70					75					80	
Arg	Ala	Leu	Met	Val	Ile	Gly	Ile	Leu	Leu	Gly	Leu	Ile	Ala	Ile	Phe	
			85					90						95		
Val	Ser	Thr	Ile	Gly	Met	Lys	Cys	Met	Arg	Cys	Leu	Glu	Asp	Asp	Glu	
		100						105					110			
Val	Gln	Lys	Met	Trp	Met	Ala	Val	Ile	Gly	Gly	Ile	Ile	Phe	Leu	Ile	
		115					120					125				
Ser	Gly	Leu	Ala	Thr	Leu	Val	Ala	Thr	Ala	Trp	Tyr	Gly	Asn	Arg	Ile	
		130				135						140				
Val	Gln	Glu	Phe	Tyr	Asp	Pro	Leu	Thr	Pro	Ile	Asn	Ala	Arg	Tyr	Glu	
145						150				155					160	
Phe	Gly	Gln	Ala	Leu	Phe	Thr	Gly	Trp	Ala	Ala	Ala	Ser	Leu	Cys	Leu	
			165						170					175		
Leu	Gly	Gly	Val	Leu	Leu	Ser	Cys	Ser	Cys	Pro	Arg	Lys	Thr	Thr	Ser	
			180					185						190		
Tyr	Pro	Thr	Pro	Arg	Pro	Tyr	Pro	Lys	Pro	Thr	Pro	Ser	Ser	Gly	Lys	
		195						200					205			
Asp	Tyr	Val														
		210														

<210> SEQ ID NO 19

<211> LENGTH: 3552

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

-continued

<220> FEATURE:

<221> NAME/KEY: CDS

<222> LOCATION: (1)...(1803)

<400> SEQUENCE: 19

```

ggg gaa gct tgc cag cag atc tgc agc tgc caa aat ggg gca gac tgt      48
Gly Glu Ala Cys Gln Gln Ile Cys Ser Cys Gln Asn Gly Ala Asp Cys
 1             5             10             15

gac agt gtg act gga aag tgc acc tgt gcc cca gga ttc aaa gga att      96
Asp Ser Val Thr Gly Lys Cys Thr Cys Ala Pro Gly Phe Lys Gly Ile
          20             25             30

gac tgc tct acc cca tgc cct ctg gga acc tat ggg ata aac tgt tcc     144
Asp Cys Ser Thr Pro Cys Pro Leu Gly Thr Tyr Gly Ile Asn Cys Ser
          35             40             45

tct cgc tgt ggc tgt aaa aat gat gca gtc tgc tct cct gtg gac ggg     192
Ser Arg Cys Gly Cys Lys Asn Asp Ala Val Cys Ser Pro Val Asp Gly
          50             55             60

tct tgt act tgc aag gca ggc tgg cac ggg gtg gac tgc tcc atc aga     240
Ser Cys Thr Cys Lys Ala Gly Trp His Gly Val Asp Cys Ser Ile Arg
          65             70             75             80

tgt ccc agt ggc aca tgg ggc ttt ggc tgt aac tta aca tgc cag tgc     288
Cys Pro Ser Gly Thr Trp Gly Phe Gly Cys Asn Leu Thr Cys Gln Cys
          85             90             95

ctc aac ggg gga gcc tgc aac acc ctg gac ggg acc tgc acg tgt gca     336
Leu Asn Gly Gly Ala Cys Asn Thr Leu Asp Gly Thr Cys Thr Cys Ala
          100            105            110

cct gga tgg cgc ggg gag aaa tgc gaa ctt ccc tgc cag gat ggc acg     384
Pro Gly Trp Arg Gly Glu Lys Cys Glu Leu Pro Cys Gln Asp Gly Thr
          115            120            125

tac ggg ctg aac tgt gct gag cgc tgc gac tgc agc cac gca gat ggc     432
Tyr Gly Leu Asn Cys Ala Glu Arg Cys Asp Cys Ser His Ala Asp Gly
          130            135            140

tgc cac cct acc acg ggc cat tgc cgc tgc ctc ccc gga tgg tca ggt     480
Cys His Pro Thr Thr Gly His Cys Arg Cys Leu Pro Gly Trp Ser Gly
          145            150            155            160

gtc cac tgt gac agc gtg tgt gct gag gga cgc tgg ggc ccc aac tgc     528
Val His Cys Asp Ser Val Cys Ala Glu Gly Arg Trp Gly Pro Asn Cys
          165            170            175

tcc ctg ccc tgc tac tgt aaa aat ggg gct tca tgc tcc cct gat gat     576
Ser Leu Pro Cys Tyr Cys Lys Asn Gly Ala Ser Cys Ser Pro Asp Asp
          180            185            190

ggc atc tgc gag tgt gca cca ggc ttc cga ggc acc act tgt cag agg     624
Gly Ile Cys Glu Cys Ala Pro Gly Phe Arg Gly Thr Thr Cys Gln Arg
          195            200            205

atc tgc tcc cct ggt ttt tat ggg cat cgc tgc agc cag aca tgc cca     672
Ile Cys Ser Pro Gly Phe Tyr Gly His Arg Cys Ser Gln Thr Cys Pro
          210            215            220

cag tgc gtt cac agc agc ggg ccc tgc cac cac atc acc ggc ctg tgt     720
Gln Cys Val His Ser Ser Gly Pro Cys His His Ile Thr Gly Leu Cys
          225            230            235            240

gac tgc ttg cct ggc ttc aca ggc gcc ctc tgc aat gaa gtg tgt ccc     768
Asp Cys Leu Pro Gly Phe Thr Gly Ala Leu Cys Asn Glu Val Cys Pro
          245            250            255

agt ggc aga ttt ggg aaa aac tgt gca gga att tgt acc tgc acc aac     816
Ser Gly Arg Phe Gly Lys Asn Cys Ala Gly Ile Cys Thr Cys Thr Asn
          260            265            270

aac gga acc tgt aac ccc att gac aga tct tgt cag tgt tac ccc ggt     864
Asn Gly Thr Cys Asn Pro Ile Asp Arg Ser Cys Gln Cys Tyr Pro Gly

```

-continued

275	280	285	
tgg att ggc agt gac tgc tct caa cca tgt cca cct gcc cac tgg ggc Trp Ile Gly Ser Asp Cys Ser Gln Pro Cys Pro Pro Ala His Trp Gly 290 295 300			912
cca aac tgc atc cac acg tgc aac tgc cat aat gga gct ttc tgc agc Pro Asn Cys Ile His Thr Cys Asn Cys His Asn Gly Ala Phe Cys Ser 305 310 315 320			960
gcc tac gat ggg gaa tgt aaa tgc act cct ggc tgg aca ggg ctc tac Ala Tyr Asp Gly Glu Cys Lys Cys Thr Pro Gly Trp Thr Gly Leu Tyr 325 330 335			1008
tgc act cag aga tgt cct cta ggg ttt tat gga aaa gat tgt gca ctg Cys Thr Gln Arg Cys Pro Leu Gly Phe Tyr Gly Lys Asp Cys Ala Leu 340 345 350			1056
ata tgc caa tgt caa aac gga gct gac tgc gac cac att tct ggg cag Ile Cys Gln Cys Gln Asn Gly Ala Asp Cys Asp His Ile Ser Gly Gln 355 360 365			1104
tgt act tgc cgc act gga ttc atg gga cgg cac tgt gag cag aag tgc Cys Thr Cys Arg Thr Gly Phe Met Gly Arg His Cys Glu Gln Lys Cys 370 375 380			1152
cct tca gga aca tat ggc tat ggc tgt cgc cag ata tgt gat tgt ctg Pro Ser Gly Thr Tyr Gly Tyr Gly Cys Arg Gln Ile Cys Asp Cys Leu 385 390 395 400			1200
aac aac tcc acc tgc gac cac atc act ggg acc tgt tac tgc agc ccc Asn Asn Ser Thr Cys Asp His Ile Thr Gly Thr Cys Tyr Cys Ser Pro 405 410 415			1248
gga tgg aag gga gcg aga tgt gat caa gct ggt gtt atc ata gtt gga Gly Trp Lys Gly Ala Arg Cys Asp Gln Ala Gly Val Ile Ile Val Gly 420 425 430			1296
aat ctg aac agc tta agc cga acc agt act gct ctc cct gct gat tcc Asn Leu Asn Ser Leu Ser Arg Thr Ser Thr Ala Leu Pro Ala Asp Ser 435 440 445			1344
tac cac atc ggg gcc att gca ggc atc atc att ctt gtc cta gtt gtt Tyr His Ile Gly Ala Ile Ala Gly Ile Ile Ile Leu Val Leu Val Val 450 455 460			1392
ctc ttc cta ctg gca ttg ttc att att tat aga cac aag cag aag gga Leu Phe Leu Leu Ala Leu Phe Ile Ile Tyr Arg His Lys Gln Lys Gly 465 470 475 480			1440
aag gaa tca agc atg cca gca gtt acc tac acc cct gct atg agg gtc Lys Glu Ser Ser Met Pro Ala Val Thr Tyr Thr Pro Ala Met Arg Val 485 490 495			1488
gtc aat gca gat tat acc att tca gga acc ctt cct cac agc aat ggt Val Asn Ala Asp Tyr Thr Ile Ser Gly Thr Leu Pro His Ser Asn Gly 500 505 510			1536
gga aac gct aat agc cac tac ttc acc aat ccc agt tac cac acg ctc Gly Asn Ala Asn Ser His Tyr Phe Thr Asn Pro Ser Tyr His Thr Leu 515 520 525			1584
acc cag tgt gcc aca tcc cct cac gtc aac aac agg gac agg atg act Thr Gln Cys Ala Thr Ser Pro His Val Asn Asn Arg Asp Arg Met Thr 530 535 540			1632
gtc acg aag tca aaa aac aat caa ctg ttt gtg aat ctt aaa aat gtg Val Thr Lys Ser Lys Asn Asn Gln Leu Phe Val Asn Leu Lys Asn Val 545 550 555 560			1680
aac cct ggg aag aga ggc cct gtg ggg gac tgc atg gga cat tgc cgg Asn Pro Gly Lys Arg Gly Pro Val Gly Asp Cys Met Gly His Cys Arg 565 570 575			1728
ctg act gga aac atg gcg gct acc tca acg agc tcg gtg ctt ttg gac Leu Thr Gly Asn Met Ala Ala Thr Ser Thr Ser Ser Val Leu Leu Asp			1776

-continued

580	585	590	
ttg aca gaa gct ata tgg gaa aat cct taaaagacct gggaaagaat			1823
Leu Thr Glu Ala Ile Trp Glu Asn Pro			
595	600		
tctgaatata attcaagtaa ctgctcccta agcagttctg agaaccata tgccactatt			1883
aaagaccac ctgtacttat ccgaaaagc tcagagtgtg gttatgtgga gatgaaatcg			1943
ccggcacgaa gagattcccc atatgcagag atcaataact caacttcagc caacaggaat			2003
gtctatgaag ttgaacctac agtgagtgtt gtccaaggag tattcagcaa taatgggcgt			2063
ctctcccag atccatatga cctcccaaag aacagtcaca tcccttgtca ttatgacctg			2123
ctgccagtcc gagacagttc atcctccct aagcaagagg acagtggagg tagcagcagc			2183
aacagcagca gcagcagtga atgacacca aggaccgctt ggtagccact ggaacccttt			2243
ccagaactgc tgtttgttc ttctccatcc tcaattttgc cactttcatg tgaatgttag			2303
tcaattcggg gggcaatttt tggacatgaa ccagaaagct gaaagctgag gctgacacgg			2363
actgtagggt ctttttgttc aggtggattc gaaggagta gagatgtgat ttgccattgc			2423
tgtagttttt agaactatac ccgtgaagca tgacttattg taagatgttg gctgaaagca			2483
tgaacttgca gaactccctc ggagacgcag gttgcagtgg acattgggat tgttgcttga			2543
aaaattaaaa tttgaatatt ttctctctca tttgcatcat acagctctac ctaggattgt			2603
acagtttacc ataaaaatta ctccatgaaa gtgggaatca ctgaacatgt agaagacaag			2663
gaacatattg ttaactcctg attcttaact ttattcaact ggactcagaa ttgtagggat			2723
aatatgaatg caggaggaaa cattctgtca ggcggtatga ctggacagac tttgaatata			2783
ctctaaaagt ggacagaaaa ttacgaaaa tcttagattt tgtttagaat gagaaaatat			2843
acaattagaa ttatttttaga aatagtagga agtattgcag aagtcaatac acaaattgtc			2903
caggcagagg tggttttctc tgtttgactc tcaaccaact tcagatctat gacattattc			2963
tgatcactgg ctccatcata catattcacc acttgagatt cataacatat caatagtatt			3023
ttcataaata tagaaatgaa ataattttat ttttgacaga ctggatggaa tgagtgtgta			3083
atgattgata aaggttgtaa attttaaatg caagatgacg cttacgttct gtaaaccatt			3143
agtaatacat gctgtaatat agaattagt gaacattttg attaattctt ccctagaagt			3203
gactgaaata tttttgtgca tatttgagaa agggaaacttt ccttttatta attgtcaatt			3263
tagagaaact atgcttaagc tggctctttg cattgctaatt gtgacatgta cccaactttt			3323
cattaatttg tatttcatt tttaaattgc atattctatg tttgtagtg tttggattgt			3383
taatgaaaaa atattatatg ttcgttattc cttgtattat tgccacttat cttttgcttg			3443
ataaaaatgc gttgttcttt ttctttttgg agggacaaga tgaaaatata taatttgaat			3503
tgattaaaat tggctcgttac taaaatagta tagtaaaaaa aaaaaaaaaa			3552
<210> SEQ ID NO 20			
<211> LENGTH: 601			
<212> TYPE: PRT			
<213> ORGANISM: Homo sapiens			
<400> SEQUENCE: 20			
Gly Glu Ala Cys Gln Gln Ile Cys Ser Cys Gln Asn Gly Ala Asp Cys			
1	5	10	15

-continued

Asp	Ser	Val	Thr	Gly	Lys	Cys	Thr	Cys	Ala	Pro	Gly	Phe	Lys	Gly	Ile
			20					25					30		
Asp	Cys	Ser	Thr	Pro	Cys	Pro	Leu	Gly	Thr	Tyr	Gly	Ile	Asn	Cys	Ser
		35					40					45			
Ser	Arg	Cys	Gly	Cys	Lys	Asn	Asp	Ala	Val	Cys	Ser	Pro	Val	Asp	Gly
	50					55				60					
Ser	Cys	Thr	Cys	Lys	Ala	Gly	Trp	His	Gly	Val	Asp	Cys	Ser	Ile	Arg
65				70					75					80	
Cys	Pro	Ser	Gly	Thr	Trp	Gly	Phe	Gly	Cys	Asn	Leu	Thr	Cys	Gln	Cys
			85					90						95	
Leu	Asn	Gly	Gly	Ala	Cys	Asn	Thr	Leu	Asp	Gly	Thr	Cys	Thr	Cys	Ala
		100						105					110		
Pro	Gly	Trp	Arg	Gly	Glu	Lys	Cys	Glu	Leu	Pro	Cys	Gln	Asp	Gly	Thr
		115					120					125			
Tyr	Gly	Leu	Asn	Cys	Ala	Glu	Arg	Cys	Asp	Cys	Ser	His	Ala	Asp	Gly
	130					135					140				
Cys	His	Pro	Thr	Thr	Gly	His	Cys	Arg	Cys	Leu	Pro	Gly	Trp	Ser	Gly
145					150					155					160
Val	His	Cys	Asp	Ser	Val	Cys	Ala	Glu	Gly	Arg	Trp	Gly	Pro	Asn	Cys
			165						170					175	
Ser	Leu	Pro	Cys	Tyr	Cys	Lys	Asn	Gly	Ala	Ser	Cys	Ser	Pro	Asp	Asp
		180						185					190		
Gly	Ile	Cys	Glu	Cys	Ala	Pro	Gly	Phe	Arg	Gly	Thr	Thr	Cys	Gln	Arg
	195						200					205			
Ile	Cys	Ser	Pro	Gly	Phe	Tyr	Gly	His	Arg	Cys	Ser	Gln	Thr	Cys	Pro
	210					215					220				
Gln	Cys	Val	His	Ser	Ser	Gly	Pro	Cys	His	His	Ile	Thr	Gly	Leu	Cys
225					230					235					240
Asp	Cys	Leu	Pro	Gly	Phe	Thr	Gly	Ala	Leu	Cys	Asn	Glu	Val	Cys	Pro
			245					250						255	
Ser	Gly	Arg	Phe	Gly	Lys	Asn	Cys	Ala	Gly	Ile	Cys	Thr	Cys	Thr	Asn
		260						265					270		
Asn	Gly	Thr	Cys	Asn	Pro	Ile	Asp	Arg	Ser	Cys	Gln	Cys	Tyr	Pro	Gly
	275						280					285			
Trp	Ile	Gly	Ser	Asp	Cys	Ser	Gln	Pro	Cys	Pro	Pro	Ala	His	Trp	Gly
	290				295						300				
Pro	Asn	Cys	Ile	His	Thr	Cys	Asn	Cys	His	Asn	Gly	Ala	Phe	Cys	Ser
305					310					315					320
Ala	Tyr	Asp	Gly	Glu	Cys	Lys	Cys	Thr	Pro	Gly	Trp	Thr	Gly	Leu	Tyr
			325					330						335	
Cys	Thr	Gln	Arg	Cys	Pro	Leu	Gly	Phe	Tyr	Gly	Lys	Asp	Cys	Ala	Leu
		340						345					350		
Ile	Cys	Gln	Cys	Gln	Asn	Gly	Ala	Asp	Cys	Asp	His	Ile	Ser	Gly	Gln
	355					360						365			
Cys	Thr	Cys	Arg	Thr	Gly	Phe	Met	Gly	Arg	His	Cys	Glu	Gln	Lys	Cys
	370					375					380				
Pro	Ser	Gly	Thr	Tyr	Gly	Tyr	Gly	Cys	Arg	Gln	Ile	Cys	Asp	Cys	Leu
385					390					395					400
Asn	Asn	Ser	Thr	Cys	Asp	His	Ile	Thr	Gly	Thr	Cys	Tyr	Cys	Ser	Pro
			405						410					415	
Gly	Trp	Lys	Gly	Ala	Arg	Cys	Asp	Gln	Ala	Gly	Val	Ile	Ile	Val	Gly

-continued

420							425					430				
Asn	Leu	Asn	Ser	Leu	Ser	Arg	Thr	Ser	Thr	Ala	Leu	Pro	Ala	Asp	Ser	
435							440					445				
Tyr	His	Ile	Gly	Ala	Ile	Ala	Gly	Ile	Ile	Ile	Leu	Val	Leu	Val	Val	
450							455					460				
Leu	Phe	Leu	Leu	Ala	Leu	Phe	Ile	Ile	Tyr	Arg	His	Lys	Gln	Lys	Gly	
465	470							475					480			
Lys	Glu	Ser	Ser	Met	Pro	Ala	Val	Thr	Tyr	Thr	Pro	Ala	Met	Arg	Val	
485							490					495				
Val	Asn	Ala	Asp	Tyr	Thr	Ile	Ser	Gly	Thr	Leu	Pro	His	Ser	Asn	Gly	
500							505					510				
Gly	Asn	Ala	Asn	Ser	His	Tyr	Phe	Thr	Asn	Pro	Ser	Tyr	His	Thr	Leu	
515							520					525				
Thr	Gln	Cys	Ala	Thr	Ser	Pro	His	Val	Asn	Asn	Arg	Asp	Arg	Met	Thr	
530							535					540				
Val	Thr	Lys	Ser	Lys	Asn	Asn	Gln	Leu	Phe	Val	Asn	Leu	Lys	Asn	Val	
545	550							555					560			
Asn	Pro	Gly	Lys	Arg	Gly	Pro	Val	Gly	Asp	Cys	Met	Gly	His	Cys	Arg	
565							570					575				
Leu	Thr	Gly	Asn	Met	Ala	Ala	Thr	Ser	Thr	Ser	Ser	Val	Leu	Leu	Asp	
580							585					590				
Leu	Thr	Glu	Ala	Ile	Trp	Glu	Asn	Pro								
595							600									

What is claimed is:

1. An isolated nucleic acid molecule selected from the group consisting of:

- a nucleic acid molecule having a nucleotide sequence which is at least 90% identical to the nucleotide sequence of any of SEQ ID NOs: 1, 3, 5, 7, 9, 10, 17, and 19, or a complement thereof;
- a nucleic acid molecule comprising at least 15 nucleotide residues and having a nucleotide sequence identical to at least 15 consecutive nucleotide residues of any of SEQ ID NOs: 1, 3, 5, 7, 9, 10, 17, and 19, or a complement thereof;
- a nucleic acid molecule which encodes a polypeptide comprising the amino acid sequence of any of SEQ ID NOs: 2, 4, 6, 8, 18, and 20;
- a nucleic acid molecule which encodes a fragment of a polypeptide comprising the amino acid sequence of any of SEQ ID NOs: 2, 4, 6, 8, 18, and 20, wherein the fragment comprises at least 10 consecutive amino acid residues of any of SEQ ID NOs: 2, 4, 6, 8, 18, and 20; and
- a nucleic acid molecule which encodes a fragment of a polypeptide comprising the amino acid sequence of any of SEQ ID NOs: 2, 4, 6, 8, 18, and 20, wherein the fragment comprises consecutive amino acid residues corresponding to at least half of the full length of any of SEQ ID NOs: 2, 4, 6, 8, 18, and 20
- a nucleic acid molecule which encodes a naturally occurring allelic variant of a polypeptide comprising

the amino acid sequence of any of SEQ ID NOs: 2, 4, 6, 8, 18, and 20, wherein the nucleic acid molecule hybridizes with a nucleic acid molecule consisting of the nucleotide sequence of any of SEQ ID NOs: 1, 3, 5, 7, 9, 10, 17, and 19, or a complement thereof under stringent conditions.

2. The isolated nucleic acid molecule of claim 1, which is selected from the group consisting of:

- a nucleic acid having the nucleotide sequence of any of SEQ ID NOs: 1, 3, 5, 7, 9, 10, 17, and 19, or a complement thereof; and
- a nucleic acid molecule which encodes a polypeptide having the amino acid sequence of any of SEQ ID NOs: 2, 4, 6, 8, 18, and 20, or a complement thereof.

3. The nucleic acid molecule of claim 1, further comprising vector nucleic acid sequences.

4. The nucleic acid molecule of claim 1 further comprising nucleic acid sequences encoding a heterologous polypeptide.

5. A host cell which contains the nucleic acid molecule of claim 1.

6. The host cell of claim 5 which is a mammalian host cell.

7. A non-human mammalian host cell containing the nucleic acid molecule of claim 1.

8. An isolated polypeptide selected from the group consisting of:

- a fragment of a polypeptide comprising the amino acid sequence of any of SEQ ID NOs: 2, 4, 6, 8, 18, and 20;
- a naturally occurring allelic variant of a polypeptide comprising the amino acid sequence of any of SEQ ID

NOs: 2, 4, 6, 8, 18, and 20, wherein the polypeptide is encoded by a nucleic acid molecule which hybridizes with a nucleic acid molecule consisting of the nucleotide sequence of any of SEQ ID NOs: 1, 3, 5, 7, 9, 10, 17, and 19, or a complement thereof under stringent conditions; and

- c) a polypeptide which is encoded by a nucleic acid molecule comprising a nucleotide sequence which is at least 90% identical to a nucleic acid consisting of the nucleotide sequence of any of SEQ ID NOs: 1, 3, 5, 7, 9, 10, 17, and 19, or a complement thereof.

9. The isolated polypeptide of claim 8 having the amino acid sequence of any of SEQ ID NOs: 2, 4, 6, 8, 18, and 20.

10. The polypeptide of claim 8, wherein the amino acid sequence of the polypeptide further comprises heterologous amino acid residues.

11. An antibody which selectively binds with the polypeptide of claim 8.

12. A method for producing a polypeptide selected from the group consisting of:

- a) a polypeptide comprising the amino acid sequence of any of SEQ ID NOs: 2, 4, 6, 8, 18, and 20;
- b) a polypeptide comprising a fragment of the amino acid sequence of any of SEQ ID NOs: 2, 4, 6, 8, 18, and 20, wherein the fragment comprises at least 10 contiguous amino acids of any of SEQ ID NOs: 2, 4, 6, 8, 18, and 20; and
- c) a naturally occurring allelic variant of a polypeptide comprising the amino acid sequence of any of SEQ ID NOs: 2, 4, 6, 8, 18, and 20, or a complement thereof, wherein the polypeptide is encoded by a nucleic acid molecule which hybridizes with a nucleic acid molecule consisting of the nucleotide sequence of any of SEQ ID NOs: 1, 3, 5, 7, 9, 10, 17, and 19, or a complement thereof under stringent conditions;

the method comprising culturing the host cell of claim 5 under conditions in which the nucleic acid molecule is expressed.

13. A method for detecting the presence of a polypeptide of claim 8 in a sample, comprising:

- a) contacting the sample with a compound which selectively binds with a polypeptide of claim 8; and
- b) determining whether the compound binds with the polypeptide in the sample.

14. The method of claim 13, wherein the compound which binds with the polypeptide is an antibody.

15. A kit comprising a compound which selectively binds with a polypeptide of claim 8 and instructions for use.

16. A method for detecting the presence of a nucleic acid molecule of claim 1 in a sample, comprising the steps of:

- a) contacting the sample with a nucleic acid probe or primer which selectively hybridizes with the nucleic acid molecule; and
- b) determining whether the nucleic acid probe or primer binds with a nucleic acid molecule in the sample.

17. The method of claim 16, wherein the sample comprises mRNA molecules and is contacted with a nucleic acid probe.

18. A kit comprising a compound which selectively hybridizes with a nucleic acid molecule of claim 1 and instructions for use.

19. A method for identifying a compound which binds with a polypeptide of claim 8 comprising the steps of:

- a) contacting a polypeptide, or a cell expressing a polypeptide of claim 8 with a test compound; and
- b) determining whether the polypeptide binds with the test compound.

20. The method of claim 19, wherein the binding of the test compound to the polypeptide is detected by a method selected from the group consisting of:

- a) detection of binding by direct detecting of test compound/polypeptide binding;
- b) detection of binding using a competition binding assay;
- c) detection of binding using an assay for an activity characteristic of the polypeptide.

21. A method for modulating the activity of a polypeptide of claim 8 comprising contacting a polypeptide or a cell expressing a polypeptide of claim 8 with a compound which binds with the polypeptide in a sufficient concentration to modulate the activity of the polypeptide.

22. A method for identifying a compound which modulates the activity of a polypeptide of claim 8, comprising:

- a) contacting a polypeptide of claim 8 with a test compound; and
- b) determining the effect of the test compound on the activity of the polypeptide to thereby identify a compound which modulates the activity of the polypeptide.

23. An antibody substance which selectively binds with the polypeptide of claim 8.

24. A method of making an antibody substance which selectively binds with the polypeptide of claim 8, the method comprising providing the polypeptide to an immunocompetent vertebrate and thereafter harvesting from the vertebrate blood or serum comprising the antibody substance.

25. A method of making an antibody substance which selectively binds with the polypeptide of claim 8, the method comprising contacting the polypeptide with a plurality of particles which individually comprise an antibody substance and a nucleic acid encoding the antibody substance, segregating a particle which selectively binds with the polypeptide, and expressing the antibody substance from the nucleic acid of the segregated particle.

26. The isolated nucleic acid of claim 1, wherein the isolated nucleic acid comprises a portion having the nucleotide sequence SEQ ID NO: 1.

27. The isolated nucleic acid of claim 1, wherein the isolated nucleic acid comprises a portion having the nucleotide sequence SEQ ID NO: 3.

28. The isolated nucleic acid of claim 1, wherein the isolated nucleic acid comprises a portion having the nucleotide sequence SEQ ID NO: 5.

29. The isolated nucleic acid of claim 1, wherein the isolated nucleic acid comprises a portion having the nucleotide sequence SEQ ID NO: 7.

30. The isolated nucleic acid of claim 1, wherein the isolated nucleic acid comprises a portion having the nucleotide sequence SEQ ID NO: 9.

31. The isolated nucleic acid of claim 1, wherein the isolated nucleic acid comprises a portion having the nucleotide sequence SEQ ID NO: 10.

32. The isolated nucleic acid of claim 1, wherein the isolated nucleic acid comprises a portion having the nucleotide sequence SEQ ID NO: 17.

33. The isolated nucleic acid of claim 1, wherein the isolated nucleic acid comprises a portion having the nucleotide sequence SEQ ID NO: 19.

34. The isolated polypeptide of claim 8, wherein the amino acid sequence of the isolated polypeptide is SEQ ID NO: 2.

35. The isolated polypeptide of claim 8, wherein the amino acid sequence of the isolated polypeptide is SEQ ID NO: 4.

36. The isolated polypeptide of claim 8, wherein the amino acid sequence of the isolated polypeptide is SEQ ID NO: 6.

37. The isolated polypeptide of claim 8, wherein the amino acid sequence of the isolated polypeptide is SEQ ID NO: 8.

38. The isolated polypeptide of claim 8, wherein the amino acid sequence of the isolated polypeptide is SEQ ID NO: 18.

39. The isolated polypeptide of claim 8, wherein the amino acid sequence of the isolated polypeptide is SEQ ID NO: 20.

* * * * *

专利名称(译)	编码具有预后，诊断，预防，治疗和其他用途的蛋白质的新基因		
公开(公告)号	US20030143632A1	公开(公告)日	2003-07-31
申请号	US10/365227	申请日	2003-02-12
[标]申请(专利权)人(译)	米伦纽姆医药公司		
申请(专利权)人(译)	千年制药，INC.		
当前申请(专利权)人(译)	千年制药，INC.		
[标]发明人	HOLTZMAN DOUGLAS A GOODEARL ANDREW D J MCCARTHY SEAN A		
发明人	HOLTZMAN, DOUGLAS A. GOODEARL, ANDREW D.J. MCCARTHY, SEAN A.		
IPC分类号	A61K38/00 C07K14/47 C07K14/705 C12N9/64 G01N33/53 C07H21/04 C12P21/02 C12N5/06		
CPC分类号	A01K2217/05 A01K2227/105 A61K38/00 C12N9/6489 C07K14/705 C07K14/70578 C07K2319/00 C07K14/47		
优先权	60/054645 1997-08-04 US 60/054966 1997-08-06 US 60/058108 1997-09-05 US		
外部链接	Espacenet USPTO		

摘要(译)

本发明涉及Tango-73, Tango-74, Tango-76, Tango-78和Tango-83多肽, 编码Tango-73, Tango-74, Tango-76, Tango-78和Tango-83的核酸分子。及其用途。本发明提供了编码多种蛋白质的分离的核酸, 所述蛋白质具有诊断, 预防, 治疗和其他用途。这些核酸和蛋白质可用于许多人和其他动物疾病的诊断, 预防和治疗。本发明还提供了反义核酸分子, 含有本发明核酸分子的表达载体, 已导入表达载体的宿主细胞, 以及已引入本发明核酸分子的非人转基因动物或破坏。本发明还进一步提供了分离的多肽, 融合多肽, 抗原肽和抗体。还提供了使用本发明组合物的诊断, 筛选和治疗方法。本发明的核酸和多肽可用作调节多种细胞过程的调节剂。

[illegible]

FIG. 1A