

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
14 June 2001 (14.06.2001)

PCT

(10) International Publication Number  
**WO 01/42289 A2**

- (51) International Patent Classification<sup>7</sup>: **C07K 14/00** [US/US]; 336 West Rock Avenue, New Haven, CT 06515 (US).
- (21) International Application Number: PCT/US00/33441
- (22) International Filing Date: 8 December 2000 (08.12.2000)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:  
60/169,848 9 December 1999 (09.12.1999) US  
Not furnished 6 December 2000 (06.12.2000) US
- (63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:  
US 60/169,848 (CIP)  
Filed on 9 December 1999 (09.12.1999)  
US Not furnished (CIP)  
Filed on 6 December 2000 (06.12.2000)
- (71) Applicant (for all designated States except US): **CURAGEN CORPORATION** [US/US]; 555 Long Wharf Drive, 11th Floor, New Haven, CT 06511 (US).
- (72) Inventor; and
- (75) Inventor/Applicant (for US only): **PEYMAN, John, A.**
- (74) Agent: **ELRIFI, Ivor, R.**; Mintz, Levin, Cohn, Ferris, Glovksy and Popeo, P.C., One Financial Center, Boston, MA 02111 (US).
- (81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).
- Published:**  
— Without international search report and to be republished upon receipt of that report.
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*



**WO 01/42289 A2**

(54) Title: NOVEL POLYPEPTIDES AND POLYNUCLEOTIDES ENCODING SAME

(57) Abstract: The invention provides polypeptides, designated herein as TETRA polypeptides, as well as polynucleotides encoding TETRA polypeptides, and antibodies that immunospecifically-bind to TETRA polypeptide or polynucleotide, or derivatives, variants, mutants, or fragments thereof. The invention additionally provides methods in which the TETRA polypeptide, polynucleotide, and antibody are used in the detection, prevention, and treatment of a broad range of pathological states.

# NOVEL POLYPEPTIDES AND POLYNUCLEOTIDES ENCODING SAME

## BACKGROUND OF THE INVENTION

5           The invention generally relates to nucleic acids and polypeptides encoded therefrom. More specifically, the invention relates to nucleic acids encoding a protein that is a member of the tetraspan superfamily, as well as vectors, host cells, antibodies, and recombinant methods for producing these nucleic acids and polypeptides.

## SUMMARY OF THE INVENTION

10           The invention is based in part upon the discovery of one novel nucleic acid sequence encoding a protein that is a member of the tetraspan superfamily. Nucleic acids encoding these polypeptides and derivatives and fragments thereof, will hereinafter be collectively designated as "TETRA."

15           In one aspect, the invention provides an isolated TETRA nucleic acid molecule encoding a TETRA polypeptide that has identity to a polypeptide sequence from the tetraspan superfamily of membrane proteins. The invention also provides an isolated TETRA5 nucleic acid molecule encoding a TETRA5 polypeptide that is a novel transmembrane polypeptide.

20           In some embodiments, the TETRA nucleic acid molecule can hybridize under stringent conditions to a nucleic acid sequence complementary to a nucleic acid molecule that includes a protein-coding sequence of the nucleic acid sequence. Also included in the invention is an oligonucleotide, e.g., an oligonucleotide which includes at least 6 contiguous nucleotides of a TETRA nucleic acid (e.g., SEQ ID NO:1) or a complement of said oligonucleotide.

25           Also included in the invention are substantially purified TETRA polypeptides (SEQ ID NO:2). The invention also features antibodies that immunoselectively-bind to TETRA polypeptides.

30           In another aspect, the invention includes pharmaceutical compositions which include therapeutically- or prophylactically-effective amounts of a therapeutic and a pharmaceutically-acceptable carrier. The therapeutic can be, e.g., a TETRA nucleic acid, a TETRA polypeptide, or an antibody specific for a TETRA polypeptide. In a further aspect, the invention includes, in one or more containers, a therapeutically- or prophylactically-effective amount of this pharmaceutical composition.

In a further aspect, the invention includes a method of producing a polypeptide by culturing a cell that includes a TETRA nucleic acid, under conditions allowing for expression of the TETRA polypeptide encoded by the DNA. If desired, the TETRA polypeptide can then be recovered.

5 In another aspect, the invention includes a method of detecting the presence of a TETRA polypeptide in a sample. In the method, a sample is contacted with a compound that selectively binds to the polypeptide under conditions allowing for formation of a complex between the polypeptide and the compound. The complex is detected, if present, thereby identifying the TETRA polypeptide within the sample.

10 Also included in the invention is a method of detecting the presence of a TETRA nucleic acid molecule in a sample by contacting the sample with a TETRA nucleic acid probe or primer, and detecting whether the nucleic acid probe or primer bound to a TETRA nucleic acid molecule in the sample.

In a further aspect, the invention provides a method for modulating the activity of a  
15 TETRA polypeptide by contacting a cell sample that includes the TETRA polypeptide with a compound that binds to the TETRA polypeptide in an amount sufficient to modulate the activity of said polypeptide. The compound can be, e.g., a small molecule, such as a nucleic acid, peptide, polypeptide, peptidomimetic, carbohydrate, lipid or other organic (carbon containing) or inorganic molecule, as further described herein.

20 Also within the scope of the invention is the use of a therapeutic in the manufacture of a medicament for treating or preventing disorders or syndromes outlined in the preferred embodiment below. The therapeutic can be, e.g., a TETRA nucleic acid, a TETRA polypeptide, or a TETRA-specific antibody, or biologically-active derivatives or fragments thereof.

25 In the preferred embodiments, the invention further includes methods for screening for a modulator of disorders or syndromes including, e.g., those involving development, differentiation, and activation of thymic immune cells; in pathologies related to spermatogenesis and male infertility; diagnosis of several human neoplasias; in diseases or pathologies of cells in blood circulation such as red blood cells and platelets; various  
30 immunological disorders and/or pathologies; autoimmune and inflammatory diseases; cardiovascular diseases; metabolic diseases; cancer growth and metastasis; viral infections,

cancer therapy, acute lymphoblastic leukemia; in gliomas; neurologic diseases; treatment of cancer; neurodegenerative disorders; Alzheimer's Disease; Parkinson's Disorder; and hematopoietic disorders. The method includes contacting a test compound with a TETRA polypeptide and determining if the test compound binds to said TETRA polypeptide. Binding  
5 of the test compound to the TETRA polypeptide indicates the test compound is a modulator of activity, or of latency or predisposition to the aforementioned disorders or syndromes.

Also within the scope of the invention is a method for screening for a modulator of activity, or of latency or predisposition to a disorders or syndromes listed above by administering a test compound to a test animal at increased risk for the aforementioned  
10 disorders or syndromes. The test animal expresses a recombinant polypeptide encoded by a TETRA nucleic acid. Expression or activity of TETRA polypeptide is then measured in the test animal, as is expression or activity of the protein in a control animal which recombinantly-expresses TETRA polypeptide and is not at increased risk for the disorder or syndrome. Next, the expression of TETRA polypeptide in both the test animal and the control animal is  
15 compared. A change in the activity of TETRA polypeptide in the test animal relative to the control animal indicates the test compound is a modulator of latency of the disorder or syndrome.

In yet another aspect, the invention includes a method for determining the presence of or predisposition to a disease associated with altered levels of a TETRA polypeptide, a  
20 TETRA nucleic acid, or both, in a subject (e.g., a human subject). The method includes measuring the amount of the TETRA polypeptide in a test sample from the subject and comparing the amount of the polypeptide in the test sample to the amount of the TETRA polypeptide present in a control sample. An alteration in the level of the TETRA polypeptide in the test sample as compared to the control sample indicates the presence of or predisposition  
25 to a disease in the subject. Preferably, the predisposition including those listed in the preferred embodiment. above.

In a further aspect, the invention includes a method of treating or preventing a pathological condition associated with a disorder in a mammal by administering to the subject a TETRA polypeptide, a TETRA nucleic acid, or a TETRA -specific antibody to a subject  
30 (e.g., a human subject), in an amount sufficient to alleviate or prevent the pathological condition.

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 the invention belongs. Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of the invention, suitable methods and materials are described  
5 below. All publications, patent applications, patents, and other references mentioned herein are incorporated by reference in their entirety. In case of conflict, the present specification, including definitions, will control. In addition, the materials, methods, and examples are illustrative purposes only, and not intended to be limiting in any manner. Other features and advantages of the invention will be apparent from the following detailed description and  
10 claims.

### DETAILED DESCRIPTION OF THE INVENTION

A TETRA nucleic acid according to the invention includes a nucleic acid encoding a polypeptide related to the tetraspan superfamily of membrane spanning proteins. Members of this protein family are implicated as cell signaling facilitators. An example of this nucleic acid  
15 and its encoded polypeptide is presented in TABLE 1. The disclosed nucleic acid (SEQ ID NO:1) is 1458 nucleotides in length and contains an ORF that begins with an ATG initiation codon at nucleotides 1-3 and ends with a TGA stop codon at nucleotides 661-663.

The representative ORF encodes a 220 amino acid polypeptide (SEQ ID NO:2), which is presented in TABLE 1. The polypeptide includes four hydrophobic putative membrane  
20 spanning segments. The encoded protein is a member of the tetraspan superfamily.

As its name implies the tetraspan superfamily of membrane proteins cross or traverse the cell membrane four times. The members of the tetraspan superfamily show about 25-30% amino acid sequence identity with one another. A number of tetraspan members have been implicated in signal transduction, control of cell adhesion, regulation of cell growth and  
25 proliferation (including development and oncogenesis) and motility (including the ability to suppress metastatic potential) and expression of a number of tetraspan members is associated with a variety of tumors (e.g., CD81/TAPA-1, L6, CD9/MRP-1, L6, CD63/ME491, CO-029, SAS, PETA-3). Indeed the expression of several tetraspan superfamily members is altered when cells are growing or activated. CD9, CD53 and CD82 are upregulated when  
30 lymphocytes are activated. Cell surface expression of CD37 is rapidly lost upon activation of B cells. Other tetraspan superfamily members are implicated in cell growth due to their

association with tumor cells. CD9 (MRP-1) is a marker for 90% of non-T acute lymphoblastic leukemia cells and 50% of acute myeloid and chronic lymphoid leukemias; CD9 is not expressed on resting B and T lymphocytes. Anti-CD9 antibodies inhibit the motility of a variety of cancer cell lines and inhibit the metastatic potential of the mouse BL6 cell line, a highly metastatic variant of B16 cells (Miyake, M. and Hakomori, S. (1991) *Biochem.* 30:3328). Expression of CD9 in transfection experiments correlated with suppression of metastatic potential and cell motility (Ikeyama, S. et al. (1993) *J. Exp. Med.* 174:1347). CD63 (ME491) is expressed in early stage melanoma but is downregulated in advanced stages of melanoma; CD63 is not expressed on normal tissue melanocytes. CO-029 is expressed on colon, gastric, pancreatic and rectal carcinomas but not on most normal tissues. The gene encoding SAS is amplified in a subset of human sarcomas.

The predicted structure of the tetraspan proteins reveals a topology where the N- and C-termini are intracellular and the major hydrophilic domain, located between transmembrane domains 3 and 4, is extracellular. Tetraspan superfamily members are most conserved in their transmembrane and cytoplasmic domains (the conservation among transmembrane domains being the highest) and most divergent in their two hydrophilic extracellular domains. The high level of conservation seen in the transmembrane and cytoplasmic domains of tetraspan members suggest an effector/signaling function common to all members. The divergence of the extracellular domains suggests these domains provide functions specific to each family member such as ligand binding or protein-protein interaction (Wright, M. D. et al. (1993) *Int. Immunol.* 5:209; Fitter, S. et al., (1995) *Blood* 86:1348.)

Biochemical data demonstrates that the tetraspan CD53 is coupled to signal transduction pathways. Anti-CD53 antibodies cause calcium fluxes in human B cells, granulocytes and monocytes along with activation of the monocyte oxidative burst (Olweus, J. et al. (1993) *J. Immunol.* 151:707). In the rat, anti-CD53 antibodies stimulate an inositol phosphate response, protein phosphorylation and an increase in intracellular free calcium; the anti-rat CD53 antibody 7D2 is mitogenic for splenic T cells (Bell, G. M. (1992) *J. Exp. Med.* 175:527). Thus, CD53 is involved in the regulation of cell proliferation and activation.

The discovery of molecules related in the tetraspan superfamily satisfies a need in the art by providing new diagnostic or therapeutic compositions useful in the treatment of disorders associated with alterations in the expression of members of the tetraspan superfamily. Tetraspan proteins are also useful in assays to identify the ligands responsible for

these cell altering events. Tetraspan proteins are useful to detect ligands to promote cell proliferation, motility, and differentiation among various cell types. Also, in some cells they are used to identify signaling complexes linked to plasma membrane proteins. Nucleic acids, polypeptides, antibodies, and other compositions of the present invention will be useful for therapeutic applications in autoimmune and inflammatory diseases, cardiovascular diseases, metabolic diseases, and cancer growth and metastasis.

**TABLE 1.**

**A representative nucleic acid sequence encoding a TETRA polypeptide according to the invention. The start and stop codons are in bold type.**

10  
15  
20  
25  
30

**ATG**AGGAGCAAAGTGACTACAATTTGCAAACATAGTAAATCCTCAACATCATCAAAGATCACCCCTAGTTCTTCCT  
CGGAAGAGTATCAGAAGGTCTGGAACCTCTTTAACCGCACGCTGCCTTTCTACTTTGTTTTCAGAAGATTGAGCGAGT  
ACAGAACCTGGCCCTCTGGGAAGTCTACCCAGTGGCAAAAAGGACAGATGCAGAAGCAGAACGGAGGGAAGGCCGTG  
GACGAGCGGCAGCTGTTCCACGGCACCAGCGCCATTTTTGTGGACGCCATCTGCCAGCAGAACTTTGACTGGCGGG  
TCTGTGGTGTTCATGGCACTTCCACGGCAAGGGGAGCTACTTTGCCCGAGATGCTGCATATTTCCACCCTACAG  
CAAATCCGACACGCAGACCCACACGATGTTCCCTGGCCCGGTGCTGGTGGGCGAGTTTCGTCAGGGGCAATGCCTCC  
TTTGTCCGTCGCCCGGCCAAGGAGGGCTGGAGCAACGCCTTCTATGATAGCTGCGTGAACAGTGTGTCCGACCCCT  
CCATCTTTGTGATCTTTGAGAAACACCAGGTCTACCCAGAGTATGTCATCCAGTACACCACCTCCTCCAAGCCCTC  
GGTCACACCCCTCCATCCTGCTGGCCTTGGGCTCCCTGTTTCCAGCAGCCGACAG**TG**AGCGCACAGGAGTGTTCAGGC  
CTTTCACCTGCTCTGCCTTGAAATGGCTATTTGGGCCCTTCCCTTTTCTTTTAAACAGAAACTTTTAATGAAGTGT  
TCTCTTAACATTGACCTCTCAATGAAGTTATGTTCTTAATCTCTTGCTAATAATGATTTTACTTTTAAGTCACTT  
TTGGGTTCACTAGTGGATTAACCAGAAGTATGTTAGTTGAGTCCAGTTTGGCTTTTAAATAATGTGTTGAAGTTT  
TAGTTTTTACTCTTTGTTGACTTTGCTGCTTATTGGCACCCAGGGACAGAGTTTCTAGATAACAATTTTATGGATTGG  
TTTTAATTTTTATGAGTTTGTCTCTGCAGTGATTGGTTTTCTCAGAGTCTCATGGCATCATAGTTTTTCCAGAATG  
ACACAGTAGCCACCGGTGGATGACAGCCCACGGGCGGCACAGTCACTTCTGCCTGTTGCTCTGACACCAACCCAGG  
CAGCTCTGCTGTGGCTTCTCCTGGGCTCTGGCATTAGTTGGTCTGTGTCACATTGTCAGAACAGGTGGCTGCTGTG  
TGGTGCCATCGAGTCCCTGCTGGTTCCCTTGTCTGGGAGGGTCACCCATGCCCCAAGGAAGTGCATCCACCTGG  
CAGGTGACCTGGAGGAGTAGCTTCCCCGAGGACCCCAAGGCTTGGCCTGTGATTGCGCAAACCCACATTTCTAAG  
CACACTGGACACCCCTTCGAGTGTGGGTTTTAACATCCCTGTGAGATTGAATACTTGTGCCACACATGTCACAAAAG  
AGTATGGAAATAAA (SEQ ID NO:1)

**A representative TETRA polypeptide according to the invention.**

35

MRSKVTTICKHSKSSTSSKITLSSSSSEYQKVVNLFNRTLPHYFVQKIERVQNLALWEVYQWQKGMQKQNGGKAV  
DERQLFHGTSIAIFVDAICQNFDRVCGVHGTSYKGSYFARDAAYSHHYSKSDTQHTMFLARVLVGEFVRGNAS  
FVRPPAKEGWSNAFYDSCVNSVSDPSIFVIFEKHQVYPEYVIQYTTSSKPSVTPSILLALGSLFSSRQ (SEQ ID  
NO:2)

**TABLE 2.**

40 **Sequences and Corresponding SEQ ID Numbers**

TETRA number	clone	SEQ ID number of nucleic acid sequence	SEQ ID number of encoded amino acid sequence	Putative function
1	5.04-y0c0-187.1	1	2	cell surface receptor and signaling

**Homology**

A BLASTX search identified two different proteins (GenBank Accession Numbers 11359916 and 10439185) to which SEQ ID NO:2 shows strong homology in the C-terminal domains. The results of this search are illustrated in Table 3 below.

**5 TABLE 3**

gi|11359916|pir||T46327 hypothetical protein DKFZp434B1813.1 - human (fragment)  
 gi|10439185|dbj|BAB15457.1| unnamed protein product [Homo sapiens]

10	Curall3_PRTSEQ gi 11359916 pir  T46327 hypoth gi 10439185 dbj BAB15457.1  un Consensus	<pre>           10      20      30      40      50      60            ... ... ... ... ... ... ... ... ... ... ...            -----           KFLRAGKNCRNshslTTEhNLSVLRthGVDhLSYNELCQLLFQNDPwLLPEICQHYNKGD           -----           </pre>
15		
20	Curall3_PRTSEQ gi 11359916 pir  T46327 hypoth gi 10439185 dbj BAB15457.1  un Consensus	<pre>           70      80      90      100     110     120            ... ... ... ... ... ... ... ... ... ... ...            -----           GPHGSCAFQKQCIKLhICQYfLQGECKfGTsCKRshDFsNsENLEKLEKLGMSsDLVSRl           -----           </pre>
25		
30	Curall3_PRTSEQ gi 11359916 pir  T46327 hypoth gi 10439185 dbj BAB15457.1  un Consensus	<pre>           130     140     150     160     170     180            ... ... ... ... ... ... ... ... ... ... ...            -----           PTiYRNahDIKkSSAPSRVpPFLVpPQGTsERkDSSGSVSPnTLsQEeGDQICLYHIRKS           -----           </pre>
35		
40	Curall3_PRTSEQ gi 11359916 pir  T46327 hypoth gi 10439185 dbj BAB15457.1  un Consensus	<pre>           190     200     210     220     230     240            ... ... ... ... ... ... ... ... ... ... ...            -----           CSfQDKChRVhFLpYrWQfLDRGkWEDLDNMELIEeAYcNPKIErILcSESAsTFHShC           -----           </pre>
45		
50	Curall3_PRTSEQ gi 11359916 pir  T46327 hypoth gi 10439185 dbj BAB15457.1  un Consensus	<pre>           250     260     270     280     290     300            ... ... ... ... ... ... ... ... ... ... ...            -----           LNFnamTYGATQARRLsTASSVTKpPHfILtTDWIWYWSDFGSwQEYGRQGTvHPVTTV           -----           </pre>
55		
60	Curall3_PRTSEQ gi 11359916 pir  T46327 hypoth gi 10439185 dbj BAB15457.1  un Consensus	<pre>           310     320     330     340     350     360            ... ... ... ... ... ... ... ... ... ... ...            -----           SSSDVEKAYLAYCTPGSDQAAALkFQAGkHNYELDFKAFVQKNLVYGTtTKVCRrPKYV           -----           MRSKVTtTCKhSKSS           MGGFGQHtTKKVCRRrPKCV           C K           </pre>
65		
70	Curall3_PRTSEQ gi 11359916 pir  T46327 hypoth gi 10439185 dbj BAB15457.1  un Consensus	<pre>           370     380     390     400     410     420            ... ... ... ... ... ... ... ... ... ... ...            -----           tSS-----KITLSSsSEEYQKvWNLfNRTLPf           SPQDVtTMQtCNTkFPgPKSIpDYWDSSALpDPGFQKITLSSsSEEYQKvWNLfNRTLPf           SPQDVtTMQtCNTkFPgPKSIpDYWDSSALpDPGFQKITLSSsSEEYQKvWNLfNRTLPf           -----           KITLSSsSEEYQKvWNLfNRTLPf           </pre>
75		
80	Curall3_PRTSEQ gi 11359916 pir  T46327 hypoth gi 10439185 dbj BAB15457.1  un Consensus	<pre>           430     440     450     460     470     480            ... ... ... ... ... ... ... ... ... ... ...            -----           YfVQKIeRVQNLALWEVYQWQkGQMQKQNGGkAVDERQLFHGTsAIFVDAICQqNFdWRV           YfVQKIeRVQNLALWEVYQWQkGQMQKQNGGkAVDERQLFHGTsAIFVDAICQqNFdWRV           YfVQKIeRVQNLALWEVYQWQkGQMQKQNGGkAVDERQLFHGTsAIFVDAICQqNFdWRV           YfVQKIeRVQNLALWEVYQWQkGQMQKQNGGkAVDERQLFHGTsAIFVDAICQqNFdWRV           -----           YfVQKIeRVQNLALWEVYQWQkGQMQKQNGGkAVDERQLFHGTsAIFVDAICQqNFdWRV           </pre>



```

                    490      500      510      520      530      540
5  Cural13_PRTSEQ    CGVHGTSYGKGSYFARDAAYSHHYSKSDTQHTMFLARVLVGEFVVRGNASFVVRPPAKEGW
   gi|11359916|pir||T46327 hypoth CGVHGTSYGKGSYFARDAAYSHHYSKSDTQHTMFLARVLVGEFVVRGNASFVVRPPAKEGW
   gi|10439185|dbj|BAB15457.1| un  CGVHGTSYGKGSYFARDAAYSHHYSKSDTQHTMFLARVLVGEFVVRGNASFVVRPPAKEGW
   Consensus          CGVHGTSYGKGSYFARDAAYSHHYSKSDTQHTMFLARVLVGEFVVRGNASFVVRPPAKEGW

                    550      560      570      580      590
10 Cural13_PRTSEQ    SNAFYDSCVNSVSDPSIFVIFEKHQVYPEYVIQYTTSSKPSVTPSILLALGSLFSSRQ
   gi|11359916|pir||T46327 hypoth SNAFYDSCVNSVSDPSIFVIFEKHQVYPEYVIQYTTSSKPSVTPSILLALGSLFSSRQ
   gi|10439185|dbj|BAB15457.1| un  SNAFYDSCVNSVSDPSIFVIFEKHQVYPEYVIQYTTSSKPSVTPSILLALGSLFSSRQ
   Consensus          SNAFYDSCVNSVSDPSIFVIFEKHQVYPEYVIQYTTSSKPSVTPSILLALGSLFSSRQ
15

```

For the aligned region, there is 92% identity or homology (*i.e.*, 180 amino acids match out of a possible 194).

```

20  pir||T46327 hypothetical protein DKFZp434B1813.1 - human (fragment)
   emb|CAB70657.1| (AL137255) hypothetical protein [Homo sapiens]
       Length = 598

   Score = 372 bits (945), Expect = e-102
   Identities = 180/194 (92%), Positives = 180/194 (92%)*
25

```

```

30  >dbj|BAB15457.1| (AK026346) unnamed protein product [Homo sapiens]
       Length = 258

   Score = 372 bits (945), Expect = e-102

   *Identities = 180/194 (92%), Positives = 180/194 (92%)

```

**TETRA Nucleic Acids and Polypeptides**

35 One aspect of the invention pertains to isolated nucleic acid molecules that encode TETRA polypeptides or biologically-active portions thereof. Also included in the invention are nucleic acid fragments sufficient for use as hybridization probes to identify TETRA-encoding nucleic acids (e.g., TETRA mRNAs) and fragments for use as PCR primers for the amplification and/or mutation of TETRA nucleic acid molecules. As used herein, the term  
40 “nucleic acid molecule” is intended to include DNA molecules (e.g., cDNA or genomic DNA), RNA molecules (e.g., mRNA), analogs of the DNA or RNA generated using nucleotide analogs, and derivatives, fragments and homologs thereof. The nucleic acid molecule can be single-stranded or double-stranded, but preferably is comprised double-stranded DNA.

45 A TETRA nucleic acid can encode a mature TETRA polypeptide. As used herein, a “mature” form of a polypeptide or protein disclosed in the present invention is the product of a naturally occurring polypeptide or precursor form or proprotein. The naturally occurring

polypeptide, precursor or proprotein includes, by way of nonlimiting example, the full length gene product, encoded by the corresponding gene. Alternatively, it can be defined as the polypeptide, precursor or proprotein encoded by an open reading frame described herein. The product "mature" form arises, again by way of nonlimiting example, as a result of one or more naturally occurring processing steps as they may take place within the cell, or host cell, in which the gene product arises. Examples of such processing steps leading to a "mature" form of a polypeptide or protein include the cleavage of the N-terminal methionine residue encoded by the initiation codon of an open reading frame, or the proteolytic cleavage of a signal peptide or leader sequence. Thus a mature form arising from a precursor polypeptide or protein that has residues 1 to N, where residue 1 is the N-terminal methionine, would have residues 2 through N remaining after removal of the N-terminal methionine. Alternatively, a mature form arising from a precursor polypeptide or protein having residues 1 to N, in which an N-terminal signal sequence from residue 1 to residue M is cleaved, would have the residues from residue M+1 to residue N remaining. Further as used herein, a "mature" form of a polypeptide or protein may arise from a step of post-translational modification other than a proteolytic cleavage event. Such additional processes include, by way of non-limiting example, glycosylation, myristoylation or phosphorylation. In general, a mature polypeptide or protein may result from the operation of only one of these processes, or a combination of any of them.

The term "probes", as utilized herein, refers to nucleic acid sequences of variable length, preferably between at least about 10 nucleotides (nt), 100 nt, or as many as approximately, e.g., 6,000 nt, depending upon the specific use. Probes are used in the detection of identical, similar, or complementary nucleic acid sequences. Longer length probes are generally obtained from a natural or recombinant source, are highly specific, and much slower to hybridize than shorter-length oligomer probes. Probes can be single- or double-stranded and designed to have specificity in PCR, membrane-based hybridization technologies, or ELISA-like technologies.

The term "isolated" nucleic acid molecule, as utilized herein, is one which is separated from other nucleic acid molecules which are present in the natural source of the nucleic acid. Preferably, an "isolated" nucleic acid is free of sequences which naturally flank the nucleic acid (i.e., sequences located at the 5'- and 3'-termini of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. For example, in various embodiments, the isolated TETRA nucleic acid molecules can contain less than about 5 kb, 4 kb, 3 kb, 2 kb,

1 kb, 0.5 kb or 0.1 kb of nucleotide sequences which naturally flank the nucleic acid molecule in genomic DNA of the cell/tissue from which the nucleic acid is derived (e.g., brain, heart, liver, spleen, etc.). Moreover, an "isolated" nucleic acid molecule, such as a cDNA molecule, can be substantially free of other cellular material or culture medium when produced by recombinant techniques, or of chemical precursors or other chemicals when chemically synthesized.

A nucleic acid molecule of the invention, e.g., a nucleic acid molecule having the nucleotide sequence of SEQ ID NO:1, or a complement of this aforementioned nucleotide sequence, can be isolated using standard molecular biology techniques and the sequence information provided herein. Using all or a portion of the nucleic acid sequence of SEQ ID NO: 1, as a hybridization probe, TETRA molecules can be isolated using standard hybridization and cloning techniques (e.g., as described in Sambrook, et al., (eds.), MOLECULAR CLONING: A LABORATORY MANUAL 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989; and Ausubel, et al., (eds.), CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York, NY, 1993.)

A nucleic acid of the invention can be amplified using cDNA, mRNA or alternatively, genomic DNA, as a template and appropriate oligonucleotide primers according to standard PCR amplification techniques. The nucleic acid so amplified can be cloned into an appropriate vector and characterized by DNA sequence analysis. Furthermore, oligonucleotides corresponding to TETRA nucleotide sequences can be prepared by standard synthetic techniques, e.g., using an automated DNA synthesizer.

As used herein, the term "oligonucleotide" refers to a series of linked nucleotide residues, which oligonucleotide has a sufficient number of nucleotide bases to be used in a PCR reaction. A short oligonucleotide sequence can be based on, or designed from, a genomic or cDNA sequence and is used to amplify, confirm, or reveal the presence of an identical, similar or complementary DNA or RNA in a particular cell or tissue. Oligonucleotides comprise portions of a nucleic acid sequence having about 10 nt, 50 nt, or 100 nt in length, preferably about 15 nt to 30 nt in length. In one embodiment of the invention, an oligonucleotide comprising a nucleic acid molecule less than 100 nt in length would further comprise at least 6 contiguous nucleotides of SEQ ID NO: 1 or a complement thereof. Oligonucleotides can be chemically synthesized and may also be used as probes.

In another embodiment, an isolated nucleic acid molecule of the invention comprises a nucleic acid molecule that is a complement of the nucleotide sequence shown in SEQ ID NO: 1, or a portion of this nucleotide sequence (e.g., a fragment that can be used as a probe or primer or a fragment encoding a biologically-active portion of a TETRA polypeptide). A  
5 nucleic acid molecule that is complementary to the nucleotide sequence shown in SEQ ID NO: 1 is one that is sufficiently complementary to the nucleotide sequence shown in SEQ ID NO: 1 that it can hydrogen bond with little or no mismatches to the nucleotide sequence shown in SEQ ID NO: 1, thereby forming a stable duplex.

As used herein, the term “complementary” refers to Watson-Crick or Hoogsteen base  
10 pairing between nucleotides units of a nucleic acid molecule, and the term “binding” means the physical or chemical interaction between two polypeptides or compounds or associated polypeptides or compounds or combinations thereof. Binding includes ionic, non-ionic, van der Waals, hydrophobic interactions, and the like. A physical interaction can be either direct or indirect. Indirect interactions can be through or due to the effects of another polypeptide or  
15 compound. Direct binding refers to interactions that do not take place through, or due to, the effect of another polypeptide or compound, but instead are without other substantial chemical intermediates.

Fragments provided herein are defined as sequences of at least 6 (contiguous) nucleic acids or at least 4 (contiguous) amino acids, a length sufficient to allow for specific  
20 hybridization in the case of nucleic acids or for specific recognition of an epitope in the case of amino acids, respectively, and are at most some portion less than a full length sequence. Fragments can be derived from any contiguous portion of a nucleic acid or amino acid sequence of choice.

Derivatives are nucleic acid sequences or amino acid sequences formed from the native  
25 compounds either directly or by modification or partial substitution. Analogs are nucleic acid sequences or amino acid sequences that have a structure similar to, but not identical to, the native compound but differs from it in respect to certain components or side chains. Analogs can be synthetic or from a different evolutionary origin and may have a similar or opposite metabolic activity compared to wild type. Homologs are nucleic acid sequences or amino acid  
30 sequences of a particular gene that are derived from different species.

Derivatives and analogs can be full length or other than full length, if the derivative or analog contains a modified nucleic acid or amino acid, as described below. Derivatives or analogs of the nucleic acids or proteins of the invention include, but are not limited to, molecules comprising regions that are substantially homologous to the nucleic acids or proteins of the invention, in various embodiments, by at least about 30%, 50%, 70%, 80%, or 95% identity (with a preferred identity of 80-95%) over a nucleic acid or amino acid sequence of identical size or when compared to an aligned sequence in which the alignment is done by a computer homology program known in the art, or whose encoding nucleic acid is capable of hybridizing to the complement of a sequence encoding the aforementioned proteins under stringent, moderately stringent, or low stringent conditions. See e.g. Ausubel, et al., CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York, NY, 1993, and below.

A "homologous nucleic acid sequence" or "homologous amino acid sequence," or variations thereof, refer to sequences characterized by a homology at the nucleotide level or amino acid level as discussed above. Homologous nucleotide sequences encode those sequences coding for isoforms of TETRA polypeptides. Isoforms can be expressed in different tissues of the same organism as a result of, for example, alternative splicing of RNA. Alternatively, isoforms can be encoded by different genes. In the invention, homologous nucleotide sequences include nucleotide sequences encoding for a TETRA polypeptide of species other than humans, including, but not limited to: vertebrates, and thus can include, e.g., frog, mouse, rat, rabbit, dog, cat, cow, horse, and other organisms. Homologous nucleotide sequences also include, but are not limited to, naturally occurring allelic variations and mutations of the nucleotide sequences set forth herein. A homologous nucleotide sequence does not, however, include the exact nucleotide sequence encoding human TETRA protein. Homologous nucleic acid sequences include those nucleic acid sequences that encode conservative amino acid substitutions (see below) in SEQ ID NO:2, as well as a polypeptide possessing TETRA biological activity. Various biological activities of the TETRA proteins are described below.

A TETRA polypeptide is encoded by the open reading frame ("ORF") of a TETRA nucleic acid. The invention includes the nucleic acid sequence comprising the stretch of nucleic acid sequences of SEQ ID NO: 1, that comprises the ORF of that nucleic acid sequence and encodes a polypeptide of SEQ ID NO: 2. The ORF can include the ORFs described above for the disclosed TETRA sequences, e.g., the ORF beginning at nucleotides

1-3 and ending at nucleotides 661-663 for the disclosed TETRA nucleotide sequence (SEQ ID NO:1).

An ORF corresponds to a nucleotide sequence that could potentially be translated into a polypeptide. A stretch of nucleic acids comprising an ORF is uninterrupted by a stop codon.

5 An ORF that represents the coding sequence for a full protein begins with an ATG "start" codon and terminates with one of the three "stop" codons, namely, TAA, TAG, or TGA. For the purposes of this invention, an ORF can be any part of a coding sequence, with or without a start codon, a stop codon, or both. For an ORF to be considered as a good candidate for coding for a *bona fide* cellular protein, a minimum size requirement is often set, e.g., a stretch  
10 of DNA that would encode a protein of 50 amino acids or more.

The nucleotide sequences determined from the cloning of the TETRA genes allows for the generation of probes and primers designed for use in identifying and/or cloning TETRA homologues in other cell types, e.g. from other tissues, as well as TETRA homologues from other vertebrates. The probe/primer typically comprises substantially purified oligonucleotide.

15 The oligonucleotide typically comprises a region of nucleotide sequence that hybridizes under stringent conditions to at least about 12, 25, 50, 100, 150, 200, 250, 300, 350 or 400 consecutive sense strand nucleotide sequence of SEQ ID NO:1; or an anti-sense strand nucleotide sequence of SEQ ID NO:1; or of a naturally occurring mutant of SEQ ID NO:1.

Probes based on the TETRA nucleotide sequences can be used to detect transcripts or  
20 genomic sequences encoding the same or homologous proteins. In various embodiments, the probe further comprises a label group attached thereto, e.g. the label group can be a radioisotope, a fluorescent compound, an enzyme, or an enzyme co-factor. Such probes can be used as a part of a diagnostic test kit for identifying cells or tissues which mis-express a TETRA protein, such as by measuring a level of a TETRA-encoding nucleic acid in a sample  
25 of cells from a subject e.g., detecting TETRA mRNA levels or determining whether a genomic TETRA gene has been mutated or deleted.

"A polypeptide having a biologically-active portion of a TETRA polypeptide" refers to polypeptides exhibiting activity similar, but not necessarily identical to, an activity of a polypeptide of the invention, including mature forms, as measured in a particular biological  
30 assay, with or without dose dependency. A nucleic acid fragment encoding a "biologically-active portion of TETRA" can be prepared by isolating a portion of SEQ ID NO:1, that

encodes a polypeptide having a TETRA biological activity (the biological activities of the TETRA proteins are described below), expressing the encoded portion of TETRA protein (e.g., by recombinant expression in vitro) and assessing the activity of the encoded portion of TETRA.

## 5 TETRA Nucleic Acid and Polypeptide Variants

The invention further encompasses nucleic acid molecules that differ from the nucleotide sequences shown in SEQ ID NO:1, due to degeneracy of the genetic code and thus encode the same TETRA proteins as that encoded by the nucleotide sequences shown in SEQ ID NO:1. In another embodiment, an isolated nucleic acid molecule of the invention has a  
10 nucleotide sequence encoding a protein having an amino acid sequence shown in SEQ ID NO:2.

In addition to the TETRA nucleotide sequences shown in SEQ ID NO:1, it will be appreciated by those skilled in the art that DNA sequence polymorphisms that lead to changes in the amino acid sequences of the TETRA polypeptides may exist within a population (e.g.,  
15 the human population). Such genetic polymorphism in the TETRA genes may exist among individuals within a population due to natural allelic variation. As used herein, the terms "gene" and "recombinant gene" refer to nucleic acid molecules comprising an open reading frame (ORF) encoding a TETRA protein, preferably a vertebrate TETRA protein. Such natural allelic variations can typically result in 1-5% variance in the nucleotide sequence of the  
20 TETRA genes. Any and all such nucleotide variations and resulting amino acid polymorphisms in the TETRA polypeptides, which are the result of natural allelic variation and that do not alter the functional activity of the TETRA polypeptides, are intended to be within the scope of the invention.

Moreover, nucleic acid molecules encoding TETRA proteins from other species, and  
25 thus that have a nucleotide sequence that differs from the human sequence of SEQ ID NO:1, are intended to be within the scope of the invention. Nucleic acid molecules corresponding to natural allelic variants and homologues of the TETRA cDNAs of the invention can be isolated based on their homology to the human TETRA nucleic acids disclosed herein using the human cDNAs, or a portion thereof, as a hybridization probe according to standard  
30 hybridization techniques under stringent hybridization conditions.

Accordingly, in another embodiment, an isolated nucleic acid molecule of the invention is at least 6 nucleotides in length and hybridizes under stringent conditions to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1. In another embodiment, the nucleic acid is at least 10, 25, 50, 100, 250, 500, 750, 1000, 1500, or 2000 or  
5 more nucleotides in length. In yet another embodiment, an isolated nucleic acid molecule of the invention hybridizes to the coding region. As used herein, the term "hybridizes under stringent conditions" is intended to describe conditions for hybridization and washing under which nucleotide sequences at least 60% homologous to each other typically remain hybridized to each other.

10 Homologs (i.e., nucleic acids encoding TETRA proteins derived from species other than human) or other related sequences (e.g., paralogs) can be obtained by low, moderate or high stringency hybridization with all or a portion of the particular human sequence as a probe using methods well known in the art for nucleic acid hybridization and cloning.

As used herein, the phrase "stringent hybridization conditions" refers to conditions  
15 under which a probe, primer or oligonucleotide will hybridize to its target sequence, but to no other sequences. Stringent conditions are sequence-dependent and will be different in different circumstances. Longer sequences hybridize specifically at higher temperatures than shorter sequences. Generally, stringent conditions are selected to be about 5°C lower than the thermal melting point ( $T_m$ ) for the specific sequence at a defined ionic strength and pH. The  
20  $T_m$  is the temperature (under defined ionic strength, pH and nucleic acid concentration) at which 50% of the probes complementary to the target sequence hybridize to the target sequence at equilibrium. Since the target sequences are generally present at excess, at  $T_m$ , 50% of the probes are occupied at equilibrium. Typically, stringent conditions will be those in which the salt concentration is less than about 1.0 M sodium ion, typically about 0.01 to 1.0 M  
25 sodium ion (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30°C for short probes, primers or oligonucleotides (e.g., 10 nt to 50 nt) and at least about 60°C for longer probes, primers and oligonucleotides. Stringent conditions may also be achieved with the addition of destabilizing agents, such as formamide.

Stringent conditions are known to those skilled in the art and can be found in Ausubel,  
30 et al., (eds.), CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. Preferably, the conditions are such that sequences at least about 65%, 70%, 75%, 85%, 90%, 95%, 98%, or 99% homologous to each other typically remain



hybridized to each other. A non-limiting example of stringent hybridization conditions are hybridization in a high salt buffer comprising 6X SSC, 50 mM Tris-HCl (pH 7.5), 1 mM EDTA, 0.02% PVP, 0.02% Ficoll, 0.02% BSA, and 500 mg/ml denatured salmon sperm DNA at 65°C, followed by one or more washes in 0.2X SSC, 0.01% BSA at 50°C. An isolated  
5 nucleic acid molecule of the invention that hybridizes under stringent conditions to the sequences of SEQ ID NO:1, corresponds to a naturally-occurring nucleic acid molecule. As used herein, a "naturally-occurring" nucleic acid molecule refers to an RNA or DNA molecule having a nucleotide sequence that occurs in nature (e.g., encodes a natural protein).

In a second embodiment, a nucleic acid sequence that can hybridize to the nucleic acid  
10 molecule comprising the nucleotide sequence of SEQ ID NO:1, or fragments, analogs or derivatives thereof, under conditions of moderate stringency is provided. A non-limiting example of moderate stringency hybridization conditions are hybridization in 6X SSC, 5X Denhardt's solution, 0.5% SDS and 100 mg/ml denatured salmon sperm DNA at 55°C, followed by one or more washes in 1X SSC, 0.1% SDS at 37°C. Other conditions of moderate  
15 stringency that can be used are well-known within the art. See, e.g., Ausubel, et al. (eds.), 1993, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, NY, and Kriegler, 1990; GENE TRANSFER AND EXPRESSION, A LABORATORY MANUAL, Stockton Press, NY.

In a third embodiment, a nucleic acid that can hybridize to the nucleic acid molecule  
20 comprising the nucleotide sequences of SEQ ID NO:1 or fragments, analogs or derivatives thereof, under conditions of low stringency, is provided. A non-limiting example of low stringency hybridization conditions are hybridization in 35% formamide, 5X SSC, 50 mM Tris-HCl (pH 7.5), 5 mM EDTA, 0.02% PVP, 0.02% Ficoll, 0.2% BSA, 100 mg/ml denatured salmon sperm DNA, 10% (wt/vol) dextran sulfate at 40°C, followed by one or more washes in  
25 2X SSC, 25 mM Tris-HCl (pH 7.4), 5 mM EDTA, and 0.1% SDS at 50°C. Other conditions of low stringency that can be used are well known in the art (e.g., as employed for cross-species hybridizations). See, e.g., Ausubel, et al. (eds.), 1993, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, NY, and Kriegler, 1990, GENE TRANSFER AND EXPRESSION, A LABORATORY MANUAL, Stockton Press, NY; Shilo and  
30 Weinberg, 1981. Proc Natl Acad Sci USA 78: 6789-6792.

### **Conservative Mutations**

In addition to naturally-occurring allelic variants of TETRA sequences that may exist in the population, the skilled artisan will further appreciate that changes can be introduced by mutation into the nucleotide sequences of SEQ ID NO:1, thereby leading to changes in the amino acid sequences of the encoded TETRA proteins, without altering the functional ability of said TETRA proteins. For example, nucleotide substitutions leading to amino acid substitutions at "non-essential" amino acid residues can be made in the sequence of SEQ ID NO:2. A "non-essential" amino acid residue is a residue that can be altered from the wild-type sequences of the TETRA proteins without altering their biological activity, whereas an "essential" amino acid residue is required for such biological activity. For example, amino acid residues that are conserved among the TETRA proteins of the invention are predicted to be particularly non-amenable to alteration. Amino acids for which conservative substitutions can be made are well-known within the art.

Another aspect of the invention pertains to nucleic acid molecules encoding TETRA proteins that contain changes in amino acid residues that are not essential for activity. Such TETRA proteins differ in amino acid sequence from SEQ ID NO:2, yet retain biological activity. In one embodiment, the isolated nucleic acid molecule comprises a nucleotide sequence encoding a protein, wherein the protein comprises an amino acid sequence at least about 45% homologous to the amino acid sequences of SEQ ID NO:2. Preferably, the protein encoded by the nucleic acid molecule is at least about 60% homologous to SEQ ID NO:2; more preferably at least about 70% homologous to SEQ ID NO:2; still more preferably at least about 80% homologous to SEQ ID NO:2; even more preferably at least about 90% homologous to SEQ ID NO:2; and most preferably at least about 95% homologous to SEQ ID NO:2.

An isolated nucleic acid molecule encoding a TETRA protein homologous to the protein of SEQ ID NO:2, can be created by introducing one or more nucleotide substitutions, additions or deletions into the nucleotide sequence of SEQ ID NO:1, such that one or more amino acid substitutions, additions or deletions are introduced into the encoded protein.

Mutations can be introduced into SEQ ID NO:2, by standard techniques, such as site-directed mutagenesis and PCR-mediated mutagenesis. Preferably, conservative amino acid substitutions are made at one or more predicted, non-essential amino acid residues. A "conservative amino acid substitution" is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having

similar side chains have been defined within the art. These families include amino acids with basic side chains (e.g., lysine, arginine, histidine), acidic side chains (e.g., aspartic acid, glutamic acid), uncharged polar side chains (e.g., glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (e.g., alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (e.g., threonine, valine, isoleucine) and aromatic side chains (e.g., tyrosine, phenylalanine, tryptophan, histidine). Thus, a predicted non-essential amino acid residue in the TETRA protein is replaced with another amino acid residue from the same side chain family. Alternatively, in another embodiment, mutations can be introduced randomly along all or part of a TETRA coding sequence, such as by saturation mutagenesis, and the resultant mutants can be screened for TETRA biological activity to identify mutants that retain activity. Following mutagenesis of SEQ ID NO:2, the encoded protein can be expressed by any recombinant technology known in the art and the activity of the protein can be determined.

In one embodiment, a mutant TETRA protein can be assayed for (i) the ability to form protein:protein interactions with other TETRA proteins, other cell-surface proteins, or biologically-active portions thereof, (ii) complex formation between a mutant TETRA protein and a TETRA ligand; or (iii) the ability of a mutant TETRA protein to bind to an intracellular target protein or biologically-active portion thereof; (e.g. avidin proteins). In yet another embodiment, a mutant TETRA protein can be assayed for the ability to regulate a specific biological function (e.g., regulation of insulin release).

### **Antisense Nucleic Acids**

Another aspect of the invention pertains to isolated antisense nucleic acid molecules that can hybridize to or complementary to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1, or fragments, analogs or derivatives thereof. An "antisense" nucleic acid comprises a nucleotide sequence that is complementary to a "sense" nucleic acid encoding a protein (e.g., complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence). In specific aspects, antisense nucleic acid molecules are provided that comprise a sequence complementary to at least about 10, 25, 50, 100, 250 or 500 nucleotides or an entire TETRA coding strand, or to only a portion thereof. Nucleic acid molecules encoding fragments, homologs, derivatives and analogs of a TETRA protein of SEQ ID NO:2; or antisense nucleic acids complementary to a TETRA nucleic acid sequence of SEQ ID NO:1, are additionally provided.

In one embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the coding strand of a nucleotide sequence encoding a TETRA protein. The term "coding region" refers to the region of the nucleotide sequence comprising codons which are translated into amino acid residues. In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence encoding the TETRA protein. The term "noncoding region" refers to 5' and 3' sequences which flank the coding region that are not translated into amino acids (i.e., also referred to as 5' and 3' untranslated regions).

Given the coding strand sequences encoding the TETRA protein disclosed herein, antisense nucleic acids of the invention can be designed according to the rules of Watson and Crick or Hoogsteen base pairing. The antisense nucleic acid molecule can be complementary to the entire coding region of TETRA mRNA, but more preferably is an oligonucleotide that is antisense to only a portion of the coding or noncoding region of TETRA mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of TETRA mRNA. An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be constructed using chemical synthesis or enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (e.g., an antisense oligonucleotide) can be chemically synthesized using naturally-occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids (e.g., phosphorothioate derivatives and acridine substituted nucleotides can be used).

Examples of modified nucleotides that can be used to generate the antisense nucleic acid include: 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxymethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, 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 methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-

N-2-carboxypyrrolyl) uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (i.e., RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

The antisense nucleic acid molecules of the invention are typically administered to a subject or generated in situ such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding a TETRA protein to thereby inhibit expression of the protein (e.g., by inhibiting transcription and/or translation). The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule that binds to DNA duplexes, through specific interactions in the major groove of the double helix. An example of a route of administration of antisense nucleic acid molecules of the invention includes direct injection at a tissue site. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then administered systemically. For example, for systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface (e.g., by linking the antisense nucleic acid molecules to peptides or antibodies that bind to cell surface receptors or antigens). The antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. To achieve sufficient nucleic acid molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a strong pol II or pol III promoter are preferred.

In yet another embodiment, the antisense nucleic acid molecule of the invention is an alpha-anomeric nucleic acid molecule. An alpha-anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual alpha-units, the strands run parallel to each other. See, e.g., Gaultier, et al., 1987. Nucl. Acids Res. 15: 6625-6641. The antisense nucleic acid molecule can also comprise a 2'-o-methylribonucleotide (see, e.g., Inoue, et al. 1987. Nucl. Acids Res. 15: 6131-6148) or a chimeric RNA-DNA analogue (see, e.g., Inoue, et al., 1987. FEBS Lett. 215: 327-330).

### **Ribozymes and PNA Moieties**

Nucleic acid modifications include, by way of non-limiting example, modified bases, and nucleic acids whose sugar phosphate backbones are modified or derivatized. These

modifications are carried out at least in part to enhance the chemical stability of the modified nucleic acid, such that they can be used, for example, as antisense binding nucleic acids in therapeutic applications in a subject.

In one embodiment, an antisense nucleic acid of the invention is a ribozyme.

5 Ribozymes are catalytic RNA molecules with ribonuclease activity that are capable of cleaving a single-stranded nucleic acid, such as an mRNA, to which they have a complementary region. Thus, ribozymes (e.g., hammerhead ribozymes as described in Haselhoff and Gerlach 1988. *Nature* 334: 585-591) can be used to catalytically cleave TETRA mRNA transcripts to thereby inhibit translation of TETRA mRNA. A ribozyme having  
10 specificity for a TETRA-encoding nucleic acid can be designed based upon the nucleotide sequence of a TETRA cDNA disclosed herein (i.e., SEQ ID NO:1). For example, a derivative of a *Tetrahymena* L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in a TETRA-encoding mRNA. See, e.g., U.S. Patent 4,987,071 to Cech, et al. and U.S. Patent 5,116,742 to Cech, et  
15 al. TETRA mRNA can also be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. See, e.g., Bartel et al., (1993) *Science* 261:1411-1418.

Alternatively, TETRA gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region of the TETRA nucleic acid (e.g., the  
20 TETRA promoter and/or enhancers) to form triple helical structures that prevent transcription of the TETRA gene in target cells. See, e.g., Helene, 1991. *Anticancer Drug Des.* 6: 569-84; Helene, et al. 1992. *Ann. N.Y. Acad. Sci.* 660: 27-36; Maher, 1992. *Bioassays* 14: 807-15.

In various embodiments, the TETRA nucleic acids can be modified at the base moiety, sugar moiety or phosphate backbone to improve, e.g., the stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of the nucleic acids can  
25 be modified to generate peptide nucleic acids. See, e.g., Hyrup, et al., 1996. *Bioorg Med Chem* 4: 5-23. As used herein, the terms "peptide nucleic acids" or "PNAs" refer to nucleic acid mimics (e.g., DNA mimics) in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of PNAs has been shown to allow for specific hybridization to DNA and RNA under  
30 conditions of low ionic strength. The synthesis of PNA oligomers can be performed using standard solid phase peptide synthesis protocols as described in Hyrup, et al., 1996. *supra*; Perry-O'Keefe, et al., 1996. *Proc. Natl. Acad. Sci. USA* 93: 14670-14675.

PNAs of TETRA can be used in therapeutic and diagnostic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, e.g., inducing transcription or translation arrest or inhibiting replication. PNAs of TETRA can also be used, for example, in the analysis of single base pair mutations in a gene (e.g., PNA directed PCR clamping; as artificial restriction enzymes when used in combination with other enzymes, e.g., S1 nucleases (see, Hyrup, et al., 1996. supra); or as probes or primers for DNA sequence and hybridization (see, Hyrup, et al., 1996, supra; Perry-O'Keefe, et al., 1996. supra).

In another embodiment, PNAs of TETRA can be modified, e.g., to enhance their stability or cellular uptake, by attaching lipophilic or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA chimeras of TETRA can be generated that may combine the advantageous properties of PNA and DNA. Such chimeras allow DNA recognition enzymes (e.g., RNase H and DNA polymerases) to interact with the DNA portion while the PNA portion would provide high binding affinity and specificity. PNA-DNA chimeras can be linked using linkers of appropriate lengths selected in terms of base stacking, number of bonds between the nucleobases, and orientation (see, Hyrup, et al., 1996. supra). The synthesis of PNA-DNA chimeras can be performed as described in Hyrup, et al., 1996. supra and Finn, et al., 1996. Nucl Acids Res 24: 3357-3363. For example, a DNA chain can be synthesized on a solid support using standard phosphoramidite coupling chemistry, and modified nucleoside analogs, e.g., 5'-(4-methoxytrityl)amino-5'-deoxy-thymidine phosphoramidite, can be used between the PNA and the 5' end of DNA. See, e.g., Mag, et al., 1989. Nucl Acid Res 17: 5973-5988. PNA monomers are then coupled in a stepwise manner to produce a chimeric molecule with a 5' PNA segment and a 3' DNA segment. See, e.g., Finn, et al., 1996. supra. Alternatively, chimeric molecules can be synthesized with a 5' DNA segment and a 3' PNA segment. See, e.g., Petersen, et al., 1975. Bioorg. Med. Chem. Lett. 5: 1119-11124.

In other embodiments, 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 (see, e.g., Letsinger, et al., 1989. Proc. Natl. Acad. Sci. U.S.A. 86: 6553-6556; Lemaitre, et al., 1987. Proc. Natl. Acad. Sci. 84: 648-652; PCT Publication No. WO88/09810) or the blood-brain barrier (see, e.g., PCT Publication No. WO 89/10134). In addition, oligonucleotides can be modified with hybridization triggered cleavage agents (see,

e.g., Krol, et al., 1988. *BioTechniques* 6:958-976) or intercalating agents (see, e.g., Zon, 1988. *Pharm. Res.* 5: 539-549). To this end, the oligonucleotide can be conjugated to another molecule, e.g., a peptide, a hybridization triggered cross-linking agent, a transport agent, a hybridization-triggered cleavage agent, and the like.

## 5 TETRA Polypeptides

A polypeptide according to the invention includes a polypeptide including the amino acid sequence of TETRA polypeptides whose sequences are provided in SEQ ID NO:2. The invention also includes a mutant or variant protein any of whose residues can be changed from the corresponding residues shown in SEQ ID NO:2, while still encoding a protein that  
10 maintains its TETRA activities and physiological functions, or a functional fragment thereof.

In general, a TETRA variant that preserves TETRA-like function includes any variant in which residues at a particular position in the sequence have been substituted by other amino acids, and further include the possibility of inserting an additional residue or residues between two residues of the parent protein as well as the possibility of deleting one or more residues  
15 from the parent sequence. Any amino acid substitution, insertion, or deletion is encompassed by the invention. In favorable circumstances, the substitution is a conservative substitution as defined above.

One aspect of the invention pertains to isolated TETRA proteins, and biologically-active portions thereof, or derivatives, fragments, analogs or homologs thereof. Also provided  
20 are polypeptide fragments suitable for use as immunogens to raise anti-TETRA antibodies. In one embodiment, native TETRA proteins can be isolated from cells or tissue sources by an appropriate purification scheme using standard protein purification techniques. In another embodiment, TETRA proteins are produced by recombinant DNA techniques. Alternative to recombinant expression, a TETRA protein or polypeptide can be synthesized chemically using  
25 standard peptide synthesis techniques.

An "isolated" or "purified" polypeptide or protein or biologically-active portion thereof is substantially free of cellular material or other contaminating proteins from the cell or tissue source from which the TETRA protein is derived, or substantially free from chemical precursors or other chemicals when chemically synthesized. The language "substantially free  
30 of cellular material" includes preparations of TETRA proteins in which the protein is separated from cellular components of the cells from which it is isolated or recombinantly-produced. In



one embodiment, the language "substantially free of cellular material" includes preparations of TETRA proteins having less than about 30% (by dry weight) of non-TETRA proteins (also referred to herein as a "contaminating protein"), more preferably less than about 20% of non-TETRA proteins, still more preferably less than about 10% of non-TETRA proteins, and most preferably less than about 5% of non-TETRA proteins. When the TETRA protein or biologically-active portion thereof is recombinantly-produced, it is also preferably substantially free of culture medium, i.e., culture medium represents less than about 20%, more preferably less than about 10%, and most preferably less than about 5% of the volume of the TETRA protein preparation.

The language "substantially free of chemical precursors or other chemicals" includes preparations of TETRA proteins in which the protein is separated from chemical precursors or other chemicals that are involved in the synthesis of the protein. In one embodiment, the language "substantially free of chemical precursors or other chemicals" includes preparations of TETRA proteins having less than about 30% (by dry weight) of chemical precursors or non-TETRA chemicals, more preferably less than about 20% chemical precursors or non-TETRA chemicals, still more preferably less than about 10% chemical precursors or non-TETRA chemicals, and most preferably less than about 5% chemical precursors or non-TETRA chemicals.

Biologically-active portions of TETRA proteins include peptides comprising amino acid sequences sufficiently homologous to or derived from the amino acid sequences of the TETRA proteins (e.g., the amino acid sequence shown in SEQ ID NO:2) that include fewer amino acids than the full-length TETRA proteins, and exhibit at least one activity of a TETRA protein. Typically, biologically-active portions comprise a domain or motif with at least one activity of the TETRA protein. A biologically-active portion of a TETRA protein can be a polypeptide which is, for example, 10, 25, 50, 100 or more amino acid residues in length. Moreover, other biologically-active portions, in which other regions of the protein are deleted, can be prepared by recombinant techniques and evaluated for one or more of the functional activities of a native TETRA protein.

In an embodiment, the TETRA protein has an amino acid sequence shown in SEQ ID NO:2. In other embodiments, the TETRA protein is substantially homologous to SEQ ID NO:2, and retains the functional activity of the protein of SEQ ID NO:2, yet differs in amino acid sequence due to natural allelic variation or mutagenesis, as described in detail, below.

Accordingly, in another embodiment, the TETRA protein is a protein that comprises an amino acid sequence at least about 45% homologous to the amino acid sequence of SEQ ID NO:2 and retains the functional activity of the TETRA proteins of SEQ ID NO:2.

### **Determining Homology Between Two or More Sequences**

5 To determine the percent homology 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  
10 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 homologous at that position (i.e., as used herein amino acid or nucleic acid "homology" is equivalent to amino acid or nucleic acid "identity").

The nucleic acid sequence homology can be determined as the degree of identity  
15 between two sequences. The homology can be determined using computer programs known in the art, such as GAP software provided in the GCG program package. See, Needleman and Wunsch, 1970. J Mol Biol 48: 443-453. Using GCG GAP software with the following settings for nucleic acid sequence comparison: GAP creation penalty of 5.0 and GAP extension penalty of 0.3, the coding region of the analogous nucleic acid sequences referred to  
20 above exhibits a degree of identity preferably of at least 70%, 75%, 80%, 85%, 90%, 95%, 98%, or 99%, with the CDS (encoding) part of the DNA sequence shown in SEQ ID NO:1.

The term "sequence identity" refers to the degree to which two polynucleotide or polypeptide sequences are identical on a residue-by-residue basis over a particular region of comparison. The term "percentage of sequence identity" is calculated by comparing two  
25 optimally aligned sequences over that region of comparison, determining the number of positions at which the identical nucleic acid base (e.g., A, T, C, G, U, or I, in the case of nucleic acids) occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the region of comparison (i.e., the window size), and multiplying the result by 100 to yield the percentage of sequence  
30 identity. The term "substantial identity" as used herein denotes a characteristic of a polynucleotide sequence, wherein the polynucleotide comprises a sequence that has at least 80

percent sequence identity, preferably at least 85 percent identity and often 90 to 95 percent sequence identity, more usually at least 99 percent sequence identity as compared to a reference sequence over a comparison region.

### **Chimeric and Fusion Proteins**

5           The invention also provides TETRA chimeric or fusion proteins. As used herein, a TETRA "chimeric protein" or "fusion protein" comprises a TETRA polypeptide operatively-linked to a non-TETRA polypeptide. An "TETRA polypeptide" refers to a polypeptide having an amino acid sequence corresponding to a TETRA protein (SEQ ID NO:2), whereas a "non-TETRA polypeptide" refers to a polypeptide having an amino acid sequence corresponding to  
10 a protein that is not substantially homologous to the TETRA protein, e.g., a protein that is different from the TETRA protein and that is derived from the same or a different organism. Within a TETRA fusion protein the TETRA polypeptide can correspond to all or a portion of a TETRA protein. In one embodiment, a TETRA fusion protein comprises at least one biologically-active portion of a TETRA protein. In another embodiment, a TETRA fusion  
15 protein comprises at least two biologically-active portions of a TETRA protein. In yet another embodiment, a TETRA fusion protein comprises at least three biologically-active portions of a TETRA protein. Within the fusion protein, the term "operatively-linked" is intended to indicate that the TETRA polypeptide and the non-TETRA polypeptide are fused in-frame with one another. The non-TETRA polypeptide can be fused to the N-terminus or C-terminus of  
20 the TETRA polypeptide.

In one embodiment, the fusion protein is a GST-TETRA fusion protein in which the TETRA sequences are fused to the C-terminus of the GST (glutathione S-transferase) sequences. Such fusion proteins can facilitate the purification of recombinant TETRA polypeptides.

25 In another embodiment, the fusion protein is a TETRA protein containing a heterologous signal sequence at its N-terminus. In certain host cells (e.g., mammalian host cells), expression and/or secretion of TETRA can be increased through use of a heterologous signal sequence.

In yet another embodiment, the fusion protein is a TETRA-immunoglobulin fusion  
30 protein in which the TETRA sequences are fused to sequences derived from a member of the immunoglobulin protein family. The TETRA-immunoglobulin fusion proteins of the

invention can be incorporated into pharmaceutical compositions and administered to a subject to inhibit an interaction between a TETRA ligand and a TETRA protein on the surface of a cell, to thereby suppress TETRA-mediated signal transduction in vivo. The TETRA-immunoglobulin fusion proteins can be used to affect the bioavailability of a TETRA cognate ligand. Inhibition of the TETRA ligand/TETRA interaction can be useful therapeutically for both the treatment of proliferative and differentiative disorders, as well as modulating (e.g. promoting or inhibiting) cell survival. Moreover, the TETRA-immunoglobulin fusion proteins of the invention can be used as immunogens to produce anti-TETRA antibodies in a subject, to purify TETRA ligands, and in screening assays to identify molecules that inhibit the interaction of TETRA with a TETRA ligand.

A TETRA chimeric or fusion protein of the invention can be produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional techniques, e.g., by employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers that give rise to complementary overhangs between two consecutive gene fragments that can subsequently be annealed and reamplified to generate a chimeric gene sequence (see, e.g., Ausubel, et al. (eds.) CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (e.g., a GST polypeptide). A TETRA-encoding nucleic acid can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the TETRA protein.

#### **TETRA Agonists and Antagonists**

The invention also pertains to variants of the TETRA proteins that function as either TETRA agonists (i.e., mimetics) or as TETRA antagonists. Variants of the TETRA protein can be generated by mutagenesis (e.g., discrete point mutation or truncation of the TETRA protein). An agonist of the TETRA protein can retain substantially the same, or a subset of, the biological activities of the naturally occurring form of the TETRA protein. An antagonist of the TETRA protein can inhibit one or more of the activities of the naturally occurring form of the TETRA protein by, for example, competitively binding to a downstream or upstream

member of a cellular signaling cascade which includes the TETRA protein. Thus, specific biological effects can be elicited by treatment with a variant of limited function. In one embodiment, treatment of a subject with a variant having a subset of the biological activities of the naturally occurring form of the protein has fewer side effects in a subject relative to  
5 treatment with the naturally occurring form of the TETRA proteins.

Variants of the TETRA proteins that function as either TETRA agonists (i.e., mimetics) or as TETRA antagonists can be identified by screening combinatorial libraries of mutants (e.g., truncation mutants) of the TETRA proteins for TETRA protein agonist or antagonist activity. In one embodiment, a variegated library of TETRA variants is generated  
10 by combinatorial mutagenesis at the nucleic acid level and is encoded by a variegated gene library. A variegated library of TETRA variants can be produced by, for example, enzymatically ligating a mixture of synthetic oligonucleotides into gene sequences such that a degenerate set of potential TETRA sequences is expressible as individual polypeptides, or alternatively, as a set of larger fusion proteins (e.g., for phage display) containing the set of  
15 TETRA sequences therein. There are a variety of methods which can be used to produce libraries of potential TETRA variants from a degenerate oligonucleotide sequence. Chemical synthesis of a degenerate gene sequence can be performed in an automatic DNA synthesizer, and the synthetic gene then ligated into an appropriate expression vector. Use of a degenerate set of genes allows for the provision, in one mixture, of all of the sequences encoding the  
20 desired set of potential TETRA sequences. Methods for synthesizing degenerate oligonucleotides are well-known within the art. See, e.g., Narang, 1983. *Tetrahedron* 39: 3; Itakura, et al., 1984. *Annu. Rev. Biochem.* 53: 323; Itakura, et al., 1984. *Science* 198: 1056; Ike, et al., 1983. *Nucl. Acids Res.* 11: 477.

### **Polypeptide Libraries**

25 In addition, libraries of fragments of the TETRA protein coding sequences can be used to generate a variegated population of TETRA fragments for screening and subsequent selection of variants of a TETRA protein. In one embodiment, a library of coding sequence fragments can be generated by treating a double stranded PCR fragment of a TETRA coding sequence with a nuclease under conditions wherein nicking occurs only about once per  
30 molecule, denaturing the double stranded DNA, renaturing the DNA to form double-stranded DNA that can include sense/antisense pairs from different nicked products, removing single stranded portions from reformed duplexes by treatment with S1 nuclease, and ligating the

resulting fragment library into an expression vector. By this method, expression libraries can be derived which encodes N-terminal and internal fragments of various sizes of the TETRA proteins.

5 Various techniques are known in the art for screening gene products of combinatorial libraries made by point mutations or truncation, and for screening cDNA libraries for gene products having a selected property. Such techniques are adaptable for rapid screening of the gene libraries generated by the combinatorial mutagenesis of TETRA proteins. The most widely used techniques, which are amenable to high throughput analysis, for screening large gene libraries typically include cloning the gene library into replicable expression vectors, 10 transforming appropriate cells with the resulting library of vectors, and expressing the combinatorial genes under conditions in which detection of a desired activity facilitates isolation of the vector encoding the gene whose product was detected. Recursive ensemble mutagenesis (REM), a new technique that enhances the frequency of functional mutants in the libraries, can be used in combination with the screening assays to identify TETRA variants. 15 See, e.g., Arkin and Yourvan, 1992. Proc. Natl. Acad. Sci. USA 89: 7811-7815; Delgrave, et al., 1993. Protein Engineering 6:327-331.

### **Anti-TETRA Antibodies**

The invention encompasses antibodies and antibody fragments, such as Fab or (Fab)<sub>2</sub>, that bind immunospecifically to any of the TETRA polypeptides of said invention.

20 An isolated TETRA protein, or a portion or fragment thereof, can be used as an immunogen to generate antibodies that bind to TETRA polypeptides using standard techniques for polyclonal and monoclonal antibody preparation. The full-length TETRA proteins can be used or, alternatively, the invention provides antigenic peptide fragments of TETRA proteins for use as immunogens. The antigenic TETRA peptides comprises at least 4 amino acid 25 residues of the amino acid sequence shown in SEQ ID NO:2, and encompasses an epitope of TETRA such that an antibody raised against the peptide forms a specific immune complex with TETRA. Preferably, the antigenic peptide comprises at least 6, 8, 10, 15, 20, or 30 amino acid residues. Longer antigenic peptides are sometimes preferable over shorter antigenic peptides, depending on use and according to methods well known to someone skilled in the 30 art.

In certain embodiments of the invention, at least one epitope encompassed by the antigenic peptide is a region of TETRA that is located on the surface of the protein (e.g., a hydrophilic region). As a means for targeting antibody production, hydropathy plots showing regions of hydrophilicity and hydrophobicity can be generated by any method well known in the art, including, for example, the Kyte Doolittle or the Hopp Woods methods, either with or without Fourier transformation (see, e.g., Hopp and Woods, 1981. Proc. Nat. Acad. Sci. USA 78: 3824-3828; Kyte and Doolittle, 1982. J. Mol. Biol. 157: 105-142, each incorporated herein by reference in their entirety).

As disclosed herein, TETRA protein sequences of SEQ ID NO:2 or derivatives, fragments, analogs or homologs thereof, can be utilized as immunogens in the generation of antibodies that immunospecifically-bind these protein components. The term "antibody" as used herein refers to immunoglobulin molecules and immunologically-active portions of immunoglobulin molecules, i.e., molecules that contain an antigen binding site that specifically-binds (immunoreacts with) an antigen, such as TETRA. Such antibodies include, but are not limited to, polyclonal, monoclonal, chimeric, single chain, Fab and F(ab')<sub>2</sub> fragments, and an Fab expression library. In a specific embodiment, antibodies to human TETRA proteins are disclosed. Various procedures known within the art can be used for the production of polyclonal or monoclonal antibodies to a TETRA protein sequence of SEQ ID NO:2 or a derivative, fragment, analog or homolog thereof. Some of these proteins are discussed below.

For the production of polyclonal antibodies, various suitable host animals (e.g., rabbit, goat, mouse or other mammal) can be immunized by injection with the native protein, or a synthetic variant thereof, or a derivative of the foregoing. An appropriate immunogenic preparation can contain, for example, recombinantly-expressed TETRA protein or a chemically-synthesized TETRA polypeptide. The preparation can further include an adjuvant. Various adjuvants used to increase the immunological response include, but are not limited to, Freund's (complete and incomplete), mineral gels (e.g., aluminum hydroxide), surface active substances (e.g., lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, dinitrophenol, etc.), human adjuvants such as Bacille Calmette-Guerin and Corynebacterium parvum, or similar immunostimulatory agents. If desired, the antibody molecules directed against TETRA can be isolated from the mammal (e.g., from the blood) and further purified by well known techniques, such as protein A chromatography to obtain the IgG fraction.

The term "monoclonal antibody" or "monoclonal antibody composition", as used herein, refers to a population of antibody molecules that contain only one species of an antigen binding site capable of immunoreacting with a particular epitope of TETRA. A monoclonal antibody composition thus typically displays a single binding affinity for a particular TETRA protein with which it immunoreacts. For preparation of monoclonal antibodies directed towards a particular TETRA protein, or derivatives, fragments, analogs or homologs thereof, any technique that provides for the production of antibody molecules by continuous cell line culture can be utilized. Such techniques include, but are not limited to, the hybridoma technique (see, e.g., Kohler & Milstein, 1975. *Nature* 256: 495-497); the trioma technique; the human B-cell hybridoma technique (see, e.g., Kozbor, et al., 1983. *Immunol. Today* 4: 72) and the EBV hybridoma technique to produce human monoclonal antibodies (see, e.g., Cole, et al., 1985. In: *MONOCLONAL ANTIBODIES AND CANCER THERAPY*, Alan R. Liss, Inc., pp. 77-96). Human monoclonal antibodies can be utilized in the practice of the invention and can be produced by using human hybridomas (see, e.g., Cote, et al., 1983. *Proc Natl Acad Sci USA* 80: 2026-2030) or by transforming human B-cells with Epstein Barr Virus in vitro (see, e.g., Cole, et al., 1985. In: *MONOCLONAL ANTIBODIES AND CANCER THERAPY*, Alan R. Liss, Inc., pp. 77-96). Each of the above citations is incorporated herein by reference in their entirety.

According to the invention, techniques can be adapted for the production of single-chain antibodies specific to a TETRA protein (see, e.g., U.S. Patent No. 4,946,778). In addition, methods can be adapted for the construction of Fab expression libraries (see, e.g., Huse, et al., 1989. *Science* 246: 1275-1281) to allow rapid and effective identification of monoclonal Fab fragments with the desired specificity for a TETRA protein or derivatives, fragments, analogs or homologs thereof. Non-human antibodies can be "humanized" by techniques well known in the art. See, e.g., U.S. Patent No. 5,225,539. Antibody fragments that contain the idiotypes to a TETRA protein can be produced by techniques known in the art including, but not limited to: (i) an F(ab')<sub>2</sub> fragment produced by pepsin digestion of an antibody molecule; (ii) an Fab fragment generated by reducing the disulfide bridges of an F(ab')<sub>2</sub> fragment; (iii) an Fab fragment generated by the treatment of the antibody molecule with papain and a reducing agent; and (iv) Fv fragments.

Additionally, recombinant anti-TETRA antibodies, such as chimeric and humanized monoclonal antibodies, comprising both human and non-human portions, which can be made using standard recombinant DNA techniques, are within the scope of the invention. Such



chimeric and humanized monoclonal antibodies can be produced by recombinant DNA techniques known in the art, for example using methods described in International Application No. PCT/US86/02269; European Patent Application No. 184,187; European Patent Application No. 171,496; European Patent Application No. 173,494; PCT International  
5 Publication No. WO 86/01533; U.S. Patent No. 4,816,567; U.S. Pat. No. 5,225,539; European Patent Application No. 125,023; Better, et al., 1988. Science 240: 1041-1043; Liu, et al., 1987. Proc. Natl. Acad. Sci. USA 84: 3439-3443; Liu, et al., 1987. J. Immunol. 139: 3521-3526; Sun, et al., 1987. Proc. Natl. Acad. Sci. USA 84: 214-218; Nishimura, et al., 1987. Cancer Res. 47: 999-1005; Wood, et al., 1985. Nature 314 :446-449; Shaw, et al., 1988. J. Natl.  
10 Cancer Inst. 80: 1553-1559; Morrison(1985) Science 229:1202-1207; Oi, et al. (1986) BioTechniques 4:214; Jones, et al., 1986. Nature 321: 552-525; Verhoeyan, et al., 1988. Science 239: 1534; and Beidler, et al., 1988. J. Immunol. 141: 4053-4060. Each of the above citations are incorporated herein by reference in their entirety.

In one embodiment, methods for the screening of antibodies that possess the desired  
15 specificity include, but are not limited to, enzyme-linked immunosorbent assay (ELISA) and other immunologically-mediated techniques known within the art. In a specific embodiment, selection of antibodies that are specific to a particular domain of a TETRA protein is facilitated by generation of hybridomas that bind to the fragment of a TETRA protein possessing such a domain. Thus, antibodies that are specific for a desired domain within a  
20 TETRA protein, or derivatives, fragments, analogs or homologs thereof, are also provided herein.

Anti-TETRA antibodies can be used in methods known within the art relating to the localization and/or quantitation of a TETRA protein (e.g., for use in measuring levels of the TETRA protein within appropriate physiological samples, for use in diagnostic methods, for  
25 use in imaging the protein, and the like). In a given embodiment, antibodies for TETRA proteins, or derivatives, fragments, analogs or homologs thereof, that contain the antibody derived binding domain, are utilized as pharmacologically-active compounds (hereinafter "Therapeutics").

An anti-TETRA antibody (e.g., monoclonal antibody) can be used to isolate a TETRA  
30 polypeptide by standard techniques, such as affinity chromatography or immunoprecipitation. An anti-TETRA antibody can facilitate the purification of natural TETRA polypeptide from cells and of recombinantly-produced TETRA polypeptide expressed in host cells. Moreover,

an anti-TETRA antibody can be used to detect TETRA protein (e.g., in a cellular lysate or cell supernatant) in order to evaluate the abundance and pattern of expression of the TETRA protein. Anti-TETRA antibodies can be used diagnostically to monitor protein levels in tissue as part of a clinical testing procedure, e.g., to, for example, determine the efficacy of a given treatment regimen. Detection can be facilitated by coupling (i.e., physically linking) the antibody to a detectable substance. Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, and radioactive materials. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, beta-galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes luminol; examples of bioluminescent materials include luciferase, luciferin, and aequorin, and examples of suitable radioactive material include  $^{125}\text{I}$ ,  $^{131}\text{I}$ ,  $^{35}\text{S}$  or  $^3\text{H}$ .

#### **TETRA Recombinant Expression Vectors and Host Cells**

Another aspect of the invention pertains to vectors, preferably expression vectors, containing a nucleic acid encoding a TETRA protein, or derivatives, fragments, analogs or homologs thereof. As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a "plasmid", which refers to a circular double stranded DNA loop into which additional DNA segments can be ligated. Another type of vector is a viral vector, wherein additional DNA segments can be ligated into the viral genome. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (e.g., bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Other vectors (e.g., non-episomal mammalian vectors) are integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along with the host genome. Moreover, certain vectors are capable of directing the expression of genes to which they are operatively-linked. Such vectors are referred to herein as "expression vectors". In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids. In the present specification, "plasmid" and "vector" can be used interchangeably as the plasmid is the most commonly used form of vector. However, the invention is intended to include such other

forms of expression vectors, such as viral vectors (e.g., replication defective retroviruses, adenoviruses and adeno-associated viruses), which serve equivalent functions.

5 The recombinant expression vectors of the invention comprise a nucleic acid of the invention in a form suitable for expression of the nucleic acid in a host cell, which means that the recombinant expression vectors include one or more regulatory sequences, selected on the basis of the host cells to be used for expression, that is operatively-linked to the nucleic acid sequence to be expressed. Within a recombinant expression vector, "operably-linked" is intended to mean that the nucleotide sequence of interest is linked to the regulatory sequence(s) in a manner that allows for expression of the nucleotide sequence (e.g., in an in vitro transcription/translation system or in a host cell when the vector is introduced into the host cell).

10 The term "regulatory sequence" is intended to include promoters, enhancers and other expression control elements (e.g., polyadenylation signals). Such regulatory sequences are described, for example, in Goeddel, GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990). Regulatory sequences include those that direct constitutive expression of a nucleotide sequence in many types of host cell and those that direct expression of the nucleotide sequence only in certain host cells (e.g., tissue-specific regulatory sequences). It will be appreciated by those skilled in the art that the design of the expression vector can depend on such factors as the choice of the host cell to be transformed, the level of expression of protein desired, etc. The expression vectors of the invention can be introduced into host cells to thereby produce proteins or peptides, including fusion proteins or peptides, encoded by nucleic acids as described herein (e.g., TETRA proteins, mutant forms of TETRA proteins, fusion proteins, etc.).

25 The recombinant expression vectors of the invention can be designed for expression of TETRA proteins in prokaryotic or eukaryotic cells. For example, TETRA proteins can be expressed in bacterial cells such as Escherichia coli, insect cells (using baculovirus expression vectors) yeast cells or mammalian cells. Suitable host cells are discussed further in Goeddel, GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990). Alternatively, the recombinant expression vector can be transcribed and translated in vitro, for example using T7 promoter regulatory sequences and T7 polymerase.

Expression of proteins in prokaryotes is most often carried out in *Escherichia coli* with vectors containing constitutive or inducible promoters directing the expression of either fusion or non-fusion proteins. Fusion vectors add a number of amino acids to a protein encoded therein, usually to the amino terminus of the recombinant protein. Such fusion vectors typically serve  
5 three purposes: (i) to increase expression of recombinant protein; (ii) to increase the solubility of the recombinant protein; and (iii) to aid in the purification of the recombinant protein by acting as a ligand in affinity purification. Often, in fusion expression vectors, a proteolytic cleavage site is introduced at the junction of the fusion moiety and the recombinant protein to enable separation of the recombinant protein from the fusion moiety subsequent to purification  
10 of the fusion protein. Such enzymes, and their cognate recognition sequences, include Factor Xa, thrombin and enterokinase. Typical fusion expression vectors include pGEX (Pharmacia Biotech Inc; Smith and Johnson, 1988. *Gene* 67: 31-40), pMAL (New England Biolabs, Beverly, Mass.) and pRIT5 (Pharmacia, Piscataway, N.J.) that fuse glutathione S-transferase (GST), maltose E binding protein, or protein A, respectively, to the target recombinant protein.

15 Examples of suitable inducible non-fusion *E. coli* expression vectors include pTrc (Amrann et al., (1988) *Gene* 69:301-315) and pET 11d (Studier et al., *GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY* 185, Academic Press, San Diego, Calif. (1990) 60-89).

One strategy to maximize recombinant protein expression in *E. coli* is to express the  
20 protein in a host bacteria with an impaired capacity to proteolytically cleave the recombinant protein. See, e.g., Gottesman, *GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY* 185, Academic Press, San Diego, Calif. (1990) 119-128. Another strategy is to alter the nucleic acid sequence of the nucleic acid to be inserted into an expression vector so that the individual codons for each amino acid are those preferentially utilized in *E. coli* (see,  
25 e.g., Wada, et al., 1992. *Nucl. Acids Res.* 20: 2111-2118). Such alteration of nucleic acid sequences of the invention can be carried out by standard DNA synthesis techniques. In another embodiment, the TETRA expression vector is a yeast expression vector. Examples of vectors for expression in yeast *Saccharomyces cerevisiae* include pYepSec1 (Baldari, et al., 1987. *EMBO J.* 6: 229-234), pMFa (Kurjan and Herskowitz, 1982. *Cell* 30: 933-943), pJRY88  
30 (Schultz et al., 1987. *Gene* 54: 113-123), pYES2 (Invitrogen Corporation, San Diego, Calif.), and picZ (Invitrogen Corp, San Diego, Calif.). Alternatively, TETRA can be expressed in insect cells using baculovirus expression vectors. Baculovirus vectors available for expression of proteins in cultured insect cells (e.g., SF9 cells) include the pAc series (Smith, et al., 1983.

Mol. Cell. Biol. 3: 2156-2165) and the pVL series (Lucklow and Summers, 1989. Virology 170: 31-39).

In yet another embodiment, a nucleic acid of the invention is expressed in mammalian cells using a mammalian expression vector. Examples of mammalian expression vectors  
5 include pCDM8 (Seed, 1987. Nature 329: 840) and pMT2PC (Kaufman, et al., 1987. EMBO J. 6: 187-195). When used in mammalian cells, the expression vector's control functions are often provided by viral regulatory elements. For example, commonly used promoters are derived from polyoma, adeTETRAirus 2, cytomegalovirus, and simian virus 40. For other  
10 suitable expression systems for both prokaryotic and eukaryotic cells see, e.g., Chapters 16 and 17 of Sambrook, et al., MOLECULAR CLONING: A LABORATORY MANUAL. 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989.

In another embodiment, the recombinant mammalian expression vector is capable of directing expression of the nucleic acid preferentially in a particular cell type (e.g., tissue-  
15 specific regulatory elements are used to express the nucleic acid). Tissue-specific regulatory elements are known in the art. Non-limiting examples of suitable tissue-specific promoters include the albumin promoter (liver-specific; Pinkert, et al., 1987. Genes Dev. 1: 268-277), lymphoid-specific promoters (Calame and Eaton, 1988. Adv. Immunol. 43: 235-275), in particular promoters of T cell receptors (Winoto and Baltimore, 1989. EMBO J. 8: 729-733)  
20 and immunoglobulins (Banerji, et al., 1983. Cell 33: 729-740; Queen and Baltimore, 1983. Cell 33: 741-748), neuron-specific promoters (e.g., the neurofilament promoter; Byrne and Ruddle, 1989. Proc. Natl. Acad. Sci. USA 86: 5473-5477), pancreas-specific promoters (Edlund, et al., 1985. Science 230: 912-916), and mammary gland-specific promoters (e.g., milk whey promoter; U.S. Pat. No. 4,873,316 and European Application Publication No.  
25 264,166). Developmentally-regulated promoters are also encompassed, e.g., the murine hox promoters (Kessel and Gruss, 1990. Science 249: 374-379) and the  $\alpha$ -fetoprotein promoter (Campes and Tilghman, 1989. Genes Dev. 3: 537-546).

The invention further provides a recombinant expression vector comprising a DNA molecule of the invention cloned into the expression vector in an antisense orientation. That  
30 is, the DNA molecule is operatively-linked to a regulatory sequence in a manner that allows for expression (by transcription of the DNA molecule) of an RNA molecule that is antisense to TETRA mRNA. Regulatory sequences operatively linked to a nucleic acid cloned in the

antisense orientation can be chosen that direct the continuous expression of the antisense RNA molecule in a variety of cell types, for instance viral promoters and/or enhancers, or regulatory sequences can be chosen that direct constitutive, tissue specific or cell type specific expression of antisense RNA. The antisense expression vector can be in the form of a recombinant  
5 plasmid, phagemid or attenuated virus in which antisense nucleic acids are produced under the control of a high efficiency regulatory region, the activity of which can be determined by the cell type into which the vector is introduced. For a discussion of the regulation of gene expression using antisense genes see, e.g., Weintraub, et al., "Antisense RNA as a molecular tool for genetic analysis," *Reviews-Trends in Genetics*, Vol. 1(1) 1986.

10 Another aspect of the invention pertains to host cells into which a recombinant expression vector of the invention has been introduced. The terms "host cell" and "recombinant host cell" are used interchangeably herein. It is understood that such terms refer not only to the particular subject cell but also to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generations due to either  
15 mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein.

A host cell can be any prokaryotic or eukaryotic cell. For example, TETRA protein can be expressed in bacterial cells such as *E. coli*, insect cells, yeast or mammalian cells (such as Chinese hamster ovary cells (CHO) or COS cells). Other suitable host cells are known to  
20 those skilled in the art.

Vector DNA can be introduced into prokaryotic or eukaryotic cells via conventional transformation or transfection techniques. As used herein, the terms "transformation" and "transfection" are intended to refer to a variety of art-recognized techniques for introducing foreign nucleic acid (e.g., DNA) into a host cell, including calcium phosphate or calcium  
25 chloride co-precipitation, DEAE-dextran-mediated transfection, lipofection, or electroporation. Suitable methods for transforming or transfecting host cells can be found in Sambrook, et al. (*MOLECULAR CLONING: A LABORATORY MANUAL*. 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989), and other laboratory manuals.

30 For stable transfection of mammalian cells, it is known that, depending upon the expression vector and transfection technique used, only a small fraction of cells may integrate

the foreign DNA into their genome. In order to identify and select these integrants, a gene that encodes a selectable marker (e.g., resistance to antibiotics) is generally introduced into the host cells along with the gene of interest. Various selectable markers include those that confer resistance to drugs, such as G418, hygromycin and methotrexate. Nucleic acid encoding a selectable marker can be introduced into a host cell on the same vector as that encoding TETRA or can be introduced on a separate vector. Cells stably transfected with the introduced nucleic acid can be identified by drug selection (e.g., cells that have incorporated the selectable marker gene will survive, while the other cells die).

A host cell of the invention, such as a prokaryotic or eukaryotic host cell in culture, can be used to produce (i.e., express) TETRA protein. Accordingly, the invention further provides methods for producing TETRA protein using the host cells of the invention. In one embodiment, the method comprises culturing the host cell of invention (into which a recombinant expression vector encoding TETRA protein has been introduced) in a suitable medium such that TETRA protein is produced. In another embodiment, the method further comprises isolating TETRA protein from the medium or the host cell.

### **Transgenic TETRA Animals**

The host cells of the invention can also be used to produce non-human transgenic animals. For example, in one embodiment, a host cell of the invention is a fertilized oocyte or an embryonic stem cell into which TETRA protein-coding sequences have been introduced. Such host cells can then be used to create non-human transgenic animals in which exogenous TETRA sequences have been introduced into their genome or homologous recombinant animals in which endogenous TETRA sequences have been altered. Such animals are useful for studying the function and/or activity of TETRA protein and for identifying and/or evaluating modulators of TETRA protein activity. As used herein, a "transgenic animal" is a non-human animal, preferably a mammal, more preferably a rodent such as a rat or mouse, in which one or more of the cells of the animal includes a transgene. Other examples of transgenic animals include non-human primates, sheep, dogs, cows, goats, chickens, amphibians, etc. A transgene is exogenous DNA that is integrated into the genome of a cell from which a transgenic animal develops and that remains in the genome of the mature animal, thereby directing the expression of an encoded gene product in one or more cell types or tissues of the transgenic animal. As used herein, a "homologous recombinant animal" is a non-human animal, preferably a mammal, more preferably a mouse, in which an endogenous

TETRA gene has been altered by homologous recombination between the endogenous gene and an exogenous DNA molecule introduced into a cell of the animal, e.g., an embryonic cell of the animal, prior to development of the animal.

5 A transgenic animal of the invention can be created by introducing TETRA-encoding nucleic acid into the male pronuclei of a fertilized oocyte (e.g., by microinjection, retroviral infection) and allowing the oocyte to develop in a pseudopregnant female foster animal. The human TETRA cDNA sequences of SEQ ID NO:1, can be introduced as a transgene into the genome of a non-human animal. Alternatively, a non-human homologue of the human TETRA gene, such as a mouse TETRA gene, can be isolated based on hybridization to the human TETRA  
10 cDNA (described further supra) and used as a transgene. Intronic sequences and polyadenylation signals can also be included in the transgene to increase the efficiency of expression of the transgene. A tissue-specific regulatory sequence(s) can be operably-linked to the TETRA transgene to direct expression of TETRA protein to particular cells. Methods for generating transgenic animals via embryo manipulation and microinjection, particularly  
15 animals such as mice, have become conventional in the art and are described, for example, in U.S. Patent Nos. 4,736,866; 4,870,009; and 4,873,191; and Hogan, 1986. In: MANIPULATING THE MOUSE EMBRYO, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. Similar methods are used for production of other transgenic animals. A transgenic founder animal can be identified based upon the presence of the TETRA transgene  
20 in its genome and/or expression of TETRA mRNA in tissues or cells of the animals. A transgenic founder animal can then be used to breed additional animals carrying the transgene. Moreover, transgenic animals carrying a transgene-encoding TETRA protein can further be bred to other transgenic animals carrying other transgenes.

To create a homologous recombinant animal, a vector is prepared which contains at  
25 least a portion of a TETRA gene into which a deletion, addition or substitution has been introduced to thereby alter, e.g., functionally disrupt, the TETRA gene. The TETRA gene can be a human gene (e.g., the cDNA of SEQ ID NO:1), but more preferably, is a non-human homologue of a human TETRA gene. For example, a mouse homologue of human TETRA gene of SEQ ID NO:1, can be used to construct a homologous recombination vector suitable  
30 for altering an endogenous TETRA gene in the mouse genome. In one embodiment, the vector is designed such that, upon homologous recombination, the endogenous TETRA gene is functionally disrupted (i.e., no longer encodes a functional protein; also referred to as a "knock out" vector).



Alternatively, the vector can be designed such that, upon homologous recombination, the endogenous TETRA gene is mutated or otherwise altered but still encodes functional protein (e.g., the upstream regulatory region can be altered to thereby alter the expression of the endogenous TETRA protein). In the homologous recombination vector, the altered portion of the TETRA gene is flanked at its 5'- and 3'-termini by additional nucleic acid of the TETRA gene to allow for homologous recombination to occur between the exogenous TETRA gene carried by the vector and an endogenous TETRA gene in an embryonic stem cell. The additional flanking TETRA nucleic acid is of sufficient length for successful homologous recombination with the endogenous gene. Typically, several kilobases of flanking DNA (both at the 5'- and 3'-termini) are included in the vector. See, e.g., Thomas, et al., 1987. Cell 51: 503 for a description of homologous recombination vectors. The vector is then introduced into an embryonic stem cell line (e.g., by electroporation) and cells in which the introduced TETRA gene has homologously-recombined with the endogenous TETRA gene are selected. See, e.g., Li, et al., 1992. Cell 69: 915.

The selected cells are then injected into a blastocyst of an animal (e.g., a mouse) to form aggregation chimeras. See, e.g., Bradley, 1987. In: TERATOCARCINOMAS AND EMBRYONIC STEM CELLS: A PRACTICAL APPROACH, Robertson, ed. IRL, Oxford, pp. 113-152. A chimeric embryo can then be implanted into a suitable pseudopregnant female foster animal and the embryo brought to term. Progeny harboring the homologously-recombined DNA in their germ cells can be used to breed animals in which all cells of the animal contain the homologously-recombined DNA by germline transmission of the transgene. Methods for constructing homologous recombination vectors and homologous recombinant animals are described further in Bradley, 1991. Curr. Opin. Biotechnol. 2: 823-829; PCT International Publication Nos.: WO 90/11354; WO 91/01140; WO 92/0968; and WO 93/04169.

In another embodiment, transgenic non-humans animals can be produced that contain selected systems that allow for regulated expression of the transgene. One example of such a system is the cre/loxP recombinase system of bacteriophage P1. For a description of the cre/loxP recombinase system, See, e.g., Lakso, et al., 1992. Proc. Natl. Acad. Sci. USA 89: 6232-6236. Another example of a recombinase system is the FLP recombinase system of *Saccharomyces cerevisiae*. See, O'Gorman, et al., 1991. Science 251:1351-1355. If a cre/loxP recombinase system is used to regulate expression of the transgene, animals containing transgenes encoding both the Cre recombinase and a selected protein are required. Such

animals can be provided through the construction of "double" transgenic animals, e.g., by mating two transgenic animals, one containing a transgene encoding a selected protein and the other containing a transgene encoding a recombinase.

Clones of the non-human transgenic animals described herein can also be produced according to the methods described in Wilmut, et al., 1997. Nature 385: 810-813. In brief, a cell (e.g., a somatic cell) from the transgenic animal can be isolated and induced to exit the growth cycle and enter G0 phase. The quiescent cell can then be fused, e.g., through the use of electrical pulses, to an enucleated oocyte from an animal of the same species from which the quiescent cell is isolated. The reconstructed oocyte is then cultured such that it develops to morula or blastocyte and then transferred to pseudopregnant female foster animal. The offspring borne of this female foster animal will be a clone of the animal from which the cell (e.g., the somatic cell) is isolated.

### **Pharmaceutical Compositions**

The TETRA nucleic acid molecules, TETRA proteins, and anti-TETRA antibodies (also referred to herein as "active compounds") of the invention, and derivatives, fragments, analogs and homologs thereof, can be incorporated into pharmaceutical compositions suitable for administration. Such compositions typically comprise the nucleic acid molecule, protein, or antibody and a pharmaceutically acceptable carrier. As used herein, "pharmaceutically acceptable carrier" is intended to include any and all solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like, compatible with pharmaceutical administration. Suitable carriers are described in the most recent edition of Remington's Pharmaceutical Sciences, a standard reference text in the field, which is incorporated herein by reference. Preferred examples of such carriers or diluents include, but are not limited to, water, saline, finger's solutions, dextrose solution, and 5% human serum albumin. Liposomes and non-aqueous vehicles such as fixed oils may also be used. The use of such media and agents for pharmaceutically active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the active compound, use thereof in the compositions is contemplated. Supplementary active compounds can also be incorporated into the compositions.

A pharmaceutical composition of the invention is formulated to be compatible with its intended route of administration. Examples of routes of administration include parenteral,

e.g., intravenous, intradermal, subcutaneous, oral (e.g., inhalation), transdermal (i.e., topical), transmucosal, and rectal administration. Solutions or suspensions used for parenteral, intradermal, or subcutaneous application can include the following components: a sterile diluent such as water for injection, saline solution, fixed oils, polyethylene glycols, glycerine, propylene glycol or other synthetic solvents; antibacterial agents such as benzyl alcohol or methyl parabens; antioxidants such as ascorbic acid or sodium bisulfite; chelating agents such as ethylenediaminetetraacetic acid (EDTA); buffers such as acetates, citrates or phosphates, and agents for the adjustment of tonicity such as sodium chloride or dextrose. The pH can be adjusted with acids or bases, such as hydrochloric acid or sodium hydroxide. The parenteral preparation can be enclosed in ampoules, disposable syringes or multiple dose vials made of glass or plastic.

Pharmaceutical compositions suitable for injectable use include sterile aqueous solutions (where water soluble) or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersion. For intravenous administration, suitable carriers include physiological saline, bacteriostatic water, Cremophor EL™ (BASF, Parsippany, N.J.) or phosphate buffered saline (PBS). In all cases, the composition must be sterile and should be fluid to the extent that easy syringeability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyethylene glycol, and the like), and suitable mixtures thereof. The proper fluidity can be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. Prevention of the action of microorganisms can be achieved by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, ascorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars, polyalcohols such as manitol, sorbitol, sodium chloride in the composition. Prolonged absorption of the injectable compositions can be brought about by including in the composition an agent which delays absorption, for example, aluminum monostearate and gelatin.

Sterile injectable solutions can be prepared by incorporating the active compound (e.g., a TETRA protein or anti-TETRA antibody) in the required amount in an appropriate solvent with one or a combination of ingredients enumerated above, as required, followed by filtered

sterilization. Generally, dispersions are prepared by incorporating the active compound into a sterile vehicle that contains a basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, methods of preparation are vacuum drying and freeze-drying that yields a powder of  
5 the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

Oral compositions generally include an inert diluent or an edible carrier. They can be enclosed in gelatin capsules or compressed into tablets. For the purpose of oral therapeutic administration, the active compound can be incorporated with excipients and used in the form  
10 of tablets, troches, or capsules. Oral compositions can also be prepared using a fluid carrier for use as a mouthwash, wherein the compound in the fluid carrier is applied orally and swished and expectorated or swallowed. Pharmaceutically compatible binding agents, and/or adjuvant materials can be included as part of the composition. The tablets, pills, capsules, troches and the like can contain any of the following ingredients, or compounds of a similar  
15 nature: a binder such as microcrystalline cellulose, gum tragacanth or gelatin; an excipient such as starch or lactose, a disintegrating agent such as alginic acid, Primogel, or corn starch; a lubricant such as magnesium stearate or Sterotes; a glidant such as colloidal silicon dioxide; a sweetening agent such as sucrose or saccharin; or a flavoring agent such as peppermint, methyl salicylate, or orange flavoring.

20 For administration by inhalation, the compounds are delivered in the form of an aerosol spray from pressured container or dispenser which contains a suitable propellant, e.g., a gas such as carbon dioxide, or a nebulizer.

Systemic administration can also be by transmucosal or transdermal means. For transmucosal or transdermal administration, penetrants appropriate to the barrier to be  
25 permeated are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration, detergents, bile salts, and fusidic acid derivatives. Transmucosal administration can be accomplished through the use of nasal sprays or suppositories. For transdermal administration, the active compounds are formulated into ointments, salves, gels, or creams as generally known in the art.

The compounds can also be prepared in the form of suppositories (e.g., with conventional suppository bases such as cocoa butter and other glycerides) or retention enemas for rectal delivery.

5 In one embodiment, the active compounds are prepared with carriers that will protect the compound against rapid elimination from the body, such as a controlled release formulation, including implants and microencapsulated delivery systems. Biodegradable, biocompatible polymers can be used, such as ethylene vinyl acetate, polyanhydrides, polyglycolic acid, collagen, polyorthoesters, and polylactic acid. Methods for preparation of such formulations will be apparent to those skilled in the art. The materials can also be  
10 obtained commercially from Alza Corporation and TETRAa Pharmaceuticals, Inc. Liposomal suspensions (including liposomes targeted to infected cells with monoclonal antibodies to viral antigens) can also be used as pharmaceutically acceptable carriers. These can be prepared according to methods known to those skilled in the art, for example, as described in U.S. Patent No. 4,522,811.

15 It is especially advantageous to formulate oral or parenteral compositions in dosage unit form for ease of administration and uniformity of dosage. Dosage unit form as used herein refers to physically discrete units suited as unitary dosages for the subject to be treated; each unit containing a predetermined quantity of active compound calculated to produce the desired therapeutic effect in association with the required pharmaceutical carrier. The  
20 specification for the dosage unit forms of the invention are dictated by and directly dependent on the unique characteristics of the active compound and the particular therapeutic effect to be achieved, and the limitations inherent in the art of compounding such an active compound for the treatment of individuals.

The nucleic acid molecules of the invention can be inserted into vectors and used as  
25 gene therapy vectors. Gene therapy vectors can be delivered to a subject by, for example, intravenous injection, local administration (see, e.g., U.S. Patent No. 5,328,470) or by stereotactic injection (see, e.g., Chen, et al., 1994. Proc. Natl. Acad. Sci. USA 91: 3054-3057). The pharmaceutical preparation of the gene therapy vector can include the gene therapy vector in an acceptable diluent, or can comprise a slow release matrix in which the gene delivery  
30 vehicle is imbedded. Alternatively, where the complete gene delivery vector can be produced intact from recombinant cells, e.g., retroviral vectors, the pharmaceutical preparation can include one or more cells that produce the gene delivery system. The pharmaceutical

compositions can be included in a container, pack, or dispenser together with instructions for administration.

### Screening and Detection Methods

The isolated nucleic acid molecules of the invention can be used to express TETRA protein (e.g., via a recombinant expression vector in a host cell in gene therapy applications), to detect TETRA mRNA (e.g., in a biological sample) or a genetic lesion in a TETRA gene, and to modulate TETRA activity, as described further, below. In addition, the TETRA proteins can be used to screen drugs or compounds that modulate the TETRA protein activity or expression as well as to treat disorders characterized by insufficient or excessive production of TETRA protein or production of TETRA protein forms that have decreased or aberrant activity compared to TETRA wild-type protein. In addition, the anti-TETRA antibodies of the invention can be used to detect and isolate TETRA proteins and modulate TETRA activity.

The invention further pertains to novel agents identified by the screening assays described herein and uses thereof for treatments as described, supra.

### 15 Screening Assays

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) that bind to TETRA proteins or have a stimulatory or inhibitory effect on, e.g., TETRA protein expression or TETRA protein activity. The invention also includes compounds identified in the screening assays described herein. In one embodiment, the invention provides assays for screening candidate or test compounds which bind to or modulate the activity of the membrane-bound form of a TETRA protein or polypeptide or biologically-active portion thereof. The test compounds of the 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. See, e.g., Lam, 1997. *Anticancer Drug Design* 12: 145.

A "small molecule" as used herein, is meant to refer to a composition that has a molecular weight of less than about 5 kD and most preferably less than about 4 kD. Small molecules can be, e.g., nucleic acids, peptides, polypeptides, peptidomimetics, carbohydrates, lipids or other organic or inorganic molecules. Libraries of chemical and/or biological mixtures, such as fungal, bacterial, or algal extracts, are known in the art and can be screened with any of the assays of the invention.

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. U.S.A. 90: 6909; Erb, et al., 1994. Proc. Natl. Acad. Sci. U.S.A. 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.

Libraries of compounds can be presented in solution (e.g., Houghten, 1992. Biotechniques 13: 412-421), or on beads (Lam, 1991. Nature 354: 82-84), on chips (Fodor, 1993. Nature 364: 555-556), bacteria (Ladner, U.S. Patent No. 5,223,409), spores (Ladner, U.S. Patent 5,233,409), plasmids (Cull, et al., 1992. Proc. Natl. Acad. Sci. USA 89: 1865-1869) or on phage (Scott and Smith, 1990. Science 249: 386-390; Devlin, 1990. Science 249: 404-406; Cwirla, et al., 1990. Proc. Natl. Acad. Sci. U.S.A. 87: 6378-6382; Felici, 1991. J. Mol. Biol. 222: 301-310; Ladner, U.S. Patent No. 5,233,409.).

In one embodiment, an assay is a cell-based assay in which a cell which expresses a membrane-bound form of TETRA protein, 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 a TETRA protein determined. The cell, for example, can be of mammalian origin or a yeast cell. Determining the ability of the test compound to bind to the TETRA protein 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 TETRA protein or biologically-active portion thereof can be determined by detecting the labeled compound in a complex. For example, test compounds can be labeled with  $^{125}\text{I}$ ,  $^{35}\text{S}$ ,  $^{14}\text{C}$ , or  $^3\text{H}$ , 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 one embodiment, the

assay comprises contacting a cell which expresses a membrane-bound form of TETRA protein, or a biologically-active portion thereof, on the cell surface with a known compound which binds TETRA to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with a TETRA protein, wherein determining the ability of the test compound to interact with a TETRA protein comprises determining the ability of the test compound to preferentially bind to TETRA protein or a biologically-active portion thereof as compared to the known compound.

In another embodiment, an assay is a cell-based assay comprising contacting a cell expressing a membrane-bound form of TETRA protein, 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 TETRA protein or biologically-active portion thereof.

Determining the ability of the test compound to modulate the activity of TETRA or a biologically-active portion thereof can be accomplished, for example, by determining the ability of the TETRA protein to bind to or interact with a TETRA target molecule. As used herein, a "target molecule" is a molecule with which a TETRA protein binds or interacts in nature, for example, a molecule on the surface of a cell which expresses a TETRA interacting 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 TETRA target molecule can be a non-TETRA molecule or a TETRA protein or polypeptide of the invention. In one embodiment, a TETRA target molecule is a component of a signal transduction pathway that facilitates transduction of an extracellular signal (e.g. a signal generated by binding of a compound to a membrane-bound TETRA molecule) through the cell membrane and into the cell. The target, for example, can be a second intercellular protein that has catalytic activity or a protein that facilitates the association of downstream signaling molecules with TETRA.

Determining the ability of the TETRA protein to bind to or interact with a TETRA target molecule can be accomplished by one of the methods described above for determining direct binding. In one embodiment, determining the ability of the TETRA protein to bind to or interact with a TETRA 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 (i.e. intracellular  $\text{Ca}^{2+}$ ,



diacylglycerol, IP3, etc.), detecting catalytic/enzymatic activity of the target an appropriate substrate, detecting the induction of a reporter gene (comprising a TETRA-responsive regulatory element operatively linked to a nucleic acid encoding a detectable marker, e.g., luciferase), or detecting a cellular response, for example, cell survival, cellular differentiation,  
5 or cell proliferation.

In yet another embodiment, an assay of the invention is a cell-free assay comprising contacting a TETRA protein or biologically-active portion thereof with a test compound and determining the ability of the test compound to bind to the TETRA protein or biologically-active portion thereof. Binding of the test compound to the TETRA protein can be determined  
10 either directly or indirectly as described above. In one such embodiment, the assay comprises contacting the TETRA protein or biologically-active portion thereof with a known compound which binds TETRA to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with a TETRA protein, wherein determining the ability of the test compound to interact with a TETRA protein  
15 comprises determining the ability of the test compound to preferentially bind to TETRA or biologically-active portion thereof as compared to the known compound.

In still another embodiment, an assay is a cell-free assay comprising contacting TETRA protein 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  
20 TETRA protein or biologically-active portion thereof. Determining the ability of the test compound to modulate the activity of TETRA can be accomplished, for example, by determining the ability of the TETRA protein to bind to a TETRA 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 TETRA protein can be  
25 accomplished by determining the ability of the TETRA protein to further modulate a TETRA target molecule. For example, the catalytic/enzymatic activity of the target molecule on an appropriate substrate can be determined as described, supra.

In yet another embodiment, the cell-free assay comprises contacting the TETRA protein or biologically-active portion thereof with a known compound which binds TETRA  
30 protein to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with a TETRA protein, wherein determining the ability of the test compound to interact with a TETRA protein comprises

determining the ability of the TETRA protein to preferentially bind to or modulate the activity of a TETRA target molecule.

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

In more than one embodiment of the above assay methods of the invention, it can be desirable to immobilize either TETRA protein 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 TETRA protein, or interaction of TETRA protein 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 microtiter plates, test tubes, and micro-centrifuge tubes. In one embodiment, a fusion protein can be provided that adds a domain that allows one or both of the proteins to be bound to a matrix. For example, GST-TETRA fusion proteins or GST-target fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, MO) or glutathione derivatized microtiter plates, that are then combined with the test compound or the test compound and either the non-adsorbed target protein or TETRA protein, and the mixture is incubated under conditions conducive to complex formation (e.g., at physiological conditions for salt and pH). Following incubation, the beads or microtiter plate wells are washed to remove any unbound components, the matrix immobilized in the case of beads, complex determined either directly or indirectly, for example, as described, supra. Alternatively, the complexes can be dissociated from the matrix, and the level of TETRA protein binding or activity determined using standard techniques.

Other techniques for immobilizing proteins on matrices can also be used in the screening assays of the invention. For example, either the TETRA protein or its target

molecule can be immobilized utilizing conjugation of biotin and streptavidin. Biotinylated TETRA protein or target molecules can be prepared from biotin-NHS (N-hydroxy-succinimide) using techniques well-known within 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 TETRA protein or target molecules, but which do not interfere with binding of the TETRA protein to its target molecule, can be derivatized to the wells of the plate, and unbound target or TETRA protein 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 TETRA protein or target molecule, as well as enzyme-linked assays that rely on detecting an enzymatic activity associated with the TETRA protein or target molecule.

In another embodiment, modulators of TETRA protein expression are identified in a method wherein a cell is contacted with a candidate compound and the expression of TETRA mRNA or protein in the cell is determined. The level of expression of TETRA mRNA or protein in the presence of the candidate compound is compared to the level of expression of TETRA mRNA or protein in the absence of the candidate compound. The candidate compound can then be identified as a modulator of TETRA mRNA or protein expression based upon this comparison. For example, when expression of TETRA mRNA or protein is greater (i.e., statistically significantly greater) in the presence of the candidate compound than in its absence, the candidate compound is identified as a stimulator of TETRA mRNA or protein expression. Alternatively, when expression of TETRA 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 TETRA mRNA or protein expression. The level of TETRA mRNA or protein expression in the cells can be determined by methods described herein for detecting TETRA mRNA or protein.

In yet another aspect of the invention, the TETRA proteins can be used as "bait proteins" in a two-hybrid assay or three hybrid assay (see, e.g., U.S. Patent 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. *Biotechniques* 14: 920-924; Iwabuchi, et al., 1993. *Oncogene* 8: 1693-1696; and Brent WO 94/10300), to identify other proteins that bind to or interact with TETRA ("TETRA-binding proteins" or "TETRA-bp") and modulate TETRA activity. Such TETRA-binding

proteins are also likely to be involved in the propagation of signals by the TETRA proteins as, for example, upstream or downstream elements of the TETRA pathway.

The two-hybrid system is based on the modular nature of most transcription factors, which consist of separable DNA-binding and activation domains. Briefly, the assay utilizes two different DNA constructs. In one construct, the gene that codes for TETRA is fused to a gene encoding the DNA binding domain of a known transcription factor (e.g., GAL-4). In the other construct, a DNA sequence, from a library of DNA sequences, that encodes an unidentified protein ("prey" or "sample") is fused to a gene that codes for the activation domain of the known transcription factor. If the "bait" and the "prey" proteins are able to interact, in vivo, forming a TETRA-dependent complex, the DNA-binding and activation domains of the transcription factor are brought into close proximity. This proximity allows transcription of a reporter gene (e.g., LacZ) that is operably linked to a transcriptional regulatory site responsive to the transcription factor. Expression of the reporter gene can be detected and cell colonies containing the functional transcription factor can be isolated and used to obtain the cloned gene that encodes the protein which interacts with TETRA.

The invention further pertains to novel agents identified by the aforementioned screening assays and uses thereof for treatments as described herein.

### **Detection Assays**

Portions or fragments of the cDNA sequences identified herein (and the corresponding complete gene sequences) can be used in numerous ways as polynucleotide reagents. By way of example, and not of limitation, 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. Some of these applications are described in the subsections, below.

### **Chromosome Mapping**

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. This process is called chromosome mapping. Accordingly, portions or fragments of the TETRA sequences, SEQ ID NO: 1, or fragments or derivatives thereof, can be used to map the location of the TETRA

genes; respectively, on a chromosome. The mapping of the TETRA sequences to chromosomes is an important first step in correlating these sequences with genes associated with disease.

Briefly, TETRA genes can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp in length) from the TETRA sequences. Computer analysis of the TETRA sequences 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 TETRA sequences will yield an amplified fragment.

Somatic cell hybrids are prepared by fusing somatic cells from different mammals (e.g., human and mouse cells). As hybrids of human and mouse cells grow and divide, they gradually lose human chromosomes in random order, but retain the mouse chromosomes. By using media in which mouse cells cannot grow, because they lack a particular enzyme, but in which human cells can, the one human chromosome that contains the gene encoding the needed enzyme will be retained. By using various media, panels of hybrid cell lines can be established. Each cell line in a panel contains either a single human chromosome or a small number of human chromosomes, and a full set of mouse chromosomes, allowing easy mapping of individual genes to specific human chromosomes. See, e.g., D'Eustachio, et al., 1983. *Science* 220: 919-924. Somatic cell hybrids containing only fragments of human chromosomes can also be produced by using human chromosomes with translocations and deletions.

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 TETRA sequences to design oligonucleotide primers, sub-localization can be achieved with panels of fragments from specific chromosomes.

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.

Chromosome spreads can be made using cells whose division has been blocked in metaphase by a chemical like colcemid that disrupts the mitotic spindle. The chromosomes can be treated briefly with trypsin, and then stained with Giemsa. A pattern of light and dark bands develops

on each chromosome, so that the chromosomes can be identified individually. The FISH technique can be used with a DNA sequence as short as 500 or 600 bases. However, clones larger than 1,000 bases have a higher likelihood of binding to a unique chromosomal location with sufficient signal intensity for simple detection. Preferably 1,000 bases, and more  
5 preferably 2,000 bases, will suffice to get good results at a reasonable amount of time. For a review of this technique, see, Verma, et al., HUMAN CHROMOSOMES: A MANUAL OF BASIC TECHNIQUES (Pergamon Press, New York 1988).

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  
10 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.

Once a sequence has been mapped to a precise chromosomal location, the physical  
15 position of the sequence on the chromosome can be correlated with genetic map data. Such data are found, e.g., in 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.  
20 Nature, 325: 783-787.

Moreover, differences in the DNA sequences between individuals affected and unaffected with a disease associated with the TETRA gene, 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  
25 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.

### 30 **Tissue Typing**

The TETRA sequences of the invention can also be used to identify individuals from minute biological samples. 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. The sequences of the invention are useful as additional DNA markers for RFLP  
5 (“restriction fragment length polymorphisms,” described in U.S. Patent No. 5,272,057).

Furthermore, the sequences of the 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 TETRA sequences described herein can be used to prepare two PCR primers from the 5'- and 3'-termini of the sequences. These primers can then be used  
10 to amplify an individual's DNA and subsequently sequence it.

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 invention can be used to obtain such identification sequences from individuals and from tissue. The TETRA sequences  
15 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 of about once per each 500 bases. Much of the allelic variation is due to single nucleotide polymorphisms (SNPs), which include restriction fragment length polymorphisms  
20 (RFLPs).

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 can comfortably provide  
25 positive individual identification with a panel of perhaps 10 to 1,000 primers that each yield a noncoding amplified sequence of 100 bases. If predicted coding sequences, such as those in SEQ ID NO:1, are used, a more appropriate number of primers for positive individual identification would be 500-2,000.

### **Predictive Medicine**

30 The 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 invention relates to diagnostic assays for determining TETRA protein and/or nucleic acid expression as well as TETRA activity, 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 TETRA expression or activity. Such diseases include those related to signal transduction, control of cell adhesion, regulation of cell growth and proliferation (including development and oncogenesis) and mobility (including the ability to suppress metastatic potential). The invention also provides for prognostic (or predictive) assays for determining whether an individual is at risk of developing a disorder associated with TETRA protein, nucleic acid expression or activity. For example, mutations in a TETRA gene 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 TETRA protein, nucleic acid expression, or biological activity.

Another aspect of the invention provides methods for determining TETRA protein, nucleic acid expression or activity 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.)

Yet another aspect of the invention pertains to monitoring the influence of agents (e.g., drugs, compounds) on the expression or activity of TETRA in clinical trials.

These and other agents are described in further detail in the following sections.

## **Diagnostic Assays**

An exemplary method for detecting the presence or absence of TETRA 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 TETRA protein or nucleic acid (e.g., mRNA, genomic DNA) that encodes TETRA protein such that the presence of TETRA is detected in the biological sample. An agent for detecting TETRA mRNA or genomic DNA is a labeled nucleic acid probe capable of hybridizing to TETRA mRNA or genomic DNA. The



nucleic acid probe can be, for example, a full-length TETRA nucleic acid, such as the nucleic acid of SEQ ID NO:1 or a portion thereof, such as an oligonucleotide of at least 15, 30, 50, 100, 250 or 500 nucleotides in length and sufficient to specifically hybridize under stringent conditions to TETRA mRNA or genomic DNA. Other suitable probes for use in the  
5 diagnostic assays of the invention are described herein.

An agent for detecting TETRA protein is an antibody capable of binding to TETRA protein, 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')<sub>2</sub>) can be used. The term "labeled", with regard to the probe or antibody, is intended to encompass  
10 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  
15 streptavidin. The term "biological 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 TETRA mRNA, protein, or genomic DNA in a biological sample in vitro as well as in vivo. For example, in vitro techniques for detection of TETRA mRNA include Northern hybridizations and in situ  
20 hybridizations. In vitro techniques for detection of TETRA protein include enzyme linked immunosorbent assays (ELISAs), Western blots, immunoprecipitations, and immunofluorescence. In vitro techniques for detection of TETRA genomic DNA include Southern hybridizations. Furthermore, in vivo techniques for detection of TETRA protein include introducing into a subject a labeled anti-TETRA antibody. For example, the antibody  
25 can be labeled with a radioactive marker whose presence and location in a subject can be detected by standard imaging techniques. 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  
30 means from a subject.

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 TETRA protein, mRNA, or genomic DNA, such that the presence of

TETRA protein, mRNA or genomic DNA is detected in the biological sample, and comparing the presence of TETRA protein, mRNA or genomic DNA in the control sample with the presence of TETRA protein, mRNA or genomic DNA in the test sample.

The invention also encompasses kits for detecting the presence of TETRA in a biological sample. For example, the kit can comprise: a labeled compound or agent capable of detecting TETRA protein or mRNA in a biological sample; means for determining the amount of TETRA in the sample; and means for comparing the amount of TETRA in the sample with a standard. The compound or agent can be packaged in a suitable container. The kit can further comprise instructions for using the kit to detect TETRA protein or nucleic acid.

## 10 **Prognostic Assays**

The diagnostic methods described herein can furthermore be utilized to identify subjects having or at risk of developing a disease or disorder associated with aberrant TETRA expression or activity. Such diseases include those related to signal transduction, control of cell adhesion, regulation of cell growth and proliferation (including development and oncogenesis) and mobility (including the ability to suppress metastatic potential). 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 TETRA protein, nucleic acid expression or activity. Alternatively, the prognostic assays can be utilized to identify a subject having or at risk for developing a disease or disorder. Thus, the invention provides a method for identifying a disease or disorder associated with aberrant TETRA expression or activity in which a test sample is obtained from a subject and TETRA protein or nucleic acid (e.g., mRNA, genomic DNA) is detected, wherein the presence of TETRA protein or nucleic acid is diagnostic for a subject having or at risk of developing a disease or disorder associated with aberrant TETRA expression or activity. 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.

Furthermore, the prognostic assays described herein can be used to determine whether a subject can be administered an agent (e.g., an agonist, antagonist, peptidomimetic, protein, peptide, nucleic acid, small molecule, or other drug candidate) to treat a disease or disorder associated with aberrant TETRA expression or activity. For example, such methods can be used to determine whether a subject can be effectively treated with an agent for a disorder.

Thus, the invention provides methods for determining whether a subject can be effectively treated with an agent for a disorder associated with aberrant TETRA expression or activity in which a test sample is obtained and TETRA protein or nucleic acid is detected (e.g., wherein the presence of TETRA protein or nucleic acid is diagnostic for a subject that can be administered the agent to treat a disorder associated with aberrant TETRA expression or activity).

The methods of the invention can also be used to detect genetic lesions in a TETRA gene, thereby determining if a subject with the lesioned gene is at risk for a disorder characterized by aberrant cell proliferation and/or differentiation. In various embodiments, the methods include detecting, in a sample of cells from the subject, the presence or absence of a genetic lesion characterized by at least one of an alteration affecting the integrity of a gene encoding a TETRA-protein, or the misexpression of the TETRA gene. For example, such genetic lesions can be detected by ascertaining the existence of at least one of: (i) a deletion of one or more nucleotides from a TETRA gene; (ii) an addition of one or more nucleotides to a TETRA gene; (iii) a substitution of one or more nucleotides of a TETRA gene, (iv) a chromosomal rearrangement of a TETRA gene; (v) an alteration in the level of a messenger RNA transcript of a TETRA gene, (vi) aberrant modification of a TETRA gene, such as of the methylation pattern of the genomic DNA, (vii) the presence of a non-wild-type splicing pattern of a messenger RNA transcript of a TETRA gene, (viii) a non-wild-type level of a TETRA protein, (ix) allelic loss of a TETRA gene, and (x) inappropriate post-translational modification of a TETRA protein. As described herein, there are a large number of assay techniques known in the art which can be used for detecting lesions in a TETRA gene. A preferred biological sample is a peripheral blood leukocyte sample isolated by conventional means from a subject. However, any biological sample containing nucleated cells can be used, including, for example, buccal mucosal cells.

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. Patent 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 the TETRA-gene (see, Abravaya, et al., 1995. *Nucl. 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 that specifically hybridize to a TETRA gene under conditions such that hybridization and amplification of the TETRA 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  
5 that PCR and/or LCR can be desirable to use as a preliminary amplification step in conjunction with any of the techniques used for detecting mutations described herein.

Alternative amplification methods include: self sustained sequence replication (see, Guatelli, et al., 1990. Proc. Natl. Acad. Sci. USA 87: 1874-1878), transcriptional amplification system (see, Kwoh, et al., 1989. Proc. Natl. Acad. Sci. USA 86: 1173-1177); Q Replicase  
10 (see, Lizardi, et al., 1988. BioTechnology 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.

In an alternative embodiment, mutations in a TETRA gene from a sample cell can be  
15 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. Patent No.  
20 5,493,531) can be used to score for the presence of specific mutations by development or loss of a ribozyme cleavage site.

In other embodiments, genetic mutations in TETRA can be identified by hybridizing a sample and control nucleic acids, e.g., DNA or RNA, to high-density arrays containing  
25 hundreds or thousands of oligonucleotide probes. See, e.g., Cronin, et al., 1996. Human Mutation 7: 244-255; Kozal, et al., 1996. Nat. Med. 2: 753-759. For example, genetic mutations in TETRA 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.  
30 This step allows the identification of point mutations. This 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.

In yet another embodiment, any of a variety of sequencing reactions known in the art can be used to directly sequence the TETRA gene and detect mutations by comparing the  
5 sequence of the sample TETRA 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  
10 procedures can be utilized when performing the diagnostic assays (see, e.g., Naeve, et al., 1995. Biotechniques 19: 448), including sequencing by mass spectrometry (see, e.g., PCT International Publication No. WO 94/16101; Cohen, et al., 1996. Adv. Chromatography 36: 127-162; and Griffin, et al., 1993. Appl. Biochem. Biotechnol. 38: 147-159).

Other methods for detecting mutations in the TETRA gene include methods in which protection from cleavage agents is used to detect mismatched bases in RNA/RNA or  
15 RNA/DNA heteroduplexes. See, e.g., Myers, et al., 1985. Science 230: 1242. In general, the art technique of "mismatch cleavage" starts by providing heteroduplexes of formed by hybridizing (labeled) RNA or DNA containing the wild-type TETRA sequence with potentially mutant RNA or DNA obtained from a tissue sample. The double-stranded duplexes are treated with an agent that cleaves single-stranded regions of the duplex such as  
20 which will exist due to basepair mismatches between the control and sample strands. For instance, RNA/DNA duplexes can be treated with RNase and DNA/DNA hybrids treated with S1 nuclease to enzymatically digesting the mismatched regions. 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  
25 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 an embodiment, the control DNA or RNA can be labeled for detection.

In still another embodiment, the mismatch cleavage reaction employs one or more  
30 proteins that recognize mismatched base pairs in double-stranded DNA (so called "DNA mismatch repair" enzymes) in defined systems for detecting and mapping point mutations in TETRA 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. See, e.g., Hsu, et al., 1994. *Carcinogenesis* 15: 1657-1662. According to an exemplary embodiment, a probe based on a TETRA sequence, e.g., a wild-type TETRA sequence, is hybridized to a cDNA or other DNA product from a test cell(s). The duplex is  
5 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. Patent No. 5,459,039.

In other embodiments, alterations in electrophoretic mobility will be used to identify mutations in TETRA genes. For example, single strand conformation polymorphism (SSCP) can be used to detect differences in electrophoretic mobility between mutant and wild type  
10 nucleic acids. See, e.g., Orita, et al., 1989. *Proc. Natl. Acad. Sci. USA*: 86: 2766; Cotton, 1993. *Mutat. Res.* 285: 125-144; Hayashi, 1992. *Genet. Anal. Tech. Appl.* 9: 73-79. Single-stranded DNA fragments of sample and control TETRA nucleic acids will be denatured and allowed to renature. The secondary structure of single-stranded nucleic acids varies according to sequence, the resulting alteration in electrophoretic mobility enables the detection of even a  
15 single base change. The DNA fragments can be labeled or detected with labeled probes. The sensitivity of the assay can be enhanced by using RNA (rather than DNA), in which the secondary structure is more sensitive to a change in sequence. In one embodiment, the subject method utilizes heteroduplex analysis to separate double stranded heteroduplex molecules on the basis of changes in electrophoretic mobility. See, e.g., Keen, et al., 1991. *Trends Genet.* 7:  
20 5.

In yet another embodiment, the movement of mutant or wild-type fragments in polyacrylamide gels containing a gradient of denaturant is assayed using denaturing gradient gel electrophoresis (DGGE). See, e.g., 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  
25 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. See, e.g., Rosenbaum and Reissner, 1987. *Biophys. Chem.* 265: 12753.

Examples of other techniques for detecting point mutations include, but are not limited  
30 to, selective oligonucleotide hybridization, selective amplification, or selective primer extension. For example, oligonucleotide primers can be prepared in which the known mutation is placed centrally and then hybridized to target DNA under conditions that permit

hybridization only if a perfect match is found. See, e.g., 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.

Alternatively, allele specific amplification technology that depends on selective PCR amplification can 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; see, e.g., Gibbs, et al., 1989. Nucl. Acids Res. 17: 2437-2448) or at the extreme 3'-terminus of one primer where, under appropriate conditions, mismatch can prevent, or reduce polymerase extension (see, e.g., Prossner, 1993. Tibtech. 11: 238). In addition it can be desirable to introduce a novel restriction site in the region of the mutation to create cleavage-based detection. See, e.g., 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. See, e.g., 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'-terminus 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.

The methods described herein can be performed, for example, by utilizing pre-packaged diagnostic kits comprising at least one probe nucleic acid or antibody reagent described herein, which can be conveniently used, e.g., in clinical settings to diagnose patients exhibiting symptoms or family history of a disease or illness involving a TETRA gene. Furthermore, any cell type or tissue, preferably peripheral blood leukocytes, in which TETRA is expressed can be utilized in the prognostic assays described herein. However, any biological sample containing nucleated cells can be used, including, for example, buccal mucosal cells.

### **Pharmacogenomics**

Agents, or modulators that have a stimulatory or inhibitory effect on TETRA activity (e.g., TETRA gene expression), as identified by a screening assay described herein, can be administered to individuals to treat (prophylactically or therapeutically) disorders (e.g., cancer or immune disorders) associated with aberrant TETRA activity. Such diseases include those related to signal transduction, control of cell adhesion, regulation of cell growth and

proliferation (including development and oncogenesis) and mobility (including the ability to suppress metastatic potential). In conjunction with such treatment, the pharmacogenomics (i.e., the study of the relationship between an individual's genotype and that individual's response to a foreign compound or drug) of the individual can 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 TETRA protein, expression of TETRA nucleic acid, or mutation content of TETRA genes in an individual can be determined to thereby select appropriate agent(s) for therapeutic or prophylactic treatment of the individual.

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., Eichelbaum, 1996. *Clin. Exp. Pharmacol. Physiol.*, 23: 983-985; Linder, 1997. *Clin. Chem.*, 43: 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 (altered drug action) or genetic conditions transmitted as single factors altering the way the body acts on drugs (altered drug metabolism). These pharmacogenetic conditions can occur either as rare defects or as polymorphisms. For example, glucose-6-phosphate dehydrogenase (G6PD) deficiency is a common inherited enzymopathy in which the main clinical complication is hemolysis after ingestion of oxidant drugs (anti-malarials, sulfonamides, analgesics, nitrofurans) and consumption of fava beans.

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, PM show no therapeutic response, as demonstrated for the analgesic effect of codeine mediated by its CYP2D6-formed metabolite morphine. At 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.

Thus, the activity of TETRA protein, expression of TETRA nucleic acid, or mutation content of TETRA genes 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 TETRA modulator, such as a modulator identified by one of the exemplary screening assays described herein.

### **Monitoring of Effects During Clinical Trials**

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

By way of example, and not of limitation, genes, including TETRA, that are modulated in cells by treatment with an agent (e.g., compound, drug or small molecule) that modulates

TETRA activity (e.g., identified in a screening assay as 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 TETRA and other genes implicated in the disorder. The levels of gene expression (i.e., a gene  
5 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 TETRA or other genes. In this manner, the gene expression pattern can serve as a marker, indicative of the physiological response of the cells to the agent. Accordingly, this response state can be determined before,  
10 and at various points during, treatment of the individual with the agent.

In one embodiment, the invention provides a method for monitoring the effectiveness of treatment of a subject with an agent (e.g., an agonist, antagonist, protein, peptide, peptidomimetic, nucleic acid, small molecule, or other drug candidate identified by the screening assays described herein) comprising the steps of (i) obtaining a pre-administration  
15 sample from a subject prior to administration of the agent; (ii) detecting the level of expression of a TETRA protein, mRNA, or genomic DNA in the preadministration sample; (iii) obtaining one or more post-administration samples from the subject; (iv) detecting the level of expression or activity of the TETRA protein, mRNA, or genomic DNA in the post-administration samples; (v) comparing the level of expression or activity of the TETRA  
20 protein, mRNA, or genomic DNA in the pre-administration sample with the TETRA protein, mRNA, or genomic DNA 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 can be desirable to increase the expression or activity of TETRA to higher levels than detected, i.e., to increase the effectiveness of the agent. Alternatively, decreased  
25 administration of the agent can be desirable to decrease expression or activity of TETRA to lower levels than detected, i.e., to decrease the effectiveness of the agent.

### **Methods of Treatment**

The invention provides for both prophylactic and therapeutic methods of treating a  
30 subject at risk of (or susceptible to) a disorder or having a disorder associated with aberrant TETRA expression or activity. Such diseases include those related to signal transduction,

control of cell adhesion, regulation of cell growth and proliferation (including development and oncogenesis) and mobility (including the ability to suppress metastatic potential). These methods of treatment will be discussed more fully, below.

### **Disease and Disorders**

5 Diseases and disorders that are characterized by increased (relative to a subject not suffering from the disease or disorder) levels or biological activity can be treated with  
Therapeutics that antagonize (i.e., reduce or inhibit) activity. Therapeutics that antagonize  
activity can be administered in a therapeutic or prophylactic manner. Therapeutics that can be  
utilized include, but are not limited to: (i) an aforementioned peptide, or analogs, derivatives,  
10 fragments or homologs thereof; (ii) antibodies to an aforementioned peptide; (iii) nucleic acids  
encoding an aforementioned peptide; (iv) administration of antisense nucleic acid and nucleic  
acids that are "dysfunctional" (i.e., due to a heterologous insertion within the coding sequences  
of coding sequences to an aforementioned peptide) that are utilized to "knockout"  
endogenous function of an aforementioned peptide by homologous recombination (see, e.g.,  
15 Capecchi, 1989. Science 244: 1288-1292); or (v) modulators ( i.e., inhibitors, agonists and  
antagonists, including additional peptide mimetic of the invention or antibodies specific to a  
peptide of the invention) that alter the interaction between an aforementioned peptide and its  
binding partner.

20 Diseases and disorders that are characterized by decreased (relative to a subject not suffering from the disease or disorder) levels or biological activity can be treated with  
Therapeutics that increase (i.e., are agonists to) activity. Therapeutics that upregulate activity  
can be administered in a therapeutic or prophylactic manner. Therapeutics that can be utilized  
include, but are not limited to, an aforementioned peptide, or analogs, derivatives, fragments  
or homologs thereof; or an agonist that increases bioavailability.

25 Increased or decreased levels can be readily detected by quantifying peptide and/or RNA, by  
obtaining a patient tissue sample (e.g., from biopsy tissue) and assaying it in vitro for RNA or  
peptide levels, structure and/or activity of the expressed peptides (or mRNAs of an  
aforementioned peptide). Methods that are well-known within the art include, but are not  
limited to, immunoassays (e.g., by Western blot analysis, immunoprecipitation followed by  
30 sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis, immunocytochemistry, etc.)

and/or hybridization assays to detect expression of mRNAs (e.g., Northern assays, dot blots, in situ hybridization, and the like).

### **Prophylactic Methods**

In one aspect, the invention provides a method for preventing, in a subject, a disease or condition associated with an aberrant TETRA expression or activity, by administering to the subject an agent that modulates TETRA expression or at least one TETRA activity. Subjects at risk for a disease that is caused or contributed to by aberrant TETRA expression or activity 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 TETRA aberrancy, such that a disease or disorder is prevented or, alternatively, delayed in its progression. Depending upon the type of TETRA aberrancy, for example, a TETRA agonist or TETRA antagonist agent can be used for treating the subject. The appropriate agent can be determined based on screening assays described herein. The prophylactic methods of the invention are further discussed in the following subsections.

### **Therapeutic Methods**

Another aspect of the invention pertains to methods of modulating TETRA expression or activity 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 TETRA protein activity associated with the cell. An agent that modulates TETRA protein activity can be an agent as described herein, such as a nucleic acid or a protein, a naturally-occurring cognate ligand of a TETRA protein, a peptide, a TETRA peptidomimetic, or other small molecule. In one embodiment, the agent stimulates one or more TETRA protein activity. Examples of such stimulatory agents include active TETRA protein and a nucleic acid molecule encoding TETRA that has been introduced into the cell. In another embodiment, the agent inhibits one or more TETRA protein activity. Examples of such inhibitory agents include antisense TETRA nucleic acid molecules and anti-TETRA 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 invention provides methods of treating an individual afflicted with a disease or disorder characterized by aberrant expression or activity of a TETRA protein or nucleic acid molecule. 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., up-regulates or down-regulates) TETRA expression or activity. In another embodiment, the method involves administering a TETRA protein or nucleic acid molecule as therapy to compensate for reduced or aberrant TETRA expression or activity.

Stimulation of TETRA activity is desirable in situations in which TETRA is abnormally downregulated and/or in which increased TETRA activity is likely to have a beneficial effect. One example of such a situation is where a subject has a disorder characterized by aberrant cell proliferation and/or differentiation (e.g., cancer or immune associated disorders). Another example of such a situation is where the subject has a gestational disease (e.g., preclampsia).

### **Determination of the Biological Effect of the Therapeutic**

In various embodiments of the invention, suitable in vitro or in vivo assays are performed to determine the effect of a specific Therapeutic and whether its administration is indicated for treatment of the affected tissue.

In various specific embodiments, in vitro assays can be performed with representative cells of the type(s) involved in the patient's disorder, to determine if a given Therapeutic exerts the desired effect upon the cell type(s). Compounds for use in therapy can be tested in suitable animal model systems including, but not limited to rats, mice, chicken, cows, monkeys, rabbits, and the like, prior to testing in human subjects. Similarly, for in vivo testing, any of the animal model system known in the art can be used prior to administration to human subjects.

### **Prophylactic and Therapeutic Uses of the Compositions of the Invention**

The TETRA nucleic acids and proteins of the invention are useful in potential prophylactic and therapeutic applications implicated in a variety of disorders including, but not limited to: those involving development, differentiation, and activation of thymic immune cells; in pathologies related to spermatogenesis and male infertility; diagnosis of several human neoplasias; in diseases or pathologies of cells in blood circulation such as red blood cells and platelets; various immunological disorders and/or pathologies; autoimmune and inflammatory diseases; cardiovascular diseases; metabolic diseases; cancer growth and

metastasis; viral infections, cancer therapy, acute lymphoblastic leukemia; gliomas; neurologic diseases; neurodegenerative disorders; Alzheimer's Disease; Parkinson's Disorder; and hematopoietic disorders.

5 As an example, a cDNA encoding the TETRA protein of the invention can be useful in gene therapy, and the protein can be useful when administered to a subject in need thereof. By way of non-limiting example, the compositions of the invention will have efficacy for treatment of patients suffering from the above mentioned disorders.

10 Both the novel nucleic acid encoding the TETRA protein, and the TETRA protein of the invention, or fragments thereof, may also be useful in diagnostic applications, wherein the presence or amount of the nucleic acid or the protein are to be assessed. A further use could be as an anti-bacterial molecule (i.e., some peptides have been found to possess anti-bacterial properties). These materials are further useful in the generation of antibodies which immunospecifically-bind to the novel substances of the invention for use in therapeutic or diagnostic methods. Antibodies prepared that bind extracellular epitopes on 5.04-y0c0-187.1  
15 protein may be useful as therapeutics that block cell-cell adhesion or block or stimulate apoptosis signaling. Applications of these therapeutic antibodies are foreseen in autoimmune and inflammatory diseases, cardiovascular diseases, metabolic diseases, and cancer growth and metastasis

### **Other Embodiments**

20 It is to be understood that while the invention has been described in conjunction with the detailed description thereof, the foregoing description is intended to illustrate and not limit the scope of the invention, which is defined by the scope of the appended claims. Other aspects, advantages, and modifications are within the scope of the following claims.

## WHAT IS CLAIMED IS:

1. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of:
  - (a) a mature form of an amino acid sequence given by SEQ ID NO:2;
  - (b) a variant of a mature form of an amino acid sequence given by SEQ ID NO: 2, wherein one or more amino acid residues in said variant differs from the amino acid sequence of said mature form, provided that said variant differs in no more than 15% of the amino acid residues from the amino acid sequence of said mature form;
  - (c) an amino acid sequence given by SEQ ID NO: 2; and
  - (d) a variant of an amino acid sequence given by SEQ ID NO: 2, wherein one or more amino acid residues in said variant differs from the amino acid sequence of said mature form, provided that said variant differs in no more than 15% of amino acid residues from said amino acid sequence.
2. The polypeptide of claim 1, wherein said polypeptide comprises the amino acid sequence of a naturally-occurring allelic variant of an amino acid sequence given by SEQ ID NO: 2.
3. The polypeptide of claim 2, wherein said allelic variant comprises an amino acid sequence that is the translation of a nucleic acid sequence differing by a single nucleotide from a nucleic acid sequence given by SEQ ID NO:1.
4. The polypeptide of claim 1, wherein the amino acid sequence of said variant comprises a conservative amino acid substitution.
5. An isolated nucleic acid molecule comprising a nucleic acid sequence encoding a polypeptide comprising an amino acid sequence selected from the group consisting of:
  - (a) a mature form of an amino acid sequence given by SEQ ID NO: 2;
  - (b) a variant of a mature form of an amino acid sequence given by SEQ ID NO: 2, wherein one or more amino acid residues in said variant differs from the amino acid

sequence of said mature form, provided that said variant differs in no more than 15% of the amino acid residues from the amino acid sequence of said mature form;

(c) an amino acid sequence given by SEQ ID NO: 2;

(d) a variant of an amino acid sequence given by SEQ ID NO: 2, wherein one or more amino acid residues in said variant differs from the amino acid sequence of said mature form, provided that said variant differs in no more than 15% of amino acid residues from said amino acid sequence;

(e) a nucleic acid fragment encoding at least a portion of a polypeptide comprising an amino acid sequence given by SEQ ID NO: 2, or a variant of said polypeptide, wherein one or more amino acid residues in said variant differs from the amino acid sequence of said mature form, provided that said variant differs in no more than 15% of amino acid residues from said amino acid sequence; and

(f) a nucleic acid molecule comprising the complement of (a), (b), (c), (d) or (e).

6. The nucleic acid molecule of claim 5, wherein the nucleic acid molecule comprises the nucleotide sequence of a naturally-occurring allelic nucleic acid variant.
7. The nucleic acid molecule of claim 5, wherein the nucleic acid molecule encodes a polypeptide comprising the amino acid sequence of a naturally-occurring polypeptide variant.
8. The nucleic acid molecule of claim 5, wherein the nucleic acid molecule differs by a single nucleotide from a nucleic acid sequence given by SEQ ID NO: 1.
9. The nucleic acid molecule of claim 5, wherein said nucleic acid molecule comprises a nucleotide sequence selected from the group consisting of

(a) a nucleotide sequence given by SEQ ID NO: 1;

(b) a nucleotide sequence differing by one or more nucleotides from a nucleotide sequence given by SEQ ID NO: 1, provided that no more than 20% of the nucleotides differ from said nucleotide sequence;



- (c) a nucleic acid fragment of (a); and
  - (d) a nucleic acid fragment of (b).
10. The nucleic acid molecule of claim 5, wherein said nucleic acid molecule hybridizes under stringent conditions to a nucleotide sequence given by SEQ ID NO: 1, or a complement of said nucleotide sequence.
11. The nucleic acid molecule of claim 5, wherein the nucleic acid molecule comprises a nucleotide sequence selected from the group consisting of
- (a) a first nucleotide sequence comprising a coding sequence differing by one or more nucleotide sequences from a coding sequence encoding said amino acid sequence, provided that no more than 20% of the nucleotides in the coding sequence in said first nucleotide sequence differ from said coding sequence;
  - (b) an isolated second polynucleotide that is a complement of the first polynucleotide; and
  - (c) a nucleic acid fragment of (a) or (b).
12. A vector comprising the nucleic acid molecule of claim 11.
13. The vector of claim 12, further comprising a promoter operably-linked to said nucleic acid molecule.
14. A cell comprising the vector of claim 12.
15. An antibody that immunospecifically-binds to the polypeptide of claim 1.
16. The antibody of claim 15, wherein said antibody is a monoclonal antibody.
17. The antibody of claim 15, wherein the antibody is a humanized antibody.
18. A method for determining the presence or amount of the polypeptide of claim 1 in a sample, the method comprising:
- (a) providing the sample;

- (b) contacting the sample with an antibody that binds immunospecifically to the polypeptide; and
  - (c) determining the presence or amount of antibody bound to said polypeptide,thereby determining the presence or amount of polypeptide in said sample.
- 19. A method for determining the presence or amount of the nucleic acid molecule of claim 5 in a sample, the method comprising:
  - (a) providing the sample;
  - (b) contacting the sample with a probe that binds to said nucleic acid molecule; and
  - (c) determining the presence or amount of the probe bound to said nucleic acid molecule,thereby determining the presence or amount of the nucleic acid molecule in said sample.
- 20. A method of identifying an agent that binds to the polypeptide of claim 1, the method comprising:
  - (a) contacting said polypeptide with said agent; and
  - (b) determining whether said agent binds to said polypeptide.
- 21. A method for identifying an agent that modulates the expression or activity of the polypeptide of claim 1, the method comprising:
  - (a) providing a cell expressing said polypeptide;
  - (b) contacting the cell with said agent; and
  - (c) determining whether the agent modulates expression or activity of said polypeptide,

whereby an alteration in expression or activity of said peptide indicates said agent modulates expression or activity of said polypeptide.

22. A method for modulating the activity of the polypeptide of claim 1, the method comprising contacting a cell sample expressing the polypeptide of said claim with a compound that binds to said polypeptide in an amount sufficient to modulate the activity of the polypeptide.
23. A method of treating or preventing a TETRA-associated disorder, said method comprising administering to a subject in which such treatment or prevention is desired the polypeptide of claim 1 in an amount sufficient to treat or prevent said TETRA-associated disorder in said subject.
24. The method of claim 23, wherein said subject is a human.
25. A method of treating or preventing a TETRA-associated disorder, said method comprising administering to a subject in which such treatment or prevention is desired the nucleic acid of claim 5 in an amount sufficient to treat or prevent said TETRA-associated disorder in said subject.
26. The method of claim 25, wherein said subject is a human.
27. A method of treating or preventing a TETRA-associated disorder, said method comprising administering to a subject in which such treatment or prevention is desired the antibody of claim 15 in an amount sufficient to treat or prevent said TETRA-associated disorder in said subject.
28. The method of claim 15, wherein the subject is a human.
29. A pharmaceutical composition comprising the polypeptide of claim 1 and a pharmaceutically-acceptable carrier.
30. A pharmaceutical composition comprising the nucleic acid molecule of claim 5 and a pharmaceutically-acceptable carrier.
31. A pharmaceutical composition comprising the antibody of claim 15 and a pharmaceutically-acceptable carrier.

32. A kit comprising in one or more containers, the pharmaceutical composition of claim 29.
33. A kit comprising in one or more containers, the pharmaceutical composition of claim 30.
34. A kit comprising in one or more containers, the pharmaceutical composition of claim 31.
35. The use of a therapeutic in the manufacture of a medicament for treating a syndrome associated with a human disease, the disease selected from a TETRA-associated disorder, wherein said therapeutic is selected from the group consisting of a TETRA polypeptide, a TETRA nucleic acid, and a TETRA antibody.
36. A method for screening for a modulator of activity or of latency or predisposition to a TETRA-associated disorder, said method comprising:
- (a) administering a test compound to a test animal at increased risk for a TETRA-associated disorder, wherein said test animal recombinantly expresses the polypeptide of claim 1;
  - (b) measuring the activity of said polypeptide in said test animal after administering the compound of step (a); and
  - (c) comparing the activity of said protein in said test animal with the activity of said polypeptide in a control animal not administered said polypeptide, wherein a change in the activity of said polypeptide in said test animal relative to said control animal indicates the test compound is a modulator of latency of or predisposition to a TETRA-associated disorder.
37. The method of claim 36, wherein said test animal is a recombinant test animal that expresses a test protein transgene or expresses said transgene under the control of a promoter at an increased level relative to a wild-type test animal, and wherein said promoter is not the native gene promoter of said transgene.
38. A method for determining the presence of or predisposition to a disease associated with altered levels of the polypeptide of claim 1 in a first mammalian subject, the method comprising:

(a) measuring the level of expression of the polypeptide in a sample from the first mammalian subject; and

(b) comparing the amount of said polypeptide in the sample of step (a) to the amount of the polypeptide present in a control sample from a second mammalian subject known not to have, or not to be predisposed to, said disease,

wherein an alteration in the expression level of the polypeptide in the first subject as compared to the control sample indicates the presence of or predisposition to said disease.

39. A method for determining the presence of or predisposition to a disease associated with altered levels of the nucleic acid molecule of claim 5 in a first mammalian subject, the method comprising:

(a) measuring the amount of the nucleic acid in a sample from the first mammalian subject; and

(b) comparing the amount of said nucleic acid in the sample of step (a) to the amount of the nucleic acid present in a control sample from a second mammalian subject known not to have or not be predisposed to, the disease;

wherein an alteration in the level of the nucleic acid in the first subject as compared to the control sample indicates the presence of or predisposition to the disease.

40. A method of treating a pathological state in a mammal, the method comprising administering to the mammal a polypeptide in an amount that is sufficient to alleviate the pathological state, wherein the polypeptide is a polypeptide having an amino acid sequence at least 95% identical to a polypeptide comprising an amino acid sequence of at least one of SEQ ID NO: 2, or a biologically active fragment thereof.

41. A method of treating a pathological state in a mammal, the method comprising administering to the mammal the antibody of claim 15 in an amount sufficient to alleviate the pathological state.