US 20020086830A1

(19) **United States**

(12) **Patent Application Publication** (10) **Pub. No.: US 2002/0086830 A1**

Hillman et al. (43) **Pub. Date: Jul. 4, 2002**

(54) **HUMAN APOPTOSIS-RELATED CALCIUM-BINDING PROTEIN**

(75) Inventors: **Jennifer L. Hillman**, San Jose, CA (US); **Surya K. Goli**, Sunnyvale, CA (US)

Correspondence Address:
INCYTE GENOMICS, INC.
PATENT DEPARTMENT
3160 Porter Drive
Palo Alto, CA 94304 (US)

(73) Assignee: **Incyte Pharmaceuticals, Inc.**

(21) Appl. No.: **09/969,836**

(22) Filed: **Oct. 1, 2001**

Related U.S. Application Data

(60) Continuation of application No. 09/328,687, filed on Jun. 9, 1999, which is a division of application No. 09/094,212, filed on Jun. 9, 1998, now patented, which is a division of application No. 08/766,605, filed on Dec. 12, 1996, now patented.

Publication Classification

(51) **Int. Cl.**[7] **A01K 67/00**; A61K 38/17; C07H 21/04; C12N 9/64; C12P 21/02; C12N 5/06; C12N 1/21

(52) **U.S. Cl.** **514/12**; 435/69.1; 435/226; 435/320.1; 435/325; 435/252.3; 800/8; 536/23.2

(57) **ABSTRACT**

The present invention provides a human apoptosis-related calcium-binding protein (HARC) and polynucleotides which identify and encode HARC. The invention also provides genetically engineered expression vectors and host cells comprising the nucleic acid sequences encoding HARC and a method for producing HARC. The invention also provides for agonists, antibodies, or antagonists specifically binding HARC, and their use, in the prevention and treatment of diseases associated with expression of HARC. Additionally, the invention provides for the use of antisense molecules to polynucleotides encoding HARC for the treatment of diseases associated with the expression of HARC. The invention also provides diagnostic assays which utilize the polynucleotide, or fragments or the complement thereof, and antibodies specifically binding HARC.

```
                9                  18                  27                  36                  45                  54
5'   T GGA CAC CCC AAT CCT GGN ATG TTC CCC TNT GGA ACT CCA GGA GNA CCA TAT
                                  M   F   P   X   G   T   P   G   X   P   Y

               63                  72                  81                  90                  99                 108
GGC GNT GCA GCT CCC GGG GGC CCC TAT GGT CAG CCA CCT CCA AGT TCC TAC GGT
G   X   A   A   P   G   G   P   Y   G   Q   P   P   P   S   S   Y   G

              117                 126                 135                 144                 153                 162
GNC CAG CAG CCT GGG NTT TAT GGA CAG GGT GGC GCC CCT CCC AAT GTG GAT CCT
X   Q   Q   P   G   X   Y   G   Q   G   G   A   P   P   N   V   D   P

              171                 180                 189                 198                 207                 216
GAG GCC TAC TCC TGG TTC CAG TCG GTG GAC TCA GAT CAC AGT GGC TAT ATC TCC
E   A   Y   S   W   F   Q   S   V   D   S   D   H   S   G   Y   I   S

              225                 234                 243                 252                 261                 270
ATG AAG GAG CTA AAG CAG GCC CTG GTC AAC TGC AAT TGG TCT TCA TTC AAT GAT
M   K   E   L   K   Q   A   L   V   N   C   N   W   S   S   F   N   D

              279                 288                 297                 306                 315                 324
GAG ACC TGC CTC ATG ATG ATA AAC ATG TTT GAC AAG ACC AAG TCA GGC CGC ATC
E   T   C   L   M   M   I   N   M   F   D   K   T   K   S   G   R   I

              333                 342                 351                 360                 369                 378
GAT GTC TAC GGC TTC TCA GCC CTG TGG AAA TTC ATC CAG CAG TGG AAG AAC CTC
D   V   Y   G   F   S   A   L   W   K   F   I   Q   Q   W   K   N   L
```

FIGURE 1A

```
                387                 396                 405                 414                 423                 432
TTC CAG CAG TAT GAC CGG GAC CGC TCG GGC TCC ATT AGC TAC ACA GAG CTG CAG
F   Q   Q   Y   D   R   D   R   S   G   S   I   S   Y   T   E   L   Q

                441                 450                 459                 468                 477                 486
CAA GCT CTG TCC CAA ATG GGC TAC AAC CTG AGC CCC CAG TTC ACC CAG CTT CTG
Q   A   L   S   Q   M   G   Y   N   L   S   P   Q   F   T   Q   L   L

                495                 504                 513                 522                 531                 540
GTC TCC CGC TAC TGC CCA CGC TCT GCC AAT CCT GCC ATG CAG CTT GAC CGC TTC
V   S   R   Y   C   P   R   S   A   N   P   A   M   Q   L   D   R   F

                549                 558                 567                 576                 585                 594
ATC CAG GTG TGC ACC CAG CTG CAG GTG CTG ACA GAG GCC TTC CGG GAG AAG GAC
I   Q   V   C   T   Q   L   Q   V   L   T   E   A   F   R   E   K   D

                603                 612                 621                 630                 639                 648
ACA GCT GTA CAA GGC AAC ATC CGG CTC AGC TTC GAG GAC TTC GTC ACC ATG ACA
T   A   V   Q   G   N   I   R   L   S   F   E   D   F   V   T   M   T

                657                 666                 675                 684                 693                 702
GCT TCT CGG ATG CTA TGA CCC AAC CAT CTG TGG AGA GTG GAG TGC ACC AGG GAC
A   S   R   M   L   *

                711                 720                 729                 738                 747                 756
CTT TCC TGG CTT CTT AGA GTG AGA GAA GTA TGT GGA CAT CTC TTC TTT TCC TGT

                765                 774
CCC TCT AGA AGA ACA TTC TCC C 3'
P   S   R   R   T   F   S
```

FIGURE 1B

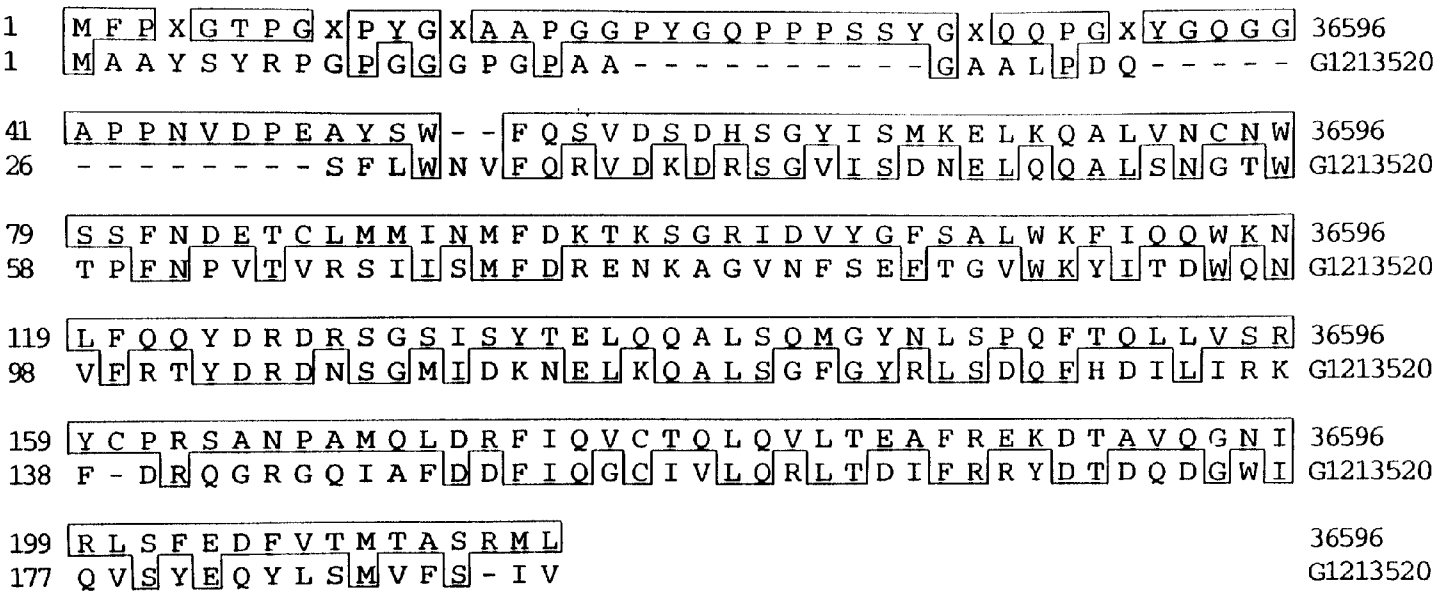

1 MFPXGTPGXPYGXAAPGGPYGQPPPSSYGXQQPGXYGQGG 36596
1 MAAYSYRPGPGGGPGPAA----------GAALPDQ----- G1213520
41 APPNVDPEAYSW--FQSVDSDHSGYISMKELKQALVNCNW 36596
26 ---------SFLWNVFQRVDKDRSGVISDNELQQALSNGTW G1213520
79 SSFNDETCLMMINMFDKTKSGRIDVYGFSALWKFIQQWKN 36596
58 TPFNPVTVRSIISMFDRENKAGVNFSEFTGVWKYITDWQN G1213520
119 LFQQYDRDRSGSISYTELQQALSQMGYNLSPQFTQLLVSR 36596
98 VFRTYDRDNSGMIDKNELKQALSGFGYRLSDQFHDILIRK G1213520
159 YCPRSANPAMQLDRFIQVCTQLQVLTEAFREKDTAVQGNI 36596
138 F-DRQGRGQIAFDDFIQGCIVLQRLTDIFRRYDTDQDGWI G1213520
199 RLSFEDFVTMTASRML 36596
177 QVSYEQYLSMVFS-IV G1213520

FIGURE 2

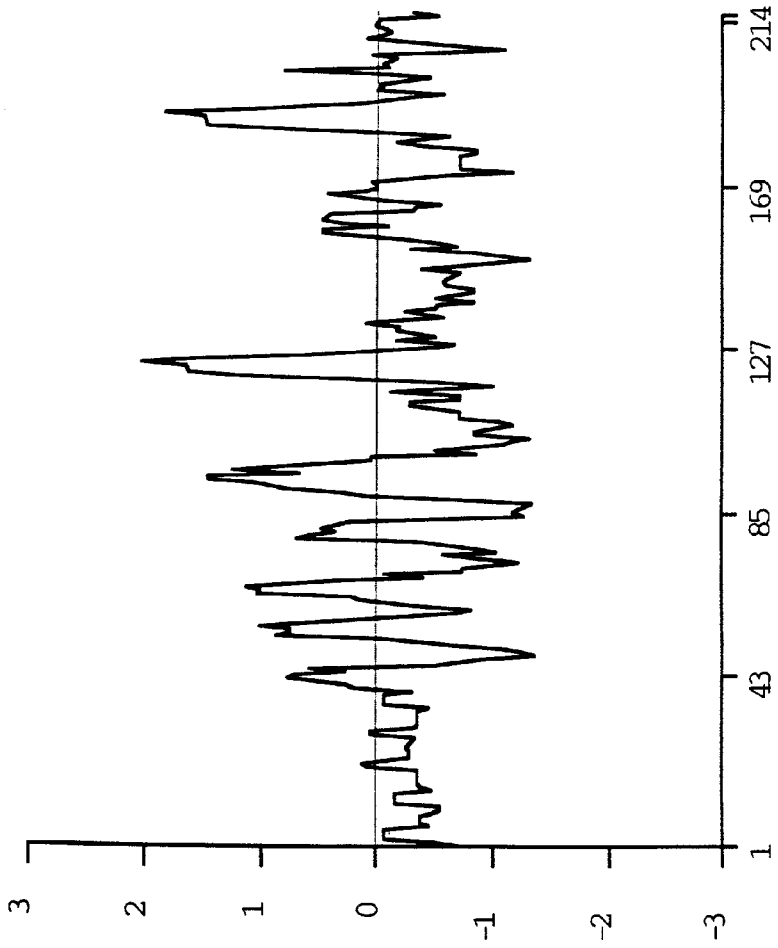


FIGURE 3

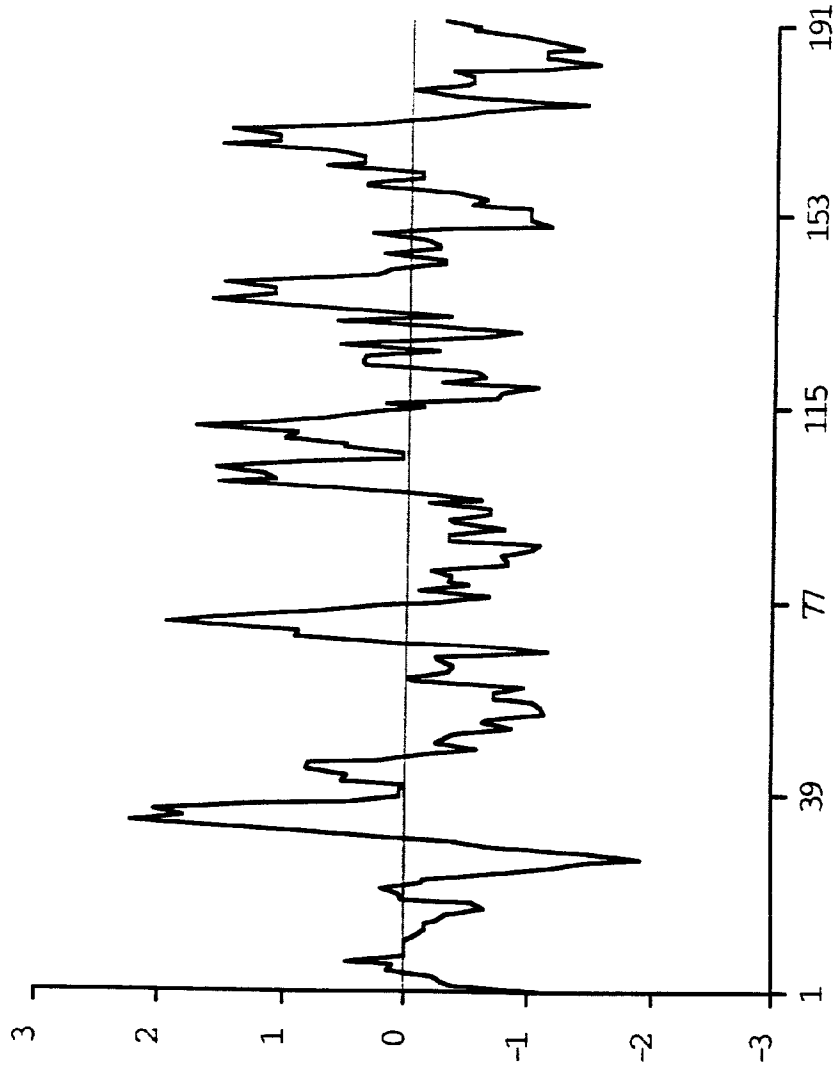


FIGURE 4

Library	Lib Description	Abun	Pct Abun
TESTNOT04	testis, 37 M	1	0.1075
HNT3AZT01	hNT2 cell line, teratocarcinoma, treated AZ	1	0.0685
ADRENOT03	adrenal gland, 17 M	2	0.0682
HIPONON01	brain, hippocampus, 72 F, NORM	2	0.0675
ENDCNOT02	endothelial cells, dermal microvascular, 30 F	1	0.0604
LIVRTUT01	liver tumor, metastasis, 51 F	2	0.0518
PROSNOT18	prostate, 58 M	2	0.0513
MUSCNOT01	muscle, skeletal	1	0.0445
COLNNOT08	colon, 60 M	1	0.0426
HUVENOB01	HUVEC endothelial cell line, control	1	0.0420
BRAINON01	brain, 26 M, NORM	1	0.0417
CRBLNOT01	brain, cerebellum, 69 M	2	0.0391
TONSNOT01	tonsil, hyperplasia, 6 M	1	0.0339
PLACNOT02	placenta, fetal F	2	0.0336
PROSNOT20	prostate, 65 M, match to PROSTUT12	1	0.0336
BSTMNON02	brain stem, 72 M, NORM	1	0.0319
LUNGTUT03	lung tumor, 69 M, match to LUNGNOT15	2	0.0318
THYRTUT03	thyroid tumor, benign, 17 M	1	0.0276
BRSTNOM01	breast, F, NORM, WM	1	0.0264
STOMFET01	stomach, fetal F	1	0.0255
HIPONOT01	brain, hippocampus, 72 F	1	0.0239
THYRNOT01	thyroid, 64 F	1	0.0229
CARDFEM01	heart, fetal, NORM, WM	2	0.0223
PANCTUT02	pancreatic tumor, carcinoma, 45 F	1	0.0202

FIGURE 5A

BRSTTUT01	breast tumor, 55 F, match to BRSTNOT02	2	0.0189
KERANOT02	keratinocytes, primary cell line, 30 F	1	0.0182
SINTBST01	small intestine, ileum, Crohn's, 18 F	1	0.0168
PLACNOM02	placenta, neonatal F, NORM, WM	3	0.0167
UTRSNOT02	uterus, 34 F	1	0.0167
LUNGFEM01	lung, fetal, NORM, WM	1	0.0148
BRSTNOT03	breast, 54 F, match to BRSTTUT02	1	0.0147
COLNFET02	colon, fetal F	1	0.0143
THYRNOT03	thyroid tumor, adenoma, 28 F	1	0.0138
BRAITUT01	brain tumor, oligoastrocytoma, 50 F	1	0.0134
TESTNOT03	testis, 37 M	1	0.0129
BLADTUT04	bladder tumor, 60 M, match to BLADNOT05	1	0.0127
SPLNFET02	spleen, fetal M	1	0.0126
PROSTUT04	prostate tumor, 57 M, match to PROSNOT06	1	0.0117
PROSNOT06	prostate, 57 M, match to PROSTUT04	1	0.0114
KIDNNOT05	kidney, neonatal F	1	0.0106
EOSIHET02	eosinophils, hypereosinophilia, 48 M	1	0.0105
LUNGAST01	lung, asthma, 17 M	1	0.0094
UCMCL5T01	mononuclear cells, treated IL-5	1	0.0084
LIVSFEM02	liver/spleen, fetal M, NORM, WM	3	0.0079

FIGURE 5B

HUMAN APOPTOSIS-RELATED CALCIUM-BINDING PROTEIN

[0001] This application is a continuation application of U.S. Ser. No. 09/328,687, filed Jun. 9, 1999, which is a divisional application of U.S. Ser. No. 09/094,212, filed Jun. 9, 1998, now U.S. Pat. No. 5,935,931, which is a divisional application of U.S. Ser. No. 08/766,605, filed Dec. 12, 1996, now U.S. Pat. No. 5,763,220, the contents each of which is hereby incorporated by reference, and each application entitled "Human Apoptosis-Related Calcium-Binding Protein".

FIELD OF THE INVENTION

[0002] This invention relates to nucleic acid and amino acid sequences of a novel human apoptosis-related calcium-binding protein and to the use of these sequences in the diagnosis, prevention, and treatment of diseases associated with decreased or increased apoptosis.

BACKGROUND OF THE INVENTION

[0003] Normal development, growth, and homeostasis in multicellular organisms require a careful balance between the production and destruction of cells in tissues throughout the body. Cell division is a carefully coordinated process with numerous checkpoints and control mechanisms. These mechanisms are designed to regulate DNA replication and to prevent inappropriate or excessive proliferation. In contrast, programmed cell death is the genetically controlled process by which unneeded or damaged cells can be eliminated without causing the tissue destruction and inflammatory responses that are often associated with acute injury and necrosis.

[0004] The term "apoptosis" was first used by Kerr, J. F. et al. (1972; Br. J. Cancer 26:239-257) to describe the morphological changes that characterize cells undergoing programmed cell death. Apoptotic cells have a shrunken appearance with an altered membrane lipid content and highly condensed nuclei. Apoptotic cells are rapidly phagocytosed by neighboring cells or macrophages without leaking their potentially damaging contents into the surrounding tissue.

[0005] The processes and mechanisms regulating apoptosis are highly conserved throughout the phylogenetic tree. Indeed, much of our current knowledge about apoptosis is derived from studies of the nematode, *Caenorhabditis elegans* and the fruit fly, *Drosophila melanogaster* (See for example, Steller, H. (1995) Science 267:1445-1449, and references therein). Dysregulation of apoptosis has recently been recognized as a significant factor in the pathogenesis of human disease. For example, inappropriate cell survival can cause or contribute to many diseases such as cancer, autoimmune diseases, and inflammatory diseases. In contrast, increased apoptosis can cause immunodeficiency diseases such as AIDS, neurodegenerative disorders, and myelodysplastic syndromes (reviewed by Thompson, C. B. (1995) Science 267:1456-1462).

[0006] A variety of ligands and their cellular receptors, enzymes, tumor suppressors, viral gene products, pharmacological agents, and inorganic ions have important positive or negative roles in regulating and implementing the apoptotic destruction of a cell (Steller, H., supra; Thompson, C. B., supra). Although many different extra- and intracellular signals can trigger apoptosis (cf. Raff, M. C. (1992) Nature 356:397-400, Raff, M. C. et al. (1993) Science 262:695-700, and Steller, H. and M. E. Grether, (1994) Neuron 13:1269-1274), these signals probably all converge on a common mechanism that ultimately causes the cell to die.

[0007] The mouse apoptosis-linked gene 2 (ALG-2) is one of six clones that has been identified using an in vitro "death trap" model for apoptosis. In this model, 3DO hybridoma cells are transformed with cloned cDNAs and then induced to undergo programmed cell death by cross-linking their T cell receptors. Certain cDNAs afford protection to the transformed 3DO cells either by expressing a protein which inhibits apoptosis or by expressing an antisense RNA which blocks the synthesis of a required protein (Vito, P. et al. (1996) Science 271:521-525).

[0008] Analysis of the recovered ALG-2 clone indicates that it is of the latter type, thereby identifying ALG-2 as a protein that is essential for programmed cell death. Northern blot analyses detect a single, ~1.3 KB ALG-2 transcript that is constitutively expressed in all adult mouse tissues; expression is highest in the thymus and liver and lowest in the testes and skeletal muscles. Constitutive expression in normal tissue implies that ALG-2 is probably in an inactive state until the apoptosis pathway is triggered. Apoptosis induced by T cell receptor cross-linking, Fas-Fas ligand interactions, and glucocorticoid treatment all depend on functional ALG-2, which is likely to be a part of the common pathway leading to cell death (Vito, P. et al., supra).

[0009] The ALG-2 sequence predicts an acidic protein of 191 amino acids with 2 EF-hand Ca^{2+}-binding domains; both EF-hand domains are required for Ca^{2+}binding. ALG-2 is the first Ca^{2+}-binding protein that has been shown to be directly required in the apoptosis pathway (Vito, P. et al., supra). A requirement for Ca^{2+}in apoptosis was previously suggested by studies showing its involvement in DNA cleavage (Hewish, D. R. and L. A. Burgoyne (1973) Biochem. Biophys. Res. Comm. 52:504-510). Other studies show that: 1) intracellular Ca^{2+}concentrations increase when apoptosis is triggered in thymocytes by either T cell receptor cross-linking or by glucocorticoids, and 2) cell death can be prevented by blocking this increase in intracellular Ca^{2+} (McConkey, D. J. et al. (1989) J. Immunol. 143:1801-1806; McConkey, D. J. et al. (1989) Arch. Biochem. Biophys. 269:365-370). Additional support for the role of Ca^{2+}in apoptosis comes from work on Fas-mediated cell death (Vignaux, F. et al. (1995) J. Exp. Med. 181:781-786; Oshimi, Y. and S. Miyazaki (1995) J. Immunol. 154:599-609).

[0010] The discovery of polynucleotides encoding human apoptosis-related calcium-binding protein, and the molecules themselves, provide a means to investigate the regulation of programmed cell death and apoptosis. Discovery of molecules related to mouse ALG-2 satisfies a need in the art by providing new diagnostic or therapeutic compositions useful in the detection, prevention, and treatment of cancer, autoimmune diseases, lymphoproliferative disorders, psoriasis, atherosclerosis, restenosis, AIDS, immunodeficiency diseases, ischemic injuries, neurodegenerative diseases, osteoporosis, myelodysplastic syndromes, toxin-induced diseases, cachexia, and viral infections.

SUMMARY OF THE INVENTION

[0011] The present invention features a novel human apoptosis-related calcium-binding protein hereinafter designated HARC and characterized as having similarity to mouse apoptosis-linked gene 2, ALG-2.

[0012] Accordingly, the invention features a substantially purified HARC having the amino acid sequence shown in SEQ ID NO: 1.

[0013] One aspect of the invention features isolated and substantially purified polynucleotides that encode HARC. In a particular aspect, the polynucleotide is the nucleotide sequence of SEQ ID NO: 2.

[0014] The invention also relates to a polynucleotide sequence comprising the complement of SEQ ID NO: 2 or variants thereof. In addition, the invention features polynucleotide sequences which hybridize under stringent conditions to SEQ ID NO: 2.

[0015] The invention additionally features nucleic acid sequences encoding polypeptides, oligonucleotides, peptide nucleic acids (PNA), fragments, portions or antisense molecules thereof, and expression vectors and host cells comprising polynucleotides that encode HARC. The present invention also features antibodies which bind specifically to HARC, and pharmaceutical compositions comprising substantially purified HARC. The invention also features the use of agonists and antagonists of HARC.

BRIEF DESCRIPTION OF THE FIGURES

[0016] **FIGS. 1A and 1B** show the amino acid sequence (SEQ ID NO: 1) and nucleic acid sequence (SEQ ID NO: 2) of HARC. The alignment was produced using MacDNASIS PRO™ software (Hitachi Software Engineering Co., Ltd., San Bruno, Calif.).

[0017] **FIG. 2** shows the amino acid sequence alignments between HARC (SEQ ID NO: 1) and mouse ALG-2 (GI1213520; SEQ ID NO: 3). The alignment was produced using the multisequence alignment program of LASERGENE software (DNASTAR Inc, Madison Wis.).

[0018] **FIG. 3** shows the hydrophobicity plot (MacDNASIS PRO software) for HARC, SEQ ID NO: 1; the positive X axis reflects amino acid position, and the negative Y axis, hydrophobicity.

[0019] **FIG. 4** shows the hydrophobicity plot for ALG-2, SEQ ID NO: 3.

[0020] **FIGS. 5A and 5B** show the northern analysis for SEQ ID NO: 2. The northern analysis was produced electronically using the LIFESEQ database (Incyte Pharmaceuticals, Inc., Palo Alto, Calif.).

DESCRIPTION OF THE INVENTION

[0021] Before the present proteins, nucleotide sequences, and methods are described, it is understood that this invention is not limited to the particular methodology, protocols, cell lines, vectors, and reagents described as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention which will be limited only by the appended claims.

[0022] It must be noted that as used herein and in the appended claims, the singular forms "a", "an", and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, reference to "a host cell" includes a plurality of such host cells, reference to the "antibody" is a reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.

[0023] Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, the preferred methods, devices, and materials are now described. All publications mentioned herein are incorporated herein by reference for the purpose of describing and disclosing the cell lines, vectors, and methodologies which are reported in the publications which might be used in connection with the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

DEFINITIONS

[0024] "Nucleic acid sequence" as used herein refers to an oligonucleotide, nucleotide, or polynucleotide, and fragments or portions thereof, and to DNA or RNA of genomic or synthetic origin which may be single- or double-stranded, and represent the sense or antisense strand. Similarly, "amino acid sequence" as used herein refers to an oligopeptide, peptide, polypeptide, or protein sequence, and fragments or portions thereof, and to naturally occurring or synthetic molecules.

[0025] Where "amino acid sequence" is recited herein to refer to an amino acid sequence of a naturally occurring protein molecule, "amino acid sequence" and like terms, such as "polypeptide" or "protein" are not meant to limit the amino acid sequence to the complete, native amino acid sequence associated with the recited protein molecule.

[0026] "Peptide nucleic acid", as used herein, refers to a molecule which comprises an oligomer to which an amino acid residue, such as lysine, and an amino group have been added. These small molecules, also designated anti-gene agents, stop transcript elongation by binding to their complementary strand of nucleic acid (Nielsen, P. E. et al. (1993) Anticancer Drug Des. 8:53-63).

[0027] HARC, as used herein, refers to the amino acid sequences of substantially purified HARC obtained from any species, particularly mammalian, including bovine, ovine, porcine, murine, equine, and preferably human, from any source whether natural, synthetic, semi-synthetic, or recombinant.

[0028] "Consensus", as used herein, refers to a nucleic acid sequence which has been resequenced to resolve uncalled bases, or which has been extended using XL-PCR kit (Perkin Elmer, Norwalk, Conn.) in the 5' and/or the 3' direction and resequenced, or which has been assembled from the overlapping sequences of more than one Incyte clone using the GELVIEW fragment assembly system (GCG, Madison, Wis.), or which has been both extended and assembled.

[0029] A "variant" of HARC, as used herein, refers to an amino acid sequence that is altered by one or more amino acids. The variant may have "conservative" changes, wherein a substituted amino acid has similar structural or chemical properties, e.g., replacement of leucine with isoleucine. More rarely, a variant may have "nonconservative" changes, e.g., replacement of a glycine with a tryptophan. Similar minor variations may also include amino acid deletions or insertions, or both. Guidance in determining which amino acid residues may be substituted, inserted, or deleted without abolishing biological or immunological activity may be found using computer programs well known in the art, for example, LASERGENE software.

[0030] A "deletion", as used herein, refers to a change in either amino acid or nucleotide sequence in which one or more amino acid or nucleotide residues, respectively, are absent.

[0031] An "insertion" or "addition", as used herein, refers to a change in an amino acid or nucleotide sequence resulting in the addition of one or more amino acid or nucleotide residues, respectively, as compared to the naturally occurring molecule.

[0032] A "substitution", as used herein, refers to the replacement of one or more amino acids or nucleotides by different amino acids or nucleotides, respectively.

[0033] The term "biologically active", as used herein, refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, "immunologically active" refers to the capability of the natural, recombinant, or synthetic HARC, or any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

[0034] The term "agonist", as used herein, refers to a molecule which, when bound to HARC, causes a change in HARC which modulates the activity of HARC. Agonists may include proteins, nucleic acids, carbohydrates, or any other molecules which bind to HARC.

[0035] The terms "antagonist" or "inhibitor", as used herein, refer to a molecule which, when bound to HARC, blocks or modulates the biological or immunological activity of HARC. Antagonists and inhibitors may include proteins, nucleic acids, carbohydrates, or any other molecules which bind to HARC.

[0036] The term "modulate", as used herein, refers to a change or an alteration in the biological activity of HARC. Modulation may be an increase or a decrease in protein activity, a change in binding characteristics, or any other change in the biological, functional or immunological properties of HARC.

[0037] The term "mimetic", as used herein, refers to a molecule, the structure of which is developed from knowledge of the structure of HARC or portions thereof and, as such, is able to effect some or all of the actions of ALG-2-like molecules.

[0038] The term "derivative", as used herein, refers to the chemical modification of a nucleic acid encoding HARC or the encoded HARC. Illustrative of such modifications would be replacement of hydrogen by an alkyl, acyl, or amino group. A nucleic acid derivative would encode a polypeptide which retains essential biological characteristics of the natural molecule.

[0039] The term "substantially purified", as used herein, refers to nucleic or amino acid sequences that are removed from their natural environment, isolated or separated, and are at least 60% free, preferably 75% free, and most preferably 90% free from other components with which they are naturally associated.

[0040] "Amplification" as used herein refers to the production of additional copies of a nucleic acid sequence and is generally carried out using polymerase chain reaction (PCR) technologies well known in the art (Dieffenbach, C. W. and G. S. Dveksler (1995) *PCR Primer, a Laboratory Manual,* Cold Spring Harbor Press, Plainview, N.Y.).

[0041] The term "hybridization", as used herein, refers to any process by which a strand of nucleic acid binds with a complementary strand through base pairing.

[0042] The term "hybridization complex", as used herein, refers to a complex formed between two nucleic acid sequences by virtue of the formation of hydrogen bonds between complementary G and C bases and between complementary A and T bases; these hydrogen bonds may be further stabilized by base stacking interactions. The two complementary nucleic acid sequences hydrogen bond in an antiparallel configuration. A hybridization complex may be formed in solution (e.g., C_ot or R_ot analysis) or between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., membranes, filters, chips, pins or glass slides to which cells have been fixed for in situ hybridization).

[0043] The terms "complementary" or "complementarity", as used herein, refer to the natural binding of polynucleotides under permissive salt and temperature conditions by base-pairing. For example, for the sequence "A-G-T" binds to the complementary sequence "T-C-A". Complementarity between two single-stranded molecules may be "partial", in which only some of the nucleic acids bind, or it may be complete when total complementarity exists between the single-stranded molecules. The degree of complementarity between nucleic acid strands has significant effects on the efficiency and strength of hybridization between nucleic acid strands. This is of particular importance in amplification reactions, which depend upon binding between nucleic acids strands.

[0044] The term "homology", as used herein, refers to a degree of complementarity. There may be partial homology or complete homology (i.e., identity). A partially complementary sequence is one that at least partially inhibits an identical sequence from hybridizing to a target nucleic acid; it is referred to using the functional term "substantially homologous." The inhibition of hybridization of the completely complementary sequence to the target sequence may be examined using a hybridization assay (Southern or northern blot, solution hybridization and the like) under conditions of low stringency. A substantially homologous sequence or probe will compete for and inhibit the binding (i.e., the hybridization) of a completely homologous sequence or probe to the target sequence under conditions of low stringency. This is not to say that conditions of low stringency are such that non-specific binding is permitted; low stringency conditions require that the binding of two sequences to one another be a specific (i.e., selective) interaction. The absence of non-specific binding may be tested by the use of a second target sequence which lacks

even a partial degree of complementarity (e.g., less than about 30% identity); in the absence of non-specific binding, the probe will not hybridize to the second non-complementary target sequence.

[0045] As known in the art, numerous equivalent conditions may be employed to comprise either low or high stringency conditions. Factors such as the length and nature (DNA, RNA, base composition) of the sequence, nature of the target (DNA, RNA, base composition, presence in solution or immobilization, etc.), and the concentration of the salts and other components (e.g., the presence or absence of formamide, dextran sulfate and/or polyethylene glycol) are considered and the hybridization solution may be varied to generate conditions of either low or high stringency different from, but equivalent to, the above listed conditions.

[0046] The term "stringent conditions", as used herein, is the "stringency" which occurs within a range from about Tm-5° C. (5° C. below the melting temperature (Tm) of the probe) to about 20° C. to 25° C. below Tm. As will be understood by those of skill in the art, the stringency of hybridization may be altered in order to identify or detect identical or related polynucleotide sequences.

[0047] The term "antisense", as used herein, refers to nucleotide sequences which are complementary to a specific DNA or RNA sequence. The term "antisense strand" is used in reference to a nucleic acid strand that is complementary to the "sense" strand. Antisense molecules may be produced by any method, including synthesis by ligating the gene(s) of interest in a reverse orientation to a viral promoter which permits the synthesis of a complementary strand. Once introduced into a cell, this transcribed strand combines with natural sequences produced by the cell to form duplexes. These duplexes then block either the further transcription or translation. In this manner, mutant phenotypes may be generated. The designation "negative" is sometimes used in reference to the antisense strand, and "positive" is sometimes used in reference to the sense strand.

[0048] The term "portion", as used herein, with regard to a protein (as in "a portion of a given protein") refers to fragments of that protein. The fragments may range in size from four amino acid residues to the entire amino acid sequence minus one amino acid. Thus, a protein "comprising at least a portion of the amino acid sequence of SEQ ID NO: 1" encompasses the full-length human HARC and fragments thereof.

[0049] "Transformation", as defined herein, describes a process by which exogenous DNA enters and changes a recipient cell. It may occur under natural or artificial conditions using various methods well known in the art. Transformation may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method is selected based on the host cell being transformed and may include, but is not limited to, viral infection, electroporation, lipofection, and particle bombardment. Such "transformed" cells include stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome. They also include cells which transiently express the inserted DNA or RNA for limited periods of time.

[0050] The term "antigenic determinant", as used herein, refers to that portion of a molecule that makes contact with a particular antibody (i.e., an epitope). When a protein or fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to a given region or three-dimensional structure on the protein; these regions or structures are referred to as antigenic determinants. An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

[0051] The terms "specific binding" or "specifically binding", as used herein, in reference to the interaction of an antibody and a protein or peptide, mean that the interaction is dependent upon the presence of a particular structure (i.e., the antigenic determinant or epitope) on the protein; in other words, the antibody is recognizing and binding to a specific protein structure rather than to proteins in general. For example, if an antibody is specific for epitope "A", the presence of a protein containing epitope A (or free, unlabeled A) in a reaction containing labeled "A" and the antibody will reduce the amount of labeled A bound to the antibody.

[0052] The term "sample", as used herein, is used in its broadest sense. A biological sample suspected of containing nucleic acid encoding HARC or fragments thereof may comprise a cell, chromosomes isolated from a cell (e.g., a spread of metaphase chromosomes), genomic DNA (in solution or bound to a solid support such as for Southern analysis), RNA (in solution or bound to a solid support such as for northern analysis), cDNA (in solution or bound to a solid support), an extract from cells or a tissue, and the like.

[0053] The term "correlates with expression of a polynucleotide", as used herein, indicates that the detection of the presence of ribonucleic acid that is similar to SEQ ID NO: 2 by northern analysis is indicative of the presence of mRNA encoding HARC in a sample and thereby correlates with expression of the transcript from the polynucleotide encoding the protein.

[0054] "Alterations" in the polynucleotide of SEQ ID NO: 2, as used herein, comprise any alteration in the sequence of polynucleotides encoding HARC including deletions, insertions, and point mutations that may be detected using hybridization assays. Included within this definition is the detection of alterations to the genomic DNA sequence which encodes HARC (e.g., by alterations in the pattern of restriction fragment length polymorphisms capable of hybridizing to SEQ ID NO: 2), the inability of a selected fragment of SEQ ID NO: 2 to hybridize to a sample of genomic DNA (e.g., using allele-specific oligonucleotide probes), and improper or unexpected hybridization, such as hybridization to a locus other than the normal chromosomal locus for the polynucleotide sequence encoding HARC (e.g., using fluorescent in situ hybridization (FISH) to metaphase chromosomes spreads).

[0055] As used herein, the term "antibody" refers to intact molecules as well as fragments thereof, such as Fab, $F(ab')_2$, and Fv, which are capable of binding the epitopic determinant. Antibodies that bind HARC polypeptides can be prepared using intact polypeptides or fragments containing small peptides of interest as the immunizing antigen. The polypeptide or peptide used to immunize an animal can be derived from the transition of RNA or synthesized chemically, and can be conjugated to a carrier protein, if desired.

Commonly used carriers that are chemically coupled to peptides include bovine serum albumin and thyroglobulin. The coupled peptide is then used to immunize the animal (e.g., a mouse, a rat, or a rabbit).

[0056] The term "humanized antibody", as used herein, refers to antibody molecules in which amino acids have been replaced in the non-antigen binding regions in order to more closely resemble a human antibody, while still retaining the original binding ability.

THE INVENTION

[0057] The invention is based on the discovery of a novel human apoptosis-related calcium-binding protein (HARC), the polynucleotides encoding HARC, and the use of these compositions for the diagnosis, prevention, and treatment of cancer, autoimmune diseases, lymphoproliferative disorders, psoriasis, atherosclerosis, restenosis, AIDS, immunodeficiency diseases, ischemic injuries, neurodegenerative diseases, osteoporosis, myelodysplastic syndromes, toxin-induced diseases, cachexia, and viral infections.

[0058] Nucleic acids encoding the human HARC of the present invention were first identified in Incyte Clone 036596 from the human umbilical cord vein endothelial cell line cDNA library (HUVENOB01) through a computer-generated search for amino acid sequence alignments. A consensus sequence, SEQ ID NO: 2, was derived from the following overlapping and/or extended nucleic acid sequences: Incyte Clones 036596 (HUVENOB01), 179146 (PLACNOB01), 334277 (EOSIHET02), 1304460 (PLACNOT02), and 1752846 (LIVRTUT01).

[0059] In one embodiment, the invention encompasses a polypeptide comprising the amino acid sequence of SEQ ID NO: 1, as shown in **FIG. 1**. HARC is 214 amino acids in length and has two potential EF-hand calcium-binding domains at residues 53-69 and 117-137. HARC has chemical and structural homology with ALG-2 (GI1213520; SEQ ID NO: 3). In particular, HARC and ALG-2 share 47% overall identity, and 58% identity in each of the two calcium-binding domains. As illustrated by **FIGS. 3 and 4**, HARC and ALG-2 have similar hydrophobicity plots, and both are very acidic proteins with isoelectric points of 5.26 and 5.01, respectively. Northern analysis (**FIG. 5**) shows the expression of this sequence in various libraries, 55% of which are from cancers, tissues associated with inflammatory diseases, fetal tissues, and cell lines. Of particular note is the expression of HARC in normal brain (5/44) and prostate (3/44) cDNA libraries.

[0060] The invention also encompasses HARC variants. A preferred HARC variant is one having at least 80%, and more preferably 90%, amino acid sequence similarity to the HARC amino acid sequence (SEQ ID NO: 1). A most preferred HARC variant is one having at least 95% amino acid sequence similarity to SEQ ID NO: 1.

[0061] The invention also encompasses polynucleotides which encode HARC. Accordingly, any nucleic acid sequence which encodes the amino acid sequence of HARC can be used to generate recombinant molecules which express HARC. In a particular embodiment, the invention encompasses the polynucleotide comprising the nucleic acid sequence of SEQ ID NO: 2 as shown in **FIGS. 1A and 1B**.

[0062] It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of nucleotide sequences encoding HARC, some bearing minimal homology to the nucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of nucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the nucleotide sequence of naturally occurring HARC, and all such variations are to be considered as being specifically disclosed.

[0063] Although nucleotide sequences which encode HARC and its variants are preferably capable of hybridizing to the nucleotide sequence of the naturally occurring HARC under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding HARC or its derivatives possessing a substantially different codon usage. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding HARC and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

[0064] The invention also encompasses production of DNA sequences, or portions thereof, which encode HARC and its derivatives, entirely by synthetic chemistry. After production, the synthetic sequence may be inserted into any of the many available expression vectors and cell systems using reagents that are well known in the art at the time of the filing of this application. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding HARC or any portion thereof.

[0065] Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed nucleotide sequences, and in particular, those shown in SEQ ID NO: 2, under various conditions of stringency. Hybridization conditions are based on the melting temperature (Tm) of the nucleic acid binding complex or probe, as taught in Wahl, G. M. and S. L. Berger (1987; Methods Enzymol. 152:399-407) and Kimmel, A. R. (1987; Methods Enzymol. 152:507-511), and may be used at a defined stringency.

[0066] Altered nucleic acid sequences encoding HARC which are encompassed by the invention include deletions, insertions, or substitutions of different nucleotides resulting in a polynucleotide that encodes the same or a functionally equivalent HARC. The encoded protein may also contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent HARC. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues as long as the biological activity of HARC is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid; positively charged amino acids may include lysine and arginine; and amino acids with uncharged polar head groups having similar hydrophilicity values may include leucine, isoleucine, and valine; glycine and alanine; asparagine and glutamine; serine and threonine; phenylalanine and tyrosine.

[0067] Also included within the scope of the present invention are alleles of the genes encoding HARC. As used herein, an "allele" or "allelic sequence" is an alternative form of the gene which may result from at least one mutation in the nucleic acid sequence. Alleles may result in altered mRNAs or polypeptides whose structure or function may or may not be altered. Any given gene may have none, one, or many allelic forms. Common mutational changes which give rise to alleles are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

[0068] Methods for DNA sequencing which are well known and generally available in the art may be used to practice any embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE (US Biochemical Corp, Cleveland, Ohio), Taq polymerase (Perkin Elmer), thermostable T7 polymerase (Amersham, Chicago, Ill.), or combinations of recombinant polymerases and proofreading exonucleases such as the amplification system (GIBCOIBRL, Gaithersburg, Md.). Preferably, the process is automated with machines such as the Hamilton MICROLAB 2200 (Hamilton, Reno, Nev.), Peltier thermal cycler (PTC200; M. J. Research, Watertown, Mass.) and the ABI 377 DNA sequencers (Perkin Elmer).

[0069] The nucleic acid sequences encoding HARC may be extended utilizing a partial nucleotide sequence and employing various methods known in the art to detect upstream sequences such as promoters and regulatory elements. For example, one method which may be employed, "restriction-site" PCR, uses universal primers to retrieve unknown sequence adjacent to a known locus (Sarkar, G. (1993) PCR Methods Applic. 2:318-322). In particular, genomic DNA is first amplified in the presence of primer to linker sequence and a primer specific to the known region. The amplified sequences are then subjected to a second round of PCR with the same linker primer and another specific primer internal to the first one. Products of each round of PCR are transcribed with an appropriate RNA polymerase and sequenced using reverse transcriptase.

[0070] Inverse PCR may also be used to amplify or extend sequences using divergent primers based on a known region (Triglia, T. et al. (1988) Nucleic Acids Res. 16:8186). The primers may be designed using OLIGO 4.06 Primer Analysis software (National Biosciences Inc., Plymouth, Minn.), or another appropriate program, to be 22-30 nucleotides in length, to have a GC content of 50% or more, and to anneal to the target sequence at temperatures about 68°-72° C. The method uses several restriction enzymes to generate a suitable fragment in the known region of a gene. The fragment is then circularized by intramolecular ligation and used as a PCR template.

[0071] Another method which may be used is capture PCR which involves PCR amplification of DNA fragments adjacent to a known sequence in human and yeast artificial chromosome DNA (Lagerstrom, M. et al. (1991) PCR Methods Applic. 1:111-119). In this method, multiple restriction enzyme digestions and ligations may also be used to place an engineered double-stranded sequence into an unknown portion of the DNA molecule before performing PCR.

[0072] Another method which may be used to retrieve unknown sequences is that of Parker, J. D. et al. (1991; Nuc. Acids Res. 19:3055-3060). Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries to walk in genomic DNA (Clontech, Palo Alto, Calif.). This process avoids the need to screen libraries and is useful in finding intron/exon junctions.

[0073] When screening for full-length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. Also, random-primed libraries are preferable, in that they will contain more sequences which contain the 5' regions of genes. Use of a randomly primed library may be especially preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into the 5' and 3' non-transcribed regulatory regions.

[0074] Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different fluorescent dyes (one for each nucleotide) which are laser activated, and detection of the emitted wavelengths by a charge coupled device camera. Output/light intensity may be converted to electrical signal using appropriate software (e.g. SEQUENCE NAVIGATOR and GENOTYPER, Perkin Elmer) and the entire process from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for the sequencing of small pieces of DNA which might be present in limited amounts in a particular sample.

[0075] In another embodiment of the invention, polynucleotide sequences or fragments thereof which encode HARC, or fusion proteins or functional equivalents thereof, may be used in recombinant DNA molecules to direct expression of HARC in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be produced and these sequences may be used to clone and express HARC.

[0076] As will be understood by those of skill in the art, it may be advantageous to produce HARC-encoding nucleotide sequences possessing non-naturally occurring codons. For example, codons preferred by a particular prokaryotic or eukaryotic host can be selected to increase the rate of protein expression or to produce a recombinant RNA transcript having desirable properties, such as a half-life which is longer than that of a transcript generated from the naturally occurring sequence.

[0077] The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter HARC encoding sequences for a variety of reasons, including but not limited to, alterations which modify the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, site-directed mutagenesis may be used to insert new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, or introduce mutations, and so forth.

[0078] In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding HARC may be ligated to a heterologous sequence to encode a fusion protein. For example, to screen peptide libraries for inhibitors of HARC activity, it may be useful to encode a chimeric HARC protein that can be recognized by a commercially available antibody. A fusion protein may also be engineered to contain a cleavage site located between the HARC encoding sequence and the heterologous protein sequence, so that HARC may be cleaved and purified away from the heterologous moiety.

[0079] In another embodiment, sequences encoding HARC may be synthesized, in whole or in part, using chemical methods well known in the art (see Caruthers, M. H. et al. (1980) Nuc. Acids Res. Symp. Ser. 7:215-223, Horn, T. et al. (1980) Nuc. Acids Res. Symp. Ser. 7:225-232). Alternatively, the protein itself may be produced using chemical methods to synthesize the amino acid sequence of HARC, or a portion thereof. For example, peptide synthesis can be performed using various solid-phase techniques (Roberge, J. Y. et al. (1995) Science 269:202-204) and automated synthesis may be achieved, for example, using the ABI 431A Peptide Synthesizer (Perkin Elmer).

[0080] The newly synthesized peptide may be substantially purified by preparative high performance liquid chromatography (e.g., Creighton, T. (1983) *Proteins, Structures and Molecular Principles,* W H Freeman and Co., New York, N.Y.). The composition of the synthetic peptides may be confirmed by amino acid analysis or sequencing (e.g., the Edman degradation procedure; Creighton, supra). Additionally, the amino acid sequence of HARC, or any part thereof, may be altered during direct synthesis and/or combined using chemical methods with sequences from other proteins, or any part thereof, to produce a variant polypeptide.

[0081] In order to express a biologically active HARC, the nucleotide sequences encoding HARC or functional equivalents, may be inserted into appropriate expression vector, i.e., a vector which contains the necessary elements for the transcription and translation of the inserted coding sequence.

[0082] Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding HARC and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. Such techniques are described in Sambrook, J. et al. (1989) *Molecular Cloning, A Laboratory Manual,* Cold Spring Harbor Press, Plainview, N.Y., and Ausubel, F. M. et al. (1989) *Current Protocols in Molecular Biology,* John Wiley & Sons, New York, N.Y.

[0083] A variety of expression vector/host systems may be utilized to contain and express sequences encoding HARC. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with virus expression vectors (e.g., baculovirus); plant cell systems transformed with virus expression vectors (e.g., cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems.

[0084] The "control elements" or "regulatory sequences" are those non-translated regions of the vector-enhancers, promoters, 5' and 3' untranslated regions--which interact with host cellular proteins to carry out transcription and translation. Such elements may vary in their strength and specificity. Depending on the vector system and host utilized, any number of suitable transcription and translation elements, including constitutive and inducible promoters, may be used. For example, when cloning in bacterial systems, inducible promoters such as the hybrid lacZ promoter of the BLUESCRIPT phagemid (Stratagene, La Jolla, Calif.) or PSPORT1 plasmid (GIBCO/BRL) and the like may be used. The baculovirus polyhedrin promoter may be used in insect cells. Promoters or enhancers derived from the genomes of plant cells (e.g., heat shock, RUBISCO; and storage protein genes) or from plant viruses (e.g., viral promoters or leader sequences) may be cloned into the vector. In mammalian cell systems, promoters from mammalian genes or from mammalian viruses are preferable. If it is necessary to generate a cell line that contains multiple copies of the sequence encoding HARC, vectors based on SV40 or EBV may be used with an appropriate selectable marker.

[0085] In bacterial systems, a number of expression vectors may be selected depending upon the use intended for HARC. For example, when large quantities of HARC are needed for the induction of antibodies, vectors which direct high level expression of fusion proteins that are readily purified may be used. Such vectors include, but are not limited to, the multifunctional *E. coli* cloning and expression vectors such as BLUESCRIPT phagemid (Stratagene), in which the sequence encoding HARC may be ligated into the vector in frame with sequences for the amino-terminal Met and the subsequent 7 residues of β-galactosidase so that a hybrid protein is produced; pIN vectors (Van Heeke, G. and S. M. Schuster (1989) J. Biol. Chem. 264:5503-5509); and the like. PGEX vectors (Promega, Madison, Wis.) may also be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to glutathione-agarose beads followed by elution in the presence of free glutathione. Proteins made in such systems may be designed to include heparin, thrombin, or factor XA protease cleavage sites so that the cloned polypeptide of interest can be released from the GST moiety at will.

[0086] In the yeast, *Saccharomyces cerevisiae,* a number of vectors containing constitutive or inducible promoters such as alpha factor, alcohol oxidase, and PGH may be used. For reviews, see Ausubel et al. (supra) and Grant et al. (1987) Methods Enzymol. 153:516-544.

[0087] In cases where plant expression vectors are used, the expression of sequences encoding HARC may be driven by any of a number of promoters. For example, viral promoters such as the 35S and 19S promoters of CaMV may be used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) EMBO J. 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used (Coruzzi, G. et al. (1984) EMBO J. 3:1671-1680; Broglie, R. et al. (1984) Science 224:838-843; and Winter, J. et al. (1991) Results Probl. Cell Differ. 17:85-105). These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. Such techniques are described in a number of generally available

reviews (see, for example, Hobbs, S. or Murry, L. E. in McGraw Hill *Yearbook of Science and Technology* (1992) McGraw Hill, New York, N.Y.; pp. 191-196.)

[0088] An insect system may also be used to express HARC. For example, in one such system, *Autographa californica* nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes in *Spodoptera frugiperda* cells or in Trichoplusia larvae. The sequences encoding HARC may be cloned into a non-essential region of the virus, such as the polyhedrin gene, and placed under control of the polyhedrin promoter. Successful insertion of HARC will render the polyhedrin gene inactive and produce recombinant virus lacking coat protein. The recombinant viruses may then be used to infect, for example, *S. frugiperda* cells or Trichoplusia larvae in which HARC may be expressed (Engelhard, E. K. et al. (1994) Proc. Nat. Acad. Sci. 91:3224-3227).

[0089] In mammalian host cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding HARC may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain a viable virus which is capable of expressing HARC in infected host cells (Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. 81:3655-3659). In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells.

[0090] Specific initiation signals may also be used to achieve more efficient translation of sequences encoding HARC. Such signals include the ATG initiation codon and adjacent sequences. In cases where sequences encoding HARC, its initiation codon, and upstream sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a portion thereof, is inserted, exogenous translational control signals including the ATG initiation codon should be provided. Furthermore, the initiation codon should be in the correct reading frame to ensure translation of the entire insert. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers which are appropriate for the particular cell system which is used, such as those described in the literature (Scharf, D. et al. (1994) Results Probl. Cell Differ. 20:125-162).

[0091] In addition, a host cell strain may be chosen for its ability to modulate the expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" form of the protein may also be used to facilitate correct insertion, folding and/or function. Different host cells such as CHO, HeLa, MDCK, HEK293, and W138, which have specific cellular machinery and characteristic mechanisms for such post-translational activities, may be chosen to ensure the correct modification and processing of the foreign protein.

[0092] For long-term, high-yield production of recombinant proteins, stable expression is preferred. For example, cell lines which stably express HARC may be transformed using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for 1-2 days in an enriched media before they are switched to selective media. The purpose of the selectable marker is to confer resistance to selection, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be proliferated using tissue culture techniques appropriate to the cell type.

[0093] Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase (Wigler, M. et al. (1977) Cell 11:223-32) and adenine phosphoribosyltransferase (Lowy, I. et al. (1980) Cell 22:817-23) genes which can be employed in tk^- or $aprt^-$ cells, respectively. Also, antimetabolite, antibiotic or herbicide resistance can be used as the basis for selection; for example, dhfr which confers resistance to methotrexate (Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. 77:3567-70); npt, which confers resistance to the aminoglycosides neomycin and (G-418 (Colbere-Garapin, F. et al (1981) J. Mol. Biol. 150:1-14) and als or pat, which confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively (Murry, supra). Additional selectable genes have been described, for example, trpB, which allows cells to utilize indole in place of tryptophan, or hisD, which allows cells to utilize histinol in place of histidine (Hartman, S. C. and R. C. Mulligan (1988) Proc. Natl. Acad. Sci. 85:8047-51). Recently, the use of visible markers has gained popularity with such markers as anthocyanins, β glucuronidase and its substrate GUS, and luciferase and its substrate luciferin, being widely used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system (Rhodes, C. A. et al. (1995) Methods Mol. Biol. 55:121-131).

[0094] Although the presence/absence of marker gene expression suggests that the gene of interest is also present, its presence and expression may need to be confirmed. For example, if the sequence encoding HARC is inserted within a marker gene sequence, recombinant cells containing sequences encoding HARC can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding HARC under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

[0095] Alternatively, host cells which contain the nucleic acid sequence encoding HARC and express HARC may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein.

[0096] The presence of polynucleotide sequences encoding HARC can be detected by DNA-DNA or DNA-RNA hybridization or amplification using probes or portions or fragments of polynucleotides encoding HARC. Nucleic acid amplification based assays involve the use of oligonucle-

otides or oligomers based on the sequences encoding HARC to detect transformants containing DNA or RNA encoding HARC. As used herein "oligonucleotides" or "oligomers" refer to a nucleic acid sequence of at least about 10 nucleotides and as many as about 60 nucleotides, preferably about 15 to 30 nucleotides, and more preferably about 20-25 nucleotides, which can be used as a probe or amplimer.

[0097] A variety of protocols for detecting and measuring the expression of HARC, using either polyclonal or monoclonal antibodies specific for the protein are known in the art. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on HARC is preferred, but a competitive binding assay may be employed. These and other assays are described, among other places, in Hampton, R. et al. (1990; *Serological Methods, a Laboratory Manual,* APS Press, St Paul, Minn.) and Maddox, D. E. et al. (1983; J. Exp. Med. 158:1211-1216).

[0098] A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding HARC include oligolabeling, nick translation, end-labeling or PCR amplification using a labeled nucleotide. Alternatively, the sequences encoding HARC, or any portions thereof may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits (Pharmacia & Upjohn, (Kalamazoo, Mich.); Promega (Madison Wis.); and U.S. Biochemical Corp., Cleveland, Ohio). Suitable reporter molecules or labels, which may be used, include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

[0099] Host cells transformed with nucleotide sequences encoding HARC may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a recombinant cell may be secreted or contained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode HARC may be designed to contain signal sequences which direct secretion of HARC through a prokaryotic or eukaryotic cell membrane. Other recombinant constructions may be used to join sequences encoding HARC to nucleotide sequence encoding a polypeptide domain which will facilitate purification of soluble proteins. Such purification facilitating domains include, but are not limited to, metal chelating peptides such as histidine-tryptophan modules that allow purification on immobilized metals, protein A domains that allow purification on immobilized immunoglobulin, and the domain utilized in the FLAGS extension/affinity purification system (Immunex Corp., Seattle, Wash.). The inclusion of cleavable linker sequences such as those specific for Factor XA or enterokinase (Invitrogen, San Diego, Calif.) between the purification domain and HARC may be used to facilitate purification. One such expression vector provides for expression of a fusion protein containing HARC and a nucleic acid encoding 6 histidine residues preceding a thioredoxin or an enterokinase cleavage site. The histidine residues facilitate purification on IMIAC (immobilized metal ion affinity chromatography as described in Porath, J. et al. (1992, Prot. Exp. Purif. 3: 263-281) while the enterokinase cleavage site provides a means for purifying HARC from the fusion protein. A discussion of vectors which contain fusion proteins is provided in Kroll, D. J. et al. (1993; DNA Cell Biol. 12:441-453).

[0100] In addition to recombinant production, fragments of HARC may be produced by direct peptide synthesis using solid-phase techniques (Merrifield J. (1963) J. Am. Chem. Soc. 85:2149-2154). Protein synthesis may be performed using manual techniques or by automation. Automated synthesis may be achieved, for example, using Applied Biosystems 431A Peptide Synthesizer (Perkin Elmer). Various fragments of HARC may be chemically synthesized separately and combined using chemical methods to produce the full length molecule.

[0101] THERAPEUTICS

[0102] Based on the chemical and structural homology between HARC (SEQ ID NO: 1) and mouse ALG-2 (SEQ ID NO: 3), HARC appears to play a role in the development of diseases related to decreased or increased apoptosis. Diseases arising from decreased apoptosis include, but are not limited to, cancer of the breast, prostate, and ovary as well as lymphomas and carcinomas, autoimmune disorders such as systemic lupus erythematosus, glomerulonephritis, Sjögren's syndrome, Graves disease, multiple sclerosis, rheumatoid arthritis, and diabetes, inflammatory diseases such as osteoarthritis, Crohn's disease, inflammatory bowel disease, and colitis, proliferative disorders such as atherosclerosis, restenosis, psoriasis, lymphadenopathy, and viral infections such as by herpesviruses, poxviruses, and adenoviruses.

[0103] HARC-associated diseases that are related to increased apoptosis include, but are not limited to, AIDS and other infectious or genetic immunodeficiencies, neurodegenerative diseases such as Alzheimer's disease, Parkinson's disease, amyotrophic lateral sclerosis, retinitis pigmentosa, and cerebellar degeneration, myelodysplastic syndromes such as aplastic anemia, ischemic injuries such as myocardial infarction, stroke, and reperfusion injury, toxin-induced diseases such as alcohol-induced liver damage, cirrhosis, and lathyrism, wasting diseases such as cachexia, viral infections such as by hepatitis B and C, and osteoporosis.

[0104] Therefore, in one embodiment, HARC or a fragment or derivative thereof may be administered to a subject to treat the disorders and diseases resulting from decreased apoptosis including, but not limited to, those listed above.

[0105] In another embodiment, a vector capable of expressing HARC, or a fragment or a derivative thereof, may also be administered to a subject to treat the disorders and diseases resulting from decreased apoptosis including, but not limited to, those described above.

[0106] In another embodiment, agonists of HARC may be administered to a subject to treat or prevent the disorders and

diseases resulting from decreased apoptosis including, but not limited to, those described above.

[0107] In another embodiment, antagonists or inhibitors of HARC may be administered to a subject to treat or prevent the disorders and diseases resulting from increased apoptosis including, but not limited to, those described above.

[0108] In another embodiment, a vector expressing antisense of the polynucleotide encoding HARC may be administered to a subject to treat or prevent the disorders and diseases resulting from increased apoptosis including, but not limited to, those described above.

[0109] In one aspect, antibodies which are specific for HARC may be used directly as an antagonist. In another aspect, antibodies may be used to deliver a pharmaceutical agent to cells or tissue which express HARC.

[0110] In other embodiments, any of the therapuetic compositions described above may be administered in combination with other conventional therapeutic agents. The combination of therapeutic agents having different mechanisms of action will have synergistic effects allowing for the use of lower effective doses of each agent and thereby lessening side effects.

[0111] Antagonists or inhibitors of HARC may be produced using methods which are generally known in the art. In particular, purified HARC may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind HARC.

[0112] Antibodies which are specific for HARC may be used directly as an antagonist, or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissue which express HARC. The antibodies may be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, single chain, Fab fragments, and fragments produced by a Fab expression library. Neutralizing antibodies, (i.e., those which inhibit dimer formation) are especially preferred for therapeutic use.

[0113] For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others, may be immunized by injection with HARC or any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanin, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and *Corynebacterium parvum* are especially preferable.

[0114] It is preferred that the peptides, fragments, or oligopeptides used to induce antibodies to HARC have an amino acid sequence consisting of at least five amino acids, and more preferably at least 10 amino acids. It is also preferable that they are identical to a portion of the amino acid sequence of the natural protein, and they may contain the entire amino acid sequence of a small, naturally occurring molecule. Short stretches of HARC amino acids may be fused with those of another protein such as keyhole limpet hemocyanin and antibody produced against the chimeric molecule.

[0115] Monoclonal antibodies to HARC may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique (Kohler, G. et al. (1975) Nature 256:495-497; Kozbor, D. et al. (1985) J. Immunol. Methods 81:31-42; Cote, R. J. et al. (1983) Proc. Natl. Acad. Sci. 80:2026-2030; Cole, S. P. et al. (1984) Mol. Cell Biol. 62:109-120).

[0116] In addition, techniques developed for the production of "chimeric antibodies", the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity can be used (Morrison, S. L. et al. (1984) Proc. Natl. Acad. Sci. 81:6851-6855; Neuberger, M. S. et al. (1984) Nature 312:604-608; Takeda, S. et al. (1985) Nature 314:452-454). Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce HARC-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobin libraries (Burton D. R. (1991) Proc. Natl. Acad. Sci. 88:11120-3).

[0117] Antibodies may also be produced by inducing in vivo production in the lymphocyte population or by screening recombinant immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature (Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. 86: 3833-3837; Winter, G. et al. (1991) Nature 349:293-299).

[0118] Antibody fragments which contain specific binding sites for HARC may also be generated. For example, such fragments include, but are not limited to, the F(ab')2 fragments which can be produced by pepsin digestion of the antibody molecule and the Fab fragments which can be generated by reducing the disulfide bridges of the F(ab')2 fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity (Huse, W. D. et al. (1989) Science 254:1275-1281).

[0119] Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between HARC and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering HARC epitopes is preferred, but a competitive binding assay may also be employed (Maddox, supra).

[0120] In another embodiment of the invention, the polynucleotides encoding HARC, or any fragment thereof, or antisense molecules, may be used for therapeutic purposes. In one aspect, antisense to the polynucleotide encoding HARC may be used in situations in which it would be desirable to block the transcription of the mRNA. In particular, cells may be transformed with sequences complementary to polynucleotides encoding HARC. Thus, antisense molecules may be used to modulate HARC activity, or to achieve regulation of gene function. Such technology is now well known in the art, and sense or antisense oligomers

or larger fragments, can be designed from various locations along the coding or control regions of sequences encoding HARC.

[0121] Expression vectors derived from retroviruses, adenovirus, herpes or vaccinia viruses, or from various bacterial plasmids may be used for delivery of nucleotide sequences to the targeted organ, tissue or cell population. Methods which are well known to those skilled in the art can be used to construct recombinant vectors which will express antisense molecules complementary to the polynucleotides of the gene encoding HARC. These techniques are described both in Sambrook et al. (supra) and in Ausubel et al. (supra).

[0122] Genes encoding HARC can be turned off by transforming a cell or tissue with expression vectors which express high levels of a polynucleotide or fragment thereof which encodes HARC. Such constructs may be used to introduce untranslatable sense or antisense sequences into a cell. Even in the absence of integration into the DNA, such vectors may continue to transcribe RNA molecules until they are disabled by endogenous nucleases. Transient expression may last for a month or more with a non-replicating vector and even longer if appropriate replication elements are part of the vector system.

[0123] As mentioned above, modifications of gene expression can be obtained by designing antisense molecules, DNA, RNA, or PNA, to the control regions of the gene encoding HARC, i.e., the promoters, enhancers, and introns. Oligonucleotides derived from the transcription initiation site, e.g., between positions −10 and +10 from the start site, are preferred. Similarly, inhibition can be achieved using "triple helix" base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature (Gee, J. E. et al. (1994) In: Huber, B. E. and B. I. Carr, *Molecular and Immunologic Approaches,* Futura Publishing Co., Mt. Kisco, N.Y.). The antisense molecules may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

[0124] Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. Examples which may be used include engineered hammerhead motif ribozyme molecules that can specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding HARC.

[0125] Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites which include the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides corresponding to the region of the target gene containing the cleavage site may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

[0126] Antisense molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences encoding HARC. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize antisense RNA constitutively or inducibly can be introduced into cell lines, cells, or tissues.

[0127] RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule or the use of phosphorothioate or 2'-O-methyl phosphodiester linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

[0128] Many methods for introducing vectors into cells or tissues are available and equally suitable for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection and by liposome injections may be achieved using methods which are well known in the art.

[0129] Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including, for example, mammals such as dogs, cats, cows, horses, rabbits, monkeys, and most preferably, humans.

[0130] An additional embodiment of the invention relates to the administration of a pharmaceutical composition, in conjunction with a pharmaceutically acceptable carrier, for any of the therapeutic effects discussed above. Such pharmaceutical compositions may consist of HARC, antibodies to HARC, mimetics, agonists, antagonists, or inhibitors of HARC. The compositions may be administered alone or in combination with at least one other agent, such as stabilizing compound, which may be administered in any sterile, biocompatible pharmaceutical carrier, including, but not limited to, saline, buffered saline, dextrose, and water. The compositions may be administered to a patient alone, or in combination with other agents, drugs or hormones.

[0131] The pharmaceutical compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

[0132] In addition to the active ingredients, these pharmaceutical compositions may contain suitable pharmaceutically-acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. Further details on techniques for formulation and administration

may be found in the latest edition of *Remington's Pharmaceutical Sciences* (Maack Publishing Co., Easton, Pa.).

[0133] Pharmaceutical compositions for oral administration can be formulated using pharmaceutically acceptable carriers well known in the art in dosages suitable for oral administration. Such carriers enable the pharmaceutical compositions to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions, and the like, for ingestion by the patient.

[0134] Pharmaceutical preparations for oral use can be obtained through combination of active compounds with solid excipient, optionally grinding a resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries, if desired, to obtain tablets or dragee cores. Suitable excipients are carbohydrate or protein fillers, such as sugars, including lactose, sucrose, mannitol, or sorbitol; starch from corn, wheat, rice, potato, or other plants; cellulose, such as methyl cellulose, hydroxypropylmethyl-cellulose, or sodium carboxymethylcellulose; gums including arabic and tragacanth; and proteins such as gelatin and collagen. If desired, disintegrating or solubilizing agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, alginic acid, or a salt thereof, such as sodium alginate.

[0135] Dragee cores may be used in conjunction with suitable coatings, such as concentrated sugar solutions, which may also contain gum arabic, talc, polyvinylpyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for product identification or to characterize the quantity of active compound, i.e., dosage.

[0136] Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a coating, such as glycerol or sorbitol. Push-fit capsules can contain active ingredients mixed with a filler or binders, such as lactose or starches, lubricants, such as talc or magnesium stearate, and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid, or liquid polyethylene glycol with or without stabilizers.

[0137] Pharmaceutical formulations suitable for parenteral administration may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks' solution, Ringer's solution, or physiologically buffered saline. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes. Optionally, the suspension may also contain suitable stabilizers or agents which increase the solubility of the compounds to allow for the preparation of highly concentrated solutions.

[0138] For topical or nasal administration, penetrants appropriate to the particular barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

[0139] The pharmaceutical compositions of the present invention may be manufactured in a manner that is known in the art, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping, or lyophilizing processes.

[0140] The pharmaceutical composition may be provided as a salt and can be formed with many acids, including but not limited to, hydrochloric, sulfuric, acetic, lactic, tartaric, malic, succinic, etc. Salts tend to be more soluble in aqueous or other protonic solvents than are the corresponding free base forms. In other cases, the preferred preparation may be a lyophilized powder which may contain any or all of the following: 1-50 mM histidine, 0.1%-2% sucrose, and 2-7% mannitol, at a pH range of 4.5 to 5.5, that is combined with buffer prior to use.

[0141] After pharmaceutical compositions have been prepared, they can be placed in an appropriate container and labeled for treatment of an indicated condition. For administration of HARC, such labeling would include amount, frequency, and method of administration.

[0142] Pharmaceutical compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

[0143] For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells, or in animal models, usually mice, rabbits, dogs, or pigs. The animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

[0144] A therapeutically effective dose refers to that amount of active ingredient, for example HARC or fragments thereof, antibodies of HARC, agonists, antagonists or inhibitors of HARC, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., ED50 (the dose therapeutically effective in 50% of the population) and LD50 (the dose lethal to 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index, and it can be expressed as the ratio, LD50/ED50.

[0145] Pharmaceutical compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies is used in formulating a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that include the ED50 with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, sensitivity of the patient, and the route of administration.

[0146] The exact dosage will be determined by the practitioner, in light of factors related to the subject that requires treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, general health of the subject, age, weight, and gender of the subject, diet, time and frequency of administration, drug combination(s), reaction

sensitivities, and tolerance/response to therapy. Long-acting pharmaceutical compositions may be administered every 3 to 4 days, every week, or once every two weeks depending on half-life and clearance rate of the particular formulation.

[0147] Normal dosage amounts may vary from 0.1 to 100,000 micrograms, up to a total dose of about 1 g, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

[0148] DIAGNOSTICS

[0149] In another embodiment, antibodies which specifically bind HARC may be used for the diagnosis of conditions or diseases characterized by expression of HARC, or in assays to monitor patients being treated with HARC, agonists, antagonists or inhibitors. The antibodies useful for diagnostic purposes may be prepared in the same manner as those described above for therapeutics. Diagnostic assays for HARC include methods which utilize the antibody and a label to detect HARC in human body fluids or extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by joining them, either covalently or non-covalently, with a reporter molecule. A wide variety of reporter molecules which are known in the art may be used, several of which are described above.

[0150] A variety of protocols including ELISA, RIA, and FACS for measuring HARC are known in the art and provide a basis for diagnosing altered or abnormal levels of HARC expression. Normal or standard values for HARC expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, preferably human, with antibody to HARC under conditions suitable for complex formation The amount of standard complex formation may be quantified by various methods, but preferably by photometric, means. Quantities of HARC expressed in subject samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

[0151] In another embodiment of the invention, the polynucleotides encoding HARC may be used for diagnostic purposes. The polynucleotides which may be used include oligonucleotide sequences, antisense RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantitate gene expression in biopsied tissues in which expression of HARC may be correlated with disease. The diagnostic assay may be used to distinguish between absence, presence, and excess expression of HARC, and to monitor regulation of HARC levels during therapeutic intervention.

[0152] In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding HARC or closely related molecules, may be used to identify nucleic acid sequences which encode HARC. The specificity of the probe, whether it is made from a highly specific region, e.g., 10 unique nucleotides in the 5' regulatory region, or a less specific region, e.g., especially in the 3' coding region, and the stringency of the hybridization or amplification (maximal, high, intermediate, or low) will determine whether the probe identifies only naturally occurring sequences encoding HARC, alleles, or related sequences.

[0153] Probes may also be used for the detection of related sequences, and should preferably contain at least 50% of the nucleotides from any of the HARC encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and derived from the nucleotide sequence of SEQ ID NO: 2 or from genomic sequence including promoter, enhancer elements, and introns of the naturally occurring HARC.

[0154] Means for producing specific hybridization probes for DNAs encoding HARC include the cloning of nucleic acid sequences encoding HARC or HARC derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, commercially available, and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, radionuclides such as 32P or 35S, or enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

[0155] Polynucleotide sequences encoding HARC may be used for the diagnosis of conditions or diseases which are associated with expression of HARC. Examples of such conditions or diseases include cancer, autoimmune diseases, lymphoproliferative disorders, psoriasis, atherosclerosis, restenosis, AIDS, immunodeficiency diseases, ischemic injuries, neurodegenerative diseases, osteoporosis, myelodysplastic syndromes, toxin-induced diseases, cachexia, and viral infections. The polynucleotide sequences encoding HARC may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; or in dip stick, pin, ELISA or chip assays utilizing fluids or tissues from patient biopsies to detect altered HARC expression. Such qualitative or quantitative methods are well known in the art.

[0156] In a particular aspect, the nucleotide sequences encoding HARC may be useful in assays that detect activation or induction of various cancers, particularly those mentioned above. The nucleotide sequences encoding HARC may be labeled by standard methods, and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantitated and compared with a standard value. If the amount of signal in the biopsied or extracted sample is significantly altered from that of a comparable control sample, the nucleotide sequences have hybridized with nucleotide sequences in the sample, and the presence of altered levels of nucleotide sequences encoding HARC in the sample indicates the presence of the associated disease. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or in monitoring the treatment of an individual patient.

[0157] In order to provide a basis for the diagnosis of disease associated with expression of HARC, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts

taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, which encodes HARC, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with those from an experiment where a known amount of a substantially purified polynucleotide is used. Standard values obtained from normal samples may be compared with values obtained from samples from patients who are symptomatic for disease. Deviation between standard and subject values is used to establish the presence of disease.

[0158] Once disease is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to evaluate whether the level of expression in the patient begins to approximate that which is observed in the normal patient. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

[0159] With respect to cancer, the presence of a relatively high amount of transcript in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the cancer.

[0160] Additional diagnostic uses for oligonucleotides designed from the sequences encoding HARC may involve the use of PCR. Such oligomers may be chemically synthesized, generated enzymatically, or produced from a recombinant source. Oligomers will preferably consist of two nucleotide sequences, one with sense orientation (5'→3') and another with antisense (3'←5'), employed under optimized conditions for identification of a specific gene or condition. The same two oligomers, nested sets of oligomers, or even a degenerate pool of oligomers may be employed under less stringent conditions for detection and/or quantitation of closely related DNA or RNA sequences.

[0161] Methods which may also be used to quantitate the expression of HARC include radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and standard curves onto which the experimental results are interpolated (Melby, P. C. et al. (1993) J. Immunol. Methods, 159:235-244; Duplaa, C. et al. (1993) Anal. Biochem. 212:229-236). The speed of quantitation of multiple samples may be accelerated by running the assay in an ELISA format where the oligomer of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

[0162] In another embodiment of the invention, the nucleic acid sequences which encode HARC may also be used to generate hybridization probes which are useful for mapping the naturally occurring genomic sequence. The sequences may be mapped to a particular chromosome or to a specific region of the chromosome using well known techniques. Such techniques include FISH, FACS, or artificial chromosome constructions, such as yeast artificial chromosomes, bacterial artificial chromosomes, bacterial P1 constructions or single chromosome cDNA libraries as reviewed in Price, C. M. (1993) Blood Rev. 7:127-134, and Trask, B. J. (1991) Trends Genet. 7:149-154.

[0163] FISH (as described in Verma et al. (1988) *Human Chromosomes: A Manual of Basic Techniques,* Pergamon Press, New York, N.Y.) may be correlated with other physical chromosome mapping techniques and genetic map data. Examples of genetic map data can be found in the 1994 Genome Issue of Science (265:1981f). Correlation between the location of the gene encoding HARC on a physical chromosomal map and a specific disease, or predisposition to a specific disease, may help delimit the region of DNA associated with that genetic disease. The nucleotide sequences of the subject invention may be used to detect differences in gene sequences between normal, carrier, or affected individuals.

[0164] In situ hybridization of chromosomal preparations and physical mapping techniques such as linkage analysis using established chromosomal markers may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the number or arm of a particular human chromosome is not known. New sequences can be assigned to chromosomal arms, or parts thereof, by physical mapping. This provides valuable information to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the disease or syndrome has been crudely localized by genetic linkage to a particular genomic region, for example, AT to 11q22-23 (Gatti, R. A. et al. (1988) Nature 336:577-580), any sequences mapping to that area may represent associated or regulatory genes for further investigation. The nucleotide sequence of the subject invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc. among normal, carrier, or affected individuals.

[0165] In another embodiment of the invention, HARC, its catalytic or immunogenic fragments or oligopeptides thereof, can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes, between HARC and the agent being tested, may be measured.

[0166] Another technique for drug screening which may be used provides for high throughput screening of compounds having suitable binding affinity to the protein of interest as described in published PCT application WO84/03564. In this method, as applied to HARC, large numbers of different small test compounds are synthesized on a solid substrate, such as plastic pins or some other surface. The test compounds are reacted with HARC, or fragments thereof, and washed. Bound HARC is then detected by methods well known in the art. Purified HARC can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

[0167] In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding HARC specifically compete with a test compound for binding HARC. In this manner, the antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with HARC.

[0168] In additional embodiments, the nucleotide sequences which encode HARC may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

[0169] The examples below are provided to illustrate the subject invention and are not included for the purpose of limiting the invention.

EXAMPLES

[0170] I HUVENOB01 cDNA Library Construction

[0171] Incyte clone 036596 was identified in the HUVENOB01 cDNA library prepared from control (or untreated) HUVEC cells. The HUVEC cell line is a normal, homogeneous, well-characterized early passage endothelial cell culture from human umbilical vein (Cell Systems Corp., Kirkland, Wash.).

[0172] The induced human endothelial cell (HUVEC) cDNA library was custom constructed by Stratagene (Cat. #937207). cDNA synthesis was primed using either oligo d(T) or random hexamers. Synthetic adaptor oligonucleotides were ligated onto cDNA ends enabling insertion into the lambda vector. The HUVEC cDNA library was constructed using the UNIZAP vector system (Stratagene). Finally, the two libraries were combined into a single library by mixing equal numbers of bacteriophage.

[0173] The custom-constructed library phage particles were infected into *E. coli* host strain XL1-BLUE (Stratagene) which has a high transformation efficiency, increasing the probability of obtaining rare, under-represented clones in the cDNA library. Alternative unidirectional vectors include, but are not limited to, pcDNAI (Invitrogen, San Diego, Calif.) and pSHlox-1 (Novagen, Madison, Wis.).

[0174] II Isolation and Sequencing of cDNA Clones

[0175] The phagemid forms of individual cDNA clones were obtained by the in vivo excision process in which the host bacterial strain was coinfected with both the lambda library phage and an f1 helper phage. Polypeptides derived from both the library-containing phage and the helper phage nicked the lambda DNA, initiated new DNA synthesis from defined sequences on the lambda target DNA and created a smaller, single-stranded circular phagemid DNA molecule that included all DNA sequences of the BLUESCRIPT plasmid (Stratagene) and the cDNA insert. The phagemid DNA was secreted from the cells, purified, and used to re-infect fresh host cells where the double-stranded phagemid DNA was produced. Because the phagemid carries the gene for β-lactamase, the newly-transformed bacteria were selected on media containing ampicillin.

[0176] Phagemid DNA was purified using the MAGIC MINIPREPS DNA Purification System (Cat. #A7100; Promega Corp., Madison, Wis.). This small-scale process provides a simple and reliable method for lysing the bacterial cells and rapidly isolating purified phagemid DNA using a proprietary DNA-binding resin.

[0177] Phagemid DNA was also purified using the QIAwell-8 Plasmid, QIAwell Plus and QIAwell Ultra DNA purification systems (QIAGEN, Chatsworth, Calif.). This product line provides a convenient, rapid, and reliable high-throughput method for lysing the bacterial cells and isolating highly purified phagemid DNA. The DNA was eluted from the purification and binding resins already prepared for DNA sequencing and other analytical manipulations.

[0178] The cDNA inserts from random isolates of the induced and control HUVEC libraries were sequenced in part using the method of Sanger et al. (1975; J. Mol. Biol. 94:441f). Conventional enzymatic methods employ DNA polymerase Klenow fragments, SEQUENASE or Taq polymersae to extend DNA chains from an oligonucleotide primer annealed to the DNA template of interest. Methods have been developed for the use of both single- and double-stranded templates. The chain termination reaction products are separated by electrophoreses through urea-acrylamide gels and are detected by fluorescence.

[0179] III Homology Searching of cDNA Clones and Their Deduced Proteins

[0180] Each cDNA was compared to sequences in GenBank using two search algorithms. The first was originally developed by D. J. Lipman and W. R. Pearson, (1985; Science 227:1435-1441). In this algorithm, the homologous regions are searched in a two-step manner. In the first step, the highest homologous regions are determined by calculating a matching score using a homology score table. The parameter 'Ktup' is used in this step to establish the minimum window size to be shifted for comparing two sequences. Ktup also sets the number of bases that must match to extract the highest homologous region among the sequences. In this step, no insertion or deletions are applied, and the homology is displayed as an initial (INIT) value.

[0181] In the second step, the homologous regions are aligned to obtain the highest matching score by inserting a gap in order to add a probable deleted portion. The matching score obtained in the first step is recalculated using the homology score table and the insertion score table to an optimized (OPT) value in the final output.

[0182] DNA homologies between two sequences can be examined graphically using the Harr method of constructing dot matrix homology plots (Needleman, S. B. and C. O. Wunsch (1970) J. Mol. Biol. 48:443-453). This method produces a two-dimensional plot which can be useful in determining regions of homology versus regions of repetition.

[0183] The second algorithm was developed by Applied Biosystems and incorporated into the INHERIT 670 sequence analysis system. In this algorithm, Pattern Specification Language (TRW Inc., Los Angeles, Calif.) was used to determine regions of homology. The three parameters that determine how the sequence comparisons run were window size, window offset, and error tolerance. Using a combination of these three parameters, the DNA database was searched for sequences containing regions of homology to the query sequence, and the appropriate sequences were scored with an initial value. Subsequently, these homologous regions were examined using dot matrix homology plots to distinguish regions of homology from chance matches. Smith-Waterman alignments were used to display the results of the homology search.

[0184] Peptide and protein sequence homologies were ascertained using the INHERIT-670 sequence analysis sys-

tem using the methods similar to those used in DNA sequence homologies. Pattern Specification Language and parameter windows were used to search protein databases for sequences containing regions of homology which were scored with an initial value. Dot-matrix homology plots were examined to distinguish regions of significant homology from chance matches.

[0185] BLAST, which stands for Basic Local Alignment Search Tool (Altschul, S. F. (1993) J. Mol. Evol. 36:290-300; Altschul et al. (1990) J. Mol. Biol. 215:403-410), was used to search for local sequence alignments. BLAST produces alignments of both nucleotide and amino acid sequences to determine sequence similarity. Because of the local nature of the alignments, BLAST is especially useful in determining exact matches or in identifying homologs. BLAST is useful for matches which do not contain gaps. The fundamental unit of BLAST algorithm output is the High-scoring Segment Pair (HSP).

[0186] An HSP consists of two sequence fragments of arbitrary but equal lengths whose alignment is locally maximal and for which the alignment score meets or exceeds a threshold or cutoff score set by the user. The BLAST approach is to look for HSPs between a query sequence and a database sequence, to evaluate the statistical significance of any matches found, and to report only those matches which satisfy the user-selected threshold of significance. The parameter E establishes the statistically significant threshold for reporting database sequence matches. E is interpreted as the upper bound of the expected frequency of chance occurrence of an HSP (or set of HSPs) within the context of the entire database search. Any database sequence whose match satisfies E is reported in the program output.

[0187] IV Northern Analysis

[0188] Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound (Sambrook et al., supra).

[0189] Analogous computer techniques using BLAST (Altschul, S. F. 1993 and 1990, supra) are used to search for identical or related molecules in nucleotide databases such as GenBank or the LIFESEQ database (Incyte Pharmaceuticals). This analysis is much faster than multiple, membrane-based hybridizations. In addition, the sensitivity of the computer search can be modified to determine whether any particular match is categorized as exact or homologous.

[0190] The basis of the search is the product score which is defined as:

$$\frac{\%\ \text{sequence identity} \times \%\ \text{maximum BLAST score}}{100}$$

[0191] The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. For example, with a product score of 40, the match will be exact within a 1-2% error; and at 70, the match will be exact. Homologous molecules are usually identified by selecting those which show product scores between 15 and 40, although lower scores may identify related molecules.

[0192] The results of northern analysis are reported as a list of libraries in which the transcript encoding HARC occurs. Abundance and percent abundance are also reported. Abundance directly reflects the number of times a particular transcript is represented in a cDNA library, and percent abundance is abundance divided by the total number of sequences examined in the cDNA library.

[0193] V Extension of HARC-Encoding Polynucleotides to Full Length or to Recover Regulatory Sequences

[0194] Full length HARC-encoding nucleic acid sequence (SEQ ID NO: 2) is used to design oligonucleotide primers for extending a partial nucleotide sequence to full length or for obtaining 5' or 3', intron or other control sequences from genomic libraries. One primer is synthesized to initiate extension in the antisense direction (XLR) and the other is synthesized to extend sequence in the sense direction (XLF). Primers are used to facilitate the extension of the known sequence "outward" generating amplicons containing new, unknown nucleotide sequence for the region of interest. The initial primers are designed from the cDNA using OLIGO 4.06 (National Biosciences), or another appropriate program, to be 22-30 nucleotides in length, to have a GC content of 50% or more, and to anneal to the target sequence at temperatures about 68°-72° C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations is avoided.

[0195] The original, selected cDNA libraries, or a human genomic library are used to extend the sequence; the latter is most useful to obtain 5' upstream regions. If more extension is necessary or desired, additional sets of primers are designed to further extend the known region.

[0196] By following the instructions for the XL-PCR kit (Perkin Elmer) and thoroughly mixing the enzyme and reaction mix, high fidelity amplification is obtained. Beginning with 40 pmol of each primer and the recommended concentrations of all other components of the kit, PCR is performed using the Peltier thermal cycler (PTC200; M. J. Research, Watertown, Mass.) and the following parameters:

Step 1	94° C. for 1 min (initial denaturation)
Step 2	65° C. for 1 min
Step 3	68° C. for 6 min
Step 4	94° C. for 15 sec
Step 5	65° C. for 1 min
Step 6	68° C. for 7 min
Step 7	Repeat step 4–6 for 15 additional cycles
Step 8	94° C. for 15 sec
Step 9	65° C. for 1 min
Step 10	68° C. for 7:15 min
Step 11	Repeat step 8–10 for 12 cycles
Step 12	72° C. for 8 min
Step 13	4° C. (and holding)

[0197] A 5-10 μl aliquot of the reaction mixture is analyzed by electrophoresis on a low concentration (about 0.6-0.8%) agarose mini-gel to determine which reactions were successful in extending the sequence. Bands thought to contain the largest products are selected and removed from the gel. Further purification involves using a commercial gel extraction method such as QIAQUICK (QIAGEN). After recovery of the DNA, Klenow enzyme is used to trim single-stranded, nucleotide overhangs creating blunt ends which facilitate religation and cloning.

[0198] After ethanol precipitation, the products are redissolved in 13 μl of ligation buffer, 1 μl T4-DNA ligase (15 units) and 1 μl T4 polynucleotide kinase are added, and the mixture is incubated at room temperature for 2-3 hours or overnight at 16° C. Competent *E. coli* cells (in 40 μl of appropriate media) are transformed with 3 μl of ligation mixture and cultured in 80 μl of SOC medium (Sambrook et al., supra). After incubation for one hour at 37° C., the whole transformation mixture is plated on Luria Bertani (LB)-agar (Sambrook et al., supra) containing 2×Carb. The following day, several colonies are randomly picked from each plate and cultured in 150 μl of liquid LB/2×Carb medium placed in an individual well of an appropriate, commercially-available, sterile 96-well microtiter plate. The following day, 5 **82** 1 of each overnight culture is transferred into a non-sterile 96-well plate and after dilution 1:10 with water, 5 μl of each sample is transferred into a PCR array.

[0199] For PCR amplification, 18 μl of concentrated PCR reaction mix (3.3×) containing 4 units of rTth DNA polymerase, a vector primer, and one or both of the gene specific primers used for the extension reaction are added to each well. Amplification is performed using the following conditions:

Step 1	94° C. for 60 sec
Step 2	94° C. for 20 sec
Step 3	55° C. for 30 sec
Step 4	72° C. for 90 sec
Step 5	Repeat steps 2–4 for an additional 29 cycles
Step 6	72° C. for 180 sec
Step 7	4° C. (and holding)

[0200] Aliquots of the PCR reactions are run on agarose gels together with molecular weight markers. The sizes of the PCR products are compared to the original partial cDNAs, and appropriate clones are selected, ligated into plasmid, and sequenced.

[0201] VI Labeling and Use of Hybridization Probes

[0202] Hybridization probes derived from SEQ ID NO: 2 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base-pairs, is specifically described, essentially the same procedure is used with larger cDNA fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 (National Biosciences), labeled by combining 50 pmol of each oligomer and 250 μCi of [γ-^{32}P] adenosine triphosphate (Amersham) and T4 polynucleotide kinase (DuPont NEN, Boston, Mass.). The labeled oligonucleotides are substantially purified with Sephadex G-25 superfine resin column (Pharmacia & Upjohn). A portion containing 10^7 counts per minute of each of the sense and antisense oligonucleotides is used in a typical membrane based hybridization analysis of human genomic DNA digested with one of the following endonucleases (Ase I, Bgl II, Eco RI, Pst I, Xba 1, or Pvu II; DuPont NEN).

[0203] The DNA from each digest is fractionated on a 0.7 percent agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham, N.H.). Hybridization is carried out for 16 hours at 40° C. To remove nonspecific signals, blots are sequentially washed at room temperature under increasingly stringent conditions up to 0.1×saline sodium citrate and 0.5% sodium dodecyl sulfate. After XOMAT AR film (Kodak, Rochester, N.Y.) is exposed to the blots in a PHOSPHOIMAGER cassette (Molecular Dynamics, Sunnyvale, Calif.) for several hours, hybridization patterns are compared visually.

[0204] VII Antisense Molecules

[0205] Antisense molecules to the HARC-encoding sequence, or any part thereof, is used to inhibit in vivo or in vitro expression of naturally occurring HARC. Although use of antisense oligonucleotides, comprising about 20 base-pairs, is specifically described, essentially the same procedure is used with larger cDNA fragments. An oligonucleotide based on the coding sequences of HARC, as shown in **FIGS. 1A and 1B**, is used to inhibit expression of naturally occurring HARC. The complementary oligonucleotide is designed from the most unique 5' sequence as shown in **FIGS. 1A and 1B** and used either to inhibit transcription by preventing promoter binding to the upstream nontranslated sequence or translation of an HARC-encoding transcript by preventing the ribosome from binding. Using an appropriate portion of the signal and 5' sequence of SEQ ID NO: 2, an effective antisense oligonucleotide includes any 15-20 nucleotides spanning the region which translates into the signal or 5' coding sequence of the polypeptide as shown in **FIGS. 1A and 1B**.

[0206] VIII Expression of HARC

[0207] Expression of HARC is accomplished by subcloning the cDNAs into appropriate vectors and transforming the vectors into host cells. In this case, the cloning vector, pSport, previously used for the generation of the cDNA library is used to express HARC in *E. coli.* Upstream of the cloning site, this vector contains a promoter for β-galactosidase, followed by sequence containing the amino-terminal Met, and the subsequent seven residues of β-galactosidase. Immediately following these eight residues is a bacteriophage promoter useful for transcription and a linker containing a number of unique restriction sites.

[0208] Induction of an isolated, transformed bacterial strain with IPTG using standard methods produces a fusion protein which consists of the first eight residues of β-galactosidase, about 5 to 15 residues of linker, and the full length protein. The signal residues direct the secretion of HARC into the bacterial growth media which can be used directly in the following assay for activity.

[0209] IX Demonstration of HARC Activity

[0210] HARC can be expressed by transforming a mammalian cell line such as COS7, HeLa or CHO with an eukaryotic expression vector encoding HARC. Eukaryotic expression vectors are commercially available, and the techniques to introduce them into cells are well known to those skilled in the art. The cells are incubated for 48-72 hours after transformation under conditions appropriate for the cell line to allow expression and accumulation of HARC.

[0211] Extracts containing solubilized proteins are prepared from cells expressing HARC by methods well known in the art. Portions of the extract containing HARC are applied to the wells of a polyacrylamide gel and electrophoresed at constant current using techniques well known in the art until a suitable tracking dye, such as bromophenol blue (Sigma), has migrated to the bottom of the gel. Appro-

priate control samples, prepared from extracts of untransformed cells and/or cells transformed with vector sequences alone, are electrophoresed in parallel lanes of the gel. Protein standards of known molecular weight (BioRad, Hercules, Calif.) are run in adjacent lanes to calibrate the gel.

[0212] The separated proteins are blotted onto a nitrocellulose membrane and assayed for binding to Ca^{2+}using the method described by Maruyama, K. et al. (1984; J. Biochem. (Tokyo) 95:511-519). After incubation in the presence of $^{45}Ca^{2+}$, the membrane is washed to remove nonspecifically bound radionuclide and exposed against Kodak XOMAT film for an appropriate period of time. A band will be visible on the film at a position that is indicative of a protein of the size predicted for HARC. A band of similar mobility will not be present in control samples prepared from extracts of untransformed cells or cells transformed with vector sequence alone.

[0213] The presence of HARC in the band is confirmed using an antibody specific for HARC by the western blot procedure that is well known in the art. The membrane is incubated with HARC-specific antibodies derived from an animal such as rabbit. After washing to remove unbound antibodies, the membrane is incubated with tagged goat anti-rabbit immunoglobulins. The tag may consist of any of a number of chromogenic, fluorescent or enzymatic molecules that can be attached to immunoglobulins by techniques well known in the art. The presence of tagged goat immunoglobulins bound to rabbit anti-HARC antibodies may be detected using techniques appropriate to the nature of the tag. Pre-immune sera or unrelated antisera may be used as suitable controls for nonspecific binding to the membrane.

[0214] X Production of HARC Specific Antibodies

[0215] HARC that is substantially purified using PAGE electrophoresis (Sambrook, supra), or other purification techniques, is used to immunize rabbits and to produce antibodies using standard protocols. The amino acid sequence deduced from SEQ ID NO: 2 is analyzed using LASERGENE software (DNASTAR Inc) to determine regions of high immunogenicity and a corresponding oligopolypeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions, is described by Ausubel et al. (supra), and others.

[0216] Typically, the oligopeptides are 15 residues in length, synthesized using an Applied Biosystems 431A peptide synthesizer using fmoc-chemistry, and coupled to keyhole limpet hemocyanin (KLH, Sigma, St. Louis, Mo.) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS; Ausubel et al., supra). Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. The resulting antisera are tested for antipeptide activity, for example, by binding the peptide to plastic, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radioiodinated, goat anti-rabbit IgG.

[0217] XI Purification of Naturally Occurring HARC Using Specific Antibodies

[0218] Naturally occurring or recombinant HARC is substantially purified by immunoaffinity chromatography using antibodies specific for HARC. An immunoaffinity column is constructed by covalently coupling HARC antibody to an activated chromatographic resin, such as CnBr-activated SEPHAROSE (Pharmacia & Upjohn). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

[0219] Media containing HARC is passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of HARC (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/HARC binding (e.g., a buffer of pH 2-3 or a high concentration of a chaotrope, such as urea or thiocyanate ion), and HARC is collected.

[0220] XII Identification of Molecules Which Interact with HARC

[0221] HARC or biologically active fragments thereof are labeled with ^{125}I Bolton-Hunter reagent (Bolton et al. (1973) Biochem. J. 133: 529). Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled HARC, washed and any wells with labeled HARC complex are assayed. Data obtained using different concentrations of HARC are used to calculate values for the number, affinity, and association of HARC with the candidate molecules.

[0222] All publications and patents mentioned in the above specification are herein incorporated by reference. Various modifications and variations of the described method and system of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific preferred embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(iii) NUMBER OF SEQUENCES: 3

(2) INFORMATION FOR SEQ ID NO: 1:

-continued

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 214 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

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

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

```
Met Phe Pro Xaa Gly Thr Pro Gly Xaa Pro Tyr Gly Xaa Ala Ala Pro
 1               5                  10                  15

Gly Gly Pro Tyr Gly Gln Pro Pro Pro Ser Ser Tyr Gly Xaa Gln Gln
            20                  25                  30

Pro Gly Xaa Tyr Gly Gln Gly Gly Ala Pro Pro Asn Val Asp Pro Glu
        35                  40                  45

Ala Tyr Ser Trp Phe Gln Ser Val Asp Ser Asp His Ser Gly Tyr Ile
    50                  55                  60

Ser Met Lys Glu Leu Lys Gln Ala Leu Val Asn Cys Asn Trp Ser Ser
65                  70                  75                  80

Phe Asn Asp Glu Thr Cys Leu Met Met Ile Asn Met Phe Asp Lys Thr
                85                  90                  95

Lys Ser Gly Arg Ile Asp Val Tyr Gly Phe Ser Ala Leu Trp Lys Phe
            100                 105                 110

Ile Gln Gln Trp Lys Asn Leu Phe Gln Gln Tyr Asp Arg Asp Arg Ser
        115                 120                 125

Gly Ser Ile Ser Tyr Thr Glu Leu Gln Gln Ala Leu Ser Gln Met Gly
    130                 135                 140

Tyr Asn Leu Ser Pro Gln Phe Thr Gln Leu Leu Val Ser Arg Tyr Cys
145                 150                 155                 160

Pro Arg Ser Ala Asn Pro Ala Met Gln Leu Asp Arg Phe Ile Gln Val
                165                 170                 175

Cys Thr Gln Leu Gln Val Leu Thr Glu Ala Phe Arg Glu Lys Asp Thr
            180                 185                 190

Ala Val Gln Gly Asn Ile Arg Leu Ser Phe Glu Asp Phe Val Thr Met
        195                 200                 205

Thr Ala Ser Arg Met Leu
    210
```

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 776 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

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

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

```
TGGACACCCC AATCCTGGNA TGTTCCCCTN TGGAACTCCA GGAGNACCAT ATGGCGNTGC     60

AGCTCCCGGG GGCCCCTATG GTCAGCCACC TCCAAGTTCC TACGGTGNCC AGCAGCCTGG    120
```

-continued

GNTTTATGGA CAGGGTGGCG CCCCTCCCAA TGTGGATCCT GAGGCCTACT CCTGGTTCTA 180

GTCGGTGGAC TCAGATCACA GTGGCTATAT CTCCATGAAG GAGCTAAAGC AGGCCCTGTT 240

CAACTGCAAT TGGTCTTCAT TCAATGATGA GACCTGCCTC ATGATGATAA ACATGTTTGA 300

CAAGACCAAG TCAGGCCGCA TCGATGTCTA CGGCTTCTCA GCCCTGTGGA AATTCATCCA 360

GCAGTGGAAG AACCTCTTCC AGCAGTATGA CCGGGACCGC TCGGGCTCCA TTAGCTACAC 420

AGAGCTGCAG CAAGCTCTGT CCCAAATGGG CTACAACCTG AGCCCCCAGT TCACCCAGCT 480

TCTGGTCTCC CGCTACTGCC CACGCTCTGC CAATCCTGCC ATGCAGCTTG ACCGCTTCAT 540

CCAGGTGTGC ACCCAGCTGC AGGTGCTGAC AGAGGCCTTC CGGGAGAAGG ACACAGCTGT 600

ACAAGGCAAC ATCCGGCTCA GCTTCGAGGA CTTCGTCACC ATGACAGCTT CTCGGATGCT 660

ATGACCCAAC CATCTGTGGA GAGTGGAGTG CACCAGGGAC CTTTCCTGGC TTCTTAGAGT 720

GAGAGAAGTA TGTGGACATC TCTTCTTTTC CTGTCCCTCT AGAAGAACAT TCTCCC 776

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 191 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vii) IMMEDIATE SOURCE:
(A) LIBRARY: GenBank
(B) CLONE: 1213520

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

```
Met Ala Ala Tyr Ser Tyr Arg Pro Gly Pro Gly Gly Gly Pro Gly Pro
 1               5                  10                  15

Ala Ala Gly Ala Ala Leu Pro Asp Gln Ser Phe Leu Trp Asn Val Phe
            20                  25                  30

Gln Arg Val Asp Lys Asp Arg Ser Gly Val Ile Ser Asp Asn Glu Leu
        35                  40                  45

Gln Gln Ala Leu Ser Asn Gly Thr Trp Thr Pro Phe Asn Pro Val Thr
    50                  55                  60

Val Arg Ser Ile Ile Ser Met Phe Asp Arg Glu Asn Lys Ala Gly Val
65                  70                  75                  80

Asn Phe Ser Glu Phe Thr Gly Val Trp Lys Tyr Ile Thr Asp Trp Gln
                85                  90                  95

Asn Val Phe Arg Thr Tyr Asp Arg Asp Asn Ser Gly Met Ile Asp Lys
            100                 105                 110

Asn Glu Leu Lys Gln Ala Leu Ser Gly Phe Gly Tyr Arg Leu Ser Asp
        115                 120                 125

Gln Phe His Asp Ile Leu Ile Arg Lys Phe Asp Arg Gln Gly Arg Gly
    130                 135                 140

Gln Ile Ala Phe Asp Asp Phe Ile Gln Gly Cys Ile Val Leu Gln Arg
145                 150                 155                 160

Leu Thr Asp Ile Phe Arg Arg Tyr Asp Thr Asp Gln Asp Gly Trp Ile
                165                 170                 175

Gln Val Ser Tyr Glu Gln Tyr Leu Ser Met Val Phe Ser Ile Val
            180                 185                 190
```

What is claimed is:

1. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of:

a) a polypeptide comprising an amino acid sequence of SEQ ID NO: 1,

b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence of SEQ ID NO: 1,

c) a biologically active fragment of a polypeptide having an amino acid sequence of SEQ ID NO: 1, and

d) an immunogenic fragment of a polypeptide having an amino acid sequence of SEQ ID NO: 1.

2. An isolated polypeptide of claim 1, having a sequence of SEQ ID NO: 1.

3. An isolated polynucleotide encoding a polypeptide of claim 1.

4. An isolated polynucleotide encoding a polypeptide of claim 2.

5. An isolated polynucleotide of claim 4, having a sequence of SEQ ID NO: 2 .

6. A recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide of claim 3.

7. A cell transformed with a recombinant polynucleotide of claim 6.

8. A transgenic organism comprising a recombinant polynucleotide of claim 6.

9. A method for producing a polypeptide of claim 1, the method comprising:

a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide, and said recombinant polynucleotide comprises a promoter sequence operably linked to a polynucleotide encoding the polypeptide of claim 1, and

b) recovering the polypeptide so expressed.

10. A method of claim 9, wherein the polypeptide has the sequence of SEQ ID NO: 1.

11. An isolated antibody which specifically binds to a polypeptide of claim 1.

12. An isolated polynucleotide comprising a sequence selected from the group consisting of:

a) a polynucleotide comprising a polynucleotide sequence of SEQ ID NO: 2,

b) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 90% identical to a polynucleotide sequence of SEQ ID NO: 2,

c) a polynucleotide having a sequence complementary to a polynucleotide of a),

d) a polynucleotide having a sequence complementary to a polynucleotide of b) and

e) an RNA equivalent of a)-d).

13. An isolated polynucleotide comprising at least 60 contiguous nucleotides of a polynucleotide of claim 12.

14. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 12, the method comprising:

a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and

b) detecting the presence or absence of said hybridization complex, and, optionally, if present, the amount thereof.

15. A method of claim 14, wherein the probe comprises at least 60 contiguous nucleotides.

16. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 12, the method comprising:

a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and

b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.

17. A composition comprising a polypeptide of claim 1 and a pharmaceutically acceptable excipient.

18. A composition of claim 17, wherein the polypeptide has an amino acid sequence of SEQ ID NO: 1.

19. A method for treating a disease or condition associated with decreased expression of functional HARC, comprising administering to a patient in need of such treatment the composition of claim 17.

20. A method for screening a compound for effectiveness as an agonist of a polypeptide of claim 1, the method comprising:

a) exposing a sample comprising a polypeptide of claim 1 to a compound, and

b) detecting agonist activity in the sample.

21. A composition comprising an agonist compound identified by a method of claim 20 and a pharmaceutically acceptable excipient.

22. A method for treating a disease or condition associated with decreased expression of functional HARC, comprising administering to a patient in need of such treatment a composition of claim 21.

23. A method for screening a compound for effectiveness as an antagonist of a polypeptide of claim 1, the method comprising:

a) exposing a sample comprising a polypeptide of claim 1 to a compound, and

b) detecting antagonist activity in the sample.

24. A composition comprising an antagonist compound identified by a method of claim 23 and a pharmaceutically acceptable excipient.

25. A method for treating a disease or condition associated with overexpression of functional HARC, comprising administering to a patient in need of such treatment a composition of claim 24.

26. A method of screening for a compound that specifically binds to the polypeptide of claim 1, said method comprising the steps of:

a) combining the polypeptide of claim 1 with at least one test compound under suitable conditions, and

b) detecting binding of the polypeptide of claim 1 to the test compound, thereby identifying a compound that specifically binds to the polypeptide of claim 1.

27. A method of screening for a compound that modulates the activity of the polypeptide of claim 1, said method comprising:

a) combining the polypeptide of claim 1 with at least one test compound under conditions permissive for the activity of the polypeptide of claim 1,

b) assessing the activity of the polypeptide of claim 1 in the presence of the test compound, and

c) comparing the activity of the polypeptide of claim 1 in the presence of the test compound with the activity of the polypeptide of claim 1 in the absence of the test compound, wherein a change in the activity of the polypeptide of claim 1 in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide of claim 1.

28. A method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a polynucleotide sequence of claim 5, the method comprising:

a) exposing a sample comprising the target polynucleotide to a compound, under conditions suitable for the expression of the target polynucleotide,

b) detecting altered expression of the target polynucleotide, and

c) comparing the expression of the target polynucleotide in the presence of varying amounts of the compound and in the absence of the compound.

29. A method for assessing toxicity of a test compound, said method comprising:

a) treating a biological sample containing nucleic acids with the test compound;

b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide of claim 12 under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide comprising a polynucleotide sequence of a polynucleotide of claim 12 or fragment thereof;

c) quantifying the amount of hybridization complex; and

d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.

30. A diagnostic test for a condition or disease associated with the expression of HARC in a biological sample comprising the steps of:

a) combining the biological sample with an antibody of claim 11, under conditions suitable for the antibody to bind the polypeptide and form an antibody:polypeptide complex; and

b) detecting the complex, wherein the presence of the complex correlates with the presence of the polypeptide in the biological sample.

31. The antibody of claim 11, wherein the antibody is:

a) a chimeric antibody,

b) a single chain antibody,

c) a Fab fragment,

d) a $F(ab')_2$ fragment, or

e) a humanized antibody.

32. A composition comprising an antibody of claim 11 and an acceptable excipient.

33. A method of diagnosing a condition or disease associated with the expression of HARC in a subject, comprising administering to said subject an effective amount of the composition of claim 32.

34. A composition of claim 32, wherein the antibody is labeled.

35. A method of diagnosing a condition or disease associated with the expression of HARC in a subject, comprising administering to said subject an effective amount of the composition of claim 34.

36. A method of preparing a polyclonal antibody with the specificity of the antibody of claim 11 comprising:

a) immunizing an animal with a polypeptide having an amino acid sequence of SEQ ID NO: 1, or an immunogenic fragment thereof, under conditions to elicit an antibody response;

b) isolating antibodies from said animal; and

c) screening the isolated antibodies with the polypeptide, thereby identifying a polyclonal antibody which binds specifically to a polypeptide having an amino acid sequence of SEQ ID NO: 1.

37. An antibody produced by a method of claim 36.

38. A composition comprising the antibody of claim 37 and a suitable carrier.

39. A method of making a monoclonal antibody with the specificity of the antibody of claim 11 comprising:

a) immunizing an animal with a polypeptide having an amino acid sequence of SEQ ID NO: 1, or an immunogenic fragment thereof, under conditions to elicit an antibody response;

b) isolating antibody producing cells from the animal;

c) fusing the antibody producing cells with immortalized cells to form monoclonal antibody-producing hybridoma cells;

d) culturing the hybridoma cells; and

e) isolating from the culture monoclonal antibody which binds specifically to a polypeptide having an amino acid sequence of SEQ ID NO: 1.

40. A monoclonal antibody produced by a method of claim 39.

41. A composition comprising the antibody of claim 40 and a suitable carrier.

42. The antibody of claim 11, wherein the antibody is produced by screening a Fab expression library.

43. The antibody of claim 11, wherein the antibody is produced by screening a recombinant immunoglobulin library.

44. A method for detecting a polypeptide having an amino acid sequence of SEQ ID NO: 1 in a sample, comprising the steps of:

a) incubating the antibody of claim 11 with a sample under conditions to allow specific binding of the antibody and the polypeptide; and

b) detecting specific binding, wherein specific binding indicates the presence of a polypeptide having an amino acid sequence of SEQ ID NO: 1 in the sample.

45. A method of purifying a polypeptide having an amino acid sequence of SEQ ID NO: 1 from a sample, the method comprising:

a) incubating the antibody of claim 11 with a sample under conditions to allow specific binding of the antibody and the polypeptide; and

b) separating the antibody from the sample and obtaining the purified polypeptide having an amino acid sequence of SEQ ID NO: 1.

46. A microarray wherein at least one element of the microarray is a polynucleotide of claim 13.

47. A method for generating a transcript image of a sample which contains polynucleotides, the method comprising the steps of:

a) labeling the polynucleotides of the sample,

b) contacting the elements of the microarray of claim 46 with the labeled polynucleotides of the sample under conditions suitable for the formation of a hybridization complex, and

c) quantifying the expression of the polynucleotides in the sample.

48. An array comprising different nucleotide molecules affixed in distinct physical locations on a solid substrate, wherein at least one of said nucleotide molecules comprises a first oligonucleotide or polynucleotide sequence specifically hybridizable with at least 30 contiguous nucleotides of a target polynucleotide, said target polynucleotide having a sequence of claim 12.

49. An array of claim 48, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to at least 30 contiguous nucleotides of said target polynucleotide.

50. An array of claim 48, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to at least 60 contiguous nucleotides of said target polynucleotide.

51. An array of claim 48, which is a microarray.

52. An array of claim 48, further comprising said target polynucleotide hybridized to said first oligonucleotide or polynucleotide.

53. An array of claim 48, wherein a linker joins at least one of said nucleotide molecules to said solid substrate.

54. An array of claim 48, wherein each distinct physical location on the substrate contains multiple nucleotide molecules having the same sequence, and each distinct physical location on the substrate contains nucleotide molecules having a sequence which differs from the sequence of nucleotide molecules at another physical location on the substrate.

55. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO: 1.

56. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO: 2.

57. An isolated antibody of claim 11 which specifically binds to a polypeptide comprising an amino acid sequence selected from the group consisting of

a) a biologically-active fragment of at least 10 amino acids of the amino acid sequence of SEQ ID NO: 1, said fragment having a calcium-binding domain and having calcium-binding activity, and

b) an immunologically-active fragment of at least 10 amino acids of the amino acid sequence of SEQ ID NO: 1, said fragment having a calcium-binding domain.

* * * * *