

PCTWORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C12N 15/12, C07K 14/47, G01N 33/50, 33/74		A2	(11) International Publication Number: WO 00/04152
			(43) International Publication Date: 27 January 2000 (27.01.00)
(21) International Application Number: PCT/US99/16122 (22) International Filing Date: 16 July 1999 (16.07.99) (30) Priority Data: 60/093,239 17 July 1998 (17.07.98) US 60/100,243 14 September 1998 (14.09.98) US (71) Applicant: UNIVERSITY OF ROCHESTER [US/US]; Office of Technology Transfer, 518 Hylan Building, Rochester, NY 14627-0140 (US). (72) Inventor: CHANG, Chawnshang; University of Rochester, 601 Elmwood Avenue, P.O. Box 626, Rochester, NY 14642 (US). (74) Agent: SEAY, Nicholas, J.; Quarles & Brady LLP, P.O. Box 2113, Madison, WI 53701-2113 (US).			(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG). Published <i>Without international search report and to be republished upon receipt of that report.</i>
(54) Title: ANDROGEN RECEPTOR COACTIVATORS			
(57) Abstract Disclosed are androgen receptor-associated proteins, designated ARA24, ARA54, ARA55, and Rb, that have been demonstrated to interact with the androgen receptor to alter levels of androgen receptor-mediated transcriptional activation. Certain of these proteins interact with the androgen receptor in an androgen-dependent manner, whereas certain proteins may induce transcriptional activation in the presence of other ligands, such as E2 or HF. Also disclosed is a method of detecting androgenic or antiandrogenic activity using these proteins in a mammalian two-hybrid transient transfection assay.			

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

ANDROGEN RECEPTOR COACTIVATORS

CROSS-REFERENCE TO RELATED APPLICATIONS

Not applicable.

STATEMENT REGARDING FEDERALLY SPONSORED RESEARCH
OR DEVELOPMENT

5

Not applicable.

BACKGROUND OF THE INVENTION

Androgens constitute a class of hormones that control the development and proper function of mammalian male reproductive systems, including the prostate and epididymis. Androgens also affect the physiology of many non-reproductive systems, including muscle, skin, pituitary, lymphocytes, hair growth, and brain. Androgens exert their effect by altering the level of gene expression of specific genes in a process that is mediated by binding of androgen to an androgen receptor. The androgen receptor, which is a member of the steroid receptor super family, plays an important role in male sexual differentiation and in prostate cell proliferation. Binding of androgen by the androgen receptor allows the androgen receptor to interact with androgen responsive element (AREs), DNA sequences found on genes whose expression is regulated by androgen.

Androgen-mediated regulation of gene expression is a complicated process that may involve multiple co-activators (Adler et al., Proc. National Acad. Sci. USA 89:6319-6325, 1992). A fundamental question in the field of steroid hormone biology is how specific androgen-activated transcription can be achieved in vivo when several different receptors recognize the same DNA sequence. For example, the androgen receptor (AR), the glucocorticoid receptor (GR), and the progesterone receptor (PR) all

recognize the same sequence but activate different transcription activities. Some have speculated that accessory factors may selectively interact with the androgen receptor to determine the specificity of gene
5 activation by the androgen receptor.

Prostate cancer is the most common malignant neoplasm in aging males in the United States. Standard treatment includes the surgical or chemical castration of the patient in combination with the administration of anti-androgens
10 such as 17 β estradiol (E2) or hydroxyflutamide (HF). However, most prostate cancers treated with androgen ablation and anti-androgens progress from an androgen-dependant to an androgen-independent state, causing a high incidence of relapse within 18 months (Crawford, Br. J.
15 Urology 70: suppl. 1, 1992). The mechanisms by which prostate cancer cells become resistant to hormonal therapy remain unclear. One hypothesis that has been advanced is that over the course of treatment, a mutation in the AR occurs which alters the receptor's sensitivity to other
20 steroid hormones or anti-androgens, such as E2 and HF, thereby causing the progression from androgen-dependent to androgen-independent prostate cancer. This hypothesis is supported by transient transfection assays in which it has been shown that anti-androgens may have an agonistic
25 activity that stimulates mutant AR (mAR)-mediated transcription.

Recently, A1B1 was identified as estrogen receptor coactivator that is expressed at higher levels in ovarian cancer cell lines and breast cancer cells than in
30 noncancerous cells (Anzick, et al. Science 277:965-968, 1997). This result suggests that steroid hormone receptor cofactors may play an important role in the progression of certain diseases, such as hormone responsive tumors.

The identification, isolation, and characterization of
35 genes that encode factors involved in the regulation of gene expression by androgen receptors will facilitate the development of screening assays to evaluate the potential

efficacy of drugs in the treatment of prostate cancers.

BRIEF SUMMARY OF THE INVENTION

The present invention includes an isolated polynucleotide that encodes a co-activator for human
5 androgen receptor, the polynucleotide comprising a sequence that encodes a polypeptide selected from the group consisting of an ARA54 polypeptide, an ARA55 polypeptide, an ARA24 polypeptide, and an Rb polypeptide.

Another aspect of the present invention is a genetic
10 construct comprising a promoter functional in a prokaryotic or eukaryotic cell operably connected to a polynucleotide that encodes a polypeptide selected from the group consisting of an ARA54 polypeptide, an ARA55 polypeptide, an ARA24 polypeptide and an Rb polypeptide.

15 The present invention provides a method for screening candidate pharmaceutical molecules for the ability to promote or inhibit the interaction of ARs and AREs to modulate androgenic activity comprising the steps of:

(a) providing a genetic construct comprising a
20 promoter functional in a eukaryotic cell operably connected to a polynucleotide comprising a sequence that encodes a polypeptide selected from the group consisting of an ARA54 polypeptide, an ARA55 polypeptide, an ARA24 polypeptide, and a retinoblastoma polypeptide;

25 (b) cotransforming a suitable eukaryotic cell with the construct of step a, and a construct comprising at least a portion of an expressible androgen receptor sequence;

(c) culturing the cells in the presence of a
30 candidate pharmaceutical molecule; and

(d) assaying the transcriptional activity induced by the androgen receptor.

It is an object of the present invention to provide a genetic construct capable of expressing a factor involved
35 in co-activation of the human androgen receptor.

It is an object of the present invention to provide a

method for evaluating the ability of candidate pharmaceutical molecules to modulate the effect of androgen receptor coactivators on gene expression.

Other objects, features, and advantages of the present invention will become apparent upon reading the specification and claims.

DETAILED DESCRIPTION OF THE INVENTION

Transactivation of genes by the androgen receptor is a complicated system that involves many different coactivators. It is not currently known just how many factors are involved in androgen receptor-mediated regulation of gene expression. The identification and/or characterization of four androgen receptor coactivators is reported herein. Inclusion of one or more of these coactivators in an assay for androgenic and antiandrogenic activity is expected to increase the sensitivity of the assay. Information about these coactivators is valuable in the design of pharmaceutical agents intended to enhance or interfere with normal coactivator function. A preliminary assessment of the efficacy of a potential therapeutic agent can be made by evaluating the effect of the agent on the ability of the coactivator to enhance transactivation by the androgen receptor.

One aspect of the present invention is an isolated polynucleotide that encodes a co-activator for human androgen receptor, the polynucleotide comprising a sequence that encodes a polypeptide selected from the group consisting of an ARA54 polypeptide, an ARA55 polypeptide, an ARA24 polypeptide and an Rb polypeptide.

Another aspect of the present invention is a genetic construct comprising a promoter functional in a prokaryotic or eukaryotic cell operably connected to a polynucleotide that encodes a polypeptide selected from the group consisting of an ARA54 polypeptide, an ARA55 polypeptide, an ARA24 polypeptide and an Rb polypeptide.

The present invention includes a method for screening

candidate pharmaceutical molecules for the ability to promote or inhibit the ARs and AREs to result in modulation of androgenic effect comprising the steps of:

- (a) providing a genetic construct comprising a promoter functional in a eukaryotic cell operably connected to a polynucleotide comprising a sequence that encodes a polypeptide selected from the group consisting of an ARA54 polypeptide, an ARA55 polypeptide, an ARA24 polypeptide, and a retinoblastoma polypeptide;
- (b) cotransforming a suitable eukaryotic cell with the construct of step a, and a construct comprising at least a portion of an expressible androgen receptor sequence;
- (c) culturing the cells in the presence of a candidate pharmaceutical molecule; and
- (d) assaying the transcriptional activity induced by the androgen receptor gene.

The human androgen receptor is comprised of a ligand binding domain (LBD), a DNA binding domain (DBD), a hinge domain containing nuclear localization signals, and a transactivation domain in the hyper-variable N-terminus. Truncation or deletion of the LBD results in constitutive transactivation by the N-terminal domain.

In certain cases, progression of prostate cancer from androgen dependent- to androgen independent-stage may be caused by a mutation in the LBD that alters the ligand specificity of the mAR (Taplan et al., *New Engl. J. Med.* 332:1393-1398 (1995); Gaddipati et al., *Cancer Res.* 54:2861-2864 (1994)). We examined whether differential steroid specificity of wild type (wt) AR and mAR involves the use of different androgen receptor-associated (ARA) proteins or coactivators by these receptors.

As described in the examples, a yeast two-hybrid system with mART887S as bait was used to screen the human prostate cDNA library. The sequences of two clones encoding a putative coactivators (designated ARA54 and ARA55) are shown in SEQ ID NO:1 and SEQ ID NO:3,

respectively. The putative amino acid sequences of ARA54 and ARA55 are shown in SEQ ID NO:2 and SEQ ID NO:4, respectively. Also provided are the DNA and amino acid sequences of ARA24 (SEQ ID NO:5 and SEQ ID NO:6, respectively) and Rb (SEQ ID NO:7 and SEQ ID NO:8, respectively). These coactivators were further characterized as detailed below. It is expected that some minor variations from SEQ ID NOs:1-8 associated with nucleotide additions, deletions, and mutations, whether naturally occurring or introduced *in vitro*, will not affect coactivation by the expression product or polypeptide.

Briefly, ARA54 is a 54 kDa protein that interacts with AR in an androgen-dependent manner. Coexpression of ARA54 and AR in a mammalian two-hybrid system demonstrated that reporter gene activity was enhanced in an androgen-dependent manner. ARA54 functions as a coactivator relatively specific for AR-mediated transcription. However, ARA54 may also function as a general coactivator of the transcriptional activity for other steroid receptors through their cognate ligands and response elements. ARA54 was found to enhance the transcriptional activity of AR and PR up to 6 fold and 3-5 fold, respectively. In contrast, ARA54 has only marginal effects (less than 2 fold) on glucocorticoid receptor (GR) and estrogen receptor (ER) in DU145 cells.

Coexpression of ARA54 with known AR coactivators SRC-1 or ARA70 revealed that each of these coactivators may contribute individually to achieve maximal AR-mediated transcriptional activity. Moreover, when ARA54 was expressed simultaneously with SRC-1 or ARA70, the increase in AR-mediated transactivation was additive but not synergistic relative to that observed in the presence of each coactivator alone.

The C-terminal domain of ARA54 (a.a. 361-471 of SEQ ID NO:1) serves as a dominant negative inhibitor of AR-mediated gene expression of target genes. Coexpression of exogenous full-length ARA54 can reduce this squelching

effect in a dose-dependent manner.

ARA54 enhanced transactivation of wtAR in the presence of DHT (10^{-10} to 10^{-8} M) by about 3-5 fold. However, transactivation of wtAR was enhanced only marginally with
5 E2 (10^{-9} - 10^{-7} M) or HF (10^{-7} - 10^{-5} M) as the ligand. The ability of ARA54 to enhance transactivation by two mutant receptors (mARt877a and mARe708k) that exhibit differential sensitivities to E2 and HF (Yeh et al., *Proc. Natl. Acad. Sci. USA*, in press (1998)) was also examined. The mutant
10 mARt877a, which is found in many prostate tumors (23), was activated by E2 (10^{-9} - 10^{-7} M) and HF (10^{-7} - 10^{-5} M), and ARA54 could further enhance E2- or HF-mediated AR transactivation. In contrast, the mutant mARe708k, first identified in a yeast genetic screening (Wang, C., Ph.D.
15 *Thesis of University of Wisconsin-Madison* (1997)), exhibited ligand specificity and response to ARE54 comparable to that of wtAR.

It is expected that any polypeptide having substantial homology to ARA54 that still actuates the same biological
20 effect can function as "an ARA54 polypeptide." With the sequence information disclosed herein, one skilled in the art can obtain any ARA54 polypeptide using standard molecular biological techniques. An ARA54 polypeptide is a polypeptide that is capable of enhancing transactivation
25 of AR in an androgen-dependent manner, enhancing E2 or HF transactivation by the mutant receptor mARt877a, and reducing inhibition of AR-mediated gene expression caused by overexpression of the C-terminal domain of ARA54 (a.a. 361-471 of SEQ ID NO:1). The sequence information
30 presented in this application can be used to identify, clone or sequence allelic variations in the ARA54 genes as well as the counterpart genes from other mammalian species. it is also contemplate that truncations of the native coding region can be made to express smaller polypeptides
35 that will retain the same biological activity.

The polynucleotide sequence of ARA55 (SEQ ID NO:3) exhibits high homology to the C-terminus of mouse hic5

(hydrogen peroxide inducible clone) (Pugh, B., *Curr. Opin. Cell Biol.* 8:303-311 (1996)), and like hic5, ARA55 expression is induced by TGF β . Cotransfection assays of transcriptional activation, which are described in detail
5 below, revealed that ARA55 is able to bind to both wtAR and mART887S in a ligand-dependent manner to enhance AR transcriptional activities. ARA55 enhanced transcriptional activation by wtAR in the presence of 10^{-9} M DHT or T, but not 10^{-9} M E2 or HF. In contrast, ARA55 can enhance
10 transcriptional activation by mART887S in the presence of DHT, testosterone (T), E2, or HF. ARA55 did not enhance transcriptional activation of mARe708k in the presence of E2, but can enhance transcription in the presence of DHT or T.

15 The C-terminal domain of ARA55 (amino acids 251-444 of SEQ ID NO:3) is sufficient for binding to ARs, but does not enhance transcriptional activation by ARs.

The invention is not limited to the particular ARA55 polypeptide disclosed in SEQ ID NO:4. It is expected that
20 any ARA55 polypeptide could be used in the practice of the present invention. By "an ARA55 polypeptide" it meant a polypeptide that is capable of enhancing transactivation of wtAR,, the mutant receptor mARt877a, in the presence of DHT, E2, or HF or intact receptor mARe708k in the presence
25 of DHT or T. Such polypeptides include allelic variants and the corresponding genes from other mammalian species as well as truncations.

The AR N-terminal domain comprises a polymorphic poly-glutamine (Q) stretch and a polymorphic poly-glycine (G)
30 stretch that account for variability in the length of human AR cDNA observed. The length of the poly-Q region (normally 11-33 residues in length) is inversely correlated with the risk of prostate cancer, and directly correlated with the SBMA, or Kennedy's disease (La Spada et al.,
35 *Nature (London)* 352:77-79 (1991)). The incidence of higher grade, distant metastatic, and fatal prostate cancer is higher in men having shorter AR poly-Q stretches.

As described in the examples, experiments undertaken to identify potential coactivators that interact with the AR poly-Q region led to the isolation of a clone encoding a coactivator, designated ARA24, that interacts with the poly-Q region. The sequences of the ARA24 clone and its putative translation product is shown in SEQ ID NO:5 and SEQ ID NO:6.

The ARA24 clone has an ORF that is identical to the published ORF for human Ran, an abundant, ras-like small GTPase (Beddow et al. *Proc. Natl. Acad. Sci. USA* 92:3328-3332, 1995). Overexpression of ARA24 in the presence of DHT does enhance transcriptional activation by AR over that observed in cells transfected with AR alone. Moreover, expression of antisense ARA24 (ARA24as) does reduce DHT-induced transcriptional activation.

An ARA24 polypeptide is one that interacts with the poly-Q region of an AR. An ARA24 polypeptide is further characterized by its ability to increase transactivation when overexpressed in eukaryotic cells having some endogenous ARA24, but expression of an ARA24 antisense RNA reduces AR receptor transactivation.

Androgen receptor mutations do not account for all cases of androgen-independent tumors, because some androgen-independent tumors retain wild-type AR. A significant percentage of androgen-insensitive tumors have been correlated with reduced expression of retinoblastoma protein (Rb) (Bookstein, et al., *Science* 247:712-715, (1990)), expression a truncated Rb protein (Bookstein, et al. *Proc. Natl. Acad. Sci. USA* 87:7762-7766 (1990)), or a missing Rb allele (Brooks, et al. *Prostate* 26:35-39, (1995)). The prostate cancer cell line DU145 has an abnormal short mRNA transcript of Rb exon 21 (Sarkar, et al. *Prostate* 21:145-152(1992)) and transfection of the wild-type Rb gene into DU145 cells was shown to repress the malignant phenotype (Bookstein, et al. *Proc. Natl. Acad. Sci. USA* 87:7762-7766 (1990)).

Rb functions in the control of cell proliferation and

differentiation(Weinberg, R.A., *Cell* 81:323-330 (1995);
Kranenburg et al., *FEBS Lett.* 367:103-106 (1995)). In
resting cells, hypophosphorylated Rb prevents inappropriate
entry of cells into the cell division cycle.

5 Phosphorylation of Rb by cyclin-dependent kinases relieves
Rb-mediated growth suppression, and allows for cell
proliferation(Dowdy et al., *Cell* 73:499-511 (1993); Chen et
al., *Cell* 58:1193-1198 (1989)). Conversely,
dephosphorylation of Rb during G1 progression induces

10 growth arrest or cell differentiation(Chen et al. (1989);
Mihara et al., *Science* 246:1300-1303 (1989)). In dividing
cells, Rb is dephosphorylated during mitotic exit and G1
entry(Ludlow et al., *Mol. Cell. Biol.* 13:367-372 (1993)).
This dephosphorylation activates Rb for the ensuing G1

15 phase of the cell cycle, during which Rb exerts its growth
suppressive effects.

We investigated the role of Rb in AR transactivation
as detailed in the examples. We found that Rb can induce
transcriptional activity of wtAR or mARs877t in the

20 presence of DHT, E2, or HF, and mARe708k in the presence of
DHT. We also discovered that Rb and ARA70 transcriptional
activity act synergistically to enhance transcriptional
activity of ARs. The sequence of the cloned Rb gene and
the deduced amino acid sequence of the ORF are shown in SEQ

25 ID NO:7 and SEQ ID NO:8, respectively. An Rb polypeptide
is a polypeptide that is substantially homologous to SEQ ID
NO:8, that interacts with the N-terminal domain of AR, and
which acts synergistically with ARA70 in enhancing
transactivation by AR.

30 In the examples, various eukaryotic cell types,
including yeast, prostate cells having mutant AR and cells
lacking AR, were used to evaluate the ability of the
putative androgen coactivators to enhance transactivation
by AR. It is expected that in the method of the present

35 invention, any eukaryotic cell could be employed in an
assay for AR activity. This feature allows the
investigator flexibility in designing assays.

As described below, cells were transfected using a calcium phosphate technique. It is expected that the method of the present invention could be practiced using any transfection means including, for example,

5 electroporation or particle bombardment.

Changes in the level of transactivation by AR can be assessed by any means, including measuring changes in the level of mRNA for a gene under the control of AR, or by quantitating the amount of a particular protein expressed
10 using an antibody specific for a protein, the expression of which is under the control of AR. Most conveniently, transactivation by AR can be assessed by means of a reporter gene.

As used herein, a reporter gene is a gene under the
15 control of an androgen receptor, the gene encoding a protein susceptible to quantitation by a colormetric or fluorescent assay. In the examples below, a chloramphenicol acetyltransferase or a luciferase gene were used as reporter genes. The gene may either be resident in
20 a chromosome of the host cell, or may be introduced into the host cell by cotransfection with the coactivator gene.

The following nonlimiting examples are intended to be purely illustrative.

EXAMPLES

25 Plasmid construction

A human prostate library in pACT2 yeast expression vector (a gift from Dr. S. Elledge) consists of the GAL4 activation domain (GAL4AD, a.a. 768-881) fused with human prostate cDNA.

30 pSG5 wtAR was constructed as described previously (Yeh and Chang, Proc. Natl. Acad. Sci USA 93:5517-5521, 1996).

pGAL0-AR (wild-type) was obtained from D. Chen (University of Massachusetts). pGAL0 contains the GAL4 DNA binding domain (DBD).

35 For construction of pAS2-wtAR or -mAR, the C-terminal fragments (aa 595-918) from wtAR, mARt877s (Dr. S.P. Balk,

Beth Israel Hospital, Boston, MA), or mARe708k (H. Shim, Hyogo Medical College, Japan) were inserted in pAS2 yeast expression vector (Clontech). Another AR mutant (mARv888m), derived from androgen insensitive syndrome patient, was constructed as previously described (Mowszowicz, et al. Endocrine 1:203-209, 1993).

pGAL4-VP16 was used to construct a fusion of ARA70. pGAL4-VP16 contains the GAL4 DBD linked to the acidic activation domain of VP16.

10 pCMX-Gal-N-RB and pCMX-VP16-AR were constructed by inserting fragments Rb (aa 370-928) and AR (aa 590-918) into pCMX-gal-N and pCMX-VP16, respectively. The sequence of construction junction was verified by sequencing.

pYX-ARA24/Ran was constructed by placing the ARA24
15 gene under the control of the gal-1 promoter of yeast expression plasmid pYX243 (Ingenus). A cDNA fragment encoding the AR poly-Q stretch and its flanking regions (AR a.a. 11-208) was ligated to a PAS2 yeast expression plasmid for use as bait in the two hybrid assay. AR cDNAs of
20 different poly-Q lengths that span the same AR poly-Q region as our bait plasmid were constructed in pAS2 in the same way, for yeast two-hybrid liquid culture β -gal assay. These AR bait plasmids with poly-Q lengths of 1, 25, 49 were all transformed into yeast Y190 and found to not be
25 autonomously active. pCMV-antisense ARA24/Ran (ARA24as) expression plasmid was constructed by inserting a 334-bp *Bgl* II fragment of ARA24/Ran, which spans 5'-untranslated region and the translation start codon of ARA24/Ran (nucleotides 1-334 of SEQ ID NO:5), into pCMV vector in the
30 antisense orientation. The MMTV-CAT and MMTV-Luc reporter genes were used for AR transactivation assay. pSG5-AR and pSV- β gal are under the regulation of SV40 promoter and β -globulin gene intron-1 enhancer. p6R-ARQ1, p6R-ARQ25, p6R-ARQ49 were kindly provided by Dr. Roger L. Meisfield
35 (Chamberlain, et al. Nucleic Acids Res. 22:3181-3186, 1994)

pSG5-GAL4DBD-ARA24 was generated by inserting the coding sequence of Gal4DBD-ARA24 hybrid protein into pSG5

vector. pVP16-ARN-Q1, pVP16-ARN-Q25, pVP16-ARN-Q25, pVP16-ARN-Q35, pVP16-ARN-Q49 were generated by inserting each poly-Q AR N-terminal domain (a.a. 34-555) into pVP16 vector (Clontech) to be expressed as a VP16AD hybrid protein.

5 GAL0AR plasmid, which contains GAL4DBD fused to E region of human AR, was a gift from Dr. D. Chen. The pSG5-CAT reporter plasmid (Clontech) contains five GAL4 binding sites upstream of the E1b TATA box, linked to the CAT gene.

pSG5-AR and pSG5-ARA70 were constructed as previously
10 described (Yeh and Chang, Proc. Natl. Acad. Sci USA 93:5517-5521, 1996). Two mutants of the AR gene (mAR877 derived from prostate cancer, codon 877 mutation Thr to Ala; and mAR708 derived from partial androgen insensitive syndrome (PIAS), codon 708 mutation Glu to
15 Lys), were provided by S. Balk (Beth Israel Hospital, Boston) and H. Shima (Hyogo Medical College, Japan), respectively.

Clones used in the two-hybrid system to evaluate the role of Rb in AR transactivation were made by ligating an
20 Rb fragment (aa 371-928) to the DBD of GAL4. Similarly, near full-length (aa 36-918) AR (nAR) and AR-LBD (aa 590-918) fragments ligated to transcriptional activator VP16.

Screening of prostate cDNA library by a yeast two-hybrid system for ARAs associated with the ligand binding domain

25 To identify ARA coactivators interact with the LBD, a pACT2-prostate cDNA library was cotransformed into Y190 yeast cells with a plasmid of pAS2mAR(mART877S) which contains GAL4DBD(aa 1-147) fused with the C-terminal domain of this mAR. Transformants were selected for growth on
30 SD plates with 3-aminotriazole (25mM) and DHT (100nM) lacking histidine, leucine and tryptophan (-3SD plates). Colonies were also filter-assayed for β -galactosidase activity. Plasmid DNA from positive cDNA clones were found to interact with mtART877s but not GAL4TR4 was isolated
35 from yeast, amplified in *E. coli*, and the inserts confirmed by DNA sequencing.

To identify clones that interact with the poly-Q region of the N-terminal domain, the AR poly-Q stretch (aa 11-208) was inserted into the pAS2 yeast expression plasmid and cotransformed into Y190 yeast cells with a human brain
5 cDNA library fused to the Gal4 activation domain. Transformants were selected for growth on SD plates lacking histidine, leucine and tryptophan and supplemented with 3-aminotriazole (40 mM).

Amplification and characterization of ARA clones

10 Full length DNA sequences comprising two coactivators, designated ARA54 (SEQ ID NO:1) and ARA55 (SEQ ID NO:3), that were found to interact with mARt877s were isolated by 5'RACE PCR using Marathon cDNA Amplification Kit (Clontech) according to the manufacturer's protocol.

15 The missing 5' coding region of the ARA54 gene was isolated from H1299 cells using the gene-specific antisense primer shown in SEQ ID NO:9 and following PCR reaction conditions: 94°C for 1 min, 5 cycles of 94°C for 5 sec-72°C for 3 min, 5 cycles of 94°C for 5 sec-70°C for 3 min, then
20 25 cycles of 94°C for 5 sec-68°C for 3 min. The PCR product was subcloned into pT7-Blue vector (Novagen) and sequenced.

ARA55 was amplified by PCR from the HeLa cell line using an ARA55-specific antisense primer (SEQ ID NO:10) and
25 the PCR reaction conditions described for isolation of ARA54.

Using the 5'RACE-PCR method, we were able to isolate a 1721 bp DNA fragment (SEQ ID NO:1) from the H1299 cell line with an open reading frame that encodes a novel protein 474
30 amino acids in length (SEQ ID NO:2). The *in-vitro* translation product is a polypeptide with an apparent molecular mass of 54±2 kDa, consistent with the calculated molecular weight (53.8 kDa). The middle portion of ARA54 (a.a. 220-265 of SEQ ID NO:2) contains a cysteine-rich
35 region that may form a zinc finger motif called the RING finger, defined as CX₂CX₉₋₂₇CXHX₂CX₂CX₆₋₁₇CX₂C (SEQ ID NO: 11),

a domain conserved among several human transcriptional factor or proto-oncogeny proteins, including BRCA1, RING1, PML and MEL-18 (Miki et al., *Science* 266:66-71 (1994); Borden et al., *EMBO J.* 14:1532-1541 (1995); Lovering et al., *Proc. Natl. Acad. Sci. USA* 90:2112-2116 (1993); Blake et al., *Oncogene* 6: 653-657 (1991); Ishida et al., *Gene* 129:249-255 (1993)). In addition, ARA54 also contains a second cysteine-rich motif which has a B box like structure located at 43 amino acids downstream from the RING finger motif. However, ARA54 differs from members of the RING finger-B-box family in that it lacks a predicted coiled-coil domain immediately C-terminal to the B box domain, which is highly conserved in the RING finger-B-box family. Therefore, ARA54 may represent a new subgroup of this family.

The full-length human ARA55 has an open reading frame that encodes a 444 aa polypeptide (SEQ ID NO:4) with a predicted molecular weight of 55 kD that ARA55 shares 91% homology with mouse hic5. Human ARA55 has four LIM motifs in the C-terminal region. An LIM motif is a cysteine-rich zinc-binding motif with consensus sequence: CX₂CX₁₆.
23HX₂CX₂CX₂CX₁₆₋₂₁CX₂(C,H,D) (SEQ ID NO:12) (Sadler, et al., *J. Cell Biol.* 119:1573-1587(1992)). Although the function of the LIM motif has not been fully defined, some data suggest that it may play a role in protein-protein interaction (Schmeichel & Beckerle, *Cell* 79:211-219, 1994). Among all identified SR associated proteins, only ARA55 and thyroid hormone interacting protein 6 (Trip 6) (Lee, et al. *Mol. Endocrinol.* 9:243-254 (1995)) have LIM motifs.

A clone that showed strong interaction with the poly-Q bait was identified and subsequently subjected to sequence analysis. This clone contains 1566 bp insert (SEQ ID NO:5) with an open reading frame encoding a 216 aa polypeptide (SEQ ID NO:6) with a calculated molecular weight of 24 kDa. GenBank sequence comparison showed that this clone has the same open reading frame sequence as Ran/TC4, an abundant ras-like small GTPase involved in nucleocytoplasmic

transport that is found in a wide variety of cell types (Beddow *et al.*, Proc. Natl. Acad. Sci. U.S.A. 92:3328-3332, 1995). Accordingly, the factor was designated ARA24/Ran. The cDNA sequence of the ARA24 clone (SEQ ID NO:5) (GenBank
5 accession number AF052578) is longer than that of the published ORF for human Ran, in that it includes 24 and 891 bp of 5'- and 3'-untranslated regions, respectively.

Northern Blotting

The total RNA (25 μ g) was fractionated on a 1%
10 formaldehyde-MOPS agarose gel, transferred onto a Hybond-N nylon membrane (Amersham) and prehybridized. A probe corresponding to the 900 bp C-terminus of ARA55 or an ARA54-specific sequence was ³²P-labeled *in vitro* using Random Primed DNA Labeling Kit (Boehringer-Mannheim)
15 according to the manufacture's protocol and hybridized overnight. After washing, the blot was exposed and quantified by Molecular Dynamics PhosphorImager. β -actin was used to monitor the amount of total RNA in each lane.

Northern blot analysis indicated the presence of a 2
20 kb ARA55 transcript in Hela and prostate PC3 cells. The transcript was not detected in other tested cell lines, including HepG2, H1299, MCF7, CHO, PC12, P19, and DU145 cells. The ARA54 transcript was found in H1299 cells, as well as in prostate cancer cell lines PC3 and LNCaP.

25 Co-immunoprecipitation of AR and ARAs

Lysates from *in-vitro* translated full-length of AR and ARA54 were incubated with or without 10⁻⁸ M DHT in the modified RIPA buffer (50mM Tris-HCL pH 7.4, 150mM NaCl, 5mM EDTA, 0.1% NP40, 1mM PMSF, aprotinin, leupeptin, pepstatin,
30 0.25% Na-deoxycholate, 0.25% gelatin) and rocked at 4°C for 2 hr. The mixture was incubated with rabbit anti-His•tag polyclonal antibodies for another 2 hr and protein A/G PLUS -Agarose (Santa Cruz) were added and incubated at 4°C for additional 2 hr. The conjugated beads were washed 4 times
35 with RIPA buffer, boiled in SDS sample buffer and analyzed

by 8% SDS/PAGE and visualized by STORM 840 (Molecular Dynamics).

ARA54 and AR were found in a complex when immunoprecipitated in the presence of 10^{-8} M DHT, but not in the absence of DHT. This result suggests that ARA54 interacts with AR in an androgen-dependent manner.

Interaction between recombinant full length human AR and ARA24/Ran proteins further examined by co-immunoprecipitation, followed by SDS-PAGE and western blotting. Results of the co-immunoprecipitation assay indicate that ARA24/Ran interacts directly with AR. The phosphorylation state of bound guanine nucleotide to the small GTPases does not affect this interaction.

AR pull-down assay using GST-Rb

Full-length Rb fused to glutathione-S-transferase (ST-Rb₁₋₉₂₈) was expressed and purified from E. coli. strain B121pLys as described recently (Zarkowska & Mitnacht, J. Biol. Chem. 272:12738-12746, 1997). Approximately 2 μ g of His-tag column purified baculovirus AR was mixed with GST-loaded glutathione-Sepharose beads in 1 ml of NET-N (20 mM Tris-HCL(pH 8.0, 100 mM NaCl, 1 mM EDTA, 0.5%(v/v) Nonidet P-40) and incubated with gentle rocking for 3 hr at 4°C.

Following low-speed centrifugation to pellet the beads, the clarified supernatant was mixed with GST-Rb-loaded glutathione-Sepharose beads in the presence or absence of 10 mM DHT and incubated for an additional 3 hr with gentle rocking at 4°C. The pelleted beads were washed 5 times with NET-N, mixed with SDS-sample buffer, boiled, and the proteins separated by electrophoresis on a 7.5% polyacrylamide gel. A Western blot of the gel was incubated with anti-AR polyclonal antibody NH27 and developed with alkaline phosphatase-conjugated secondary antibodies.

AR was coprecipitated with GST-Rb, but not GST alone, indicating that AR and Rb are associated in a complex together.

Transfection Studies

Human prostate cancer DU145 or PC3 cells, or human lung carcinoma cells NCI H1299 were grown in Dulbecco's minimal essential medium (DMEM) containing penicillin (25U/ml), streptomycin (25 μ g/ml), and 5% fetal calf serum (FCS). One hour before transfection, the medium was changed to DMEM with 5% charcoal-stripped FCS. Phenol red-free and serum-free media were used on the experiments employing E2 or TGF β , respectively. A β -galactosidase expression plasmid, pCMV- β -gal, was used as an internal control for transfection efficiency.

Cells were transfected using the calcium phosphate technique (Yeh, et al. Molec. Endocrinol. 8:77-88, 1994). The medium was changed 24 hr posttransfection and the cells treated with either steroid hormones or hydroxyflutamide, and cultured for an additional 24 hr. Cells were harvested and assayed for CAT activity after the cell lysates were normalized by using β -galactosidase as an internal control. Chloramphenicol acetyltransferase (CAT) activity was visualized by PhosphorImager (Molecular Dynamics) and quantitated by ImageQuant software (Molecular Dynamics).

Mammalian Two-Hybrid Assay

The mammalian two-hybrid system employed was essentially the protocol of Clontech (California), with the following modifications. In order to obtain better expression, the GAL4DBD (a.a. 1-147) was fused to pSG5 under the control of an SV40 promoter, and named pGAL0. The hinge and LBD of wtAR were then inserted into pGAL0. Similarly, the VP16 activation domain was fused to pCMX under the control of a CMV promoter, and designated pCMX-VP16 (provided by Dr. R.M. Evan).

The DHT-dependent interaction between AR and ARA54 was confirmed in prostate DU145 cells using two-hybrid system with CAT reporter gene assay. Transient transfection of either ARA54 or wtAR alone showed negligible transcriptional activity. However, coexpression of AR with

ARA54 in the presence of 10^{-8} M DHT significantly induced CAT activity.

ARA54 functions as a coactivator relatively specific for AR-mediated transcription. ARA54 induces the transcriptional activity of AR and PR by up to 6 fold and 3-5 fold, respectively. In contrast, ARA54 showed only marginal effects (less than 2 fold) on GR and ER in DU145 cells. These data suggest that ARA54 is less specific to AR as relative to ARA70, which shows higher specificity to AR. However, we can not rule out the possibility that ARA54 might be more general to other steroid receptors in other cell types under different conditions.

Coexpression of ARA54 with SRC-1 or ARA70 was found to enhance AR transcriptional activity additively rather than synergistically. These results indicate that these cofactors may contribute individually to the proper or maximal AR-mediated transcriptional activity.

Since the C-terminal region of ARA54 (a.a. 361-471 of SEQ ID NO:2) isolated from prostate cDNA library has shown to be sufficient to interact with AR in yeast two-hybrid assays, we further investigated whether it could squelch the effect of ARA54 on AR-activated transcription in H1299 cells, which contain endogenous ARA54. The C-terminal region of ARA54 inhibits AR-mediated transcription by up to 70%; coexpression of exogenous full-length ARA54 reverses this squelching effect in a dose-dependent manner. These results demonstrate that the C-terminal domain of ARA54 can serve as a dominant negative inhibitor, and that ARA54 is required for the proper or maximal AR transactivation in human H1299 cells.

Examination of the effect of ARA54 on the transcriptional activities of wtAR and mtARs in the presence of DHT, E2 and HF revealed differential ligand specificity. Translational activation of wtAR occurred in the presence of DHT (10^{-10} to 10^{-8} M); coexpression of ARA54 enhanced transactivation by another 3-5 fold. However, wtAR responded only marginally to E2 (10^{-9} - 10^{-7} M) or HF

(10^{-7} - 10^{-5} M) in the presence or absence of ARA54. As expected, the positive control, ARA70, is able to enhance the AR transcriptional activity in the presence of 10^{-9} - 10^{-7} M E2 and 10^{-7} - 10^{-5} HF, that matches well with previous reports (Yeh, PNAS, Miyamoto, PNAS).

The AR mutants Art877a, which is found in many prostate tumors (23), and Are708k, found in a yeast genetic screening (24) and a patient with partial androgen insensitivity, exhibited differential specificity for ligands. In the absence of ARA54, Art877a responded to E2 (10^{-9} - 10^{-7} M) and HF (10^{-7} - 10^{-5} M), and ARA54 could further enhance E2- or HF-mediated AR transactivation. These results suggested that mtARs might also require cofactors for the proper or maximal DHT-, E2-, or HF-mediated AR transcriptional activity. The DHT response of mAre708k was only a slightly less sensitive than that of wtAR or mArt877s, whereas E2 and HF exhibited no agonistic activity toward Are708k. Together, these results imply that the change of residue 708 on AR might be critical for the interaction of the antiandrogen-ARE708k-ARA54 complex, and that both AR structure and coactivators may play a role in determining ligand specificity.

CAT activity in DU145 cells cotransfected with a plasmid encoding the hormone binding domain of wtAR fused to the GAL4 DBD (GAL0AR) and a plasmid encoding full-length ARA55 fused to the activation domain of VP16 (VP16-ARA55) was significantly induced by the cotransfection of VP16-ARA55 and GAL0AR in the presence of 10 nM DHT, but not induced by E2 or HF. Combination of GAL0 empty vector and VP16-ARA55 did not show any CAT activity. Combination of GAL0AR and VP16 vector showed negligible CAT activity. These results indicate that ARA55 interacts with AR in an androgen-dependent manner.

Transient transfection assays were conducted to investigate the role of ARA55 in the transactivation activity of AR. DU145 cells were cotransfected with MMTV-CAT reporter, increasing amounts of ARA55 and wtAR under

eukaryotic promoter control. Ligand-free AR has minimal MMTV-CAT reporter activity in the presence or absence of ARA55. ARA55 alone also has only minimal reporter activity. Addition of 10 nM DHT resulted in 4.3 fold increase of AR transcriptional activity and ARA55 further increased this induction by 5.3 fold (from 4.3 fold to 22.8 fold) in a dose-dependent manner. The induced activity reached a plateau at the ratio of AR:ARA55 to 1:4.5. Similar results were obtained using PC3 cells with DU145 cells, or using a CAT reporter gene under the control of a 2.8 kb promoter region of a PSA gene. The C-terminus of ARA55 (ARA55₂₅₁₋₄₄₄) (a.a. 251-444 of SEQ ID NO:4) did not enhance CAT activity. Cotransfection of PC3 cells, which contain endogenous ARA55, with ARA55₂₅₁₋₄₄₄, AR and MMTV-CAT reporter in the presence of 10 nM DHT demonstrated dramatically reduced AR transcriptional activity relative to cells transfected with AR and MMTV-CAT alone. These results demonstrate that ARA55 is required for the proper or maximal AR transcriptional activity in PC3 cells, and that the C-terminus of ARA55 can serve as a dominant negative inhibitor.

The effect of ARA55 on mARt877s and mARe708k in the presence of DHT and its antagonists, E2, and HF. The mARt877s receptor is found in LNCaP cells and/or advanced prostate cancers and has a point mutation at codon 877 (Thr to Ser) (Gaddipati et al., *Cancer Res.* 54:2861-2864 (1994); Veldscholte et al., *Biochem. Biophys. Commun.* 173:534-540 (1990)). The mARe708k receptor, has a point mutation at codon 708 (Glu to Lys), was isolated by a yeast genetic screening and exhibits reduced sensitivity to HF and E2 relative to wtAR (Wang, C., *PhD thesis of University of Wisconsin -Madison* (1997)). The transcriptional activities of wtAR, mARt877s, and mARe708k are induced by DHT (10^{-11} to 10^{-8} M). ARA55 enhanced the transactivation of all three receptors by 4-8 fold. In the presence of E2 or HF, wtAR responded marginally only at higher concentrations (10^{-7} M for E2 and 10^{-5} M for HF). Cotransfection of wtAR with

ARA55 at a 1:4.5 ratio, however, increases AR transcriptional activity at 10^{-8} - 10^{-7} M for E2 or 10^{-6} to 10^{-5} M for HF. Compared to wtAR, the LNCaP mAR responded much better to E2 and HF and ARA55 significantly enhanced its transcriptional activity. ARA55 may be needed for the proper or maximal DHT-, E2-, or HF-mediated AR transcriptional activity.

The effect of ARA55 on transcriptional activation by GR, PR, and ER was tested in DU145 cells. ARA55 is relatively specific to AR, although it may also enhance GR and PR to a lesser degree, and has only a marginal effect on ER. ARA70 shows much higher specificity to AR than ARA55, relative to the other tested steroid receptors. Although ARA55 enhances AR-mediated transcription to a greater degree than GR-, PR-, or ER-mediated transcription, it appears to be less specific than ARA70.

Because the amino acid sequence of ARA55 has very high homology to mouse hic5, and early studies hic5 suggested this mouse gene expression can be induced by the negative TGF β (Shibanuma et al., *J. Biol. Chem.* 269:26767-26774 (1994)), we were interested to see whether ARA55 could serve as a bridge between TGF β and AR steroid hormone system. Northern blot analysis indicated that TGF β treatment (5 ng/ml) could induce ARA55 mRNA by 2-fold in PC3 cells. In the same cells, TGF β treatment increased AR transcriptional activity by 70%. This induction is weak relative to the affect achieved upon transfection of PC3 cells with exogenous ARA55 (70% vs. 4 fold). This may be related to the differences in the ratios of AR and ARA55. The best ratio of AR:ARA55 for maximal AR transcriptional activity is 1:4.5. Whether other mechanisms may also be involve in this TGF β -induced AR transcriptional activity will be an interesting question to investigate. The unexpected discovery that TGF β may increase AR transcriptional activity via induction of ARA55 in prostate may represent the first evidence to link a negative regulatory protein function in a positive manner, by

inducing the transcriptional activity of AR, the major promoter for the prostate tumor growth.

The ability of ARA55 to induce transcriptional activity of both wtAR and mARt877s in the presence of DHT, E2, and HF suggests an important role for ARA55 in the progression of prostate cancer and the development of resistance to hormonal therapy. Evaluation of molecules that interfere with the function of ARA55 may aid in the identification of potential chemotherapeutic
10 pharmaceuticals.

Human small lung carcinoma H1299 cell line, which has no endogenous AR protein, were transfected with AR and ARA24/Ran. Because ARA24/Ran is one of the most abundant and ubiquitously expressed proteins in various cells, both
15 sense and antisense ARA24/Ran mammalian expression plasmids were tested. Overexpression of sense ARA24/Ran did not significantly enhance the AR transactivation, a result that is not surprising, in view of the abundance of endogenous ARA24/RAN. However, expression of antisense ARA24/Ran
20 (ARA24as) markedly decreased DHT-induced CAT activity in a dose dependent manner. Furthermore, increasing the DHT concentration from 0.1 nM to 10 nM DHT resulted in strong induction of AR transactivation and decreased the inhibitory effect of ARA24as effect, indicating that
25 increased DHT concentration can antagonize the negative effect of ARA24as.

The affinity between ARA24/Ran and AR is inversely related to the length of AR poly-Q stretch. AR transactivation decreases with increasing AR poly-Q length.
30 Reciprocal two-hybrid assays with exchanged fusion partners, Gal4DBD-ARA24/Ran and VP16AD-ARs (a.a. 34-555 with poly-Q lengths of 1, 25, 35, 49 residues) were conducted using mammalian CHO cells. These results consistently show that the affinity between ARA24/Ran and
35 AR poly-Q region is inversely correlated with AR poly-Q length in both yeast and mammalian CHO cells.

The regulation of AR transactivation by ARA24/Ran

correlates with their affinity. These results suggest that ARA24/Ran could achieve differential transactivation of AR, with ARs having different poly-Q length could existing in a single cell or cell system. ARA24as was again used in the ARE-Luc transfection assays to address the role of AR poly-Q length in the regulation of AR by ARA24/Ran. ARs of poly-Q lengths 1, 25, and 49 residues, and increasing amounts (1, 2, and 4 μ g) of ARA24as expression vectors were co-transfected with equal amounts of reporter plasmid (pMMTV-Luc) in CHO cells. Although the basal reporter activity is slightly affected by increasing amounts of antisense ARA24/Ran, ARA24as showed a more significant decrease of AR transactivation. As AR poly-Q length increased, the ARA24as effect on AR transactivation decreased. These results suggest that the affinity of ARA24/Ran for AR and the effect of decreasing ARA24/Ran on AR transactivation faded over the expansion of AR poly-Q length.

Coexpression of Rb and AR expression plasmids in DU145 cells using the mammalian two-hybrid system resulted in a 3 fold increase in CAT activity by cotransfection of near full length AR (nAR, amino acids 36-918) and Rb. Cells cotransfected with nAR and PR-LBD or Rb and ARA70 did not show increased CAT activity. Surprisingly, addition of 10 nM DHT made very little difference in the interaction between Rb and nAR. The inability of Rb to interact with AR-LBD suggest that interaction site of AR is located in N-terminal domain (aa 36 to 590). Together, our data suggest the interaction between Rb and AR is unique in the following ways: first, the interaction is androgen-independent and binding is specific but relatively weak as compared to other AR associated protein, such as ARA70 (3 fold vs. 12 fold induced CAT activity in mammalian two-hybrid assay, data not shown). Second, unlike most identified steroid receptor associated proteins that bind to C-terminal domain of steroid receptor, Rb binds to N-terminal domain of AR. Third, no interaction occurred

between Rb and ARA70, two AR associated proteins in DU145 cells.

DU145 cells containing mutated Rb (Singh et al., Nature 374: 562-565 (1995)) were cultured with charcoal-
5 stripped FCS in the presence or absence of 1 nM DHT. No AR transcriptional activity was observed in DU145 cells transiently transfected with wild type AR and Rb at the ratio of 1:3 in the absence of DHT. When However, AR transcriptional activity could be induced 5-fold when wild
10 type AR was expressed in the presence of 1 nM DHT. Cotransfection of Rb with AR can further enhance the AR transcriptional activity from 5-fold to 21-fold in the presence of 1 nM DHT. As a control, cotransfection of ARA70, the first identified AR coactivator, can further
15 enhance in DU145 cells transcriptional activity from 5-fold to 36-fold. In DU145 cells transfected with Rb, ARA70, and AR, the induction of AR transcriptional activity was synergistically increased from 5-fold to 64-fold. Upon transfection of wild type AR without Rb or ARA70, only
20 marginal induction (less than 2-fold) was detected in the presence of 10 nM E2 or 1 μ M HF. In contrast, cotransfection of the wild type AR with Rb or ARA70 can enhance the AR transcriptional activity to 12-fold (E2) or 3-4 fold (HF), and cotransfection of Rb and ARA₇₀ with AR
25 can further enhance the AR transcriptional activity to 36-fold (E2 or 12-fold (HF)). We then extended these findings to two different AR mutants: mARt877s from a prostate cancer patient and mARe708k from a partial-androgen-insensitive patient. Similar inductions were obtained when
30 wild type AR was replaced by mARt877s. In contrast, while similar induction was also detected in the presence of 1 nM DHT when we replace wild type AR with mARe708k, there was almost no induction by cotransfection of mARe708k with Rb and/or ARA70 in the presence of 10 nM E2 or 1 μ M HF. These
35 results indicated that Rb and ARA70 can synergistically induce the transcriptional activity of wild type AR and mAR877 in the presence of 1 nM DHT, 10 nM E2 or 1 μ M HF.

However, Rb and ARA70 synergistically induce the transcriptional activity of mAR708 only in the presence of 1 nM DHT, but not 10 nM E2 or 1 μ M HF. The fact that Rb and ARA70 can induce transcriptional activity of both wild
5 type AR and mutated AR that occur in many prostate tumors may also argue strongly the importance of Rb and ARA70 in normal prostate as well as prostate tumor. Also, the differential induction of DHT vs. E2/HF may suggest the position of 708 in AR may play vital role for the
10 recognition of androgen vs anti-androgens to AR.

We also examined the effect of Rb and ARA70 on the transcriptional activity of other steroid receptors through their cognate DNA response elements [MMTV-CAT for AR, glucocorticoid receptor (GR), and progesterone receptor
15 (PR); ERE-CAT for estrogen receptor (ER)]. Although Rb and ARA70 can synergistically induce AR transcriptional activity up to 64-fold, Rb and ARA70 can only have marginal induction on the transcriptional activity of GR, PR, and ER in DU145 cells. These results suggest that Rb and ARA70
20 are more specific coactivators for AR in prostate DU145 cells. However, it cannot be ruled out that possibly the assay conditions in prostate DU145 cells are particularly favorable for Rb and ARA70 to function as coactivators for AR only, and Rb and ARA70 may function as stronger
25 coactivators for ER, PR, and GR in other cells or conditions. Failure of Rb to induce transactivation by mutant AR888, which is unable to bind androgen, suggests that while interaction between Rb and AR is androgen-independent, the AR-Rb (and AR-ARA70) complexes require a
30 ligand for the transactivation activity.

The activity of Rb in cell cycle control is related essentially to its ability to bind to several proteins, thus modulating their activity. To date, many cellular proteins have been reported which bind to Rb (Weinberg,
35 R.A., *Cell* 81:323-330 (1995)). These include a number of transcription factors, a putative regulator of ras, a nuclear structural protein, a protein phosphatase, and

several protein kinases. Whether all of these proteins actually complex, and are regulated by Rb, in cells remains to be seen.

Much attention has been given to the functional
5 interaction between Rb and transcription factors. To date, several of these factors have been shown to form complexes with Rb in cells. Such complex formation and subsequent function studies have revealed that the modulating activity of Rb can take the form of repression of transcription as
10 with E2F (Weintraub et al., *Nature* 375:812-815 (1995)), or activation as with NF-IL6 (Chen et al., *Proc. Natl. Acad. Sci. USA* 93:465-469 (1996)) and the hBrm/BRG1 complex (Singh et al., (1995)). Here, we show that Rb can bind to AR and induce the AR transcriptional activity. To our
15 knowledge, this is the first demonstration of a negative growth regulatory protein functioning in a positive manner, by initiating transcription via a signal transduction mechanism involving binding to a nuclear receptor. When
20 place in the context of regulating the cell cycle and differentiation, these data suggest a previously undescribed function for Rb which underscores the importance of this protein in regulating transcription by direct binding to transcription factor, but this protein can also regulate transcription by stimulating at least one
25 type of signal transduction mechanism.

A relationship between Rb expression and response to endocrine therapy of human breast tumor has been suggested (Anderson et al., *J. Pathology* 180:65-70 (1996)). Other studies indicate that Rb gene alterations can occur in all
30 grades and stages of prostate cancer, in localized as well as metastatic disease (Brooks et al., *Prostate* 26:35-39 (1995)). How Rb function may be linked to androgen-dependent status in prostate tumor progression remains unclear. One possible explanation is that Rb alteration
35 may be a necessary event in prostate carcinogenesis for a subset of prostatic neoplasms, which may be also true for the AR expression in prostate tumors.

All publications cited in this application are incorporated by reference.

The present invention is not limited to the exemplified embodiment, but is intended to encompass all
5 such modifications and variations as come within the scope of the following claims.

CLAIMS

WE CLAIM:

1. An isolated polynucleotide comprising a sequence
selected from the group consisting of SEQ ID NO:1 and SEQ
5 ID NO:3.

2. A genetic construct comprising a promoter capable
of causing expression of a protein coding region in a cell,
the promoter operably connected to a protein coding region
encoding the expression of a polypeptide from coding
10 regions of ARA54 or ARA55.

3. The genetic construct of claim 2 wherein the
polypeptide encoded by the protein coding sequence
comprises a sequence selected from the group consisting of
SEQ ID NO:2 and SEQ ID NO:4.

15 4. A eukaryotic host cell comprising the genetic
construct of claim 2.

5. A method for testing the androgenic or
antiandrogenic effect of a chemical compound comprising the
steps of:

- 20 (a) transfecting a host cell with at least one
genetic construct capable of producing in the host cell a
polypeptide selected from the group consisting of ARA54,
ARA55, ARA24, and Rb, the host cell also producing human
androgen receptor protein;
- 25 (b) exposing the cell to the chemical compound; and
(c) measuring the level of transcriptional activity
caused by the androgen receptor.

6. The method of claim 5 wherein the host cell is a prostate cell.

7. The method of claim 5, wherein the cell is a eukaryotic cell that lacks native endogenous androgen
5 receptor, the cell having also an introduced genetic construct producing androgen receptor protein.

8. The method of claim 5, wherein the genetic construct comprises a DNA sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, and
10 SEQ ID NO:7.

9. The method of claim 5, wherein the cell is transfected with a genetic construct comprising a reporter gene expressible in the cell, the expression of said reporter gene being susceptible to detection and
15 quantitation.

10. The method of claim 9, wherein the reporter gene is selected from the group consisting of a chloramphenicol acetyltransferase gene and a luciferase gene.

11. A method for testing the androgenic or antiandrogenic effect of a chemical compound comprising the steps of:

- (a) transfecting a host cell with at least one
5 genetic construct capable of producing in the host cell human androgen receptor protein and a polypeptide selected from the group consisting of ARA54, ARA55, ARA24, and Rb;
- (b) exposing the cell to the chemical compound; and
- (c) measuring the interaction between AR and an AR
10 co-activator.

12. A method as claimed in claim 11 wherein the co-activator is selected from the group consisting of ARA54, ARA55, ARA24 and Rb.

SEQUENCE LISTING

<110> Chang, Chawnshang

<120> Androgen Receptor Coactivators

<130> 920920.90011

<140>

<141>

<150> US 60/100,243

<151> 1998-09-14

<160> 12

<170> PatentIn Ver. 2.0

<210> 1

<211> 1721

<212> DNA

<213> Homo sapien

<220>

<221> CDS

<222> (40)..(1464)

<220>

<221> misc_feature

<222> (1120)..(1452)

<223> Coding sequence and polypeptide region for the
C-terminal domain.

<220>

<221> misc_feature

<222> (697)..(834)

<223> Coding sequence and polypeptide region which may
form a cystein-rich RING finger motif.

<220>

<221> misc_feature

<222> (964)..(1089)

<223> Coding sequence and polypeptide region for a
cystein-rich B box like structure.

<400> 1

```

gggtctctgggt ctccccctctc tgagcaactct gaggtcctt atg tcg tca gaa gat      54
                               Met Ser Ser Glu Asp
                               1           5

cga gaa gct cag gag gat gaa ttg ctg gcc ctg gca agt att tac gat      102
Arg Glu Ala Gln Glu Asp Glu Leu Leu Ala Leu Ala Ser Ile Tyr Asp
                10                15                20

gga gat gaa ttt aga aaa gca gag tct gtc caa ggt gga gaa acc agg      150
Gly Asp Glu Phe Arg Lys Ala Glu Ser Val Gln Gly Gly Glu Thr Arg
                25                30                35

atc tat ttg gat ttg cca cag aat ttc aag ata ttt gtg agc ggc aat      198
Ile Tyr Leu Asp Leu Pro Gln Asn Phe Lys Ile Phe Val Ser Gly Asn
                40                45                50

tca aat gag tgt ctc cag aat agt ggc ttt gaa tac acc att tgc ttt      246
Ser Asn Glu Cys Leu Gln Asn Ser Gly Phe Glu Tyr Thr Ile Cys Phe
                55                60                65

ctg cct cca ctt gtg ctg aac ttt gaa ctg cca cca gat tat cca tcc      294
Leu Pro Pro Leu Val Leu Asn Phe Glu Leu Pro Pro Asp Tyr Pro Ser
                70                75                80                85

tct tcc cca cct tca ttc aca ctt agt ggc aaa tgg ctg tca cca act      342
Ser Ser Pro Pro Ser Phe Thr Leu Ser Gly Lys Trp Leu Ser Pro Thr
                90                95                100

cag cta tct gct cta tgc aag cac tta gac aac cta tgg gaa gaa cac      390
Gln Leu Ser Ala Leu Cys Lys His Leu Asp Asn Leu Trp Glu Glu His
                105                110                115

cgt ggc agc gtg gtc ctg ttt gcc tgg atg caa ttt ctt aag gaa gag      438
Arg Gly Ser Val Val Leu Phe Ala Trp Met Gln Phe Leu Lys Glu Glu
                120                125                130

acc cta gca tac ttg aat att gtc tct cct ttt gag ctc aag att ggt      486
Thr Leu Ala Tyr Leu Asn Ile Val Ser Pro Phe Glu Leu Lys Ile Gly
                135                140                145

tct cag aaa aaa gtg cag aga agg aca gct caa gct tct ccc aac aca      534
Ser Gln Lys Lys Val Gln Arg Arg Thr Ala Gln Ala Ser Pro Asn Thr
                150                155                160                165

gag cta gat ttt gga gga gct gct gga tct gat gta gac caa gag gaa      582
Glu Leu Asp Phe Gly Gly Ala Ala Gly Ser Asp Val Asp Gln Glu Glu
                170                175                180

```

att gtg gat gag aga gca gtg cag gat gtg gaa tca ctg tca aat ctg	630
Ile Val Asp Glu Arg Ala Val Gln Asp Val Glu Ser Leu Ser Asn Leu	
185 190 195	
atc cag gaa atc ttg gac ttt gat caa gct cag cag ata aaa tgc ttt	678
Ile Gln Glu Ile Leu Asp Phe Asp Gln Ala Gln Gln Ile Lys Cys Phe	
200 205 210	
aat agt aaa ttg ttc ctg tgc agt atc tgt ttc tgt gag aag ctg ggt	726
Asn Ser Lys Leu Phe Leu Cys Ser Ile Cys Phe Cys Glu Lys Leu Gly	
215 220 225	
agt gaa tgc atg tac ttc ttg gag tgc agg cat gtg tac tgc aaa gcc	774
Ser Glu Cys Met Tyr Phe Leu Glu Cys Arg His Val Tyr Cys Lys Ala	
230 235 240 245	
tgt ctg aag gac tac ttt gaa atc cag atc aga gat ggc cag gtt caa	822
Cys Leu Lys Asp Tyr Phe Glu Ile Gln Ile Arg Asp Gly Gln Val Gln	
250 255 260	
tgc ctc aac tgc cca gaa cca aag tgc cct tcg gtg gcc act cct ggt	870
Cys Leu Asn Cys Pro Glu Pro Lys Cys Pro Ser Val Ala Thr Pro Gly	
265 270 275	
cag gtc aaa gag tta gtg gaa gca gag tta ttt gcc cgt tat gac cgc	918
Gln Val Lys Glu Leu Val Glu Ala Glu Leu Phe Ala Arg Tyr Asp Arg	
280 285 290	
ctt ctc ctc cag tcc tcc ttg gac ctg atg gca gat gtg gtg tac tgc	966
Leu Leu Leu Gln Ser Ser Leu Asp Leu Met Ala Asp Val Val Tyr Cys	
295 300 305	
ccc cgg cgg tgc tgc cag ctg cct gtg atg cag gaa cct ggc tgc acc	1014
Pro Arg Pro Cys Cys Gln Leu Pro Val Met Gln Glu Pro Gly Cys Thr	
310 315 320 325	
atg ggt atc tgc tcc agc tgc aat ttt gcc ttc tgt act ttg tgc agg	1062
Met Gly Ile Cys Ser Ser Cys Asn Phe Ala Phe Cys Thr Leu Cys Arg	
330 335 340	
ttg acc tac cat ggg gtc tcc cca tgt aag gtg act gca gag aaa tta	1110
Leu Thr Tyr His Gly Val Ser Pro Cys Lys Val Thr Ala Glu Lys Leu	
345 350 355	
atg gac tta cga aat gaa tac ctg caa gcg gat gag gct aat aaa aga	1158
Met Asp Leu Arg Asn Glu Tyr Leu Gln Ala Asp Glu Ala Asn Lys Arg	
360 365 370	
ctt ttg gat caa agg tat ggt aag aga gtg att cag aag gca ctg gaa	1206
Leu Leu Asp Gln Arg Tyr Gly Lys Arg Val Ile Gln Lys Ala Leu Glu	
375 380 385	

gag atg gaa agt aag gag tgg cta gag aag aac tca aag agc tgc cca 1254
 Glu Met Glu Ser Lys Glu Trp Leu Glu Lys Asn Ser Lys Ser Cys Pro
 390 395 400 405

tgt tgt gga act ccc ata gag aaa tta gac gga tgt aac aag atg aca 1302
 Cys Cys Gly Thr Pro Ile Glu Lys Leu Asp Gly Cys Asn Lys Met Thr
 410 415 420

tgt act ggc tgt atg caa tat ttc tgt tgg att tgc atg ggt tct ctc 1350
 Cys Thr Gly Cys Met Gln Tyr Phe Cys Trp Ile Cys Met Gly Ser Leu
 425 430 435

tct aga gca aac cct tac aaa cat ttc aat gac cct ggt tca cca tgt 1398
 Ser Arg Ala Asn Pro Tyr Lys His Phe Asn Asp Pro Gly Ser Pro Cys
 440 445 450

ttt aac cgg ctg ttt tat gct gtg gat gtt gac gac gat att tgg gaa 1446
 Phe Asn Arg Leu Phe Tyr Ala Val Asp Val Asp Asp Asp Ile Trp Glu
 455 460 465

gat gag gta gaa gac tag ttaactactg ctcaagatat ttaactactg 1494
 Asp Glu Val Glu Asp
 470 475

ctcaagatat ggaagtggat tggttttccc taatcttccg tcaagtacac aaagtaactt 1554

tgcggggatat ttaggggtact attcattcac tcttctgctg tagaagatat ggaagaacga 1614

ggttttatatt ttcattgtgt actactgaag aaggtgcatt gatacatattt taaatgtaag 1674

ttgagaaaaa tttataagcc aaagggttcag aaaattaaac tacagaa 1721

<210> 2

<211> 474

<212> PRT

<213> Homo sapien

<400> 2

Met Ser Ser Glu Asp Arg Glu Ala Gln Glu Asp Glu Leu Leu Ala Leu
 1 5 10 15

Ala Ser Ile Tyr Asp Gly Asp Glu Phe Arg Lys Ala Glu Ser Val Gln
 20 25 30

Gly Gly Glu Thr Arg Ile Tyr Leu Asp Leu Pro Gln Asn Phe Lys Ile
 35 40 45

Phe Val Ser Gly Asn Ser Asn Glu Cys Leu Gln Asn Ser Gly Phe Glu
 50 55 60

Tyr Thr Ile Cys Phe Leu Pro Pro Leu Val Leu Asn Phe Glu Leu Pro
 65 70 75 80
 Pro Asp Tyr Pro Ser Ser Ser Pro Pro Ser Phe Thr Leu Ser Gly Lys
 85 90 95
 Trp Leu Ser Pro Thr Gln Leu Ser Ala Leu Cys Lys His Leu Asp Asn
 100 105 110
 Leu Trp Glu Glu His Arg Gly Ser Val Val Leu Phe Ala Trp Met Gln
 115 120 125
 Phe Leu Lys Glu Glu Thr Leu Ala Tyr Leu Asn Ile Val Ser Pro Phe
 130 135 140
 Glu Leu Lys Ile Gly Ser Gln Lys Lys Val Gln Arg Arg Thr Ala Gln
 145 150 155 160
 Ala Ser Pro Asn Thr Glu Leu Asp Phe Gly Gly Ala Ala Gly Ser Asp
 165 170 175
 Val Asp Gln Glu Glu Ile Val Asp Glu Arg Ala Val Gln Asp Val Glu
 180 185 190
 Ser Leu Ser Asn Leu Ile Gln Glu Ile Leu Asp Phe Asp Gln Ala Gln
 195 200 205
 Gln Ile Lys Cys Phe Asn Ser Lys Leu Phe Leu Cys Ser Ile Cys Phe
 210 215 220
 Cys Glu Lys Leu Gly Ser Glu Cys Met Tyr Phe Leu Glu Cys Arg His
 225 230 235 240
 Val Tyr Cys Lys Ala Cys Leu Lys Asp Tyr Phe Glu Ile Gln Ile Arg
 245 250 255
 Asp Gly Gln Val Gln Cys Leu Asn Cys Pro Glu Pro Lys Cys Pro Ser
 260 265 270
 Val Ala Thr Pro Gly Gln Val Lys Glu Leu Val Glu Ala Glu Leu Phe
 275 280 285
 Ala Arg Tyr Asp Arg Leu Leu Leu Gln Ser Ser Leu Asp Leu Met Ala
 290 295 300
 Asp Val Val Tyr Cys Pro Arg Pro Cys Cys Gln Leu Pro Val Met Gln
 305 310 315 320
 Glu Pro Gly Cys Thr Met Gly Ile Cys Ser Ser Cys Asn Phe Ala Phe
 325 330 335

Cys Thr Leu Cys Arg Leu Thr Tyr His Gly Val Ser Pro Cys Lys Val
 340 345 350

Thr Ala Glu Lys Leu Met Asp Leu Arg Asn Glu Tyr Leu Gln Ala Asp
 355 360 365

Glu Ala Asn Lys Arg Leu Leu Asp Gln Arg Tyr Gly Lys Arg Val Ile
 370 375 380

Gln Lys Ala Leu Glu Glu Met Glu Ser Lys Glu Trp Leu Glu Lys Asn
 385 390 395 400

Ser Lys Ser Cys Pro Cys Cys Gly Thr Pro Ile Glu Lys Leu Asp Gly
 405 410 415

Cys Asn Lys Met Thr Cys Thr Gly Cys Met Gln Tyr Phe Cys Trp Ile
 420 425 430

Cys Met Gly Ser Leu Ser Arg Ala Asn Pro Tyr Lys His Phe Asn Asp
 435 440 445

Pro Gly Ser Pro Cys Phe Asn Arg Leu Phe Tyr Ala Val Asp Val Asp
 450 455 460

Asp Asp Ile Trp Glu Asp Glu Val Glu Asp
 465 470

<210> 3

<211> 1335

<212> DNA

<213> Homo sapien

<220>

<221> CDS

<222> (1)..(1335)

<220>

<221> misc_feature

<222> (750)..(1332)

<223> Coding sequence and polypeptide region for the
 C-terminal binding domain

<220>

<221> misc_feature

<222> (631)..(783)

<223> Coding sequence and polypeptide region for a
 cystein rich LIM motif

<220>

<221> misc_feature

<222> (808)..(996)

<223> Coding sequence and polypeptide region for a
cystein rich LIM motif

<220>

<221> misc_feature

<222> (985)..(1137)

<223> Coding sequence and polypeptide region for a
cystein rich LIM motif

<220>

<221> misc_feature

<222> (1162)..(1314)

<223> Coding sequence and polypeptide region for a
cystein rich LIM motif

<400> 3

atg cca agg tca ggg gct ccc aaa gag cgc cct gcg gag cct ctc acc	48
Met Pro Arg Ser Gly Ala Pro Lys Glu Arg Pro Ala Glu Pro Leu Thr	
1 5 10 15	
cct ccc cca tcc tat ggc cac cag cca aca ggg cag tct ggg gag tct	96
Pro Pro Pro Ser Tyr Gly His Gln Pro Thr Gly Gln Ser Gly Glu Ser	
20 25 30	
tca gga gcc tcg ggg gac aag gac cac ctg tac agc acg gta tgc aag	144
Ser Gly Ala Ser Gly Asp Lys Asp His Leu Tyr Ser Thr Val Cys Lys	
35 40 45	
cct cgg tcc cca aag cct gca gcc ccg gcc gcc cct cca ttc tcc tct	192
Pro Arg Ser Pro Lys Pro Ala Ala Pro Ala Ala Pro Pro Phe Ser Ser	
50 55 60	
tcc agc ggt gtc ttg ggt acc ggg ctc tgt gag cta gat cgg ttg ctt	240
Ser Ser Gly Val Leu Gly Thr Gly Leu Cys Glu Leu Asp Arg Leu Leu	
65 70 75 80	
cag gaa ctt aat gcc act cag ttc aac atc aca gat gaa atc atg tct	288
Gln Glu Leu Asn Ala Thr Gln Phe Asn Ile Thr Asp Glu Ile Met Ser	
85 90 95	
cag ttc cca tct agc aag gtg gct tca gga gag cag aag gag gac cag	336
Gln Phe Pro Ser Ser Lys Val Ala Ser Gly Glu Gln Lys Glu Asp Gln	
100 105 110	
tct gaa gat aag aaa aga ccc agc ctc cct tcc agc ccg tct cct ggc	384
Ser Glu Asp Lys Lys Arg Pro Ser Leu Pro Ser Ser Pro Ser Pro Gly	
115 120 125	

ctc cca aag gct tct gcc acc tca gcc act ctg gag ctg gat aga ctg	432
Leu Pro Lys Ala Ser Ala Thr Ser Ala Thr Leu Glu Leu Asp Arg Leu	
130 135 140	
atg gcc tca ctc cct gac ttc cgc gtt caa aac cat ctt cca gcc tct	480
Met Ala Ser Leu Pro Asp Phe Arg Val Gln Asn His Leu Pro Ala Ser	
145 150 155 160	
ggg cca act cag cca ccg gtg gtg agc tcc aca aat gag ggc tcc cca	528
Gly Pro Thr Gln Pro Pro Val Val Ser Ser Thr Asn Glu Gly Ser Pro	
165 170 175	
tcc cca cca gag ccg act gca aag ggc agc cta gac acc atg ctg ggg	576
Ser Pro Pro Glu Pro Thr Ala Lys Gly Ser Leu Asp Thr Met Leu Gly	
180 185 190	
ctg ctg cag tcc gac ctc agc cgc cgg ggt gtt ccc acc cag gcc aaa	624
Leu Leu Gln Ser Asp Leu Ser Arg Arg Gly Val Pro Thr Gln Ala Lys	
195 200 205	
ggc ctc tgt ggc tcc tgc aat aaa cct att gct ggg caa gtg gtg acg	672
Gly Leu Cys Gly Ser Cys Asn Lys Pro Ile Ala Gly Gln Val Val Thr	
210 215 220	
gct ctg ggc cgc gcc tgg cac ccc gag cac ttc gtt tgc gga ggc tgt	720
Ala Leu Gly Arg Ala Trp His Pro Glu His Phe Val Cys Gly Gly Cys	
225 230 235 240	
tcc acc gcc ctg gga ggc agc agc ttc ttc gag aag gat gga gcc ccc	768
Ser Thr Ala Leu Gly Gly Ser Ser Phe Phe Glu Lys Asp Gly Ala Pro	
245 250 255	
ttc tgc ccc gag tgc tac ttt gag cgc ttc tcg cca aga tgt ggc ttc	816
Phe Cys Pro Glu Cys Tyr Phe Glu Arg Phe Ser Pro Arg Cys Gly Phe	
260 265 270	
tgc aac cag ccc atc cga cac aag atg gtg acc gcc ttg ggc act cac	864
Cys Asn Gln Pro Ile Arg His Lys Met Val Thr Ala Leu Gly Thr His	
275 280 285	
tgg cac cca gag cat ttc tgc tgc gtc agt tgc ggg gag ccc ttc gga	912
Trp His Pro Glu His Phe Cys Cys Val Ser Cys Gly Glu Pro Phe Gly	
290 295 300	
gat gag ggt ttc cac gag cgc gag ggc cgc ccc tac tgc cgc cgg gac	960
Asp Glu Gly Phe His Glu Arg Glu Gly Arg Pro Tyr Cys Arg Arg Asp	
305 310 315 320	
ttc ctg cag ctg ttc gcc ccg cgc tgc cag ggc tgc cag ggc ccc atc	1008
Phe Leu Gln Leu Phe Ala Pro Arg Cys Gln Gly Cys Gln Gly Pro Ile	
325 330 335	

ctg gat aac tac atc tcg gcg ctc agc ctg ctc tgg cac ccg gac tgt 1056
 Leu Asp Asn Tyr Ile Ser Ala Leu Ser Leu Leu Trp His Pro Asp Cys
 340 345 350

ttc gtc tgc agg gaa tgc ttc gcg ccc ttc tcg gga ggc agc ttt ttc 1104
 Phe Val Cys Arg Glu Cys Phe Ala Pro Phe Ser Gly Gly Ser Phe Phe
 355 360 365

gag cac gag ggc cgc ccg ttg tgc gag aac cac ttc cac gca cga cgc 1152
 Glu His Glu Gly Arg Pro Leu Cys Glu Asn His Phe His Ala Arg Arg
 370 375 380

ggc tcg ctg tgc ccc acg tgt ggc ctc cct gtg acc ggc cgc tgc gtg 1200
 Gly Ser Leu Cys Pro Thr Cys Gly Leu Pro Val Thr Gly Arg Cys Val
 385 390 395 400

tcg gcc ctg ggt cgc cgc ttc cac ccg gac cac ttc gca tgc acc ttc 1248
 Ser Ala Leu Gly Arg Arg Phe His Pro Asp His Phe Ala Cys Thr Phe
 405 410 415

tgc ctg cgc ccg ctc acc aag ggg tcc ttc cag gag cgc gcc ggc aag 1296
 Cys Leu Arg Pro Leu Thr Lys Gly Ser Phe Gln Glu Arg Ala Gly Lys
 420 425 430

ccc tac tgc cag ccc tgc ttc ctg aag ctc ttc ggc tga 1335
 Pro Tyr Cys Gln Pro Cys Phe Leu Lys Leu Phe Gly
 435 440 445

<210> 4

<211> 444

<212> PRT

<213> Homo sapien

<400> 4

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

Pro Pro Pro Ser Tyr Gly His Gln Pro Thr Gly Gln Ser Gly Glu Ser
 20 25 30

Ser Gly Ala Ser Gly Asp Lys Asp His Leu Tyr Ser Thr Val Cys Lys
 35 40 45

Pro Arg Ser Pro Lys Pro Ala Ala Pro Ala Ala Pro Pro Phe Ser Ser
 50 55 60

Ser Ser Gly Val Leu Gly Thr Gly Leu Cys Glu Leu Asp Arg Leu Leu
 65 70 75 80

Gln Glu Leu Asn Ala Thr Gln Phe Asn Ile Thr Asp Glu Ile Met Ser
 85 90 95

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

Glu His Glu Gly Arg Pro Leu Cys Glu Asn His Phe His Ala Arg Arg
 370 375 380

Gly Ser Leu Cys Pro Thr Cys Gly Leu Pro Val Thr Gly Arg Cys Val
 385 390 395 400

Ser Ala Leu Gly Arg Arg Phe His Pro Asp His Phe Ala Cys Thr Phe
 405 410 415

Cys Leu Arg Pro Leu Thr Lys Gly Ser Phe Gln Glu Arg Ala Gly Lys
 420 425 430

Pro Tyr Cys Gln Pro Cys Phe Leu Lys Leu Phe Gly
 435 440

<210> 5
 <211> 1566
 <212> DNA
 <213> Homo sapien

<220>
 <221> CDS
 <222> (25) .. (675)

<220>
 <221> 3'UTR
 <222> (676) .. (1566)

<220>
 <221> 5'UTR
 <222> (1) .. (24)

<400> 5
 ggcgcttctg gaaggaacgc cgcg atg gct gcg cag gga gag ccc cag gtc 51
 Met Ala Ala Gln Gly Glu Pro Gln Val
 1 5

cag ttc aaa ctt gta ttg gtt ggt gat ggt ggt act gga aaa acg acc 99
 Gln Phe Lys Leu Val Leu Val Gly Asp Gly Gly Thr Gly Lys Thr Thr
 10 15 20 25

ttc gtg aaa cgt cat ttg act ggt gaa ttt gag aag aag tat gta gcc 147
 Phe Val Lys Arg His Leu Thr Gly Glu Phe Glu Lys Lys Tyr Val Ala
 30 35 40

acc ttg ggt gtt gag gtt cat ccc cta gtg ttc cac acc aac aga gga 195
 Thr Leu Gly Val Glu Val His Pro Leu Val Phe His Thr Asn Arg Gly
 45 50 55

```

cct att aag ttc aat gta tgg gac aca gcc ggc cag gag aaa ttc ggt 243
Pro Ile Lys Phe Asn Val Trp Asp Thr Ala Gly Gln Glu Lys Phe Gly
      60              65              70

gga ctg aga gat ggc tat tat atc caa gcc cag tgt gcc atc ata atg 291
Gly Leu Arg Asp Gly Tyr Tyr Ile Gln Ala Gln Cys Ala Ile Ile Met
      75              80              85

ttt gat gta aca tcg aga gtt act tac aag aat gtg cct aac tgg cat 339
Phe Asp Val Thr Ser Arg Val Thr Tyr Lys Asn Val Pro Asn Trp His
      90              95              100              105

aga gat ctg gta cga gtg tgt gaa aac atc ccc att gtg ttg tgt ggc 387
Arg Asp Leu Val Arg Val Cys Glu Asn Ile Pro Ile Val Leu Cys Gly
      110              115              120

aac aaa gtg gat att aag gac agg aaa gtg aag gcg aaa tcc att gtc 435
Asn Lys Val Asp Ile Lys Asp Arg Lys Val Lys Ala Lys Ser Ile Val
      125              130              135

ttc cac cga aag aag aat ctt cag tac tac gac att tct gcc aaa agt 483
Phe His Arg Lys Lys Asn Leu Gln Tyr Tyr Asp Ile Ser Ala Lys Ser
      140              145              150

aac tac aac ttt gaa aag ccc ttc ctc tgg ctt gct agg aag ctc att 531
Asn Tyr Asn Phe Glu Lys Pro Phe Leu Trp Leu Ala Arg Lys Leu Ile
      155              160              165

gga gac cct aac ttg gaa ttt gtt gcc atg cct gct ctc gcc cca cca 579
Gly Asp Pro Asn Leu Glu Phe Val Ala Met Pro Ala Leu Ala Pro Pro
      170              175              180              185

gaa gtt gtc atg gac cca gct ttg gca gca cag tat gag cac gac tta 627
Glu Val Val Met Asp Pro Ala Leu Ala Ala Gln Tyr Glu His Asp Leu
      190              195              200

gag gtt gct cag aca act gct ctc ccg gat gag gat gat gac ctg tga 675
Glu Val Ala Gln Thr Thr Ala Leu Pro Asp Glu Asp Asp Asp Leu
      205              210              215

gaatgaagct ggagcccagc gtcagaagtc tagttttata ggcagctgtc ctgtgatgtc 735

agcgggtgcag cgtgtgtgcc acctcattat tatctagcta agcgggaacat gtgctttatc 795

tgtgggatgc tgaaggagat gagtgggctt cggagtgaat gtggcagttt aaaaaataac 855

ttcattgttt ggacctgcat atttagctgt ttggacgcag ttgattcctt gagtttcata 915

tataagactg ctgcagtcac atcacaatat tcagtgggtga aatcttgttt gttactgtca 975

ttcccatgcc ttttcttttag aatcagaata aagttgtatt tcaaatatct aagcaagtga 1035

```

actcatccct tgtttataaa tagcatttgg aaaccactaa agtagggaag ttttatgcc 1095
 tgtaaatatt tgaattgcct tgcttttatt acttaatttg aaatctattg ggtaatttc 1155
 tccctatggt tatttttgta catttgagcc atgtcacaca aactgatgat gacaggtcag 1215
 cagtattcta tttggttaga agggttacat ggtgtaaata ttagtgagcag taagctaaag 1275
 cagtgtttgc tccaccttca tattggctag gtagggtcac ctagggaagc acttgctcaa 1335
 aatctgtgac ctgtcagaat aaaaatgtgg tttgtacata tcaaatagat attttaaggg 1395
 taatattttc ttttatggca aaagtaatca tgttttaatg tagaacctca aacaggatgg 1455
 aacatcagtg gatggcagga ggttggaat tcttgctgtt aaaaataatt acaaattttg 1515
 cactttttgt ttgaatgtta gatgcttagt gtgaagttga tacgcaagcc g 1566

<210> 6

<211> 216

<212> PRT

<213> Homo sapien

<400> 6

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

Gln Tyr Tyr Asp Ile Ser Ala Lys Ser Asn Tyr Asn Phe Glu Lys Pro
 145 150 155 160

Phe Leu Trp Leu Ala Arg Lys Leu Ile Gly Asp Pro Asn Leu Glu Phe
 165 170 175

Val Ala Met Pro Ala Leu Ala Pro Pro Glu Val Val Met Asp Pro Ala
 180 185 190

Leu Ala Ala Gln Tyr Glu His Asp Leu Glu Val Ala Gln Thr Thr Ala
 195 200 205

Leu Pro Asp Glu Asp Asp Asp Leu
 210 215

<210> 7

<211> 4839

<212> DNA

<213> Homo sapien

<220>

<221> CDS

<222> (138) .. (2924)

<400> 7

tccgggttttt ctcaggggac gttgaaatta tttttgtaac gggagtcggg agaggacggg 60

gcgtgccccg cgtgcgcgcg cgtcgtcctc cccggcgctc ctccacagct cgctggctcc 120

cgccgcggaa aggcgtc atg ccg ccc aaa acc ccc cga aaa acg gcc gcc 170
 Met Pro Pro Lys Thr Pro Arg Lys Thr Ala Ala
 1 5 10

acc gcc gcc gct gcc gcc gcg gaa ccc ccg gca ccg ccg ccg ccg ccc 218
 Thr Ala Ala Ala Ala Ala Ala Glu Pro Pro Ala Pro Pro Pro Pro Pro
 15 20 25

cct cct gag gag gac cca gag cag gac agc ggc ccg gag gac ctg cct 266
 Pro Pro Glu Glu Asp Pro Glu Gln Asp Ser Gly Pro Glu Asp Leu Pro
 30 35 40

ctc gtc agg ctt gag ttt gaa gaa aca gaa gaa cct gat ttt act gca 314
 Leu Val Arg Leu Glu Phe Glu Glu Thr Glu Glu Pro Asp Phe Thr Ala
 45 50 55

tta tgt cag aaa tta aag ata cca gat cat gtc aga gag aga gct tgg 362
 Leu Cys Gln Lys Leu Lys Ile Pro Asp His Val Arg Glu Arg Ala Trp
 60 65 70 75

tta act tgg gag aaa gtt tca tct gtg gat gga gta ttg gga ggt tat	410
Leu Thr Trp Glu Lys Val Ser Ser Val Asp Gly Val Leu Gly Gly Tyr	
80 85 90	
att caa aag aaa aag gaa ctg tgg gga atc tgt atc ttt att gca gca	458
Ile Gln Lys Lys Lys Glu Leu Trp Gly Ile Cys Ile Phe Ile Ala Ala	
95 100 105	
gtt gac cta gat gag atg tcg ttc act ttt act gag cta cag aaa aac	506
Val Asp Leu Asp Glu Met Ser Phe Thr Phe Thr Glu Leu Gln Lys Asn	
110 115 120	
ata gaa atc agt gtc cat aaa ttc ttt aac tta cta aaa gaa att gat	554
Ile Glu Ile Ser Val His Lys Phe Phe Asn Leu Leu Lys Glu Ile Asp	
125 130 135	
acc agt acc aaa gtt gat aat gct atg tca aga ctg ttg aag aag tat	602
Thr Ser Thr Lys Val Asp Asn Ala Met Ser Arg Leu Leu Lys Lys Tyr	
140 145 150 155	
gat gta ttg ttt gca ctc ttc agc aaa ttg gaa agg aca tgt gaa ctt	650
Asp Val Leu Phe Ala Leu Phe Ser Lys Leu Glu Arg Thr Cys Glu Leu	
160 165 170	
ata tat ttg aca caa ccc agc agt tcg ata tct act gaa ata aat tct	698
Ile Tyr Leu Thr Gln Pro Ser Ser Ser Ile Ser Thr Glu Ile Asn Ser	
175 180 185	
gca ttg gtg cta aaa gtt tct tgg atc aca ttt tta tta gct aaa ggg	746
Ala Leu Val Leu Lys Val Ser Trp Ile Thr Phe Leu Leu Ala Lys Gly	
190 195 200	
gaa gta tta caa atg gaa gat gat ctg gtg att tca ttt cag tta atg	794
Glu Val Leu Gln Met Glu Asp Asp Leu Val Ile Ser Phe Gln Leu Met	
205 210 215	
cta tgt gtc ctt gac tat ttt att aaa ctc tca cct ccc atg ttg ctc	842
Leu Cys Val Leu Asp Tyr Phe Ile Lys Leu Ser Pro Pro Met Leu Leu	
220 225 230 235	
aaa gaa cca tat aaa aca gct gtt ata ccc att aat ggt tca cct cga	890
Lys Glu Pro Tyr Lys Thr Ala Val Ile Pro Ile Asn Gly Ser Pro Arg	
240 245 250	
aca ccc agg cga ggt cag aac agg agt gca cgg ata gca aaa caa cta	938
Thr Pro Arg Arg Gly Gln Asn Arg Ser Ala Arg Ile Ala Lys Gln Leu	
255 260 265	
gaa aat gat aca aga att att gaa gtt ctc tgt aaa gaa cat gaa tgt	986
Glu Asn Asp Thr Arg Ile Ile Glu Val Leu Cys Lys Glu His Glu Cys	
270 275 280	

aat ata gat gag gtg aaa aat gtt tat ttc aaa aat ttt ata cct ttt	1034
Asn Ile Asp Glu Val Lys Asn Val Tyr Phe Lys Asn Phe Ile Pro Phe	
285 290 295	
atg aat tct ctt gga ctt gta aca tct aat gga ctt cca gag gtt gaa	1082
Met Asn Ser Leu Gly Leu Val Thr Ser Asn Gly Leu Pro Glu Val Glu	
300 305 310 315	
aat ctt tct aaa cga tac gaa gaa att tat ctt aaa aat aaa gat cta	1130
Asn Leu Ser Lys Arg Tyr Glu Glu Ile Tyr Leu Lys Asn Lys Asp Leu	
320 325 330	
gat gca aga tta ttt ttg gat cat gat aaa act ctt cag act gat tct	1178
Asp Ala Arg Leu Phe Leu Asp His Asp Lys Thr Leu Gln Thr Asp Ser	
335 340 345	
ata gac agt ttt gaa aca cag aga aca cca cga aaa agt aac ctt gat	1226
Ile Asp Ser Phe Glu Thr Gln Arg Thr Pro Arg Lys Ser Asn Leu Asp	
350 355 360	
gaa gag gtg aat gta att cct cca cac act cca gtt agg act gtt atg	1274
Glu Glu Val Asn Val Ile Pro Pro His Thr Pro Val Arg Thr Val Met	
365 370 375	
aac act atc caa caa tta atg atg att tta aat tca gca agt gat caa	1322
Asn Thr Ile Gln Gln Leu Met Met Ile Leu Asn Ser Ala Ser Asp Gln	
380 385 390 395	
cct tca gaa aat ctg att tcc tat ttt aac aac tgc aca gtg aat cca	1370
Pro Ser Glu Asn Leu Ile Ser Tyr Phe Asn Asn Cys Thr Val Asn Pro	
400 405 410	
aaa gaa agt ata ctg aaa aga gtg aag gat ata gga tac atc ttt aaa	1418
Lys Glu Ser Ile Leu Lys Arg Val Lys Asp Ile Gly Tyr Ile Phe Lys	
415 420 425	
gag aaa ttt gct aaa gct gtg gga cag ggt tgt gtc gaa att gga tca	1466
Glu Lys Phe Ala Lys Ala Val Gly Gln Gly Cys Val Glu Ile Gly Ser	
430 435 440	
cag cga tac aaa ctt gga gtt cgc ttg tat tac cga gta atg gaa tcc	1514
Gln Arg Tyr Lys Leu Gly Val Arg Leu Tyr Tyr Arg Val Met Glu Ser	
445 450 455	
atg ctt aaa tca gaa gaa gaa cga tta tcc att caa aat ttt agc aaa	1562
Met Leu Lys Ser Glu Glu Glu Arg Leu Ser Ile Gln Asn Phe Ser Lys	
460 465 470 475	
ctt ctg aat gac aac att ttt cat atg tct tta ttg gcg tgc gct ctt	1610
Leu Leu Asn Asp Asn Ile Phe His Met Ser Leu Leu Ala Cys Ala Leu	
480 485 490	

gag gtt gta atg gcc aca tat agc aga agt aca tct cag aat ctt gat	1658
Glu Val Val Met Ala Thr Tyr Ser Arg Ser Thr Ser Gln Asn Leu Asp	
495 500 505	
tct gga aca gat ttg tct ttc cca tgg att ctg aat gtg ctt aat tta	1706
Ser Gly Thr Asp Leu Ser Phe Pro Trp Ile Leu Asn Val Leu Asn Leu	
510 515 520	
aaa gcc ttt gat ttt tac aaa gtg atc gaa agt ttt atc aaa gca gaa	1754
Lys Ala Phe Asp Phe Tyr Lys Val Ile Glu Ser Phe Ile Lys Ala Glu	
525 530 535	
ggc aac ttg aca aga gaa atg ata aaa cat tta gaa cga tgt gaa cat	1802
Gly Asn Leu Thr Arg Glu Met Ile Lys His Leu Glu Arg Cys Glu His	
540 545 550 555	
cga atc atg gaa tcc ctt gca tgg ctc tca gat tca cct tta ttt gat	1850
Arg Ile Met Glu Ser Leu Ala Trp Leu Ser Asp Ser Pro Leu Phe Asp	
560 565 570	
ctt att aaa caa tca aag gac cga gaa gga cca act gat cac ctt gaa	1898
Leu Ile Lys Gln Ser Lys Asp Arg Glu Gly Pro Thr Asp His Leu Glu	
575 580 585	
tct gct tgt cct ctt aat ctt cct ctc cag aat aat cac act gca gca	1946
Ser Ala Cys Pro Leu Asn Leu Pro Leu Gln Asn Asn His Thr Ala Ala	
590 595 600	
gat atg tat ctt tct cct gta aga tct cca aag aaa aaa ggt tca act	1994
Asp Met Tyr Leu Ser Pro Val Arg Ser Pro Lys Lys Lys Gly Ser Thr	
605 610 615	
acg cgt gta aat tct act gca aat gca gag aca caa gca acc tca gcc	2042
Thr Arg Val Asn Ser Thr Ala Asn Ala Glu Thr Gln Ala Thr Ser Ala	
620 625 630 635	
ttc cag acc cag aag cca ttg aaa tct acc tct ctt tca ctg ttt tat	2090
Phe Gln Thr Gln Lys Pro Leu Lys Ser Thr Ser Leu Ser Leu Phe Tyr	
640 645 650	
aaa aaa gtg tat cgg cta gcc tat ctc cgg cta aat aca ctt tgt gaa	2138
Lys Lys Val Tyr Arg Leu Ala Tyr Leu Arg Leu Asn Thr Leu Cys Glu	
655 660 665	
cgc ctt ctg tct gag cac cca gaa tta gaa cat atc atc tgg acc ctt	2186
Arg Leu Leu Ser Glu His Pro Glu Leu Glu His Ile Ile Trp Thr Leu	
670 675 680	
ttc cag cac acc ctg cag aat gag tat gaa ctc atg aga gac agg cat	2234
Phe Gln His Thr Leu Gln Asn Glu Tyr Glu Leu Met Arg Asp Arg His	
685 690 695	

ttg gac caa att atg atg tgt tcc atg tat ggc ata tgc aaa gtg aag	2282
Leu Asp Gln Ile Met Met Cys Ser Met Tyr Gly Ile Cys Lys Val Lys	
700 705 710 715	
aat ata gac ctt aaa ttc aaa atc att gta aca gca tac aag gat ctt	2330
Asn Ile Asp Leu Lys Phe Lys Ile Ile Val Thr Ala Tyr Lys Asp Leu	
720 725 730	
cct cat gct gtt cag gag aca ttc aaa cgt gtt ttg atc aaa gaa gag	2378
Pro His Ala Val Gln Glu Thr Phe Lys Arg Val Leu Ile Lys Glu Glu	
735 740 745	
gag tat gat tct att ata gta ttc tat aac tcg gtc ttc atg cag aga	2426
Glu Tyr Asp Ser Ile Ile Val Phe Tyr Asn Ser Val Phe Met Gln Arg	
750 755 760	
ctg aaa aca aat att ttg cag tat gct tcc acc agg ccc cct acc ttg	2474
Leu Lys Thr Asn Ile Leu Gln Tyr Ala Ser Thr Arg Pro Pro Thr Leu	
765 770 775	
tca cca ata cct cac att cct cga agc cct tac aag ttt cct agt tca	2522
Ser Pro Ile Pro His Ile Pro Arg Ser Pro Tyr Lys Phe Pro Ser Ser	
780 785 790 795	
ccc tta cgg att cct gga ggg aac atc tat att tca ccc ctg aag agt	2570
Pro Leu Arg Ile Pro Gly Gly Asn Ile Tyr Ile Ser Pro Leu Lys Ser	
800 805 810	
cca tat aaa att tca gaa ggt ctg cca aca cca aca aaa atg act cca	2618
Pro Tyr Lys Ile Ser Glu Gly Leu Pro Thr Pro Thr Lys Met Thr Pro	
815 820 825	
aga tca aga atc tta gta tca att ggt gaa tca ttc ggg act tct gag	2666
Arg Ser Arg Ile Leu Val Ser Ile Gly Glu Ser Phe Gly Thr Ser Glu	
830 835 840	
aag ttc cag aaa ata aat cag atg gta tgt aac agc gac cgt gtg ctc	2714
Lys Phe Gln Lys Ile Asn Gln Met Val Cys Asn Ser Asp Arg Val Leu	
845 850 855	
aaa aga agt gct gaa gga agc aac cct cct aaa cca ctg aaa aaa cta	2762
Lys Arg Ser Ala Glu Gly Ser Asn Pro Pro Lys Pro Leu Lys Lys Leu	
860 865 870 875	
cgc ttt gat att gaa gga tca gat gaa gca gat gga agt aaa cat ctc	2810
Arg Phe Asp Ile Glu Gly Ser Asp Glu Ala Asp Gly Ser Lys His Leu	
880 885 890	
cca gga gag tcc aaa ttt cag cag aaa ctg gca gaa atg act tct act	2858
Pro Gly Glu Ser Lys Phe Gln Gln Lys Leu Ala Glu Met Thr Ser Thr	
895 900 905	

cga aca cga atg caa aag cag aaa atg aat gat agc atg gat acc tca 2906
 Arg Thr Arg Met Gln Lys Gln Lys Met Asn Asp Ser Met Asp Thr Ser
 910 915 920

aac aag gaa gag aaa tga ggatctcagg accttgggtg acactgtgta 2954
 Asn Lys Glu Glu Lys
 925

cacctctgga ttcattgtct ctcacagatg tgactgtata actttcccag gttctgttta 3014
 tggccacatt taatatcttc agctcttttt gtggatataa aatgtgcaga tgcaattgtt 3074
 tgggtgattc ctaagccact tgaaatgtta gtcattgtta tttatacaag attgaaaatc 3134
 ttgtgtaaat cctgccattt aaaaagtgtg agcagattgt ttcctcttcc aaagtaaaat 3194
 tgctgtgctt tatggatagt aagaatggcc ctagagtggg agtcctgata acccaggcct 3254
 gtctgactac tttgccttct tttgtagcat ataggtgatg tttgtctctg tttttattaa 3314
 tttatatgta tattttttta atttaacatg aacaccctta gaaaatgtgt cctatctatc 3374
 ttccaaatgc aatttgattg actgcccatt caccaaaatt atcctgaact cttctgcaaa 3434
 aatggatatt attagaaatt agaaaaaat tactaatttt acacattaga ttttatttta 3494
 ctattggaat ctgatatact gtgtgcttgt tttataaaat tttgctttta attaaataaa 3554
 agctggaagc aaagtataac catatgatac tatkatacta ctgaaacaga tttcatacct 3614
 cagaatgtaa aagaacttac tgattatfff cttcatccaa cttatgtfff taaatgagga 3674
 ttattgatag tactcttggg ttttatacca ttcagatcac tgaatttata aagtacccat 3734
 ctagtacttg aaaaagtaaa gtgttctgcc agatcttagg tatagaggac cctaacacag 3794
 tatatcccaa gtgcactttc taatgtttct gggtcctgaa gaattaagat acaaattaat 3854
 tttactccat aaacagactg ttaattatag gagccttaat ttttttttca tagagatttg 3914
 tctaattgca tctcaaaatt attctgccct ccttaatttg ggaaggtttg tgttttctct 3974
 ggaatggtac atgtcttcca tgtatctttt gaactggcaa ttgtctatft atcttttatt 4034
 tttttaagtc agtatggtct aacactggca tgttcaaagc cacattatft ctagtccaaa 4094
 attacaagta atcaagggtc attatgggtt aggcattaat gtttctatct gattttgtgc 4154
 aaaagcttca aattaaaaca gctgcattag aaaaagaggc gcttctcccc tcccctacac 4214
 ctaaagggtg atttaaaacta tcttgtgtga ttaacttatt tagagatgct gtaacttaaa 4274

ataggggata tttaaggtag cttcagctag cttttaggaa aatcactttg tctaactcag 4334
 aattatTTTT aaaaagaaat ctggctctgt tagaaaacaa aattttattt tgtgctcatt 4394
 taagtttcaa acttactatt ttgacagtta ttttgataac aatgacacta gaaaacttga 4454
 ctccatttca tcattgtttc tgcataaata tcatacaaat cagttagttt ttaggtcaag 4514
 ggcttactat ttctgggtct tttgctacta agttcacatt agaattagtg ccagaatttt 4574
 aggaacttca gagatcgtgt attgagattt cttaaataat gcttcagata ttattgcttt 4634
 attgcttttt tgtattggtt aaaactgtac atttaaaatt gctatgttac tattttctac 4694
 aattaatagt ttgtctattt taaaataaat tagttgttaa gagtcctaat ggtctgatgt 4754
 tgtgttcttt gtattaagta cactaatgtt ctcttttctg tctaggagaa gatagataga 4814
 agataactct cctagtatct catcc 4839

<210> 8

<211> 928

<212> PRT

<213> Homo sapien

<400> 8

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

His Lys Phe Phe Asn Leu Leu Lys Glu Ile Asp Thr Ser Thr Lys Val
 130 135 140
 Asp Asn Ala Met Ser Arg Leu Leu Lys Lys Tyr Asp Val Leu Phe Ala
 145 150 155 160
 Leu Phe Ser Lys Leu Glu Arg Thr Cys Glu Leu Ile Tyr Leu Thr Gln
 165 170 175
 Pro Ser Ser Ser Ile Ser Thr Glu Ile Asn Ser Ala Leu Val Leu Lys
 180 185 190
 Val Ser Trp Ile Thr Phe Leu Leu Ala Lys Gly Glu Val Leu Gln Met
 195 200 205
 Glu Asp Asp Leu Val Ile Ser Phe Gln Leu Met Leu Cys Val Leu Asp
 210 215 220
 Tyr Phe Ile Lys Leu Ser Pro Pro Met Leu Leu Lys Glu Pro Tyr Lys
 225 230 235 240
 Thr Ala Val Ile Pro Ile Asn Gly Ser Pro Arg Thr Pro Arg Arg Gly
 245 250 255
 Gln Asn Arg Ser Ala Arg Ile Ala Lys Gln Leu Glu Asn Asp Thr Arg
 260 265 270
 Ile Ile Glu Val Leu Cys Lys Glu His Glu Cys Asn Ile Asp Glu Val
 275 280 285
 Lys Asn Val Tyr Phe Lys Asn Phe Ile Pro Phe Met Asn Ser Leu Gly
 290 295 300
 Leu Val Thr Ser Asn Gly Leu Pro Glu Val Glu Asn Leu Ser Lys Arg
 305 310 315 320
 Tyr Glu Glu Ile Tyr Leu Lys Asn Lys Asp Leu Asp Ala Arg Leu Phe
 325 330 335
 Leu Asp His Asp Lys Thr Leu Gln Thr Asp Ser Ile Asp Ser Phe Glu
 340 345 350
 Thr Gln Arg Thr Pro Arg Lys Ser Asn Leu Asp Glu Glu Val Asn Val
 355 360 365
 Ile Pro Pro His Thr Pro Val Arg Thr Val Met Asn Thr Ile Gln Gln
 370 375 380
 Leu Met Met Ile Leu Asn Ser Ala Ser Asp Gln Pro Ser Glu Asn Leu
 385 390 395 400

Ile Ser Tyr Phe Asn Asn Cys Thr Val Asn Pro Lys Glu Ser Ile Leu
 405 410 415
 Lys Arg Val Lys Asp Ile Gly Tyr Ile Phe Lys Glu Lys Phe Ala Lys
 420 425 430
 Ala Val Gly Gln Gly Cys Val Glu Ile Gly Ser Gln Arg Tyr Lys Leu
 435 440 445
 Gly Val Arg Leu Tyr Tyr Arg Val Met Glu Ser Met Leu Lys Ser Glu
 450 455 460
 Glu Glu Arg Leu Ser Ile Gln Asn Phe Ser Lys Leu Leu Asn Asp Asn
 465 470 475 480
 Ile Phe His Met Ser Leu Leu Ala Cys Ala Leu Glu Val Val Met Ala
 485 490 495
 Thr Tyr Ser Arg Ser Thr Ser Gln Asn Leu Asp Ser Gly Thr Asp Leu
 500 505 510
 Ser Phe Pro Trp Ile Leu Asn Val Leu Asn Leu Lys Ala Phe Asp Phe
 515 520 525
 Tyr Lys Val Ile Glu Ser Phe Ile Lys Ala Glu Gly Asn Leu Thr Arg
 530 535 540
 Glu Met Ile Lys His Leu Glu Arg Cys Glu His Arg Ile Met Glu Ser
 545 550 555 560
 Leu Ala Trp Leu Ser Asp Ser Pro Leu Phe Asp Leu Ile Lys Gln Ser
 565 570 575
 Lys Asp Arg Glu Gly Pro Thr Asp His Leu Glu Ser Ala Cys Pro Leu
 580 585 590
 Asn Leu Pro Leu Gln Asn Asn His Thr Ala Ala Asp Met Tyr Leu Ser
 595 600 605
 Pro Val Arg Ser Pro Lys Lys Lys Gly Ser Thr Thr Arg Val Asn Ser
 610 615 620
 Thr Ala Asn Ala Glu Thr Gln Ala Thr Ser Ala Phe Gln Thr Gln Lys
 625 630 635 640
 Pro Leu Lys Ser Thr Ser Leu Ser Leu Phe Tyr Lys Lys Val Tyr Arg
 645 650 655
 Leu Ala Tyr Leu Arg Leu Asn Thr Leu Cys Glu Arg Leu Leu Ser Glu
 660 665 670

His Pro Glu Leu Glu His Ile Ile Trp Thr Leu Phe Gln His Thr Leu
 675 680 685
 Gln Asn Glu Tyr Glu Leu Met Arg Asp Arg His Leu Asp Gln Ile Met
 690 695 700
 Met Cys Ser Met Tyr Gly Ile Cys Lys Val Lys Asn Ile Asp Leu Lys
 705 710 715 720
 Phe Lys Ile Ile Val Thr Ala Tyr Lys Asp Leu Pro His Ala Val Gln
 725 730 735
 Glu Thr Phe Lys Arg Val Leu Ile Lys Glu Glu Glu Tyr Asp Ser Ile
 740 745 750
 Ile Val Phe Tyr Asn Ser Val Phe Met Gln Arg Leu Lys Thr Asn Ile
 755 760 765
 Leu Gln Tyr Ala Ser Thr Arg Pro Pro Thr Leu Ser Pro Ile Pro His
 770 775 780
 Ile Pro Arg Ser Pro Tyr Lys Phe Pro Ser Ser Pro Leu Arg Ile Pro
 785 790 795 800
 Gly Gly Asn Ile Tyr Ile Ser Pro Leu Lys Ser Pro Tyr Lys Ile Ser
 805 810 815
 Glu Gly Leu Pro Thr Pro Thr Lys Met Thr Pro Arg Ser Arg Ile Leu
 820 825 830
 Val Ser Ile Gly Glu Ser Phe Gly Thr Ser Glu Lys Phe Gln Lys Ile
 835 840 845
 Asn Gln Met Val Cys Asn Ser Asp Arg Val Leu Lys Arg Ser Ala Glu
 850 855 860
 Gly Ser Asn Pro Pro Lys Pro Leu Lys Lys Leu Arg Phe Asp Ile Glu
 865 870 875 880
 Gly Ser Asp Glu Ala Asp Gly Ser Lys His Leu Pro Gly Glu Ser Lys
 885 890 895
 Phe Gln Gln Lys Leu Ala Glu Met Thr Ser Thr Arg Thr Arg Met Gln
 900 905 910
 Lys Gln Lys Met Asn Asp Ser Met Asp Thr Ser Asn Lys Glu Glu Lys
 915 920 925

<210> 9
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
Oligonucleotide

<400> 9
ttctgtagtt taattttctg aacctttggc

30

<210> 10
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
Oligonucleotide

<400> 10
tcagccgaag agcttcagga agcaggg

27

<210> 11
<211> 32
<212> PRT
<213> Homo sapien

<220>
<221> VARIANT
<222> (2)..(3)

<220>
<221> VARIANT
<222> (6)..(13)

<220>
<221> VARIANT
<222> (15)

<220>
<221> VARIANT
<222> (17)..(18)

<220>
<221> VARIANT
<222> (20)..(21)

<220>

<221> VARIANT

<222> (23) .. (28)

<220>

<221> VARIANT

<222> (30) .. (31)

<400> 11

Cys Xaa Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys Xaa His
1 5 10 15

Xaa Xaa Cys Xaa Xaa Cys Xaa Xaa Xaa Xaa Xaa Cys Xaa Xaa Cys
20 25 30

<210> 12

<211> 50

<212> PRT

<213> Homo sapien

<220>

<221> VARIANT

<222> (2) .. (3)

<220>

<221> VARIANT

<222> (5) .. (20)

<220>

<221> VARIANT

<222> (22) .. (23)

<220>

<221> VARIANT

<222> (25) .. (26)

<220>

<221> VARIANT

<222> (28) .. (29)

<220>

<221> VARIANT

<222> (31) .. (46)

<220>

<221> VARIANT

<222> (48) .. (49)

<400> 12

Cys Xaa Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
1 5 10 15

Xaa Xaa Xaa Xaa His Xaa Xaa Cys Xaa Xaa Cys Xaa Xaa Cys Xaa Xaa
20 25 30
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys Xaa
35 40 45
Xaa Cys
50