US 20090181003A1

(19) United States
(12) Patent Application Publication
Gout et al.

(10) Pub. No.: US 2009/0181003 A1
(43) Pub. Date: Jul. 16, 2009

(54) **REGULATION OF S6 KINSASE PROTEIN ACTIVITY AND RELATED METHODS**

(76) Inventors: **Ivan Gout**, London (GB); **Timothy Robert Fenton**, San Diego, CA (US); **Heba Ismail**, London (GB)

Correspondence Address:
MORGAN LEWIS & BOCKIUS LLP
1111 PENNSYLVANIA AVENUE NW
WASHINGTON, DC 20004 (US)

(21) Appl. No.: **11/990,226**

(22) PCT Filed: **Aug. 8, 2006**

(86) PCT No.: **PCT/US2006/030675**

§ 371 (c)(1),
(2), (4) Date: **Feb. 13, 2009**

Related U.S. Application Data

(60) Provisional application No. 60/706,010, filed on Aug. 8, 2005, provisional application No. 60/785,981, filed on Mar. 27, 2006.

Publication Classification

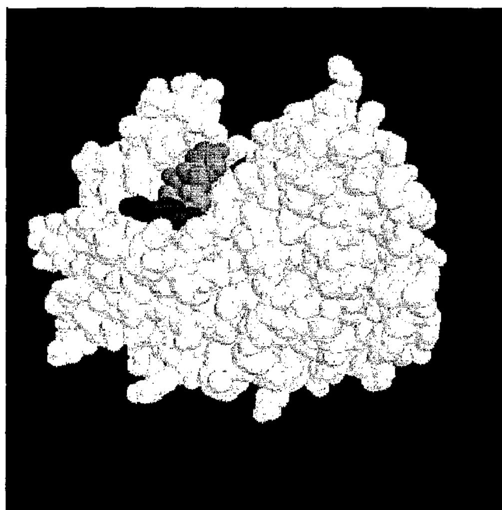
(51) **Int. Cl.**

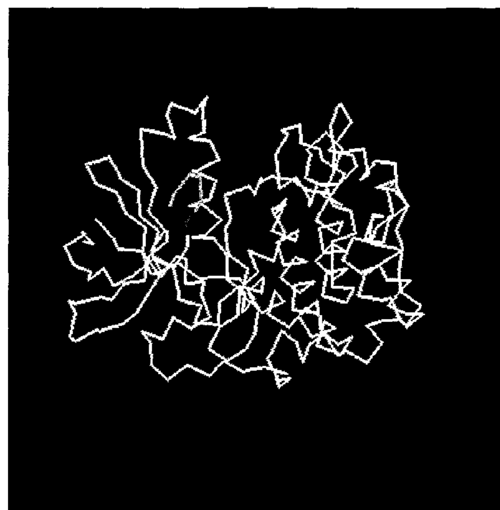
A61K 38/45	(2006.01)
C12N 9/10	(2006.01)
C12N 15/11	(2006.01)
G01N 33/53	(2006.01)
C12Q 1/48	(2006.01)
C12N 9/99	(2006.01)
C07K 16/18	(2006.01)
C12Q 1/68	(2006.01)

(52) **U.S. Cl.** **424/94.5**; 435/193; 536/23.2; 435/7.8; 435/15; 435/184; 530/387.9; 435/6

(57) **ABSTRACT**

The invention encompasses agents and their methods of use for modulating the activity of kinases by effecting their acetylation or their binding to nucleic acids. The invention thus encompasses the modulation of S6 kinase by effecting its acetylation by p300. The invention further encompasses the modulation of S6 kinase 2 by affecting its binding to DNA.





Accessibilty data: Lys 99 - 63.8
Lys 104: 75.3

Red & Blue: P-loop lysines
Green: Rest of P-Loop
Yellow: Thr 252

S6K1 Human	VLGKGGYGKVF
S6K2 Human	VLGKGGYGKVF
S6K D.melanogaster	VLGKGGYGKVF
S6K C.elegans	VLGKGGYGKVF
S6K A.thaliana	VVGKGAFGKVY
PKB Human	LLGKGTFGKVY
PKC Human	VLGKGSFGKVM
PKA Human	TLGTGSFGRVM

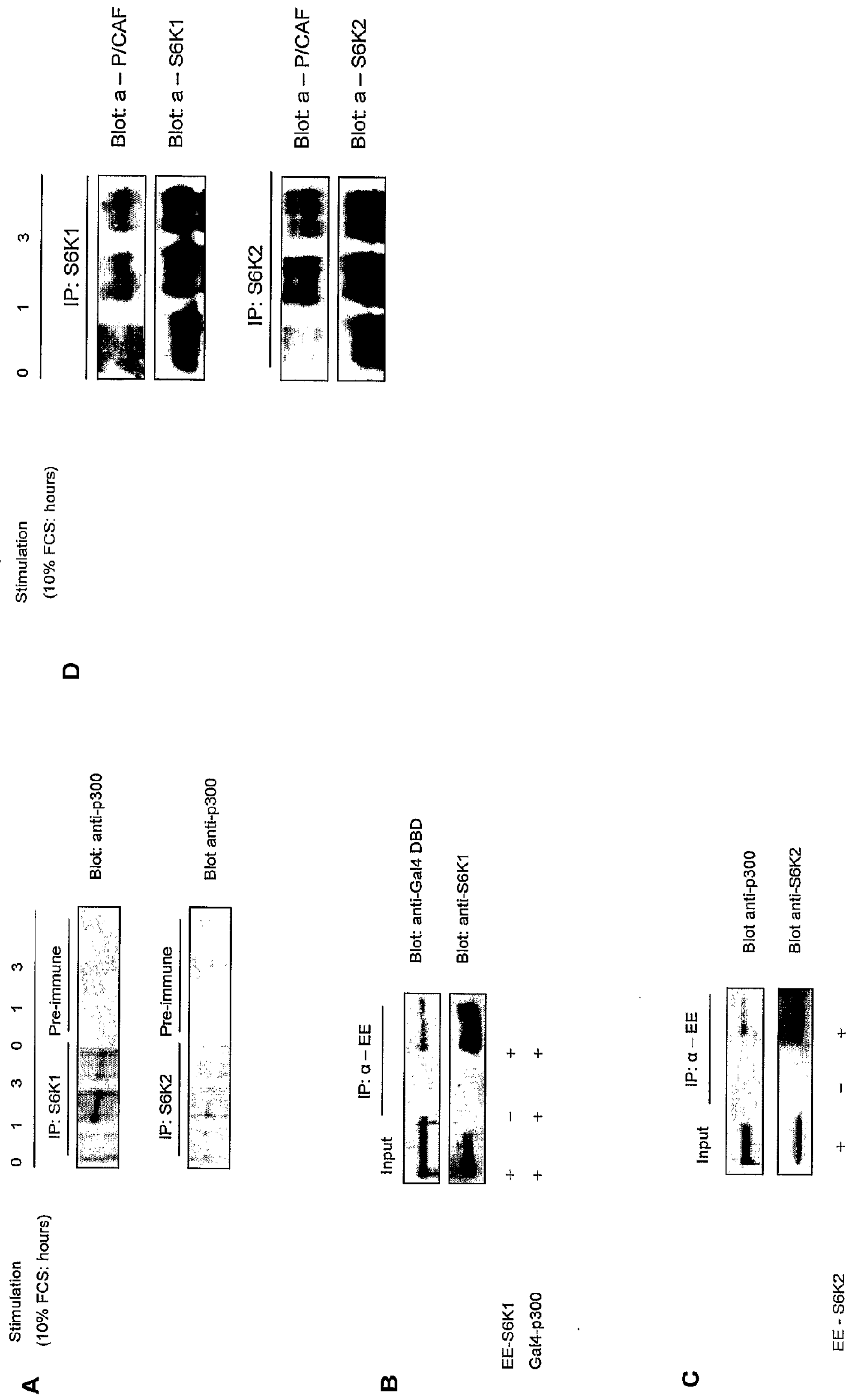


Figure 1

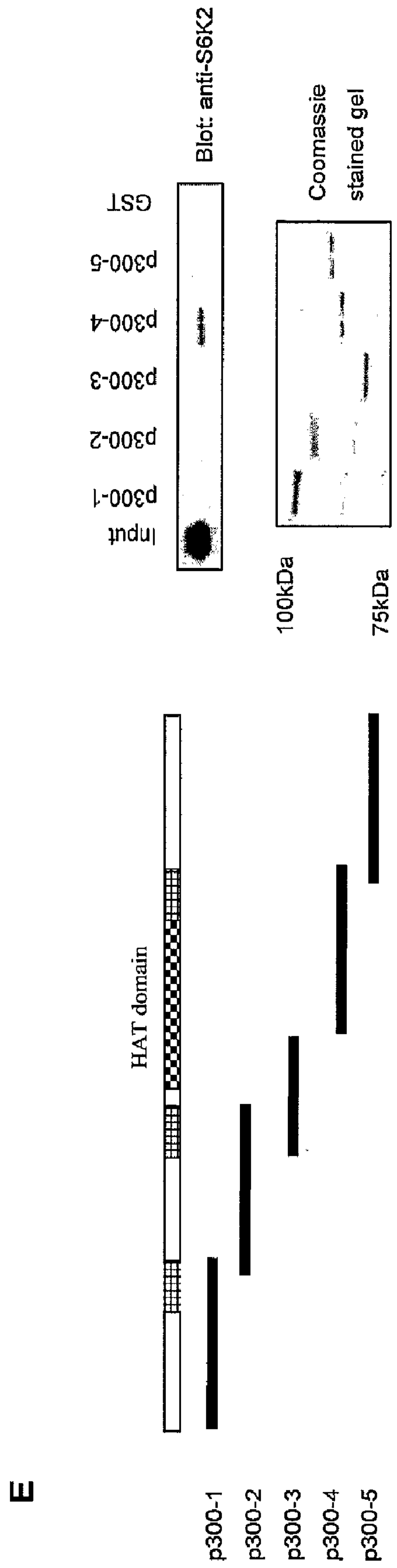


Figure 1

A

S6K1	+	+	+	−	−	−
S6K2	−	−	−	+	+	+
P300-HAT	−	+	−	−	+	−
GST-P/CAF	−	−	+	−	−	+

MW (kDa) 100 75 50

Ac-P/CAF
Ac-p300
Ac-S6K1
Ac-S6K2

^{14}C Autoradiograph

100 75 50

P/CAF
p300 HAT
S6K1
S6K2

Coomassie stain

B

p300	+	+
GST	+	−
S6K	−	+

MW (kDa) 75 50 37 25

Ac-S6K2

^{14}C Autoradiograph

75 50 37 25

S6K2
GST

Coomassie stain

Figure 2

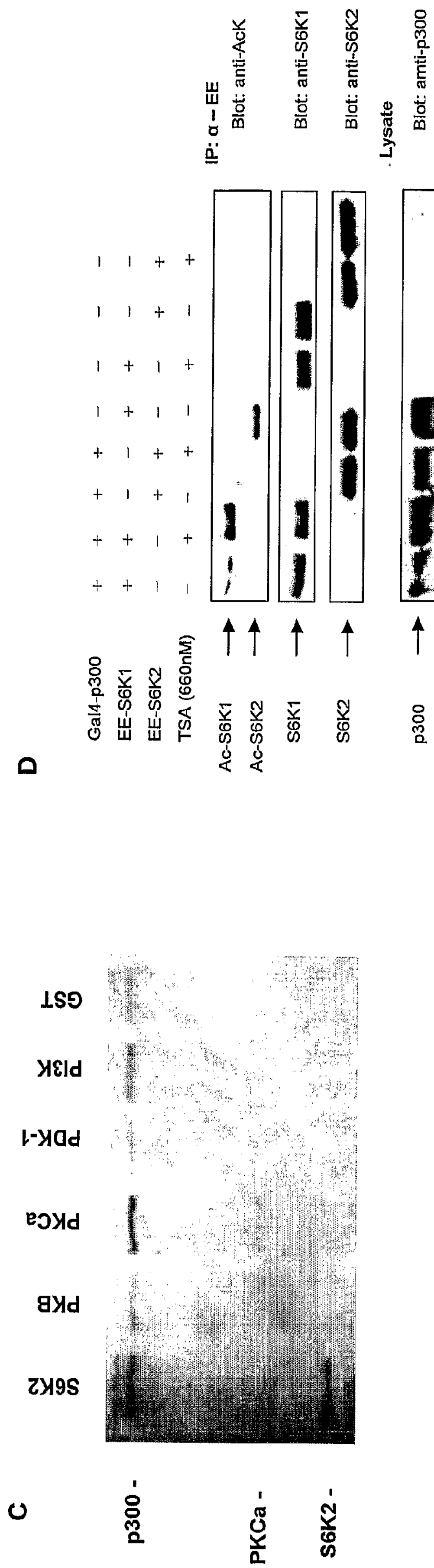


Figure 2

A

$QAFPM_{ox}ISK_{ac}RPEHLR$

Intensity

m /z

y_3 b_5 y_5 y_6 a_8-NH_3 y_7 b_9-64 b_9 y_9 b_{11}-64 b_{12}-64 b_{13}-64

M+H
M+H-64
M-OH-64
$b_{13}+H_2O$-64
$b_{13}+H_2O$

B

	P-loop
H. sapiens (S6K1)	FELLRVLGKGGYGKVFQVRK
H. sapiens (S6K2)	FELLRVLGKGGYGKVFQVRK
R. norvegicus	FELLRVLGKGGYGKVFQVRK
D. melanogastor	FELLRVLGKGGYGKVFQVRK
X. laevis	FELLRVLGKGGYGKVFQVRK
C. elegans	FQLLKVLGKGGYGKVFQVRK

C

	P-loop
S6K1	FELLRVLGKGGYGKVFQVRK
S6K2	FELLRVLGKGGYGKVFQVRK
PKBα	FEYLKLLGKGTFGKVILVKE
PKBβ	FDYLKLLGKGTFGKVILVRE
PKBγ	FDYLKLLGKGTFGKVILVRE
PKCα	FNFLMVLGKGSFGKVMLADR
PKCβ	FNFLMVLGKGSFGKVMLSER
PKCγ	FSFLMVLGKGSFGKVMLAER
PKCε	FNFIKVLGKGSFGKVMLAEL
PKCη	FEFIRVLGKGSFGKVMLARV
SGK	FHFLKVIGKGSFGKVLLARH
PDK-1	FKFGKILGEGSFSTVVLARE
RSK	FELLKVLGQGSFGKVFLVRK

Figure 3

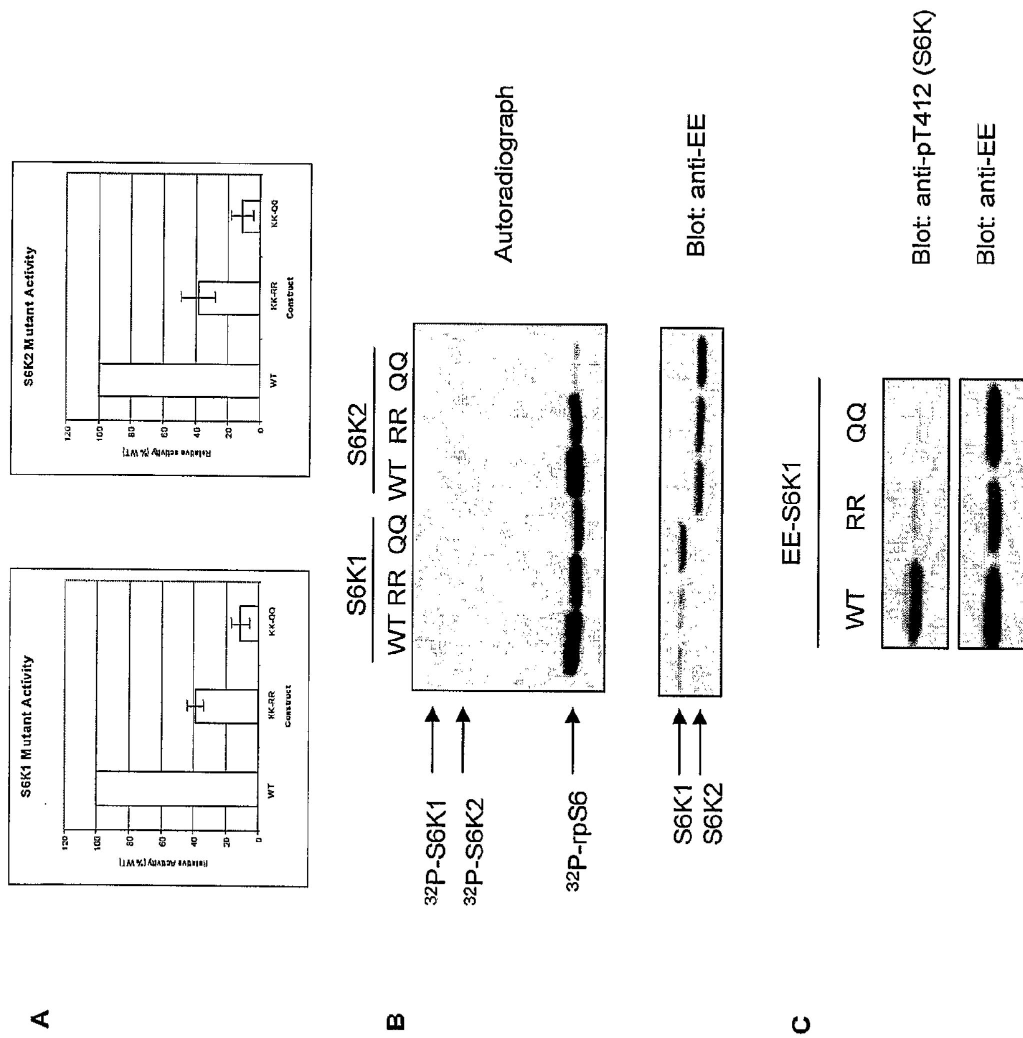
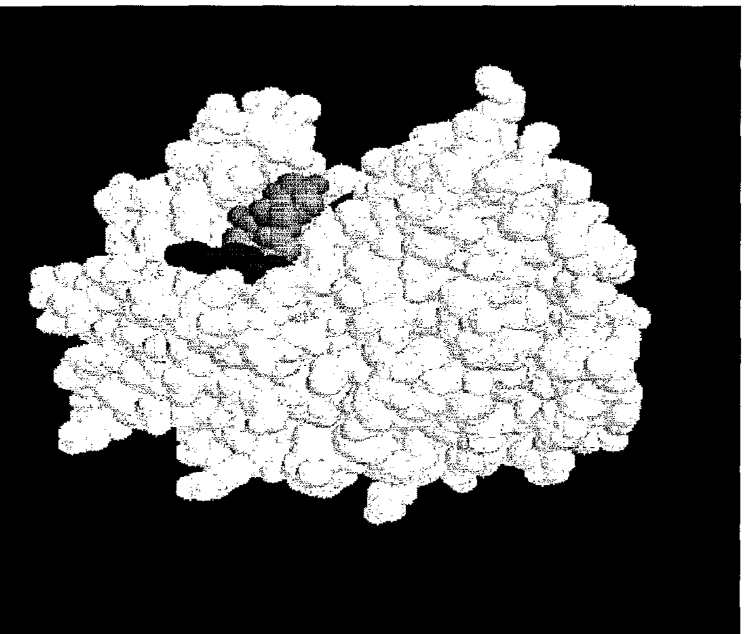
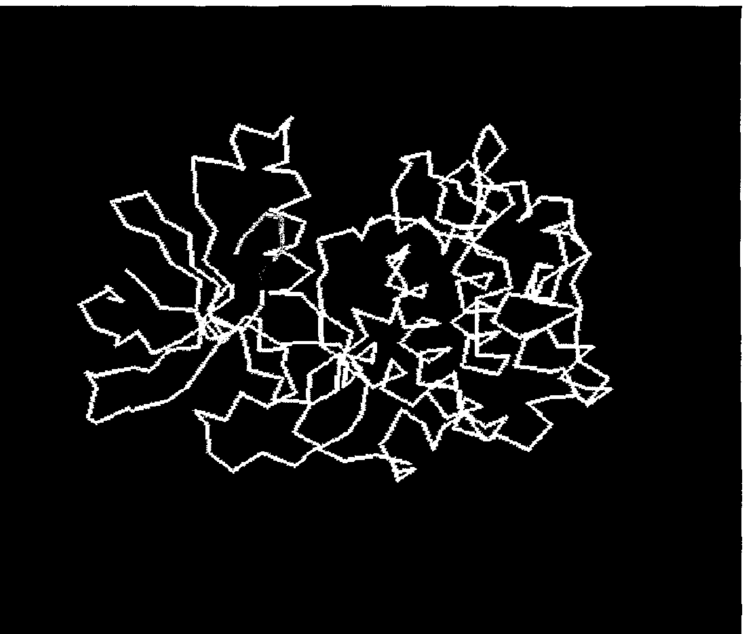


Figure 4





Accessibilty data: Lys 99 - 63.8
Lys 104: 75.3

Red & Blue: P-loop lysines
Green: Rest of P-Loop
Yellow: Thr 252

S6K1 Human	VLGKGGYGKVF
S6K2 Human	VLGKGGYGKVF
S6K D.melanogaster	VLGKGGYGKVF
S6K C.elegans	VLGKGGYGKVF
S6K A.thaliana	VVGKGAFGKVY
PKB Human	LLGKGTFGKVY
PKC Human	VLGKGSFGKVM
PKA Human	TLGTGSFGRVM

Figure 5

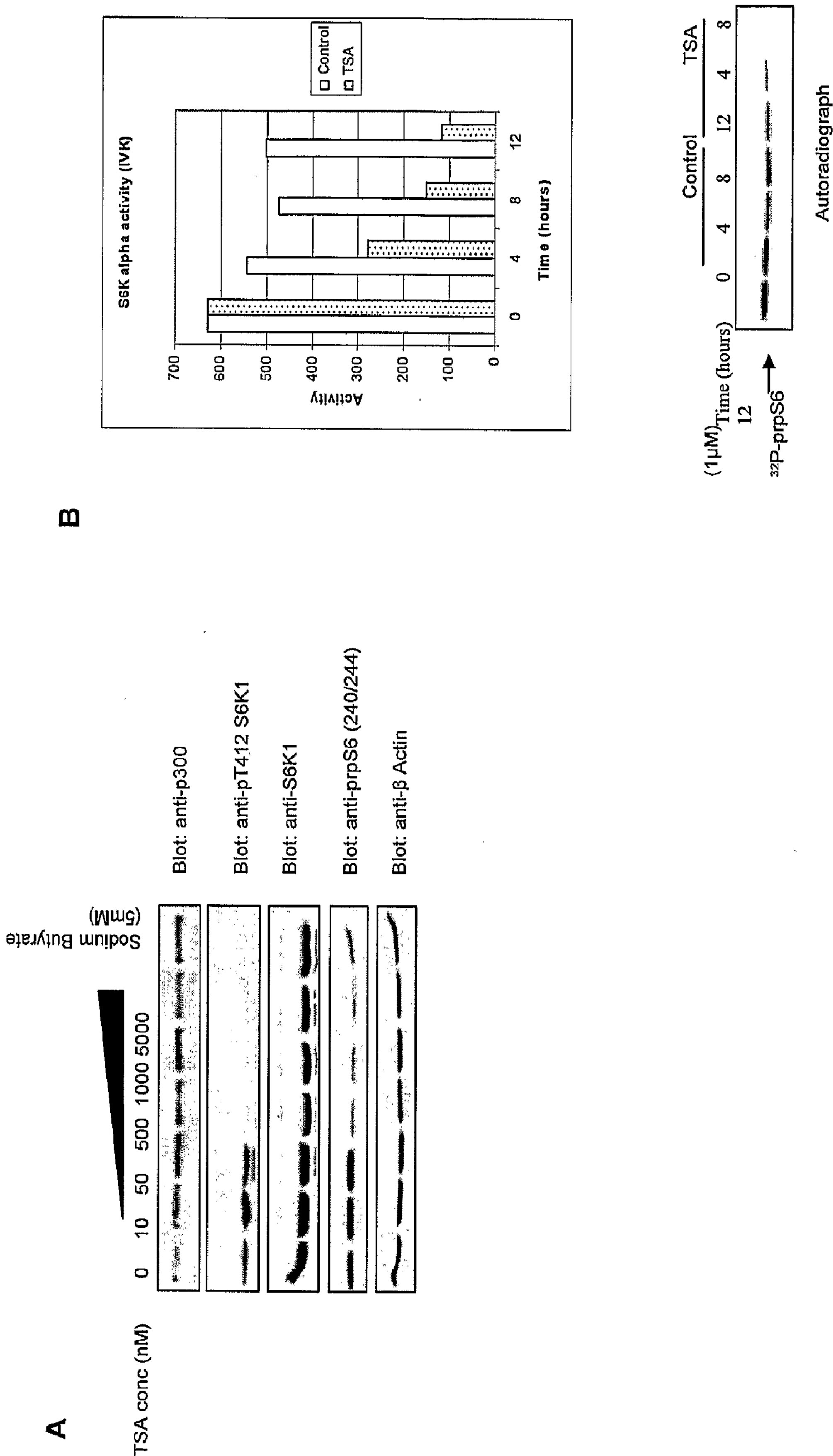


Figure 6

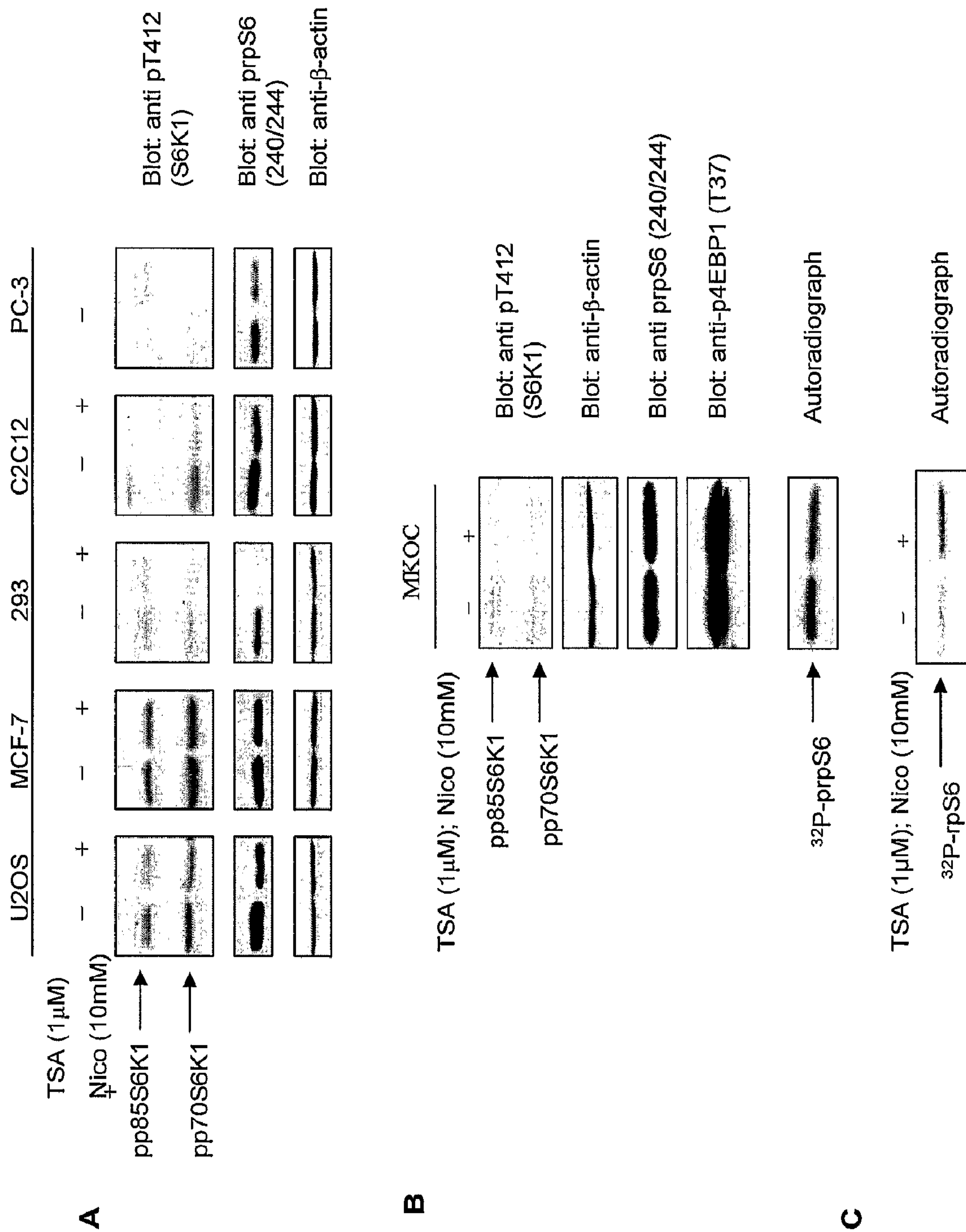


Figure 7

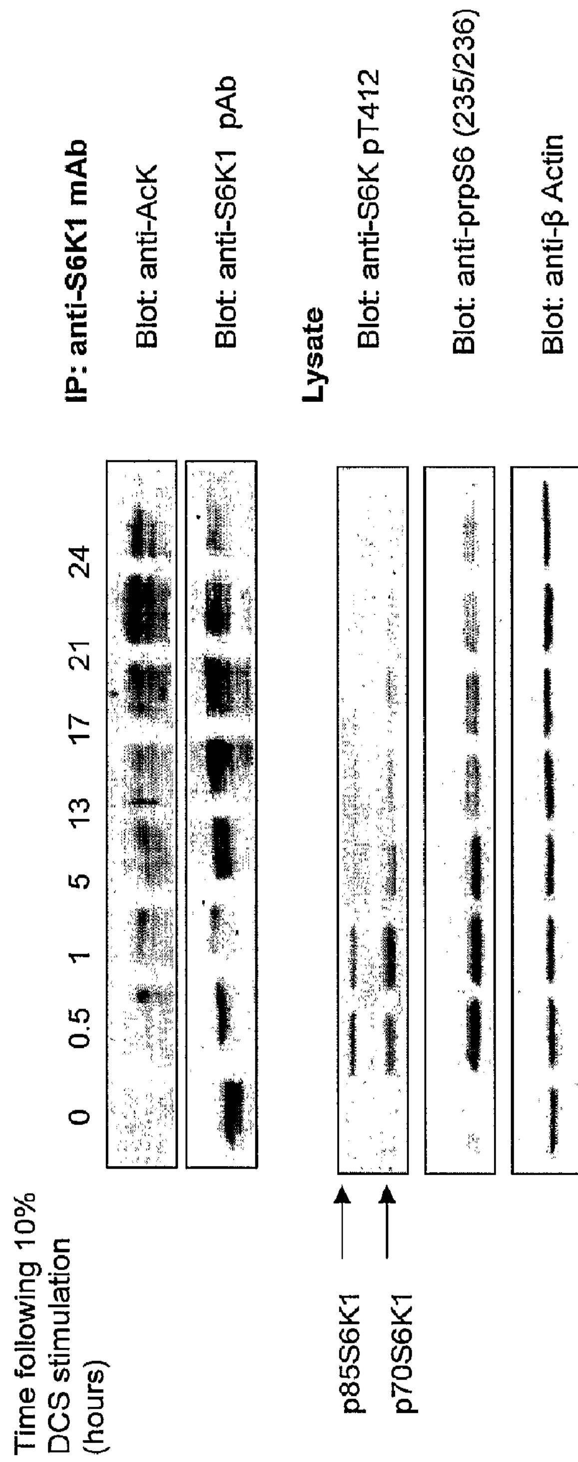


Figure 8

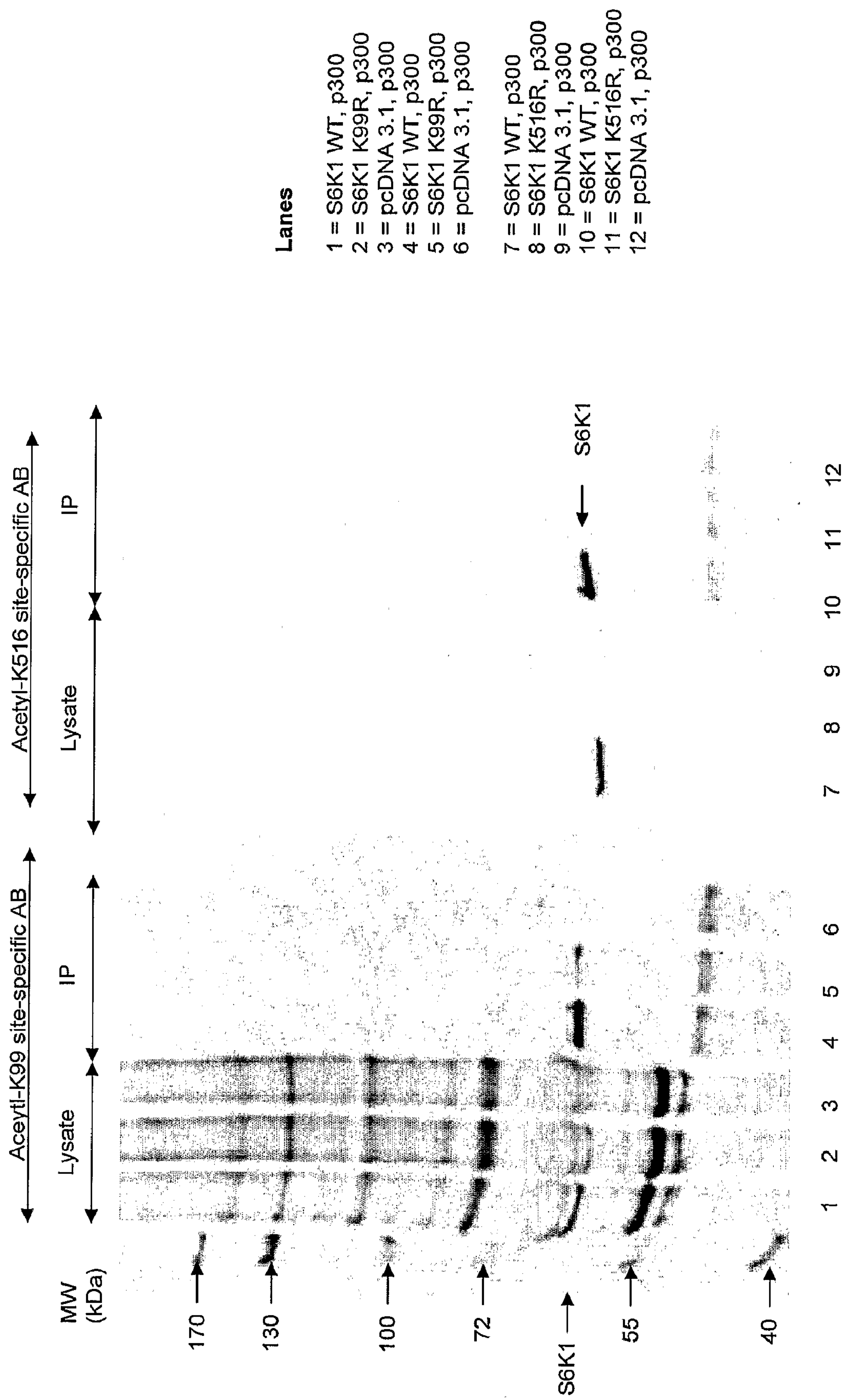


Figure 9

A

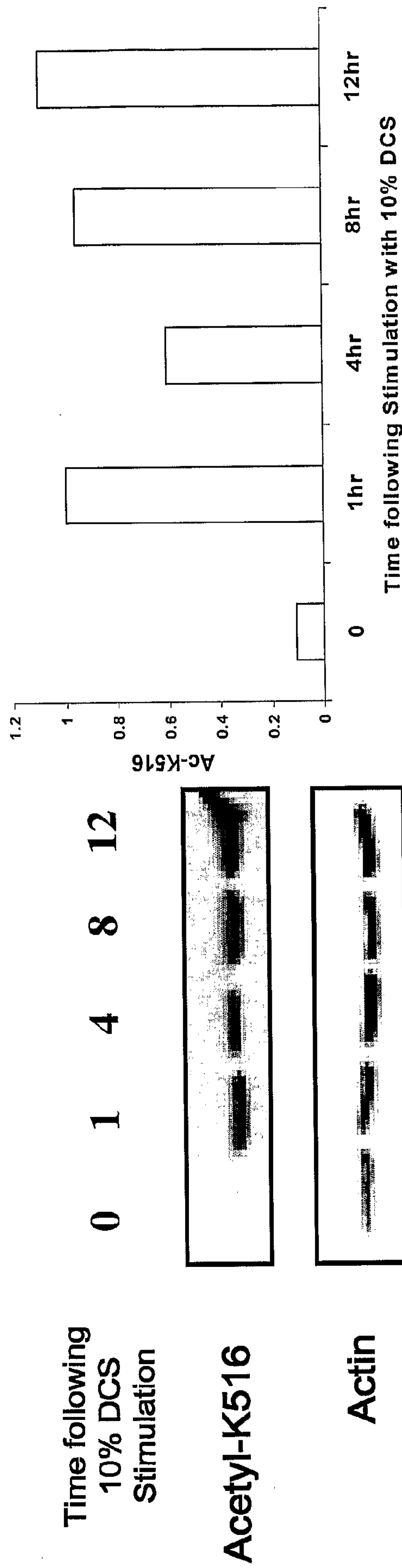


Figure 10A

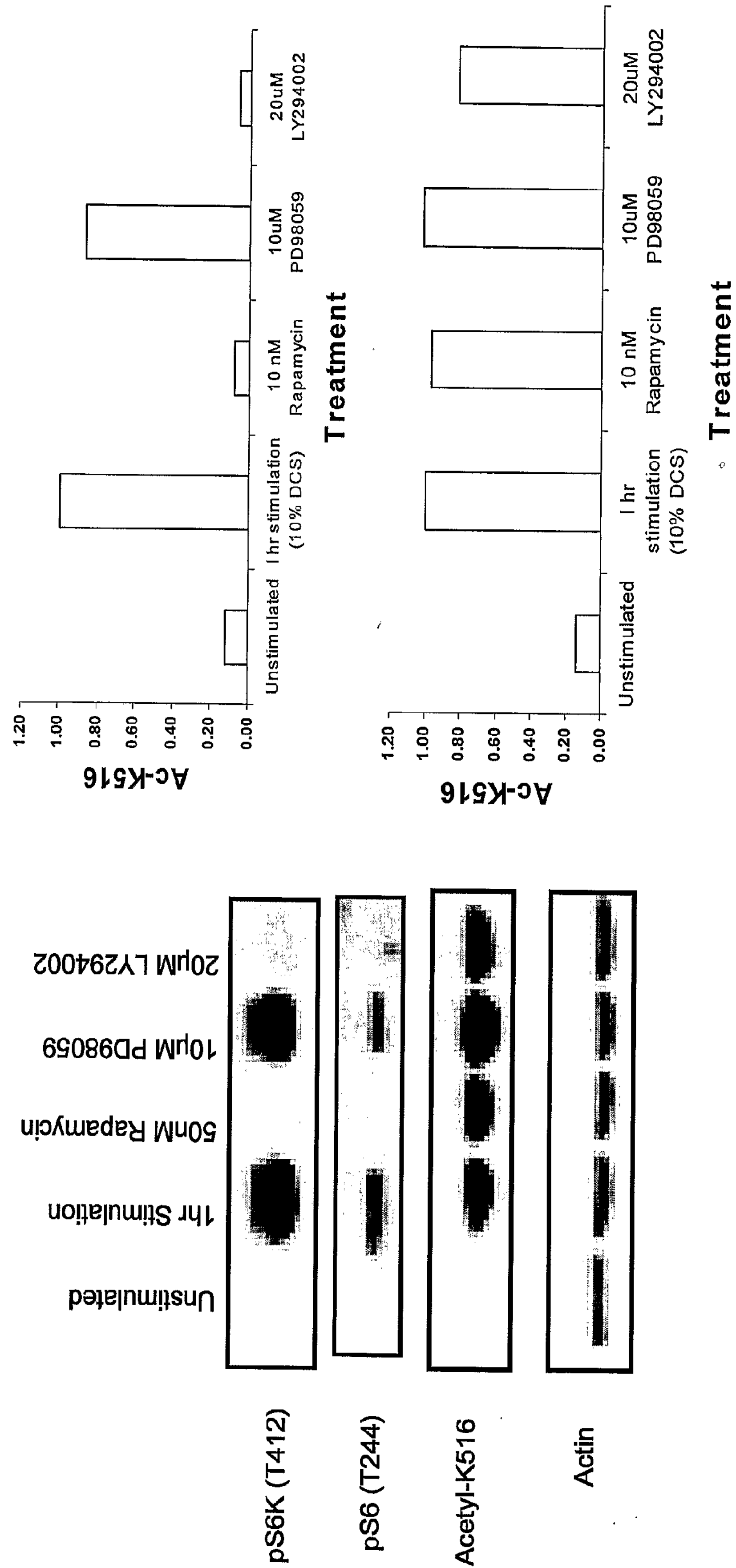


Figure 10B

C

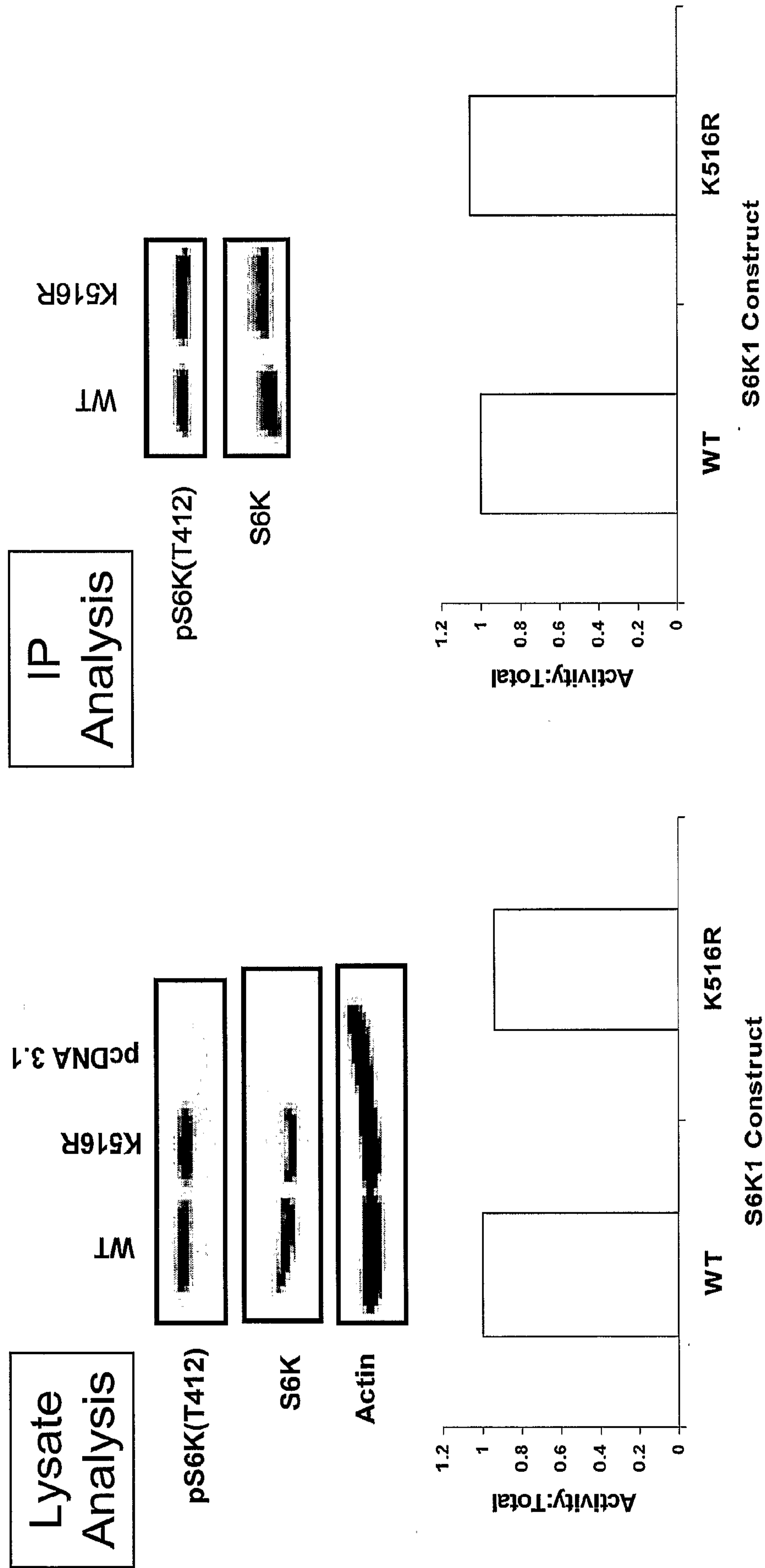


Figure 10C

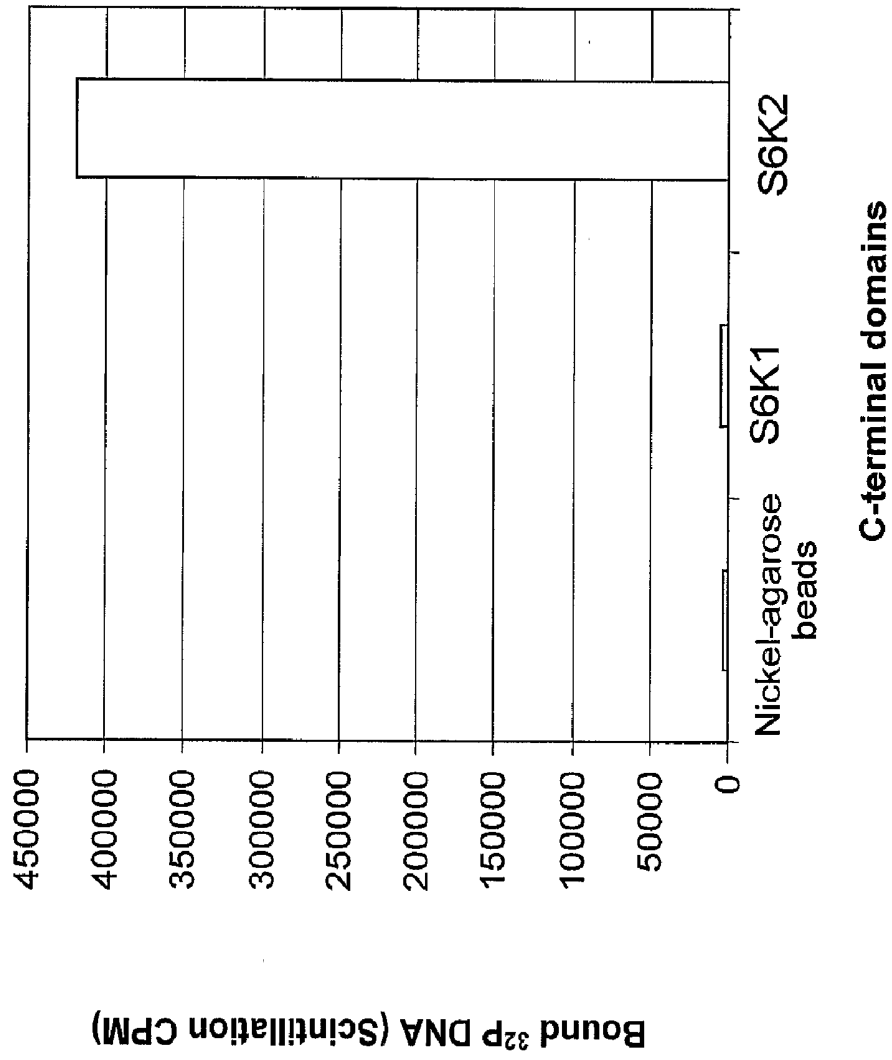
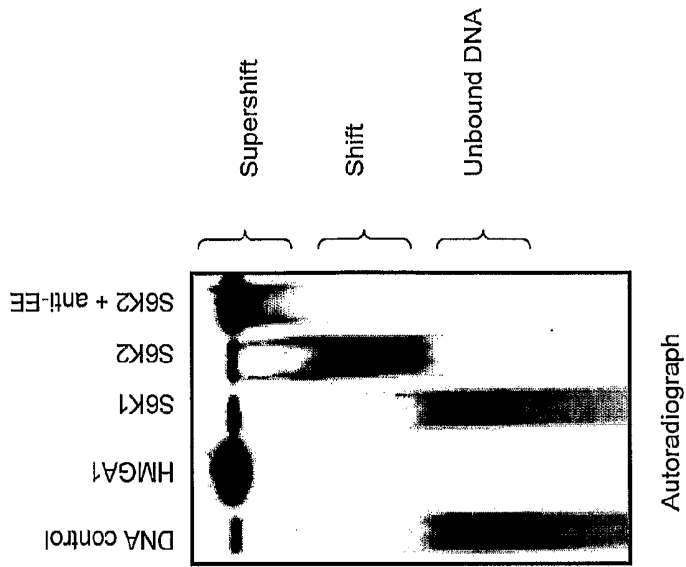


Figure 11

B



A

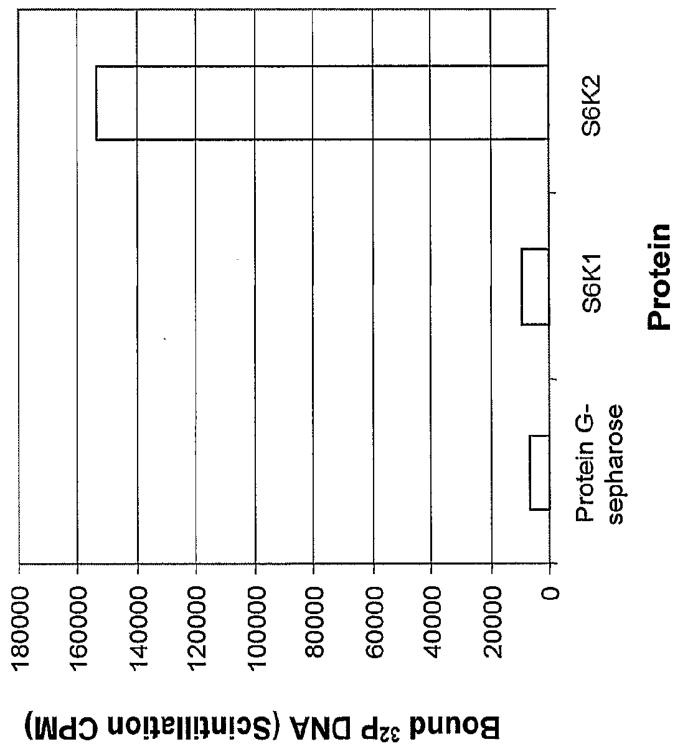


Figure 12

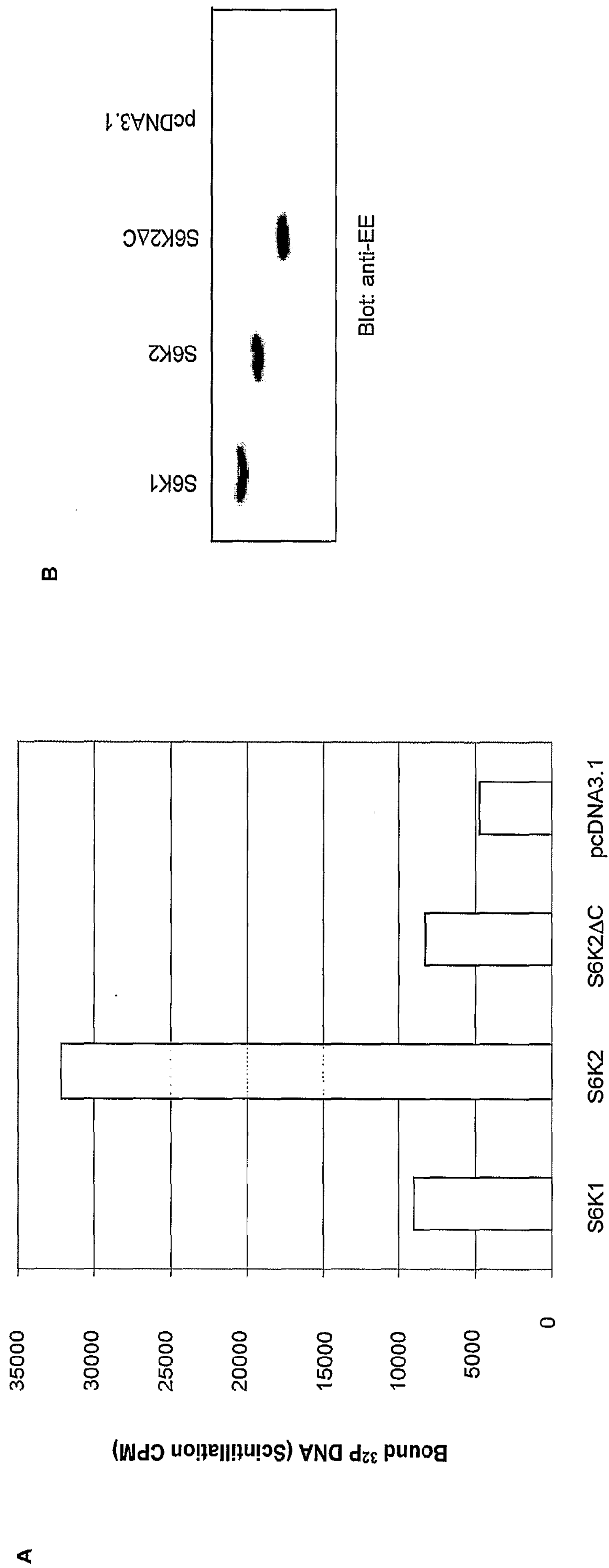


Figure 13

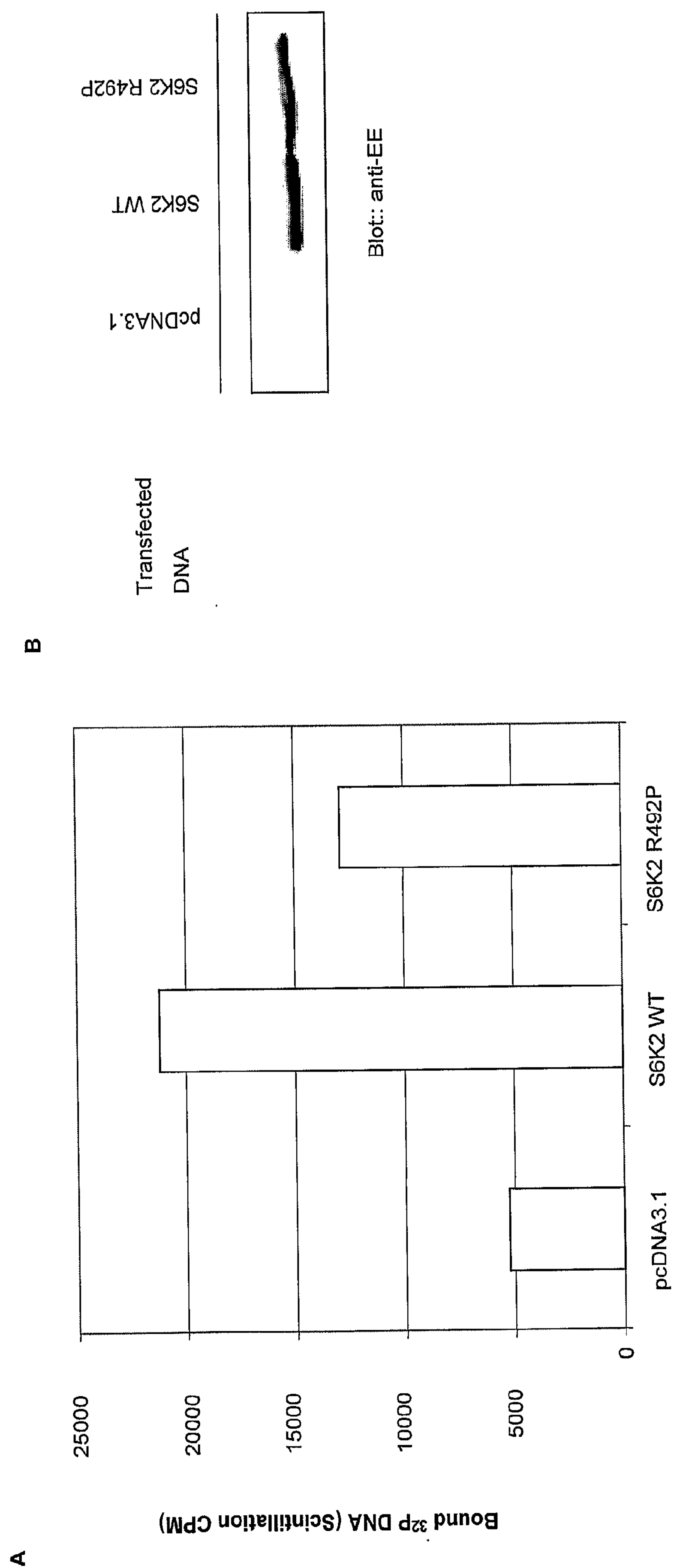


Figure 14

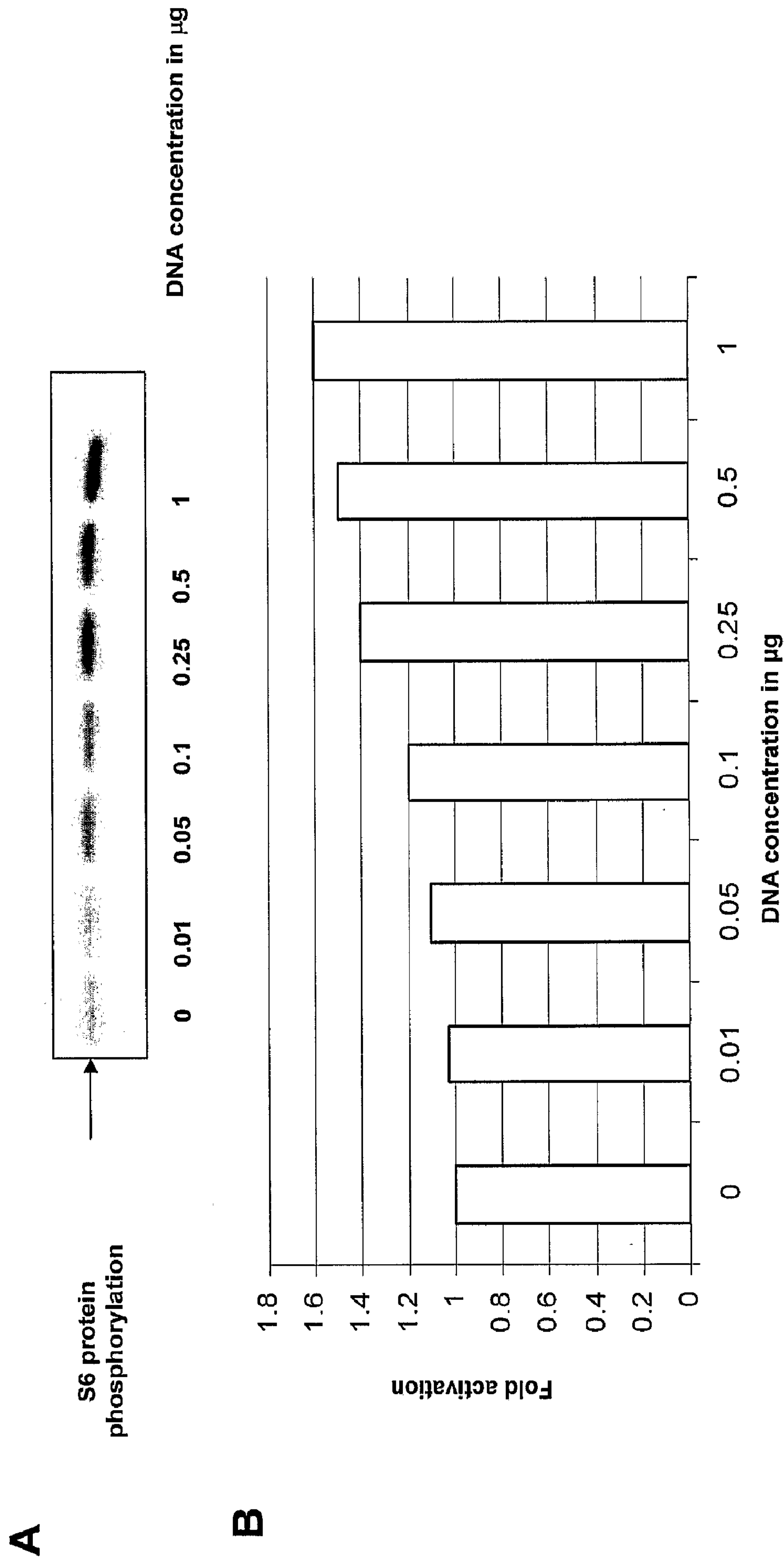


Figure 15

REGULATION OF S6 KINSASE PROTEIN ACTIVITY AND RELATED METHODS

RELATED APPLICATIONS

[0001] This application claims priority to U.S. Provisional Application 60/785,981 (filed on Mar. 27, 2006) and U.S. Provisional Application 60/706,010 (filed on Aug. 8, 2005) both of which are incorporated by reference in their entirety.

FIELD OF THE INVENTION

[0002] The present invention relates to methods of modulating protein kinase activity by an acetylating agent, to methods of treating and/or preventing protein kinase-related disorders and to molecules that catalyze the acetylation of protein kinases. The present invention further relates to methods of modulating protein kinase activity by binding to nucleic acids, to methods of treating and/or preventing protein kinase-related disorders and to molecules that can block DNA/protein kinase interactions.

BACKGROUND OF THE INVENTION

[0003] The 40S ribosomal protein S6 kinases (S6Ks) belong to AGC family of serine/threonine protein kinases, which includes key players in the regulation of cellular functions such as PKB/Akt, PKA, PDK1 and PKC. There are two isoforms of S6K, termed S6K1 (also known as p70α S6 kinase or S6Kα) and S6K2 (also known as p70β S6 kinase or S6Kβ), which have cytoplasmic and nuclear splicing variants. The 23- and 13-amino-acid extensions at the N-termini of S6K1 and S6K2 contain nuclear localization signals (NLSs) that target these isoforms constitutively to the nucleus. The cytoplasmic form of S6K1 (S6K1-II) is predominantly cytosolic in serum starved cells, but translocates to the nucleus upon mitogenic stimulation. The presence of an additional nuclear localisation signal at the C-terminus of the cytoplasmic form of S6K2 (S6K2-II) determines its predominant nuclear localisation. Nucleaocytoplasmic shuttling has been reported for both cytoplasmic forms of S6K in cellular response to mitogenic stimuli, but the mechanisms of its regulation have not been elucidated.

[0004] S6 kinases are activated through mitogen- and nutrient-mediated pathways. Growth factor-activated receptor tyrosine kinases recruit PI3K, which, via its effectors PKB/Akt and PDK1, mediates S6K activation (Chung et al. (1994) Nature 1370, 71-75). Another major player in the activation of S6K is the mammalian target of rapamycin, mTor (FRAP) which senses the level of amino acids and possibly other nutrients within a cell (Hara et al. (1998) J. Biol. Chem. 273, 14484-14494). Mitogen- and nutrient-induced activation of S6K occurs through a complex series of serine/threonine phosphorylation events and requires precise interplay between specific domains and multiple phosphorylation sites (Dennis et al. (1998) J. Biol. Chem. 273, 14845-14852; Pullen and Thomas (1997) FEBS Lett. 410, 78-82; Weng et al. (1998) J. Biol. Chem. 273, 16621-16629.). Based on numerous studies a model for S6K activation has been proposed which implies that active conformation of the kinase is achieved by coordinated phosphorylations at regulatory residues clustered in the carboxyl-terminal autoinhibitory segment, the conserved kinase extension domain and the activation loop of the catalytic domain (Alessi et al. (1997) Curr. Biol. 7, 776-789; Dennis et al. (1998) J. Biol. Chem. 273, 14845-14852; Pullen and Thomas (1997) FEBS Lett. 410, 78-82). Regarding this model, the process of mitogen-induced S6K activation can be divided into several steps, regulated by divers signaling inputs. The initial step involves the action of an array of proline-directed kinases which phosphorylate the S6K carboxyl-terminal segment (Ser434, Ser441, Thr444, Ser447 and Ser452 in S6K1), and perhaps Ser394 in the linker domain (Price et al. (1991) Mol. Cell. Biol. 11, 5541-5550). This disrupts the inhibitory interaction between the carboxyl and amino termini and subsequently releases the phosphorylated carboxyl-terminal autoinhibitory segment from the catalytic domain, providing access to Thr412 and Thr252. The activation loop phosphorylation sites, Thr252 and Thr412, in the kinase extension domain of S6K1, are common regulatory elements found in most members of the AGC family of serine/threonine kinases (Hanks and Hunter (1995) FASEB J. 9, 576-96; Marshall, 1994). The kinase responsible for phosphorylation of Thr252 in vivo has been identified as PDK1 (Alessi et al. (1997); Pullen et al. (1998) Science 30 279, 707-710; Williams et al., 2000), which is know to phosphorylate several other substrates, such as PKB, PKCs and PKA. The identity of the kinase which can phosphorylate S6K at T412 is still unclear. It has been recently shown that Thr 412 in S6K1 can be directly phosphorylated by the mTOR kinase in vitro and that this phosphorylation induces substantial increase in S6 kinase activity (Burnett et al. (1998) Proc. Natl. Acad. Sci. USA 95, 8351-8356; Isotani et al. (1999) J. Biol. Chem. 274, 34493-34498). Subsequent phosphorylation of S6K1 by PDK1 gives a strong synergistic increase in S6 kinase activity (Burnett et al. (1998)). In addition, the NIMA-related kinases NEK6 and NEK7 have recently been identified and shown to phosphorylate Thr412, as well as some other sites, and activate S6K1 in vitro and in vivo, in a manner synergistic with PDK1 (Belham et al. (2001) Curr. Biol. 11, 1155-67).

[0005] Most of the phosphorylation sites identified in S6K1 are conserved in S6K2 and include: a set of (Ser/Thr)-Pro motifs clustering in the autoinhibitory sequence and region located immediately carboxyl-terminal to this sequence (Ser423, Ser430, Ser436 and Ser441 in S6K2-I, which correspond to Ser434, Ser441, Ser447 and Ser452 in S6K1-I, respectively); Thr379, Ser383, Thr401 and Ser416 located in the kinase extension domain, which correspond to Thr390, Ser394, Thr412 and S427 in S6KαI, respectively; and Thr241 located in the catalytic loop, which corresponds to Thr252 in S6K1 (Gout et al. (1998) J. Biol. Chem. 273, 30061-30064). However, one of the known S6K1 regulatory residues, represented by Thr444, is not conserved and substituted to valine in S6K2. Based on a comparison of the phosphorylation sites between these two kinases, it has been predicted that S6K2 might be activated through mechanisms which are similar, but not identical, to those proposed for S6K1 (Gout et al. (1998) J. Biol. Chem. 273, 30061-30064; Shima et al. (1998) EMBO J. 17, 6649-6659).

[0006] S6K is inactivated by complexing with protein phosphatases, such as PP2A and PP1 (Andres et al. (1987) J. Biol. Chem. 262, 14389-14393; Peterson et al. (1999) Proc. Natl. Acad. Sci. USA 96, 4438-4442). This interaction leads to dephosphorylation and subsequent inactivation of S6K. Two tumor suppressors, PTEN and TSC1/2 complex, are known to inhibit indirectly the activity of S6K. PTEN is a lipid phosphatase, which reduces the level of PIP3, a key mediator of PI3-K signaling pathway (Sansal et al. (2004) J. Clin. Oncol. 22, 2954-2963). The activation of tumor sup-

pressor complex TSC1/2, a negative regulator of mTOR, leads to full inactivation of S6K (Harrington et al., (2004) J. Cell. Biol. 166, 213-23).

[0007] Both kinases share a high level of homology between their kinase domains, but differ significantly in the N- and C-terminal regulatory regions. It is believed that conformational changes induced by multiple S/T phosphorylations open the structure of S6K, making both N- and C-terminal regulatory domains available for protein-protein interactions. The list of S6K-binding partners includes small GTPases Rac1 and cdc42, protein kinases PDK1 and PKC zeta, protein phosphatase PP2A and PP1 and cytoskeletal protein neurabin.

[0008] Ribosomal protein S6 is the most widely studied physiological substrate of S6K. The phosphorylation of S6 protein was shown to closely correlate with the initiation of protein synthesis induced by various extracellular stimuli. Phosphorylation of IRS-1 by S6K was shown to be essential to multiple pathways that make cells unresponsive to insulin, including those initiating from continual exposure to insulin or free fatty acids (Giraud et al. (2004) J. Biol. Chem. 279, 3447-3454). The transcriptional activator CREM, elongation factor 2 kinase, the regulator of apoptosis Bad 1 and cytoskeletal protein neurabin have also been shown to be phosphorylated by S6K in vitro and in vivo (de Groot et al. (1994) Cell. Signal. 10, 619-628; Harada et al. (2001) Proc. Natl. Acad. Sci. USA 98, 9666-9670; Wang et al. (2001) Biochem J. 358, 497-503; Burnett et al. (1998) Proc. Natl. Acad. Sci. USA 95, 1432-1437). However, the physiological relevance of these phosphorylations requires further investigation, since other protein kinases can phosphorylate these molecules at identical sites.

[0009] Studies from numerous laboratories indicated that S6Ks are key players in ribosomal biogenesis and protein biosynthesis. The expression of translational components which make up the protein biosynthetic apparatus is controlled via S6K pathway (Thomas (2002) Biol. Res. 35, 305-313). In addition, a functional link between mTor/S6K signaling and Coenzyme A biosynthesis has been recently established (Nemazanyy et al. (2005) FEBS J. 272, E1-046P). Cell cycle progression requires S6K function at G1/S and G2/M phases (Banatti et al., 1998). S6K was found to be involved in regulating the expression of genes essential for cell cycle progression. In addition to playing an essential role in regulating growth and cell cycle progression, S6K1 appears to be a multifunctional protein involved in other cellular processes such as cell survival and RNA processing. These links are mediated through phosphorylation of anti-apoptotic protein Bad1 and complexing with SKAR protein, which has been proposed to couple transcription with pre-mRNA splicing and mRNA export (Harada et al. (2001) Proc. Natl. Acad. Sci. USA 98, 9666-9670; Richardson et al. (2004) Curr. Biol. 14, 1540-1549). So far, there are only sporadic data which implicate S6K in the regulation of transcription. It was reported that transcriptional activator CREM is the substrate for S6K in vitro and in vivo (de Groot et al. (1994) Cell. Signal. 10, 619-628). In addition, the function of transcription factor E2F was found to be regulated by mTor/S6K pathway in lymphocytes (Brennan et al. Mol. Cell. Biol. 19, 4729-4738).

[0010] Much of what is known about the physiological roles of S6Ks has come from genetic studies in *Drosophila* and mice. *Drosophila* possess only one S6K gene, dS6K, the disruption of which results in the death of most flies at the larval stage or during early pupation (Montagne et al. (1999) Science 285, 2126-2129). Those $S6K1^{-/-}$ flies that do survive, live for only a few weeks and the females are sterile. These flies are also much smaller than wild type and this reduction in body size is due almost entirely to a reduction in individual cell size rather than cell number (Montagne et al. (1999) Science 285, 2126-2129). Therefore, it appears that a major physiological role of dS6K is as a positive regulator of cell growth rather than cell proliferation. Indeed, studies in which dPI(3)K or dPKB were knocked out again produced small flies, but a deficit in cell number was also seen, suggesting that the pathway controlling cell growth and proliferation bifurcates downstream of PI(3)K/PKB but upstream of dS6K (reviewed in Weinkove and Leevers (2000) Curr. Opin. Genet. Dev. 10, 75-80).

[0011] Disruption of the S6K1 gene in mice also produced a small size phenotype. This was again largely due to a defect in cell growth rather than in proliferation, although a slight reduction in the rate of proliferation was observed during development (Shima et al. (1998) EMBO J. 17, 6649-6659; Kawasome et al. (1998) Proc. Natl. Acad. Sci. USA 95, 5033-5038). As mentioned earlier, in S6K1 knockout mice, S6K2 expression is upregulated and appears able to restore phosphorylation of rpS6 to levels close to those of wild type mice in all tissues examined (Shima et al. (1998) EMBO J. 17, 6649-6659). This indicated that phosphorylation of rpS6 is not the mechanism by which S6K1 exerts its positive effect on cell growth and suggested the existence of an S6K1-specific substrate (or substrates) involved in this process.

[0012] S6K has been recently implicated in ageing. Since caloric restriction increases lifespan, Kapahi and co-workers investigated whether the nutrient sensing TOR-S6K pathway may be involved in the regulation of lifespan in *Drosophila*. Overexpression of the upstream inhibitor of TOR/S6K signalling, dTsc2, extended lifespan, as did overexpression of a dominant negative dTOR construct or kinase dead dS6K. Conversely, overexpression of a constitutively active dS6K mutant decreased lifespan (Kapahi et al. (2004) Curr. Biol. 14, 885-90). This study demonstrates that like other positive regulators of cell growth, dS6K plays a role in the aging process of *Drosophila*. A recent study using primary human endothelial cells has shown that PKB-mediated inhibition of FOXO3a is involved in the onset of senescence and that inhibition of PKB extends the lifespan of these cells in culture (Miyauchi et al. (2004) EMBO J. 23, 212-20). This suggests that the insulin signalling pathway is also important in regulating the lifespan of human cells. Interestingly, S6K1 was recently identified as a positive mediator of Ras-induced senescence, suggesting that it too may be a negative regulator of cellular lifespan in humans.

[0013] Deregulation of S6K function has been implicated with human pathologies, like cancer and diabetes. Mice lacking S6K1 display hypoinsulinaemia exhibit glucose intolerance (a phenotype closely resembling that of preclinical type II diabetes mellitus). Examination of the pancreatic beta cells of these mice revealed that they were significantly smaller than those of the wild type littermates and that therefore their capacity for insulin production was impaired (Pende et al. (2000) Nature 408, 994-997). This underlines the importance of S6K I-mediated signalling and demonstrates how defects in cell growth can have secondary effects on cellular function.

[0014] A recent study has also shown that $S6K1^{-/-}$ mice are resistant to age- and diet-induced obesity. The rate of oxidative phosphorylation in white adipose tissue and muscle of

S6K1$^{-/-}$ mice is higher than in wild type animals. Accordingly, the mRNA levels of several genes involved in energy combustion and oxidative phosphorylation are highly upregulated in these tissues from the S6K1$^{-/-}$ mice (Um et al. (2004) Nature 431, 200-205). As mentioned above, the hypoinsulinaemia and glucose intolerance manifest in the S6K1$^{-/-}$ mice, would normally be expected to result in the development of insulin resistance seen in type II diabetes (Pende et al. (2000) Nature 408, 994-997). However, when mice were fed on high fat diets, the wild type animals developed insulin-resistance, while the S6K1$^{-/-}$ mice remained insulin-sensitive. This was associated with a decrease in insulin signalling via PKB in wild type mice not seen in the knockouts (Um et al. (2004) Nature 431, 200-205). A major advance in linking S6K signalling to cancer has come from the study of two tumour suppressor genes, TSC1 and TSC2, which encode the proteins hamartin and tuberin, respectively (reviewed in Marygold and Leevers (2002) Curr. Biol. 12, R785-7; McManus and Alessi (2002) Nat. Cell. Biol. 214, E214-216). Patients suffering from the relatively common autosomal dominant disorder, tuberous sclerosis complex (TSC) all have mutations in either TSC1 or TSC2. Inheritance of germ line mutations or acquisition of sporadic mutations in these genes leads to the development of benign tumours (hamartomas) in multiple tissues, particularly in the kidneys, skin, heart, eyes and brain, in which they often cause mental retardation and seizures. Numerous observations have identified the Tsc1/2 heterodimer as the link between PI(3)K/PKB signalling and mTOR/S6K activation. They provide a mechanism by which growth factor signals and nutrient signals feed into a common downstream pathway and explain the high sensitivity of S6K activity to PI(3)K inhibitors such as wortmannin and LY-294002 (reviewed in Manning and Cantley, (2003) Trends Biochem. Sci. 28, 573-6). It has been clearly established that aberrant S6K signalling is a major feature of TSC (Goncharova et al. (2002) J. Cell. Biol. 167, 1171-82; Kenerson et al. (2002) Cancer Res. 62, 5645-50). However, the fact that TSC patients develop benign tumours confirms data suggesting that constitutive TOR-S6K signalling alone is not sufficient to induce complete cellular transformation. There is however, an accumulating body of evidence suggesting that constitutive S6K signalling is important to the growth of many tumours. A recent report shows a reduction in the level of tuberin in a number of sporadic squamous and basal cell carcinomas. This effect was shown to be epigenetic, suggesting that cancer cells that downregulate tuberin expression at the level of transcription or translation obtain a selective growth advantage over the surrounding normal cells (Wienecke et al. (2002) J. Cutan. Pathol. 29, 287-290).

[0015] The gene encoding S6K1 was found to be located to chromosomal region 17q23, which is a frequent site of gene amplification in breast cancer (Barlund et al. (2000) Cancer Res. 60, 5340-5344). The gene encoding S6K2 was mapped to chromosomal region 11q13, which is associated with acute myeloid leukemia, Non-Hodgkin's lymphoma, chronic lymphoblastic disorder. Furthermore, S6K1 amplifications were identified in both breast cancer tissues and cell lines. In addition, an elevated level of S6K1 expression have been demonstrated for other human malignancies, such as thyroid, endometrial and ovary cancer. S6K overexpression in human meningioma cell lines resulted in increased tumor size in vivo (Surace et al. (2004) Ann. Neurol. 56, 295-298). Taken together, these studies clearly indicate the involvement of S6K1 and 2 in malignant transformation.

[0016] So far, no direct and specific inhibitors for S6K have been identified. Under these circumstances, mTor inhibitor Rapamycin and PI3-K inhibitor Wortmannin have been used extensively to elucidate the function and regulation of S6Ks. These two unrelated inhibitors fully abrogate the activation of S6K induced by mitogenic stimuli. The development of water-soluble esters of rapamycin, such as CCI-779 and RAD001, which are potentially suitable for administration as anti-cancer agents has allowed such experiments to progress into clinical trials (reviewed in Hidalgo and Rowinsly (2000) Oncogene 19, 6680-6686; Bjornsti and Houghton (2004) Nat. Rev. Cancer 4, 335-348). In the USA, CCI-779 is currently being tested in thirteen phase II clinical trials for the treatment of a wide range of cancers and one phase III trial, for renal cell carcinoma (http://clinical trials.gov). These developments have encouraged further interest in the mTOR/S6K signalling pathway in the context of cancer. It is important to note, that Rapamycin derivatives demonstrate impressive activity against a broad range of human cancers and exhibit low toxicity (Hidalgo and Rowinsky (2000) Oncogene 19, 6680-6686).

[0017] Recent data from our laboratory implicate S6K not only in the regulation of gene expression at the level of translation, but also transcription. Using DNA microarrays and stable cells lines overexpressing activated forms of S6K1 and S6K2, we identified a panel of genes which are strongly induced by S6Ks. The mechanism of transcriptional activation by S6Ks is not known.

[0018] Gene expression in mammalian cells is achieved through the formation of multienzyme complexes around DNA sequences located in the regulatory regions of the gene, which is usually distinct from the coding region. The formation of regulatory complexes is facilitated by DNA binding proteins, such as transcription factors. DNA binding proteins possess binding domains which recognize and associate with DNA in sequence specific or sequence-independent manner (A. Klug, (1995) Ann. N.Y. Acad. Sci. 758, 143-160). A myriad of sequence specific DNA binding domains have been identified, including homeobox, zinc finger, leucine zipper, helix-loop-helix etc. DNA-binding HMG-box domains A and B or AT-hook motifs bind DNA without sequence specificity. However, HMG-box domains exhibit a high affinity for bent, distorted or bend linear DNA, while AT-hooks recognize DNA minor grove which is AT rich (Thomas (2001) Biochem. Soc. Trans, 29, 395-401).

[0019] The AT-hook is a small DNA-binding protein motif which was first described in the high mobility group non-histone chromosomal protein HMG-I(Y) (Reeves (2000) Environ. Health Perspect. 108, Suppl. 5, 803-809). Since its discovery, this motif has been observed in other DNA-binding proteins from a wide range of organisms. Unlike several of the other well-characterized DNA-binding motifs, the AT-hook is a small motif, which has a typical sequence pattern centered around a glycine-arginine-proline (GRP) tripeptide. The importance of this short conserved sequence is stressed by the observation that it is necessary and sufficient to bind DNA. In some proteins containing AT-hook motifs, there is a possibility that these may serve as accessory DNA-binding domains for several transcription factors, presumably to anchor them to particular DNA structures (e.g. minor grooves and four-way junctions). These AT-hook motifs seem to be auxiliary elements necessary for cooperation with other DNA-binding activities in the same or different proteins (Sus-

bielle et al. (2005) Curr. Med. Chem. Anticancer Agents 5, 409-420; Reeves et al. (2003) J. Biol. Chem. 278, 42106-42114).

[0020] The AT-hook forms a C-shaped structure with the 'rear' of the concave surface inserted into the minor groove of the DNA. The proline residues in the AT-hook are probably responsible for maintenance of the rigid structure of this domain in the presence of DNA while the characteristic RGR sequence motif of the AT-hook adopts an extended structure and participates in DNA-protein interactions. The arginine residues are seen to penetrate deep into the B-DNA helix and interact with the bases of DNA.

[0021] Multiple or single AT-hook motifs are also found in several multi-domain proteins which associate with chromatin. Striking examples include the human HRX (ALL-1) protein, $TAF_{II}250$, *Saccharomyces cerevisiae* ASH1, *S. cerevisiae* SWI2, *Drosophila melanogaster* ISWI, ENBP1, M33 (vertebrate polycomb), doom/Mod(Mdg4) and tramtrack (Wassarman et al. (2001) J. Cell. Sci. 114, 2895-2902; Mohrmann et al. (2005) Biochim. Biophys. Acta. 1681, 59-73, Harvey et al. (1997) Mol Cell Biol. 17, 2835-2843). Several of these proteins are involved in the organization of chromatin into specific states. HRX (ALL-1) and ASH1 are members of the trithorax group of regulatory genes which are involved in chromatin decondensation. SWI2 and ISWI are ATP-dependent chromatin remodeling proteins. $TAF_{II}250$ is a basal transcription factor which is involved in transcription regulation and also contains histone acetyltransferase activity. In contrast, the POZ domain proteins Mod(Mdg4) and tramtrack act as regulators of transcriptional insulation and mediators of transcription repression, respectively.

[0022] The presence of AT-hooks in these proteins suggests that these chromatin architecture regulators probably use DNA sequence information in the form of AT-rich tracts, to target specific regions on the chromosomes. The AT-richness of SARs and their presumptive involvement in transcription activation suggest that the AT-hooks may indeed play a role in tethering these proteins to specific chromosomal regions. Both ARBP and its ortholog, the mammalian methylated DNA binding protein MeCP2, contain both an AT-hook motif and a methylated CpG DNA binding domain (Klose et al. (2005) Mol. Cell. 19, 667-78).

[0023] There are several examples where a single AT-hook motif was detected in proteins which have additional larger DNA-binding domains. In addition to the AT-hook *D. melanogaster* apterous, mammalian LH2 and Barx1 also contain homeodomains; human RFX5 protein also contains the DNA-binding RFX box; human ESE-1 protein also contains an ETS domain; *S. cerevisiae* SWI5, *Drosophila* castor and mouse CTCF1 also contain zinc fingers; the fungal metal binding transcription factors of the AMT1 family of proteins contain a zinc cluster. This suggests that in several cases, the AT-hook may serve as an additional peptide contact in the minor groove even as the principal or larger DNA-binding protein contacts the major groove. It is tempting to speculate that in these cases the AT-hook may serve as a 'built-in' cofactor like the HMG-I(Y) protein and subtly alter the affinity and specificity of the 'principal or larger' DNA-binding domain by recognizing sites in the vicinity of the principle DNA binding site. This may again be a general feature of eukaryotic DNA-binding proteins: small basic patches such as the N-terminus of homeodomains and the recently described GRIP box of nuclear hormone receptors, which contact the minor groove in addition to the main DNA-binding activity in the major groove.

[0024] In the protein kinase family, which is highly diverse and includes over 500 proteins, only few members were found to possess functional DNA binding domains. These are cABL tyrosine kinase, DNA-dependent protein kinase (DNA-PK), ATM kinase (Ataxia Telangiectasia Mutated). Numerous studies from various laboratories implicated these kinases in the regulation of transcription, genome stability, DNA repair, cell cycle control and survival (Taylor et al. (2005) J. Clin. Pathol. 58, 1009-1015; Meek et al. (2004) Immunol. Rev. 200, 132-141; Giles et al. (2005)). Deregulation of their function leads to cancer. Therefore, they are excellent targets for drug development (Hannah (2005) Curr. Mol. Med. 5, 625-42).

SUMMARY OF THE INVENTION

[0025] The invention encompasses an isolated peptide or derivative thereof comprising a binding domain of p300 acetyltransferase of less than 200 amino acids of SEQ ID NO: 4 wherein the binding domain is capable of binding to and modulating an S6 kinase protein activity. In one embodiment, said binding domain binds to an activated form of S6 kinase protein and can comprise a histone acetyltransferase domain capable of acetylating the S6 kinase protein. One site of acetylation can be at an amino acid corresponding to residue 516 of an S6 kinase protein. In some embodiments, the acetylation site comprises a lysine residue.

[0026] The invention also includes isolated nucleic acid molecules encoding the above isolated peptides or derivatives thereof. In some embodiments, the isolated nucleic acid molecule specifically hybridizes under stringent conditions to a complementary nucleic acid molecule.

[0027] The invention also encompasses an isolated protein or fragments thereof comprising the A-T hook motif of S6 kinase 2. The isolated peptide may lack the ability to bind DNA. In one embodiment of the invention the isolate protein has one or more mutations, such as a substitution and/or deletion, at any amino acid residue corresponding to amino acids 487 to 495 of SEQ ID NO: 8. In another embodiment there is a substitution at amino acid residue 492 of SEQ ID NO: 8, which may be e.g. a substitution from Arginine to Proline.

[0028] The invention also includes nucleic acid molecules encoding such isolated proteins or fragments thereof. These nucleic acid molecules may be operably linked to one or more expression control elements. The nucleic acid molecules or fragments thereof may be part of a vector. Such vectors may be used to transform a host cell line which may consist of prokaryotic or eukaryotic cells.

[0029] The invention also encompasses methods of producing a polypeptide comprising culturing a host cell under conditions in which the polypeptide encoded by said nucleic acid molecule or fragment thereof is expressed.

[0030] The invention further encompasses methods of identifying an agent which binds to the S6 kinase protein, comprising exposing the S6 kinase protein to the agent in the presence of the isolated p300-4 peptide, and measuring the specific binding of the isolated peptide to the S6 kinase protein, wherein a decrease in the binding of the isolated peptide compared to a control is indicative of an agent capable of binding to the S6 kinase protein.

[0031] The invention also encompasses a method of identifying an agent which modulates the acetylation of the S6

kinase protein, comprising exposing the S6 kinase protein to the agent, and measuring the acetylation of the S6 kinase protein by the isolated peptide; wherein a change in the acetylation of S6 kinase compared to a control is indicative of an agent capable of modulating the acetylation of S6 kinase by the isolated peptide. In some embodiments of these methods, they further comprise contacting the S6 kinase protein with the above isolated peptide prior to measuring acetylation.

[0032] The invention also includes a method of identifying an agent which modulates an activity of the S6 kinase protein, comprising exposing the S6 kinase protein to the agent, measuring the acetylation at a lysine residue at position 516 of the S6 kinase protein by the agent, and measuring the activity of the S6 kinase protein, wherein a change in activity compared to a control is indicative of an agent capable of modulating an activity of the S6 kinase protein. In some embodiments of this method, the activity of the S6 kinase protein is determined by measuring phosphorylation of one or more additional proteins.

[0033] In further embodiments of all of these methods, the control includes a negative control. In one embodiment, the negative control does not comprise the agent. In all of these methods, a positive control can be included. In some embodiments, the positive control comprises a class I/II HDAC inhibitor such as Trichostatin-A.

[0034] The invention further encompasses methods of identifying an agent that modulates the binding of S6 kinases, preferably S6 kinase to a nucleic acid (e.g. DNA). In one embodiment, the method comprises exposing S6 kinase 2 protein to an agent; and detecting the binding of the S6 kinase 2 protein to the nucleic acid, wherein change in the level of binding compared to a control is indicative of an agent capable of modulating S6 kinase 2 binding to the nucleic acid. The method may further comprise assaying the activity of S6 kinase 2 by such means as e.g. phosphorylation of other proteins. The methods may utilize both positive and negative controls. In one embodiment, the positive control protein comprises the AT hook domain of S6K2 (corresponding to amino acids 487 to 495 of SEQ ID NO: 8). The invention also encompasses agents identified by such methods.

[0035] The invention also includes a method of modulating the activity of an S6 kinase protein comprising contacting the S6 kinase protein with the isolated peptide or derivative thereof as described herein. In some embodiments, the S6 kinase protein activity is inhibited. The invention further includes a method of treating a protein kinase related disease in a patient in need of such treatment comprising administering a therapeutically effective amount of a pharmaceutical composition comprising an agent capable of acetylating the protein kinase. In some embodiments, the agent is the isolated peptide described herein while in other embodiments the agent is a deacetylase inhibitor and/or a kinase inhibitor including S6 kinase inhibitors.

[0036] The invention further encompasses methods of identifying an agent that modulates the binding of S6 kinases, preferably S6 kinase to a nucleic acid (e.g. DNA). In one embodiment, the method comprises exposing S6 kinase 2 protein to an agent; and detecting the binding of the S6 kinase 2 protein to the nucleic acid, wherein change in the level of binding compared to a control is indicative of an agent capable of modulating S6 kinase 2 binding to the nucleic acid. The method may further comprise assaying the activity of S6 kinase 2 by such means as e.g. phosphorylation of other proteins. The methods may utilize both positive and negative controls. In one embodiment, the positive control protein comprises the AT hook domain of S6K2 (corresponding to amino acids 487 to 495 of SEQ ID NO: 8). The invention also encompasses agents identified by such methods.

[0037] The invention also includes methods of treating a disease associated with protein kinase (e.g. S6 kinase 2 protein) in a patient comprising administering to the patient of an effective amount of an agent to modulate the activity of the S6 kinase 2 protein. Such an agent may downregulate the activity of S6 kinase protein. Such an agent may be an antibody (e.g. a monoclonal antibody. In one embodiment, the S6 kinase 2 protein is human S6 kinase 2 (SEQ ID NO: 8). In another embodiment, the S6 kinase 2 protein has 95% homology to SEQ ID NO: 8.

[0038] Exemplary protein kinase related diseases include, but are not limited to, cancer, blood vessel proliferative disorders, autoimmune disorders, and metabolic diseases. Types of cancer which can be treated by the methods of the invention include, but are not limited to, squamous cell carcinoma, astrocytoma, Kaposi's sarcoma, glioblastoma, multiple myeloma, lung cancer, bladder cancer, head and neck cancer, melanoma, ovarian cancer, prostate cancer, breast cancer, small-cell lung cancer, glioma, colorectal cancer, genitourinary cancer, gastrointestinal cancer. Types of blood vessel proliferative disorders which can be treated by the methods of the invention include, but are not limited to, diabetic retinopathy, age-related macular degeneration, retinopathy of prematurity, arthritis and restenosis. Types of autoimmune disorders which can be treated by the methods of the invention include, but are not limited to, lupus erythematosus, discoid lupus erythematosus, subacute cutaneous lupus erythematosus, drug-induced lupus erythematosus, and systemic lupus erythematosus. Types of metabolic disorders which can be treated by the methods of the invention include, but are not limited to, psoriasis, diabetes mellitus, wound healing, inflammation and neurodegenerative diseases.

[0039] The invention also encompasses an antibody which specifically binds to an acetylation site on a kinase protein and is capable of inhibiting acetylation of the kinase protein by an acetyltransferase protein. In one embodiment, the antibody binds to S6 kinase on or near amino acid residue.

BRIEF DESCRIPTION OF THE DRAWINGS

[0040] FIG. **1** shows binding of S6 kinases to acetyltransferases p300 and P/CAF. In FIG. **1**A, S6K1 and S6K2 were immunoprecipitated from MCF-7 cells using polyclonal antibodies or pre-immune serum control. Co-immunoprecipitated p300 was detected by western blotting with specific antibody. In FIG. **1**B, 293 cells were cotransfected with either EE-S6K1 and Gal4-p300 or empty vector and Gal4-p300. S6K1 was immunoprecipitated using anti-EE antibody and associated p300 was detected by western blotting with anti-Gal4 antibody. Input: approximately ten percent of soluble lysate. In FIG. **1**C, NIH-3T3 cells were transfected with empty vector or EE-S6K2. Anti-EE immunoprecipitates were subjected to western blotting with anti-p300 antibody. Input: approximately one percent of soluble lysate. FIG. **1**D demonstrates the interaction between S6 kinases with P/CAF in cells following serum stimulation. Endogenous S6K1 and S6K2 were immunoprecipitated from MCF-7 cells using polyclonal antibodies. Associated P/CAF was detected by western blotting with polyclonal antibodies (Upstate). In FIG. **1**E, the full length p300 was used to produce several GST-fusion fragments as shown (Hasan et al. (2001) Mol.

Cell. 7, 1221-1231). Expressed fusion proteins were coupled to glutathione beads and used in an in vitro GST pulldown assay with recombinant S6K2. After extensive washing, bound S6K2 was detected by western blotting.

[0041] FIG. **2** shows acetylation of S6 kinases by p300 and P/CAF in vitro and in vivo. In FIG. **2**A, recombinant S6 kinases were acetylated in vitro using the purified acetyltransferase domain of p300 or full length PCAF and ^{14}C-labelled acetyl CoA. Proteins were resolved by SDS-PAGE and acetylated proteins were analysed by phosphorimaging (Bio-rad). In FIG. **2**B, recombinant S6K2 or GST were subject to in vitro acetylation reactions with p300 and 14C AcCoA. In vitro acetylation of other protein kinases is shown in FIG. **2**C. Recombinant PKB/Akt, PKCa, PDK1 and PI3K (p85a/p110a) were purified using baculovirus expression system. In vitro acetylation assay was carried out as previously described. In FIG. **2**D, HEK 293 cells were transfected with S6 kinases in the absence or presence of Gal4-p300 as shown. Cells were treated with medium containing either solvent control or TSA to a final concentration of 660 nM for twelve hours prior to harvesting. Recombinant S6 kinases were immunoprecipitated using the EE-antibody and acetylated S6 kinase was detected by western blotting with polyclonal acetylated lysine antibodies (Upstate). The level of immunoprecipitated S6K1 and S6K2 was analysed by reprobing the membranes with specific antibodies. The expression of Gal4-p300 was tested in total lysates using anti-Gal4 antibodies.

[0042] FIG. **3** shows identification of acetylation sites in S6K1 by mass spectrometry. In this study recombinant S6 kinases were acetylated in vitro using p300 immunoprecipitated from cells. Proteins were resolved by SDS-PAGE and S6 kinase bands from acetylation and control (no p300) reactions were excised after Coomassie staining. Proteins were subject to in-gel trypsin digest and MALDI-TOF mass spectrometry. An example of MS/MS analysis of an acetylated peptide in shown in FIG. **3**A. The peptide shown was detected in the acetylated sample in the form of two peaks separated by 43 Da indicating presence of an acetyl group. The identified peptide (shown in the box) corresponds to the C-terminal region of S6K1, where K516 was found to be acetylated. FIG. **3**B shows alignment of P-loop lysines in S6Ks from different species and FIG. **3**C shows alignment with P-loops of other AGC kinases.

[0043] FIG. **4** shows substitution of S6K P-loop lysines to arginines or glutamines significantly inhibits S6 kinase activity. Wild type or mutant S6Ks were expressed in 293 cells, immunoprecipitated using anti-EE antibody and subjected to in vitro kinase assays. (A) Graphs showing the S6K activity of wild type S6K1 and S6K2 compared with constructs bearing substitutions of arginine (RR) or glutamine (QQ) for the P-loop lysines (P=0.1, n=3) (B) Representative autoradiograph of one such experiment. (C) Equal quantities of lysates from cells overexpressing wild type S6K1 or the P-loop mutants were blotted with anti-phospho-T412 or with anti-EE to compare S6K phosphorylation between wild type and mutant constructs.

[0044] FIG. **5** shows the position of the P-loop and conserved lysines in a putative model of the S6 kinase domain. The modelling of the S6 kinase domain was based on a crystal structure of Akt/PKB, a member of the AGC family. The model was generated using MODELLER7 programme with standard comparative modelling parameters. The location of two conserved lysines in the P-loop is indicated in red and blue. The rest of the loop is shown in green. Both lysines are well accessible on the surface of S6 kinase domain. The alignment of the P-loop sequences from various members of AGC family is shown.

[0045] FIG. **6** shows that acetylation affects S6K1 activity in cells. In FIG. **6**A (upper panel), C2C12 myoblasts were treated for 18 hours with increasing concentrations of TSA or with the structurally unrelated HDAC inhibitor, sodium butyrate. Endogenous S6 kinase activity was assessed by western blotting against pT412 and against phosho-rpS6. In another experiment, C2C12 cells were treated with 1 μM TSA or control for the times shown and S6 kinase activity was measured by in vitro kinase assay using purified ribosomes as substrate (FIG. **6**B). A representative autoradiograph is shown beneath the graph.

[0046] FIG. **7** shows the effect of histone deacetylase inhibitor TSA on S6K activity in a range of different cell types. (A) The cell lines indicated (see text) were treated with TSA and nicotinamide for eight hours. Following lysis, proteins were resolved by SDS-PAGE and western blots were probed as labelled to evaluate S6K activation. (B) MKOC cells (see text) were treated as in (A) and analysed by western blotting. S6K1 was also immunoprecipitated from lysates and subjected to immune complex kinase assay. (C) To check whether TSA or nicotinamide directly inhibit S6K1, recombinant purified S6K1 was used for kinase assays in the absence or presence of deacetylase inhibitors as indicated.

[0047] FIG. **8** shows the analysis of endogenous S6K1 acetylation during the cell cycle. NIH-3T3 cells were synchronized by serum starvation for twenty-four hours followed by addition of serum for the indicated times. Endogenous S6K1 was immunoprecipitated and acetylation was detected by western blotting with anti-acetyl lysine antibodies. Blots were reprobed with polyclonal antibodies against total S6K1. Lysates were resolved by SDS-PAGE and western blotting was carried out to measure S6K activation at each stage using anti-S6K pT412 and anti-pS6 (235/236) antibodies.

[0048] FIG. **9** shows testing the specificity of acetyl-Lysine specific S6K1 antibodies. Synthetic peptides, possessing acetylated lysines identified by mass spectrometry, were used for the immunisation of rabbits. Generated antibodies were initially purified on a column containing acetylated peptide. Then, purified antibodies were loaded on a column containing corresponding non-acetylated peptides in order to remove IgGs which recognise non-acetylated form of S6K. The specificity of purified antibodies was tested by Western blotting of total lysates or immunoprecipitates from Hek293 cells transfected with wtS6K1, K516R S6K mutant and vector alone.

[0049] FIG. **10** shows S6K1 is acetylated at K516 in vivo. FIG. **10**A shows that acetylation of S6K1 at K516 is induced by serum stimulation. In this study, NIN3T3 cells were starved and stimulated with serum for various period of time. Total cell lysates were separated by SDS-PAGE and immunoblotted with acetyl-K516 antibodies. In FIG. **10**B, the effect of various cell signalling inhibitors on K516 acetylation is shown. In this study, NIH3T3 cells were starved and treated with 50 nM Rapamycin (mTor inhibitor), 20 μM LY294002 (phosphatidylinositol 3'-kinase inhibitor), 10 μM PD98059 (MAP kinase inhibitor) or vehicle alone thirty minutes before stimulation with serum. Then, cells were stimulated with serum for one hour and lysed. The supernatants of total lysates were separated by SDS-PAGE and Western blotted with acetyl-K516 antibodies. FIG. **10**C shows that acetylation of S6K1 at K516 does not affect its kinase activity. The

wtS6K and K516R mutants were transiently expressed in Hek293 cells. Immunoprecipitated recombinant S6Ks and total cell lysates were resolved by gel electrophoresis and immunoblotted with antibodies specific for EE tag, pT412 and actin.

[0050] FIG. **11** shows S6K2CT binding to DNA in vitro. S6K1 and S6K2 C-terminal domains were coupled to nickel-NTA beads and assayed for binding to ^{32}P labeled genomic DNA. Bound DNA was measured by scintillation counting of beads following extensive washing.

[0051] FIG. **12** shows the in vitro DNA binding of S6K2 and S6K1. FIG. **12**A shows that full length S6K2 but not S6K1 binds DNA in vitro. Recombinant EE-S6K1 and EE-S6K2 were immobilized on protein-G sepharose using anti-EE antibody and assayed for binding to ^{32}P labeled genomic DNA. Bound DNA was measured by scintillation counting of beads following extensive washing. FIG. **12**B shows an electrophoretic mobility shift assay mobility shift assay (EMSA), thus demonstrating binding of recombinant EE-S6K2 to genomic DNA in vitro.

[0052] FIG. **13** shows that the S6K2 C-terminal domain is required for DNA binding. FIG. **3**A shows the in vitro binding of DNA to the C-terminal domain of S6K2. HEK cells were transiently transfected with EE-S6K1, EE-S6K2 or with EE-S6K2ΔCT. Proteins were then immunoprecipitated using anti-EE and used in an in vitro DNA binding assay with genomic DNA. FIG. **13**B shows expression of the recombinant protein. Total protein (30 μg) from soluble lysates were subjected to western blotting with anti-EE to measure expression of the recombinant proteins.

[0053] FIG. **14** shows that mutation of the S6K2 AT-hook motif inhibits binding of the full length protein to DNA. FIG. **14**A compares the DNA binding abilities of S6K2T WT and R492P. S6K2 WT and R492P were overexpressed in 293 cells and immunoprecipitated using anti-EE antibody. DNA binding to genomic DNA was assayed. FIG. **14**B shows the expression of the recombinant proteins. 30 μg of total protein from soluble lysates were subjected to western blotting with anti-EE to measure expression of the recombinant proteins.

[0054] FIG. **15** shows the effect of DNA binding on S6K2 kinase activity. FIG. **15**A is an analysis of S6K2 kinase assay in vitro in the presence of increasing concentration of DNA. FIG. **15**B is a bar chart representing the results from the in vitro kinase assay in **15**A. DNA-induced stimulation of S6K2 activity was observed in at least four different experiments.

DETAILED DESCRIPTION

[0055] The present invention is based on the discovery that, in addition to phosphorylation, protein kinases can also be modulated by acetylation. This is the first time that any kinase has been shown to be modulated by acetylation. The present invention is based upon the finding that S6 kinase protein is acetylated both in vivo and in vitro by p300 and P/CAF acetyltransferases. More accurate to say that in conditions under which S6K acetylation is increased (i.e. presence of HDAC inhibitors/overexpression of p300), S6 kinase activity and 412 phosphorylation are reduced. Mutation of the P-loop lysines (which are acetylated by p300 in vitro) to glutamine, which mimics acetylation, results in the complete inactivation of S6Ks and loss of 412 phosphorylation. The present invention is also based on the discovery that S6K2 possesses an AT-hook DNA binding motif. This is also the first time that any protein kinase has been shown to contain a functional AT-hook DNA binding motif. The present invention is based upon the finding that S6 kinase 2 protein binds to DNA and is thereby activated which in turn stimulates its kinase activity. Based upon these data a novel mechanism is proposed by which S6K2 can transduce growth-promoting effects in response to mitogens and nutrients. This may involve the regulation of transcription factors and/or chromatin remodeling proteins by phosphorylation, when S6K2 complexes with DNA and is activated by this interaction. This discovery has important implications for the development of drugs which can block S6K2-DNA interaction and subsequently its growth promoting functions, which are deregulated in cancer and diabetes.

[0056] It has been determined that this active domain of p300 acetyltransferase (p300-4) extends from amino acid 1,459 to 1,892 of this protein (SEQ ID NO: 6). Thus, any agent which modulates the acetylation of a protein kinase, including but not limited to the p300-4 fragment and its derivatives, can be used for the treatment of cell proliferative disorders associated with aberrant protein kinase activity.

[0057] The present invention also encompasses methods of modulating acetylation of a protein kinase utilizing acetyltransferase binding domains from other proteins in the histone acetyltransferase family of proteins which are similar to p300 acetyltransferase. In one embodiment, the acetyltransferase binding domain is from CBP (CREB-binding protein), which is a highly homologous to p300 acetyltransferase.

DEFINITIONS

[0058] As used herein, the term "p300 acetyltransferase binding domain" refers to a region of the p300 acetyltransferase (SEQ ID NO: 4) which interacts or binds to one or more species of S6 kinase protein including human. The interaction or binding may be by one or more of a variety of mechanisms including, but not limited to ionic interactions, covalent interactions, hydrophobic interactions, van der Waals interactions, etc. In one embodiment, the binding domain includes the fragment of p300 acetyltransferase known as p300-4 (i.e., p300-4 fragment, SEQ ID NO: 6). In another embodiment, the binding domain contains a histone acetyltransferase domain. The terms "p300 acetyltransferase binding domain" or "binding domain" are synonymous and used interchangeably throughout the specification. As used herein, "p300 acetyltransferase binding domain" does not include the full-length p300 acetyltransferase protein.

[0059] As used herein, the term "p300 acetyltransferase binding domain polypeptide" refers to any polypeptide containing a p300 acetyltransferase binding domain including fusion proteins, chimeras, variants, etc. In one embodiment, the p300 acetyltransferase binding domain polypeptide is capable of binding to a S6 kinase, including human S6 kinase. In a further embodiment, the p300 acetyltransferase binding domain polypeptide is capable of acetylating an S6 kinase, including human S6 kinase.

[0060] As used herein, the term "acetyltransferase" is used to refer to an enzyme responsible for the covalent modification of proteins via acetylation. In one embodiment, the acetyltransferase is a p300 acetyltransferase as set forth in SEQ ID NO: 4, a family member of the p300/CBP family of co-activators which have histone acetyltransferase activity. For the purposes of the present invention, reference to p300/CBP refers to human allelic and synthetic variants of p300 acetyltransferase, and to other mammalian variants and allelic and synthetic variants thereof, as well as fragments of said human and mammalian forms of p300 acetyltransferase.

The p300 acetyltransferase molecule will however retain the ability to physically associate both in vitro and in vivo with S6 kinase protein. For the purposes of the present invention, the precise form and structure of a p300-acetyltransferase protein or fragment thereof may be varied by those of skill in the art, having regard to the particular assay format to be used.

[0061] As used herein, the term "S6 kinase protein" refers to p70 S6 kinase protein that is capable of transferring a phosphate group from ATP to an acceptor group of protein. It also refers to the full sequence of the protein or any suitable derivative or fragment thereof or fusion of S6 kinase protein or derivative, or fragment thereof. Examples include, but are not limited to, subtypes of this protein such as S6K1 and S6K2.

[0062] As used herein, the term "S6K2 derivative" refers to a S6K2 protein comprising one or more amino acid mutations in the AT-hook domain to prevent DNA binding. The mutations may be substitutions and/or deletions.

[0063] As used herein, the term "AT-hook motif" refers to an amino acid sequence, which contains a short DNA binding motif capable of binding to the minor groove of short (4 to 6 nucleotide base pairs) 'AT-rich' (adenine-thymine) DNA. Examples of AT-hook motifs, include, but are not limited to the amino acid sequences set forth in Table 2. As used herein, the term "AT-hook motif" is used interchangeably with the term "AT-hook domain" throughout the specification.

[0064] Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, the preferred methods, and materials are described.

Polypeptides Containing p300 Acetyltransferase Binding Domain

[0065] As used herein, the term "p300 acetyltransferase binding domain derivative" refers to derivatives, analogs, variants, polypeptide fragments and mimetics of the p300 acetyltransferase binding domain and related peptides which retain the same activity as the p300-4 fragment, such as binding specifically to S6 kinase protein. Such p300 acetyltransferase binding domain derivatives can also be used for practicing the methods of the invention. Examples of derivatives include, but are not limited to, peptide variants of p300 acetyltransferase binding domain, for example, fragments comprising or consisting of contiguous n-mer peptides of SEQ ID NO: 4 or comprising at least about amino acid residues 50 to 100 or 150 to 200 of SEQ ID NO: 4, core binding sequences, and peptide mimetics.

[0066] Peptides containing a p300 acetyltransferase binding domain and derivatives thereof can be prepared using standard solid phase (or solution phase) peptide synthesis methods, as is known in the art. In addition, the nucleic acids encoding these peptides may be synthesized using commercially available oligonucleotide synthesis instrumentation and produced recombinantly using standard recombinant production systems. Production using solid phase peptide synthesis is necessitated if non-gene-encoded amino acids are to be included. The term "p300 acetyltransferase binding domain derivative" as used herein is synonymous with "variant" and also includes modifications to the p300 acetyltransferase binding domain by one or more deletions of up to ten (e.g., 1 to 2 or 1 to 5) amino acids; insertions of a total of up to ten (e.g., 1 to 5) amino acids internally within the amino acid sequence of p300-4 fragment; or up to a total of 100 amino acids at either terminus of the p300 acetyltransferase binding domain sequence; or conservative substitutions of a total of up to fifteen (e.g., 1 to 5) amino acids.

[0067] Derivatives of p300 acetyltransferase binding domain include polypeptides comprising a conservative or non-conservative substitution of at least one amino acid residue when the derivative sequence and the p300 acetyltransferase binding domain sequence are maximally aligned. The substitution may be one which enhances at least one property or function of the p300 acetyltransferase binding domain, inhibits at least one property or function of the p300 acetyltransferase binding domain, or is neutral to at least one property or function of the p300 acetyltransferase binding domain. As used herein, a "property or function" of the p300 acetyltransferase binding domain includes, but is not limited to, at least one selected from the group consisting of the ability to inhibit S6 kinase protein activity, arrest abnormal cell growth, specific binding to a benign or malignant cancer cell when compared to a non-cancer cell (i.e., normal), and killing of a benign or malignant cancer cell. In terms of the present disclosure, the cancer cell may be in vivo, ex vivo, in vitro, a primary isolate from a subject, a cultured cell, or a cell line.

[0068] Peptide variants include, but are not limited to, deletion or conservative amino acid substitution variants of SEQ ID NO: 6. As used herein, a conservative variant refers to alterations in the amino acid sequence that do not adversely substantially affect the biological functions of the peptide. A substitution, insertion, or deletion is said to adversely affect the peptide when the altered sequence substantially prevents or disrupts a biological function associated with the peptide (e.g., acetylation). For example, the overall charge, structure, or hydrophobic/hydrophilic properties of the peptide can be altered without adversely affecting a biological activity. Accordingly, the amino acid sequence can be altered, for example to render the peptide more hydrophobic or hydrophilic, without adversely affecting the biological activities of the peptide.

[0069] As discussed above, various changes can often be made to the amino acid sequence of a polypeptide which has a desired property in order to produce variants which still have that property. Such variants of the polypeptides are within the scope of the present invention. An example of a variant of the present invention is a polypeptide with the conservative substitution of one or more amino acids with one or more other amino acids. One skilled in the art is aware that various amino acids have similar characteristics. One or more such amino acids of a polypeptide can often be substituted by one or more other such amino acids without eliminating a desired property of that polypeptide (such as acetyltransferase activity). For example, the amino acids glycine, alanine, valine, leucine and isoleucine can often be substituted for one another (amino acids having aliphatic side chains). Of these possible substitutions it is preferred that glycine and alanine are used to substitute for one another (since they have relatively short side chains) and that valine, leucine and isoleucine are used to substitute for one another (since they have larger aliphatic side chains which are hydrophobic). Other amino acids that can often be substituted for one another include phenylalanine, tyrosine and tryptophan (amino acids having aromatic side chains), lysine, arginine and histidine (amino acids having basic side chains); aspartate and glutamate (amino acids having acidic side chains); aspar-

agine and glutamine (amino acids having amide side chains); and cysteine and methionine (amino acids having sulphur containing side chains).

[0070] Deletion and insertion variants of the polypeptides are within the scope of the present invention. Amino acid deletions can be advantageous since the overall length and the molecular weight of a polypeptide can be reduced whilst still retaining a desired activity. This can enable the amount of polypeptide required for a particular purpose to be reduced. Amino acid insertions relative to a polypeptide can also be made. This may be done to alter the nature of the polypeptide (e.g., to assist in identification, purification or expression). Polypeptides incorporating amino acid changes (whether substitutions, deletions or insertions) relative to the sequence of a polypeptide as defined above can be provided using any suitable techniques. For example, a nucleic acid sequence incorporating a desired sequence change can be provided by site-directed mutagenesis. This can then be used to allow the expression of a polypeptide having a corresponding change in its amino acid sequence.

[0071] The methods of the invention include corresponding polypeptides that display similar or related activity to p300-4 fragment for the diagnosis and treatment of diseases associated S6 kinase expression and activity including those with abnormal cell proliferation as described herein, including cancer. For purposes of the specification, "similar or related activity" is defined as binding to S6 kinase protein expressed by cells displaying abnormal cell growth, including benign cells exhibiting abnormal growth and malignant cancer cells. Examples of such polypeptide include, but are not limited to, polypeptides which contain one or more of the binding domains set forth in SEQ ID NO: 6.

[0072] Homology or sequence identity at the nucleotide or amino acid sequence level is determined by BLAST (Basic Local Alignment Search Tool) analysis using the algorithm employed by the programs blastp, blastn, blastx, tblastn and tblastx (Altschul et al. (1997) Nucleic Acids Res. 25, 3389-3402 and Karlin et al. (1990) Proc. Natl. Acad. Sci. USA 87, 2264-2268, both fully incorporated by reference) which are tailored for sequence similarity searching. The approach used by the BLAST program is to first consider similar segments, with gaps (non-contiguous) and without gaps (contiguous), between a query sequence and a database sequence, then to evaluate the statistical significance of all matches that are identified and finally to summarize only those matches which satisfy a preselected threshold of significance. For a discussion of basic issues in similarity searching of sequence databases, see Altschul et al. (1994) Nature Genetics 6, 119-129 which is fully incorporated by reference. The search parameters for histogram, descriptions, alignments, expect (i.e., the statistical significance threshold for reporting matches against database sequences), cutoff, matrix and filter (low complexity) are at the default settings. The default scoring matrix used by blastp, blastx, tblastn, and tblastx is the BLOSUM62 matrix (Henikoff et al. (1992) Proc. Natl. Acad. Sci. USA 89, 10915-10919, fully incorporated by reference), recommended for query sequences over eighty-five nucleotides or amino acids in length.

[0073] For blastn, the scoring matrix is set by the ratios of M (i.e., the reward score for a pair of matching residues) to N (i.e., the penalty score for mismatching residues), wherein the default values for M and N are +5 and −4, respectively. Four blastn parameters were adjusted as follows: Q=10 (gap creation penalty); R=10 (gap extension penalty); wink=1 (generates word hits at every winkth position along the query); and gapw=16 (sets the window width within which gapped alignments are generated). The equivalent Blastp parameter settings were Q=9; R=2; wink=1; and gapw=32. A Bestfit comparison between sequences, available in the GCG package version 10.0, uses DNA parameters GAP=50 (gap creation penalty) and LEN=3 (gap extension penalty) and the equivalent settings in protein comparisons are GAP=8 and LEN=2.

[0074] The present invention encompasses allelic variants, conservative substitution variants, and members of the p300 acetyltransferase binding domain having an amino acid sequence identity or homology of at least about seventy-five percent, at least about eighty-five percent, at least about ninety percent sequence, at least about ninety-five percent, at least about ninety-nine percent sequence identity with the entire p300-4 sequence set forth in SEQ ID NO: 6. As used herein, the terms "sequence identity" and "sequence homology" are synonymous. Identity or homology with respect to such sequences is defined herein as the percentage of amino acid residues in the candidate sequence that are identical with the known peptides, after alignment the sequences.

[0075] The polypeptide variants include peptides having a fragment of the amino acid sequence set forth in SEQ ID NO: 4, having at least about 10, 20, 30, 40, 50, 100, 200, 300, or 400 contiguous amino acid residues. The polypeptide variants further include those fragments associated with the activity of containing the p300 acetyltransferase binding domain or fragments thereof. Such fragments may contain functional regions of the p300 acetyltransferase binding as well as regions of pronounced hydrophilicity. Variants may also include a polypeptide containing the p300 acetyltransferase binding domain or fragments thereof linked to a second polypeptide, in any order, with intervening amino acids removed or replaced by a linker sequence.

S6K2 Protein Derivatives

[0076] The present invention provides isolated proteins, allelic variants of the proteins, and conservative amino acid substitutions of the protein comprising the amino acids sequence of SEQ ID NO: 8 which contain one or more mutations at amino acids corresponding to residues 487 to 495. As used herein, the "protein" or "polypeptide" refers, in part, to a protein that has the human amino acid sequence depicted in SEQ ID NO: 8 and contains any of the aforementioned mutations. The terms also refer to proteins that have a slightly different amino acid sequence than that specifically recited above.

[0077] The proteins of the present invention are preferably in isolated form. As used herein, a protein is said to be isolated when physical, mechanical, or chemical methods are employed to remove the protein from cellular constituents that are normally associated with the protein. A skilled artisan can readily employ standard purification methods to obtain an isolated protein.

[0078] The proteins of the present invention further include insertion, deletion, or conservative amino acid substitution variants of SEQ ID NO: 8 including those defined herein at amino acids corresponding to residues 487 to 495. As used herein, a conservative variant refers to alterations in the amino acid sequence that does not adversely affect the biological functions of the protein. Here conservative variants would retain their ability to be acetylated, the ability to phosphorylate second proteins and the ability to bind DNA. A substitution, insertion, or deletion is said to adversely affect the

protein when the altered sequence prevents or disrupts a biological function associated with the protein. For example, the overall charge, structure, or hydrophobic/hydrophilic properties of the protein can be altered without adversely affecting a biological activity. Accordingly, the amino acid sequence can be altered, for example to render the peptide more hydrophobic or hydrophilic, without adversely affecting the biological activities of the protein.

[0079] Ordinarily, the conservative substitution variants, and the members of the protein family, will have an amino acid sequence having at least about 75% amino acid sequence identity with the entire sequence set forth in SEQ ID NO: 8, more preferably at least about 80%, even more preferably at least about 90%, and most preferably at least about 95% sequence identity. Identity or homology with respect to such sequences is defined herein as the percentage of amino acid residues in the candidate sequence that are identical with the known peptides, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent homology, and not considering any conservative substitutions as part of the sequence identity. Fusion proteins, or N-terminal, C-terminal or internal extensions, deletions, or insertions into the peptide sequence shall not be construed as affecting homology.

[0080] Thus, the proteins of the present invention include molecules having the amino acid sequence disclosed in SEQ ID NO: 8 and amino acid sequence variants wherein one or more amino acid residues has been inserted N- or C-terminal to, or within, the disclosed coding sequence; and amino acid sequence variants of the disclosed sequence, or their fragments as defined above, that have been substituted by at least one residue. Such fragments, also referred to as peptides or polypeptides, may contain functional regions of the protein identified as regions of the amino acid sequence which correspond to known protein domains, such as the AT-hook domain, as well as regions of pronounced hydrophilicity. The regions are all easily identifiable by using commonly available protein sequence analysis software such as MacVector (Oxford Molecular).

[0081] The present invention further provides compositions comprising a protein or polypeptide of the invention and a diluent. Suitable diluents can be aqueous or non-aqueous solvents or a combination thereof, and can comprise additional components, for example water-soluble salts or glycerol, that contribute to the stability, solubility, activity, and/or storage of the protein or polypeptide.

[0082] As described below, members of the family of S6K2 proteins can be used: (1) to identify agents which modulate the level of expression or at least one activity of the protein, (2) to identify binding partners for the protein, and (3) as a therapeutic agent or target.

[0083] Contemplated variants of polypeptides containing p300 acetyltransferase binding domain and/or of S6K2 protein derivatives further include those containing predetermined mutations by, e.g., homologous recombination, site-directed or PCR mutagenesis, and the corresponding proteins of other animal species, including but not limited to rabbit, mouse, rat, porcine, bovine, ovine, equine and non-human primate species, and the alleles or other naturally occurring variants of the family of proteins; and derivatives wherein the protein has been covalently modified by substitution, chemical, enzymatic, or other appropriate means with a moiety other than a naturally occurring amino acid (for example a detectable moiety such as an enzyme or radioisotope).

Nucleic Acids Molecules

[0084] The present invention further provides nucleic acid molecules that encode the p300 acetyltransferase binding domain and derivative polypeptides of the invention. The present invention also provides nucleic acids that encode proteins comprising one or more mutations at any amino acid residue corresponding to amino acids 487 to 495 of SEQ ID NO: 8 and derivative polypeptides of the invention In one embodiment, these nucleic acids encode a p300 acetyltransferase binding domain polypeptide capable of binding to S6 kinase, including human S6 kinase. In another embodiment, these nucleic acids encode a p300 acetyltransferase binding domain polypeptide capable of acetylating a S6 kinase, including human S6 kinase. As used herein, “nucleic acid” is defined as RNA or DNA that encodes a protein or peptide as defined above, is complementary to a nucleic acid sequence encoding such peptides, hybridizes to a nucleic acid encoding a p300 acetyltransferase binding domain or to amino acids 487 to 495 of SEQ ID NO: 8 across the open reading frame under appropriate stringency conditions, or encodes a polypeptide that shares at least about 75% sequence identity, preferably at least about 80%, more preferably at least about 85%, and even more preferably at least about 90% or even 95% or more identity with the entire contiguous amino acid sequence of SEQ ID NO: 6 or 8. The “nucleic acid” of the invention further includes nucleic acid molecules that share at least 80%, preferably at least about 85%, and more preferably at least about 90% or 95% or more identity with the nucleotide sequence of SEQ ID NO: 5 or 7 particularly across the open reading frame. Specifically contemplated are genomic DNA, cDNA, mRNA and antisense molecules, as well as nucleic acids based on alternative backbones or including alternative bases whether derived from natural sources or synthesized. Such nucleic acids, however, are defined further as being novel and unobvious over any prior art nucleic acid including that which encodes, hybridizes under appropriate stringency conditions, or is complementary to nucleic acid encoding a protein according to the present invention.

[0085] “Stringent conditions” are those that (1) employ low ionic strength and high temperature for washing, for example, 0.015 M NaCl/0.0015 M sodium citrate/0.1% SDS at a temperature of at least 50° C., or (2) employ during hybridization a denaturing agent such as formamide, for example, 50% (vol/vol) formamide with 0.1% bovine serum albumin/0.1% Ficoll/0.1% polyvinylpyrrolidone/50 mM sodium phosphate buffer (pH 6.5) with 750 mM NaCl, 75 mM sodium citrate at 42° C. Another example is hybridization in 50% formamide, 5×SSC (0.75 M NaCl, 0.075 M sodium citrate), 50 mM sodium phosphate (pH 6.8), 0.1% sodium pyrophosphate, 5×Denhardt’s solution, sonicated salmon sperm DNA (50 μg/ml), 0.1% SDS, and 10% dextran sulfate at a temperature of at least 42° C., with washes at 42° C. in 0.2×SSC and 0.1% SDS. A skilled artisan can readily determine and vary the stringency conditions appropriately to obtain a clear and detectable hybridization signal. Preferred molecules are those that hybridize under the above conditions to the complement of SEQ ID NO: 5 or 7 and which encode a functional protein. Even more preferred hybridizing molecules are those that hybridize under the above conditions to the complement strand of the open reading frame of SEQ ID NO: 5 or 7.

[0086] As used herein, a nucleic acid molecule is said to be "isolated" when the nucleic acid molecule is substantially separated from contaminant nucleic acid molecules encoding other polypeptides.

[0087] The present invention further provides fragments of the encoding nucleic acid molecule. As used herein, a fragment of an encoding nucleic acid molecule refers to a small portion of the entire protein coding sequence. The size of the fragment will be determined by the intended use. For example, if the fragment is chosen so as to encode an active portion of the protein, the fragment will need to be large enough to encode the functional regions of the protein. For instance, fragments which encode peptides corresponding to predicted antigenic regions may be prepared. If the fragment is to be used as a nucleic acid probe or PCR primer, then the fragment length is chosen so as to obtain a relatively small number of false positives during probing/priming.

[0088] Fragments of the encoding nucleic acid molecules of the present invention (i.e., synthetic oligonucleotides) that are used as probes or specific primers for the polymerase chain reaction (PCR), or to synthesize gene sequences encoding proteins of the invention, can easily be synthesized by chemical techniques, for example, the phosphotriester method of Matteucci et al. (1981) J. Am. Chem. Soc. 103, 3185-3191 or using automated synthesis methods. In addition, larger DNA segments can readily be prepared by well known methods, such as synthesis of a group of oligonucleotides that define various modular segments of the gene, followed by ligation of oligonucleotides to build the complete modified gene. In a preferred embodiment, the nucleic acid molecule of the present invention contains a contiguous open reading frame of at least about three-thousand and forty-five nucleotides.

[0089] The encoding nucleic acid molecules of the present invention may further be modified so as to contain a detectable label for diagnostic and probe purposes. A variety of such labels are known in the art and can readily be employed with the encoding molecules herein described. Suitable labels include, but are not limited to, biotin, radiolabeled nucleotides and the like. A skilled artisan can readily employ any such label to obtain labeled variants of the nucleic acid molecules of the invention. Modifications to the primary structure itself by deletion, addition, or alteration of the amino acids incorporated into the protein sequence during translation can be made without destroying the activity of the protein. Such substitutions or other alterations result in proteins having an amino acid sequence encoded by a nucleic acid falling within the contemplated scope of the present invention.

[0090] As used herein, a "nucleic acid" refers to the polymeric form of ribonucleotide or deoxyribonucleotides (adenine, guanine, thymine, and/or cytosine) in either its single stranded form, or in double-stranded helix. This term refers only to the primary and secondary structure of the molecule and is not limited to any particular tertiary form. Thus, this term includes single-stranded RNA or DNA, double-stranded DNA found in linear DNA molecules (e.g., restriction fragments), viruses, plasmids, and chromosomes. Sequences of double-stranded DNA molecules, sequences may be described herein according to the normal convention of giving only the sequence in the 5' to 3' direction along the non-transcribed strand of DNA (e.g., the strand having a sequence homologous to the mRNA).

[0091] A nucleic acid "coding sequence" is a double-stranded DNA sequence which is transcribed and translated into a polypeptide in vivo when placed under the control of appropriate regulatory sequences. The boundaries of the coding sequence are determined by a start codon at the 5' (amino) terminus and a translation stop codon at the 3' (carboxy) terminus. A polyadenylation signal and transcription termination sequence will usually be located 3' to the coding sequence.

[0092] Transcriptional and translational control sequences are DNA regulatory sequences, such as promoters, enhancers, polyadenylation signals, terminators, and the like, that provide for the expression of a coding sequence in a host cell.

[0093] As used herein, a "promoter sequence" is a DNA regulatory region capable of binding RNA polymerase in a cell and initiating transcription of a downstream (3' direction) coding sequence. For purposes of defining the present invention, the promoter sequence is bounded (inclusively) at its 3' terminus by the transcription initiation site and extends upstream (5' direction) to include the minimum number of bases or elements necessary to initiate transcription at levels detectable above background. Within the promoter sequence will be found a transcription initiation site, as well as protein binding domains responsible for the binding of RNA polymerase. Eukaryotic promoters will often, but not always, contain "TATA" boxes and "CAT" boxes.

[0094] A coding sequence is "under the control" of transcriptional and translational control sequences in a cell when RNA polymerase transcribes the coding sequence into mRNA, which is then translated into the protein encoded by the coding sequence.

[0095] A "signal sequence" can be included before the coding sequence. This sequence encodes a signal peptide, N-terminal to the polypeptide that directs the polypeptide to the cell surface or instructs the host cell to secrete the polypeptide to the extracellular space. This signal peptide is clipped off by the host cell before the protein leaves the cell. Signal sequences can be found associated with a variety of proteins native to prokaryotes and eukaryotes. For instance, alpha-factor, a native yeast protein, is secreted from yeast, and its signal sequence can be attached to heterologous proteins to be secreted into the media (see U.S. Pat. No. 4,546,082). Further, the alpha-factor and its analogs have been found to secrete heterologous proteins from a variety of yeast, such as *Saccharomyces* and *Kluyveromyces* (EP 88312306.9, EP 0324274, and EP 0301669). An example for use in mammalian cells is the tPA signal used for expressing Factor VIIc light chain.

[0096] A cell has been "transformed" by a exogenous or heterologous nucleic acid when such nucleic acid as been introduced inside the cell. The transforming nucleic acid may or may not be integrated (covalently linked) into chromosomal DNA making up the genome of the cell. In prokaryotes, for example, the transforming nucleic acid may be maintained on an episomal element such as a plasmid or viral vector. With respect to eukaryotic cells, a stably transformed cell is one in which the transforming DNA has become integrated into a chromosome so that it is inherited by daughter cells through chromosome replication. This stability is demonstrated by the ability of the eukaryotic cell to establish cell lines or clones comprised of a population of daughter cells containing the transforming nucleic acid.

[0097] As used herein, a "cell line" is a clone of a primary cell that is capable of stable growth in vitro for many generations. As used herein, nucleic acid sequences display "substantial identity" when at least about 85 percent (preferably at

least about 90 percent and most preferably at least about 95 percent) of the nucleotides match over the defined length of the nucleotide sequences. Sequences that are substantially identical can be identified in a Southern hybridization experiment under, for example, stringent conditions as defined for that particular system. Defining appropriate hybridization conditions is within the skill of the art.

[0098] A "heterologous" region of the nucleic acid construct is an identifiable segment of a nucleic acid within a larger nucleic acid molecule that is not found in association with the larger molecule in nature. Thus, when the heterologous region encodes a mammalian gene, the gene will usually be flanked by DNA that does not flank the mammalian genomic DNA in the genome of the source organism. Another example of a heterologous coding sequence is a construct where the coding sequence itself is not found in nature (e.g., a cDNA where the genomic coding sequence contains introns, or synthetic sequences having codons different than the native gene).

[0099] Vectors are used to simplify manipulation of the nucleic acids, which encode proteins of the invention, either for preparation of large quantities of nucleic acids for further processing (cloning vectors) or for expression of the polypeptides (expression vectors). Such proteins include the p300-4 fragment, proteins comprising one or more mutations at any amino acid residue corresponding to amino acids 487 to 495 of SEQ ID NO: 8 and derivative polypeptides, thereof. Vectors comprise plasmids, viruses (including phage), and integrated DNA fragments (i.e., fragments that are integrated into the host genome by recombination). Cloning vectors need not contain expression control sequences. However, control sequences in an expression vector include transcriptional and translational control sequences such as a transcriptional promoter, a sequence encoding suitable ribosome binding sites, and sequences, which control termination of transcription and translation. In one embodiment, the expression vector should preferably include a selection gene to facilitate the stable expression of the p300-4 fragment gene and/or to identify transformed cells. In another embodiment, the expression vector should preferably include a selection gene to facilitate the stable expression of the S6K2 fragment gene and/or to identify transformed cells. However, the selection gene for maintaining expression can be supplied by a separate vector in co-transformation systems using eukaryotic host cells.

[0100] Suitable vectors generally will contain a replicon (origins of replication, for use in non-integrative vectors) and control sequences which are derived from species compatible with the intended expression host. By the term "replicable" vector as used herein, it is intended to encompass vectors containing such replicons as well as vectors which are replicated by integration into the host genome. Transformed host cells are cells which have been transformed or transfected either with vectors containing p300-4 fragment or p300-4 fragment derivative polypeptide encoding nucleic acid or with vectors containing S6K2 fragment or S6K2 fragment derivative polypeptide encoding nucleic acid. The expressed polypeptides may be secreted into the culture supernatant, under the control of suitable processing signals in the expressed peptide (e.g. homologous or heterologous signal sequences).

[0101] Expression vectors for host cells ordinarily include an origin of replication, a promoter located upstream from the p300 acetyltransferase binding domain polypeptide coding sequence or the S6K2 coding sequence, together with a ribosome binding site, a polyadenylation site, and a transcriptional termination sequence. Those of ordinary skill will appreciate that certain of these sequences are not required for expression in certain hosts. An expression vector for use with microbes need only contain an origin of replication recognized by the host, a promoter which will function in the host, and a selection gene.

[0102] Commonly used promoters are derived from polyoma, bovine papilloma virus, CMV (cytomegalovirus, either murine or human), Rouse sarcoma virus, adenovirus, and simian virus 40 (SV40). Other control sequences (e.g., terminator, polyA, enhancer, or amplification sequences) can also be used.

[0103] An expression vector is constructed so that the p300-4 fragment or S6K2 protein or derivative polypeptide coding sequence is located in the vector with the appropriate regulatory sequences. The positioning and orientation of the coding sequence with respect to the control sequences is such that the coding sequence is transcribed and translated under the "control" of the control sequences i.e., RNA polymerase which binds to the DNA molecule at the control sequences transcribes the coding sequence. The control sequences may be ligated to the coding sequence prior to insertion into a vector, such as the cloning vectors described above. Alternatively, the coding sequence can be cloned directly into an expression vector which already contains the control sequences and an appropriate restriction site. If the selected host cell is a mammalian cell, the control sequences can be heterologous or homologous to the p300-4 fragment, S6K2 protein or derivative polypeptide coding sequence, and the coding sequence can either be genomic DNA containing introns or cDNA.

[0104] Higher eukaryotic cell cultures may be used to express the proteins of the present invention, whether from vertebrate or invertebrate cells, including insects, and the procedures of propagation thereof are known.

[0105] Other expression vectors are those for use in eukaryotic systems. An exemplary eukaryotic expression system is that employing vaccinia virus, which is well-known in the art (see, for example, WO 86/07593). Yeast expression vectors are known in the art (see, for example, U.S. Pat. Nos. 4,446,235 and 4,430,428). Another expression system is vector pHSI, which transforms Chinese hamster ovary cells (see WO 87/02062). Mammalian tissue may be cotransformed with DNA encoding a selectable marker such as dihydrofolate reductase (DHFR) or thymidine kinase and DNA encoding the p300-4 fragment or p300-4 fragment derivative polypeptide or S6K2 protein or S6K2 derivative peptide. If wild type DHFR gene is employed, it is preferable to select a host cell which is deficient in DHFR, thus permitting the use of the DHFR coding sequence as marker for successful transfection in hgt medium, which lacks hypoxanthine, glycine, and thymidine.

[0106] Depending on the expression system and host selected, the peptide of interest, such as p300 acetyltransferase binding domain polypeptides and S6K2 protein, are produced by growing host cells transformed by an exogenous or heterologous DNA construct, such as an expression vector described above under conditions whereby the polypeptide is expressed. The peptide of interest, such as p300-4 fragment, S6K2 protein or derivative polypeptides thereof, is then isolated from the host cells and purified. If the expression system secretes the protein or peptide into the growth media, the protein can be purified directly from cell-free media. The

selection of the appropriate growth conditions and initial crude recovery methods are within the skill of the art.

[0107] Once a coding sequence for a p300 acetyltransferase binding domain polypeptide or S6K2 protein of the invention has been prepared or isolated, it can be cloned into any suitable vector and thereby maintained in a composition of cells which is substantially free of cells that do not contain any p300-4 fragment or S6K2 protein coding sequence, respectively. As described above, numerous cloning vectors are known to those of skill in the art.

Binding Domain Peptide Mimetics

[0108] The present invention includes another class of p300 acetyltransferase binding domain derivatives called peptide mimetics that mimic the three-dimensional structure of p300-4 fragment. The present invention also includes peptide mimetics that mimic the three-dimensional structure of S6K2. Such peptide mimetics may have significant advantages over naturally occurring peptides including, for example, more economical production, greater chemical stability, enhanced pharmacological properties (half-life, absorption, potency, efficacy, etc.), altered specificity (e.g., a broad-spectrum of biological activities), reduced antigenicity and others.

[0109] In one form, mimetics are peptide-containing molecules that mimic elements of p300 acetyltransferase binding domain secondary structure. In another embodiment, the mimetics are capable of binding to S6 kinase, including human S6 kinase. In a further embodiment, the mimetics are capable of modulating the activity of the S6 kinase by effecting acetylation at the binding domain. In one embodiment, the mimetics facilitate acetylation while in another embodiment, they inhibit acetylation by another acetyltransferase (e.g., p300).

[0110] In an alternate form, mimetics are peptide-containing molecules that mimic elements of S6K2 AT-hook motif secondary structure. In yet another embodiment, the mimetics are capable of binding to S6K2, including human S6K2. In an alternate embodiment, the mimetics are capable of modulating the activity of the S6K2 by effecting DNA binding.

[0111] The underlying rationale behind the use of peptide mimetics is that the peptide backbone of proteins exists chiefly to orient amino acid side chains in such a way as to facilitate molecular interactions, similar to those of antibody and antigen. A peptide mimetic is expected to permit molecular interactions similar to the natural molecule. In another form, peptide analogs are commonly produced in the pharmaceutical industry as non-peptide drugs with properties analogous to those of the template peptide. These types of non-peptide compounds are also referred to as peptide mimetics or peptidomimetics (Fauchere (1986) Adv. Drug Res. 15, 29-69; Veber & Freidinger (1985) Trends Neurosci. 8, 392-396; Evans et al. (1987) J. Med. Chem. 30, 1229-1239 which are incorporated herein by reference) and are usually developed with the aid of computerized molecular modeling.

[0112] Peptide mimetics that are structurally similar to therapeutically useful peptides may be used to produce an equivalent therapeutic or prophylactic effect. Generally, peptide mimetics are structurally similar to a paradigm polypeptide (i.e., a polypeptide that has a biochemical property or pharmacological activity), but have one or more peptide linkages optionally replaced by a chemical linkage not normally found in peptides by methods known in the art. Labeling of peptide mimetics usually involves covalent attachment of one or more labels, directly or through a spacer (e.g., an amide group), to non-interfering positions on the peptide mimetic that are predicted by quantitative structure-activity data and molecular modeling. Such non-interfering positions generally are positions that do not form direct contacts with the macromolecules to which the peptide mimetic binds to produce the therapeutic effect. Derivitization (e.g., labeling) of peptide mimetics should not substantially interfere with the desired biological or pharmacological activity of the peptide mimetic.

[0113] The use of peptide mimetics can be enhanced through the use of combinatorial chemistry to create drug libraries. The design of peptide mimetics can be aided by identifying amino acid mutations that increase or decrease binding of a peptide to, for instance, a tumor cell. Approaches that can be used include the yeast two hybrid method (see Chien et al. (1991) Proc. Natl. Acad. Sci. USA 88, 9578-9582) and using the phage display method. The two hybrid method detects protein-protein interactions in yeast (Fields et al. (1989) Nature 340, 245-246). The phage display method detects the interaction between an immobilized protein and a protein that is expressed on the surface of phages such as lambda and M13 (Amberg et al. (1993) Strategies 6, 2-4; Hogrefe et al. (1993) Gene 128, 119-126). These methods allow positive and negative selection for peptide-protein interactions and the identification of the sequences that determine these interactions.

Recombinant DNA Molecules Containing a Nucleic Acid Molecule

[0114] The present invention further provides recombinant DNA molecules (rDNA) that contain a coding sequence. As used herein, a rDNA molecule is a DNA molecule that has been subjected to molecular manipulation in situ. Methods for generating rDNA molecules are well known in the art, for example, see Sambrook et al. (2001) Molecular Cloning—A Laboratory Manual, Cold Spring Harbor Laboratory Press. In the preferred rDNA molecules, a coding DNA sequence is operably linked to expression control sequences and/or vector sequences.

[0115] The choice of vector and/or expression control sequences to which one of the protein family encoding sequences of the present invention is operably linked depends directly, as is well known in the art, on the functional properties desired, e.g., protein expression, and the host cell to be transformed. A vector contemplated by the present invention is at least capable of directing the replication or insertion into the host chromosome, and preferably, also expression, of the structural gene included in the rDNA molecule.

[0116] Expression control elements that are used for regulating the expression of an operably linked protein encoding sequence are known in the art and include, but are not limited to, inducible promoters, constitutive promoters, secretion signals, and other regulatory elements. Preferably, the inducible promoter is readily controlled, such as being responsive to a nutrient in the host cell's medium.

[0117] In one embodiment, the vector containing a coding nucleic acid molecule will include a prokaryotic replicon, i.e., a DNA sequence having the ability to direct autonomous replication and maintenance of the recombinant DNA molecule extrachromosomally in a prokaryotic host cell, such as a bacterial host cell, transformed therewith. Such replicons are well known in the art. In addition, vectors that include a prokaryotic replicon may also include a gene whose expres-

sion confers a detectable marker such as a drug resistance. Typical bacterial drug resistance genes are those that confer resistance to ampicillin or tetracycline.

[0118] Vectors that include a prokaryotic replicon can further include a prokaryotic or bacteriophage promoter capable of directing the expression (transcription and translation) of the coding gene sequences in a bacterial host cell, such as *E. coli*. A promoter is an expression control element formed by a DNA sequence that permits binding of RNA polymerase and transcription to occur. Promoter sequences compatible with bacterial hosts are typically provided in plasmid vectors containing convenient restriction sites for insertion of a DNA segment of the present invention. Typical of such vector plasmids are pUC8, pUC9, pBR322 and pBR329 (BioRad), pPL and pKK223 (Pharmacia).

[0119] Expression vectors compatible with eukaryotic cells, preferably those compatible with vertebrate cells, can also be used to form rDNA molecules that contain a coding sequence. Eukaryotic cell expression vectors, including viral vectors, are well known in the art and are available from several commercial sources. Typically, such vectors are provided containing convenient restriction sites for insertion of the desired DNA segment. Typical of such vectors are pSVL and pKSV-10 (Pharmacia), pBPV-1/pML2d (International Biotechnologies, Inc.), pTDT1 (ATCC), the vector pCDM8 described herein, and the like eukaryotic expression vectors.

[0120] Eukaryotic cell expression vectors used to construct the rDNA molecules of the present invention may further include a selectable marker that is effective in an eukaryotic cell, preferably a drug resistance selection marker. A preferred drug resistance marker is the gene whose expression results in neomycin resistance, i.e., the neomycin phosphotransferase (neo) gene. (Southern et al. (1982) J. Mol. Anal. Genet. 1, 327-341). Alternatively, the selectable marker can be present on a separate plasmid, and the two vectors are introduced by co-transfection of the host cell, and selected by culturing in the appropriate drug for the selectable marker. The present invention further provides host cells transformed with a nucleic acid molecule that encodes a protein of the present invention. The host cell can be either prokaryotic or eukaryotic. Eukaryotic cells useful for expression of a protein of the invention are not limited, so long as the cell line is compatible with cell culture methods and compatible with the propagation of the expression vector and expression of the gene product. Preferred eukaryotic host cells include, but are not limited to, yeast, insect and mammalian cells, preferably vertebrate cells such as those from a mouse, rat, monkey or human cell line. Preferred eukaryotic host cells include Chinese hamster ovary (CHO) cells available from the ATCC as CCL61, NIH Swiss mouse embryo cells (NIH-3T3) available from the ATCC as CRL 1658, baby hamster kidney cells (BHK), and the like eukaryotic tissue culture cell lines.

[0121] Any prokaryotic host can be used to express a rDNA molecule encoding a protein of the invention. The preferred prokaryotic host is *E. coli*.

[0122] Transformation of appropriate cell hosts with a rDNA molecule of the present invention is accomplished by well known methods that typically depend on the type of vector used and host system employed. With regard to transformation of prokaryotic host cells, electroporation and salt treatment methods are typically employed, see, for example, Cohen et al. (1972) Proc. Natl. Acad. Sci. USA 69, 2110; and Sambrook et al. (1989) Molecular —A Laboratory Manual, Cold Spring Harbor Laboratory Press. With regard to transformation of vertebrate cells with vectors containing rDNAs, electroporation, cationic lipid or salt treatment methods are typically employed, see, for example, Graham et al. (1973) Virol. 52, 456; Wigler et al. (1979) Proc. Natl. Acad. Sci. USA 76, 1373-1376.

[0123] Successfully transformed cells, i.e., cells that contain a rDNA molecule of the present invention, can be identified by well known techniques including the selection for a selectable marker. For example, cells resulting from the introduction of an rDNA of the present invention can be cloned to produce single colonies. Cells from those colonies can be harvested, lysed and their DNA content examined for the presence of the rDNA using a method such as that described by Southern (1975) J. Mol. Biol. 98, 503-504 or Berent et al. (1985) Biotech. 3, 208-209 or the proteins produced from the cell assayed via an immunological method.

Production of Recombinant Proteins Using a rDNA Molecule

[0124] The present invention further provides methods for producing a protein of the invention using nucleic acid molecules herein described. In general terms, the production of a recombinant form of a protein typically involves the following steps:

[0125] A nucleic acid molecule is first obtained that encodes a protein of the invention, such as a nucleic acid molecule comprising, consisting essentially of or consisting of the nucleic acid sequence encoding SEQ ID NO: 8 or derivatives thereof.

[0126] The nucleic acid molecule is then preferably placed in operable linkage with suitable control sequences, as described above, to form an expression unit containing the protein open reading frame. The expression unit is used to transform a suitable host and the transformed host is cultured under conditions that allow the production of the recombinant protein. Optionally the recombinant protein is isolated from the medium or from the cells; recovery and purification of the protein may not be necessary in some instances where some impurities may be tolerated.

[0127] Each of the foregoing steps can be done in a variety of ways. For example, the desired coding sequences may be obtained from genomic fragments and used directly in appropriate hosts. The construction of expression vectors that are operable in a variety of hosts is accomplished using appropriate replicons and control sequences, as set forth above. The control sequences, expression vectors, and transformation methods are dependent on the type of host cell used to express the gene and were discussed in detail earlier. Suitable restriction sites can, if not normally available, be added to the ends of the coding sequence so as to provide an excisable gene to insert into these vectors. A skilled artisan can readily adapt any host/expression system known in the art for use with the nucleic acid molecules of the invention to produce recombinant protein.

Methods to Identify Binding Partners

[0128] Another embodiment of the present invention provides methods for use in isolating and identifying binding partners of proteins of the invention. In general, a protein of the invention is mixed with a potential binding partner or an extract or fraction of a cell under conditions that allow the association of potential binding partners with the protein of the invention. After mixing, peptides, polypeptides, proteins or other molecules that have become associated with a protein of the invention are separated from the mixture. The binding partner that bound to the protein of the invention can then be

removed and further analyzed. To identify and isolate a binding partner, the entire protein, for instance a protein comprising the entire amino acid sequence of SEQ ID NO: 8 can be used. Alternatively, a fragment of the protein can be used.

[0129] As used herein, a cellular extract refers to a preparation or fraction that is made from a lysed or disrupted cell. The preferred source of cellular extracts will be cells derived from human skin tissue or the human respiratory tract or cells derived from a biopsy sample of human lung tissue in patients with allergic hypersensitivity. Alternatively, cellular extracts may be prepared from normal tissue or available cell lines, particularly granulocytic cell lines.

[0130] A variety of methods can be used to obtain an extract of a cell. Cells can be disrupted using either physical or chemical disruption methods. Examples of physical disruption methods include, but are not limited to, sonication and mechanical shearing. Examples of chemical lysis methods include, but are not limited to, detergent lysis and enzyme lysis. A skilled artisan can readily adapt methods for preparing cellular extracts in order to obtain extracts for use in the present methods.

[0131] Once an extract of a cell is prepared, the extract is mixed with the protein of the invention under conditions in which association of the protein with the binding partner can occur. A variety of conditions can be used, the most preferred being conditions that closely resemble conditions found in the cytoplasm of a human cell. Features such as osmolarity, pH, temperature, and the concentration of cellular extract used, can be varied to optimize the association of the protein with the binding partner.

[0132] After mixing under appropriate conditions, the bound complex is separated from the mixture. A variety of techniques can be utilized to separate the mixture. For example, antibodies specific to a protein of the invention can be used to immunoprecipitate the binding partner complex. Alternatively, standard chemical separation techniques such as chromatography and density/sediment centrifugation can be used.

[0133] After removal of non-associated cellular constituents found in the extract, the binding partner can be dissociated from the complex using conventional methods. For example, dissociation can be accomplished by altering the salt concentration or pH of the mixture. To aid in separating associated binding partner pairs from the mixed extract, the protein of the invention can be immobilized on a solid support. For example, the protein can be attached to a nitrocellulose matrix or acrylic beads. Attachment of the protein to a solid support aids in separating peptide/binding partner pairs from other constituents found in the extract. The identified binding partners can be either a single protein or a complex made up of two or more proteins. Alternatively, binding partners may be identified using a Far-Western assay according to the procedures of Takayama et al. (1997) Methods Mol. Biol. 69, 171-184 or Sauder et al. (1996) J. Gen. Virol. 77, 991-996 or identified through the use of epitope tagged proteins or GST fusion proteins.

[0134] Alternatively, the nucleic acid molecules of the invention can be used in a yeast two-hybrid system. The yeast two-hybrid system has been used to identify other protein partner pairs and can readily be adapted to employ the nucleic acid molecules herein described.

Methods to Identify Agents that Modulate Activity

[0135] The present invention provides methods for identifying agents that modulate at least one activity of a protein of SEQ ID NO: 8. Such methods or assays may utilize any means of monitoring or detecting the desired activity. These assays may include detection of the acetylation of S6K2 or the monitoring of the ability of S6K2 to phosphorylate a second protein.

[0136] In one format, the specific activity of a protein of the invention, normalized to a standard unit, between a cell population that has been exposed to the agent to be tested compared to an un-exposed control cell population may be assayed. Cell lines or populations are exposed to the agent to be tested under appropriate conditions and time. Cellular lysates may be prepared from the exposed cell line or population and a control, unexposed cell line or population. The cellular lysates are then analyzed with the probe.

[0137] Antibody probes can be prepared by immunizing suitable mammalian hosts utilizing appropriate immunization protocols using the proteins of the invention or antigen-containing fragments thereof. To enhance immunogenicity, these proteins or fragments can be conjugated to suitable carriers. Methods for preparing immunogenic conjugates with carriers such as BSA, KLH or other carrier proteins are well known in the art. In some circumstances, direct conjugation using, for example, carbodiimide reagents may be effective; in other instances linking reagents such as those supplied by Pierce Chemical Co. may be desirable to provide accessibility to the hapten. The hapten peptides can be extended at either the amino or carboxy terminus with a cysteine residue or interspersed with cysteine residues, for example, to facilitate linking to a carrier. Administration of the immunogens is conducted generally by injection over a suitable time period and with use of suitable adjuvants, as is generally understood in the art. During the immunization schedule, titers of antibodies are taken to determine adequacy of antibody formation.

[0138] While the polyclonal antisera produced in this way may be satisfactory for some applications, for pharmaceutical compositions, use of monoclonal preparations is preferred. Immortalized cell lines which secrete the desired monoclonal antibodies may be prepared using standard methods, see e.g., Kohler & Milstein (1992) Biotechnology 24, 524-526 or modifications which effect immortalization of lymphocytes or spleen cells, as is generally known. The immortalized cell lines secreting the desired antibodies can be screened by immunoassay in which the antigen is the peptide hapten, polypeptide or protein. When the appropriate immortalized cell culture secreting the desired antibody is identified, the cells can be cultured either in vitro or by production in ascites fluid.

[0139] The desired monoclonal antibodies may be recovered from the culture supernatant or from the ascites supernatant. Fragments of the monoclonal antibodies or the polyclonal antisera that contain the immunologically significant portion can be used as antagonists, as well as the intact antibodies. Use of immunologically reactive fragments, such as Fab or Fab' fragments, is often preferable, especially in a therapeutic context, as these fragments are generally less immunogenic than the whole immunoglobulin.

[0140] The antibodies or fragments may also be produced, using current technology, by recombinant means. Antibody regions that bind specifically to the desired regions of the protein can also be produced in the context of chimeras with multiple species origin.

[0141] Antibody regions that bind specifically to the desired regions of the protein can also be produced in the

context of chimeras with multiple species origin, for instance, humanized antibodies. The antibody can therefore be a humanized antibody or human a antibody, as described in U.S. Pat. No. 5,585,089 or Riechmann et al. (1988) Nature 332, 323-327.

[0142] Agents that are assayed in the above method can be randomly selected or rationally selected or designed. As used herein, an agent is said to be randomly selected when the agent is chosen randomly without considering the specific sequences involved in the association of the a protein of the invention alone or with its associated substrates, binding partners, etc. An example of randomly selected agents is the use a chemical library or a peptide combinatorial library, or a growth broth of an organism.

[0143] As used herein, an agent is said to be rationally selected or designed when the agent is chosen on a non-random basis which takes into account the sequence of the target site or its conformation in connection with the agent's action. Agents can be rationally selected or rationally designed by utilizing the peptide sequences that make up these sites. For example, a rationally selected peptide agent can be a peptide whose amino acid sequence is identical to or a derivative of any functional consensus site.

[0144] The agents of the present invention can be, as examples, peptides, peptide mimetics, antibodies, antibody fragments, small molecules, vitamin derivatives, as well as carbohydrates. Peptide agents of the invention can be prepared using standard solid phase (or solution phase) peptide synthesis methods, as is known in the art. In addition, the DNA encoding these peptides may be synthesized using commercially available oligonucleotide synthesis instrumentation and produced recombinantly using standard recombinant production systems. The production using solid phase peptide synthesis is necessitated if non-gene-encoded amino acids are to be included.

[0145] Another class of agents of the present invention are antibodies or fragments thereof that bind to a protein of SEQ ID NO: 8. Antibody agents can be obtained by immunization of suitable mammalian subjects with peptides, containing as antigenic regions, those portions of the protein intended to be targeted by the antibodies.

Methods to Identify Agents that Modulate the Expression

[0146] Another embodiment of the present invention provides methods for identifying agents that modulate the expression of a nucleic acid encoding an 56 kinase protein such as a protein having the amino acid sequence of SEQ ID NO: 8. Such assays may utilize any available means of monitoring for changes in the expression level of the nucleic acids of the invention. As used herein, an agent is said to modulate the expression of a nucleic acid of the invention, for instance a nucleic acid encoding the protein having the sequence of SEQ ID NO: 8, if it is capable of up- or down-regulating expression of the nucleic acid in a cell.

[0147] In one assay format, cell lines that contain reporter gene fusions between the open reading frame of the S6K2 (SEQ ID NO: 7) and any assayable fusion partner may be prepared. Numerous assayable fusion partners are known and readily available including the firefly luciferase gene and the gene encoding chloramphenicol acetyltransferase (Alam et al., 1990 Anal. Biochem. 188, 245-254). Cell lines containing the reporter gene fusions are then exposed to the agent or a polypeptide comprising a p300 acetyltransferase binding domain (i.e., control) to be tested under appropriate conditions and time. Differential expression of the reporter gene between samples exposed to the agent and control samples identifies agents which modulate the expression of a nucleic acid encoding the S6 kinase protein.

[0148] Additional assay formats may be used to monitor the ability of the agent to modulate the expression of a nucleic acid encoding a protein of the invention such as the S6 kinase protein. For instance, mRNA expression may be monitored directly by hybridization to the nucleic acids of the invention. Cell lines are exposed to the agent and/or a polypeptide comprising a p300 acetyltransferase binding domain to be tested under appropriate conditions and time and total RNA or mRNA is isolated by standard procedures such those disclosed in Sambrook et al. (2001) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory Press).

[0149] Probes to detect differences in RNA expression levels between cells exposed to the agent and control cells may be prepared from the nucleic acids of the invention. It is preferable, but not necessary, to design probes which hybridize only with target nucleic acids under conditions of high stringency. Only highly complementary nucleic acid hybrids form under conditions of high stringency. Accordingly, the stringency of the assay conditions determines the amount of complementarity which should exist between two nucleic acid strands in order to form a hybrid. Stringency should be chosen to maximize the difference in stability between the probe:target hybrid and potential probe:non-target hybrids.

[0150] Probes may be designed from the nucleic acids of the invention through methods known in the art. For instance, the G+C content of the probe and the probe length can affect probe binding to its target sequence. Methods to optimize probe specificity are commonly available in Sambrook et al. (2001) or Ausubel et al. (1995) Current Protocols in Molecular Biology, Greene Publishing Company.

[0151] Hybridization conditions are modified using known methods, such as those described by Sambrook et al. (2001) and Ausubel et al. (1995) as required for each probe. Hybridization of total cellular RNA or RNA enriched for polyA RNA can be accomplished in any available format. For instance, total cellular RNA or RNA enriched for polyA RNA can be affixed to a solid support and the solid support exposed to at least one probe comprising at least one, or part of one of the sequences of the invention under conditions in which the probe will specifically hybridize. Alternatively, nucleic acid fragments comprising at least one, or part of one of the sequences of the invention can be affixed to a solid support, such as a porous glass wafer. The glass wafer can then be exposed to total cellular RNA or polyA RNA from a sample under conditions in which the affixed sequences will specifically hybridize. Such glass wafers and hybridization methods are widely available, for example, those disclosed by Beattie (WO 95/11755). By examining for the ability of a given probe to specifically hybridize to an RNA sample from an untreated cell population and from a cell population exposed to the agent, agents which up or down regulate the expression of a nucleic acid encoding the S6 kinase protein.

Pharmaceutical Compositions

[0152] This invention includes a pharmaceutical composition for the treatment of abnormal cell growth or for a disease characterized by cell proliferation in a mammal, including a human, comprising, consisting essentially of, or consisting of an amount of polypeptide comprising p300 acetyltransferase binding domain or derivatives thereof that is effective in inhibiting an S6 kinase protein mediated activity and a phar-

maceutically acceptable carrier. In one embodiment, the S6 kinase protein mediated activity is abnormal cell growth. As used herein, the terms "abnormal cell growth" and "disease characterized by cell proliferation" unless otherwise indicated, refer to cell growth or proliferation that is independent of normal regulatory mechanisms (e.g., loss of contact inhibition). This includes the abnormal growth and/or proliferation of cells in both benign and malignant cells of neoplastic diseases. Inhibition of abnormal cell growth can occur by a variety of mechanisms including, but not limited to, cell death, apoplosis, arrest of mitosis, inhibition of cell division, transcription, translation, transduction, etc.

[0153] As used herein, an "effective amount" of p300 acetyltransferase binding domain fragment is an amount which exerts an inhibitory effect on S6 kinase protein activity, including phosphorylation of a substrate protein. An effective amount is also an amount which is effective to inhibit the proliferation of cells exhibiting abnormal growth, such as cancer cells. As used herein, compositions comprising, treatment with or administration of a polypeptide containing a "p300 acetyltransferase binding domain" includes to the same degree treatment with analogues, derivatives, fragments, variants, related peptides and mimetics disclosed herein.

[0154] Pharmaceutical compositions comprising the active agents (e.g., polypeptides, small molecules, etc.) of the present invention can be administered via parenteral, subcutaneous, intravenous, intramuscular, intraperitoneal, intrathecal, intracranial or transdermal or buccal routes. For example, an agent may be administered locally to a tumor via microinfusion. Alternatively, or concurrently, administration may be by the oral route. The dosage administered will be dependent upon the age, health, and weight of the recipient, kind of concurrent treatment, if any, frequency of treatment, and the nature of the effect desired.

[0155] While individual needs vary, determination of optimal ranges of effective amounts of each component is within the skill of the art. Dosages of p300-4 fragment and/or derivatives thereof of the present invention typically comprise about 1.0 ng/kg body weight to about 0.13 mg/kg body weight. In one embodiment, dosages of p300-4 fragment and/or derivatives thereof comprise about 1.0 ng/kg body weight to about 0.1 mg/kg body weight. In a preferred embodiment, dosages for systemic administration comprise about 0.01 μg/kg body weight to about 0.1 mg/kg body weight. In another embodiment, the dosage of p300-4 fragment and/or derivatives thereof comprises less than about 0.1 mg/kg body weight. More preferred dosages for systemic administration comprise about 0.1 μg/kg body weight to about 0.05 mg/kg body weight. In another preferred embodiment, the dosage of p300-4 fragment and/or derivatives thereof comprises less than about 0.05 mg/kg body weight. The most preferred dosages for systemic administration comprise between about 1.0 μg/kg body weight to about 0.01 mg/kg body weight. In other embodiments, the amount of p300-4 fragment and/or derivatives thereof administered is an amount effective to bring the concentration of p300-4 fragment and/or derivatives thereof in the serum to a concentration of about 20.0, 10.0, 5.0, 2.50, 1.25, 0.625, 0.3125, 0.156, 0.078, 0.039, 0.020, 0.010, 0.005, 0.003, 0.0015, 0.0008, 0.0003, or 0.0001 nM.

[0156] Dosages of S6K2/or derivatives thereof of the present invention typically comprise about 1.0 ng/kg body weight to about 0.13 mg/kg body weight. In one embodiment, dosages of S6K2 and/or derivatives thereof comprise about 1.0 ng/kg body weight to about 0.1 mg/kg body weight. In a preferred embodiment, dosages for systemic administration comprise about 0.01 μg/kg body weight to about 0.1 mg/kg body weight. In another embodiment, the dosage of S6K2 and/or derivatives thereof comprises less than about 0.1 mg/kg body weight. More preferred dosages for systemic administration comprise about 0.1 mg/kg body weight to about 0.05 mg/kg body weight. In another preferred embodiment, the dosage of S6K2 and/or derivatives thereof comprises less than about 0.05 mg/kg body weight. The most preferred dosages for systemic administration comprise between about 1.0 mg/kg body weight to about 0.01 mg/kg body weight. In other embodiments, the amount of S6K2 and/or derivatives thereof administered is an amount effective to bring the concentration of S6K2 and/or derivatives thereof in the serum to a concentration of about 20.0, 10.0, 5.0, 2.50, 1.25, 0.625, 0.3125, 0.156, 0.078, 0.039, 0.020, 0.010, 0.005, 0.003, 0.0015, 0.0008, 0.0003 or 0.0001 nM. The preferred dosages for direct administration to a site via microinfusion comprise 1 ng/kg body weight to 1 mg/kg body weight.

[0157] In addition to S6K2 polypeptide or derivatives thereof or in the alternative to p300 acetyltransferase binding domain polypeptides or derivatives thereof, the pharmaceutical compositions of the present invention may contain suitable pharmaceutically acceptable carriers comprising excipients and auxiliaries that facilitate processing of the active compounds into preparations which can be used pharmaceutically for delivery to the site of action. Suitable formulations for parenteral administration include aqueous solutions of the active compounds in water-soluble form, for example, water-soluble salts. In addition, suspensions of the active compounds as appropriate oily injection suspensions may be administered. Suitable lipophilic solvents or vehicles include fatty oils, for example, sesame oil or synthetic fatty acid esters, for example, ethyl oleate or triglycerides. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension include, for example, sodium carboxymethyl cellulose, sorbitol and dextran. Optionally, the suspension may also contain stabilizers. Liposomes can also be used to encapsulate the agent for delivery into the cell.

[0158] The pharmaceutical formulation for systemic administration according to the invention may be formulated for enteral, parenteral or topical administration. Indeed, all three types of formulations may be used simultaneously to achieve systemic administration of the active ingredient.

[0159] As mentioned above for some methods of the invention, topical administration may be used. Any common topical formulation such as a solution, suspension, gel, ointment or salve and the like may be employed. Preparation of such topical formulations are described in the art of pharmaceutical formulations as exemplified, for example, by Gennaro et al. (2000) Remington's Pharmaceutical Sciences, Mack Publishing. For topical application, the compositions could also be administered as a powder or spray, particularly in aerosol form. In some embodiments, the compositions of this invention may be administered by inhalation. For inhalation therapy, the active ingredients may be in a solution useful for administration by metered dose inhalers or in a form suitable for a dry powder inhaler. In another embodiment, the compositions are suitable for administration by bronchial lavage.

[0160] Suitable formulations for oral administration include hard or soft gelatin capsules, pills, tablets, including coated tablets, elixirs, suspensions, syrups or inhalations and controlled release forms thereof.

[0161] The invention also includes isotopically-labeled agents that bind to S6 kinase, including human S6 kinase, that have one or more atoms replaced by an atom having an atomic mass or mass number different from the atomic mass or mass number usually found in nature. In one embodiment, the isotopically labeled agents bind to the acetylated form of S6 kinase while in other embodiments, they bind to the deacetylated form of S6 kinase.

[0162] In one embodiment of the composition and methods of the invention, the abnormal cell growth is cancer. As used herein, the term "cancer" unless otherwise indicated, refers to diseases that are characterized by uncontrolled, abnormal cell growth and/or proliferation. Types of cancer where the compositions are useful include, but are not limited to, prostate cancer, breast cancer, multiple myeloma, lung cancer, non-small cell lung carcinoma, bone cancer, liver cancer, pancreatic cancer, skin cancer, cancer of the head or neck, cutaneous or intraocular melanoma, uterine cancer, ovarian cancer, rectal cancer, cancer of the anal region, stomach cancer, colon cancer, uterine cancer, carcinoma of the fallopian tubes, carcinoma of the endometrium, carcinoma of the cervix, carcinoma of the vagina, carcinoma of the vulva, Hodgkin's Disease, cancer of the esophagus, cancer of the small intestine, cancer of the endocrine system, cancer of the thyroid gland, cancer of the parathyroid gland, cancer of the adrenal gland, sarcoma of soft tissue, cancer of the urethra, cancer of the penis, chronic or acute leukemia, lymphocytic lymphomas, cancer of the bladder, cancer of the kidney or ureter, renal cell carcinoma, carcinoma of the renal pelvis, neoplasms of the central nervous system (CNS), neuroectodermal cancers, spinal axis tumors, glioma, meningioma, pituitary adenoma, or a combination of one or more of the foregoing cancers.

[0163] In another embodiment of the compositions and methods of the invention, the abnormal cell growth is a benign proliferative disease, including, but not limited to, benign prostatic hyperplasia, hypertrophy or restinosis.

[0164] As discussed above, the invention includes compositions and methods for the treatment of abnormal cell growth in a mammal, including a human, which comprises, consists of, or consists essentially of an effective amount of a polypeptide containing a p300 acetyltransferase binding domain, in combination with a pharmaceutically acceptable carrier.

[0165] The invention further includes isotopically-labeled agents that bind to S6K2 at the AT-hook domain, including human S6K2, that have one or more atoms replaced by an atom having an atomic mass or mass number different from the atomic mass or mass number usually found in nature. In one embodiment, the isotopically-labeled agents bind to the activated form of S6K2 while in other embodiments, they bind to the inactivated form of S6K2.

[0166] Examples of isotopes that can be incorporated into compounds of the invention include, but are not limited to, isotopes of hydrogen, carbon, phosphorous, iodine, rhenium, indium, yttrium, technetium and lutetium (i.e., including, but not limited to, ^{3}H, ^{14}C, ^{31}P, ^{32}P, ^{35}S, ^{131}I, ^{125}I, ^{123}I, ^{187}Re, ^{64}Cu, ^{111}In, ^{90}Y, ^{99m}Tc, ^{177}Lu), others isotopes of these elements, and other isotopes known in the art. Agents of the present invention, prodrugs thereof, and pharmaceutically acceptable salts of said agents or of said prodrugs which contain the aforementioned isotopes and/or other isotopes of other atoms are within the scope of this invention. Tritium and carbon-14 isotopes are particularly preferred for their ease of preparation and detectability. Further, substitution with heavier isotopes such as deuterium can afford certain therapeutic advantages resulting from greater metabolic stability, for example increased in vivo half-life or reduced dosage requirements and, hence, may be preferred in some circumstances.

Fusion Proteins

[0167] The invention includes fusion polypeptides containing a p300 acetyltransferase binding domain and salts thereof, comprising at least one second polypeptide. In one embodiment, the fusion polypeptides are capable of binding to S6 kinase, including human S6 kinase. In a further embodiment, the fusion polypeptides are capable of modulating the activity of the S6 kinase by effecting acetylation at the binding domain. In one embodiment, the fusion polypeptides facilitate acetylation while in another embodiment, they inhibit acetylation by another acetyltransferase (e.g., p300).

[0168] In some embodiments, the second polypeptide includes a cancer cell-binding domain which specifically binds to an epitope expressed only on cells exhibiting abnormal growth (i.e. cancer cells). The term "cancer cell-binding domain" refers to an amino acid sequence capable of binding or otherwise specifically associating with a cell displaying abnormal growth (e.g., benign and malignant cancer cells). In some embodiments the cancer cell binding-domain is an antibody while in other embodiments it is a ligand which specifically binds to a receptor expressed only on cancer cells. Examples of antibodies include, but are not limited to, antibodies which specifically bind to B-cells or T-cells. Examples of receptor ligands include, but are not limited to, cytokines and growth factors including epidermal growth factor.

[0169] The second polypeptide can also include a stabilization domain which increases the in vitro and in vivo half-life of the fusion polypeptide. As used herein, the term "stabilization domain" refers to an amino acid sequence capable of extending the in vitro and in vivo half-life of the p300 acetyltransferase binding domain polypeptide when compared to polypeptide alone. The stabilization domain can comprise human proteins (e.g., full length or truncated, soluble proteins from extracellular fragments, etc) such as human serum albumin, transferrin or other proteins which stabilize the in vivo or in vitro half-life of the polypeptide or derivative thereof. These additional functional domains may themselves serve as linker peptides, for example, for joining a cancer cell-binding domain to the p300-4 fragment or derivative. Alternatively, they may be located elsewhere in the fusion molecule (e.g., at the amino or carboxy terminus thereof). In alternative embodiments, the stabilization domain is a chemical moiety (e.g., PEG (polyethylene glycol) or a dextran).

[0170] The term "fused" or "fusion polypeptide" as used herein refers to polypeptides in which: (i) a given functional domain (i.e. a cancer cell-binding domain) is bound at its carboxy terminus by a covalent bond either to the amino terminus of another functional domain (i.e., an human serum albumin component) or to a linker peptide which itself is bound by a covalent bond to the amino terminus of p300 acetyltransferase binding domain; or (ii) a given functional domain (i.e. a cancer cell-binding domain) is bound at its amino terminus by a covalent bond either to the carboxy terminus of another functional domain (i.e., an human serum albumin component) or to a linker peptide which itself is bound by a covalent bond to the carboxy terminus of the p300 acetyltransferase binding domain.

[0171] Similarly, "fused" when used in connection with the nucleic acid intermediates of the invention means that the 3'-[or 5'-] terminus of a nucleotide sequence encoding a first functional domain is bound to the respective 3'-[or 5'-] terminus of a nucleotide sequence encoding a second functional domain, either by a covalent bond or indirectly via a nucleotide linker which itself is covalently bound preferably at its termini to the first functional domain-encoding polynucleotide and optionally, a second functional domain-encoding nucleic acid.

[0172] Examples of fusion polypeptides of the invention may be represented by, but are not limited by, the following formulas:

R1-L-R2 (i)

R2-L-R1 (ii)

R1-L-R2-L-R1 (iii)

R1-L-R1-L-R2 (iv)

R2-L-R1-L-R1 (v)

wherein R1 is the amino acid sequence of a cancer cell-binding domain, R2 is the amino acid sequence of a stabilizing domain (e.g., human serum albumin), each L is p300 acetyltransferase binding domain which is bound by a covalent bond to a terminus of R1 and/or R2, whereby the above molecule fragments are read directionally (i.e., with the left side corresponding to the amino terminus and the right side to the carboxy terminus of the molecule).

[0173] The present invention also includes compositions where a cytotoxic agent is linked to a p300 acetyltransferase binding domain polypeptide or derivative thereof. Examples of cytotoxic agents include, but are not limited to, gelonin, ricin, saponin, pseudonomas exotoxin, pokeweed antiviral protein, diphtheria toxin, complement proteins, or any other agent known in the art which is capable of killing a cell upon contact with that cell.

Methods of Treatment

[0174] This invention includes methods for the treatment of abnormal cell growth in a mammal, including a human, comprising administering to said mammal an amount of an agent capable of modulating the acetylation of a kinase protein. In one embodiment, the acetylation is inhibited while in another embodiment the agent is capable of acetylating, or facilitating the acetylation of, the kinase protein. In another embodiment, the kinase protein is S6 kinase, including human S6 kinase. This invention also includes methods for the treatment of abnormal cell growth in a mammal, including a human, comprising administering to said mammal an amount of an agent capable of modulating the activity of a kinase protein. In one embodiment, the DNA binding is inhibited while in another embodiment the agent is capable of facilitating the DNA binding of a kinase protein. In another embodiment, the kinase protein is S6K2, including human S6K2.

[0175] In one aspect of the invention, the above described agents are deacetylase inhibitors which can be administered to inhibit the activity of a kinase which is modulated by an acetyltransferase. In one embodiment, the deacetylase inhibitor is an S6 kinase deacetylase inhibitor. In another embodiment, the deacetylase inhibitor is an PKCa kinase deacetylase inhibitor. In yet another embodiment, the deacetylase inhibitors will have the same activity as the p300-4 fragment described herein (i.e., inhibition of S6 kinase activity). Types of deacetylase inhibitors which can be used in the methods of the invention include, but are not limited to, low molecular weight carboxylates (e.g., sodium butyrate, valproic acid, sodium phenylbutyrate), hydroxamic acids (e.g., CBHA, TSA, SAHA, LAQ824), benzamides (e.g., MS-275, CI-994), epoxyketones (e.g., AOE, trapoxin B), cyclic peptides (e.g., depsipeptide, apicidin), and hybrid molecules such as CHAP31 and CHAP50. Exemplary examples of deacetylase inhibitors include, but are not limited to, hydroxamate-type HDAC inhibitors such as LAQ824 (Atadja et al. (2004) Cancer Res. 64, 689-695), suberoylanilide hydroxamic acid (Kelly et al. (2003) Clin. Cancer Res. 9, 3578-3588); CI-994 (Loprevite et al. (2005) Oncol Res. 15, 39-48); JNJ1641199 (Arts et al. (2003) AACR-NCI-EORTC Int. Conf. Mol. Targets. Cancer Ther.) and those set forth in Guo et al. (2005) Ann. Rev. Pharmacol. Toxicol. 45, 495-528).

[0176] The invention also includes methods of treatment of the above referenced kinase-related disorders with a deacetylase inhibitor in combination with kinase inhibitor. Types of kinase inhibitors which can be used in the methods of the invention include, but are not limited to, rampamycin derivatives such as RAD001 (Lane (2003) Mol. Targets. Cancer Therapeut. 259-260); serine-threonine kinase inhibitors such as mTor inhibitors, including but not limited to, AP23573 (Mita et al. (2004) Proc. American Soc. Clin. Oncology, 23, 3076) and CCI-779 (Yu et al. (2001) Endocrine-Related Cancer, 8, 249-258). In one embodiment, the deacetylase inhibitor is an S6 kinase inhibitor.

[0177] In some embodiments of the invention, the agent will have activity similar any of the following: a polypeptide comprising p300 acetyltransferase binding domain or derivative thereof, an fusion protein comprising p300 acetyltransferase binding domain or a derivative thereof, or a pharmaceutical composition comprising an amount of p300 acetyltransferase binding domain or derivatives thereof, that is effective in inhibiting S6 kinase mediated activity associated with abnormal cell growth, without the addition of other therapeutic agents.

[0178] In some embodiments of this method, the abnormal cell growth is cancer, including, but not limited to, prostate cancer, breast cancer, multiple myeloma, lung cancer, non-small cell lung carcinoma, bone cancer, liver cancer, pancreatic cancer, skin cancer, cancer of the head or neck, cutaneous or intraocular melanoma, uterine cancer, ovarian cancer, rectal cancer, cancer of the anal region, stomach cancer, colon cancer, uterine cancer, carcinoma of the fallopian tubes, carcinoma of the endometrium, carcinoma of the cervix, carcinoma of the vagina, carcinoma of the vulva, Hodgkin's Disease, cancer of the esophagus, cancer of the small intestine, cancer of the endocrine system, cancer of the thyroid gland, cancer of the parathyroid gland, cancer of the adrenal gland, sarcoma of soft tissue, cancer of the urethra, cancer of the penis, chronic or acute leukemia, lymphocytic lymphomas, cancer of the bladder, cancer of the kidney or ureter, renal cell carcinoma, carcinoma of the renal pelvis, neoplasms of the central nervous system (CNS), neuroectodermal cancer, spinal axis tumors, glioma, meningioma, pituitary adenoma, or a combination of one or more of the foregoing cancers. In another embodiment of said method, said abnormal cell growth is a benign proliferative disease, including, but not limited to, psoriasis, benign prostatic hyperplasia, hypertrophy or restinosis.

[0179] This invention also includes methods for the treatment of abnormal cell growth in a mammal which comprises administering to said mammal, including a human, a pharmaceutical composition comprising an amount of an agent capable of modulating the acetylation of a kinase protein. The methods may alternatively comprise administration of an amount of agent capable of modulating the DNA binding of a kinase protein. In one embodiment, the acetylation is inhibited while in another embodiment the agent itself is capable of acetylating the kinase protein. In an alternate embodiment, the DNA binding is inhibited while in another embodiment the agent itself is capable of facilitating DNA binding by the kinase protein. In another embodiment, the kinase protein is S6 kinase, including human S6 kinase. In yet another embodiment, the kinase protein is S6K2, including human S6K2. In some embodiments, the agent is a polypeptide comprising a p300 acetyltransferase binding domain or derivatives thereof that is effective in inhibiting abnormal cell growth associate with S6 kinase expression and activity.

[0180] In some other embodiments, the agent is effective in inhibiting abnormal cell growth associate with S6K2 expression and activity. This includes the abnormal growth and/or proliferation of cancer cells including benign and malignant cells of neoplastic diseases. Inhibition of abnormal cell growth can occur by a variety of mechanisms including, but not limited to, cell death, apoptosis, inhibition of cell division, transcription, translation, transduction, etc.

[0181] In practicing the methods of this invention, these agents may be used alone or in combination with other inactive ingredients. As discussed above, the present invention includes compositions and methods where a drug or cytotoxic agent is linked to an agent capable of modulating acetylation of a kinase, including S6 kinase. In addition, as discussed above, the present invention includes compositions and methods where a drug or cytotoxic agent is linked to an agent capable of modulating DNA binding of a kinase, including S6K2. The methods of the invention therefore include administration of a polypeptide comprising these agents linked to a cytotoxic agent for the treatment of a disease associated with abnormal cell growth, including cancer. Examples of cytotoxic agents include, but are not limited to, gelonin, ricin, saponin, pseudonomas exotoxin, pokeweed antiviral protein, diphtheria toxin, complement proteins, or any other agent known in the art which is capable of killing a cell upon contact with that cell.

[0182] The compositions and methods of the invention for therapeutic purposes can be utilized in vivo, ordinarily in mammals, such as humans, sheep, horses, cattle, pigs, dogs, cats, rats and mice or ex vivo. The invention is particularly useful in the treatment of human subjects.

Diagnostic Methods

[0183] This invention includes methods for the diagnosis of a disorder associated with abnormal kinase activity, including cancer, in a mammal, including a human, comprising administering to said mammal an effective amount of antibody which is capable of measuring acetylation of one or more kinases in a cell, such as a cancer cell. The diagnostic methods of the invention include in vitro methods where a biological sample is first isolated from a mammal, including humans, and then contacted with an amount of the antibody, that is effective in specifically and selectively binding to cells expressing an acetylated form of a kinase, such as cancer cells.

[0184] In one embodiment of this diagnostic method, the kinase related disorder is cancer. Types of cancer included in the methods of the invention include, but are not limited to, prostate cancer, breast cancer, multiple myeloma, lung cancer, non-small cell lung carcinoma, bone cancer, liver cancer, pancreatic cancer, skin cancer, cancer of the head or neck, cutaneous or intraocular melanoma, uterine cancer, ovarian cancer, rectal cancer, cancer of the anal region, stomach cancer, colon cancer, uterine cancer, carcinoma of the fallopian tubes, carcinoma of the endometrium, carcinoma of the cervix, carcinoma of the vagina, carcinoma of the vulva, Hodgkin's Disease, cancer of the esophagus, cancer of the small intestine, cancer of the endocrine system, cancer of the thyroid gland, cancer of the parathyroid gland, cancer of the adrenal gland, sarcoma of soft tissue, cancer of the urethra, cancer of the penis, chronic or acute leukemia, lymphocytic lymphomas, cancer of the bladder, cancer of the kidney or ureter, renal cell carcinoma, carcinoma of the renal pelvis, neoplasms of the central nervous system (CNS), neuroectodermal cancer, spinal axis tumors, glioma, meningioma, pituitary adenoma, or a combination of one or more of the foregoing cancers. In another embodiment of said method, said abnormal cell growth is a benign proliferative disease, including, but not limited to, psoriasis, benign prostatic hyperplasia, hypertrophy or restinosis.

[0185] In practicing the diagnostic methods of this invention, the antibody may be used alone or in combination with other inactive ingredients. As discussed above, the present invention includes compositions and methods where a label is linked to the antibody. The methods of the invention therefore include administration of labeled antibody the diagnosis of a disease associated with abnormal kinase activity, including cancer.

[0186] The compositions and methods of the invention for diagnostic purposes can be utilized both in vivo, ordinarily in mammals, such as humans, sheep, horses, cattle, pigs, dogs, cats, rats and mice or in vitro. The invention is particularly useful in the diagnosis of human subjects.

Methods to Identify Agents that Modulate Activity

[0187] The present invention provides methods for identifying agents that modulate at least one activity of a kinase (e.g., phosphorylation) by measuring acetylation of the kinase protein by an acetyltransferase protein. Such methods or assays may utilize any means of monitoring or detecting the desired activity (i.e., acetylation) and include changes in phosphorylation activity following acetylation of the kinase protein.

[0188] In one format, the specific activity of particular kinase, normalized to a standard unit, between a cell population that has been exposed to the agent to be tested compared to an unexposed control cell population may be assayed. Cell lines or populations are exposed to the agent to be tested under appropriate conditions and time. In this format, probes such as specific antibodies are used to monitor the phosphorylation of the kinase protein substrates in the different cell populations. In another embodiment, acetylation of the kinase protein in each of the samples can also be measured to determine the effect on kinase activity. Cellular lysates may be prepared from the exposed cell line or population and a control, unexposed cell line or population. The cellular lysates are then analyzed with the probe.

[0189] Agents that are assayed in the above method can be randomly selected or rationally selected or designed. As used herein, an agent is said to be randomly selected when the

agent is chosen randomly without considering the specific sequences involved in the association of the protein of the invention alone or with its associated substrates, binding partners, etc. An example of randomly selected agents is the use a chemical library or a peptide combinatorial library, or a growth broth of an organism.

[0190] The agents of the present invention can be, as examples, peptides, peptide mimetics, antibodies, antibody fragments, small molecules, vitamin derivatives, as well as carbohydrates. A skilled artisan can readily recognize that there is no limit as to the structural nature of the agents of the present invention. In one embodiment, the small molecules are deacetylase inhibitors as described herein.

[0191] The peptide agents of the invention can be prepared using standard solid phase (or solution phase) peptide synthesis methods, as is known in the art. In addition, the DNA encoding these peptides may be synthesized using commercially available oligonucleotide synthesis instrumentation and produced recombinantly using standard recombinant production systems. The production using solid phase peptide synthesis is necessitated if non-gene-encoded amino acids are to be included.

[0192] Another class of agents of the present invention are antibodies or fragments thereof that bind to the either the kinase protein or the acetyltransferase. Antibody agents can be obtained by immunization of suitable mammalian subjects with peptides, containing as antigenic regions, those portions of the protein intended to be targeted by the antibodies as described herein.

Methods to Identify Agents that Modulate the Expression

[0193] Another embodiment of the present invention provides methods for identifying agents that modulate the expression of a nucleic acid encoding an S6 kinase protein such as a protein having the amino acid sequence of SEQ ID NO: 2 or 8. Such assays may utilize any available means of monitoring for changes in the expression level of the nucleic acids of the invention. As used herein, an agent is said to modulate the expression of a nucleic acid of the invention, for instance a nucleic acid encoding the protein having the sequence of SEQ ID NO: 1, if it is capable of up- or down-regulating expression of the nucleic acid in a cell.

[0194] In one assay format, cell lines that contain reporter gene fusions between the open reading frame of the S6 kinase protein alpha (SEQ ID NO: 1) or S6 kinase protein beta (SEQ ID NO: 7) and any assayable fusion partner may be prepared. Numerous assayable fusion partners are known and readily available including the firefly luciferase gene and the gene encoding chloramphenicol acetyltransferase (Alam et al., 1990 Anal. Biochem. 188, 245-254). Cell lines containing the reporter gene fusions are then exposed to the agent or a polypeptide comprising a p300 acetyltransferase binding domain (i.e., control) to be tested under appropriate conditions and time. Differential expression of the reporter gene between samples exposed to the agent and control samples identifies agents which modulate the expression of a nucleic acid encoding the S6 kinase protein.

[0195] Additional assay formats may be used to monitor the ability of the agent to modulate the expression of a nucleic acid encoding a protein of the invention such as the S6 kinase protein. For instance, mRNA expression may be monitored directly by hybridization to the nucleic acids of the invention. Cell lines are exposed to the agent and/or a polypeptide comprising a p300 acetyltransferase binding domain to be tested under appropriate conditions and time and total RNA or mRNA is isolated by standard procedures such those disclosed in Sambrook et al. (2001) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory Press).

[0196] Probes to detect differences in RNA expression levels between cells exposed to the agent and control cells may be prepared from the nucleic acids of the invention. It is preferable, but not necessary, to design probes which hybridize only with target nucleic acids under conditions of high stringency. Only highly complementary nucleic acid hybrids form under conditions of high stringency. Accordingly, the stringency of the assay conditions determines the amount of complementarity which should exist between two nucleic acid strands in order to form a hybrid. Stringency should be chosen to maximize the difference in stability between the probe:target hybrid and potential probe:non-target hybrids.

[0197] Probes may be designed from the nucleic acids of the invention through methods known in the art. For instance, the G+C content of the probe and the probe length can affect probe binding to its target sequence. Methods to optimize probe specificity are commonly available in Sambrook et al. (2001) or Ausubel et al (1995) Current Protocols in Molecular Biology, Greene Publishing Company).

[0198] Hybridization conditions are modified using known methods, such as those described by Sambrook et al. (2001) and Ausubel et al. (1995) as required for each probe. Hybridization of total cellular RNA or RNA enriched for polyA RNA can be accomplished in any available format. For instance, total cellular RNA or RNA enriched for polyA RNA can be affixed to a solid support and the solid support exposed to at least one probe comprising at least one, or part of one of the sequences of the invention under conditions in which the probe will specifically hybridize. Alternatively, nucleic acid fragments comprising at least one, or part of one of the sequences of the invention can be affixed to a solid support, such as a porous glass wafer. The glass wafer can then be exposed to total cellular RNA or polyA RNA from a sample under conditions in which the affixed sequences will specifically hybridize. Such glass wafers and hybridization methods are widely available, for example, those disclosed by Beattie (WO 95/11755). By examining for the ability of a given probe to specifically hybridize to an RNA sample from an untreated cell population and from a cell population exposed to the agent, agents which up or down regulate the expression of a nucleic acid encoding the S6 kinase protein.

Methods to Identify Binding Partners

[0199] Another embodiment of the present invention provides methods for identifying agents that modulate at least one activity of S6 kinase protein. Such methods or assays may utilize any means of monitoring or detecting the desired activity.

[0200] In one format, the relative amounts of a S6 kinase protein between a cell population that has been exposed to the agent to be tested and/or a polypeptide comprising a p300 acetyltransferase binding domain compared to an unexposed control cell population may be assayed. In this format, probes such as specific antibodies are used to monitor the differential expression of the protein in the different cell populations. Cell lines or populations are exposed to the agent to be tested under appropriate conditions and time. Cellular lysates may be prepared from the exposed cell line or population and a control, unexposed cell line or population. The cellular lysates are then analyzed with the probe.

[0201] For example, fragments comprising lysine 516 of S6 kinase protein can be expressed in bacteria and used to search for agents which bind to these fragments. Fusion proteins, such as His-tag or GST fusion to the N- or C-terminal regions of S6 kinase protein comprising lysine 516 can be prepared for use as a S6 kinase protein fragment substrate. These fusion proteins can be coupled to Talon or Glutathione-Sepharose beads and then probed with cell lysates. Prior to lysis, the cells may be treated with rapamycin or other drugs which may modulate S6 kinase protein or proteins that interact with S6 kinase protein. Lysate proteins binding to the fusion proteins can be resolved by SDS-PAGE, isolated and identified by protein sequencing or mass spectroscopy, as is known in the art.

[0202] Kinase assays to measure the ability of the agent to modulate the kinase activity of a protein of the invention are widely available such as the assays disclosed by Mishima et al. (1996) J. Biochem. 119, 906-913) and Michnoff et al. (1986) J. Biol. Chem. 261, 8320-8326. Alternative assay formats include actin-myosin motility assays such as those disclosed by Kohama et al. (1996) TIPS 17, 284-287 or Warrick et al. (1987) Ann. Rev. Cell. Biol. 3, 379-421.

[0203] Agents that are assayed in the above method can be randomly selected or rationally selected or designed. As used herein, an agent is said to be randomly selected when the agent is chosen randomly without considering the specific sequences involved in the association of the protein of the invention alone or with its associated substrates, binding partners, etc. An example of randomly selected agents are agents present in a chemical library or a peptide combinatorial library, or agents present in the medium in which an organism has been grown.

[0204] As used herein, an agent is said to be rationally selected or designed when the agent is chosen on a non-random basis which takes into account the sequence of the target site and/or its conformation in connection with the agent's action. There are proposed binding sites for ATP/GTP and calmodulin as well as cAMP/cGMP kinase sites, TyrP sites and Ser/Thr kinase (catalytic) sites in the protein having SEQ ID NO: 2 or SEQ ID NO: 8. Agents can be rationally selected or rationally designed by utilizing the peptide sequences that make up these sites. For example, a rationally selected peptide agent can be a peptide whose amino acid sequence is identical to the ATP or calmodulin binding sites or domains.

[0205] The agents of the present invention can be, as examples, peptides, small molecules, vitamin derivatives, as well as carbohydrates. A skilled artisan can readily recognize that there is no limit as to the structural nature of the agents of the present invention.

[0206] The peptide agents of the invention can be prepared using standard solid phase (or solution phase) peptide synthesis methods, as is known in the art. In addition, the DNA encoding these peptides may be synthesized using commercially available oligonucleotide synthesis instrumentation and produced recombinantly using standard recombinant production systems. The production using solid phase peptide synthesis is necessitated if non-gene-encoded amino acids are to be included.

Antibodies

[0207] Another class of agents of the present invention are antibodies. These antibodies may be immunoreactive with critical positions of S6 kinase protein where acetylation occurs (e.g., Lys516). In an alternate embodiment, these antibodies are immunoreactive with critical positions of the S6K2 protein where DNA binding occurs (e.g. amino acids corresponding to position 487 to 495 of SEQ ID NO: 8). Antibody agents are obtained by immunization of suitable mammalian subjects with peptides, containing as antigenic regions, those portions of the protein intended to be targeted by the antibodies.

[0208] Antibodies which bind to an acetylation site on a kinase protein are encompassed in the invention. Such antibodies include, but are not limited to, antibodies which bind to the acetylation site on the S6 kinase protein (i.e., Lysine at 516). Antibodies which also bind to the acetyltransferase domain on an acetyltransferase protein capable of acetylating a kinase are also encompassed in the scope of the invention. Such antibodies include, but are not limited to, antibodies which bind to the acetylation site on p300 acetyltransferase (i.e., residues 1,459 to 1,892), and capable of inhibiting the activity of this protein are also encompassed in the invention.

[0209] Antibodies are prepared by immunizing suitable mammalian hosts in appropriate immunization protocols using the peptides, polypeptides, or proteins of the invention, such as S6 kinase protein, variants and isolated binding partners, if they are of sufficient length, or, if desired, or if required to enhance immunogenicity, conjugated to suitable carriers. Methods for preparing immunogenic conjugates with carriers such as bovine serum albumin (BSA), keyhole limpet hemocyanin (KLH), or other carrier proteins are well known in the art. In some circumstances, direct conjugation using, for example, carbodiimide reagents may be effective; in other instances linking reagents well known in the art, may be desirable to provide accessibility to the hapten. The hapten peptides can be extended at either the amino or carboxy terminus with a cysteine residue or interspersed with cysteine residues, for example, to facilitate linking to a carrier. Administration of the immunogens is conducted generally by injection over a suitable time period and with use of suitable adjuvants, as is generally understood in the art. During the immunization schedule, titers of antibodies are taken to determine adequacy of antibody formation.

[0210] Anti-peptide antibodies can be generated using synthetic peptides corresponding to, for example, the carboxy terminal 15 amino acids S6 kinase protein or amino acids 487 to 495 of SEQ ID NO: 8. Antibodies which bind to the acetylation site on a kinase can also be generated and include antibodies which bind to the acetylation site on the S6 kinase protein (i.e., Lysine at 516). Synthetic peptides can be as small as 1 to 3 amino acids in length, but are preferably at least 4 or more amino acid residues long. The peptides are coupled to KLH using standard methods and can be immunized into animals such as rabbits. Polyclonal antibodies (e.g. anti-S6 kinase protein peptide, anti-S6K2 protein peptide) can then be purified, for example using Actigel beads containing the covalently bound peptide.

[0211] While the polyclonal antisera produced in this way may be satisfactory for some applications, for pharmaceutical compositions, use of monoclonal preparations is preferred. Immortalized cell lines which secrete the desired monoclonal antibodies may be prepared using the standard method of Kohler and Milstein or modifications which effect immortalization of lymphocytes or spleen cells, as is generally known. The immortalized cell lines secreting the desired antibodies are screened by immunoassay in which the antigen is the peptide hapten. polypeptide or protein. When the appropriate

immortalized cell culture secreting the desired antibody is identified, the cells can be cultured either in vitro or by production in ascites fluid. Of particular interest, are monoclonal antibodies which recognize the proline-rich domain of S6 kinase protein.

[0212] The desired monoclonal antibodies are then recovered from the culture supernatant or from the ascites supernatant. Fragments of the monoclonals or the polyclonal antisera which contain the immunologically significant portion can be used as antagonists, as well as the intact antibodies. Use of immunologically reactive fragments, such as Fav, $_{SC}$FV, Fab, Fab' or F(ab')$_2$ fragments is often preferable, especially in a therapeutic context, as these fragments are generally less immunogenic than the whole immunoglobulin. Antibodies may preferably also be human, humanized or chimeric variants of the foregoing. Such antibodies can be less immunogenic when administered to a subject. Methods of producing humanized or chimeric antibodies are well known in the art. The antibodies contemplated also include different isotypes and isotype subclasses (e.g., IgG_1, IgG_2, IgM). These antibodies can be prepared by raising them in vertebrates, in hybridoma cell lines or other cell lines, or by recombinant means. For references on how to prepare these antibodies, see Harlow & Lane (1988), Antibodies: A Laboratory Manual, Cold Spring Harbor Press.

[0213] The antibodies or fragments may also be produced, using current technology, by recombinant means. Regions that bind specifically to the desired regions of receptor can also be produced in the context of chimeras with multiple species origin.

[0214] Without further description, it is believed that one of ordinary skill in the art can, using the preceding description and the following illustrative examples, make and utilize the agents of the present invention and practice the claimed methods. The following working examples describe embodiments of the present invention, and are not to be construed as limiting in any way the remainder of the disclosure.

Example 1

[0215] This example demonstrates that serum stimulation induces complex formation between p300 acetyltransferase and S6K1. Since S6 kinases appear to play a role in the regulation of transcription, it would be of interest to identify components of the transcriptional machinery with which they may interact. To this end, S6K1 and 2 were immunoprecipitated from MCF-7 cells before and after serum stimulation and blotting was carried out with a panel of antibodies to search for binding partners (data not shown). Interestingly, both isoforms of S6K show serum-inducible binding to the acetyltransferases, p300 and P/CAF (FIGS. **1**A and D). Importantly, no p300 was detected in control immunoprecipitations using pre-immune sera from the rabbits used to produce the S6 kinase polyclonal antibodies. This demonstrates that the p300 interacts with the S6 kinases themselves rather than the anti-S6 kinase antibodies or protein A sepharose. The specificity of these interactions was further confirmed by overexpression of tagged S6 kinase and p300 constructs in HEK 293 and NIH-3T3 cells (FIGS. **1**B and **1**C). In FIG. **1**B, a p300/Gal4 DNA binding domain (DBD) fusion was transfected into HEK 293 cells together with EE-tagged S6K1 or empty pcDNA 3.1 vector. EE-S6K1 was immunoprecipitated using anti-EE antibody and immunoprecipitates were resolved by SDS-PAGE. Following transfer, the top part of the nitrocellulose membrane was blotted with an antibody against the Gal4 DBD to detect coprecipitated p300, while the lower part was blotted with S6K1 antibodies. Note that lysates from cells transfected with p300 and the empty pcDNA 3.1 plasmid were also subjected to immunoprecipitation using the EE-antibody. The lack of a p300 band in this lane demonstrates that the p300 is binding specifically to S6K1 and not to the anti-EE antibody, or to the protein A sepharose. FIG. **1**C shows binding of EE-tagged S6K2 to endogenous p300 in NIH 3T3 cells. Here, cells were transfected whether with EE-S6K2 or with empty pcDNA 3.1. Anti-EE immunoprecipitates were resolved by SDS-PAGE and this time, the blot was probed with p300 antibody. Again, p300 was only detected in the presence of S6K2 and not in the negative control anti-EE immunoprecipitate. In FIGS. **1**B and **1**C, cell lysate equal to approximately 10 percent of the input material used for immunoprecipitations was run alongside the samples (labeled 'Input'). This gives an idea of the proportion of total p300 that is complexed with S6 kinases under these conditions and also serves as a marker to ensure the band detected in the IP lanes is of the correct molecular weight to be p300.

[0216] Having established that S6 kinases can interact with p300 in cells, it was interesting to see whether this interaction occurs directly and to determine to which region of p300 the S6 kinases bind. In order to investigate this, p300 was divided into five domains (FIG. **1**D, upper panel). Each domain was cloned into the pGEX4T2 vector to allow expression as N-terminal GST fusion proteins in bacteria. These fusion proteins were expressed and purified and used for GST pulldown experiments with highly purified recombinant S6 kinases. FIG. **1**D (lower panel) shows selective binding of S6K2 to p300 domain 4, which contains part of the acetyltransferase domain (FIG. **1**D). Similar results were obtained for S6K1 (data not shown). These data show that S6 kinases can bind directly to p300 without the need for accessory factors such as other proteins or DNA, however, the interaction is weak, and is probably further stabilised in cells by other components.

Example 2

[0217] Since S6Ks can interact with a region within, or adjacent to the acetyltransferase domain of p300 we hypothesized that they may in fact be targeted by its enzymatic activity. This is indeed the case for two other proteins which bind to the p300-domain 4: Flap endonuclease-1 and DNA polymerase-β (Hasan et al. 2001b; Hasan et al. 2002) and for a number of proteins that bind to the third zinc finger region within domain 4, such as E1A, E2F, MyoD, SV40 large T and GATA-1 (Vo and Goodman, 2001). In order to test this possibility, recombinant S6K1 or 2 (1 μg) were incubated with GST-p300 HAT domain 4 or GST-P/CAF in HAT buffer containing ^{14}C labelled acetyl coenzyme A (AcCoA). FIG. **2**A shows that when S6K1 and S6K2 were incubated in the presence of GST-p300 HAT domain 4 or GST-P/CAF, ^{14}C labeled acetyl groups were indeed incorporated into the proteins as visualized by autoradiographic detection of specific S6K bands following SDS-PAGE. Note that no label was incorporated into S6K1 or 2 in the absence of GST-p300 HAT domain 4 or GST-P/CAF (lanes 1 and 3). FIG. **2**A shows that PCAF acetylates S6 kinases in vitro with a similar efficiency to that of the p300 HAT domain. In order to further test the specificity of S6K acetylation in vitro, acetylation reactions were conducted using recombinant S6K2 or GST (FIG. **2**B). This experiment shows that while S6K2 is acetylated by p300, no acetylation of GST is seen under these conditions.

Furthermore, in vitro acetylation of other recombinant proteins was tested, including protein kinases from AGC family. In order to test this possibility, full-length p300 was overexpressed in HEK 293 cells and immunoprecipitated. Recombinant kinases produced in insect cells (1 μg) were added to the immunoprecipitates in HAT buffer containing ^{14}C labelled AcCoA. FIG. **2**C shows that S6K2 is acetylated in the presence of p300 and AcCoA, but no acetylation is observed for PKB/Akt, PDK1, PI3-K (p85a/p110a complex) under conditions used. We detected weak, but reproducible acetylation of PKCa. These data further demonstrate the specificity of in vitro acetylation assay and indicate that other protein kinases could be acetylated in vitro.

[0218] In order to test whether S6 kinase proteins are subject to acetylation within cells, HEK 293 cells were cotransfected with EE-tagged S6 kinases and p300. Two days after transfection, cells were treated with Trichostatin A or control for 12 hours as indicated. Recombinant S6 kinases were immunoprecipitated using anti-EE and the immunocomplexes were subject to stringent washing to remove non-specifically bound proteins prior to SDS-PAGE. Immunoblotting was conducted with the anti-acetylated lysine antibodies from Upstate. FIG. **2**D clearly shows that acetylated S6K1 and 2 can be detected in 293 cells in the presence of overexpressed p300. The level of S6 kinase acetylation was significantly increased in the presence of Trichostatin A (TSA). TSA is a specific inhibitor of class I and class II histone deacetylases (HDACs), which are responsible for the deacetylation of histones associated with gene silencing. Importantly, re-probing of blots following stripping of the acetyl lysine antibodies demonstrated similar levels of S6 kinases present in all lanes (FIG. **2**D). This confirms the specificity of the acetyl lysine antibodies since they clearly bind to S6 kinases only when they are in an acetylated form. The finding that one or more of the class I/II HDACs targets S6 kinases in cells is in agreement with the published data on the other non-histone acetyltransferase substrates, which are also subject to deacetylation by these enzymes. Thus, it appears that S6 kinase acetylation (similarly to phosphorylation) is subject to dynamic regulation within the cell.

Example 3

[0219] Having established that S6 kinases can be acetylated by p300 in vitro and in cells, attempts were made to find specific sites of acetylation in order to investigate the function (s) of S6 kinase acetylation. Recombinant S6 kinases were acetylated in vitro using either full length p300 or the recombinant p300 HAT domain. Reactions were carried out as before except that unlabelled AcCoA was used and the amount of S6 kinase was increased to 2 μg. S6 kinases were also incubated in the absence of p300 (termed un-acetylated samples). Reaction products were separated by SDS-PAGE, gels were stained with Coomassie blue and S6 kinase bands from both acetylated and un-acetylated samples were excised and subjected to in-gel digestion with trypsin or Arg-C prior to analysis by MALDI TOF mass spectroscopy.

[0220] Initially, peptide mass fingerprints were acquired to check that the in-gel digestion was satisfactory, to establish the protein identification, and to look for putative acetylated peptide ion signals (data not shown). From the S6K1 sample treated with p300 and digested with trypsin, one ion signal was assigned to a tryptic acetylated S6K1 peptide. MALDI tandem mass spectrometry measurements using collision-induced dissociation unambiguously identified this ion as the 509Q-R522S6K1 peptide. The fragment ions of these measurements further revealed that this peptide was acetylated on the Lysine 516 (FIG. **3**A).

[0221] In a complementary approach, samples were digested with Arg-C. This endoproteinase cleaves proteins exclusively at amide bonds directly C-terminal to arginine residues. Following digestion of S6K2 with this enzyme, two peptides predicted to contain acetylated lysines were detected from the acetylated sample. The first peptide, residues 297 to s310 contains a lysine at position 2 (residue 298 in S6K2). This peptide was sequenced and the acetylated lysine was confirmed (data not shown). A second peptide (residues 85 to 98) was predicted to be doubly acetylated. This peptide indeed contained two lysines (positions 88 and 93). It was not possible to sequence this peptide by MALDI tandem mass spectrometry so these lysines must be regarded as potential acetylation sites. It is important to note that both K88 and 93 in S6K2 are evolutionary conserved and are equivalent to K99 and K104 in S6K1.

TABLE 1

S6K Acetylation Sites

Peptide (showing Ac-Lysine)	Enzyme
S6K1:K516 ^{509} **QAFPMIS**K**RPEHLR**^{522}	Trypsin
S6K2:K298 ^{297} **G**K**LALPPYLTPDAR**^{310}	Arg-C
S6K2:K88,K93 ^{85} VLG**K**GGYG**K**VFQVR^{98}	Arg-C

[0222] Alignment of the S6K1 and two amino acid sequences revealed that while K298 in S6K2 is present in S6K1, in this isoform, it is preceded by a cysteine rather than a glycine residue. Many published p300 acetylation sites conform to a consensus in which the target lysine is directly preceded by either a glycine or serine residue and this G/S-K motif has been suggested to be an optimal p300/CBP target sequence (Bannister et al. (2000) Curr. Biol. 10, 467-70). Note that all acetylation sites identified in the experiments presented here conform to this consensus (see Table 1). It is possible therefore, that the lysine in S6K1 equivalent to K298 in S6K2 is a poor substrate for p300 due to the preceding cysteine.

[0223] The sequence surrounding the double acetylation site, K88/93 is highly conserved between S6K2 and S6K1, suggesting that it could be acetylated in both kinases. Inspection of this region, located at the N-terminus of the kinase domain revealed that it is a feature highly conserved in all kinases (also found in other nucleotide binding proteins) called the P-loop. Alignment of this motif with S6Ks from other organisms revealed that both lysines are conserved throughout evolution, from *C. elegans* to humans (FIG. **3**B). Comparison of the S6K P-loop with P-loops from other related AGC kinases showed that the lysines are highly conserved in all members of this family with the exception of PDK1 (FIG. **3**B). The main feature of the P-loop is the invariant GXGXXG motif, which forms an 'elbow' over the bound ATP. The first glycine is in contact with the ribose moiety, while the second glycine lies close to the terminal pyrophosphate of the ATP (Hanks et al. (1987) Proc. Natl. Acad. Sci. USA 84, 388-392).

Example 4

[0224] Mutation of the P-loop lysines severely abrogates S6K activity in vitro. Several recent reports have demonstrated that mutation of lysine to glutamine can mimic acetylation in a manner analogous to substitution of serines or threonines with acidic residues to mimic phosphorylation (Zhang et al. (1998) EMBO J. 17, 3155-3167; Gronroos et al. (2002) Mol. Cell. 10, 483-93). The glutamine residue is neutrally charged at physiological pH and contains a carbonyl group at the end of the side chain, two characteristics possessed by acetylated lysine. Arginine is most commonly substituted for lysine to produce a non-acetylated mutant since, like lysine, it possesses a basic side chain but cannot be acetylated. Since the P-loop is involved in ATP binding, we hypothesized that acetylation of the P-loop lysines may modulate the kinase activity of S6Ks in some way. In order to investigate this possibility, double mutants in which both P-loop lysines were replaced either with arginine or with glutamine were expressed in 293 cells and immunoprecipitated along with wild type S6Ks. Immune complex S6K assays were conducted and gels were exposed to a phosphorimager screen. These experiments were carried out on three separate occasions and in each case it was clear that mutation of these lysine residues has a significant inhibitory effect on the activity of S6Ks (FIGS. **4**A and B). While the activity of the arginine mutant is significantly lower than that of the wild type, mutation to glutamine almost completely abrogates the S6K activity. These data suggest that while it is critical to have lysine at these positions (even the relatively subtle change to arginine reduces activity), acetylation of the lysines (mimicked by the glutamine mutant) may be a way in which S6K activity can be down-regulated by acetyltransferases. Note that phosphorylation of the critical rapamycin-sensitive S6K activation site, threonine 412 is much reduced in the P-loop mutants (FIG. **4**C).

[0225] Taking these data into account, single-site mutants corresponding in S6K1 to K99 and K104. The wtS6K and both mutants were made, K99R and K104R, were cotransfected with p300 into Hek293 cells and immunoprecipitated using anti-EE-tag antibody. The immune complexes and total cell lysates were separated by SDS-PAGE and immunoblotted with pS6K(T412), anti-EE tag or anti-actin antibodies. The data shown in FIG. **4**D clearly indicate that substitution of K99 to R does not effect T412 phosphorylation, while K104R mutant is defective in T412 phosphorylation. The level of T412 phosphorylation was standardized to the level of recombinant S6K1 in total lysates or immunoprecipitates. In addition, reprobing of membranes with anti-actin antibody indicated that equal amount of total protein was loaded from each sample.

Example 5

[0226] The identification of several S6K acetylation sites allowed the study of site-specific acetylation using point mutants. It has been demonstrated that mutation of lysine to glutamine can mimic acetylation in a manner analogous to substitution of serines or threonines with acidic residues to mimic phosphorylation. The glutamine residue is neutrally charged at physiological pH and contains a carbonyl group at the end of the side chain, two characteristics possessed by acetylated lysine. Arginine is most commonly substituted for lysine to produce a non-acetylated mutant since, like lysine, it possesses a basic side chain but cannot be acetylated. All acetylation sites identified by mass spectrometry were mutated to arginine and to glutamine in order to study the possible effects of site-specific acetylation on S6K function. Since the P-loop is involved in ATP binding, we hypothesized that acetylation of the P-loop lysines may modulate the kinase activity of S6Ks in some way. In order to investigate this possibility, double mutants in which both P-loop lysines were replaced either with arginine or with glutamine were expressed in 293 cells and immunoprecipitated along with wild type S6Ks. Immune complex S6K assays were conducted and gels were exposed to a phosphorimager screen. These experiments were carried out on three separate occasions and in each case it was clear that mutation of these lysine residues has a significant inhibitory effect on the activity of S6Ks (FIGS. **4** A and B). While the activity of the arginine mutant is significantly lower than that of the wild type, mutation to glutamine almost completely abrogates the S6K activity. These data suggest that while it is critical to have lysine at these positions (even the relatively subtle change to arginine reduces activity), acetylation of the lysines (mimicked by the glutamine mutant) may be a way in which S6K activity can be down-regulated by acetyltransferases. Note that phosphorylation of the critical rapamycin-sensitive S6K activation site, threonine 412 is much reduced in the P-loop mutants (FIG. **4**C).

[0227] To further elucidate the importance of individual lysine residues in the P loop, we have made mutants which have single K to R substitutions at positions K99 and K104 in S6K1. The wt S6K1 and generated mutants, K99R and K104R, were transiently expressed in Hek 293 cells and immunoprecipitated using anti-EE tag antibody. Total cell lysate and immune complexes were resolved by SDS-PAGE and immunoblotted with antibodies specific for pT412, EE-tag and actin. FIG. **4**D demonstrate that the phosphorylation of K99R mutant is equal to that of wtS6K1 in total lysates and immune complexes. It important to state that the level of T412 phosphorylation in different samples is standardized to the level of actin or recombinant S6K1. In contrast to K99R mutant, the phosphorylation of K104R mutant at T412 was significantly inhibited. This is an intriguing observation, however it is not yet clear under what conditions these particular modifications may act to regulate S6K activity.

[0228] We are currently making K99Q and K104Q mutants, which should mimic the acetylation status of S6K1. The analysis of these mutants would allow us to further elucidate in vivo the role of acetylation of S6K in the P loop.

[0229] Immune complex kinase assays conducted using the S6K1 K516R/Q and the S6K2 K298R/Q mutants immunoprecipitated from asynchronous growing cells revealed no obvious differences from the wild type proteins (data not shown).

Example 6

[0230] The model of S6 kinase domain indicates that P loop lysines are well accessible for posttranslational modifications. The 3D structure of S6K has not been solved so far. Therefore, it was decided to generate a model of S6 kinase domain, based on the crystal structure of PKB/Akt. A model was then generated using MODELLER7 (Sali & Blundell (1993) J. Mol. Biol. 234, 779-815) with standard comparative modelling parameters. Solvent accessibilities were then calculated using the method of Lee & Richards (1971) J. Mol. Biol. 55, 379-400 as implemented in NACCESS (Hubbard and Thornton, unpublished). In the model, the kinase domain

adopts the bilobal architecture characteristic of other members of the protein kinase family (FIG. **5**). The P loop is located in the cleft between N- the C-terminal lobes of the kinase domain. In the model, K99 and K104 are shown in blue and red respectively and the rest of the loop in green. Solvent accessibility analysis clearly indicates that both lysines are well accessible for posttranslational modifications, such as acetylation.

[0231] Interestingly, sequence analysis of the B-RAF gene associated with human cancers has identified numerous single site missense mutations, mostly within the kinase domain. Most of identified mutations were found in two regions: the glycine-rich P loop of the N lobe and the activation segment. Structure analysis of kinase domains from B-RAF and insulin receptor in their inactive state showed that the P loop and the activation segment form an array of hydrophobic interactions. These interactions lock the kinase in the inactive conformation, which is not appropriate for the phosphotransfer reaction. Therefore, any posttranslational modifications in the P loop might affect the interaction between P loop and the activation segment.

Example 7

[0232] This example demonstrates that deacetylase inhibitors exhibit a strong inhibitory effect on S6K activity in a dose- and time-dependent manner. Since phosphorylation of multiple residues in S6 kinases has a potent effect on their kinase activity, it was interesting to see whether acetylation, an alternative post-translational modification, may also modulate the S6 kinase enzyme activity. Although kinase acetylation has not been described to date, there are examples of enzymes whose activity is regulated by p300-mediated acetylation, including Fen 1 (Hasan et al. (2001) Mol. Cell. 7, 1221-1231) and DNA polymerase beta (Hasan et al. (2002) Mol. Cell. 10, 1213-1222). In order to test whether S6 kinase acetylation may affect its kinase activity in vivo, C2C12 myoblasts were treated with a dose course of TSA. S6 kinase activation was analysed by western blotting with an antibody specific for phospho-threonine 412 and also with an antibody against the phosphorylated form of rpS6. FIG. **6**A clearly shows that at doses of 500 nM TSA and above, S6 kinase activation is potently inhibited, as measured by both a reduction in phospho-threonine 412 and S6 protein phosphorylation. In order to further investigate this effect, time course experiments were conducted in which C2C12 cells were treated with 1 μM TSA in combination with 10 mM nicotinamide for various time periods prior to analysis of S6 kinase activity. In this case, S6K1 and 2 were immunoprecipitated from the cell lysates and subjected to immunocomplex kinase assays using purified ribosomes as substrate. Lysates were again blotted for pT412 and prpS6. FIG. **6**B shows that treatment of C2C12 cells with a combination of the class I/II HDAC inhibitor, TSA and the SIRT1 deacetylase inhibitor, nicotinamide, causes a significant time-dependent reduction in S6 kinase activity as assessed by in vitro kinase assay and by western blotting with phospho-specific antibodies. Interestingly, the activation status of an upstream pathway member, PKB, as analysed by an antibody specific for phospho-Foxo3a (a well defined PKB substrate) appears to be unaffected by TSA/nicotinamide treatment (data not shown). This would suggest that TSA affects the S6 kinase pathway downstream of PKB thus reducing the likelihood that receptor tyrosine kinases, PI3K, PTEN, PDK-1, or indeed PKB itself, are the targets through which the TSA effect is mediated.

Example 8

[0233] Histone deacetylase inhibitors downregulate S6K activity in a range of different cell types. In order to test the model that S6K acetylation may be playing a role in its de-activation, the effect of TSA on S6K activity was analysed in a number of cell types, some of which contain mutations in the pathways leading to S6K activation, thus affording a genetic analysis of this phenomenon. FIG. **6**A shows that S6K activity was inhibited (to a greater or lesser extent) by deacetylase inhibition in all cell types tested. Interestingly, the effect was retained in PC-3 cells. PC-3 is a PTEN-null prostate cancer cell line in which basal PKB and S6K activation is high due to deregulation of PI3K signalling. The fact that S6K is inhibited by TSA treatment of these cells suggests that this effect is exerted downstream of PTEN, i.e. the inhibitors do not act by affecting a receptor tyrosine kinase or PI(3)K. Furthermore, S6K activity is affected by deacetylase treatment of MKOC cells (FIG. **6**B). The MKOC cell line was derived from a renal tumour removed from a $Tsc2^{-/-}$ mouse. Since these cells lack Tsc2, S6K activity is uncoupled from both PKB and the tuberous sclerosis complex heterodimer. The fact that TSA causes reduction of S6K activity in these cells suggests that mTOR or S6K itself may well be the point at which deacetylase inhibition is acting. Blotting with an antibody against phospho-4EBP1, the other major target of mTOR revealed that its phosphorylation status is unaffected by TSA treatment, suggesting that S6K itself and not mTOR, is affected by deacetylase inhibition (FIG. **6**B). It should be noted that, although the same trend (reduced S6K phosphorylation and activity following treatment with deacetylase inhibitors) is seen in all the cell lines tested, the magnitude of this effect does vary. The relatively minor decrease in rpS6 phosphorylation status in MCF-7 cells and MKOC cells may be indicative of another functional, 'deacetylase inhibitor-insensitive' S6 kinase, in these cells, possibly S6K2. Indeed, MCF-7 cells express high levels of S6K2. Significantly, phosphorylation of S6K1 on T412 and S6K1 activity as measured by immune complex kinase assay is reduced in MKOC cells upon treatment with HDAC inhibitors (FIG. **6**B).

[0234] These observations suggest two obvious explanations. Firstly, that TSA is not a specific inhibitor of HDAC activity but that it is also a direct S6K inhibitor in the classical sense. The second explanation is that through inhibiting HDACs, TSA causes accumulation of an acetylated species of S6K that is catalytically inactive. In order to test the former explanation, that TSA is a direct S6K inhibitor, in vitro kinase assays were conducted using recombinant purified S6K in the presence or absence of TSA and nicotinamide (FIG. **6**C). This experiment does not indicate that TSA inhibits S6K activity in vitro and so is unlikely to be acting as a direct kinase inhibitor.

Example 9

[0235] Analysis of endogenous S6K1 acetylation during the cell cycle. In order to learn more about the potential role of S6K acetylation, we sought to determine the physiological conditions under which it becomes acetylated. Since stimulation of quiescent cells induces complex formation between S6K1 and p300, it was pertinent to establish whether this

event results in S6K1 acetylation. NIH-3T3 cells were used for this analysis, since, in our experience, they are amenable to synchronisation and express relatively high levels of S6K1. For synchronisation, the cells were grown to 50% confluence and the medium was replaced with DMEM containing 0.5% serum for a starvation period of 24 hours. Cells were then stimulated by re-addition of DMEM/10% serum for various time periods. Time points were taken up to 24 hours in order to analyse how S6K1 acetylation may be regulated throughout the cell cycle. S6K1 was immunoprecipitated from lysates using a monoclonal antibody and immune complexes were washed extensively prior to SDS-PAGE. Acetylated S6K1 was detected by western blotting with anti-acetyl lysine antibodies as described in chapter four. Blots were then stripped and reprobed with polyclonal antibodies against S6K1. FIG. **6** shows that in the absence of serum, no S6K1 acetylation was detected. Note that no interaction between S6Ks and acetyltransferases p300 or P/CAF was detected in starved MCF-7 cells (FIG. **1**A). S6K1 acetylation is clearly detected one hour post-stimulation and the acetylation signal appears to increase towards the later stages of the cell cycle, particularly between 17 hours and 21 hours. This correlates with the end of S-phase and the passage through G_2 to M-phase, as determined by FACS analysis of these cells (data not shown). Interestingly, speckled nuclear staining of S6K1 has been reported during late S-phase in Swiss-3T3 cells, suggesting a possible association with chromatin at this time. The lysates from these cells were subjected to western blotting with antibodies against phospho-T412 (S6K) and phospho-rpS6, to assess the S6K activity at each time point. The profile obtained (FIG. **7**, lower panels) matches that seen for S6K activity throughout the cell cycle of a related mouse fibroblast cell line, Swiss 3T3: S6K1 pT412 phosphorylation and rpS6 phosphorylation peak 30 minutes to one hour following serum stimulation and gradually fall to around 40% of the maximal activity. It appears then, that S6K acetylation peaks at a time when phosphorylation and activity are low.

Example 10

[0236] Testing the specificities of antibodies directed against identified acetylated lysines. Acetyl-lysine specific antibodies are a powerful tool in investigating the physiological importance of protein acetylations. A number of highly specific acetyl-lysine site-directed antibodies have been raised and used successfully to study the physiological relevance of these modifications in histones, transcription factors and signaling molecules. We therefore attempted the production of a panel of antibodies that would specifically recognize acetylated forms of S6K1 at K99, K104 and K516. The antibodies were raised in rabbits by injecting short synthetic peptides containing acetylated lysines in the middle of the peptide chain. Produced antibodies were affinity purified in two steps. Initially, Affigel beads coupled with antigenic peptides were used as matrices for purifying generated antibodies. Then, in order to remove IgGs which are directed to non-acetylated peptides, the antibodies obtained in the first step of purification were loaded onto the column containing corresponding non-acetylated peptides. The non-bound fractions were used for testing the specificity of generated antibodies. We have purified very small amount of antibodies against Ac-K104 peptide, so they were not included in the specificity test. The wtS6K1 or corresponding mutants (K99R and K516R) were transiently overexpressed in Hek293 cells. Cells transfected with vector alone were used as a negative control. Total cell lysates or immunoprecipitated recombinant S6Ks were resolved by SDS-PAGE and probed with Ac-K99 or Ac-K516 antibodies. As shown in FIG. **8**, affinity purified Ac-K516 antibodies specifically recognized wtS6K1 in immune complexes and in total lysates, but not the mutant K99R form of the kinase. These data clearly indicate that generated antibodies are highly specific towards the acetylated S6K1 at K516 and could be used on total cell and tissue lysates. As shown in FIG. **9**, Ac-K99 antibodies have to be further purified, since they recognize not only wtS6K, but also K99R mutant (to a lesser extent).

Example 11

[0237] Endogenous S6K1 is acetylated at K516 in serum-stimulated cells. The availability of highly specific Ac-K516 antibodies has allowed us to study the acetylation status of S6K1 at K516 in response to mitogenic stimuli and in cells treated with inhibitors of major signal transduction pathways. Initially, acetylation of K516 in starved and serum stimulated NIH 3T3 cells was tested. In this study, NIH 3T3 cells were starved for 24 hrs and stimulated with serum for various period of time. Total cell lysates were separated by SDS-PAGE and immunoblotted with anti-actin and Ac-K516 antibodies. The results presented in FIG. **10**A clearly indicate that acetylation at K516 does not occur in starved cells, but is strongly observed after serum stimulation.

[0238] Furthermore, the effect of signal transduction inhibitors on K516 acetylation was studied in NIH 3T3 cell. In this experiment, cells were serum starved and treated with rapamycin, LY and PD for 30 minutes before serum stimulation. Cells were lysed 1 hr hour after stimulation and the lysates separated by SDS-PAGE and immunoblotted with Ac-K516, pT412 antibodies. As shown in FIG. **10**B, acetylation of S6K1 at K516 was induced by serum stimulation, but not affected by any of signal transduction inhibitors used in this study. It is important to note that Rapamycin and LY294002 strongly inhibited phosphorylation of S6K1 at T412 and ribosomal protein S6 at T244. These data suggest that acetylation of S6K1 at K516 does not affect its kinase activity.

[0239] Mutational analysis of this site further supported this hypothesis. In this study, wtS6K1 and K516R mutant were expressed in Hek293 cells. Cells transfected with vector alone were used as a control. Transiently expressed S6Ks were immunoprecipitated with anti-EE tag antibodies and the immune complexes together with the total cell lysates were separated by SDS-PAGE. The Western blot analysis showed no significant changes in the level of T412 phosphorylation between the wild type and mutant form of S6K1 (FIG. **10**C). The level of T412 signal was standardized to the level of actin and recombinant S6K in each sample.

Example 12

[0240] S6K2 possesses a putative AT-hook DNA Binding Motif. Screening of the NCBI protein sequence databases with short segments of the S6K2 revealed the presence of a sequence with high homology to a DNA binding motif known as the AT hook. This motif is found at the extreme C-terminus of S6K2 (residues 487 to 495), adjacent to the nuclear localization sequence. The AT hook has been described in only a few proteins, the canonical example being the high mobility group protein HMG1A (formerly HMGI(Y) (reviewed in

Reeves (2001) Gene 277, 63-81). The AT hook is a short DNA binding motif that binds to the minor groove of short (4 to 6 base pairs) 'AT-rich' DNA.

[0241] Table 2 shows the alignment of AT hook sequences from several cellular proteins, including S6K2. The AT hook is a non-classical DNA binding motif found in a growing number of proteins with a variety of functions, including transcription factors and other transcriptional regulators, protein kinases and phosphatases. AT hook-containing proteins are known to play a number of roles, including the repression or potentiation of specific gene transcription via binding to gene promoter/enhancer elements (Reeves et al. (2001). Among the several reported roles for HMGA1 is participation in the enhanceosome, which controls expression of the interferon-β gene (Thanos and Maniatis (1995) Cell 83, 1101-1111). Another example is the recently described AKNA transcription factor, which controls the expression of CD40 and CD40 ligand in B-lymphocytes (Siddiqa et al. (2001) Nature 410, 383-387). Other functions include the initiation of DNA replication in yeast via binding at origins of replication, maintenance of heterochromatin and signal transduction to the nucleus.

TABLE 2

Alignment of A-T hook sequences

Protein	AT-hook sequence	Protein function
AT-hook consensus	KRK**RGRP**RK	
S6K2	KRG**RGRP**GR	Ser/Thr kinase
HMGA	KRG**RGRP**RG	Transcription factor
D-1	-KG**RGRP**KK	Unclear
AKNA	-RT**RGRP**AD	Transcription factor
HP2	KRK**RGRP**RK	Gene silencing
Orp4p	KRG**RGRP**RL	*S. pombe* DNA replication
DNA-PK	KRV**RGRP**CL	DNA repair
VRK1	KKS**RGRP**KK	Ser/Thr kinase
PSK-C3	KRR**RGRP**PK	Phosphorylase kinase
Methyltransferase 8	KRG**RGRP**KG	Histone H3 methylation
MeCP-2 protein	GRG**RGRP**KG	Methyl-CpG-binding protein
Zinc finger prot 278	KRG**RGRP**RK	Transcription factor

Example 13

[0242] The AT-hook motif is evolutionary conserved in S6K2. To gain the insight into functional importance of S6K2 AT-hook motif, it was examined to see whether the motif is conserved in evolution. The alignment of the S6K2 C-terminal protein sequences from various species clearly indicated that the AT-hook motif is highly conserved (see Table 3). The central RGRP motif is flanked with characteristic basic residues, which contributes to DNA binding. Three AT-hook subtypes have been described, all bearing the central RGRP motif. The S6K2 sequence is very similar to the type II AT-hook consensus.

TABLE 3

Evolutionary conservation of the AT hook motif in S6K2

Human S6K2	KRG**RGRP**GR
Rat S6K2	KKG**RGRP**GR
Bovine S6K2	KRG**RGRP**GR
Pig S6K2	KRG**RGRP**GR
Dog S6K2	KRG**RGRP**GR

Example 14

[0243] The C-terminal domain of S6K2 possesses DNA binding ability. Since a putative DNA binding motif was identified in the S6K2 C-terminal domain, it was pertinent to establish whether this motif may function in a simple biochemical assay. Two approaches were adopted in order to assess the DNA binding activity of S6K2: an on-beads DNA binding assay in which the DNA binding activity of immobilised recombinant protein to radiolabelled DNA is measured by scintillation counting and also an electrophoretic mobility shift assay (EMSA), in which binding of recombinant protein to DNA was measured by retardation of the DNA mobility in a non-denaturing gel, relative to free DNA control. Since the putative DNA binding activity is found at the S6K2 C-terminus, initial experiments were carried out using recombinant 6×His-tagged S6K2 C-terminal domain purified after expression in bacteria (hereafter referred to as S6K2CT). The C-terminal domain of S6K1 (S6K1 CT) was also included in these experiments as a negative control. It can be seen from FIG. **13** that the S6K2 C-terminal domain binds to DNA in vitro and that background binding of DNA to nickel-NTA agarose beads or to the S6K1 C-terminus was very low by comparison.

Example 15

[0244] The full-length S6K2 but not S6K1 binds to DNA in vitro. Having established that the C-terminal domain of S6K2 can bind directly to DNA in vitro, it was next investigated whether this DNA binding activity is retained in the context of the full-length protein. Purified EE-tagged S6Ks derived from baculoviral expression in sf9 cells were assayed for their DNA binding activity using the methods described for the C-terminal domains earlier. Once again, S6K1 was used as a negative control. FIG. **15**A shows that the DNA binding activity displayed by S6K2CT appears to be preserved in the full length protein, while full length S6K1 appears unable to bind DNA. Furthermore, the DNA binding activity of S6K2 was able to be demonstrated using the EMSA approach. FIG. **14** clearly shows a shift in the DNA when it is pre-incubated with S6K2 relative to the free DNA control (compare lanes 1 and 4). This indicates that a protein-DNA complex is formed, thus increasing the size and decreasing the mobility of the labeled DNA. The specificity of this interaction is indicated by both the lack of a mobility shift when DNA is incubated with S6K1

(lane 3) and by the 'supershift' produced by addition of the anti-EE antibody to the DNA/S6K2 mixture (lane 5). In this case, the antibody binds to the epitope tag on S6K2, thus forming a tertiary complex that further decreases the mobility. As discussed earlier, this technique is more commonly applied to EMSA experiments in which a mixture of proteins (typically a nuclear extract or total cell lysate) is incubated with a specific DNA probe. In this case, a number of proteins in the mixture may be expected to bind the DNA, so a shift cannot be attributed to a particular protein. Here, addition of an antibody specific to the protein of interest is used to produce a supershift, indicating binding of the protein of interest to the DNA probe.

[0245] FIG. **14**B also shows binding of the positive control AT-hook protein, HMG1A to DNA. In this case, the protein-DNA complexes failed to enter the gel and remained in the well. This is effect may be overcome by reducing the amount of protein.

Example 16

[0246] The deletion of the S6K2 C-terminal domain abolishes DNA binding. The experiment described in FIG. **15** demonstrates that the S6K2 C-terminal domain is sufficient for DNA binding but it still needed to be determined whether this domain is required for DNA binding of the full-length protein. In order to investigate whether this is indeed the case, or whether there may be other DNA binding regions present in full length S6K2, a truncation mutant of S6K2 lacking the C-terminal domain (S6K2ΔCT) was assayed for its DNA binding activity. Truncation mutants of S6K1 and S6K2, lacking regulatory N- or C-terminus were made. These truncated mutants were fused with the EE tag epitope sequence at the N-terminus. EE-S6K1, EE-S6K2 or EE-S6K2ΔCT were expressed in HEK 293 cells by transient transfection and immunoprecipitated the recombinant proteins using anti-EE antibodies. The immunocomplexes were washed under stringent conditions before incubation with labeled DNA as described earlier. FIG. **15**A shows that the DNA binding activity of S6K2ΔCT is significantly lower than that of the full-length protein, suggesting that the DNA binding activity of S6K2 is indeed conferred by the C-terminal domain. Equal amounts of the wild type and S6K2ΔCT were used in the experiment as determined by western blotting of the cell lysates with EE monoclonal antibodies (FIG. **15**B) Note that even with the stringent washing conditions employed, it could be argued that full length S6K2 co-precipitates with a second DNA-binding protein that is not able to bind S6K2ΔCT. The fact that in vitro binding of highly purified S6K2 has been demonstrated would argue against this conclusion.

Example 17

[0247] Substitution of a conservative residue within S6K2 AT-hook core motif inhibits DNA binding. Having demonstrated that the DNA binding activity of S6K2 is contained within the C-terminal domain of the protein, it was interesting to establish whether this activity is due to the putative AT-hook identified by sequence homology. In order to test this hypothesis, it was decided to disrupt the S6K2 AT-hook by site directed mutagenesis of a residue highly conserved in all known AT-hooks. The critical determinant residues for AT-hook function are the Arg-Gly-Arg-Pro sequence. Binding of this motif to the minor groove of AT rich sequences has been studied in detail and the interactions between these residues and the DNA are known (Reeves and Nissen (1990) Biol. Chem. 265, 8573-8582). The N—H group on the amide bonds between the glycine and the second arginine residue in this core motif forms a hydrogen bond with the DNA phosphate backbone. In the absence of any published data on mutational analysis of AT-hook motifs, it was reasoned that replacement of this arginine with proline would disrupt the overall structure of this motif and also remove the N—H group from the amide bond. Site directed mutagenesis was carried out on the full-length S6K2 sequence in pcDNA3.1. The substitution of Arg 492 to Pro was confirmed by DNA sequencing. The mutant cDNA (termed S6K2 R492P) was transfected into HEK 293 cells in parallel with S6K2 wild type. The proteins were immunoprecipitated and subjected to DNA binding assays. FIG. **15**B shows that mutation of Arg 492 to proline results in a considerable loss of DNA binding activity. Taken together, these data suggest that S6K2 does indeed possess a functional AT-hook motif that bestows a non-specific DNA binding activity on the protein. Loss of the region encoding this motif, or direct disruption of the motif by site directed mutagenesis abrogates the DNA binding activity of S6K2.

Example 18

[0248] The binding of DNA to S6K2 stimulates its kinase activity. The regulatory N and C-terminal regions of S6K1 and S6K2 are known to regulate their kinase activities. Interaction between regulatory N-terminal acidic and C-terminal basic domains is thought to maintain S6K1 in an inactive conformation in unstimulated cells. In this inactive state, a pseudosubstrate region in the C terminus may hinder the catalytic site. Phosphorylation of the C-terminal proline-directed motifs is thought to disrupt the autoinhibitory interaction of the N- and C-terminal domains, exposing other critical regulatory sites in the linker and catalytic regions. It has been shown that deletion of the C-terminal domain of S6K2 enhances kinase activity whereas analogous deletion of S6K1 is inhibitory. The authors also found that the inhibitory effects of the S6K2 C-terminal domain are partly attributable to the nuclear localization signal and to the C-terminal proline-directed phosphorylation sites. The nuclear localization signal overlaps with the AT-hook DNA binding motif. Considering this, it could predicted that binding of DNA to the basic region of S6K2 C-terminal domain which possesses NLS and an AT-hook motif may affect the autoinhibitory interaction of the N- and C-terminal regions. To test this hypothesis, EE-tagged forms of S6K1 and S6K2 were expressed in insect cells using recombinant baculoviruses. Both kinases were purified nearly to homogeneity using affinity matrixes with EE monoclonal antibodies. Purified preparations of S6K1 and S6K2 were used in in vitro kinase in the presence or absence of sonicated genomic DNA. Ribosomal S6 protein was used as a substrate in the kinase assay. The level of S6 phosphorylation was measured by SDS-PAGE and phosphorimager. The level of S6 phosphorylation by S6K2 was found to be reproducibly higher in the presence of DNA. The presence of absence of DNA did not affect the activity of S6K1 towards ribosomal S6 protein. Dose-course stimulation of S6K2 activity by DNA was carried out. The S6 kinase reaction was carried out in the presence of various concentrations of DNA. In several independent experiments, a dose-dependent activation of S6K2 activity by DNA was found (FIG. **15**B). A 1.5 to 2 fold increase of S6K2 activity in the presence of 1 mg of DNA. It is an interesting observation, since only

few protein kinases are known to possess DNA-binding potential and more importantly to be activated by interaction with DNA. The best example is the DNA dependent protein kinase (DNA-PK) (Lees-Miller et al. (1990) Mol. Cell. Biol. 10, 6472-6481). Using a nuclear fractionation protocol, a small fraction of S6K2 in the nuclear matrix fraction was detected. This is consistent with an earlier report describing nuclear S6K2 isoform that is insoluble in standard detergents (Lee-Fruman et al. (1999) Oncogene 18, 5108-5114). Digestion of this fraction with DNase1 appeared to solubilize a portion of S6K2, providing an indication that is indeed associated with chromatin in cells. Since scaffold attachment, regions are known to be sites of transcriptional activity, containing highly acetylated histones (Davie (1997) Mol. Biol. Rep. 24, 197-207), localization of S6K2 to these regions would be consistent with a role in transcriptional regulation.

[0249] Based on the examples, it was found for the first time that S6K2, but not S6K1 possesses an AT-hook DNA binding motif. Furthermore, it was demonstrated that S6K2 AT-hook motif mediates specific association with DNA. The consequence of S6K2-DNA interaction is the induction of its kinase activity. This is the first time the presence of a functional AT-hook DNA binding motif in the family of protein kinases has been described. Furthermore, these findings provide new insight into the mechanism of S6K2 signaling in the nucleus. Based on these data, a novel mechanism by which S6K2 can transduce growth-promoting effects in response to mitogenes and nutrients is proposed. This may involve the regulation of transcription factors and/or chromatin remodeling proteins by phosphorylation, when S6K2 complexes with DNA and is activated by this interaction. It is important to note that this discovery has an important implication for the development of drugs, which can block S6K2-DNA interaction and subsequently its growth promoting functions, which are deregulated in cancer and diabetes.

[0250] While the invention has been described and illustrated herein by references to various specific materials, procedures and examples, it is understood that the invention is not restricted to the particular combinations of material and procedures selected for that purpose. Numerous variations of such details can be implied as will be appreciated by those skilled in the art. It is intended that the specification and examples be considered as exemplary, only, with the true scope and spirit of the invention being indicated by the following claims. All references, patents, and patent applications referred to in this application are herein incorporated by reference in their entirety.

SEQUENCE LISTING

```
<160> NUMBER OF SEQ ID NOS: 29

<210> SEQ ID NO 1
<211> LENGTH: 2346
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (28)..(1602)
<223> OTHER INFORMATION: p70(alpha) S6 Kinase gene

<400> SEQUENCE: 1

gcacgaggct gcggcgggtc cgggccc atg agg cga cga agg agg cgg gac ggc      54
                              Met Arg Arg Arg Arg Arg Arg Asp Gly
                              1               5

ttt tac cca gcc ccg gac ttc cga gac agg gaa gct gag gac atg gca      102
Phe Tyr Pro Ala Pro Asp Phe Arg Asp Arg Glu Ala Glu Asp Met Ala
10                  15                  20                  25

gga gtg ttt gac ata gac ctg gac cag cca gag gac gcg ggc tct gag      150
Gly Val Phe Asp Ile Asp Leu Asp Gln Pro Glu Asp Ala Gly Ser Glu
30                  35                  40

gat gag ctg gag gag ggg ggt cag tta aat gaa agc atg gac cat ggg      198
Asp Glu Leu Glu Glu Gly Gly Gln Leu Asn Glu Ser Met Asp His Gly
45                  50                  55

gga gtt gga cca tat gaa ctt ggc atg gaa cat tgt gag aaa ttt gaa      246
Gly Val Gly Pro Tyr Glu Leu Gly Met Glu His Cys Glu Lys Phe Glu
60                  65                  70

atc tca gaa act agt gtg aac aga ggg cca gaa aaa atc aga cca gaa      294
Ile Ser Glu Thr Ser Val Asn Arg Gly Pro Glu Lys Ile Arg Pro Glu
75                  80                  85

tgt ttt gag cta ctt cgg gta ctt ggt aaa ggg ggc tat gga aag gtt      342
Cys Phe Glu Leu Leu Arg Val Leu Gly Lys Gly Gly Tyr Gly Lys Val
90                  95                  100                 105

ttt caa gta cga aaa gta aca gga gca aat act ggg aaa ata ttt gcc      390
```

-continued

```
Phe Gln Val Arg Lys Val Thr Gly Ala Asn Thr Gly Lys Ile Phe Ala
110                 115                 120

atg aag gtg ctt aaa aag gca atg ata gta aga aat gct aaa gat aca      438
Met Lys Val Leu Lys Lys Ala Met Ile Val Arg Asn Ala Lys Asp Thr
125                 130                 135

gct cat aca aaa gca gaa cgg aat att ctg gag gaa gta aag cat ccc      486
Ala His Thr Lys Ala Glu Arg Asn Ile Leu Glu Glu Val Lys His Pro
140                 145                 150

ttc atc gtg gat tta att tat gcc ttt cag act ggt gga aaa ctc tac      534
Phe Ile Val Asp Leu Ile Tyr Ala Phe Gln Thr Gly Gly Lys Leu Tyr
155                 160                 165

ctc atc ctt gag tat ctc agt gga gga gaa cta ttt atg cag tta gaa      582
Leu Ile Leu Glu Tyr Leu Ser Gly Gly Glu Leu Phe Met Gln Leu Glu
170                 175                 180                 185

aga gag gga ata ttt atg gaa gac act gcc tgc ttt tac ttg gca gaa      630
Arg Glu Gly Ile Phe Met Glu Asp Thr Ala Cys Phe Tyr Leu Ala Glu
190                 195                 200

atc tcc atg gct ttg ggg cat tta cat caa aag ggg atc atc tac aga      678
Ile Ser Met Ala Leu Gly His Leu His Gln Lys Gly Ile Ile Tyr Arg
205                 210                 215

gac ctg aag ccg gag aat atc atg ctt aat cac caa ggt cat gtg aaa      726
Asp Leu Lys Pro Glu Asn Ile Met Leu Asn His Gln Gly His Val Lys
220                 225                 230

cta aca gac ttt gga cta tgc aaa gaa tct att cat gat gga aca gtc      774
Leu Thr Asp Phe Gly Leu Cys Lys Glu Ser Ile His Asp Gly Thr Val
235                 240                 245

aca cac aca ttt tgt gga aca ata gaa tac atg gcc cct gaa atc ttg      822
Thr His Thr Phe Cys Gly Thr Ile Glu Tyr Met Ala Pro Glu Ile Leu
250                 255                 260                 265

atg aga agt ggc cac aat cgt gct gtg gat tgg tgg agt ttg gga gca      870
Met Arg Ser Gly His Asn Arg Ala Val Asp Trp Trp Ser Leu Gly Ala
270                 275                 280

tta atg tat gac atg ctg act gga gca ccc cca ttc act ggg gag aat      918
Leu Met Tyr Asp Met Leu Thr Gly Ala Pro Pro Phe Thr Gly Glu Asn
285                 290                 295

aga aag aaa aca att gac aaa atc ctc aaa tgt aaa ctc aat ttg cct      966
Arg Lys Lys Thr Ile Asp Lys Ile Leu Lys Cys Lys Leu Asn Leu Pro
300                 305                 310

ccc tac ctc aca caa gaa gcc aga gat ctg ctt aaa aag ctg ctg aaa     1014
Pro Tyr Leu Thr Gln Glu Ala Arg Asp Leu Leu Lys Lys Leu Leu Lys
315                 320                 325

aga aat gct gct tct cgt ctg gga gct ggt cct ggg gac gct gga gaa     1062
Arg Asn Ala Ala Ser Arg Leu Gly Ala Gly Pro Gly Asp Ala Gly Glu
330                 335                 340                 345

gtt caa gct cat cca ttc ttt aga cac att aac tgg gaa gaa ctt ctg     1110
Val Gln Ala His Pro Phe Phe Arg His Ile Asn Trp Glu Glu Leu Leu
350                 355                 360

gct cga aag gtg gag ccc ccc ttt aaa cct ctg ttg caa tct gaa gag     1158
Ala Arg Lys Val Glu Pro Pro Phe Lys Pro Leu Leu Gln Ser Glu Glu
365                 370                 375

gat gta agt cag ttt gat tcc aag ttt aca cgt cag aca cct gtc gac     1206
Asp Val Ser Gln Phe Asp Ser Lys Phe Thr Arg Gln Thr Pro Val Asp
380                 385                 390

agc cca gat gac tca act ctc agt gaa agt gcc aat cag gtc ttt ctg     1254
Ser Pro Asp Asp Ser Thr Leu Ser Glu Ser Ala Asn Gln Val Phe Leu
395                 400                 405

ggt ttt aca tat gtg gct cca tct gta ctt gaa agt gtg aaa gaa aag     1302
```

-continued

```
Gly Phe Thr Tyr Val Ala Pro Ser Val Leu Glu Ser Val Lys Glu Lys
410                 415                 420                 425

ttt tcc ttt gaa cca aaa atc cga tca cct cga aga ttt att ggc agc     1350
Phe Ser Phe Glu Pro Lys Ile Arg Ser Pro Arg Arg Phe Ile Gly Ser
430                 435                 440

cca cga aca cct gtc agc cca gtc aaa ttt tct cct ggg gat ttc tgg     1398
Pro Arg Thr Pro Val Ser Pro Val Lys Phe Ser Pro Gly Asp Phe Trp
445                 450                 455

gga aga ggt gct tcg gcc agc aca gca aat cct cag aca cct gtg gaa     1446
Gly Arg Gly Ala Ser Ala Ser Thr Ala Asn Pro Gln Thr Pro Val Glu
460                 465                 470

tac cca atg gaa aca agt ggc ata gag cag atg gat gtg aca atg agt     1494
Tyr Pro Met Glu Thr Ser Gly Ile Glu Gln Met Asp Val Thr Met Ser
475                 480                 485

ggg gaa gca tcg gca cca ctt cca ata cga cag ccg aac tct ggg cca     1542
Gly Glu Ala Ser Ala Pro Leu Pro Ile Arg Gln Pro Asn Ser Gly Pro
490                 495                 500                 505

tac aaa aaa caa gct ttt ccc atg atc tcc aaa cgg cca gag cac ctg     1590
Tyr Lys Lys Gln Ala Phe Pro Met Ile Ser Lys Arg Pro Glu His Leu
510                 515                 520

cgt atg aat cta tgacagagca atgcttttaa tgaatttaag gcaaaaaggt         1642
Arg Met Asn Leu
525

ggagagggag atgtgtgagc atcctgcaag gtgaaacaag actcaaaatg acagtttcag   1702

agagtcaatg tcattacata gaacacttcg gacacaggaa aataaacgt ggattttaaa     1762

aaatcaatca atggtgcaaa aaaaaactta aagcaaaata gtattgctga actcttaggc   1822

acatcaatta attgattcct cgcgacatct ttctcaacct tatcaaggat tttcatgttg   1882

atgactcgaa actgacagta ttaagggtag gatgttgctc tgaatcactg tgagtctgat   1942

gtgtgaagaa gggtatcctt tcattaggca agtacaaatt gcctataata cttgcaacta   2002

aggacaaatt agcatgcaag cttggtcaaa cttttcccag gcaaaatggg aaggcaaaga   2062

caaaagaaac ttaccaattg atgttttacg tgcaaacaac ctgaatcttt tttttatata   2122

aatatatatt tttcaaatag atttttgatt cagctcatta tgaaaaacat cccaaacttt   2182

aaaatgcgaa attattggtt ggtgtgaaga aagccagaca acttctgttt cttctcttgg   2242

tgaaataata aaatgcaaat gaatcattgt taacacagct gtggctcgtt tgagggattg   2302

gggtggacct ggggtttatt ttcagtaacc cagctgcgga gcct                     2346


<210> SEQ ID NO 2
<211> LENGTH: 525
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 2

Met Arg Arg Arg Arg Arg Arg Asp Gly Phe Tyr Pro Ala Pro Asp Phe
1               5                   10                  15

Arg Asp Arg Glu Ala Glu Asp Met Ala Gly Val Phe Asp Ile Asp Leu
20                  25                  30

Asp Gln Pro Glu Asp Ala Gly Ser Glu Asp Glu Leu Glu Glu Gly Gly
35                  40                  45

Gln Leu Asn Glu Ser Met Asp His Gly Gly Val Gly Pro Tyr Glu Leu
50                  55                  60

Gly Met Glu His Cys Glu Lys Phe Glu Ile Ser Glu Thr Ser Val Asn
```

-continued

```
65                  70                  75                  80

Arg Gly Pro Glu Lys Ile Arg Pro Glu Cys Phe Glu Leu Leu Arg Val
85                  90                  95

Leu Gly Lys Gly Gly Tyr Gly Lys Val Phe Gln Val Arg Lys Val Thr
100                 105                 110

Gly Ala Asn Thr Gly Lys Ile Phe Ala Met Lys Val Leu Lys Lys Ala
115                 120                 125

Met Ile Val Arg Asn Ala Lys Asp Thr Ala His Thr Lys Ala Glu Arg
130                 135                 140

Asn Ile Leu Glu Glu Val Lys His Pro Phe Ile Val Asp Leu Ile Tyr
145                 150                 155                 160

Ala Phe Gln Thr Gly Gly Lys Leu Tyr Leu Ile Leu Glu Tyr Leu Ser
165                 170                 175

Gly Gly Glu Leu Phe Met Gln Leu Glu Arg Glu Gly Ile Phe Met Glu
180                 185                 190

Asp Thr Ala Cys Phe Tyr Leu Ala Glu Ile Ser Met Ala Leu Gly His
195                 200                 205

Leu His Gln Lys Gly Ile Ile Tyr Arg Asp Leu Lys Pro Glu Asn Ile
210                 215                 220

Met Leu Asn His Gln Gly His Val Lys Leu Thr Asp Phe Gly Leu Cys
225                 230                 235                 240

Lys Glu Ser Ile His Asp Gly Thr Val Thr His Thr Phe Cys Gly Thr
245                 250                 255

Ile Glu Tyr Met Ala Pro Glu Ile Leu Met Arg Ser Gly His Asn Arg
260                 265                 270

Ala Val Asp Trp Trp Ser Leu Gly Ala Leu Met Tyr Asp Met Leu Thr
275                 280                 285

Gly Ala Pro Pro Phe Thr Gly Glu Asn Arg Lys Lys Thr Ile Asp Lys
290                 295                 300

Ile Leu Lys Cys Lys Leu Asn Leu Pro Pro Tyr Leu Thr Gln Glu Ala
305                 310                 315                 320

Arg Asp Leu Leu Lys Lys Leu Leu Lys Arg Asn Ala Ala Ser Arg Leu
325                 330                 335

Gly Ala Gly Pro Gly Asp Ala Gly Glu Val Gln Ala His Pro Phe Phe
340                 345                 350

Arg His Ile Asn Trp Glu Glu Leu Leu Ala Arg Lys Val Glu Pro Pro
355                 360                 365

Phe Lys Pro Leu Leu Gln Ser Glu Glu Asp Val Ser Gln Phe Asp Ser
370                 375                 380

Lys Phe Thr Arg Gln Thr Pro Val Asp Ser Pro Asp Asp Ser Thr Leu
385                 390                 395                 400

Ser Glu Ser Ala Asn Gln Val Phe Leu Gly Phe Thr Tyr Val Ala Pro
405                 410                 415

Ser Val Leu Glu Ser Val Lys Glu Lys Phe Ser Phe Glu Pro Lys Ile
420                 425                 430

Arg Ser Pro Arg Arg Phe Ile Gly Ser Pro Arg Thr Pro Val Ser Pro
435                 440                 445

Val Lys Phe Ser Pro Gly Asp Phe Trp Gly Arg Gly Ala Ser Ala Ser
450                 455                 460

Thr Ala Asn Pro Gln Thr Pro Val Glu Tyr Pro Met Glu Thr Ser Gly
465                 470                 475                 480
```

-continued

```
Ile Glu Gln Met Asp Val Thr Met Ser Gly Glu Ala Ser Ala Pro Leu
485                 490                 495

Pro Ile Arg Gln Pro Asn Ser Gly Pro Tyr Lys Lys Gln Ala Phe Pro
500                 505                 510

Met Ile Ser Lys Arg Pro Glu His Leu Arg Met Asn Leu
515                 520                 525


<210> SEQ ID NO 3
<211> LENGTH: 8765
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (403)..(7647)
<223> OTHER INFORMATION: p300 acetyltransferase

<400> SEQUENCE: 3

ggacagcgcc gaggaggaag aggttgatgg cggcggcgga gctccgagag acctcggctg      60

ggcaggggcc ggccgtggcg ggccggggac tgcgcctcta gagccgcgag ttctcgggaa     120

ttcgccgcag cggacgcgct cggcgaattt gtgctcttgt gccctcctcc gggcttgggc     180

ccaggcccgg cccctcgcac ttgcccttac cttttctatc gagtccgcat ccctctccag     240

ccactgcgac ccggcgaaga gaaaaaggaa cttcccccac cccctcgggt gccgtcggag     300

ccccccagcc caccccctggg tgcggcgcgg ggaccccggg ccgaagaaga gatttcctga     360

ggattctggt ttcctcgct tgtatctccg aaagaattaa aa atg gcc gag aat        414
                                              Met Ala Glu Asn
                                              1

gtg gtg gaa ccg ggg ccg cct tca gcc aag cgg cct aaa ctc tca tct      462
Val Val Glu Pro Gly Pro Pro Ser Ala Lys Arg Pro Lys Leu Ser Ser
5                   10                  15                  20

ccg gcc ctc tcg gcg tcc gcc agc gat ggc aca gat ttt ggc tct cta      510
Pro Ala Leu Ser Ala Ser Ala Ser Asp Gly Thr Asp Phe Gly Ser Leu
25                  30                  35

ttt gac ttg gag cac gac tta cca gat gaa tta atc aac tct aca gaa      558
Phe Asp Leu Glu His Asp Leu Pro Asp Glu Leu Ile Asn Ser Thr Glu
40                  45                  50

ttg gga cta acc aat ggt ggt gat att aat cag ctt cag aca agt ctt      606
Leu Gly Leu Thr Asn Gly Gly Asp Ile Asn Gln Leu Gln Thr Ser Leu
55                  60                  65

ggc atg gta caa gat gca gct tct aaa cat aaa cag ctg tca gaa ttg      654
Gly Met Val Gln Asp Ala Ala Ser Lys His Lys Gln Leu Ser Glu Leu
70                  75                  80

ctg cga tct ggt agt tcc cct aac ctc aat atg gga gtt ggt ggc cca      702
Leu Arg Ser Gly Ser Ser Pro Asn Leu Asn Met Gly Val Gly Gly Pro
85                  90                  95                  100

ggt caa gtc atg gcc agc cag gcc caa cag agc agt cct gga tta ggt      750
Gly Gln Val Met Ala Ser Gln Ala Gln Gln Ser Ser Pro Gly Leu Gly
105                 110                 115

ttg ata aat agc atg gtc aaa agc cca atg aca cag gca ggc ttg act      798
Leu Ile Asn Ser Met Val Lys Ser Pro Met Thr Gln Ala Gly Leu Thr
120                 125                 130

tct ccc aac atg ggg atg ggc act agt gga cca aat cag ggt cct acg      846
Ser Pro Asn Met Gly Met Gly Thr Ser Gly Pro Asn Gln Gly Pro Thr
135                 140                 145

cag tca aca ggt atg atg aac agt cca gta aat cag cct gcc atg gga      894
Gln Ser Thr Gly Met Met Asn Ser Pro Val Asn Gln Pro Ala Met Gly
```

-continued

```
150                 155                 160

atg aac aca ggg atg aat gcg ggc atg aat cct gga atg ttg gct gca      942
Met Asn Thr Gly Met Asn Ala Gly Met Asn Pro Gly Met Leu Ala Ala
165                 170                 175                 180

ggc aat gga caa ggg ata atg cct aat caa gtc atg aac ggt tca att      990
Gly Asn Gly Gln Gly Ile Met Pro Asn Gln Val Met Asn Gly Ser Ile
185                 190                 195

gga gca ggc cga ggg cga cag aat atg cag tac cca aac cca ggc atg     1038
Gly Ala Gly Arg Gly Arg Gln Asn Met Gln Tyr Pro Asn Pro Gly Met
200                 205                 210

gga agt gct ggc aac tta ctg act gag cct ctt cag cag ggc tct ccc     1086
Gly Ser Ala Gly Asn Leu Leu Thr Glu Pro Leu Gln Gln Gly Ser Pro
215                 220                 225

cag atg gga gga caa aca gga ttg aga ggc ccc cag cct ctt aag atg     1134
Gln Met Gly Gly Gln Thr Gly Leu Arg Gly Pro Gln Pro Leu Lys Met
230                 235                 240

gga atg atg aac aac ccc aat cct tat ggt tca cca tat act cag aat     1182
Gly Met Met Asn Asn Pro Asn Pro Tyr Gly Ser Pro Tyr Thr Gln Asn
245                 250                 255                 260

cct gga cag cag att gga gcc agt ggc ctt ggt ctc cag att cag aca     1230
Pro Gly Gln Gln Ile Gly Ala Ser Gly Leu Gly Leu Gln Ile Gln Thr
265                 270                 275

aaa act gta cta tca aat aac tta tct cca ttt gct atg gac aaa aag     1278
Lys Thr Val Leu Ser Asn Asn Leu Ser Pro Phe Ala Met Asp Lys Lys
280                 285                 290

gca gtt cct ggt gga gga atg ccc aac atg ggt caa cag cca gcc ccg     1326
Ala Val Pro Gly Gly Gly Met Pro Asn Met Gly Gln Gln Pro Ala Pro
295                 300                 305

cag gtc cag cag cca ggc ctg gtg act cca gtt gcc caa ggg atg ggt     1374
Gln Val Gln Gln Pro Gly Leu Val Thr Pro Val Ala Gln Gly Met Gly
310                 315                 320

tct gga gca cat aca gct gat cca gag aag cgc aag ctc atc cag cag     1422
Ser Gly Ala His Thr Ala Asp Pro Glu Lys Arg Lys Leu Ile Gln Gln
325                 330                 335                 340

cag ctt gtt ctc ctt ttg cat gct cac aag tgc cag cgc cgg gaa cag     1470
Gln Leu Val Leu Leu Leu His Ala His Lys Cys Gln Arg Arg Glu Gln
345                 350                 355

gcc aat ggg gaa gtg agg cag tgc aac ctt ccc cac tgt cgc aca atg     1518
Ala Asn Gly Glu Val Arg Gln Cys Asn Leu Pro His Cys Arg Thr Met
360                 365                 370

aag aat gtc cta aac cac atg aca cac tgc cag tca ggc aag tct tgc     1566
Lys Asn Val Leu Asn His Met Thr His Cys Gln Ser Gly Lys Ser Cys
375                 380                 385

caa gtg gca cac tgt gca tct tct cga caa atc att tca cac tgg aag     1614
Gln Val Ala His Cys Ala Ser Ser Arg Gln Ile Ile Ser His Trp Lys
390                 395                 400

aat tgt aca aga cat gat tgt cct gtg tgt ctc ccc ctc aaa aat gct     1662
Asn Cys Thr Arg His Asp Cys Pro Val Cys Leu Pro Leu Lys Asn Ala
405                 410                 415                 420

ggt gat aag aga aat caa cag cca att ttg act gga gca ccc gtt gga     1710
Gly Asp Lys Arg Asn Gln Gln Pro Ile Leu Thr Gly Ala Pro Val Gly
425                 430                 435

ctt gga aat cct agc tct cta ggg gtg ggt caa cag tct gcc ccc aac     1758
Leu Gly Asn Pro Ser Ser Leu Gly Val Gly Gln Gln Ser Ala Pro Asn
440                 445                 450

cta agc act gtt agt cag att gat ccc agc tcc ata gaa aga gcc tat     1806
Leu Ser Thr Val Ser Gln Ile Asp Pro Ser Ser Ile Glu Arg Ala Tyr
```

-continued

```
455                 460                 465

gca gct ctt gga cta ccc tat caa gta aat cag atg ccg aca caa ccc      1854
Ala Ala Leu Gly Leu Pro Tyr Gln Val Asn Gln Met Pro Thr Gln Pro
470                 475                 480

cag gtg caa gca aag aac cag cag aat cag cag cct ggg cag tct ccc      1902
Gln Val Gln Ala Lys Asn Gln Gln Asn Gln Gln Pro Gly Gln Ser Pro
485                 490                 495                 500

caa ggc atg cgg ccc atg agc aac atg agt gct agt cct atg gga gta      1950
Gln Gly Met Arg Pro Met Ser Asn Met Ser Ala Ser Pro Met Gly Val
505                 510                 515

aat gga ggt gta gga gtt caa acg ccg agt ctt ctt tct gac tca atg      1998
Asn Gly Gly Val Gly Val Gln Thr Pro Ser Leu Leu Ser Asp Ser Met
520                 525                 530

ttg cat tca gcc ata aat tct caa aac cca atg atg agt gaa aat gcc      2046
Leu His Ser Ala Ile Asn Ser Gln Asn Pro Met Met Ser Glu Asn Ala
535                 540                 545

agt gtg ccc tcc ctg ggt cct atg cca aca gca gct caa cca tcc act      2094
Ser Val Pro Ser Leu Gly Pro Met Pro Thr Ala Ala Gln Pro Ser Thr
550                 555                 560

act gga att cgg aaa cag tgg cac gaa gat att act cag gat ctt cga      2142
Thr Gly Ile Arg Lys Gln Trp His Glu Asp Ile Thr Gln Asp Leu Arg
565                 570                 575                 580

aat cat ctt gtt cac aaa ctc gtc caa gcc ata ttt cct acg ccg gat      2190
Asn His Leu Val His Lys Leu Val Gln Ala Ile Phe Pro Thr Pro Asp
585                 590                 595

cct gct gct tta aaa gac aga cgg atg gaa aac cta gtt gca tat gct      2238
Pro Ala Ala Leu Lys Asp Arg Arg Met Glu Asn Leu Val Ala Tyr Ala
600                 605                 610

cgg aaa gtt gaa ggg gac atg tat gaa tct gca aac aat cga gcg gaa      2286
Arg Lys Val Glu Gly Asp Met Tyr Glu Ser Ala Asn Asn Arg Ala Glu
615                 620                 625

tac tac cac ctt cta gct gag aaa atc tat aag atc cag aaa gaa cta      2334
Tyr Tyr His Leu Leu Ala Glu Lys Ile Tyr Lys Ile Gln Lys Glu Leu
630                 635                 640

gaa gaa aaa cga agg acc aga cta cag aag cag aac atg cta cca aat      2382
Glu Glu Lys Arg Arg Thr Arg Leu Gln Lys Gln Asn Met Leu Pro Asn
645                 650                 655                 660

gct gca ggc atg gtt cca gtt tcc atg aat cca ggg cct aac atg gga      2430
Ala Ala Gly Met Val Pro Val Ser Met Asn Pro Gly Pro Asn Met Gly
665                 670                 675

cag ccg caa cca gga atg act tct aat ggc cct cta cct gac cca agt      2478
Gln Pro Gln Pro Gly Met Thr Ser Asn Gly Pro Leu Pro Asp Pro Ser
680                 685                 690

atg atc cgt ggc agt gtg cca aac cag atg atg cct cga ata act cca      2526
Met Ile Arg Gly Ser Val Pro Asn Gln Met Met Pro Arg Ile Thr Pro
695                 700                 705

caa tct ggt ttg aat caa ttt ggc cag atg agc atg gcc cag ccc cct      2574
Gln Ser Gly Leu Asn Gln Phe Gly Gln Met Ser Met Ala Gln Pro Pro
710                 715                 720

att gta ccc cgg caa acc cct cct ctt cag cac cat gga cag ttg gct      2622
Ile Val Pro Arg Gln Thr Pro Pro Leu Gln His His Gly Gln Leu Ala
725                 730                 735                 740

caa cct gga gct ctc aac ccg cct atg ggc tat ggg cct cgt atg caa      2670
Gln Pro Gly Ala Leu Asn Pro Pro Met Gly Tyr Gly Pro Arg Met Gln
745                 750                 755

cag cct tcc aac cag ggc cag ttc ctt cct cag act cag ttc cca tca      2718
Gln Pro Ser Asn Gln Gly Gln Phe Leu Pro Gln Thr Gln Phe Pro Ser
```

-continued

```
760                 765                 770

cag gga atg aat gta aca aat atc cct ttg gct ccg tcc agc ggt caa      2766
Gln Gly Met Asn Val Thr Asn Ile Pro Leu Ala Pro Ser Ser Gly Gln
775                 780                 785

gct cca gtg tct caa gca caa atg tct agt tct tcc tgc ccg gtg aac      2814
Ala Pro Val Ser Gln Ala Gln Met Ser Ser Ser Ser Cys Pro Val Asn
790                 795                 800

tct cct ata atg cct cca ggg tct cag ggg agc cac att cac tgt ccc      2862
Ser Pro Ile Met Pro Pro Gly Ser Gln Gly Ser His Ile His Cys Pro
805                 810                 815                 820

cag ctt cct caa cca gct ctt cat cag aat tca ccc tcg cct gta cct      2910
Gln Leu Pro Gln Pro Ala Leu His Gln Asn Ser Pro Ser Pro Val Pro
825                 830                 835

agt cgt acc ccc acc cct cac cat act ccc cca agc ata ggg gct cag      2958
Ser Arg Thr Pro Thr Pro His His Thr Pro Pro Ser Ile Gly Ala Gln
840                 845                 850

cag cca cca gca aca aca att cca gcc cct gtt cct aca cct cct gcc      3006
Gln Pro Pro Ala Thr Thr Ile Pro Ala Pro Val Pro Thr Pro Pro Ala
855                 860                 865

atg cca cct ggg cca cag tcc cag gct cta cat ccc cct cca agg cag      3054
Met Pro Pro Gly Pro Gln Ser Gln Ala Leu His Pro Pro Pro Arg Gln
870                 875                 880

aca cct aca cca cca aca aca caa ctt ccc caa caa gtg cag cct tca      3102
Thr Pro Thr Pro Pro Thr Thr Gln Leu Pro Gln Gln Val Gln Pro Ser
885                 890                 895                 900

ctt cct gct gca cct tct gct gac cag ccc cag cag cag cct cgc tca      3150
Leu Pro Ala Ala Pro Ser Ala Asp Gln Pro Gln Gln Gln Pro Arg Ser
905                 910                 915

cag cag agc aca gca gcg tct gtt cct acc cca aca gca ccg ctg ctt      3198
Gln Gln Ser Thr Ala Ala Ser Val Pro Thr Pro Thr Ala Pro Leu Leu
920                 925                 930

cct ccg cag cct gca act cca ctt tcc cag cca gct gta agc att gaa      3246
Pro Pro Gln Pro Ala Thr Pro Leu Ser Gln Pro Ala Val Ser Ile Glu
935                 940                 945

gga cag gta tca aat cct cca tct act agt agc aca gaa gtg aat tct      3294
Gly Gln Val Ser Asn Pro Pro Ser Thr Ser Ser Thr Glu Val Asn Ser
950                 955                 960

cag gcc att gct gag aag cag cct tcc cag gaa gtg aag atg gag gcc      3342
Gln Ala Ile Ala Glu Lys Gln Pro Ser Gln Glu Val Lys Met Glu Ala
965                 970                 975                 980

aaa atg gaa gtg gat caa cca gaa cca gca gat act cag ccg gag gat      3390
Lys Met Glu Val Asp Gln Pro Glu Pro Ala Asp Thr Gln Pro Glu Asp
985                 990                 995

att tca gag tct  aaa gtg gaa gac tgt  aaa atg gaa tct acc  gaa       3435
Ile Ser Glu Ser  Lys Val Glu Asp Cys  Lys Met Glu Ser Thr  Glu
1000                 1005                 1010

aca gaa gag aga  agc act gag tta aaa  act gaa ata aaa gag  gag       3480
Thr Glu Glu Arg  Ser Thr Glu Leu Lys  Thr Glu Ile Lys Glu  Glu
1015                 1020                 1025

gaa gac cag cca  agt act tca gct acc  cag tca tct ccg gct  cca       3525
Glu Asp Gln Pro  Ser Thr Ser Ala Thr  Gln Ser Ser Pro Ala  Pro
1030                 1035                 1040

gga cag tca aag  aaa aag att ttc aaa  cca gaa gaa cta cga  cag       3570
Gly Gln Ser Lys  Lys Lys Ile Phe Lys  Pro Glu Glu Leu Arg  Gln
1045                 1050                 1055

gca ctg atg cca  act ttg gag gca ctt  tac cgt cag gat cca  gaa       3615
Ala Leu Met Pro  Thr Leu Glu Ala Leu  Tyr Arg Gln Asp Pro  Glu
```

-continued

```
1060                1065                1070

tcc ctt ccc ttt  cgt caa cct gtg gac  cct cag ctt tta gga  atc       3660
Ser Leu Pro Phe  Arg Gln Pro Val Asp  Pro Gln Leu Leu Gly  Ile
1075                1080                1085

cct gat tac ttt  gat att gtg aag agc  ccc atg gat ctt tct  acc       3705
Pro Asp Tyr Phe  Asp Ile Val Lys Ser  Pro Met Asp Leu Ser  Thr
1090                1095                1100

att aag agg aag  tta gac act gga cag  tat cag gag ccc tgg  cag       3750
Ile Lys Arg Lys  Leu Asp Thr Gly Gln  Tyr Gln Glu Pro Trp  Gln
1105                1110                1115

tat gtc gat gat  att tgg ctt atg ttc  aat aat gcc tgg tta  tat       3795
Tyr Val Asp Asp  Ile Trp Leu Met Phe  Asn Asn Ala Trp Leu  Tyr
1120                1125                1130

aac cgg aaa aca  tca cgg gta tac aaa  tac tgc tcc aag ctc  tct       3840
Asn Arg Lys Thr  Ser Arg Val Tyr Lys  Tyr Cys Ser Lys Leu  Ser
1135                1140                1145

gag gtc ttt gaa  caa gaa att gac cca  gtg atg caa agc ctt  gga       3885
Glu Val Phe Glu  Gln Glu Ile Asp Pro  Val Met Gln Ser Leu  Gly
1150                1155                1160

tac tgt tgt ggc  aga aag ttg gag ttc  tct cca cag aca ctg  tgt       3930
Tyr Cys Cys Gly  Arg Lys Leu Glu Phe  Ser Pro Gln Thr Leu  Cys
1165                1170                1175

tgc tac ggc aaa  cag ttg tgc aca ata  cct cgt gat gcc act  tat       3975
Cys Tyr Gly Lys  Gln Leu Cys Thr Ile  Pro Arg Asp Ala Thr  Tyr
1180                1185                1190

tac agt tac cag  aac agg tat cat ttc  tgt gag aag tgt ttc  aat       4020
Tyr Ser Tyr Gln  Asn Arg Tyr His Phe  Cys Glu Lys Cys Phe  Asn
1195                1200                1205

gag atc caa ggg  gag agc gtt tct ttg  ggg gat gac cct tcc  cag       4065
Glu Ile Gln Gly  Glu Ser Val Ser Leu  Gly Asp Asp Pro Ser  Gln
1210                1215                1220

cct caa act aca  ata aat aaa gaa caa  ttt tcc aag aga aaa  aat       4110
Pro Gln Thr Thr  Ile Asn Lys Glu Gln  Phe Ser Lys Arg Lys  Asn
1225                1230                1235

gac aca ctg gat  cct gaa ctg ttt gtt  gaa tgt aca gag tgc  gga       4155
Asp Thr Leu Asp  Pro Glu Leu Phe Val  Glu Cys Thr Glu Cys  Gly
1240                1245                1250

aga aag atg cat  cag atc tgt gtc ctt  cac cat gag atc atc  tgg       4200
Arg Lys Met His  Gln Ile Cys Val Leu  His His Glu Ile Ile  Trp
1255                1260                1265

cct gct gga ttc  gtc tgt gat ggc tgt  tta aag aaa agt gca  cga       4245
Pro Ala Gly Phe  Val Cys Asp Gly Cys  Leu Lys Lys Ser Ala  Arg
1270                1275                1280

act agg aaa gaa  aat aag ttt tct gct  aaa agg ttg cca tct  acc       4290
Thr Arg Lys Glu  Asn Lys Phe Ser Ala  Lys Arg Leu Pro Ser  Thr
1285                1290                1295

aga ctt ggc acc  ttt cta gag aat cgt  gtg aat gac ttt ctg  agg       4335
Arg Leu Gly Thr  Phe Leu Glu Asn Arg  Val Asn Asp Phe Leu  Arg
1300                1305                1310

cga cag aat cac  cct gag tca gga gag  gtc act gtt aga gta  gtt       4380
Arg Gln Asn His  Pro Glu Ser Gly Glu  Val Thr Val Arg Val  Val
1315                1320                1325

cat gct tct gac  aaa acc gtg gaa gta  aaa cca ggc atg aaa  gca       4425
His Ala Ser Asp  Lys Thr Val Glu Val  Lys Pro Gly Met Lys  Ala
1330                1335                1340

agg ttt gtg gac  agt gga gag atg gca  gaa tcc ttt cca tac  cga       4470
Arg Phe Val Asp  Ser Gly Glu Met Ala  Glu Ser Phe Pro Tyr  Arg
```

-continued

```
1345                1350                1355

acc aaa gcc ctc  ttt gcc ttt gaa gaa  att gat ggt gtt gac  ctg      4515
Thr Lys Ala Leu  Phe Ala Phe Glu Glu  Ile Asp Gly Val Asp  Leu
1360                1365                1370

tgc ttc ttt ggc  atg cat gtt caa gag  tat ggc tct gac tgc  cct      4560
Cys Phe Phe Gly  Met His Val Gln Glu  Tyr Gly Ser Asp Cys  Pro
1375                1380                1385

cca ccc aac cag  agg aga gta tac ata  tct tac ctc gat agt  gtt      4605
Pro Pro Asn Gln  Arg Arg Val Tyr Ile  Ser Tyr Leu Asp Ser  Val
1390                1395                1400

cat ttc ttc cgt  cct aaa tgc ttg agg  act gca gtc tat cat  gaa      4650
His Phe Phe Arg  Pro Lys Cys Leu Arg  Thr Ala Val Tyr His  Glu
1405                1410                1415

atc cta att gga  tat tta gaa tat gtc  aag aaa tta ggt tac  aca      4695
Ile Leu Ile Gly  Tyr Leu Glu Tyr Val  Lys Lys Leu Gly Tyr  Thr
1420                1425                1430

aca ggg cat att  tgg gca tgt cca cca  agt gag gga gat gat  tat      4740
Thr Gly His Ile  Trp Ala Cys Pro Pro  Ser Glu Gly Asp Asp  Tyr
1435                1440                1445

atc ttc cat tgc  cat cct cct gac cag  aag ata ccc aag ccc  aag      4785
Ile Phe His Cys  His Pro Pro Asp Gln  Lys Ile Pro Lys Pro  Lys
1450                1455                1460

cga ctg cag gaa  tgg tac aaa aaa atg  ctt gac aag gct gta  tca      4830
Arg Leu Gln Glu  Trp Tyr Lys Lys Met  Leu Asp Lys Ala Val  Ser
1465                1470                1475

gag cgt att gtc  cat gac tac aag gat  att ttt aaa caa gct  act      4875
Glu Arg Ile Val  His Asp Tyr Lys Asp  Ile Phe Lys Gln Ala  Thr
1480                1485                1490

gaa gat aga tta  aca agt gca aag gaa  ttg cct tat ttc gag  ggt      4920
Glu Asp Arg Leu  Thr Ser Ala Lys Glu  Leu Pro Tyr Phe Glu  Gly
1495                1500                1505

gat ttc tgg ccc  aat gtt ctg gaa gaa  agc att aag gaa ctg  gaa      4965
Asp Phe Trp Pro  Asn Val Leu Glu Glu  Ser Ile Lys Glu Leu  Glu
1510                1515                1520

cag gag gaa gaa  gag aga aaa cga gag  gaa aac acc agc aat  gaa      5010
Gln Glu Glu Glu  Glu Arg Lys Arg Glu  Glu Asn Thr Ser Asn  Glu
1525                1530                1535

agc aca gat gtg  acc aag gga gac agc  aaa aat gct aaa aag  aag      5055
Ser Thr Asp Val  Thr Lys Gly Asp Ser  Lys Asn Ala Lys Lys  Lys
1540                1545                1550

aat aat aag aaa  acc agc aaa aat aag  agc agc ctg agt agg  ggc      5100
Asn Asn Lys Lys  Thr Ser Lys Asn Lys  Ser Ser Leu Ser Arg  Gly
1555                1560                1565

aac aag aag aaa  ccc ggg atg ccc aat  gta tct aac gac ctc  tca      5145
Asn Lys Lys Lys  Pro Gly Met Pro Asn  Val Ser Asn Asp Leu  Ser
1570                1575                1580

cag aaa cta tat  gcc acc atg gag aag  cat aaa gag gtc ttc  ttt      5190
Gln Lys Leu Tyr  Ala Thr Met Glu Lys  His Lys Glu Val Phe  Phe
1585                1590                1595

gtg atc cgc ctc  att gct ggc cct gct  gcc aac tcc ctg cct  ccc      5235
Val Ile Arg Leu  Ile Ala Gly Pro Ala  Ala Asn Ser Leu Pro  Pro
1600                1605                1610

att gtt gat cct  gat cct ctc atc ccc  tgc gat ctg atg gat  ggt      5280
Ile Val Asp Pro  Asp Pro Leu Ile Pro  Cys Asp Leu Met Asp  Gly
1615                1620                1625

cgg gat gcg ttt  ctc acg ctg gca agg  gac aag cac ctg gag  ttc      5325
Arg Asp Ala Phe  Leu Thr Leu Ala Arg  Asp Lys His Leu Glu  Phe
```

-continued

```
1630                1635                1640

tct tca ctc cga  aga gcc cag tgg tcc  acc atg tgc atg ctg  gtg      5370
Ser Ser Leu Arg  Arg Ala Gln Trp Ser  Thr Met Cys Met Leu  Val
1645                1650                1655

gag ctg cac acg  cag agc cag gac cgc  ttt gtc tac acc tgc  aat      5415
Glu Leu His Thr  Gln Ser Gln Asp Arg  Phe Val Tyr Thr Cys  Asn
1660                1665                1670

gaa tgc aag cac  cat gtg gag aca cgc  tgg cac tgt act gtc  tgt      5460
Glu Cys Lys His  His Val Glu Thr Arg  Trp His Cys Thr Val  Cys
1675                1680                1685

gag gat tat gac  ttg tgt atc acc tgc  tat aac act aaa aac  cat      5505
Glu Asp Tyr Asp  Leu Cys Ile Thr Cys  Tyr Asn Thr Lys Asn  His
1690                1695                1700

gac cac aaa atg  gag aaa cta ggc ctt  ggc tta gat gat gag  agc      5550
Asp His Lys Met  Glu Lys Leu Gly Leu  Gly Leu Asp Asp Glu  Ser
1705                1710                1715

aac aac cag cag  gct gca gcc acc cag  agc cca ggc gat tct  cgc      5595
Asn Asn Gln Gln  Ala Ala Ala Thr Gln  Ser Pro Gly Asp Ser  Arg
1720                1725                1730

cgc ctg agt atc  cag cgc tgc atc cag  tct ctg gtc cat gct  tgc      5640
Arg Leu Ser Ile  Gln Arg Cys Ile Gln  Ser Leu Val His Ala  Cys
1735                1740                1745

cag tgt cgg aat  gcc aat tgc tca ctg  cca tcc tgc cag aag  atg      5685
Gln Cys Arg Asn  Ala Asn Cys Ser Leu  Pro Ser Cys Gln Lys  Met
1750                1755                1760

aag cgg gtt gtg  cag cat acc aag ggt  tgc aaa cgg aaa acc  aat      5730
Lys Arg Val Val  Gln His Thr Lys Gly  Cys Lys Arg Lys Thr  Asn
1765                1770                1775

ggc ggg tgc ccc  atc tgc aag cag ctc  att gcc ctc tgc tgc  tac      5775
Gly Gly Cys Pro  Ile Cys Lys Gln Leu  Ile Ala Leu Cys Cys  Tyr
1780                1785                1790

cat gcc aag cac  tgc cag gag aac aaa  tgc ccg gtg ccg ttc  tgc      5820
His Ala Lys His  Cys Gln Glu Asn Lys  Cys Pro Val Pro Phe  Cys
1795                1800                1805

cta aac atc aag  cag aag ctc cgg cag  caa cag ctg cag cac  cga      5865
Leu Asn Ile Lys  Gln Lys Leu Arg Gln  Gln Gln Leu Gln His  Arg
1810                1815                1820

cta cag cag gcc  caa atg ctt cgc agg  agg atg gcc agc atg  cag      5910
Leu Gln Gln Ala  Gln Met Leu Arg Arg  Arg Met Ala Ser Met  Gln
1825                1830                1835

cgg act ggt gtg  gtt ggg cag caa cag  ggc ctc cct tcc ccc  act      5955
Arg Thr Gly Val  Val Gly Gln Gln Gln  Gly Leu Pro Ser Pro  Thr
1840                1845                1850

cct gcc act cca  acg aca cca act ggc  caa cag cca acc acc  ccg      6000
Pro Ala Thr Pro  Thr Thr Pro Thr Gly  Gln Gln Pro Thr Thr  Pro
1855                1860                1865

cag acg ccc cag  ccc act tct cag cct  cag cct acc cct ccc  aat      6045
Gln Thr Pro Gln  Pro Thr Ser Gln Pro  Gln Pro Thr Pro Pro  Asn
1870                1875                1880

agc atg cca ccc  tac ttg ccc agg act  caa gct gct ggc cct  gtg      6090
Ser Met Pro Pro  Tyr Leu Pro Arg Thr  Gln Ala Ala Gly Pro  Val
1885                1890                1895

tcc cag ggt aag  gca gca ggc cag gtg  acc cct cca acc cct  cct      6135
Ser Gln Gly Lys  Ala Ala Gly Gln Val  Thr Pro Pro Thr Pro  Pro
1900                1905                1910

cag act gct cag  cca ccc ctt cca ggg  ccc cca cct gca gca  gtg      6180
Gln Thr Ala Gln  Pro Pro Leu Pro Gly  Pro Pro Pro Ala Ala  Val
```

-continued

```
1915                1920                1925

gaa atg gca atg  cag att cag aga gca  gcg gag acg cag cgc  cag      6225
Glu Met Ala Met  Gln Ile Gln Arg Ala  Ala Glu Thr Gln Arg  Gln
1930                1935                1940

atg gcc cac gtg  caa att ttt caa agg  cca atc caa cac cag  atg      6270
Met Ala His Val  Gln Ile Phe Gln Arg  Pro Ile Gln His Gln  Met
1945                1950                1955

ccc ccg atg act  ccc atg gcc ccc atg  ggt atg aac cca cct  ccc      6315
Pro Pro Met Thr  Pro Met Ala Pro Met  Gly Met Asn Pro Pro  Pro
1960                1965                1970

atg acc aga ggt  ccc agt ggg cat ttg  gag cca ggg atg gga  ccg      6360
Met Thr Arg Gly  Pro Ser Gly His Leu  Glu Pro Gly Met Gly  Pro
1975                1980                1985

aca ggg atg cag  caa cag cca ccc tgg  agc caa gga gga ttg  cct      6405
Thr Gly Met Gln  Gln Gln Pro Pro Trp  Ser Gln Gly Gly Leu  Pro
1990                1995                2000

cag ccc cag caa  cta cag tct ggg atg  cca agg cca gcc atg  atg      6450
Gln Pro Gln Gln  Leu Gln Ser Gly Met  Pro Arg Pro Ala Met  Met
2005                2010                2015

tca gtg gcc cag  cat ggt caa cct ttg  aac atg gct cca caa  cca      6495
Ser Val Ala Gln  His Gly Gln Pro Leu  Asn Met Ala Pro Gln  Pro
2020                2025                2030

gga ttg ggc cag  gta ggt atc agc cca  ctc aaa cca ggc act  gtg      6540
Gly Leu Gly Gln  Val Gly Ile Ser Pro  Leu Lys Pro Gly Thr  Val
2035                2040                2045

tct caa caa gcc  tta caa aac ctt ttg  cgg act ctc agg tct  ccc      6585
Ser Gln Gln Ala  Leu Gln Asn Leu Leu  Arg Thr Leu Arg Ser  Pro
2050                2055                2060

agc tct ccc ctg  cag cag caa cag gtg  ctt agt atc ctt cac  gcc      6630
Ser Ser Pro Leu  Gln Gln Gln Gln Val  Leu Ser Ile Leu His  Ala
2065                2070                2075

aac ccc cag ctg  ttg gct gca ttc atc  aag cag cgg gct gcc  aag      6675
Asn Pro Gln Leu  Leu Ala Ala Phe Ile  Lys Gln Arg Ala Ala  Lys
2080                2085                2090

tat gcc aac tct  aat cca caa ccc atc  cct ggg cag cct ggc  atg      6720
Tyr Ala Asn Ser  Asn Pro Gln Pro Ile  Pro Gly Gln Pro Gly  Met
2095                2100                2105

ccc cag ggg cag  cca ggg cta cag cca  cct acc atg cca ggt  cag      6765
Pro Gln Gly Gln  Pro Gly Leu Gln Pro  Pro Thr Met Pro Gly  Gln
2110                2115                2120

cag ggg gtc cac  tcc aat cca gcc atg  cag aac atg aat cca  atg      6810
Gln Gly Val His  Ser Asn Pro Ala Met  Gln Asn Met Asn Pro  Met
2125                2130                2135

cag gcg ggc gtt  cag agg gct ggc ctg  ccc cag cag caa cca  cag      6855
Gln Ala Gly Val  Gln Arg Ala Gly Leu  Pro Gln Gln Gln Pro  Gln
2140                2145                2150

cag caa ctc cag  cca ccc atg gga ggg  atg agc ccc cag gct  cag      6900
Gln Gln Leu Gln  Pro Pro Met Gly Gly  Met Ser Pro Gln Ala  Gln
2155                2160                2165

cag atg aac atg  aac cac aac acc atg  cct tca caa ttc cga  gac      6945
Gln Met Asn Met  Asn His Asn Thr Met  Pro Ser Gln Phe Arg  Asp
2170                2175                2180

atc ttg aga cga  cag caa atg atg caa  cag cag cag caa cag  gga      6990
Ile Leu Arg Arg  Gln Gln Met Met Gln  Gln Gln Gln Gln Gln  Gly
2185                2190                2195

gca ggg cca gga  ata ggc cct gga atg  gcc aac cat aac cag  ttc      7035
Ala Gly Pro Gly  Ile Gly Pro Gly Met  Ala Asn His Asn Gln  Phe
```

-continued

```
2200                2205                2210

cag caa ccc caa  gga gtt ggc tac cca  cca cag cag cag cag  cgg      7080
Gln Gln Pro Gln  Gly Val Gly Tyr Pro  Pro Gln Gln Gln Gln  Arg
2215                2220                2225

atg cag cat cac  atg caa cag atg caa  caa gga aat atg gga  cag      7125
Met Gln His His  Met Gln Gln Met Gln  Gln Gly Asn Met Gly  Gln
2230                2235                2240

ata ggc cag ctt  ccc cag gcc ttg gga  gca gag gca ggt gcc  agt      7170
Ile Gly Gln Leu  Pro Gln Ala Leu Gly  Ala Glu Ala Gly Ala  Ser
2245                2250                2255

cta cag gcc tat  cag cag cga ctc ctt  cag caa cag atg ggg  tcc      7215
Leu Gln Ala Tyr  Gln Gln Arg Leu Leu  Gln Gln Gln Met Gly  Ser
2260                2265                2270

cct gtt cag ccc  aac ccc atg agc ccc  cag cag cat atg ctc  cca      7260
Pro Val Gln Pro  Asn Pro Met Ser Pro  Gln Gln His Met Leu  Pro
2275                2280                2285

aat cag gcc cag  tcc cca cac cta caa  ggc cag cag atc cct  aat      7305
Asn Gln Ala Gln  Ser Pro His Leu Gln  Gly Gln Gln Ile Pro  Asn
2290                2295                2300

tct ctc tcc aat  caa gtg cgc tct ccc  cag cct gtc cct tct  cca      7350
Ser Leu Ser Asn  Gln Val Arg Ser Pro  Gln Pro Val Pro Ser  Pro
2305                2310                2315

cgg cca cag tcc  cag ccc ccc cac tcc  agt cct tcc cca agg  atg      7395
Arg Pro Gln Ser  Gln Pro Pro His Ser  Ser Pro Ser Pro Arg  Met
2320                2325                2330

cag cct cag cct  tct cca cac cac gtt  tcc cca cag aca agt  tcc      7440
Gln Pro Gln Pro  Ser Pro His His Val  Ser Pro Gln Thr Ser  Ser
2335                2340                2345

cca cat cct gga  ctg gta gct gcc cag  gcc aac ccc atg gaa  caa      7485
Pro His Pro Gly  Leu Val Ala Ala Gln  Ala Asn Pro Met Glu  Gln
2350                2355                2360

ggg cat ttt gcc  agc ccg gac cag aat  tca atg ctt tct cag  ctt      7530
Gly His Phe Ala  Ser Pro Asp Gln Asn  Ser Met Leu Ser Gln  Leu
2365                2370                2375

gct agc aat cca  ggc atg gca aac ctc  cat ggt gca agc gcc  acg      7575
Ala Ser Asn Pro  Gly Met Ala Asn Leu  His Gly Ala Ser Ala  Thr
2380                2385                2390

gac ctg gga ctc  agc acc gat aac tca  gac ttg aat tca aac  ctc      7620
Asp Leu Gly Leu  Ser Thr Asp Asn Ser  Asp Leu Asn Ser Asn  Leu
2395                2400                2405

tca cag agt aca  cta gac ata cac tag agacaccttg tagtattttg          7667
Ser Gln Ser Thr  Leu Asp Ile His
2410

ggagcaaaaa aattattttc tcttaacaag actttttgta ctgaaaacaa ttttttttgaa   7727

tctttcgtag cctaaaagac aattttcctt ggaacacata agaactgtgc agtagccgtt   7787

tgtggtttaa agcaaacatg caagatgaac ctgagggatg atagaataca aagaatatat   7847

ttttgttatg gctggttacc accagccttt cttccccttt gtgtgtgtgg ttcaagtgtg   7907

cactgggagg aggctgaggc ctgtgaagcc aaacaatatg ctcctgcctt gcacctccaa   7967

taggttttat tatttttttt aaattaatga acatatgtaa tattaatagt tattatttac   8027

tggtgcagat ggttgacatt tttccctatt ttcctcactt tatggaagag ttaaaacatt   8087

tctaaaccag aggacaaaag ggttaatgt tactttaaaa ttacattcta tatatatata    8147

aatatatata aatatatatt aaaataccag ttttttttct ctgggtgcaa agatgttcat   8207
```

-continued

```
tcttttaaaa aatgtttaaa aaaaaaaaaa aactgccttt cttcccctca agtcaacttt   8267

tgtgctccag aaaattttct attctgtaag tctgagcgta aaacttcaag tattaaaata   8327

atttgtacat gtagagagaa aaatgacttt ttcaaaaata tacaggggca gctgccaaat   8387

tgatgtatta tatattgtgg tttctgtttc ttgaaagaat ttttttcgtt atttttacat   8447

ctaacaaagt aaaaaaatta aaaagagggt aagaaacgat tccggtggga tgattttaac   8507

atgcaaaatg tccctggggg tttcttcttt gcttgctttc ttcctcctta ccctaccccc   8567

cactcacaca cacacacaca cacacacaca cacacacaca cacacacttt ctataaaact   8627

tgaaaatagc aaaaaccctc aactgttgta aatcatgcaa ttaaagttga ttacttataa   8687

atatgaactt tggatcactg tatagactgt taaatttgat ttcttattac ctattgttaa   8747

ataaactgtg tgagacag                                                 8765


<210> SEQ ID NO 4
<211> LENGTH: 2414
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 4

Met Ala Glu Asn Val Val Glu Pro Gly Pro Pro Ser Ala Lys Arg Pro
1               5                   10                  15

Lys Leu Ser Ser Pro Ala Leu Ser Ala Ser Ala Ser Asp Gly Thr Asp
20                  25                  30

Phe Gly Ser Leu Phe Asp Leu Glu His Asp Leu Pro Asp Glu Leu Ile
35                  40                  45

Asn Ser Thr Glu Leu Gly Leu Thr Asn Gly Gly Asp Ile Asn Gln Leu
50                  55                  60

Gln Thr Ser Leu Gly Met Val Gln Asp Ala Ala Ser Lys His Lys Gln
65                  70                  75                  80

Leu Ser Glu Leu Leu Arg Ser Gly Ser Ser Pro Asn Leu Asn Met Gly
85                  90                  95

Val Gly Gly Pro Gly Gln Val Met Ala Ser Gln Ala Gln Gln Ser Ser
100                 105                 110

Pro Gly Leu Gly Leu Ile Asn Ser Met Val Lys Ser Pro Met Thr Gln
115                 120                 125

Ala Gly Leu Thr Ser Pro Asn Met Gly Met Gly Thr Ser Gly Pro Asn
130                 135                 140

Gln Gly Pro Thr Gln Ser Thr Gly Met Met Asn Ser Pro Val Asn Gln
145                 150                 155                 160

Pro Ala Met Gly Met Asn Thr Gly Met Asn Ala Gly Met Asn Pro Gly
165                 170                 175

Met Leu Ala Ala Gly Asn Gly Gln Gly Ile Met Pro Asn Gln Val Met
180                 185                 190

Asn Gly Ser Ile Gly Ala Gly Arg Gly Arg Gln Asn Met Gln Tyr Pro
195                 200                 205

Asn Pro Gly Met Gly Ser Ala Gly Asn Leu Leu Thr Glu Pro Leu Gln
210                 215                 220

Gln Gly Ser Pro Gln Met Gly Gly Gln Thr Gly Leu Arg Gly Pro Gln
225                 230                 235                 240

Pro Leu Lys Met Gly Met Met Asn Asn Pro Asn Pro Tyr Gly Ser Pro
245                 250                 255
```

-continued

```
Tyr Thr Gln Asn Pro Gly Gln Gln Ile Gly Ala Ser Gly Leu Gly Leu
260                 265                 270

Gln Ile Gln Thr Lys Thr Val Leu Ser Asn Asn Leu Ser Pro Phe Ala
275                 280                 285

Met Asp Lys Lys Ala Val Pro Gly Gly Gly Met Pro Asn Met Gly Gln
290                 295                 300

Gln Pro Ala Pro Gln Val Gln Gln Pro Gly Leu Val Thr Pro Val Ala
305                 310                 315                 320

Gln Gly Met Gly Ser Gly Ala His Thr Ala Asp Pro Glu Lys Arg Lys
325                 330                 335

Leu Ile Gln Gln Gln Leu Val Leu Leu Leu His Ala His Lys Cys Gln
340                 345                 350

Arg Arg Glu Gln Ala Asn Gly Glu Val Arg Gln Cys Asn Leu Pro His
355                 360                 365

Cys Arg Thr Met Lys Asn Val Leu Asn His Met Thr His Cys Gln Ser
370                 375                 380

Gly Lys Ser Cys Gln Val Ala His Cys Ala Ser Ser Arg Gln Ile Ile
385                 390                 395                 400

Ser His Trp Lys Asn Cys Thr Arg His Asp Cys Pro Val Cys Leu Pro
405                 410                 415

Leu Lys Asn Ala Gly Asp Lys Arg Asn Gln Gln Pro Ile Leu Thr Gly
420                 425                 430

Ala Pro Val Gly Leu Gly Asn Pro Ser Ser Leu Gly Val Gly Gln Gln
435                 440                 445

Ser Ala Pro Asn Leu Ser Thr Val Ser Gln Ile Asp Pro Ser Ser Ile
450                 455                 460

Glu Arg Ala Tyr Ala Ala Leu Gly Leu Pro Tyr Gln Val Asn Gln Met
465                 470                 475                 480

Pro Thr Gln Pro Gln Val Gln Ala Lys Asn Gln Gln Asn Gln Gln Pro
485                 490                 495

Gly Gln Ser Pro Gln Gly Met Arg Pro Met Ser Asn Met Ser Ala Ser
500                 505                 510

Pro Met Gly Val Asn Gly Gly Val Gly Val Gln Thr Pro Ser Leu Leu
515                 520                 525

Ser Asp Ser Met Leu His Ser Ala Ile Asn Ser Gln Asn Pro Met Met
530                 535                 540

Ser Glu Asn Ala Ser Val Pro Ser Leu Gly Pro Met Pro Thr Ala Ala
545                 550                 555                 560

Gln Pro Ser Thr Thr Gly Ile Arg Lys Gln Trp His Glu Asp Ile Thr
565                 570                 575

Gln Asp Leu Arg Asn His Leu Val His Lys Leu Val Gln Ala Ile Phe
580                 585                 590

Pro Thr Pro Asp Pro Ala Ala Leu Lys Asp Arg Arg Met Glu Asn Leu
595                 600                 605

Val Ala Tyr Ala Arg Lys Val Glu Gly Asp Met Tyr Glu Ser Ala Asn
610                 615                 620

Asn Arg Ala Glu Tyr Tyr His Leu Leu Ala Glu Lys Ile Tyr Lys Ile
625                 630                 635                 640

Gln Lys Glu Leu Glu Glu Lys Arg Arg Thr Arg Leu Gln Lys Gln Asn
645                 650                 655

Met Leu Pro Asn Ala Ala Gly Met Val Pro Val Ser Met Asn Pro Gly
```

-continued

```
660                 665                 670

Pro Asn Met Gly Gln Pro Gln Pro Gly Met Thr Ser Asn Gly Pro Leu
675                 680                 685

Pro Asp Pro Ser Met Ile Arg Gly Ser Val Pro Asn Gln Met Met Pro
690                 695                 700

Arg Ile Thr Pro Gln Ser Gly Leu Asn Gln Phe Gly Gln Met Ser Met
705                 710                 715                 720

Ala Gln Pro Pro Ile Val Pro Arg Gln Thr Pro Pro Leu Gln His His
725                 730                 735

Gly Gln Leu Ala Gln Pro Gly Ala Leu Asn Pro Pro Met Gly Tyr Gly
740                 745                 750

Pro Arg Met Gln Gln Pro Ser Asn Gln Gly Gln Phe Leu Pro Gln Thr
755                 760                 765

Gln Phe Pro Ser Gln Gly Met Asn Val Thr Asn Ile Pro Leu Ala Pro
770                 775                 780

Ser Ser Gly Gln Ala Pro Val Ser Gln Ala Gln Met Ser Ser Ser Ser
785                 790                 795                 800

Cys Pro Val Asn Ser Pro Ile Met Pro Pro Gly Ser Gln Gly Ser His
805                 810                 815

Ile His Cys Pro Gln Leu Pro Gln Pro Ala Leu His Gln Asn Ser Pro
820                 825                 830

Ser Pro Val Pro Ser Arg Thr Pro Thr Pro His His Thr Pro Pro Ser
835                 840                 845

Ile Gly Ala Gln Gln Pro Pro Ala Thr Thr Ile Pro Ala Pro Val Pro
850                 855                 860

Thr Pro Pro Ala Met Pro Pro Gly Pro Gln Ser Gln Ala Leu His Pro
865                 870                 875                 880

Pro Pro Arg Gln Thr Pro Thr Pro Pro Thr Thr Gln Leu Pro Gln Gln
885                 890                 895

Val Gln Pro Ser Leu Pro Ala Ala Pro Ser Ala Asp Gln Pro Gln Gln
900                 905                 910

Gln Pro Arg Ser Gln Gln Ser Thr Ala Ala Ser Val Pro Thr Pro Thr
915                 920                 925

Ala Pro Leu Leu Pro Pro Gln Pro Ala Thr Pro Leu Ser Gln Pro Ala
930                 935                 940

Val Ser Ile Glu Gly Gln Val Ser Asn Pro Pro Ser Thr Ser Ser Thr
945                 950                 955                 960

Glu Val Asn Ser Gln Ala Ile Ala Glu Lys Gln Pro Ser Gln Glu Val
965                 970                 975

Lys Met Glu Ala Lys Met Glu Val Asp Gln Pro Glu Pro Ala Asp Thr
980                 985                 990

Gln Pro Glu Asp Ile Ser Glu Ser  Lys Val Glu Asp Cys  Lys Met Glu
995                 1000                 1005

Ser Thr  Glu Thr Glu Glu Arg  Ser Thr Glu Leu Lys  Thr Glu Ile
1010                 1015                 1020

Lys Glu  Glu Glu Asp Gln Pro  Ser Thr Ser Ala Thr  Gln Ser Ser
1025                 1030                 1035

Pro Ala  Pro Gly Gln Ser Lys  Lys Lys Ile Phe Lys  Pro Glu Glu
1040                 1045                 1050

Leu Arg  Gln Ala Leu Met Pro  Thr Leu Glu Ala Leu  Tyr Arg Gln
1055                 1060                 1065
```

-continued

```
Asp Pro  Glu Ser Leu Pro Phe  Arg Gln Pro Val Asp  Pro Gln Leu
1070                 1075                 1080

Leu Gly  Ile Pro Asp Tyr Phe  Asp Ile Val Lys Ser  Pro Met Asp
1085                 1090                 1095

Leu Ser  Thr Ile Lys Arg Lys  Leu Asp Thr Gly Gln  Tyr Gln Glu
1100                 1105                 1110

Pro Trp  Gln Tyr Val Asp Asp  Ile Trp Leu Met Phe  Asn Asn Ala
1115                 1120                 1125

Trp Leu  Tyr Asn Arg Lys Thr  Ser Arg Val Tyr Lys  Tyr Cys Ser
1130                 1135                 1140

Lys Leu  Ser Glu Val Phe Glu  Gln Glu Ile Asp Pro  Val Met Gln
1145                 1150                 1155

Ser Leu  Gly Tyr Cys Cys Gly  Arg Lys Leu Glu Phe  Ser Pro Gln
1160                 1165                 1170

Thr Leu  Cys Cys Tyr Gly Lys  Gln Leu Cys Thr Ile  Pro Arg Asp
1175                 1180                 1185

Ala Thr  Tyr Tyr Ser Tyr Gln  Asn Arg Tyr His Phe  Cys Glu Lys
1190                 1195                 1200

Cys Phe  Asn Glu Ile Gln Gly  Glu Ser Val Ser Leu  Gly Asp Asp
1205                 1210                 1215

Pro Ser  Gln Pro Gln Thr Thr  Ile Asn Lys Glu Gln  Phe Ser Lys
1220                 1225                 1230

Arg Lys  Asn Asp Thr Leu Asp  Pro Glu Leu Phe Val  Glu Cys Thr
1235                 1240                 1245

Glu Cys  Gly Arg Lys Met His  Gln Ile Cys Val Leu  His His Glu
1250                 1255                 1260

Ile Ile  Trp Pro Ala Gly Phe  Val Cys Asp Gly Cys  Leu Lys Lys
1265                 1270                 1275

Ser Ala  Arg Thr Arg Lys Glu  Asn Lys Phe Ser Ala  Lys Arg Leu
1280                 1285                 1290

Pro Ser  Thr Arg Leu Gly Thr  Phe Leu Glu Asn Arg  Val Asn Asp
1295                 1300                 1305

Phe Leu  Arg Arg Gln Asn His  Pro Glu Ser Gly Glu  Val Thr Val
1310                 1315                 1320

Arg Val  Val His Ala Ser Asp  Lys Thr Val Glu Val  Lys Pro Gly
1325                 1330                 1335

Met Lys  Ala Arg Phe Val Asp  Ser Gly Glu Met Ala  Glu Ser Phe
1340                 1345                 1350

Pro Tyr  Arg Thr Lys Ala Leu  Phe Ala Phe Glu Glu  Ile Asp Gly
1355                 1360                 1365

Val Asp  Leu Cys Phe Phe Gly  Met His Val Gln Glu  Tyr Gly Ser
1370                 1375                 1380

Asp Cys  Pro Pro Pro Asn Gln  Arg Arg Val Tyr Ile  Ser Tyr Leu
1385                 1390                 1395

Asp Ser  Val His Phe Phe Arg  Pro Lys Cys Leu Arg  Thr Ala Val
1400                 1405                 1410

Tyr His  Glu Ile Leu Ile Gly  Tyr Leu Glu Tyr Val  Lys Lys Leu
1415                 1420                 1425

Gly Tyr  Thr Thr Gly His Ile  Trp Ala Cys Pro Pro  Ser Glu Gly
1430                 1435                 1440
```

-continued

```
Asp Asp  Tyr Ile Phe His Cys  His Pro Pro Asp Gln  Lys Ile Pro
1445                 1450                 1455

Lys Pro  Lys Arg Leu Gln Glu  Trp Tyr Lys Lys Met  Leu Asp Lys
1460                 1465                 1470

Ala Val  Ser Glu Arg Ile Val  His Asp Tyr Lys Asp  Ile Phe Lys
1475                 1480                 1485

Gln Ala  Thr Glu Asp Arg Leu  Thr Ser Ala Lys Glu  Leu Pro Tyr
1490                 1495                 1500

Phe Glu  Gly Asp Phe Trp Pro  Asn Val Leu Glu Glu  Ser Ile Lys
1505                 1510                 1515

Glu Leu  Glu Gln Glu Glu Glu  Glu Arg Lys Arg Glu  Glu Asn Thr
1520                 1525                 1530

Ser Asn  Glu Ser Thr Asp Val  Thr Lys Gly Asp Ser  Lys Asn Ala
1535                 1540                 1545

Lys Lys  Lys Asn Asn Lys Lys  Thr Ser Lys Asn Lys  Ser Ser Leu
1550                 1555                 1560

Ser Arg  Gly Asn Lys Lys Lys  Pro Gly Met Pro Asn  Val Ser Asn
1565                 1570                 1575

Asp Leu  Ser Gln Lys Leu Tyr  Ala Thr Met Glu Lys  His Lys Glu
1580                 1585                 1590

Val Phe  Phe Val Ile Arg Leu  Ile Ala Gly Pro Ala  Ala Asn Ser
1595                 1600                 1605

Leu Pro  Pro Ile Val Asp Pro  Asp Pro Leu Ile Pro  Cys Asp Leu
1610                 1615                 1620

Met Asp  Gly Arg Asp Ala Phe  Leu Thr Leu Ala Arg  Asp Lys His
1625                 1630                 1635

Leu Glu  Phe Ser Ser Leu Arg  Arg Ala Gln Trp Ser  Thr Met Cys
1640                 1645                 1650

Met Leu  Val Glu Leu His Thr  Gln Ser Gln Asp Arg  Phe Val Tyr
1655                 1660                 1665

Thr Cys  Asn Glu Cys Lys His  His Val Glu Thr Arg  Trp His Cys
1670                 1675                 1680

Thr Val  Cys Glu Asp Tyr Asp  Leu Cys Ile Thr Cys  Tyr Asn Thr
1685                 1690                 1695

Lys Asn  His Asp His Lys Met  Glu Lys Leu Gly Leu  Gly Leu Asp
1700                 1705                 1710

Asp Glu  Ser Asn Asn Gln Gln  Ala Ala Ala Thr Gln  Ser Pro Gly
1715                 1720                 1725

Asp Ser  Arg Arg Leu Ser Ile  Gln Arg Cys Ile Gln  Ser Leu Val
1730                 1735                 1740

His Ala  Cys Gln Cys Arg Asn  Ala Asn Cys Ser Leu  Pro Ser Cys
1745                 1750                 1755

Gln Lys  Met Lys Arg Val Val  Gln His Thr Lys Gly  Cys Lys Arg
1760                 1765                 1770

Lys Thr  Asn Gly Gly Cys Pro  Ile Cys Lys Gln Leu  Ile Ala Leu
1775                 1780                 1785

Cys Cys  Tyr His Ala Lys His  Cys Gln Glu Asn Lys  Cys Pro Val
1790                 1795                 1800

Pro Phe  Cys Leu Asn Ile Lys  Gln Lys Leu Arg Gln  Gln Gln Leu
1805                 1810                 1815

Gln His  Arg Leu Gln Gln Ala  Gln Met Leu Arg Arg  Arg Met Ala
```

-continued

```
1820                1825                1830

Ser Met  Gln Arg Thr Gly Val  Val Gly Gln Gln Gln  Gly Leu Pro
1835                1840                1845

Ser Pro  Thr Pro Ala Thr Pro  Thr Thr Pro Thr Gly  Gln Gln Pro
1850                1855                1860

Thr Thr  Pro Gln Thr Pro Gln  Pro Thr Ser Gln Pro  Gln Pro Thr
1865                1870                1875

Pro Pro  Asn Ser Met Pro Pro  Tyr Leu Pro Arg Thr  Gln Ala Ala
1880                1885                1890

Gly Pro  Val Ser Gln Gly Lys  Ala Ala Gly Gln Val  Thr Pro Pro
1895                1900                1905

Thr Pro  Pro Gln Thr Ala Gln  Pro Pro Leu Pro Gly  Pro Pro Pro
1910                1915                1920

Ala Ala  Val Glu Met Ala Met  Gln Ile Gln Arg Ala  Ala Glu Thr
1925                1930                1935

Gln Arg  Gln Met Ala His Val  Gln Ile Phe Gln Arg  Pro Ile Gln
1940                1945                1950

His Gln  Met Pro Pro Met Thr  Pro Met Ala Pro Met  Gly Met Asn
1955                1960                1965

Pro Pro  Pro Met Thr Arg Gly  Pro Ser Gly His Leu  Glu Pro Gly
1970                1975                1980

Met Gly  Pro Thr Gly Met Gln  Gln Gln Pro Pro Trp  Ser Gln Gly
1985                1990                1995

Gly Leu  Pro Gln Pro Gln Gln  Leu Gln Ser Gly Met  Pro Arg Pro
2000                2005                2010

Ala Met  Met Ser Val Ala Gln  His Gly Gln Pro Leu  Asn Met Ala
2015                2020                2025

Pro Gln  Pro Gly Leu Gly Gln  Val Gly Ile Ser Pro  Leu Lys Pro
2030                2035                2040

Gly Thr  Val Ser Gln Gln Ala  Leu Gln Asn Leu Leu  Arg Thr Leu
2045                2050                2055

Arg Ser  Pro Ser Ser Pro Leu  Gln Gln Gln Gln Val  Leu Ser Ile
2060                2065                2070

Leu His  Ala Asn Pro Gln Leu  Leu Ala Ala Phe Ile  Lys Gln Arg
2075                2080                2085

Ala Ala  Lys Tyr Ala Asn Ser  Asn Pro Gln Pro Ile  Pro Gly Gln
2090                2095                2100

Pro Gly  Met Pro Gln Gly Gln  Pro Gly Leu Gln Pro  Pro Thr Met
2105                2110                2115

Pro Gly  Gln Gln Gly Val His  Ser Asn Pro Ala Met  Gln Asn Met
2120                2125                2130

Asn Pro  Met Gln Ala Gly Val  Gln Arg Ala Gly Leu  Pro Gln Gln
2135                2140                2145

Gln Pro  Gln Gln Gln Leu Gln  Pro Pro Met Gly Gly  Met Ser Pro
2150                2155                2160

Gln Ala  Gln Gln Met Asn Met  Asn His Asn Thr Met  Pro Ser Gln
2165                2170                2175

Phe Arg  Asp Ile Leu Arg Arg  Gln Gln Met Met Gln  Gln Gln Gln
2180                2185                2190

Gln Gln  Gly Ala Gly Pro Gly  Ile Gly Pro Gly Met  Ala Asn His
2195                2200                2205
```

-continued

```
Asn Gln  Phe Gln Gln Pro Gln  Gly Val Gly Tyr Pro  Pro Gln Gln
2210                 2215                 2220

Gln Gln  Arg Met Gln His His  Met Gln Gln Met Gln  Gln Gly Asn
2225                 2230                 2235

Met Gly  Gln Ile Gly Gln Leu  Pro Gln Ala Leu Gly  Ala Glu Ala
2240                 2245                 2250

Gly Ala  Ser Leu Gln Ala Tyr  Gln Gln Arg Leu Leu  Gln Gln Gln
2255                 2260                 2265

Met Gly  Ser Pro Val Gln Pro  Asn Pro Met Ser Pro  Gln Gln His
2270                 2275                 2280

Met Leu  Pro Asn Gln Ala Gln  Ser Pro His Leu Gln  Gly Gln Gln
2285                 2290                 2295

Ile Pro  Asn Ser Leu Ser Asn  Gln Val Arg Ser Pro  Gln Pro Val
2300                 2305                 2310

Pro Ser  Pro Arg Pro Gln Ser  Gln Pro Pro His Ser  Ser Pro Ser
2315                 2320                 2325

Pro Arg  Met Gln Pro Gln Pro  Ser Pro His His Val  Ser Pro Gln
2330                 2335                 2340

Thr Ser  Ser Pro His Pro Gly  Leu Val Ala Ala Gln  Ala Asn Pro
2345                 2350                 2355

Met Glu  Gln Gly His Phe Ala  Ser Pro Asp Gln Asn  Ser Met Leu
2360                 2365                 2370

Ser Gln  Leu Ala Ser Asn Pro  Gly Met Ala Asn Leu  His Gly Ala
2375                 2380                 2385

Ser Ala  Thr Asp Leu Gly Leu  Ser Thr Asp Asn Ser  Asp Leu Asn
2390                 2395                 2400

Ser Asn  Leu Ser Gln Ser Thr  Leu Asp Ile His
2405                 2410


<210> SEQ ID NO 5
<211> LENGTH: 1302
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (1)..(1302)
<223> OTHER INFORMATION: Active domain of p300 acetyltransferase
      (p300-4)

<400> SEQUENCE: 5

aag ccc aag cga ctg cag gaa tgg tac aaa aaa atg ctt gac aag gct       48
Lys Pro Lys Arg Leu Gln Glu Trp Tyr Lys Lys Met Leu Asp Lys Ala
1               5                   10                  15

gta tca gag cgt att gtc cat gac tac aag gat att ttt aaa caa gct       96
Val Ser Glu Arg Ile Val His Asp Tyr Lys Asp Ile Phe Lys Gln Ala
20                  25                  30

act gaa gat aga tta aca agt gca aag gaa ttg cct tat ttc gag ggt      144
Thr Glu Asp Arg Leu Thr Ser Ala Lys Glu Leu Pro Tyr Phe Glu Gly
35                  40                  45

gat ttc tgg ccc aat gtt ctg gaa gaa agc att aag gaa ctg gaa cag      192
Asp Phe Trp Pro Asn Val Leu Glu Glu Ser Ile Lys Glu Leu Glu Gln
50                  55                  60

gag gaa gaa gag aga aaa cga gag gaa aac acc agc aat gaa agc aca      240
Glu Glu Glu Glu Arg Lys Arg Glu Glu Asn Thr Ser Asn Glu Ser Thr
65                  70                  75                  80
```

-continued

```
gat gtg acc aag gga gac agc aaa aat gct aaa aag aag aat aat aag      288
Asp Val Thr Lys Gly Asp Ser Lys Asn Ala Lys Lys Lys Asn Asn Lys
85                  90                  95

aaa acc agc aaa aat aag agc agc ctg agt agg ggc aac aag aag aaa      336
Lys Thr Ser Lys Asn Lys Ser Ser Leu Ser Arg Gly Asn Lys Lys Lys
100                 105                 110

ccc ggg atg ccc aat gta tct aac gac ctc tca cag aaa cta tat gcc      384
Pro Gly Met Pro Asn Val Ser Asn Asp Leu Ser Gln Lys Leu Tyr Ala
115                 120                 125

acc atg gag aag cat aaa gag gtc ttc ttt gtg atc cgc ctc att gct      432
Thr Met Glu Lys His Lys Glu Val Phe Phe Val Ile Arg Leu Ile Ala
130                 135                 140

ggc cct gct gcc aac tcc ctg cct ccc att gtt gat cct gat cct ctc      480
Gly Pro Ala Ala Asn Ser Leu Pro Pro Ile Val Asp Pro Asp Pro Leu
145                 150                 155                 160

atc ccc tgc gat ctg atg gat ggt cgg gat gcg ttt ctc acg ctg gca      528
Ile Pro Cys Asp Leu Met Asp Gly Arg Asp Ala Phe Leu Thr Leu Ala
165                 170                 175

agg gac aag cac ctg gag ttc tct tca ctc cga aga gcc cag tgg tcc      576
Arg Asp Lys His Leu Glu Phe Ser Ser Leu Arg Arg Ala Gln Trp Ser
180                 185                 190

acc atg tgc atg ctg gtg gag ctg cac acg cag agc cag gac cgc ttt      624
Thr Met Cys Met Leu Val Glu Leu His Thr Gln Ser Gln Asp Arg Phe
195                 200                 205

gtc tac acc tgc aat gaa tgc aag cac cat gtg gag aca cgc tgg cac      672
Val Tyr Thr Cys Asn Glu Cys Lys His His Val Glu Thr Arg Trp His
210                 215                 220

tgt act gtc tgt gag gat tat gac ttg tgt atc acc tgc tat aac act      720
Cys Thr Val Cys Glu Asp Tyr Asp Leu Cys Ile Thr Cys Tyr Asn Thr
225                 230                 235                 240

aaa aac cat gac cac aaa atg gag aaa cta ggc ctt ggc tta gat gat      768
Lys Asn His Asp His Lys Met Glu Lys Leu Gly Leu Gly Leu Asp Asp
245                 250                 255

gag agc aac aac cag cag gct gca gcc acc cag agc cca ggc gat tct      816
Glu Ser Asn Asn Gln Gln Ala Ala Ala Thr Gln Ser Pro Gly Asp Ser
260                 265                 270

cgc cgc ctg agt atc cag cgc tgc atc cag tct ctg gtc cat gct tgc      864
Arg Arg Leu Ser Ile Gln Arg Cys Ile Gln Ser Leu Val His Ala Cys
275                 280                 285

cag tgt cgg aat gcc aat tgc tca ctg cca tcc tgc cag aag atg aag      912
Gln Cys Arg Asn Ala Asn Cys Ser Leu Pro Ser Cys Gln Lys Met Lys
290                 295                 300

cgg gtt gtg cag cat acc aag ggt tgc aaa cgg aaa acc aat ggc ggg      960
Arg Val Val Gln His Thr Lys Gly Cys Lys Arg Lys Thr Asn Gly Gly
305                 310                 315                 320

tgc ccc atc tgc aag cag ctc att gcc ctc tgc tgc tac cat gcc aag     1008
Cys Pro Ile Cys Lys Gln Leu Ile Ala Leu Cys Cys Tyr His Ala Lys
325                 330                 335

cac tgc cag gag aac aaa tgc ccg gtg ccg ttc tgc cta aac atc aag     1056
His Cys Gln Glu Asn Lys Cys Pro Val Pro Phe Cys Leu Asn Ile Lys
340                 345                 350

cag aag ctc cgg cag caa cag ctg cag cac cga cta cag cag gcc caa     1104
Gln Lys Leu Arg Gln Gln Gln Leu Gln His Arg Leu Gln Gln Ala Gln
355                 360                 365

atg ctt cgc agg agg atg gcc agc atg cag cgg act ggt gtg gtt ggg     1152
Met Leu Arg Arg Arg Met Ala Ser Met Gln Arg Thr Gly Val Val Gly
370                 375                 380
```

-continued

```
cag caa cag ggc ctc cct tcc ccc act cct gcc act cca acg aca cca      1200
Gln Gln Gln Gly Leu Pro Ser Pro Thr Pro Ala Thr Pro Thr Thr Pro
385                 390                 395                 400

act ggc caa cag cca acc acc ccg cag acg ccc cag ccc act tct cag      1248
Thr Gly Gln Gln Pro Thr Thr Pro Gln Thr Pro Gln Pro Thr Ser Gln
405                 410                 415

cct cag cct acc cct ccc aat agc atg cca ccc tac ttg ccc agg act      1296
Pro Gln Pro Thr Pro Pro Asn Ser Met Pro Pro Tyr Leu Pro Arg Thr
420                 425                 430

caa gct                                                              1302
Gln Ala


&lt;210&gt; SEQ ID NO 6
&lt;211&gt; LENGTH: 434
&lt;212&gt; TYPE: PRT
&lt;213&gt; ORGANISM: Homo sapiens

&lt;400&gt; SEQUENCE: 6

Lys Pro Lys Arg Leu Gln Glu Trp Tyr Lys Lys Met Leu Asp Lys Ala
1               5                   10                  15

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

Thr Glu Asp Arg Leu Thr Ser Ala Lys Glu Leu Pro Tyr Phe Glu Gly
35                  40                  45

Asp Phe Trp Pro Asn Val Leu Glu Glu Ser Ile Lys Glu Leu Glu Gln
50                  55                  60

Glu Glu Glu Glu Arg Lys Arg Glu Glu Asn Thr Ser Asn Glu Ser Thr
65                  70                  75                  80

Asp Val Thr Lys Gly Asp Ser Lys Asn Ala Lys Lys Lys Asn Asn Lys
85                  90                  95

Lys Thr Ser Lys Asn Lys Ser Ser Leu Ser Arg Gly Asn Lys Lys Lys
100                 105                 110

Pro Gly Met Pro Asn Val Ser Asn Asp Leu Ser Gln Lys Leu Tyr Ala
115                 120                 125

Thr Met Glu Lys His Lys Glu Val Phe Phe Val Ile Arg Leu Ile Ala
130                 135                 140

Gly Pro Ala Ala Asn Ser Leu Pro Pro Ile Val Asp Pro Asp Pro Leu
145                 150                 155                 160

Ile Pro Cys Asp Leu Met Asp Gly Arg Asp Ala Phe Leu Thr Leu Ala
165                 170                 175

Arg Asp Lys His Leu Glu Phe Ser Ser Leu Arg Arg Ala Gln Trp Ser
180                 185                 190

Thr Met Cys Met Leu Val Glu Leu His Thr Gln Ser Gln Asp Arg Phe
195                 200                 205

Val Tyr Thr Cys Asn Glu Cys Lys His His Val Glu Thr Arg Trp His
210                 215                 220

Cys Thr Val Cys Glu Asp Tyr Asp Leu Cys Ile Thr Cys Tyr Asn Thr
225                 230                 235                 240

Lys Asn His Asp His Lys Met Glu Lys Leu Gly Leu Gly Leu Asp Asp
245                 250                 255

Glu Ser Asn Asn Gln Gln Ala Ala Ala Thr Gln Ser Pro Gly Asp Ser
260                 265                 270

Arg Arg Leu Ser Ile Gln Arg Cys Ile Gln Ser Leu Val His Ala Cys
275                 280                 285
```

-continued

```
Gln Cys Arg Asn Ala Asn Cys Ser Leu Pro Ser Cys Gln Lys Met Lys
290                 295                 300

Arg Val Val Gln His Thr Lys Gly Cys Lys Arg Lys Thr Asn Gly Gly
305                 310                 315                 320

Cys Pro Ile Cys Lys Gln Leu Ile Ala Leu Cys Cys Tyr His Ala Lys
325                 330                 335

His Cys Gln Glu Asn Lys Cys Pro Val Pro Phe Cys Leu Asn Ile Lys
340                 345                 350

Gln Lys Leu Arg Gln Gln Gln Leu Gln His Arg Leu Gln Gln Ala Gln
355                 360                 365

Met Leu Arg Arg Arg Met Ala Ser Met Gln Arg Thr Gly Val Val Gly
370                 375                 380

Gln Gln Gln Gly Leu Pro Ser Pro Thr Pro Ala Thr Pro Thr Thr Pro
385                 390                 395                 400

Thr Gly Gln Gln Pro Thr Thr Pro Gln Thr Pro Gln Pro Thr Ser Gln
405                 410                 415

Pro Gln Pro Thr Pro Pro Asn Ser Met Pro Pro Tyr Leu Pro Arg Thr
420                 425                 430

Gln Ala


<210> SEQ ID NO 7
<211> LENGTH: 1816
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (77)..(1561)
<223> OTHER INFORMATION: p70(beta) S6 Kinase gene

<400> SEQUENCE: 7

gagagagaga gagagagaga gagagagaga gagagagaga gagagagaga gagagagaga      60

gagagactcg tgccga atg gca cga ggc cga cgg gcc cgc ggg gcc ggc gcc     112
                  Met Ala Arg Gly Arg Arg Ala Arg Gly Ala Gly Ala
                  1               5                   10

gcc atg gcg gcc gtg ttt gat ttg gat ttg gag acg gag gaa ggc agc       160
Ala Met Ala Ala Val Phe Asp Leu Asp Leu Glu Thr Glu Glu Gly Ser
15                  20                  25

gag ggc gag ggc gag cca gag ctc agc ccc gcg gac gca tgt ccc ctt       208
Glu Gly Glu Gly Glu Pro Glu Leu Ser Pro Ala Asp Ala Cys Pro Leu
30                  35                  40

gcc gag ttg agg gca gct ggc cta gag cct gtg gga cac tat gaa gag       256
Ala Glu Leu Arg Ala Ala Gly Leu Glu Pro Val Gly His Tyr Glu Glu
45                  50                  55                  60

gtg gag ctg act gag acc agc gtg aac gtt ggc cca gag cgc atc ggg       304
Val Glu Leu Thr Glu Thr Ser Val Asn Val Gly Pro Glu Arg Ile Gly
65                  70                  75

ccc cac tgc ttt gag ctg ctg cgt gtg ctg ggc aag ggg ggc tat ggc       352
Pro His Cys Phe Glu Leu Leu Arg Val Leu Gly Lys Gly Gly Tyr Gly
80                  85                  90

aag gtg ttc cag gtg cga aag gtg caa ggc acc aac ttg ggc aaa ata       400
Lys Val Phe Gln Val Arg Lys Val Gln Gly Thr Asn Leu Gly Lys Ile
95                  100                 105

tat gcc atg aaa gtc cta agg aag gcc aaa att gtg cgc aat gcc aag       448
Tyr Ala Met Lys Val Leu Arg Lys Ala Lys Ile Val Arg Asn Ala Lys
110                 115                 120
```

-continued

```
gac aca gca cac aca cgg gct gag cgg aac att cta gag tca gtg aag      496
Asp Thr Ala His Thr Arg Ala Glu Arg Asn Ile Leu Glu Ser Val Lys
125                 130                 135                 140

cac ccc ttt att gtg gaa ctg gcc tat gcc ttc cag act ggt ggc aaa      544
His Pro Phe Ile Val Glu Leu Ala Tyr Ala Phe Gln Thr Gly Gly Lys
145                 150                 155

ctc tac ctc atc ctt gag tgc ctc agt ggt ggc gag ctc ttc acg cat      592
Leu Tyr Leu Ile Leu Glu Cys Leu Ser Gly Gly Glu Leu Phe Thr His
160                 165                 170

ctg gag cga gag ggc atc ttc ctg gaa gat acg gcc tgc ttc tac ctg      640
Leu Glu Arg Glu Gly Ile Phe Leu Glu Asp Thr Ala Cys Phe Tyr Leu
175                 180                 185

gct gag atc acg ctg gcc ctg ggc cat ctc cac tcc cag ggc atc atc      688
Ala Glu Ile Thr Leu Ala Leu Gly His Leu His Ser Gln Gly Ile Ile
190                 195                 200

tac cgg gac ctc aag ccc gag aac atc atg ctc agc agc cag ggc cac      736
Tyr Arg Asp Leu Lys Pro Glu Asn Ile Met Leu Ser Ser Gln Gly His
205                 210                 215                 220

atc aaa ctg acc gac ttt gga ctc tgc aag gag tct atc cat gag ggc      784
Ile Lys Leu Thr Asp Phe Gly Leu Cys Lys Glu Ser Ile His Glu Gly
225                 230                 235

gcc gtc act cac acc ttc tgc ggc acc att gag tac atg gcc cct gag      832
Ala Val Thr His Thr Phe Cys Gly Thr Ile Glu Tyr Met Ala Pro Glu
240                 245                 250

att ctg gtg cgc agt ggc cac aac cgg gct gtg gac tgg tgg agc ctg      880
Ile Leu Val Arg Ser Gly His Asn Arg Ala Val Asp Trp Trp Ser Leu
255                 260                 265

ggg gcc ctg atg tac gac atg ctc act gga tcg ccg ccc ttt acc gca      928
Gly Ala Leu Met Tyr Asp Met Leu Thr Gly Ser Pro Pro Phe Thr Ala
270                 275                 280

gag aac cgg aag aaa acc atg gat aag atc atc agg ggc aag ctg gca      976
Glu Asn Arg Lys Lys Thr Met Asp Lys Ile Ile Arg Gly Lys Leu Ala
285                 290                 295                 300

ctg ccc ccc tac ctc acc cca gat gcc cgg gac ctt gtc aaa aag ttt     1024
Leu Pro Pro Tyr Leu Thr Pro Asp Ala Arg Asp Leu Val Lys Lys Phe
305                 310                 315

ctg aaa cgg aat ccc agc cag cgg att ggg ggt ggc cca ggg gat gct     1072
Leu Lys Arg Asn Pro Ser Gln Arg Ile Gly Gly Gly Pro Gly Asp Ala
320                 325                 330

gct gat gtg cag aga cat ccc ttt ttc cgg cac atg aat tgg gac gac     1120
Ala Asp Val Gln Arg His Pro Phe Phe Arg His Met Asn Trp Asp Asp
335                 340                 345

ctt ctg gcc tgg cgt gtg gac ccc cct ttc agg ccc tgt ctg cag tca     1168
Leu Leu Ala Trp Arg Val Asp Pro Pro Phe Arg Pro Cys Leu Gln Ser
350                 355                 360

gag gag gac gtg agc cag ttt gat acc cgc ttc aca cgg cag acg ccg     1216
Glu Glu Asp Val Ser Gln Phe Asp Thr Arg Phe Thr Arg Gln Thr Pro
365                 370                 375                 380

gtg gac agt cct gat gac aca gcc ctc agc gag agt gcc aac cag gcc     1264
Val Asp Ser Pro Asp Asp Thr Ala Leu Ser Glu Ser Ala Asn Gln Ala
385                 390                 395

ttc ctg ggc ttc aca tac gtg gcg ccg tct gtc ctg gac agc atc aag     1312
Phe Leu Gly Phe Thr Tyr Val Ala Pro Ser Val Leu Asp Ser Ile Lys
400                 405                 410

gag ggc ttc tcc ttc cag ccc aag ctg cgc tca ccc agg cgc ctc aac     1360
Glu Gly Phe Ser Phe Gln Pro Lys Leu Arg Ser Pro Arg Arg Leu Asn
415                 420                 425
```

-continued

```
agt agc ccc cgg gtc ccc gtc agc ccc ctc aag ttc tcc cct ttt gag     1408
Ser Ser Pro Arg Val Pro Val Ser Pro Leu Lys Phe Ser Pro Phe Glu
430                 435                 440

ggg ttt cgg ccc agc ccc agc ctg ccg gag ccc acg gag cta cct cta     1456
Gly Phe Arg Pro Ser Pro Ser Leu Pro Glu Pro Thr Glu Leu Pro Leu
445                 450                 455                 460

cct cca ctc ctg cca ccg ccg ccg ccc tcg acc acc gcc cct ctc ccc     1504
Pro Pro Leu Leu Pro Pro Pro Pro Pro Ser Thr Thr Ala Pro Leu Pro
465                 470                 475

atc cgt ccc ccc tca ggg acc aag aag tcc aag agg ggc cgt ggg cgt     1552
Ile Arg Pro Pro Ser Gly Thr Lys Lys Ser Lys Arg Gly Arg Gly Arg
480                 485                 490

cca ggg cgc taggaagccg ggtgggggtg agggtagccc ttgagccctg             1601
Pro Gly Arg
495

tccctgcggc tgtgagagca gcaggaccct gggccagttc cagagacctg ggggtgtgtc   1661

tgggggtggg gtgtgagtgc gtatgaaagt gtgtgtctgc tggggcagct gtgcccctga   1721

atcatgggca cggagggccg cccgccacac cccgcgctca actgctcccg tggaagatta   1781

aagggctgaa tcatgaaaaa aaaaaaaaaa aaaaa                              1816


<210> SEQ ID NO 8
<211> LENGTH: 495
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 8

Met Ala Arg Gly Arg Arg Ala Arg Gly Ala Gly Ala Ala Met Ala Ala
1               5                   10                  15

Val Phe Asp Leu Asp Leu Glu Thr Glu Glu Gly Ser Glu Gly Glu Gly
20                  25                  30

Glu Pro Glu Leu Ser Pro Ala Asp Ala Cys Pro Leu Ala Glu Leu Arg
35                  40                  45

Ala Ala Gly Leu Glu Pro Val Gly His Tyr Glu Glu Val Glu Leu Thr
50                  55                  60

Glu Thr Ser Val Asn Val Gly Pro Glu Arg Ile Gly Pro His Cys Phe
65                  70                  75                  80

Glu Leu Leu Arg Val Leu Gly Lys Gly Gly Tyr Gly Lys Val Phe Gln
85                  90                  95

Val Arg Lys Val Gln Gly Thr Asn Leu Gly Lys Ile Tyr Ala Met Lys
100                 105                 110

Val Leu Arg Lys Ala Lys Ile Val Arg Asn Ala Lys Asp Thr Ala His
115                 120                 125

Thr Arg Ala Glu Arg Asn Ile Leu Glu Ser Val Lys His Pro Phe Ile
130                 135                 140

Val Glu Leu Ala Tyr Ala Phe Gln Thr Gly Gly Lys Leu Tyr Leu Ile
145                 150                 155                 160

Leu Glu Cys Leu Ser Gly Gly Glu Leu Phe Thr His Leu Glu Arg Glu
165                 170                 175

Gly Ile Phe Leu Glu Asp Thr Ala Cys Phe Tyr Leu Ala Glu Ile Thr
180                 185                 190

Leu Ala Leu Gly His Leu His Ser Gln Gly Ile Ile Tyr Arg Asp Leu
195                 200                 205

Lys Pro Glu Asn Ile Met Leu Ser Ser Gln Gly His Ile Lys Leu Thr
```

-continued

```
210                 215                 220

Asp Phe Gly Leu Cys Lys Glu Ser Ile His Glu Gly Ala Val Thr His
225                 230                 235                 240

Thr Phe Cys Gly Thr Ile Glu Tyr Met Ala Pro Glu Ile Leu Val Arg
245                 250                 255

Ser Gly His Asn Arg Ala Val Asp Trp Trp Ser Leu Gly Ala Leu Met
260                 265                 270

Tyr Asp Met Leu Thr Gly Ser Pro Pro Phe Thr Ala Glu Asn Arg Lys
275                 280                 285

Lys Thr Met Asp Lys Ile Ile Arg Gly Lys Leu Ala Leu Pro Pro Tyr
290                 295                 300

Leu Thr Pro Asp Ala Arg Asp Leu Val Lys Lys Phe Leu Lys Arg Asn
305                 310                 315                 320

Pro Ser Gln Arg Ile Gly Gly Gly Pro Gly Asp Ala Ala Asp Val Gln
325                 330                 335

Arg His Pro Phe Phe Arg His Met Asn Trp Asp Asp Leu Leu Ala Trp
340                 345                 350

Arg Val Asp Pro Pro Phe Arg Pro Cys Leu Gln Ser Glu Glu Asp Val
355                 360                 365

Ser Gln Phe Asp Thr Arg Phe Thr Arg Gln Thr Pro Val Asp Ser Pro
370                 375                 380

Asp Asp Thr Ala Leu Ser Glu Ser Ala Asn Gln Ala Phe Leu Gly Phe
385                 390                 395                 400

Thr Tyr Val Ala Pro Ser Val Leu Asp Ser Ile Lys Glu Gly Phe Ser
405                 410                 415

Phe Gln Pro Lys Leu Arg Ser Pro Arg Arg Leu Asn Ser Ser Pro Arg
420                 425                 430

Val Pro Val Ser Pro Leu Lys Phe Ser Pro Phe Glu Gly Phe Arg Pro
435                 440                 445

Ser Pro Ser Leu Pro Glu Pro Thr Glu Leu Pro Leu Pro Pro Leu Leu
450                 455                 460

Pro Pro Pro Pro Pro Ser Thr Thr Ala Pro Leu Pro Ile Arg Pro Pro
465                 470                 475                 480

Ser Gly Thr Lys Lys Ser Lys Arg Gly Arg Gly Arg Pro Gly Arg
485                 490                 495


<210> SEQ ID NO 9
<211> LENGTH: 14
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic peptide

<400> SEQUENCE: 9

Gln Ala Phe Pro Met Ile Ser Lys Arg Pro Glu His Leu Arg
1               5                   10


<210> SEQ ID NO 10
<211> LENGTH: 14
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic peptide

<400> SEQUENCE: 10
```

-continued

```
Gly Lys Leu Ala Leu Pro Pro Tyr Leu Thr Pro Asp Ala Arg
1               5                   10


<210> SEQ ID NO 11
<211> LENGTH: 14
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic peptide

<400> SEQUENCE: 11

Val Leu Gly Lys Gly Gly Tyr Gly Lys Val Phe Gln Val Arg
1               5                   10


<210> SEQ ID NO 12
<211> LENGTH: 9
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic peptide

<400> SEQUENCE: 12

Lys Arg Lys Arg Gly Arg Pro Arg Lys
1               5


<210> SEQ ID NO 13
<211> LENGTH: 9
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic peptide

<400> SEQUENCE: 13

Lys Arg Gly Arg Gly Arg Pro Gly Arg
1               5


<210> SEQ ID NO 14
<211> LENGTH: 9
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic peptide

<400> SEQUENCE: 14

Lys Arg Gly Arg Gly Arg Pro Arg Gly
1               5


<210> SEQ ID NO 15
<211> LENGTH: 8
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic peptide

<400> SEQUENCE: 15

Lys Gly Arg Gly Arg Pro Lys Lys
1               5


<210> SEQ ID NO 16
<211> LENGTH: 8
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic peptide
```

-continued

<400> SEQUENCE: 16

Arg Thr Arg Gly Arg Pro Ala Asp
1   5

<210> SEQ ID NO 17
<211> LENGTH: 9
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic peptide

<400> SEQUENCE: 17

Lys Arg Lys Arg Gly Arg Pro Arg Lys
1   5

<210> SEQ ID NO 18
<211> LENGTH: 9
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic peptide

<400> SEQUENCE: 18

Lys Arg Gly Arg Gly Arg Pro Arg Leu
1   5

<210> SEQ ID NO 19
<211> LENGTH: 9
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic peptide

<400> SEQUENCE: 19

Lys Arg Val Arg Gly Arg Pro Cys Leu
1   5

<210> SEQ ID NO 20
<211> LENGTH: 9
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic peptide

<400> SEQUENCE: 20

Lys Lys Ser Arg Gly Arg Pro Lys Lys
1   5

<210> SEQ ID NO 21
<211> LENGTH: 9
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic peptide

<400> SEQUENCE: 21

Lys Arg Arg Arg Gly Arg Pro Pro Lys
1   5

<210> SEQ ID NO 22
<211> LENGTH: 9
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:

-continued

<223> OTHER INFORMATION: Synthetic peptide

<400> SEQUENCE: 22

Lys Arg Gly Arg Gly Arg Pro Lys Gly
1   5

<210> SEQ ID NO 23
<211> LENGTH: 9
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic peptide

<400> SEQUENCE: 23

Gly Arg Gly Arg Gly Arg Pro Lys Gly
1   5

<210> SEQ ID NO 24
<211> LENGTH: 9
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic peptide

<400> SEQUENCE: 24

Lys Arg Gly Arg Gly Arg Pro Arg Lys
1   5

<210> SEQ ID NO 25
<211> LENGTH: 9
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic peptide

<400> SEQUENCE: 25

Lys Arg Gly Arg Gly Arg Pro Gly Arg
1   5

<210> SEQ ID NO 26
<211> LENGTH: 9
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic peptide

<400> SEQUENCE: 26

Lys Lys Gly Arg Gly Arg Pro Gly Arg
1   5

<210> SEQ ID NO 27
<211> LENGTH: 9
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic peptide

<400> SEQUENCE: 27

Lys Arg Gly Arg Gly Arg Pro Gly Arg
1   5

<210> SEQ ID NO 28
<211> LENGTH: 9
<212> TYPE: PRT

-continued

<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic peptide

<400> SEQUENCE: 28

Lys Arg Gly Arg Gly Arg Pro Gly Arg
1 5

<210> SEQ ID NO 29
<211> LENGTH: 9
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic peptide

<400> SEQUENCE: 29

Lys Arg Gly Arg Gly Arg Pro Gly Arg
1 5

1. An isolated peptide or derivative thereof comprising a binding domain of p300 acetyltransferase of less than 200 amino acids of SEQ ID NO: 4 wherein the binding domain is capable of binding to and modulating an S6 kinase protein activity.

2-4. (canceled)

5. The isolated peptide or derivative thereof of claim **4**, wherein the site of acetylation is at an amino acid corresponding to residue 516 of an S6 kinase protein.

6. The isolated peptide or derivative thereof of claim **5**, wherein said residue is lysine.

7. The isolated peptide or derivative thereof of claim **1**, wherein the isolated peptide comprises the amino acid sequence of SEQ ID NO: 6 (p300-4).

8. (canceled)

9. An isolated nucleic acid molecule encoding the isolated peptide or derivative thereof of claim **1**.

10. (canceled)

11. A method of identifying an agent which binds to and/or modulates the acetylation of the S6 kinase protein, comprising:

(a) exposing the S6 kinase protein to the agent in the presence of the isolated peptide of claim **1**, and

(b) measuring the specific binding of the isolated peptide to the S6 kinase protein and/or measuring the acetylation of the S6 kinase protein by the isolated peptide, wherein a decrease in the binding of the isolated peptide and/or a change in acetylation of S6 kinase compared to a control is indicative of an agent capable of binding to the S6 kinase protein.

12. (canceled)

13. The method of claim **11** further comprising contacting the S6 kinase protein with the isolated peptide of claim **1** prior to measuring acetylation.

14. The method of claim **11** wherein the acetylation is at a lysine residue at position 516 of the S6 kinase protein.

15. The method of claim **14** wherein the activity of the S6 kinase protein is determined by measuring phosphorylation of one or more additional proteins.

16-20. (canceled)

21. A method of modulating an S6 kinase protein activity, comprising contacting the S6 kinase protein with the isolated peptide or derivative of claim **1**.

22. The method of claim **21**, wherein the S6 kinase protein activity is inhibited.

23. A method of treating a protein kinase related disease in a patient in need of such treatment comprising administering a therapeutically effective amount of a pharmaceutical composition comprising the isolated peptide of claim **1** capable of acetylating the protein kinase.

24-27. (canceled)

28. The method of claim **23**, wherein said protein kinase related disease is selected from the group consisting of cancer, blood vessel proliferative disorders, autoimmune disorders, and metabolic diseases.

29. The method of claim **28** wherein said cancer is selected from the group consisting of squamous cell carcinoma, astrocytoma, Kaposi's sarcoma, glioblastoma, multiple myeloma, lung cancer, bladder cancer, head and neck cancer, melanoma, ovarian cancer, prostate cancer, breast cancer, small-cell lung cancer, glioma, colorectal cancer, genitourinary cancer, gastrointestinal cancer.

30. The method of claim **28** wherein said blood vessel proliferative disorder is selected from the group consisting of diabetic retinopathy, age-related macular degeneration, retinopathy of prematurity, arthritis and restenosis.

31. The method of claim **28**, wherein said autoimmune disorder is selected from the group consisting of lupus erythematosus, discoid lupus erythematosus, subacute cutaneous lupus erythematosus, drug-induced lupus erythematosus, and systemic lupus erythematosus.

32. The method of claim **28**, wherein said metabolic disorder is selected from the group consisting of psoriasis, diabetes mellitus, wound healing, inflammation and neurodegenerative diseases.

33. An antibody which specifically binds to an acetylation site on a kinase protein and is capable of inhibiting acetylation of the kinase protein by an acetyltransferase protein.

34. (canceled)

35. The antibody of claim **33** wherein the S6 kinase comprises the amino acid sequence of SEQ ID NO: 2.

36. The antibody of claim **33** wherein the antibody binds to a region in the S6 kinase comprising amino acid residue 516 of SEQ ID NO: 2.

37. The antibody of claim **34** wherein the amino acid residue is a lysine.

38. A method of identifying an agent that modulates the binding of S6 kinase 2 ("S6K2") protein to a nucleic acid comprising:

(a) exposing the S6 kinase 2 protein to the agent; and

(b) detecting the binding of the S6 kinase 2 protein to the nucleic acid, wherein change in the level of binding compared to a control is indicative of an agent capable of modulating S6 kinase 2 binding to the nucleic acid.

39. The method of claim **38**, wherein the binding to the nucleic acid is inhibited.

40. The method of claim **38** further comprising measuring of the activity of S6 kinase 2 protein.

41-50. (canceled)

51. The method according to claim **38**, wherein the S6 kinase 2 protein is human S6 kinase 2 (SEQ ID NO: 8).

52-54. (canceled)

55. An isolated protein comprising one or more mutations at any amino acid residue corresponding to amino acids 487 to 495 of SEQ ID NO: 8.

56-60. (canceled)

61. A nucleic acid molecule encoding the isolated protein of claim **55** or a fragment thereof.

62-67. (canceled)

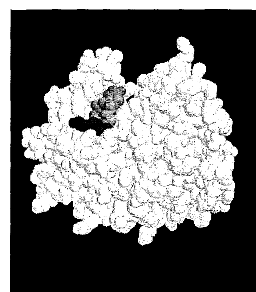
* * * * *

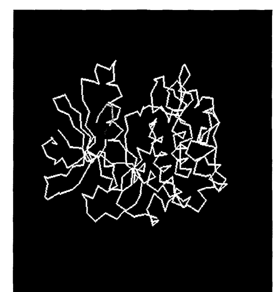
patsnap

专利名称(译)	S6激酶蛋白活性的调节及相关方法		
公开(公告)号	US20090181003A1	公开(公告)日	2009-07-16
申请号	US11/990226	申请日	2006-08-08
[标]申请(专利权)人(译)	痛风IVAN FENTON TIMOTHY ROBERT 伊斯梅尔HEBA		
申请(专利权)人(译)	痛风IVAN FENTON TIMOTHY ROBERT 伊斯梅尔HEBA		
当前申请(专利权)人(译)	痛风IVAN FENTON TIMOTHY ROBERT 伊斯梅尔HEBA		
[标]发明人	GOUT IVAN FENTON TIMOTHY ROBERT ISMAIL HEBA		
发明人	GOUT, IVAN FENTON, TIMOTHY ROBERT ISMAIL, HEBA		
IPC分类号	A61K38/45 C12N9/10 C12N15/11 G01N33/53 C12Q1/48 C12N9/99 C07K16/18 C12Q1/68		
CPC分类号	C12N9/1029		
优先权	60/706010 2005-08-08 US 60/785981 2006-03-27 US		
外部链接	Espacenet USPTO		

摘要(译)

本发明包括通过实现它们的乙酰化或它们与核酸的结合来调节激酶活性的试剂及其使用方法。因此，本发明包括通过p300实现其乙酰化来调节S6激酶。本发明还包括通过影响其与DNA的结合来调节S6激酶2。





Accessibilty data: Lys 99 - 63.8
Lys 104: 75.3

Red & Blue: P-loop lysines
Green: Rest of P-Loop
Yellow: Thr 252

S6K1 Human	VLGKGGYGKVF
S6K2 Human	VLGKGGYGKVF
S6K D.melanogaster	VLGKGGYGKVF
S6K C.elegans	VLGKGGYGKVF
S6K A.thaliana	VVGKGAFGKVY
PKB Human	LLGKGTFGKVY
PKC Human	VLGKGSFGKVM
PKA Human	TLGTGSFGRVM