



(72) DENTON, REX, US

(72) BARON, ROLAND, US

(71) YALE UNIVERSITY, US

(51) Int.Cl.⁷ C07K 14/435, C07K 14/705, C07K 14/47, C12N 1/21,
C12N 15/11, C12N 15/10, C12N 5/10

(30) 1997/07/28 (60/053,869) US

(30) 1997/07/30 (60/054,210) US

(54) **RECEPTEUR β D'OESTROGENES ET ISOFORMES**

(54) **NOVEL ESTROGEN β RECEPTOR AND ISOFORMS**

(57) La présente invention concerne l'acide aminé et la séquence nucléotidique du gène complet du récepteur .beta. d'oestrogène (ER β_c) ainsi que les séquences de protéine associées. Au vu de cette découverte, la présente invention concerne des procédés d'identification d'agents bloquant ou augmentant la régulation transcriptionnelle induite par l'ER β , des procédés permettant de déterminer si l'ER β_c , ou ses isoformes, sont exprimés dans des tissus ou des cellules, et des procédés d'identification et d'utilisation des agents bloquant la régulation transcriptionnelle des gènes par l'ER β_c , ou ses isoformes, qui, à leur tour, modulent d'autres processus biologiques et pathologiques.

(57) The present invention provides the amino acid and nucleotide sequence of the complete estrogen receptor beta (ER β_c) gene and related protein sequences. Based on this disclosure, the present invention provides methods for identifying agents that block or augment ER β_c mediated transcriptional regulation, methods to determine wheter ER β_c or its isoforms is being expressed in tissues or cells and methods of identifying and using agents that block the transcriptional regulation of genes by ER β_c or its isoforms, which in turn modulates other biological and pathological processes.

PCTWORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C07K 14/435, 14/47, 14/705, C12N 1/21, 5/10, 15/10, 15/11	A1	(11) International Publication Number: WO 99/05170 (43) International Publication Date: 4 February 1999 (04.02.99)
(21) International Application Number: PCT/US98/15539 (22) International Filing Date: 28 July 1998 (28.07.98) (30) Priority Data: 60/053,869 28 July 1997 (28.07.97) US 60/054,210 30 July 1997 (30.07.97) US (71) Applicant (for all designated States except US): YALE UNIVERSITY [US/US]; 451 College Street, New Haven, CT 06520 (US). (72) Inventors; and (75) Inventors/Applicants (for US only): DENTON, Rex [US/US]; 129 Hunters Trail, Madison, CT 06443 (US). GALIEN, Rene [-/FR]; Hoechst Marion Roussel, 102, route de Noisey, F-93235 Romainville Cedex (FR). RISCHE-RIGON, Michele [-/FR]; Hoechst Marion Roussel, 102, route de Noisey, F-93235 Romainville Cedex (FR). FRASION, Caroline [-/FR]; Hoechst Marion Roussel, 102, route de Noisey, F-93235 Romainville Cedex (FR). (74) Agents: ADLER, Reid, G. et al.; Morgan, Lewis & Bockius LLP, 1800 M Street, N.W., Washington, DC 20036 (US).		(81) Designated States: CA, JP, US, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE). Published <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>
(54) Title: NOVEL ESTROGEN β RECEPTOR AND ISOFORMS (57) Abstract The present invention provides the amino acid and nucleotide sequence of the complete estrogen receptor beta ($ER\beta_c$) gene and related protein sequences. Based on this disclosure, the present invention provides methods for identifying agents that block or augment $ER\beta_c$ mediated transcriptional regulation, methods to determine whether $ER\beta_c$ or its isoforms is being expressed in tissues or cells and methods of identifying and using agents that block the transcriptional regulation of genes by $ER\beta_c$ or its isoforms, which in turn modulates other biological and pathological processes.		

NOVEL ESTROGEN RECEPTOR β AND ISOFORMS

FIELD OF THE INVENTION

The present invention relates to a novel complete mammalian estrogen receptor β , referred to as ER β_c , its polypeptide sequence, the nucleic acid sequence encoding ER β_c and methods of making or expressing ER β_c . The present invention also relates to methods of screening for drugs which modulate the interaction of estrogens and ER β_c as well as methods of diagnosing and/or treating diseases involving ER β_c or its isoforms. This application is related to U.S. Provisional applications 60/053,869 and 60/054,210, which are herein incorporated by reference.

BACKGROUND OF THE INVENTION

A. Estrogen

Estrogens are a class of naturally occurring steroid hormones which are produced in the ovaries and other tissues of the body including the testis. Estrogens are known to directly influence the growth, differentiation and function of specific target tissues and organs in humans and animals. These specific tissues and organs also include the mammary gland, uterus, prostate, pituitary, brain and liver. Estrogens also play an important role in bone maintenance and in the cardiovascular system, where estrogens have certain cardio-protective effects. In bone, both osteoclasts and osteoblasts have been reported to respond to estrogens with estrogen withdrawal leading to increased turnover and bone loss. A variety of naturally occurring and chemically synthesized estrogens have been identified and characterized, perhaps the best known of which is the endogenous estrogen, estradiol-17 beta (also known as E₂).

B. Estrogen Receptors

Estrogens, as a class of hormones, act by binding to the ligand binding domain (LBD) of an intracellular protein identified as an "estrogen receptor" (ER). The presence of this intracellular ER provides and accounts for both cell proliferation and protein synthesis by estrogen-dependent cells. In the absence of the estrogen hormone, the estrogen receptor is biologically inactive both *in vivo* and *in vitro*; and, if the cells or tissues are homogenized

and fractionated into cytosol and nuclear fractions, the estrogen receptor is found in the nucleus and may also be detected the cytosol.

The known estrogen receptors are members of the well studied family of gene regulatory proteins referred to as the steroid hormone receptor family. Nuclear receptors, such as steroid hormone receptors, have a modular structure with six distinct regions. The N-terminal domain is the A/B region which includes a non-ligand dependent activation function (See Fig. 1a). The C region is the DNA binding domain (DBD). The D region contains nuclear localization signals. Finally, the E domain contains the ligand binding domain (LBD) and the ligand-dependent transaction function. Kuiper *et al.*, Endocrinology 138(3): 863 (1997); Tremblay *et al.*, Mol. Endocrin 11(3): 333 (1997). The central DBD is typically about 100 amino acids. Like the other members of the steroid-hormone receptor family, estrogen receptors are activated by the binding of estrogen to the C-terminal LBD. The receptor proteins enable cells to respond to various lipid-soluble hormones by activating or repressing specific genes, through the interaction between the steroid hormone and its receptor. Steroid hormone receptors are distinguishable from other nuclear receptors in a number of respects, including the nature of their ligands, their association (in the unliganded state) with a repertoire of heat-shock proteins and the fact that they may bind to hormone response elements as homodimers. Mosselman *et al.*, FEBS Letters 392: 49 (1996); Komm *et al.*, Science 241: 81 (1988); Burch *et al.*, Mol. Cell. Biol. 8(3): 1123 (1988).

The conventional model for steroid hormone action has assumed that steroid hormone receptors act as transcriptional regulators only when complexed with their ligands. It has, however, become evident that the majority of steroid receptors are present in the cell nucleus even in the absence of ligand. The presence of the receptors in the nucleus, despite the absence of hormone, suggests possible additional regulatory functions for the receptor in its unbound state. For example, the thyroid hormone receptors (TR) have a dual regulatory role: in the presence of hormone they function as transcriptional activators, whereas in the absence of hormone, TRs are response element (TRE) specific transcriptional repressors.

The first estrogen receptor discovered was ER α , which was known for the past ten years merely as ER. The human ER (hER) is composed of 595 amino acids in its unbound

state and is approximately 67,000 Daltons. In the absence of estrogen-binding, the ER α protein can be located *in vitro* within the cytosol.

Transcription of ER α occurs from two separate promoters, P0 and P1, although no functional mapping has been previously published. P1 represents the major ER α transcriptional start site. The P1 start site is predominantly utilized in human mammary epithelial cells (HMEC) and is the major start site in ER α -positive human breast carcinomas. Multiple start sites have been identified for the P0 promoter. Studies of the murine ER α gene identified 10 start sites spanning approximately 60 bases, and there is a start site at - 1,994 (from the P1 start site) in human cells, which would agree closely with the major murine P0 start site. Transcription from the P0 promoter is characteristic of human endometrial tissue and can account for 12 to 33% of ER α transcription in breast carcinoma cells. The genetic regulatory control elements of the recently discovered ER β gene have yet to be delineated. Kuiper *et al.*, (1996 and 1997); Tremblay *et al.*, (1997); and Mosselman *et al.*, (1996). It remains to be determined whether the ER β gene contains regulatory elements, such as promoters and enhancers, that are similar or function in a manner analogous to those described for ER α .

In soluble systems and under set conditions, the ER α protein can be found in various molecular forms with sedimentation coefficients of 8S, 5S or 4S as determined by sucrose density gradient analysis. The 8S form of ER α protein is believed to be the inactivated, untransformed form of ER α protein associated with the unbound, inactive state of estrogen receptor in the absence of estrogen. The 4S ER α protein is a monomeric protein molecule that can be generated from the 8S form *in vitro*. The 4S form binds to both nuclei and DNA-cellulose *in vitro*; it is generally termed the "activated but untransformed" estrogen receptor protein. The 5S form of ER α is a dimeric protein molecule, which is created by the conversion of the 4S ER α protein via a bimolecular reaction. It is generally believed that the 5S form of ER α protein is both "activated and transformed," and therefore is the biologically active entity which binds to the DNA within the nuclei. Moreover, it is also this 5S form which is found associated with the nuclei subsequent to the administration of estradiol *in vivo*. Already it has been demonstrated that both ER α and ER β can form heterodimers (Kuiper and Gustafsson, FEBS 410: 87 (1997)).

Although the precise interactions between ER α and estrogens remain poorly understood, the generally accepted mechanism of action and sequence of events is believed to be as follows: When an estrogen, such as estradiol 17 beta (E₂), is introduced to the target cells and tissues, there is specific binding between the estrogen and the ER α protein which results in the formation of an estrogen/ER α protein complex. Also, at a time subsequent to hormone binding, a process termed activation and/or transformation ensues leading to the formation of functional estrogen/hormone receptor complexes possessing a high affinity for the nuclear components, the DNA of the target cell. Once the hormone/receptor protein complex is physically formed, it binds to the chromatin at specific binding sites on the chromosomes and regulates messenger ribonucleic acid (mRNA) transcription. If transcription is up-regulated, new messenger RNA (mRNA) is synthesized, chemically modified and exported from the nucleus into the cytoplasm of the cell where ribosomes then translate the mRNA into new proteins; the hormone/receptor protein complex can also down-regulate mRNA transcription. This constitutes the well recognized estrogenic effect that occurs within cells and involves the regulation of new protein synthesis and concomitant new cell growth/proliferation or differentiation. It remains unclear whether ER β shares the same mechanisms of action and in the same order as have been demonstrated for ER α . Certainly, the localization of ER β along with the manner in which it modulates transcription will be at least grossly similar to ER α ; however, affinities for certain DNA sequences, as well as receptor ligands likely will differ between ER α and ER β , as there is a 97% and 60% identity respectively between the DBD and LBD sequences between the two estrogen receptors. Tremblay *et al.*, (1997).

The clinical significance of the ER in the management of breast cancer is well known. Tamoxifen, a substituted triphenylethylene antiestrogen, is a partial antagonist that is used in the management of ER-positive breast tumors. Gallo *et al.*, Semin. Oncol. 24: S1 (1997). Generally, the expression of the receptor is usually associated with a better prognosis and is less metastatic. Bonetti *et al.*, Breast Cancer Res. Treat. 38(3): 289 (1996). However, in many cases the tumors are either ER-negative or contain splice variants that are commonly biologically inactive. Hence, there is interest to understand how ER gene regulation, as well as the editing of the ER message, contribute to the development of mammary cancer and its clinical outcome with chemotherapy drugs such as tamoxifen. For

more background, see Gallo *et al.*, (1997); Kangas, Acta Oncol. 31(2): 143 (1992); Evans *et al.*, Bone 17(4S): 1815 (1995); Safarians *et al.*, Cancer Res. 56(15): 3560 (1996).

More recently, estrogen receptors have been linked to bone loss associated with postmenopausal osteoporosis. Paralleling that discovery has been the fact that certain
5 antiestrogens (*e.g.*, tamoxifen, raloxifene, droloxifene and tamoxifen methiodide), which by definition block the actions of estrogens, stimulate only the skeletal muscle tissues and have no corresponding stimulatory effect in the uterus or mesometrial fat. Somjen *et al.*, J. Steroid Biochem. Mol. Biol. 59: 389 (1996); Grasser *et al.*, J. Cell Biochem. 65: 159 (1997). These antiestrogens have been termed selective estrogen receptor modulators
10 (SERMs); they typically possess estrogen agonist-like activity on bone tissues and serum lipids, while displaying potent estrogen antagonist properties in the breast and uterus. The observed paradoxical effects observed between the different estrogen receptor agonists and antagonists most likely corresponds to response differences to the antiestrogens between ER α and ER β . Tremblay *et al.*, 361 (1997).

15 Estrogen receptors are also present in human and rat prostate, as evidenced by ligand binding studies. In contrast to androgen receptors, the major part of the estrogen receptors are localized in the stroma of the rat prostate, although the epithelial cells of the secreting alveoli contain ER. Estrogens are, in addition to androgens, implicated in the growth of the prostate, and consequently estrogens have been implicated in the pathogenesis of benign
20 prostatic hyperplasia. Habenich *et al.*, J. Steroid Biochem. Mol Biol. 44: 557 (1993); Kuiper *et al.*, PNAS 93: 5930 (1996). Diethylstilbesterol (DES), a stilbene estrogen with an increased affinity for ER, is used to treat prostatic hyperplasia and carcinoma. Goethuys *et al.*, Am. J. Clin. Oncol. 20(1): 40 (1997); Aprikian *et al.*, Cancer 71(12): 3952 (1993). Therefore, identifying the tissues and diseases that express ER β likely will prove helpful in
25 the treatment of diseases involving ER β .

Estrogen has also been demonstrated to prevent osteoporosis. Postmenopausal osteoporosis, the most common bone disease in the developed world, is associated with estrogen deficiency. This deficiency increases generation and activity of osteoclasts, large multi-nuclear cells involved with bone resorption. Estrogen has been demonstrated to
30 down-regulate osteoclast formation and function. Tamoxifen has been demonstrated to possess estrogenic effects on bone resorption likely through tamoxifen-induced osteoclast

apoptosis. Hughes *et al.*, Nat. Med. 2(10): 1132 (1996). Isolation of additional reagents that inhibit progression of osteoporosis would be beneficial in treating postmenopausal women suffering from the disease.

5 C. ER β

Following the cloning of estrogen receptor α (ER α) 10 years ago, there was general acceptance that only one ER gene existed and consequently only one subtype of ER, ER α . This contrasted sharply with other members of the nuclear receptor superfamily, where multiple forms have been reported, *e.g.*, thyroid hormone receptor (TR) α and β and retinoic acid receptor (RAR) α , β , and γ .
10

Recently, a novel rat ER cDNA was cloned from rat prostate and ovary tissues and named ER β subtype to distinguish it from the previously cloned ER cDNA, now named the ER α subtype. ER β was partially isolated from cDNA libraries from human testis, mouse ovaries and rat prostate, which are not generally considered to be major estrogen target
15 tissues. The estrogen receptor subtype initially discovered was termed ER β , but for purposes of this invention will be termed the incomplete ER β (ER β_i) to differentiate it from the complete ER β (ER β_c or ER β -3) of the present invention, or the three claimed alternatively spliced isoforms (ER β -1, ER β -2 and ER β -4) of this invention. Mosselman *et al.*, (1996); Kuiper *et al.*, (1996 and 1997); Tremblay *et al.*, (1997). "ER β -3" refers to the
20 sequence as isolated from mouse ovaries or its analogous sequence in other mammalian species. "ER β_c " refers to the sequence that encodes the complete ER β , which includes the novel 192 bp at the 5' terminus of exon 1 and the newly described exon 5B; ER β_c includes ER β -3, the complete sequence that encodes the nine exons of murine ER β . ER β_i , as characterized using the clones obtained from mouse ovary tissue, encodes a protein that has
25 a molecular weight of approximately 62 kDa and has a 60 kilobase (Kb) gene size. The isolated mouse ER β_i gene, called *Estrb* by Tremblay *et al.*, mapped to the central region of chromosome 12; the central region of mouse chromosome 12 shares homology with human chromosome 14q, suggesting that *Estrb* may lie here as well. Tremblay *et al.*, (1997).

The ER β_i cDNA encoded a predicted protein of 485 amino acids and had a
30 calculated molecular weight of 54,200. Kuiper *et al.*, (1997). This protein, described by Kuiper *et al.*, as Clone 29 (herein ER β_i), displays high affinity binding of estrogens, and in a

transactivation assay system, it activates expression of an estrogen response element (ERE) containing a reporter gene construct in the presence of estrogens. Kuiper *et al.*, (1996). Alignment of the ligand binding domain (LBD) of ER α (rat, mouse and human) and ER β_1 (rat) uncovered various regions of conservation, whereas other segments are non-conserved. Kuiper *et al.*, (1996). The DNA binding domain (DBD) and C-terminal LBD of ER β_1 is highly homologous to the rat ER α . Kuiper *et al.*, (1997); Tremblay *et al.*, (1997).

ER β_1 was isolated in an effort to clone and characterize novel nuclear receptors or unknown isoforms of existing receptors. Degenerate primers were designed based on conserved regions within the DBD and LBD of nuclear receptors. Using these primers in conjunction with Polymerase Chain Reaction (PCR), rat prostate mRNA was amplified. One targeted tissue was the prostate, an organ of interest given the high incidence of prostate cancer and benign prostatic hyperplasia. Nearly all prostate tumors eventually become androgen-independent, at which point they are beyond clinical control. Kuiper *et al.*, (1996). Therefore, the possibility existed that, apart from androgens, other (steroid) hormones or locally produced factors interacted with nuclear receptors and modulated the cell proliferation, differentiation and apoptosis of the normal prostate. For example, human testicular receptors 2 and 4 (TR2 and TR4) and the estrogen receptor (ER)-related receptors (ERR1 and ERR2) are examples of orphan nuclear genes whose receptors are expressed in the prostate. Kuiper *et al.*, (1996).

Tissue expression of ER β_1 revealed additional differences from ER α expression. Some tissues contain exclusively ER α (*i.e.*, uterus, pituitary, epididymis, and kidney). Other tissues display equal or greater levels of ER β_1 RNA and may be expressed preferentially in the different cell types of an organ (*i.e.*, ovary and prostate). Kuiper *et al.*, (1996). In brain, ER β_1 appears to be a conspicuous fraction of the ER subtype RNA. Although Northern blots did not detect ER β_1 expression in peripheral blood lymphocytes, the initial PCR fragment of ER β_1 cloned by Mosselman was acquired from these cells. Thus, the ER β_1 subtype may play a significant role in estrogen action in brain, ovary, prostate, hypothalamus and possibly other tissues. Mosselman *et al.*, at 52 (1996); Byers *et al.*, Mol. Endocrinol. 11(2): 172 (1997); and Shughrue *et al.*, Steroids 61(12): 678 (1996).

In addition to differences in tissue expression, the order of competition for physiological estrogens and stilbene estrogens, which form a diphenolic resonance structure,

for ER α versus the ER β _i isoform was also observed to vary. Kuiper *et al.*, (1997). These differences may result from the protein sequences differences observed between ER α and ER β _c, as demonstrated in the comparison of the two murine subtypes in Figure 6. However, the order of affinity for the tested triphenylethylene antiestrogens however, was the same for both subtypes: 4-OH-tamoxifen >> nafoxidine > clomifene > tamoxifen. Kuiper *et al.*, (1997). Such ligand binding differences most likely, portend different drug therapies for ER α versus ER β dependent disease. Another incongruity between the two ER subtypes is the agonistic-antagonistic difference in response to estrogens. This observed paradoxical disparity may relate not only to binding affinity, but, more importantly to the presence of specific activation function domains located in ER β _c (*e.g.*, AF-1 or AF-2). Tremblay *et al.*, (1997).

Substantial interest exists in determining the individual and perhaps combined roles that both ER β and ER α play in carcinomas, increased estrogen turnover and bone loss. For example, tamoxifen augments bone growth, whereas it is an antagonist to ER α positive breast cancer. Gallo *et al.*, (1997); Delmas *et al.*, J. Clin. Oncol. 15: 955 (1997). This incongruous observation with tamoxifen administration may arise from the drug's different interactions with ER α and ER β . The mixed agonist-antagonist or pure antagonist actions observed with antiestrogens may result more specifically from binding differences between the activation function domains, AF-1 and AF-2. For example, studies using the estrogen antagonists 4-hydroxytamoxifen (OHT) and ICI 164,384 indicated that although both compounds blocked estrogen effects, their mode of action differed: the mixed agonist/antagonist OHT inhibited only AF-2 function, while the pure antiestrogen ICI 164,384 inhibited activation by both AF-1 and AF-2. Using the mouse ER β _i, all antagonists tested effectively inhibited E₂ activity. In contrast to ER α , OHT displayed no agonistic activity on ER β _i. Tremblay *et al.*, (1997). Therefore, once the underlying ER subtype responsible for a particular disease state is determined (*e.g.*, ER α positive breast cancer), one may have a more accurate means of prognosticating the estrogen receptor related disease outcome; one may accurately follow therapies; one may develop gene specific and isoform specific therapies targeting diseases influenced by ER α and/or ER β ; and one may provide for opportunities for varying the aggressiveness of the therapy.

SUMMARY OF THE INVENTION

The present invention is based, in part, on the isolation and identification of the complete murine (m) estrogen receptor β gene (*mER β -3*) and two alternatively spliced isoforms, *e.g.*, *mER β -1* and *mER β -2* and a third isoform isolated from rat (r) ovaries, *rER β -4*. More broadly, the invention relates to the corresponding ER β_c gene (including the human gene) and to certain mammalian receptors (denoted herein as ER β -1, ER β -2, ER β -3 and ER β -4). The ER β_i sequence has been published by other laboratories, which had prematurely claimed that ER β_i represented the complete *ER β* gene (*ER β_c*).

The present invention further provides nucleic acid molecules that encode the *mER β -1*, *mER β -2*, *mER β -3* and *mER β -4* proteins. Such nucleic acid molecules can be in an isolated form or can be operably linked to expression control elements or vector sequences.

The present invention also provides methods of identifying other alternatively spliced forms of the *mER β -3*, the analogous *mER β -3* and corresponding ER β_c as expressed in different animal species or additional ER subtypes. Specifically, the nucleic acid sequence of *mER β -3* can be used as a probe or to generate PCR primers to identify nucleic acid molecules that encode other members of the ER β_c family of proteins. The nucleic acid molecules encoding *mER β -1*, *mER β -2*, *mER β -3* or *rER β -4* can be used to identify and isolate the *ER β -3* gene or corresponding ER β_c in other mammalian species, and has been used to isolate the *ER β -3* analog in human DNA.

The present invention further provides antibodies that recognize and bind to the ER β_c protein or the *mER β -3* protein or its isoforms. Such antibodies can be either polyclonal or monoclonal. Particularly preferred are antibodies that are specific for the complete receptor protein, ER β_c , as opposed to antibodies against the previously known receptors, *e.g.*, ER α and ER β_i . More specifically, the invention claims an anti-peptide antibody that distinguishes between ER β_i and ER β_c . Antibodies that bind to the ER β_c protein can be utilized in a variety of diagnostic and prognostic formats and therapeutic methods. Alternatively, antibodies that can distinguish between the complete form, ER β_c , and its isoforms may also be useful for purposes of diagnosis and treatment of ER β subtype based disease.

The present invention further provides methods for reducing, blocking or augmenting the association of an estrogen and other agonists and antagonists with the ER β_c protein. For example, the association of an ER β -3 protein with a cytoplasmic signaling partner, such as estradiol, can be blocked or reduced by contacting the ER β -3 protein with a compound that blocks the binding of estradiol or other estrogen-like agonists or antagonists (e.g., estrogens, stilbene estrogens or triphenylethylene antiestrogens). Tora *et al.*, Cell 59: 447 (1989); Berry *et al.*, EMBO 9: 2811 (1990). Additionally, as the proteins are allosteric, the association of the ligand with ER β can also be influenced, in theory, by the dimer partner. Therefore, identifying agents that modulate ER β dimerization may pose another means of manipulating ER β regulation.

Blocking the interaction between the ligand and ER β -3 or one of its isoforms can be used to modulate biological and pathological processes that require such a ligand bound complex to mediate transcription. Such methods and agents can be used to modulate cellular proliferation, differentiation, DNA synthesis or cell cycle distribution.

The present invention further provides methods for isolating ER β_c or ER β_c protein isoforms (e.g., ER β -1, ER β -2, ER β -3 and ER β -4) that regulate transcription. For example, ER β -3 ligand binding partners, e.g., estrogen, are isolated using the ER β -3 protein or ligand binding portions thereof. Alternatively, for example, the DNA sequences that the ER β -3 protein binds can be determined, for example, utilizing electrophoretic mobility shift assays (EMSA), yeast two hybrid assays, or by affinity selection and degenerate ERE consensus sequences using the DNA binding domains (DBDs) of ER β_c or its isoforms. Berkowitz and Evans, J. Biol. Chem. 267(10): 7134 (1992); Nawaz *et al.*, Gene Expr. 2(1): 39 (1992); Mosselman *et al.*, (1996).

The invention also describes methods to screen compounds that can distinguish between ER α and ER β_c and its isoforms (e.g., ER β -1, ER β -2, ER β -3 and ER β -4). These methods will include methods of determining whether the compound binds and either functionally acts as an agonist or an antagonist with regard to each estrogen receptor. One method to determine whether compounds act in an agonistic or antagonistic fashion would use ER β_c in a yeast two hybrid system. Such methods have been previously employed to test the interaction of certain drugs with ER α and recognized by those of ordinary skill in

the art. See Ichinose *et al.*, Gene 188: 95 (1997); Collins *et al.*, Steroids 62: 365 (1997); Jackson *et al.*, Mol. Endocrinol. 11: 693 (1997).

The biological and pathological processes that require estrogen/ER β_c complex can be modulated further by using gene therapy methods. Additional genetic manipulation within an organism can be used to alter ER β_c gene expression or the production of a ER β_c protein. For example, an ER β -3 gene can be introduced into a mammal deficient for ER β -3 protein to correct the genetic deficiency; peptide modulators of ER β -3 activity can be produced within a target cell using genetic transfection methods to introduce into the target cells nucleic acid molecules encoding the modulators; and the ER β -3 gene can be introduced or deleted in a non-human mammal to produce animal models expressing ER β -3 gene abnormalities or delete the gene entirely (*e.g.*, knock-out mice). The latter application, ER β -3 transgenic animals, is particularly useful for identifying agents *in vivo* that modulate ER β -3 activity and perhaps even other genes that encode proteins that influence ER β -3 actions. The use of nucleic acids for antisense and triple helix therapies and interventions are also expressly contemplated.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1

Nucleotide sequence, deduced amino acid sequence and putative domain structure of the complete murine ER β_c gene (*mER β -3*)

Fig. 1 (a). Illustrates the location of each of the nine exons comprising the clone of the complete murine ER β_c , *mER β -3*, and the splicing domains that yield the different alternatively spliced isoforms of *mER β -3*. The numbers directly above the lines signifying the exons represented by terminal nucleotides of the exon. The sizes of the nine exons in base pairs (bp) and the encoded amino acid (a.a.) sequence for each of the exons and splice variants derived from mouse ovaries is indicated. The 1,704 nucleotides of *mER β -3* encodes a 567 amino acid protein. The letters (A through F) refer to regions of homology shared by all members of the steroid receptor super family. Green *et al.*, Cold Spring Harbor Symposia on Quantitative Biology 51(2): 751-8 (1986). Region C corresponds to

the DNA binding domain (DBD). Region E is the ligand binding domain (LBD). The newly described exon 5B lies within the LBD. Exon 5B starts with GTCCTCA and stops with CCCAAG. The shaded regions in the rendering depict the amino terminus that is included in all mER β and rER β isoforms and the additional exon (exon 5B) that is included in the full length (*mER β -3*) as well as the alternative spliced rat isoform, *rER β -4*. The deletion of exon 6 in *rER β -4* results in a frame-shift and the juxtaposition of an in-frame stop codon causing the protein to be truncated, as indicated.

Clone mER β -1 is 1,650 bp in length. It contains a previously undescribed 192 bp located in the 5' end of exon 1 as well as the 7 other described exons; *mER β -1* lacks the newly described exon 5B. The isolated isoform *mER β -2* is 1,533 bp and lacks both exon 3 and exon 5B. Isoform *rER β -4*, isolated from rat tissue, is 1,570 bp. Although *rER β -4* possesses the new exon 5B, it lacks exon 6. The loss of exon 6 results in a frame shift that causes translation to terminate at a stop codon located in exon 7.

Fig. 1 (b). The full length sequence of murine mER β_c (mER β -3 clone). The additional sequence included in all mER β clones (mER β -1, mER β -2), as well as the alternatively spliced rat isoform, (*rER β -4*) is noted in underlined bold type. The sequence included in the ninth exon, exon 5B, is presented in lower case letters beginning at base 1,149.

Figure 2

Amino acid sequences of the alternatively spliced isoforms of the *mER β -3*

Fig. 2 (a). Deduced amino acid sequences for alternative splice variant mER β -1. The polypeptide sequence shared by all 3 of the alternatively spliced isoforms is indicated by the underlined sequence in bold characters. The mER β -1 protein contains 549 amino acid residues.

Fig. 2 (b). Deduced amino acid sequences for alternative splice variant mER β -2. The alternatively spliced mER β -2 is 510 amino acid residues in length.

Fig. 2 (c). Deduced amino acid sequences for a rat alternative splice variant rER β -4. This splice variant was obtained from rat ovaries. The deletion of exon 6 produces a frame shift causing a truncation that terminates 13 amino acids beyond the translated exon 5B; the resulting rER β -4 protein likely is 414 residues long. The italicized, underlined, bold

characters (residues 1-64) represent the polypeptide encoded by the novel 192 nucleotides located at the 5' terminus of exon 1. The characters indicated in bold and underlined represent the polypeptide encoded by exon 5B. The "*" refers to a translated stop codon.

5

Figure 3

Tissue Specific Expression of mER β -3 Protein Detected by Western Blot

Fig. 3 (a). Using anti-peptide antibodies (Antibody 1068) raised against N-CSSDPHWHVAQTKSAVPR-OH (the sequence encoded by exon 5B), the mER β -3 protein was observed in a Western blot of human ovary, mouse ovary, rat ovary, ROS 17/2.8 cells, and murine primary osteoblasts protein extracts.

10

Fig. 3 (b). Western blot of human ovary, mouse ovary, rat ovary, ROS 17/2.8 cells, and murine primary osteoblasts protein extracts probed with antibody 1068 pre-immune sera. The protein extracts of each lane of both Figures 3(a) and 3(b) are: lane 1, human ovary; lane 2, mouse ovary; lane 3, rat ovary; lane 4, ROS 17/2.8 cells; lane 5, ROS 17/2.8 cells treated with 100 nM estradiol for 16 hours; lane 6, murine primary osteoblasts.

15

Figure 4

Tissue Specific Expression of rER β DNA Detected by Southern Blot of RT-PCR Products

Fig. 4 (a). Total RNA from rat ovarian and ROS 17/2.8 cells amplified for 35 cycles using an oligo that can detect rER β . Each lane in Fig. 4 (a) contains PCR products derived from the following types of RNA: lane 1, control, no RNA; lane 2, rat ovarian RNA (0.1 μ g); lane 3, ROS 17/2.8 cells (0.1 μ g); lane 4, rat ovarian RNA control (0.1 μ g), no reverse transcriptase (RT); and lane 5, ROS 17/2.8 total RNA (0.1 μ g), no RT.

20

Fig. 4 (b). Total RNA from rat ovarian, ROS 17/2.8 cells, bone marrow RNA treated with estradiol and total RNA from primary osteoblasts in co-culture amplified for 25 cycles using an oligo derived from rER β (Accession #U57439). Each lane in Fig. 4 (b) contains the following types and amounts of RNA: lane 1, control, no RNA; lane 2, rat ovarian RNA (2 ng); lane 3, ROS 17/2.8 total RNA (0.1 μ g), lane 4, total (cultured) bone marrow RNA (0.1 μ g); lane 5, total cultured bone marrow RNA (0.1 μ g) where the cells had been treated with estradiol for 16 hours; lane 6, total RNA from primary osteoblasts in co-

25

30

culture (0.1 μ g); lanes 7-11, control reactions without reverse transcriptase (RT) for lanes 2-6, respectively.

Figure 5

Gel Shift Assay

Fig. 5(a). Gel shift analysis of mER β -3. The receptor-DNA complex was disrupted using the anti-peptide antibody 1067, which recognizes polypeptides encoded by exon 5B.

Fig. 5(b). Gel shift analysis of the human alpha form of the estrogen receptor (ER α). Disruption of the ER α -DNA complex was assayed using the two anti-peptide antibodies specific to exon 5B.

Both Fig. 5 (a) and (b) contain the following: lanes 1 and 2, extract alone; antibody 1067, lanes 3 and 4; antibody 1067 pre-immune serum, lanes 5 and 6; antibody 1068, lanes 7 and 8; antibody 1068 pre-immune serum, lanes 9 and 10; lanes 11 and 12 are control lanes that contain 16 μ g of untransfected COS-7 nuclear extract.

Figure 6

Comparison of mER β -3 protein with the murine ER α

The upper sequence is the protein sequence of mER β -3, whereas the lower sequence is that of the mouse (m) mER α . The “|” between the matched sequences indicates residue identity. The “:” between the matched sequences represents similar amino acids. The “.” observed in the sequences is a “gap” added by the sequence alignment program. The lines bisecting the paired sequences delineate the six domains (A-F) found in ER β_c and ER α . There is 99% similarity and 97% identity between the C domains, which contain the DBD, of the two murine estrogen receptor subtypes. There is 79% similarity and 59% identity between the E domains, which contains the LBD.

Figure 7

Comparison between mER β -3 and mER β_i nucleotide sequences

The upper paired sequence (which starts at nucleotide 151) is the nucleotide sequence of mER β -3, whereas the lower sequence is the nucleotide sequence of mER β_i published by Tremblay *et al.*, (1997). There is a one nucleotide difference between the

mERβ-3 sequence (an adenine at 1,244) and *mERβ_i* (a guanine at 1,009). This nucleotide difference results in an aspartic acid (D) residue in *mERβ-3* and a glycine (G) residue in *mERβ_i*.

5

Figure 8

Activity of ERβ Isoforms in the presence of various estrogens

Reporter constructs expressing ERβ-1 (B1), ERβ-3 (B3), ERα (alpha), or both ERβ-1 and ERβ-3 (B1+B3) were exposed to clomiphene, diethylstilbestrol (DES), 4 OH-tamoxifen (4-OHT), or 17 β estradiol (E2). Expression was standardized to ERα response to 100 nM drug.

10

Figure 9

Transactivation Profiles - cV2ERE

The four panels display the ability of the different estrogen receptors to transactivate cV2ERE. The cellular response of ERα (ER Alpha), murine ERβ-1 (mER-B1), murine ERβ-3 (mER-B3), or coexpression of both murine ERβ-1 and ERβ-3 isoforms (mER B1+B3) in COS-7 cells to E2, clomid, DES and 4-OHT were compared.

15

Figure 10

In situ Hybridization of Various Estrogen Target Tissues

20

Ovaries (upper panels), uteri (middle panels), and E-15 rat embryos were serially sectioned and probed using anti-sense (left panels) and sense (right panels) probes from ERβ and ERα. Cervical spine is shown in the lower panels.

25

Figure 11

3' RACE of Human Products Detected with mERβ-3

30

Figure 11 is a southern blot of human ERβ (3' RACE) using mouse probes. 3' RACE was performed using the Marathon cDNA kit and RNA derived from stromal cells derived of human osteoclastoma (lane 1), human ovary (lane 2), human prostate (lane 3) and human ovary (lane 4). The amplified products were detected using a random-primed probe

derived from mER β -3. Lanes 5-8 represent the corresponding amplifications for lanes 1-4 in the absence of reverse transcriptase (RT) during cDNA synthesis.

Figure 12

5

DNA and Corresponding Protein Sequence of Human Exon 5B

Fig. 12A. This figure aligns the rat, mouse (murine) and human ER β nucleic acid sequences of exon 5B. The last "TGA" on the second line of the human sequence is the stop codon (indicated by a "†"). The human polypeptide product of exon 5B is truncated as compared to the exon 5B nucleotide sequences for rat and mouse. In both Figure 12 A and B, a "-" indicates that the sequence is homologous with the other sequences, whereas a "*" indicates a non-homologous substitution only in Figure 12A.

10

15

Fig. 12B. This figure aligns the putative translation products of exon 5B. The "*" in the human amino acid sequence indicates the location of the termination codon, TGA (underlined). Both the amino acid and nucleotide sequences of the rat, murine and human sequences are displayed.

I. General Description

Estrogen receptors are members of the nuclear hormone receptor family.

20

Biologically, these proteins are intracellular receptors which mediate the effects of steroid hormones. Upon hormone binding, estrogen receptors control the transcriptional expression of certain hormone-responsive genes. This involves the binding of the receptors, often in homo- or heterodimeric form, to specific sequences, hormone response elements, located in the target gene promoter.

25

The compositions and methods of this invention provide for the screening of candidate compounds to be used to treat ER β_c related diseases. The compositions are based on the isolation of an ER β_c sequence, *ER β -3*, and the three alternatively spliced isoforms, *ER β -1*, *ER β -2* and *ER β -4*. Additionally, these compositions can be used to screen for ER β_c based disease to facilitate disease prognosis and to monitor disease-related aberrant expression of ER β_c or its isoforms.

30

II. Specific Embodiments

The specific embodiments disclosed in this invention relate to the isolation of the nucleic acid sequence that encodes the $ER\beta_c$ gene, $ER\beta-3$. The murine (m) form of $ER\beta-3$ is composed of 1,704 base pairs (bp) from the ATG start codon to TGA (Fig. 1a and b) and encodes a 567 amino acid protein; this sequence contains nine exons, including the newly described exon 5B, which is located in the region encoding the LBD. Also isolated were three alternatively spliced isoforms: $mER\beta-1$, $mER\beta-2$ and $rER\beta-4$. $mER\beta-1$ is 1,650 bp and encodes a 549 residue long polypeptide; $ER\beta-1$ lacks exon 5B (see Figs. 1a and 2a). $mER\beta-2$ is composed of 1,533 base pairs (bp); it lacks both exon 5B and exon 3, which contains 117 bp (see Figs. 1a and 2b). The sequence encoding $rER\beta-4$, an alternatively spliced isoform isolated from rat ovaries, is composed of 1,570 bp; it contains exon 5B, and the 54 bp it comprises, but exon 6, which contains 134 bp, has been deleted (see Figs. 1a and 2c).

Also described herein are methods of making and using the nucleic acid sequences corresponding to $ER\beta-3$, its isoforms and to the proteins encoded by these nucleic acid sequences. The methods of using the nucleic acid sequences of $ER\beta-3$ or its isoforms include determination of what tissues express $ER\beta_c$ and its isoforms (e.g., $ER\beta-1$, $ER\beta-2$ and $ER\beta-4$), function characterization for the proteins and nucleic acid sequences of $ER\beta-3$ and its isoforms, development of methods to recombinantly express $ER\beta_c$ nucleic acid molecules and their associated protein products, development of an $ER\beta-3$ reporter system, identification of $ER\beta-3$ ligands such as estrogen that influence $ER\beta-3$ or its isoforms and identification of compounds that modulate the influence exerted by $ER\beta-3$ or an isoform thereof on transcriptional regulation of other genes and determining the corresponding physiological effects of such influence.

A. Isolation of the complete ($ER\beta_c$) DNA and protein

Through such methods as reverse transcriptase (RT)-PCR and/or 5' RACE (rapid amplification of cDNA ends), the complete estrogen receptor β and three isoforms were isolated. Genomic primers were used for RT-PCR on mouse ovary RNAs to clone murine (m) $mER\beta-1$, $mER\beta-2$ and $mER\beta-3$. The sequences for $mER\beta-1$, $mER\beta-3$ and the rat (r)

isoform *rERβ-4* were obtained by 5' RACE using the Marathon system and a different set of primers. The primers and vectors chosen to isolate and clone these sequences would have been commonly known to an individual skilled in the art.

Using these techniques, a novel upstream sequence was found at the 5' end of exon 1. In mice and rats, this is an 192 bp sequence which is located at the 5' end of exon 1 of the *mERβ-3* gene, two alternatively spliced isoforms of *mERβ-3* and the alternatively spliced isoform isolated from rat ovaries (*rERβ-4*), as a result of an additional open reading frame (ORF) located upstream of the published clones. Kuiper *et al.*, (1996 and 1997); Mosselman *et al.*, (1996); Tremblay *et al.*, (1997). Further analysis by RT-PCR of mRNA derived from osteoblast and bone marrow co-cultures revealed a ninth exon, exon 5B, comprised of 54 bp and located within the LBD, as depicted in Figure 1a and b. As a result, the previously published human, rat and mouse sequences, all of which are referred to herein as $ER\beta_i$, are probably 5' truncated splice variants of this larger complete $ER\beta_c$ form, which in the murine system is *mERβ-3* (see Fig. 1a and b). The nucleic acid sequence information for *ERβ-3* predicts a 567 amino acid protein with a molecular weight of approximately 63 kDa, instead of 54 kD predicted for $ER\beta_i$.

To obtain the analogous sequences in other mammals, such as humans, the heretofore unknown *mERβ-3* gene or portions thereof can be used as probes. These probes should be of at least 18 nucleotides and preferably should be redundant for one or more sequences encoding the $ER\beta-3$ protein; the probes are to be designed from the $ER\beta_c$ amino acid sequence and should account for the degenerate genetic code. An appropriate cDNA library, such as that for ovary, testes or prostate cells, may then be screened with the probes for cDNAs which hybridize under standard conditions to one or more of the probe compositions. For examples of such general methods, see Sambrook *et al.*, MOLECULAR CLONING: A LABORATORY MANUAL (1989). The cDNAs may then be isolated and sequenced to determine whether they code for the $ER\beta_c$ protein. In this manner, the cDNA encoding the human $ER\beta_c$ protein or other mammalian *ERβ_c* genes and their respective species specific isoforms may be isolated.

A method of isolating other $ER\beta_c$ related genes is also described herein. Briefly, the nucleic acid sequences can be isolated by probing a DNA library such as that for prostate, ovary or testes, which is comprised of either genomic DNA or cDNA. Libraries may be

from either commercial sources or prepared from mammalian tissue by techniques known to those skilled in the art. The preferred cDNA libraries are human cDNA libraries which are available from commercial sources such as Stratagene.

The DNA libraries can be probed by plaque hybridization using oligonucleotide probes of at least 20 nucleic acid residues in length, which are complementary to unique sequences of murine or other *ERβ-3* genes. The preferred probes are the sequences for Primer 1 and Primer 2. The nucleic acid probes may be labeled to facilitate isolation of the hybridized clones. Labeling can be by any of the techniques known to those skilled in the art, but typically the probes are labeled with [³²P] using terminal deoxynucleotidyl-transferase as disclosed in Sambrook *et al.*, (1989).

Alternatively, those of skill may use polymerase chain reaction (PCR) technology to amplify nucleic acid sequences of the *ERβ-3* gene directly from mRNA, cDNA, genomic libraries or cDNA libraries. PCR or other *in vitro* amplification methods may also be useful, for example, to clone nucleic acid sequences that code for proteins to be expressed, to make nucleic acids to use as probes for detecting the presence of *ERβ-3* DNA or *ERβ-3* mRNA in tissue samples, for nucleic acid sequencing, or for other purposes.

Appropriate primers and probes for identifying *ERβ-3* from alternative mammalian tissues are generated from comparisons of the sequences provided herein. For a general overview of PCR techniques, see PCR PROTOCOLS: A GUIDE TO METHODS AND APPLICATIONS (Innis, M, Gelfand, D., Sninsky, J. and White, T., eds.), Academic Press, San Diego (1990), incorporated herein by reference.

The present invention further provides nucleic acid molecules that encode *ERβ-1*, and the related *ERβ-3* isoform proteins herein described, preferably in isolated form. As used herein, "nucleic acid" is defined as RNA or DNA that encodes a *ERβ-3* polypeptide, or is complementary to nucleic acid sequence encoding such peptides, or hybridizes to such nucleic acid and remains stably bound to it under appropriate stringency conditions, or encodes a polypeptide sharing at least 75% sequence identity, preferably at least 80%, and more preferably at least 85%, with the peptide sequences. Specifically contemplated are genomic DNA, cDNA, mRNA and antisense molecules, as well as nucleic acids based on alternative backbone or including alternative bases whether derived from natural sources or synthesized. Such a hybridizing or complementary nucleic acid, however, is defined further

as being novel and nonobvious over any prior art nucleic acid including that encodes, hybridizes under stringent conditions or other appropriate stringency conditions, or is complementary to a nucleic acid encoding an ER β -3 protein according to the present invention.

5 "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 titrate, 0.1% SDS at 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 (BSA), 0.1% Ficoll, 0.1% polyvinylpyrrolidone, 50 mM sodium phosphate buffer at pH 6.5 with 750 mM NaCl, 10 75 mM sodium citrate at 42°C. Another example is use of 50% formamide, 5 x SSC (0.75 M NaCl, 0.075 M sodium citrate), 50 mM sodium phosphate (pH 6.8), 0.1% sodium pyrophosphate, 5 x Denhardt's solution, sonicated salmon sperm DNA (50 μ g/ml), 0.1% SDS, and 10% dextran sulfate at 42°C, with washes at 42°C in 0.2 x SSC and 0.1% SDS. A skilled artisan can readily determine and vary the stringency conditions appropriately to 15 obtain a clear and detectable hybridization signal.

B. Characterization of the new sequences

The complete estrogen receptor β , such as *mER β -3*, contains nine exons. The three isoforms that have been isolated include *mER β -1*, *mER β -2*, *mER β -3* and the alternatively 20 spliced isoform from rat ovaries, *rER β -4*. *mER β -1* is 1,650 bp; it contains the previously identified eight exons, lacks the new exon 5B, but contains the previously undescribed 192 bp located at the 5' end of exon 1 (see Figs. 1a and 2a). In addition to having the novel 192 bp sequence located in the 5' terminus of exon 1 and the newly described exon 5B, there is a one nucleotide difference between the sequence published by Tremblay *et al.*, (1997) and 25 the sequence disclosed here: nucleotide 1,244 in exon 6 of the *mER β -3* sequence is an adenine whereas in the sequence by Tremblay *et al.*, (1997) it is a guanine (nucleotide 1,009).

mER β -2 contains 1,533 base pairs (bp); *ER β -2* lacks both exon 3 and exon 5B (see Figs. 1a and 2b). *rER β -4* includes 1,570 bp and has exon 5B, but exon 6 is deleted (see 30 Figs. 1a and 2c). The full length *mER β -3* contains a previously unidentified 192

nucleotides at its 5' terminus as well as the sequences of exon 5B and exon 6. All three isoforms, as well as *mERβ-3*, contain the novel 192 bp located at the 5' terminus of exon 1.

One embodiment of this invention includes using *ERβ-3* nucleic acid sequences containing the heretofore unknown 192 bp or 54 bp (exon 5B) domains or portions thereof and placing these sequences in appropriate vectors for purposes of replication. Such vectors can then be introduced into the appropriate cell expression systems to express the proteins for use either in an assay system or to help to characterize the function of particular portions of the *ERβ_c* gene or its corresponding protein.

Characterization of the *ERβ_c* protein can be performed by creating mutants, using antibodies that recognize specific domains on *ERβ_c* and using polypeptide sequences to specific regions of the protein to determine their function through competition assays. This invention proposes using such techniques to characterize the specific functions of the sequences or isoforms containing the novel 192 bp and/or exon 5B (54 bp) sequences.

Another method of characterizing *ERβ_c* and its isoform proteins includes the use of antibodies to map out specific functional domains on the *ERβ_c* protein, including the LBD, the dimerization site, and the DNA binding domain (DBD) of the *ERβ_c* protein. Antibodies could also be utilized to determine whether the *ERβ_c* or its isoforms is in a functional or non-functional conformation.

C. Creating antibodies to *ERβ_c* protein sequences

Antibodies are useful in several areas, including determining tissue expression of *ERβ_c*, such as *ERβ-3* or its isoforms (e.g., *ERβ-1*, *ERβ-2* or *ERβ-4*), and determining the functional domains of *ERβ-3* or its isoforms. Once the amino acid sequence of *ERβ-3* or its isoforms is known, another embodiment of this invention includes using polypeptides to create antibodies. Polypeptide sequences can be assessed using computer software to determine the antigenicity of certain polypeptide sequences for the purpose of creating antibodies to these *ERβ_c* specific polypeptides. Hopp and Woods, PNAS 78: 3824 (1981); Garnier *et al.*, J. Mol. Bio. 120: 97 (1978).

One antibody that has been created is an anti-peptide antibody that can distinguish between the *mERβ-3* and *ERβ_c*. Other antibodies can be created to distinguish between the *ERβ-3* isoforms, in addition to being able to distinguish between the active and inactive

states of ER β resulting from allosteric-induced ligand interactions with the receptor. The anti-peptide antibodies that distinguish between ER β -3 and ER β_i were prepared using conventional methods and were raised to the polypeptide sequence encoded by exon 5B with a cysteine group at the amino terminus: N - CSSEDPHWHVAQTKSAVPR - OH (Antibodies 1067 and 1068). This antibody contains all of the exon 5B polypeptide. The Jameson-Wolf antigenicity program determined that this polypeptide possesses a high degree of antigenicity. Garnier *et al.*, (1978). This program or the Hopp and Wood algorithm can also be employed to determine sequences of antigenicity in the novel amino terminus of ER β -3 and its isoforms to develop additional antibodies.

Two other antibodies were created that recognize both ER β -3 and ER β_i . These antibodies (Antibodies 1069 and 1070) were created against the following sequence: N - CSSTEDSKNKESQ - OH. This polypeptide sequence is located in the carboxy terminus of the published rat ER β_i . Kuiper *et al.*, (1996 and 1997). Antibodies 1067 and 1068 or 1069 and 1070 were obtained from the eggs of different chickens. Antibodies can also be created to polypeptides comprising the sequence in Figure 12B or fragments thereof.

An alternative method to create antibodies to ER β -3 polypeptide sequences involves isolating ER β -3 proteins and digesting them with various proteases. The cleavage fragments can then be purified by size and used to raise antibodies against specific portions of ER β -3. Finally, ER β -3 polypeptide sequences can be created recombinantly through fusion protein techniques. ER β -3 polypeptide sequences can be expressed by fusing the desired ER β -3 nucleotide sequence to, for example, the gene expressing glutathione S-transferase (GST). The expressed ER β -3 polypeptide sequences created as a fusion ER β -3/GST fusion product can then be used to create antibodies to the specific portion of ER β -3 encoded in the ER β -3 containing fusion gene construct. Antibodies raised to such recombinant proteins can be either monoclonal or polyclonal and such preparation techniques are generally known.

Polyclonal antibodies 1067, 1068, 1069 and 1070 were raised in chickens. Other animals could also be utilized. Pre-immune sera was purified from 2-3 eggs collected prior to hen immunization. Immunizations were prepared with 2 mg of antigen conjugated to 2 mg Imject Keyhole limpet hemocyanin (KLH) via maleimide to the extra cysteine residue located at the amino terminus of each peptide as recommended in the manufacturer's

(Pierce) instructions. The coupled carrier-antigen complex (0.5 ml) was emulsified with Complete Freund's adjuvant (0.5 ml) and 1.0 ml was used for the initial injection. The chickens were subsequently boosted every 2 weeks with coupled immunogen as described by Aves Laboratory, except that Incomplete Freund's Adjuvant was used. Six eggs were collected and the IgY was purified from the yolks. Other immunoglobulin isotypes and isotype subclasses can also be used (*e.g.*, IgG₁, IgG₂, IgM).

Monoclonal antibodies may be obtained by various techniques familiar to those skilled in the art. Briefly, spleen cells from an animal immunized with a desired antigen are immortalized, commonly by fusion with a myeloma cell (see, Kohler and Milstein, Eur. J. Immunol. 6: 511 (1976)). Alternative methods of immortalization include transformation with Epstein Barr Virus, oncogenes, or retroviruses or other methods well known to those of ordinary skill in the art. Colonies arising from single immortalized cells are screened for production of antibodies of the desired antigen specificity and affinity. The yield of the monoclonal antibodies produced by such cells may be enhanced by various techniques, including injection into the peritoneal cavity of a vertebrate host.

Alternatively, peptide specific antibodies, such as antibodies 1067 and 1068, are prepared by immunizing suitable mammalian hosts (*e.g.*, chickens or rabbits) under appropriate immunization protocols using the peptide haptens alone, if they are of sufficient length, 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, such as those supplied by Pierce Chemical Co., Rockford, IL, 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 (Cys) 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 a suitable adjuvant, as is generally understood in the art. During the immunization schedule, titers of antibodies are taken to determine adequacy of antibody formation. For more information, refer to Harlow and Lane, ANTIBODIES: A LABORATORY

MANUAL, Cold Spring Harbor Pubs., N.Y. (1988), which is incorporated herein by reference.

While the polyclonal antisera produced in this way may be satisfactory for some applications, for pharmaceutical compositions, the 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 with modifications which effect immortalization of lymphocytes or spleen cells, as is generally known. Kohler and Milstein, (1976). The immortalized cell lines secreting the desired antibodies are screened by immunoassay in which the antigen is the peptide hapten or is the ER β_c protein itself. When the appropriate immortalized cell culture secreting the desired antibody is identified, the cells can be cultured either *in vitro* or by production from ascites fluid.

The desired monoclonal antibodies are then recovered from the culture supernatant or from the ascites supernatant. Fragments of the monoclonal 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 the 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. The antibodies or fragments may also be produced, using current technology, by recombinant means. Regions that bind specifically to the desired regions of the receptor can also be produced in the context of chimeras with multiple species origin. Alternatively, the ER β_c specific antibody can be humanized antibodies or human antibodies, as described in U. S. Patent No. 5,585,089 by Queen *et al.* See also Riechmann *et al.*, Nature 332: 323 (1988). Finally, given that ER β_c protein may be involved in certain cancers, it would be useful to create bispecific antibodies capable of recognizing both the ER β_c protein and, for example, cytotoxic T cells to facilitate the killing of tumor cells which may be useful in treating cancer. Berg *et al.*, PNAS 88: 4723 (1991).

The antibodies thus produced are useful not only as modulators of ER β_c -estrogen interaction, but are also useful in immunoassays to detect ER β_c protein or its isoforms and for the purification of ER β_c protein or its protein isoforms. One can use immunoassays to detect the ER β_c protein or its alternatively spliced isoforms. Immunoassays can be used to qualitatively and quantitatively analyze the ER β_c protein. A general overview of the

applicable technology can be found in Harlow and Lane, (1988). In brief, ER β_c protein or a fragment or isoform thereof is expressed in transfected cells, preferably bacterial cells, and purified as generally described above and in the examples. The product is then injected into a mammal capable of producing antibodies. Either monoclonal or polyclonal
5 antibodies specific for the gene product can be used in various immunoassays; such assays include enzyme linked immunoabsorbant assays (ELISAs), competitive immunoassays, radioimmunoassays, Western blots (Fig. 3), indirect immunofluorescent assays, gel shift assays (Fig. 5) and the like.

10 D. Creating polypeptides that interfere with the binding domains of ER β_c

One embodiment of this invention utilizes ER β_c polypeptide sequences to assay their ability to interfere with ER β_c protein mediated transcription regulation. Such interference can be created by preventing ER β_c activating agents, such as estradiol, from binding to the ER β_c protein. Alternatively, a polypeptide could be designed to inhibit dimerization and
15 subsequent signaling from occurring. Such polypeptides could be created using peptide synthesizers or by creating fusion protein expressing gene constructs or other expression systems for either prokaryotic or eukaryotic cell systems.

In brief summary, the expression of natural or synthetic nucleic acids encoding mammalian ER β_c will typically be achieved by operably linking the gene or cDNA to a
20 promoter (which is either constitutive or inducible) and incorporating it into an expression vector. The vectors preferably are suitable for replication and integration in either prokaryotes or eukaryotes. Typical cloning vectors contain transcription and translation terminators, initiation sequences and promoters useful for regulation of the expression of the ER β_c gene. The vectors may also comprise generic expression cassettes containing at least
25 one independent terminator sequence, sequences permitting replication of the plasmid in both eukaryotes and prokaryotes, *i.e.*, shuttle vectors and selectable markers for both prokaryotic and eukaryotic systems.

Methods for the expression of cloned genes in bacteria are also well known. In general, to obtain high level expression of a cloned gene in a prokaryotic system, it is
30 essential to construct expression vectors which contain, at a minimum, a strong promoter to direct DNA replication. The inclusion of selectable markers in DNA vectors transformed in

E. coli is also useful. Examples of such markers include genes specifying resistance to ampicillin, tetracycline or chloramphenicol.

Suitable eukaryote hosts may include plant cells, insect cells, mammalian cells, yeast and filamentous fungi. In a preferred embodiment of this invention, the baculovirus/insect cell system is used for gene expression.

The protein encoded by the $ER\beta_c$ gene, which can be produced by recombinant DNA technology, may be purified by standard techniques well known to those of skill in the art. Recombinantly produced $ER\beta_c$ can be directly expressed or expressed as a fusion protein. The protein is then purified by a combination of cell lysis (e.g., sonication) and affinity chromatography. For fusion products, subsequent digestion of the fusion protein with an appropriate proteolytic enzyme releases the desired $ER\beta_c$, its isoforms or a fragment thereof.

The purified $ER\beta_c$, when described as "isolated" and or "substantially pure", describes a protein that has been separated from components which naturally accompany it. Typically, a monomeric protein is substantially pure when at least about 85% or more of a sample exhibits a single polypeptide backbone. Minor variants or chemical modifications may typically share the same polypeptide sequence. Depending on the purification procedure, purities of 85%, and preferably over 95% pure are possible. Protein purity or homogeneity may be indicated by a number of means well known in the art, such as polyacrylamide gel electrophoresis of a protein sample, followed by visualizing a single polypeptide band on a polyacrylamide gel upon staining. For certain purposes, high resolution will be needed and high performance liquid chromatography (HPLC) or a similar means for purification utilized.

The $ER\beta_c$ protein or its isoforms of this invention may be purified to substantial purity by standard techniques well known in the art, including selective precipitation with such substances as ammonium sulfate, column chromatography, immunopurification methods, and others. See, for instance, R. Scopes, PROTEIN PURIFICATION: PRINCIPLES AND PRACTICE, Springer-Verlag: New York (1982).

The $ER\beta_c$ polypeptides or isoform polypeptides could then be used in various assays, such as gel shift assays or yeast two hybrid systems wherein these polypeptide sequences can be tested to observe their binding ability to the hormone response elements

(HRE) on DNA sequences, dimerization binding ability, and agonist/antagonist binding ability.

E. Determining tissue localization of $ER\beta_c$ or its isoforms by nucleic acid hybridization

Using portions of the newly isolated $ER\beta_c$ gene, probes can be synthesized either using polymerase chain reaction (PCR) techniques or using *in vitro* transcription, of which both techniques are known to skilled artisans. These probes, which are typically radiolabeled, can be utilized to determine which tissues express a particular $ER\beta_c$ transcript either via Northern blot analysis or dot blots of RNA samples or by Southern blots wherein the mRNA has been reverse transcribed into DNA, which is then further amplified using polymerase chain reaction (PCR) as demonstrated in Fig. 4. Southern analysis of DNA is also useful in determining whether the $ER\beta_c$ gene is present or disrupted. For example, it is known that $ER\alpha$ is disrupted in certain breast tumors; such information may in turn be beneficial in determining the course of chemotherapy to be utilized on a patient. Using nucleic acid sequences unique to the $ER\beta_c$, it can be readily determined what tissues express the gene.

The present invention also provides methods for detecting the presence, absence and or abnormal expression of $ER\beta_c$ gene products in a physiological specimen, as well as in other tissue samples. One method for evaluating the presence or absence of $ER\beta_c$ in a sample involves a Southern transfer and is well known to those of skill in the art (Fig. 4). Briefly, the digested genomic DNA is run on agarose slab gels in buffer and transferred to membranes. Hybridization is carried out using the probes discussed above. Visualization of the hybridized portions allows the qualitative determination of the presence or absence of the $ER\beta_c$ gene or its isoforms. Southern blotting will also distinguish, depending on the stringency conditions used for hybridization, whether the $ER\beta_c$ gene is normal or contains gene deletions or rearrangements.

Similarly, a Northern transfer may be used for the detection of $ER\beta_c$ messenger RNA (mRNA) in tissue samples of mRNA. This procedure is also well known in the art. See Maniatis *et al.*, MOLECULAR CLONING: A LABORATORY MANUAL, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y. (1982). In brief, the mRNA is isolated from a given

cell sample using an acid guanidinium-phenol-chloroform extraction method. The mRNA is then electrophoresed to separate the mRNA species, and the mRNA is transferred from the gel to a nitrocellulose membrane. Labeled probes are used to identify the presence or absence of the $ER\beta_c$ transcript.

5 An alternative means for determining the level of expression of the $ER\beta_c$ gene is *in situ* hybridization. *In situ* hybridization assays are well known and are generally described in Angerer *et al.*, Methods Enzymol., 152: 649 (1987). This hybridization technique has already been used to study $ER\beta_i$ expression in rat hypothalamus. Shughrue *et al.*, (1997). In an *in situ* hybridization, cells are fixed to a solid support, typically a glass slide. If DNA
10 is to be probed, the cells are denatured with heat or alkali. The cells are then contacted with a hybridization solution at a moderate temperature to permit annealing of $ER\beta_c$ specific probes that are labeled. The probes are preferably labeled with radioisotopes or fluorescent reporters.

15 F. Recombinant nucleic acid molecules containing $ER\beta_c$ sequences

This invention relates to recombinant sequences that express the entire $ER\beta_c$ gene, its isoforms or portions thereof that include the newly described 5' terminus or the newly described 54 bp of exon 5B, as described in Figures 1 and 2. The invention includes all methods of expressing such recombinant constructs both in prokaryotic and eukaryotic
20 replication systems, which would have been known to one skilled in the art.

The methods for using recombinant deoxyribonucleic acid (rDNA) or recombinant ribonucleic acid (rRNA) sequences would include sequences that form triple helixes or sequences that are antisense to the $ER\beta_c$ mRNA or isoform mRNA. Additional methods would include expressions of these recombinant nucleic acid sequences to express the
25 encoded protein.

G. Host cells containing an exogenously supplied $ER\beta_c$ encoding nucleic acid molecule

The invention also relates to a method of introducing the recombinant full length
30 form of $ER\beta_c$, such as $ER\beta-3$ or one of the other isolated isoforms, $ER\beta-1$, $ER\beta-2$ or $ER\beta-4$ into non- $ER\beta-3$ expressing cells and assaying the effect said rDNA and its associated

protein product have on transcriptional regulation. Cells transfected with either the full-length ($ER\beta$ -3) or alternatively spliced isoforms of $ER\beta$ -3 (e.g., $ER\beta$ -1, $ER\beta$ -2 or $ER\beta$ -4) can then be utilized to assay the transfected cells' ability to form colonies in soft agar, different rates of DNA synthesis, differences in cell-cycle distribution in cells expressing different $ER\beta$ -3 isoforms and altered morphology of the transfected cells.

The present invention further provides host cells transformed or transfected with a nucleic acid molecule encoding an $ER\beta$ -3 protein. The host cell can be either prokaryotic or eukaryotic. Eukaryotic cells useful for expression of a $ER\beta$ -3 protein are not limited, so long as the cell line is compatible with cell culture methods and with the propagation of the expression vector and expression of the $ER\beta$ -3 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 fibroblastic cell line. Particularly preferred eukaryotic host cells include insect cells. Any prokaryotic host can be used to express an $ER\beta$ -3 encoding recombinant DNA (rDNA) molecule. The preferred prokaryotic host is *E. coli*.

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.*, PNAS 69: 2110 (1972); and Maniatis *et al.*, (1982); Sambrook *et al.*, (1989); or CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, F. Ausubel *et al.*, ed. Greene Publishing and Wiley-Interscience, New York (1987). 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.*, Virology 52: 456 (1973); Wigler *et al.*, PNAS 76: 1373 (1979); and Sambrook *et al.*, (1989).

Successfully transformed cells, *i.e.*, cells that contain a rDNA molecule of the present invention, can be identified by well known techniques. 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, J. Mol. Biol. 98: 503 (1975) or Berent *et al.*, Biotech. 3: 208 (1985) or the proteins produced from the cell assayed via an immunological method, as discussed above.

Utilization of the full-length, fragments, or isoforms of *ERβ-3* to determine their ability to regulate the formation of colonies in soft agar is useful in assessing whether a particular isoform of the *ERβ-3* gene is responsible for cellular proliferation and/or differentiation. The ability of a particular *ERβ-3* isoform to spur proliferation and/or differentiation may in turn correspond to the gene's involvement in certain ERβ associated diseases.

Alternatively, the isoforms may be used in transfected cell lines to assay [³H]-thymidine incorporation to test the effect of a particular *ERβ_c* isoform on DNA synthesis. Fluorescent activated cell sorting (FACS) could be utilized to determine differences in cell growth between cells bearing one isoform over another. Finally, transfected cells could be examined for morphological changes due to the expression of different *ERβ_c* isoforms. Once the properties of ERβ_c are determined with respect to impact on DNA expression, changes in morphology, and effects on cellular proliferation and/or differentiation, the same assays can be implemented to identify compounds that regulate the observed effects induced by isoforms of ERβ_c. Identification of putative drugs, which are discussed in greater detail herein, would be valuable in modulating concentrations of ERβ_c proteins or its isoforms in diseases involving such proteins.

H. Production of *ERβ_c* protein using recombinant methods

This invention also describes the methods used to express the *ERβ_c* protein, such as by using recombinant DNA (rDNA) of the *ERβ-3* gene, such as using its novel isoforms (*ERβ-1*, *ERβ-2* and *ERβ-4*) or portions thereof. Methods for generating rDNA molecules are well known in the art, for example, see Sambrook *et al.*, (1989). The preferred rDNA molecules would contain an *ERβ-3* encoding DNA or a DNA encoding one of its isoforms operably linked to expression control sequences and/or vector sequences.

The choice of vector and/or expression control sequences to which one of the ERβ_c nucleic acid molecules 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. Any vector contemplated by the present invention should be at least capable of directing replication, insertion into the eukaryote's chromosome or replicating extrachromasomally in a prokaryote, and preferably also expression of the ERβ-3 protein encoded in the rDNA molecule.

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.

In one embodiment, the vector containing a ER β -3 encoding 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 genes which confer such detectable markers as a drug resistance marker. Typical bacterial drug resistance genes are those that confer resistance to ampicillin or tetracycline. Vectors that include a prokaryotic replicon can further include a prokaryotic or viral promoter capable of directing the expression (transcription and translation) of the ER β -3 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 available from Biorad Laboratories, (Richmond, CA), pPL and pKK223 available from Pharmacia, Piscataway, N.J.

Expression vectors compatible with eukaryotic cells, preferably those compatible with vertebrate cells, can also be used to form rDNA molecules that contain ER β -3 sequences. Eukaryotic cell expression 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, #31255), the vector pCDM8 described herein, and like eukaryotic expression vectors.

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 for

which expression results in neomycin resistance, *i.e.*, the neomycin phosphotransferase (*neo*) gene as described by Southern *et al.*, J. Mol. Anal. Genet. 1: 327 (1982).

I. Diagnostic technique for measuring $ER\beta_c$ mRNA transcript levels

5 Another embodiment of the present invention is the use of $ER\beta_c$ nucleic acid sequences to measure changes in cells' mRNA concentrations. Methods of quantitatively and/or qualitatively assessing mRNA levels includes Northern blotting, *in situ* hybridization, nucleic acid hybridization and RT-PCR. Raval, J. Pharmacol. Toxicol. Methods 32(3): 125 (1994). One may use the coding sequence of $ER\beta_c$ or its isoforms, particularly the sequences of $ER\beta_o$,
10 not found in $ER\beta_i$, to determine the level of mRNA present in the cell. By lysing cells obtained by a biopsy, under conditions which inhibit RNases in accordance with conventional methodologies, the mRNA may be reverse transcribed into DNA and the DNA expanded using PCR. The expanded DNA may then be quantified. Less conveniently, Northern blot analysis may be used as described above. Reverse transcription PCR (RT-PCR) has been used to
15 ascertain specific mRNA concentrations in breast cancer cells and would be commonly known to individuals skilled in the art. See for example, Bartlett *et al.*, Br. J. Cancer 73(12): 1538 (1996). The benefits of using RT-PCR is that sample sizes do not have to be great to obtain valuable and sensitive results, as was observed in a study looking at mRNA levels of heart muscle biopsies. Engelhardt *et al.*, J. Am. Coll. Cardiol. 27(1): 146 (1996).

20 Determination of $ER\beta_c$ mRNA expression can also be assessed using *in situ* hybridization. This *in situ* labeling technique, which would employ labeled nucleic acid sequences capable of hybridizing to $ER\beta_c$ mRNA or its alternatively spliced isoforms and subsequent detection by a imaging device, would be useful in localizing tissues that have increased or decreased expression of $ER\beta_c$ or its isoforms' mRNA. This technique also would
25 be commonly known to individuals skilled in the art. For example, see Guldenaar *et al.*, Brain Res. 700 (1-2): 107 (1995).

By employing any of the above diagnostic techniques, the presence and amount of transcription and expression of $ER\beta-3$ or its isoforms may be determined, as a measure of the expression of $ER\beta-3$ protein, as well as other proteins for which transcription is regulated by the
30 $ER\beta-3$ protein. This information is related to the aggressive nature of a particular cancer, the change in the nature of the cancer in relation to treatments, such as irradiation, chemotherapy, or

surgery, the metastatic nature of the cancer, as well as the aggressiveness of metastases, and the like. For example, see Maas *et al.*, Cancer Lett. 97(1): 107 (1995), which discussed changes of specific mRNA levels in breast cancer cells using RT-PCR after treatment with different anti-cancer agents. This relationship may be useful for determining the level of therapeutic treatment, monitoring the response of the tumor (or other ER β_c related diseases) to the therapeutic treatment, and in providing a prognosis for the patient concerning the course of the disease.

J. Methods to identify agents that block ER β_c transcriptional regulation

Another embodiment of the present invention provides methods for identifying agents that inhibit or block the association of an estrogen or estrogen-like agonists/antagonists with ER β_c protein. For example, estrogen can be mixed with the ER β_c protein or a cellular extract containing the ER β_c , in the presence and absence of the compound to be tested. After mixing under conditions that allow association of the estrogen or estrogen-like agonist/antagonist with ER β_c , the two mixtures are analyzed and compared to determine if the compound augmented, reduced or completely blocked the association of the estrogen or estrogen-like agonist/antagonist with the ER β_c protein or its isoforms. Agents that block or reduce the association of an estrogen or estrogen-like agonist/antagonist with the ER β_c protein will be identified as decreasing the concentration of estrogen-ER β_c binding present in the sample containing the tested compound.

The receptor protein likely must undergo allosteric change in its conformation before the estrogen-ER β_c complex has the ability to bind to DNA. Once inside the nucleus, the activated receptor initiates transcription of genetic information from the DNA to mRNA, which is in turn a template for the linking of amino acids into proteins.

The antiestrogen effects produced by drugs such as tamoxifen (Nolvadex Registered TM) appear to be one of preventing the estrogen receptor from interacting with DNA in the nucleus to stimulate RNA and protein synthesis. This action initiates a block in the synthesis of macromolecules such as proteins, causing cell damage and the ultimate death of the cell. Antiestrogens are believed to be lipophilic molecules having a portion of the molecule which resembles naturally occurring estrogens. This portion of the antiestrogen selectively binds to the estrogen receptors. The antiestrogens, however, have a side chain arm (*e.g.*, dimethylaminophenyl ethoxy) which distorts the three-dimensional configuration of the estrogen

receptor preventing translocation of the receptor to the nucleus. Morgan, U.S. Patent No. 4,732,904 (1988). Another method of determining whether candidate reagents inhibit estrogen action on the complete estrogen receptor β subtype would be by determining whether ER β_c has undergone an allosteric transformation as a result of interacting with a candidate reagent such that ER β_c or its isoforms can no longer combine with the native substrate, estrogen. Changes in the conformation of ER β_c or homodimers of ER β_c can be detected using antibodies, either monoclonal or polyclonal, to conformational epitopes that exist on ER β_c or homodimers of the receptor. Antibodies were used to determine the functional state of ER α and a similar method could be used in determining whether compounds augment transformation into the activated allosteric conformation or inhibit the conformation all together. See Wotiz *et al.*, U.S. Patent No. 5,312,752 (1994).

Antibodies can not only be used to determine whether the ER β_c is functionally in an active or inactive state. Antibodies could also be screened to determine whether their binding to either the ligand or to the receptor itself enhanced the binding of the ligand to the receptor. Methods of determining said enhancement are known to the art. See Aguilar *et al.*, Mol. Cell. Biochem. 136(1): 35 (1994). Another method of determining whether a particular reagent augments or inhibits dimerization of ER β_c or augments or inhibits ER β_c from assuming the activated state would be to utilize a yeast two hybrid system. Yeast two hybrid systems have been successfully used to determine whether ER α dimerization is ligand-dependent (Wang *et al.*, J. Biol. Chem. 270(40): 23,322 (1992)); to isolate agents such as proteins or antibodies that enhance transcriptional activity of hormone receptors (Onate *et al.*, Science 270(5240): 1354 (1995)); to isolate compounds that are antagonistic to ER β_c action in a manner comparable to what has been done with ER α (Ichinose *et al.*, (1997) and Collins *et al.*, (1997)); and to determine whether ER β_c can form heterodimers in a manner analogous to what has been observed for retinoic acid receptors. See for further discussion Forman *et al.*, Cell 81(4): 541 (1995) and Walfish *et al.*, PNAS 94(8): 3697 (1997).

Another method to screen agents is to use a reporter gene such as β -galactosidase (β -gal) or luciferase. These transactivation experiments can be performed in yeast or in mammalian cell lines. The cells would contain ER β_c along with an appropriate estrogen responsive element (ERE) upstream of the reporter gene (*e.g.*, luciferase or β -gal), such as cV2ERE. Both antagonists and agonists of ER β can be assayed in this fashion. Gaido *et al.*, Toxicol. Appl.

Pharmacol. 143(1): 205 (1997); Hafner *et al.*, J. Steroid Biochem. Mol. Biol. 58(4): 385 (1996); Muhn *et al.*, Ann. N.Y. Acad. Sci. 761: 311 (1995). Such assays can also be utilized to determine whether cross-talk occurs for ER β_c and progesterone (PR) as has been demonstrated for ER α and PR. Kraus *et al.*, Mol. Cell. Biol. 15(4): 1847 (1995). Identifying novel ER β_c responsive elements can be done rapidly using libraries of degenerate oligonucleotides. The protocol requires no purified protein and specifically selects for functional response elements. Nawaz *et al.*, (1992).

Compounds that are assayed by the above methods 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 arbitrarily, without considering the specific sequences involved in the association of the estrogen or estrogen-like agonist/antagonist to the ER β_c protein. An example of such randomly selected agents is the use a chemical library, a peptide combinatorial library or a growth broth of an organism.

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. Agents can be rationally selected or rationally designed by utilizing the peptide sequences that recognize and bind to either the estrogen or estrogen-like agonist/antagonist or to the steroid hormone binding site on the ER β_c protein.

The agents of this embodiment can be, by way of example, peptides or other small molecules, antibodies (*e.g.*, monoclonal or polyclonal), fragments of antibodies (*e.g.*, Fv), or drugs with antiestrogenic or estrogenic activity (*e.g.*, narigenin, kaempferide, phloretin, biochanin A, flavone, ICI 182,780, raloxifene, tamoxifen, [6-hydroxy-3-[4-[2-(1-piperidinyl)ethoxy]phenoxy]-2-]4-hydroxybenzo[b]thiophene, raloxifene HCl, and ethynyl estradiol). Collins *et al.*, (1998); Palkowitz *et al.*, J. Chem. Med. 40(10): 1407 (1997). A skilled artisan can readily recognize that there is no limit as to the structural nature of the agents of the present invention. One class of compounds of the present invention includes polypeptide agents whose amino acid sequences are chosen based on the amino acid sequence of the ER β_c LBD.

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, rDNAs encoding these polypeptides may be synthesized using commercially available oligonucleotide synthesis

instrumentation and produced recombinantly using standard recombinant production systems. These rDNA molecules can then be utilized to recombinantly express polypeptides that bind to the ER β -3 protein or its isoforms. The production using solid phase peptide synthesis is necessitated if non-recombinantly produced polypeptide sequences are to be used.

5

K. Administration of agents that affect ER β _c signaling

The agents of the present invention can be provided alone, or in combination with additional agents that modulate a particular pathological process. For example, an agent of the present invention that reduces or otherwise modulates ER β _c transcriptional regulation, by
10 blocking estrogen or other agonist/antagonists from binding and transforming the ER β _c protein or its isoforms into an active state, can be administered in combination with other similar agents. As used herein, two agents are said to be administered in combination when the two agents are administered simultaneously or are administered independently in a fashion such that the agents will act at the same time.

15

The agents of the present invention can be administered via parenteral, subcutaneous, intravenous, intramuscular, intraperitoneal, transdermal, or buccal routes. 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. The present invention further
20 provides compositions containing one or more agents which block transcriptional regulation by the ER β -1 protein. While individual needs may vary, determination of optimal ranges of effective amounts of each component is within the skill of the art.

25

30

In addition to the pharmacologically active agents, the compositions of the present invention may contain pharmaceutically acceptable carriers comprising excipients and auxiliaries which 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 (*e.g.*, sesame oil) or synthetic fatty acid esters (*e.g.*, 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/or dextran. Antiestrogens are typically characterized as having limited solubility, therefore the use of agents such as dimethylformamide increases the solubility of such agonists/antagonists thus increasing their effect on, in this instance, $ER\beta_c$ or its isoforms. Sasson *et al.*, J. Steroid Biochem. 29(5): 491 (1988). Optionally, the suspension may also contain stabilizers. Liposomes can also be used to encapsulate the agent for delivery into the cell. For certain drugs such as tamoxifen, agents (*e.g.*, acetone and polyethylene glycol 4000) may be required to enhance the drug's solubility. Cotreau-Bibbo *et al.*, J. Pharm. Sci. 85(11): 1180 (1996).

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. 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.

In practicing the methods of this invention, the compounds of this invention may be used alone or in combination with other therapeutic or diagnostic agents. In certain preferred embodiments, the compounds of this invention may be co-administered along with other compounds typically prescribed for these conditions according to generally accepted medical practice.

L. Gene therapy expression of $ER\beta_c$

The $ER\beta_c$ gene, for example $ER\beta-3$ gene and the $ER\beta-3$ protein can also serve as a target for gene therapy in a variety of contexts. For example, in one application, $ER\beta-3$ deficient animals can be generated using standard knock-out procedures to inactivate a $ER\beta-3$ gene. In such a use, a non-human mammal (*e.g.*, a mouse or a rat) is generated in which the $ER\beta-3$ gene is inactivated or deleted. This can be accomplished using a variety of procedures known in the art, such as targeted recombination. Once generated, the $ER\beta-3$ deficient animal can be used to (1) identify biological and pathological processes mediated by the $ER\beta-3$ gene; (2) identify proteins and other genes that interact with $ER\beta-3$; (3) identify agents that can be exogenously supplied to overcome $ER\beta-3$ deficiency; and (4) serve as an appropriate screen for identifying mutations within $ER\beta-3$ gene that increase or decrease activity.

In addition to animal models, human $ER\beta_c$ deficiencies or mutations can be corrected by supplying to a patient a genetic construct encoding the necessary $ER\beta_c$ protein. A variety of techniques are presently available, and others are being developed, for introducing nucleic acid molecules into human subjects to correct genetic deficiencies and mutations. Such methods can be readily adapted to employ the $ER\beta_c$ encoding nucleic acid molecules of the present invention.

In another embodiment, genetic therapy can be used as a means for modulating an $ER\beta_c$ mediated biological or pathological process. For example, during osteoporosis, it may be desirable to introduce into the patient a genetic expression unit that encodes a modulator of $ER\beta_c$ mediated transcriptional regulation, such as a nucleic acid molecule that is antisense to the $ER\beta_c$ mRNA. Alternatively, tissue specific co-activators or co-repressors could be identified and introduced into a recipient to augment modulation of $ER\beta_c$ or its isoforms. Such a modulator can either be constitutively produced or inducible within a cell or specific target cell. This allows a continual or inducible supply of a therapeutic agent within the patient.

M. Conformational analysis of $ER\beta_c$ using antibodies

Using site-specific polyclonal and monoclonal antibodies against the DNA binding domain (DBD) of the $ER\beta_c$ protein, one can determine the active state of the protein. On this basis, the user is able to determine whether the DBD of $ER\beta_c$ is present in a functional or non-functional altered state and whether the $ER\beta_c$ protein has been activated or not. The invention, therefore, includes specifically prepared immunogens, polyclonal antisera and monoclonal antibodies which bind specifically to the DBD of $ER\beta_c$ or its isoforms, and immunoassays employing these site-specific antibodies with cellular samples on a functional and correlative test basis, as described above. Similar procedures and methods have been utilized in determining whether $ER\alpha$ is in its active or inactive state. Wotiz *et al.*

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 compounds of the present invention and practice the claimed methods; additionally, all the preceding description involving $ER\beta-3$ or, alternatively spliced isoforms of the complete gene, (*e.g.*, $ER\beta-1$, $ER\beta-2$ and $ER\beta-4$) can be applied to their analogs in other mammalian species. The following working examples therefore, specifically point out preferred embodiments of the

present invention, and are not to be construed as limiting in any way the remainder of the disclosure. Other generic configurations will be apparent to one skilled in the art.

EXAMPLES

5

Example 1

Cloning of the complete murine *mERβ-3* cDNA

The *mERβ-3* clone was twice isolated using two separate procedures: (1) reverse transcriptase PCR (RT-PCR) of mRNA, and (2) amplification from a mouse embryonic stem (ES) cell genomic DNA library. A mouse ES cell genomic DNA library was screened using a cDNA probe and RT-PCR. The oligonucleotides chosen, corresponded to regions in the D and E domains of rat *ERβ*, published by Kuiper *et al.*, (1996). These oligonucleotides were: 5'- ATG ACA TTC TAC AGT CCT GCT GTG ATG-3' (Primer 1) and 5'-GAA GTG AGC ATC CCT CTT TGC GTT TGG-3' (Primer 2). Using these oligonucleotides five clones were obtained. Two primers were chosen in these genomic DNAs, one around the first ATG, which is 192 bp upstream from the published ATG (Kuiper *et al.*, (1996); Mosselman *et al.*, (1996); and Tremblay *et al.*, (1997)), 5'-TCT CTG AGA GCA TCA TGT CC-3' (Primer 3), and one around the TGA, 5'-CAG CCT GGC CGT CAC TGT GA-3' (Primer 4). The RT-PCR was performed on 10 and 100 ng samples of mouse ovary RNA using the Titan™ RT-PCR System of Boehringer Mannheim according to manufacturer's instructions. The amplified products obtained using Primers 3 and 4 underwent a second amplification using: 5'-TGC TCT AGA CCA CCA TGT CCA TCT GTG CCT CT-3' (Primer 5) and 5'-CCG GAA TTC TCA CTG TGA CTG GAG GTT CTG 3' (Primer 6). The products obtained using Primers 6 and 7 were then inserted into Bluescript® vector. The same conditions were used to clone *mERβ-1*, *mERβ-2* and *mERβ-3*.

25

30

The *mERβ-3* clone was also isolated from mRNA using the Marathon RT-PCR system from Clontech. For 5' RACE, poly A+ RNA was prepared from total RNA derived from mouse ovaries according to the methods described in Sambrook *et al.*, (1989). Approximately 0.5 µg of the poly A+ RNA was reverse transcribed using 200 U Superscript II exogenase- (exo-) using the Marathon cDNA synthesis primer, 5'-TTC TAG AAT TCA GCG GCC GC(T₃₀)-3', according to manufacturer instructions (GIBCO). The second strand synthesis and all

subsequent steps, except PCR, were performed according to the conditions described by Marathon. The cDNA (0.5 μ l of a 10 μ l reaction) was then amplified using the Marathon adaptor primer, 5'-CCA TCC TAA TAC GAC TCA CTA TAG GC-3', with one of two gene specific reverse primers in the presence of Advantage *Taq* polymerase: 5'-GCA GTA GCT CCT TCA CCC G-3' (Primer 7) or 5'-GCA CTT CAT GCT GAG CAG-3' (Primer 8). The following four step thermocycling program was used to amplify the two products: (1) 5 cycles, 30 sec at 94°C, 4 min at 72°C; (2) 5 cycles, 30 sec at 94°C; (3) 25 cycles, 4 min at 70°C; and (4) 20 sec at 94°C, 4 min at 68°C. Single, predominant amplicons corresponding to the 5' end of the cDNA were then digested with restriction enzymes, cloned and sequenced. The clone was then inserted into a Bluescript® vector as described above.

Once the *mER β -3* gene was cloned, both the nucleic acid and the amino acid sequences were deduced for the complete estrogen receptor β sequence (see Figs. 1a and b). In addition to having the novel 192 bp sequence located in the 5' terminus of exon 1 and the newly described exon 5B, nucleotide 1,244, an adenine, in exon 6 of the *mER β -3* sequence differs from the guanine (nucleotide 1,009) found in the sequence by Tremblay *et al.*, (1997). The *mER β -3* gene is 1,704 nucleotides long and encodes a 567 amino acid protein. Using these techniques, it was ascertained that the 5' end of *mER β -3* included an additional open reading (192 bp) frame as well as a ninth exon, exon 5B (54 bp); both of these sequences are not found in ER β . Kuiper *et al.*, (1996); Mosselman *et al.*, (1996); and Tremblay *et al.*, (1997). Exon 5B is located in the ligand binding domain (LBD) of *mER β -3* and likely plays a significant role in *mER β -3* function.

Example 2

Isolation of three alternatively spliced isoforms

In addition to the complete murine estrogen receptor β gene sequence, *mER β -3*, two other alternatively spliced murine forms of *mER β -3* were identified (*mER β -1* and *mER β -2*), as well as a fourth alternatively spliced isoform isolated from rat ovaries, *rER β -4*. The first alternatively spliced form of *mER β -3*, *mER β -1*, contains the novel 192 bp at the 5' terminus of exon 1, but lacks the 54 bp of exon 5B; it is 1,650 nucleotides in length and putatively encodes a 549 amino acid long polypeptide (Fig. 2a). Preliminary data indicates that the *mER β -1* isoform may be more active than the full length *mER β -3*. The *mER β -1* isoform was isolated using both methods described for the isolation of *mER β -3*.

Isoform *mERβ-2* is composed of 1,533 bp, which would encode 510 amino acids (Fig. 2b); *mERβ-2* lacks exon 3, which contains 117 bp. The *mERβ-2* isoform was isolated only from the mouse ES cell genomic library.

Isoform *rERβ-4* was obtained from rat (r) ovaries whereas *mERβ-1* and *mERβ-2* as well as the full length *mERβ-3* were obtained from mouse (m) ovaries; it is 1,570 nucleotides in length and contains exon 5B, but exon 6 is deleted. Exon 6 is comprised (as shown in Fig. 1a) of 134 bp. The putative protein product of *rERβ-4* would be 414 amino acids (Fig. 2c).

All the nucleic acid sequences discussed relate to the coding regions and sequences for the corresponding mRNAs would be longer in both their 5' and 3' regions. It is likely that the published incomplete estrogen receptor β genes (*ERβ*) isolated from human, rat and mouse libraries are splice variants of this complete form, which in mice is *mERβ-3*, and contains the 54 bp of exon 5B and the 192 bp located at the 5' terminus of exon 1. Mosselman *et al.*, (1996); Kuiper *et al.*, (1996); and Tremblay *et al.*, (1997). All four sequences contain the 192 bp located at the 5' terminus of exon 1 and not described in the previously published sequences. *Id.* A sequence similar to the novel 192 bp region located in the 5' terminus of exon 1 may also exist in human *ERβ_c* and its isoforms.

The alternatively spliced isoforms (*e.g.*, *mERβ-1*, *mERβ-2* and *rERβ-4*) of the full length murine *ERβ_c* gene, *mERβ-3*, were twice isolated using the same two different procedures used to acquire *mERβ-3*. The primers used in both Examples 1 and 2 were selected based on the assumption that variants, if any, would occur within the boundaries of these selected primers. Historically, similar primers have produced analogous results with *ERα*.

Once the isoforms were isolated the DNA sequences could be sequenced and the amino acid sequence encoded by each could be determined. The proteins for the three alternatively spliced isoforms are shown in Figures 2a, 2b and 2c.

Example 3

Tissue specific expression of *mERβ-3* protein using Western Blotting

Anti-peptide antibodies raised against a sequence specific to the mouse *ERβ_c* (*mERβ-3*) specifically recognized a protein of 64 kDa in ovary and in bone, as well as in other tissues.

Two anti-peptide antibodies were raised in chickens to N-CSSEDPHWHVAQTKSAVPR-OH (Antibodies 1067 and 1068); this polypeptide is encoded by exon 5B and recognizes *mERβ-3* as

well as the isoforms that express the exon 5B coding region. Antibody 1067 and 1068 were obtained from the eggs of two different chickens, as were antibodies 1069 and 1070. These antibodies recognize the protein produced by *mERβ-3*, but not the *ERβ_i* protein discovered by Kuiper *et al.*, (1996 and 1997), which lacks exon 5B.

5 Total proteins (60 μg) obtained from ovarian tissue or bone tissue samples were resolved by electrophoresis in 10% SDS acrylamide gels; the gels were electrophoresed for 16 hours at 40 V. The proteins were transferred from the gels onto nitrocellulose membranes; the transfer was done for 4 hours at 100 mA. The blots were probed using a 1:1,000 dilution of the chicken antisera to *mERβ-3* (Antibody 1068) in conjunction with a 1:1,000 dilution of a secondary
10 antibody conjugated to horseradish peroxidase (Promega). The proteins were visualized using the ECL chemiluminescent substrate, and exposed to film (BMR film, Kodak) for one minute.

Figure 3 shows the results of the Western blot obtained using Antibody 1068, which detects the polypeptide encoded by exon 5B. Total protein (60 μg) was resolved by electrophoresis. The proteins were transferred to nitrocellulose membrane and probed with a
15 1:1,000 dilution of Antibody 1068 (Fig. 3a). Figure 3(b) is the blot probed with antibody 1068 pre-immune sera. The protein extracts of each lane of both Figures 3(a) and 3(b) are: lane 1, human ovary; lane 2, mouse ovary; lane 3, rat ovary; lane 4, ROS 17/2.8 cells; lane 5, ROS 17/2.8 cells treated with 100 nM estradiol for 16 hours; lane 6, murine primary osteoblasts. The antibody specifically recognizes a 64 kDa protein, which closely approximates the predicted size
20 of *mERβ-3*. The question mark refers to a protein migrating at approximately 58 kDa that may be immune specific but is otherwise unidentified. ROS 17/2.8 cells are a line characterized by Gideon Rodan; it is a rat osteoblast-like osteosarcoma cell line (ROS).

Example 4

Tissue specific expression of rat *ERβ* determined by Southern Blotting of RT-PCR Products

25 Cell expression of rat (*r*) *ERβ* (*rERβ*) mRNA was examined Southern blotting of RT-PCR products. Total RNA (2-100 ng) obtained from rat ovary, rat total bone marrow (100 ng), and ROS 17/2.8 cells (100 ng) were reverse transcribed using 200 U of Superscript (exo-) reverse transcriptase (Gibco-BRL) and 100 pmol random hexamer probe according to the
30 manufacturer's recommended conditions. ROS 17/2.8 cells are a rat osteoblast-like osteosarcoma cell line (ROS). The rat cDNA was then amplified by PCR in 100 μl reactions

using 2 U *Taq* polymerase and 1 μ M 5'-GTC AAG TGT GGA TCC AGG-3' (Primer 9; beginning at base 924 of Accession U57439 and corresponding to base 700 of *mER β -3*) and 5'-GCT CAC TAG CAC ATT GGG-3' (Primer 10; beginning at base 1,130 of *rER β* , by Kuiper *et al.*, Accession U57439, and corresponding to base 906 of *mER β -3*) per each individual reaction. Products were amplified using 25-40 cycles of the following amplification program: 90°C x 1 min; 55°C x 45 sec; 72°C x 2 min. The product was allowed to be extended at 72°C x 5 min at the end of the program.

Following amplification, the PCR products were resolved in a 4% NuSieve agarose (FMP)/TBE gel; the DNA was transferred to nylon membranes (Boehringer Mannheim) and cross-linked by UV irradiation for Southern analysis. Ten pmol of an oligonucleotide internal to the predicted amplicon 5'-AGC AGG TAC ACT GCC TGA GCA AAG CCA AGA-3' (Primer 11; beginning at base 991 of Accession U57439 and corresponding to bases beginning at 767 of *mER β -3*) was end-labeled using T4 polynucleotide and used to probe the immobilized DNA amplicon. Following pre-hybridization at 58°C in Quick-Hyb hybridization solution (Stratagene), the probe was added and allowed to hybridize for 1 hr. The blot was then washed twice with 2x SSC containing 0.1% SDS at room temperature, and then twice with 0.1 X SSC and 0.1% SDS at 58°C. The blot was then exposed to film.

Figure 4 is an autoradiograph of Southern blot of rat ER β (*rER β*) products amplified by RT-PCR. Total RNA from a variety of tissues was reverse transcribed, amplified by PCR, transferred to nylon membranes and probed using a ^{32} P labeled *mER β -3* oligonucleotide. Figure 4 (a) was amplified for 35 cycles. Each lane in Figure 4 (a) contains the following types and amounts of RNA: lane 1, control, no RNA; lane 2, rat ovarian RNA (0.1 μ g); lane 3, ROS 17/2.8 cells (0.1 μ g); lane 4, rat ovarian RNA control (0.1 μ g), no reverse transcriptase (RT); and lane 5, ROS 17/2.8 total RNA (0.1 μ g), no RT.

Figure 4 (b) is a Southern blot of total RNA. The ER β products were amplified for 25 cycles by RT-PCR. Each lane in Figure 4 (b) contains the following types and amounts of RNA: lane 1, control, no RNA; lane 2, rat ovarian RNA (2 ng); lane 3, ROS 17/2.8 total RNA (0.1 μ g), lane 4, total (cultured) bone marrow RNA (0.1 μ g); lane 5, total cultured bone marrow RNA (0.1 μ g) where the cells had been treated with estradiol for 16 hours; lane 6, total RNA from primary osteoblasts in co-culture (0.1 μ g); lanes 7-11, control reactions for lanes 2-6, respectively.

Discrimination analysis for the relative expression of ER β _c isoforms may be done utilizing random primers and reverse transcriptase (RT) to synthesize the cDNA from various rat or mouse or other mammalian tissues. The cDNAs so obtained are then amplified by PCR using the completely homologous rat and mouse primers 5'-GTC AAG TGT GGA TCC AGG-3' (Primer 9), which corresponds to base 700 of mER β -3 or base 924 of *Rattus norvegicus* estrogen receptor β mRNA, accession U57439 (Kuiper *et al.*, 1996), and 5'-GCA CTT CAT GCT GAG CAG-3' (Primer 8) corresponding to base 1,554 of mER β -3 and 1,724 of accession U57439. Following amplification, the PCR products are purified and digested with *Fsp I*, a restriction endonuclease with a consensus site within exon 5B (TGCGCA at base 1,176 of mER β and also present in rER β -4). Digestion of the mouse or rat amplicons bearing the exon 5B sequence thus yields smaller products. The digested PCR products are resolved by agarose gel electrophoresis, transferred to nylon membranes, and probed with complementary oligonucleotide probes specific to either rat or murine sequences, or both. The specific sizes of the hybridized DNA present determines what isoform is present in a particular tissue or cell sample. Additionally, the intensity of the band allows quantitation of the relative abundance of the isoform(s) in a particular sample.

Example 5

Gel Shift Assays

Gel shift analysis of mER β -3 is demonstrated in Figure 5(a). The results obtained by the mER β -3 gel shift (Fig. 5a) were compared to that obtained for the human estrogen receptor alpha (ER α) form, as displayed in Fig. 5(b). The receptor-DNA complexes formed were then disrupted using anti-peptide antibodies directed toward the novel exon 5B (Antibodies 1067 and 1068). Nuclear extracts (16 μ g) derived from COS-7 cells transfected with expression plasmids containing mER β -3 (Fig. 5a) or human alpha estrogen receptor (pHEGO) (Fig. 5b) were incubated with 5 fmols of 27 base pair perfect ERE end-labeled with ³²P isotope. The description for the lanes in both figures are the same. The lanes for both Fig. 5 (a) and (b) contain the following: lanes 1 and 2, extract alone; antibody 1067, lanes 3 and 4; antibody 1067 pre-immune serum, lanes 5 and 6; antibody 1068, lanes 7 and 8; antibody 1068 pre-immune serum, lanes 9 and 10; lanes 11 and 12 are control lanes that contain 16 μ g of untransfected COS-7 nuclear extract.

Example 6Relative Affinities of Various Estrogens for ER Subtypes

The experimental data displayed in Table I demonstrates the different affinities that various estrogens have for the ER β subtypes (e.g., ER α , ER β -1 and ER β -3).

E2 binding affinity was determined by incubating transfected COS-7 cell cytosol with different concentrations of [^{3} H]-E2 (0-200nM) and with or without unlabeled E2 500X for 4 h at 4°C in 40 mM Tris HCl pH 7.4, 150 mM KCl, PMSF 0.1 mM, DTT 2 mM. COS-7 cells were transfected as described in Example 7. Bound receptor was separated by the hydroxy apatite method (Obourn *et al.*, Biochemistry 32(24): 6229-6236 (1993)) or the ligand was removed by the dextran coated charcoal method (Garcia *et al.*, Mol. Endocrinol. 6(12): 2071-2078 (1992)), and bound hormone measured by liquid scintillation counting. Dissociation constants (kd) were obtained by Scatchard plots. Similarly, the relative affinities of different estrogenic ligands were determined using 1 nM receptor and 5 nM [^{3} H]-E2 with or without various concentrations of the described competitor steroids. The concentration of ligand necessary to displace 50% of the bound, labeled [^{3} H]-17 β -estradiol from the receptor, was used in the denominator to express the values shown in table (with a constant value of 1 nM in the numerator).

The differences in the relative affinities (results of the IC-50 experiments) show that the mER β -1 and mER β -3 receptors have different affinities for different ligands. This suggests that the transcriptional responsiveness to different ligands is a function of both the expression pattern of the receptor sub-type, and the estrogenic ligand used to stimulate transcription.

Table I shows the different affinities of estrogens to human ER α , mouse ER β -1 and mouse ER β -3 (which contains exon 5B). As indicated, the affinity of the different estrogens varies as to the receptor. The larger the number, the greater the affinity the estrogen has for the estrogen receptor target. Diethylstilbestrol (DES) has a greater affinity for the ER β isoforms than for ER α .

TABLE I. Relative Affinities of Various Estrogens for ER Subtypes

Estrogen	ER α Human	ER β -1 Mouse	ER β -3 Mouse
17 β Estradiol (standard)	100 Kd = 0.40 nM	100 Kd = 0.57 nM	100 Kd = 7.14 nM
Diethylstilbestrol	283	432	1009

Estrogen	ER α Human	ER β -1 Mouse	ER β -3 Mouse
Ethynyl Estradiol	216	34	189
RU 58668	208	111	475
Raloxifene	143	22	279
RU 39411	126	147	900
Moxestrol	87	16	132
4 OH-Tamoxifen	49	94	291
Estriol	5	14	38
17- α -estradiol	2.5	0.8	7
Tamoxifen	2	5	10
Estrone	1.3	2.3	21.5

The method used in this experiment can be utilized for screening reagents with different affinities for each of the ER β isoforms and comparing them to the ER α for determination of the affinity a particular drug may have for the other estrogen receptor proteins and their isoforms.

Example 7

Transactivation Profiles of ER β -1 and ER β -3 Isoforms

This experiment assessed the effect of ER β -1 and ER β -3 isoforms when expressed both individually and when expressed together as compared to the effect of ER α .

The ability of estrogens to stimulate transcription via an estrogen response element (ERE) functionally linked to tk-CAT (a construct described by Metzger *et al.*, *J. Biol. Chem.* 270(16): 9535 (1995)) was measured by transient transfection of the expression vectors for mER β -1 and mER β -3 in COS-7 cells. For transfection, COS-7 cells were seeded into six-well plates in phenol red free, low glucose DMEM. At approximately 50-80% confluency, the cells were transfected using lipofectamine according to the manufacturer's instructions (GIBCO-BRL). The expression constructs were transfected with a total of 2 μ g DNA containing 500 ng of reporter, 100-500 ng expression plasmid, and the remainder (1-1.4 μ g) as pBluescript as a carrier DNA. After 24 h, the cells were washed with DMEM and replaced with fresh medium containing drug (17- β estradiol, 4-hydroxy tamoxifen, clomiphene or DES at

1-300nM concentrations) or vehicle (ethanol). After 24 h the cells were lysed, and the CAT activity determined by liquid scintillation counting of converted chloramphenicol (as described in CURRENT PROTOCOLS IN MOLECULAR BIOLOGY). For co-transfection of both mER β -1 and mER β -3 isoforms, equivalent amounts of expression constructs were transfected (usually 100 ng each) (Fig. 8). To analyze the effect of other drugs on transcription, similar experiments were performed as described above, with the exception that drug concentrations were varied (Fig. 9).

The results of transfection analysis (Fig. 8) show that mER β is capable of stimulating transcription from a reporter containing a canonical responsive element. The mER β -1 can stimulate transcription to approximately 50-70% of that observed in similar cells transfected with the ER α construct, pHEGO, at estradiol concentrations of 100 nM. The mER β -3 isoform is capable of stimulating transcription to only approximately 40% of that observed in pHEGO at the 100 nM drug concentration.

As can be seen in Figures 8 and 9, the mER β -1 responsiveness is similar to that observed in cells transfected with pHEGO, which encodes ER α . mER β -3 only stimulated transcription at very high estrogen concentrations (100-200 nM). By contrast, when both receptors are co-expressed together, the magnitude of the response is ablated, showing that mER β -3 functions as a dominant negative modulator of the action of mER β -1. (Similar results were corroborated in an article published by K. Maruyama *et al.*, Biochemical Biophysical Research Communications 246(1):142-147(1998)).

The results in Figure 8 show that mER β -1 has a transactivation profile similar to ER α when exposed to E2, clomiphene (clomid), diethylstilbestrol (DES) and 4-OHT. The mER β -3 isoform has a decreased ability to transactivate cV2ERE as compared to either ER α or mER β -1. However, when coexpressed, the transactivation activity is reduced when the isoforms are co-expressed (Fig. 8, panel indicated as mER B1+B3). The assay utilized in this example can be similarly used to determine what agents can modulate homodimers of ER β isoforms, as well as heterodimers of the ER β isoforms or heterodimers composed of ER β and ER α isoforms.

Figure 9 demonstrates that ER β -1 (displayed as B1 in Fig. 9) and ER β -3 (B3) both possess similar activity when exposed to clomiphene, DES, 4-OHT, and E2. However when ER β -1 and ER β -3 are co-expressed in a reporter system, their activity is down regulated as compared to individual expression of the ER β isoforms or to ER α . This assay system can be

utilized to screen other estrogens or compounds that modulate the activity of the various ER β isoforms.

Example 8

5

In Situ Hybridization of various tissues

10

15

20

In situ hybridization analysis was performed using anti-sense cRNA probes to both the ER α and ER β to localize the message for each of the ER subtypes. The tissue was treated with 0.1 M TEA, pH 8.0, plus 0.25% acetic anhydride for 10 min at room temperature, rinsed three times in 2X SSC, dehydrated through a series of alcohols and air dried. cRNA riboprobes corresponding to the ER α or ER β -3 isoforms were prepared and used to probe tissue sections. The hybridization solution was removed, the sections washed and air dried. For riboprobes, an 801 base pair insert corresponding to the ligand binding domain of the mER β -1 plasmid (bases 931-1731 of the rat sequence) was linearized using the restriction enzyme *Apa*LI and transcribed using RNA polymerase *in vitro* in the presence of [³⁵S]-UTP and [³⁵S]-CTP according to methods of Goldstein *et al.*, Neuroscience 71(1): 243 (1996). The riboprobes were purified by ethanol precipitation, and the dried tissue sections hybridized with probe in hybridization buffer overnight at 55°C. The hybridization solution was removed, the sections were incubated briefly with RNase, then washed, dehydrated, and air dried. The dried sections were exposed to film for normalization of subsequent exposure times and dipped in NTB3 emulsion to determine the cellular and anatomical localization of each mRNA.

25

The results demonstrate abundant and wide spread distribution of the ER β message within the developing ovarian follicle (Fig. 10, top panel) and in the lung, kidney cortex, and specific regions of the brain (not shown). The pattern of distribution of ER α was quite different and was highly expressed only in the uterus (Fig. 10, middle panel), in the medullary regions of the kidney (not shown) and specific regions of the brain. Preliminary data also indicated that ER β is expressed in ossification center that appear to correspond with mesenchymal condensation zones in developing rat bone (12 days), especially in the spine (Fig. 10).

30

The ER β message is observed in developing Graffian follicles (GA), but not in resorbing follicles (FA) undergoing atresia (Fig. 10, top panel, antisense). In the ovary, the ER α message receptor was only abundant within the uterine tube (not shown). ER α mRNA was observed to be widely expressed throughout the uterus (Fig. 10, middle panel, antisense). In the cervical

spine, the ER β mRNA was localized to zones of mesenchymal ossification (M) similar to the expression pattern of Osf/Cbfa1, an osteoblast differentiation protein (Fig. 10, bottom panel, antisense, arrows). Controls corresponding to serial sections hybridized using sense riboprobe controls are also shown in the panels on the right.

5

Example 9

Methods of Screening for Drugs

A. Phosphorylation of ER β . Most of the members of the steroid receptor superfamily, including ER α , undergo post-translational modifications (*e.g.*, phosphorylation) as a function of their basal state or in response to ligand binding. With ER α , there are a variety of sites on the molecule that are phosphorylated in response to ligand binding. Post-translational modification of mER β or human ER β can be accomplished using the same methods as previously utilized for ER α . Methods of analyzing phosphorylation include transient or stable expression of the various cDNA constructs in COS-7 cells, or by immunoprecipitation of [^{32} P]-labeled ER β from cells metabolically labeled with [^{32} P]-orthophosphate. Tryptic maps from ligand stimulated or unstimulated cells can be obtained using ER β proteins isolated by immunoprecipitation of the mER β or human ER β molecule using our antibodies (*e.g.*, directed towards products of exon 5B such as the antibody used to obtain Fig. 3) or commercially available antibodies. For studies performed using *in vitro* transient transfection, a triple-myc tag or GST tag can also be linked to the carboxyl or amino termini by cloning the appropriate coding sequence into the expression plasmid. The expressed (phosphorylated) protein can then be immunoprecipitated using a very reliable, and commercially available anti-myc antibody (if using the triple-myc tag) or anti-GST antibodies. In addition, exon 5B amino acid residues can be substituted with other residues to prevent phosphorylation. In contrast to either ER α or mER β -1, exon 5B, which is unique to mER β -3, is located within a region of the molecule that otherwise is extremely hydrophobic. The exon 5B region, however, is unusually hydrophilic and contains a consensus casein kinase II (CKII) phosphorylation site (VLDRSSEDP) that arises as direct consequence of the location of the exon 5-exon 5B-splice junction. Many of the steroid receptors, including the ER α subtype, are phosphorylated on CKII sites. The serines present in the portion of ER β encoded by exon 5B can be substituted with alanine residues (or other uncharged amino acids) or with residues which mimic constitutively phosphorylated molecules (*e.g.*, aspartic acid residues). Such forms

of ER β can be utilized in screening and isolating drugs which modulate the activity of the various ER β isoforms. Alternatively, these mutant forms of ER β or polypeptide fragments containing this region can themselves be tested for agonist or antagonist activity in the ER β signal pathways.

5 B. Domain Switching. The amino terminus of the ER α contains an autonomous transcriptional activity (AF-1) that is only fully active when "integrated" with the ligand-dependent transcriptional domain (AF-2) present within the ligand binding domain of the ER α molecule. While these domains have yet to be described for the ER β molecule, the high degree of sequence homology at the protein level between ER α and ER β molecules logically suggests that ER β is similarly organized. Many of these domains have been identified and characterized using portions of the ligand binding domain (LBD) fused to convenient and reliable epitope tags, such as GST and myc. Such constructs can then be utilized to identify, in whole cell lysates or other expression models such as expression libraries, proteins that functionally alter the transcriptional responsiveness of the ER complex. We postulate that specific integrator molecules may be found using the LBD of mER β -3 fused to such convenient epitope tags as probes for protein-protein interactions. Such proteins can then alter the transcriptional responsiveness of the functional ER complex, (defined as the homo-dimers of ER β 3 with ER β 3 or hetero-dimers of ER β 3 with ER β 1, or ER β 3 with ER α) portion of the amino terminus fused with such epitope tags as probes for proteins that interact with the ER-complex. These complexes in turn can be used in drug screening assays to identify drugs which modulate ER β isoform activity. Alternatively the complexes themselves may be used to regulate pathways mediated by estrogen receptors.

Example 10

3' RACE of Human Products Detected with mER β -3

25 cDNA was prepared using the Marathon kit, (Clontech) as per the manufacturer's instructions and as discussed in the examples above, and 5 μ g total RNA derived from a human osteoclastoma (Fig. 11, lane 1), human ovary (lanes 2 and 4) or human prostate (lane 3). The cDNA products were resolved by electrophoresis in an agarose/TAE gel and transferred to membranes. The gene specific primer (5'-GTC AAG TGT GGA TCC AGG-3'), corresponding to bases 502-516 of GenBank Accession No. X99101 (Mosselman *et al.*, 1996), was used in

conjunction with the AP-1 adaptor primer for amplification of the 3' end of the human sequence. The amplification conditions were 92°C x 40 sec, 60°C x 40 sec, 75 x 1:30 min for a total of 40 cycles. For detection of the amplicon, a random primed probe was made using mER β -3 as template, and the blot hybridized in Church's buffer. Following hybridization, the blot was washed twice, 15 min each in 2X SSC and 1.0% SDS at 21°C, and two times for 15 min in 0.2X SSC and 1.0% SDS at 50°C. Background and overall signal can be modulated by the level of stringency utilized. Autoradiography was performed on the washed blot.

Figure 11 shows hybridization of the mER β -3 probe to sequences amplified from human ovary (lanes 2 and 4) and human prostate (lane 3) but not from a human osteoclastoma cell line.

Example 11

Human Sequence of ER β Exon 5B

Utilizing the rat and mouse exon 5B domains, human intron 5 was sequenced and examined for the presence of an exon 5B-like domain. Such a region was identified at the 5' end of intron 5, and the sequence is shown in Fig. 12A. The putative translation product is depicted in Figure 12B. The human exon 5B equivalent to the exon 5B observed in rat and mouse ER β_c also can be obtained using the hybridization methods described above, followed by sequencing.

A comparison of the human exon 5B ER β sequence and the sequences surrounding the human exon 5B sequence shows that there is high homology to the corresponding exon 5B and surrounding sequences of the rat and mouse (Fig. 12A). The sequences ending the introns (intron 5) of the three sequences are highly homologous (indicated by "*" in Fig. 12B), as are the regions beginning SSED in Figure 12B. The corresponding 3' end of the human exon 5B is also similar to the mouse and rat 3' ends. Previously published sequences of human ER β do not contain this exon 5B, which may be expressed only in certain human populations. We note that there is an one nucleotide difference in the 3' splice acceptor site of the human exon 5B as compared to that observed in the mouse and rat populations. Thus, the human gene containing the intron 5 sequence in which the exon 5B sequence was found may not produce an mRNA containing the exon 5B sequence.

We also note that the human ER β_c does not contain the additional 5' domain found in the murine sequence described above. Therefore, as used herein specifically, the term "human

ER β_c " contains the putative exon 5B as shown in Figure 12A or allelic variants or conservative substitutions thereof.

Referring to Figures 12A and 12B, while the translation products of both the mouse and rat nucleic acid sequences continue through the ligand binding domain (LBD), the translation product of human exon 5B is truncated due to the presence of an in-frame stop codon (TGA) in the exon. It is interesting to note that the SSED domain, which is 100% conserved between human, rat and mouse, contains two serines. These two serines, as discussed in Example 9, could be phosphorylated thereby regulating the activity of ER β isoforms containing the polypeptide encoded by exon 5B. The roles of these serines are discussed in Example 9, which sets forth phosphorylation assays. These assays may be used to studying ER β isoforms comprising the product encoded by exon 5B (e.g., the human equivalent to mER β -3) and their regulation. As described above, these human exon 5B and human ER β_c sequences may be used in screening assays to select compounds that modulate their activity and may be useful in regulating other ER β isoforms.

REFERENCES

The following references and all articles, texts and patents referred to above and below, are hereby incorporated by reference in their entirety:

Aguilar *et al.*, Mol. Cell. Biochem. 136(1): 35 (1994).

Angerer *et al.*, Methods Enzymol., 152: 649 (1987).

Aprikian *et al.*, Cancer 71(12): 3952 (1993).

Bartlett *et al.*, Br. J. Cancer 73(12): 1538 (1996).

Berent *et al.*, Biotech. 3: 208 (1985).

Berg *et al.*, PNAS 88: 4723 (1991).

Berkowitz and Evans, J. Biol. Chem. 267(10): 7134 (1992).

Berry *et al.*, EMBO 9: 2811 (1990).

Bonetti *et al.*, Breast Cancer Res. Treat. 38(3): 289 (1996).

Burch *et al.*, Mol. Cell. Biol. 8(3): 1123 (1988).

Byers *et al.*, Mol. Endocrinol. 11(2): 172 (1997).

Cohen *et al.*, PNAS 69: 2110 (1972).

Collins *et al.*, Steroids 62: 365 (1997).

Cotreau-Bibbo *et al.*, J. Pharm. Sci. 85(11): 1180 (1996).

CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, F. Ausubel *et al.*, ed. Greene Publishing and Wiley-Interscience, New York (1987, 1995).

Delmas *et al.*, J. Clin. Oncol. 15: 955 (1997).

Engelhardt *et al.*, J. Am. Coll. Cardiol. 27(1): 146 (1996).

5 Evans *et al.*, Bone 17(4S): 1815 (1995).

Forman *et al.*, Cell 81(4): 541 (1995).

Gaido *et al.*, Toxicol. Appl. Pharmacol. 143(1): 205 (1997).

Gallo *et al.*, Semin. Oncol. 24: S1 (1997).

Garcia *et al.*, Mol. Endocrinol. 6(12): 2071 (1992).

10 Garnier *et al.*, J. Mol. Bio. 120: 97 (1978).

Goethuys *et al.*, Am. J. Clin. Oncol. 20 (1): 40 (1997).

Goldstein *et al.*, Neuroscience 71(1): 243 (1996).

Graham *et al.*, Virology 52: 456 (1973).

Grasser *et al.*, J. Cell Biochem. 65: 159 (1997).

15 Green *et al.*, Cold Spring Harbor Symposia on Quantitative Biology 51(2): 751 (1986).

Guldenaar *et al.*, Brain Res. 700(1-2): 107 (1995).

Habenich *et al.*, J. Steroid Biochem. Mol. Biol. 44: 557 (1993).

Hafner *et al.*, J. Steroid Biochem. Mol. Biol. 58(4): 385 (1996).

Harlow and Lane, ANTIBODIES: A LAB MANUAL, Cold Spring Harbor Pubs., N.Y. (1988).

20 Hopp and Woods, PNAS 78: 3824 (1981).

Hughes *et al.*, Nat. Med. 2(10): 1132 (1996).

Ichinose *et al.*, Gene 188: 95 (1997).

Jackson *et al.*, Mol. Endocrinol. 11: 693 (1997).

Kangas, Acta Oncol. 31(2): 143 (1992).

25 Kohler and Milstein, Eur. J. Immunol. 6: 511 (1976).

Komm *et al.*, Science 241: 81 (1988).

Kraus *et al.*, Mol. Cell. Biol. 15(4): 1847 (1995).

Kuiper *et al.*, Endocrinology 138(3): 863 (1997).

Kuiper and Gustafsson, FEBS 410: 87 (1997).

30 Kuiper *et al.*, PNAS 93: 5925-30 (1996).

Maas *et al.*, Cancer Lett. 97(1): 107 (1995).

- Maniatis *et al.*, MOLECULAR CLONING, A LABORATORY MANUAL, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1982).
- K. Maruyama *et al.*, Biochemical Biophysical Research Communications 246(1):142 (1998).
- Metzger *et al.*, J. Biol. Chem. 270(16): 9535 (1995).
- 5 Morgan, U.S. Patent No. 4,732,904 (1988).
- Mosselman *et al.*, FEBS Letters 392: 49 (1996).
- Muhn *et al.*, Ann. N.Y. Acad. Sci. 761: 311 (1995).
- Nawaz *et al.*, Gene Expr. 2(1): 39 (1992).
- Obourn *et al.*, Biochemistry 32(24): 6229 (1993).
- 10 Onate *et al.*, Science 270(5240): 1354 (1995).
- Palkowitz *et al.*, J. Chem. Med. 40(10): 1407 (1997).
- PCR PROTOCOLS: A GUIDE TO METHODS AND APPLICATIONS, Innis, M, Gelfand, D., Sninsky, J. and White, T., eds., (Academic Press, San Diego 1990).
- Queen *et al.*, U. S. Patent No. 5,585,089.
- 15 Raval, J. Pharmacol. Toxicol. Methods 32(2): 125 (1994).
- Riechmann *et al.*, Nature 332: 323 (1988).
- Safarians *et al.*, Cancer Res. 56(15): 3560 (1996).
- Sambrook *et al.*, MOLECULAR CLONING: A LABORATORY MANUAL (2nd ed.), Vols. 1-3, Cold Spring Harbor Laboratory, (1989).
- 20 Sasson *et al.*, J. Steroid Biochem. 29(5): 491 (1988).
- Scopes, PROTEIN PURIFICATION: PRINCIPLES AND PRACTICE, Springer-Verlag: New York (1982).
- Somjen *et al.*, J. Steroid Biochem. Mol. Biol. 59: 389 (1996).
- Southern, J. Mol. Biol. 98: 503 (1975).
- Southern *et al.*, J. Mol. Anal. Genet. 1: 327 (1982).
- 25 Shughrue *et al.*, Steroids 61(12): 678 (1996).
- Tora *et al.*, Cell 59: 447 (1989).
- Tremblay *et al.*, Mol. Endocrin. 11(3): 353 (1997).
- Wang *et al.*, J. Biol. Chem. 270(40): 23,322 (1992).
- Walfish *et al.*, PNAS 94(8): 3697 (1997).
- 30 Wigler *et al.*, PNAS 76: 1373 (1979).
- Wotiz *et al.*, U.S. Patent No. 5,312,752 (1994).

WHAT IS CLAIMED:

1. An isolated nucleic acid molecule that encodes human *ERβ_c* and its allelic variants or isoforms.
2. The isolated nucleic acid molecule of claim 1, wherein said nucleic acid molecule is operably linked to one or more expression control elements.
3. An isolated nucleic acid molecule of claim 1 or a portion thereof, wherein said nucleic acid molecule hybridizes to a nucleic acid molecule of claim 1 under conditions of sufficient stringency to produce a clear signal.
4. A host transformed to contain a nucleic acid molecule of claim 3.
5. The host of claim 4, wherein said host is selected from the group consisting of prokaryotic hosts and eukaryotic hosts.
6. A method of producing human *ERβ_c* allelic variants or fragments thereof, comprising the step of culturing a host of claim 5 under conditions in which the *ERβ_c* protein or polypeptides thereof are expressed.
7. An isolated protein comprising the amino acid sequences depicted in Figure 12B.
8. An isolated antibody that binds to the protein of claim 7.
9. The antibody of claim 8, wherein said antibody is selected from the group consisting of a monoclonal antibody, a humanized antibody, a human antibody, a bispecific antibody, a chimeric antibody, and an antibody fragment such as Fv, Fab, Fab', or F(ab')₂.

10. An antibody of claim 9, wherein said antibody that recognizes and binds to N-CSSDPVLHR-OH or polypeptide sequences that contain these sequences or portions of these sequences.

11. A method for blocking the interaction of an estrogen selected from the group consisting of physiological estrogens, stilbene estrogens or triphenylethylene (anti)estrogens with the protein of claim 7 comprising the step of contacting said protein with an agent that blocks the binding of the protein to physiological estrogens, stilbene estrogens or triphenylethylene (anti)estrogens.

12. The method of claim 11, wherein said agent blocks the binding of physiological estrogens, stilbene estrogens or triphenylethylene (anti)estrogens to said protein by selectively binding to a hormone binding domain, to a DNA binding domain, or to a dimerization domain of the protein of claim 7.

13. The method of claim 12, wherein said agent is selected from the group consisting of a fragment of said protein, an antibody that binds to said protein, and an antibody fragment, such as Fv, Fab, Fab', or F(ab')₂ that binds to said protein, and a humanized antibody, human antibody, bispecific antibody or a chimeric antibody that bind to said protein.

14. The method of claim 13, wherein said binding reduces/enhances cellular proliferation, differentiation, DNA synthesis or cell cycle diversity.

15. A method for identifying agents that modulate the interaction or the effect of physiological estrogens, stilbene estrogens or triphenylethylene (anti)estrogens with or of the protein of claim 7 comprising the steps of:

- a) incubating said protein with a putative inhibitory compound;
- b) adding a physiological estrogen, stilbene estrogen or triphenylethylene (anti)estrogen; and

c) determining whether said agent modulates the binding of the protein of claim 9 to said physiological estrogens, stilbene estrogens or triphenylethylene (anti)estrogens.

16. A method to assay for the activation of a protein of claim 7 comprising the step of determining whether said protein is expressed or the step of determining whether the mRNA encoding said protein of claim 7 is expressed.

17. A method to localize a protein of claim 7 *in situ* which method comprises administering to a subject an amount of antibody that binds to said protein.

18. The method of claim 17, wherein the antibody is conjugated to a radioactive isotope.

19. An isolated nucleic acid molecule which is complementary to a nucleic acid molecule of claim 1.

20. A method to inhibit expression of the protein of claim 7 in a cell, which comprises providing to said cell the nucleic acid molecule of claim 19 in an amount sufficient to inhibit expression of said protein.

21. A pharmaceutical composition comprising a nucleic acid molecule of claim 19 together with a pharmaceutically acceptable carrier or excipient.

22. A nucleic acid molecule encoding a human ER β_c isoform comprising the nucleotide sequence of human exon 5B depicted in Figure 12A.

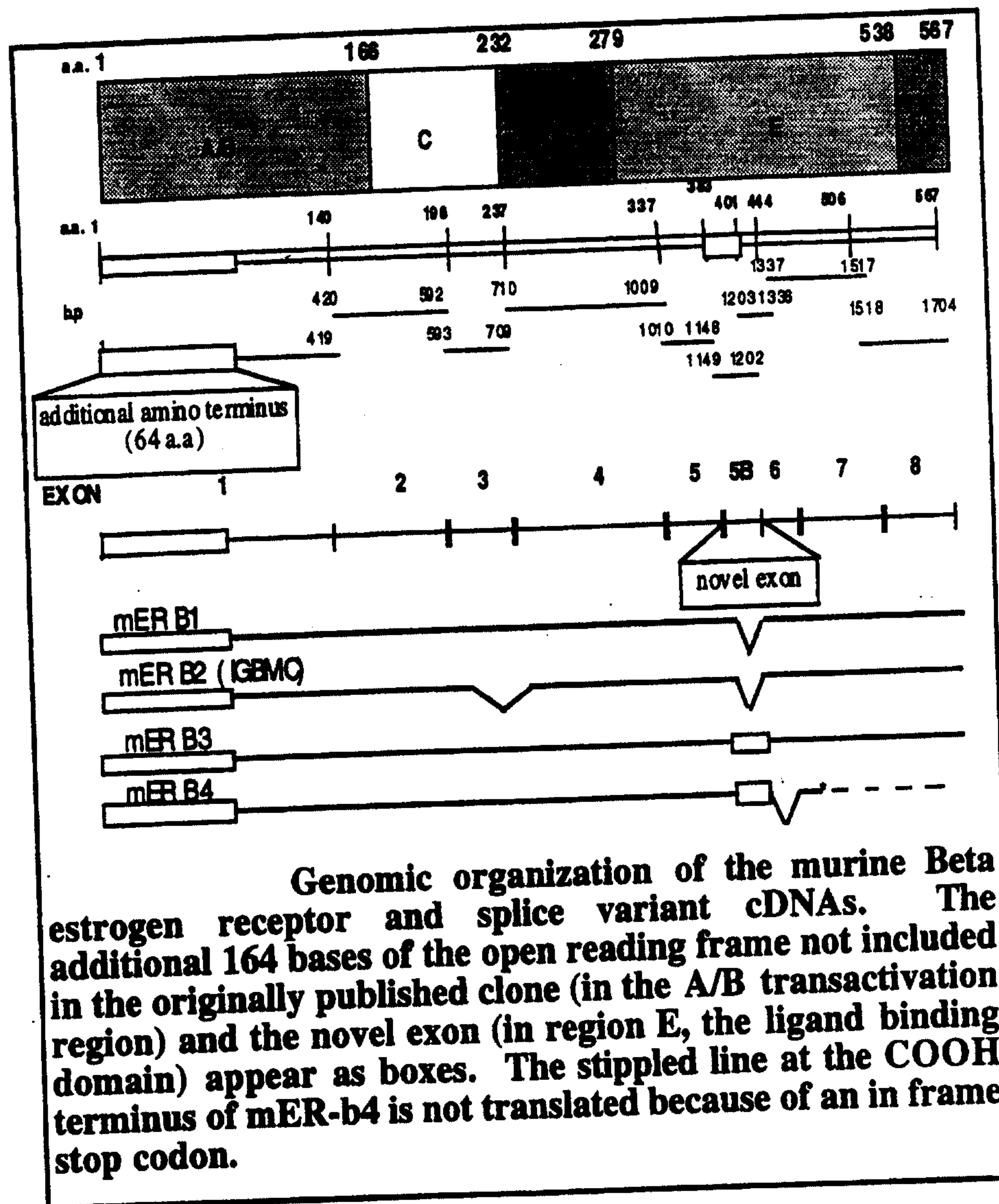


FIGURE 1A

WO 99/05170

2/20

PCT/US98/15539

1 ATGTCCATCTGTGCCTCTTCTCACAAGGATTTTTCTCAGCTGAGACCTACGCAAGACATG 60
 -----+-----+-----+-----+-----+-----+-----+-----
 TACAGGTAGACACGGAGAAGAGTGTTCTAAAAGAGTCGACTCTGGATGCGTTCTGTAC
M S I C A S S H K D F S O L R P T O D M -
 GAGATCAAAACTCACCATCAAGCCTTACTTCCCCTGCTTCTTATAACTGTAGCCAGTCC 120
 61 -----+-----+-----+-----+-----+-----+-----+-----
 CTCTAGTTTTTTGAGTGGTAGTTCGGAATGAAGGGGACGAAGAATATTGACATCGGTCAGG
E I K N S P S S L T S P A S Y N C S O S -
 ATCCTACCCTTGGAGCATGGTCCCATCTATATCCCTTCCTCCTATGTAGAGAGCCGTCAC 180
 121 -----+-----+-----+-----+-----+-----+-----+-----
 TAGGATGGGAACCTCGTACCAGGGTAGATATAGGGAAGGAGGATACATCTCTCGGCAGTG
I L P L E H G P I Y I P S S Y V E S R H -
 GAATACTCAGCCATGACATTCTACAGTCCTGCTGTGATGAACTACAGTGTTCCCAGCAGC 240
 181 -----+-----+-----+-----+-----+-----+-----+-----
 CTTATGAGTCGGTACTGTAAGATGTCAGGACGACACTACTTGATGTCACAAGGGTCGTCG
E Y S A M T F Y S P A V M N Y S V P S S -
 ACCGGTAACCTGGAAGGTGGGCCTGTTCCGCCAGACTGCAAGCCCAAATGTGCTATGGCCA 300
 241 -----+-----+-----+-----+-----+-----+-----+-----
 TGGCCATTGGACCTTCCACCCGGACAAGCGGTCTGACGTTCCGGGTTTACACGATACCGGT
T G N L E G G P V R Q T A S P N V L W P -
 ACTTCTGGACACCTCTCTCCTTTAGCCACCCACTGCCAATCATCGCTTCTCTATGCAGAA 360
 301 -----+-----+-----+-----+-----+-----+-----+-----
 TGAAGACCTGTGGAGAGAGGAAATCGGTGGGTGACGGTTAGTAGCGAAGAGATACGTCTT
T S G H L S P L A T H C Q S S L L Y A E -
 CCTCAAAAGAGTCCTTGGTGTGAAGCAAGATCACTAGAACACACCTTGCCTGTAAACAGA 420
 361 -----+-----+-----+-----+-----+-----+-----+-----
 GGAGTTTTTCTCAGGAACCACACTTCGTTCTAGTGATCTTGTGTGGAACGGACATTTGTCT
P Q K S P W C E A R S L E H T L P V N R -
 GAGACCCTGAAGAGGAAGCTTGGCGGGAGCGGTTGTGCCAGCCCTGTTACTAGTCCAAGC 480
 421 -----+-----+-----+-----+-----+-----+-----+-----
 CTCTGGGACTTCTCCTTCGAACCGCCCTCGCCAACACGGTCGGGACAATGATCAGGTTCTG
E T L K R K L G G S G C A S P V T S P S -
 GCCAAGAGGGATGCTCACTTCTGCGCCGTCTGCAGTGATTATGCATCTGGGTATCATTAC 540
 481 -----+-----+-----+-----+-----+-----+-----+-----
 CGGTTCTCCCTACGAGTGAAGACGCGGCAGACGTCCTAATACGTAGACCCATAGTAATG
A K R D A H F C A V C S D Y A S G Y H Y -
 GGTGTCTGGTCCTGTGAAGGATGTAAGGCCTTTTTTAAAAGAAGCATTCAAGGACATAAT 600
 541 -----+-----+-----+-----+-----+-----+-----+-----
 CCACAGACCAGGACACTTCCTACATTCCGGAAAAAATTTTCTTCGTAAGTTCCTGTATTA
G V W S C E G C K A F F K R S I Q G H N -

Figure 1b

WO 99/05170

3/20

PCT/US98/15539

601 GACTATATCTGTCCAGCCACGAATCAGTGTACCATAGACAAGAACCGGCGTAAAAGCTGC 660
-----+-----+-----+-----+-----+-----+-----+-----
CTGATATAGACAGGTCGGTGCTTAGTACATGGTATCTGTTCTTGGCCGCATTTTCGACG
D Y I C P A T N Q C T I D K N R R K S C -

661 CAGGCCTGCCGACTTCGCAAGTGTTACGAAGTAGGAATGGTCAAGTGTGGATCCAGGAGA 720
-----+-----+-----+-----+-----+-----+-----+-----
GTCCGGACGGCTGAAGCGTTCACAATGCTTCATCCTTACCAGTTCACACCTAGGTCCTCT
Q A C R L R K C Y E V G M V K C G S R R -

721 GAAAGGTGTGGGTACCGAATAGTACGAAGACAGAGAAGTGCCAGCGAGCAGGTGCATTGC 780
-----+-----+-----+-----+-----+-----+-----+-----
CTTTCCACACCCATGGCTTATCATGCTTCTGTCTCTTCACGGTCGCTCGTCCACGTAACG
E R C G Y R I V R R Q R S A S E Q V H C -

781 CTGAACAAAGCCAAGAGAACCAGTGGGCACACACCCCGGGTGAAGGAGCTACTGCTGAAC 840
-----+-----+-----+-----+-----+-----+-----+-----
GACTTGTTTTCGGTTCTCTTGGTCACCCGTGTGTGGGGCCCACTTCCTCGATGACGACTTG
L N K A K R T S G H T P R V K E L L L N -

841 TCTCTGAGTCCCGAGCAGCTGGTGCTCACCCTGCTGGAAGCTGAGCCACCCAATGTGCTA 900
-----+-----+-----+-----+-----+-----+-----+-----
AGAGACTCAGGGCTCGTCGACCACGAGTGGGACGACCTTCGACTCGGTGGGTACACGAT
S L S P E Q L V L T L L E A E P P N V L -

901 GTGAGCCGTCCCAGCATGCCCTTCACCGAGGCCTCCATGATGATGTCCCTCACGAAGCTG 960
-----+-----+-----+-----+-----+-----+-----+-----
CACTCGGCAGGGTCGTACGGGAAGTGGCTCCGGAGGTACTACTACAGGGAGTGCTTCGAC
V S R P S M P F T E A S M M M S L T K L -

961 GCTGACAAGGAAGTGGTGCACATGATTGGCTGGGCCAAGAAAATCCCTGGCTTTGTGGAG 1020
-----+-----+-----+-----+-----+-----+-----+-----
CGACTGTTCTTGACCACGTGTACTAACCGACCCGGTTCTTTTAGGGACCGAAACACCTC
A D K E L V H M I G W A K K I P G F V E -

1021 CTCAGCCTGTTGGACCAAGTCCGCCTCTTGGAAAGCTGCTGGATGGAGGTGCTGATGGTG 1080
-----+-----+-----+-----+-----+-----+-----+-----
GAGTCGGACAACCTGGTTCAGGCGGAGAACCTTTCGACGACCTACCTCCACGACTACCAC
L S L L D Q V R L L E S C W M E V L M V -

1081 GGGCTGATGTGGCGCTCCATCGACCACCCCGGCAAGCTCATCTTTGCTCCAGACCTCGTT 1140
-----+-----+-----+-----+-----+-----+-----+-----
CCCGACTACACCGCGAGGTAGCTGGTGGGGCCGTTTCGAGTAGAAACGAGGTCTGGAGCAA
G L M W R S I D H P G K L I F A P D L V -

1141 CTGGACAGgtctctcagaagaccctcactggcaggttgccgacagacgaagagtgcctgtccca 1200
-----+-----+-----+-----+-----+-----+-----+-----
GACCTGTCCaggagtcttctgggagtgaccgtgcaacgcgtctgcttctcagacagggt
L D R S S E D P H H V A O T K S A V R -

Figure 1b
(Continued)

WO 99/05170

4/20

PCT/US98/15539

```

1201  aggGATGAGGGGAAGTGCGTGGAAGGGATTCTGGAAATCTTTGACATGCTCCTGGCGACG 1260
-----+-----+-----+-----+-----+-----+-----+-----
      cccCTACTCCCCCTTCACGCACCTTCCCTAAGACCTTTAGAAACTGTACGAGGACCGCTGC
      B D E G K C V E G I L E I F D M L L A T -
1261  ACGGCACGGTTCCGTGAGTTAAAACTGCAGCACAAAGAATATCTGTGTGTGAAGGCCATG 1320
-----+-----+-----+-----+-----+-----+-----+-----
      TGCCGTGCCAAGGCACTCAATTTTGACGTCGTGTTTCTTATAGACACACACTTCCGGTAC
      T A R F R E L K L Q H K E Y L C V K A M -
1321  ATTCTCCTCAACTCCAGTATGTACCCCTTGGCTACCGCAAGCCAGGAAGCAGAGAGTAGC 1380
-----+-----+-----+-----+-----+-----+-----+-----
      TAAGAGGAGTTGAGGTCATACATGGGGAACCGATGGCGTTCGGTCCTTCGTCTCTCATCG
      I L L N S S M Y P L A T A S Q E A E S S -
1381  CGGAAGCTGACACACCTATTGAACGCAGTGACAGATGCCCTGGTCTGGGTGATTTCGAAG 1440
-----+-----+-----+-----+-----+-----+-----+-----
      GCCTTCGACTGTGTGGATAACTTGCGTCACTGTCTACGGGACCAGACCCACTAAAGCTTC
      R K L T H L L N A V T D A L V W V I S K -
1441  AGTGGAATCTCTTCCCAGCAGCAGTCAGTCCGTCTGGCCAACCTCCTGATGCTTCTTTCT 1500
-----+-----+-----+-----+-----+-----+-----+-----
      TCACCTTAGAGAAGGGTCGTCGTCAGTCAGGCAGACCGGTTGGAGGACTACGAAGAAAGA
      S G I S S Q Q Q S V R L A N L L M L L S -
1501  CATGTCAGGCACATCAGTAACAAGGGCATGGAACATCTGCTCAGCATGAAGTGCAAAAAT 1560
-----+-----+-----+-----+-----+-----+-----+-----
      GTACAGTCCGTGTAGTCATTGTTCCCGTACCTTGTAGACGAGTCGTACTTCACGTTTTTA
      H V R H I S N K G M E H L L S M K C K N -
1561  GTGGTCCCGGTGTACGACCTGCTGCTGGAGATGCTGAATGCTCACACGCTTCGAGGGTAC 1620
-----+-----+-----+-----+-----+-----+-----+-----
      CACCAGGGCCACATGCTGGACGACGACCTCTACGACTTACGAGTGTGCGAAGCTCCCATG
      V V P V Y D L L L E M L N A H T L R G Y -
1621  AAGTCCTCAATCTCGGGGTCTGAGTGCTGCTCGACAGAGGACAGTAAGAGCAAAGAGGGC 1680
-----+-----+-----+-----+-----+-----+-----+-----
      TTCAGGAGTTAGAGCCCCAGACTCACGACGAGCTGTCTCCTGTCATTCTCGTTTCTCCCG
      K S S I S G S E C C S T E D S K S K E G -
1681  TCCCAGAACCTCCAGTCACAGTGA 1704
-----+-----+-----+-----+-----+-----+-----+-----
      AGGGTCTTGGAGGTCAGTGTCACT

```

Figure 1b
(Continued)

1 MSICASSHKD ESOLRPTODM EIKNSPSSLT SPASYNCSOS ILELEHGPIY
51 IPSSYYESRE EYSAMTFYSP AVMNYSVPSS TGNLEGGPVR QTASPNVLWF
101 TSGHLSPLAT HCQSSLLYAE POKSPWCEAR SLEHTLPVNR ETLKRKLGGG
151 GCASPVTSPP AKRDAHECAV CSDYASGYHY GVNSCEGCKA FFKRSIQGHN
201 DYICPATNQC TIDKNRRKSC QACRLRKCYE VGMVKCGSRR ERCCGYRIVRE
251 QRSASEQVHC LNKAKRTSGH TPRVKELLN SLSPEQLVLT LLEAEPPNVI
301 VSRPSMPFTE ASMMMSLTKL ADKELVHMIG WAKKIPGFVE LSLLDQVRLI
351 ESCWMEVLMV GLMWRSIDHP GKLIFAPDLV LDRDEGKCV E GILEIFDMLI
401 ATTARFRELK LQHKEYLCVK AMILLNSSMY PLATASQEA E SSRKLTHLLN
451 AVTDALVWVI SKSGISSQQQ SVRLANLLML LSHVRHISNK GMEHLLSMKC
501 KNVVPVYDLL LEMLNAHTLR GYKSSISGSE CCSTEDSKSK EGSQNLSQ*

Figure 2a

1 MSICASSHKD FSQLRPTQDM EIKNSPSSLT SPASYNCSQS ILPLEHGPIY
51 IPSSYVESRH EYSAMTFYSP AVMNYSVPSS TGNLEGGPVR QTASPNVLWP
101 TSGHLSPLAT HCQSSLLYAE PQKSPWCEAR SLEHTLPVNR ETLKRKLGGS
151 GCASPVTSPTS AKRDAHFCV CSDYASGYHY GVWSCEGCKA FFKRSIQGSR
201 RERCGYRIVR RQRSASEQVH CLNKAKRTSG HTPRVKELL NSLSPEQLVL
251 TLLEAEPPNV LVS RPSMPFT EASMMMSLTK LADKELVHMI GWAKKIPGFV
301 ELSLLDQVRL LESCWMEVLM VGLMWRSIDH PGKLIFAPDL VLDRDEGKCV
351 EGILEIFDML LATTARFREL KLQHKEYLCV KAMILLNSSM YPLATASQEA
401 ESSRKLTHLL NAVTDALVWV ISKSGISSQQ QSVRLANLLM LLSHVRHISN
451 KGMEHLLSMK CKNVVPVYDL LLEMLNAHTL RGYKSSISGS ECCSTEDSKS
501 KEGSQNLQSQ

Figure 2b

WO 99/05170

PCT/US98/15539

7/20

1 ~~MSICTSGHTE~~ ~~TSOLRPSEDM~~ ~~EIKNSPSSLS~~ ~~SPASYNCSQS~~ ~~LLPLEHGRIY~~
 51 ~~TPSSXVDNRH~~ ~~EYSAMTEYSP~~ AVMNYSVFGS TSNLDGGPVR LSTSPNVLWP
 101 TSGHLSPLAT HCQSSLLYAE PQKSPWCEAR SLEHTLPVNR ETLKRKLSGS
 151 SCASPVTSPN AKRDAHFCEPV CSDYASGYHY GVWSCEGCKA EFKRSIQGHN
 201 DYICPATNQC TIDKNRRKSC QACRLRKCYE VGMVKCGSRR ERGGRIVRR
 251 QRSSSEQVHC LSKAKRNGGH APRVKELLS TLSPEQLVLT LLEAEPNVL
 301 VSRPSMPFTE ASMMMSLTKL ADKELVHMIG WAKKI?GFVE LSLDQVRL
 351 ESCWMEVLMV GLMWRSIDHP GKLI?APDLV LDRSSEDPHE ~~HVAOMKSAAP~~
 401 ~~RYVPLGFCCKP~~ GGRK*PEADT PTERGDRCPG LGDCSEWYLL PAAVSPTGCP
 451 PDASF SRQAH Q*QGHG TSAQ REVQKCGPGV *PAAGDAECS HASRVQVLNL
 501 GV*VQLNRGO *EQRELPEPT VSVMARPEAD RLQRNSKVEH VP*HLGVPLR
 551 AALVTHPLPT LHFPGVRVVV WRCSSYQDVP PNAKF*LV*P *RLSVYLLSV
 601 SLPTNFKHLKG SGTGQSLIW KDCP*SGKGI WHVTQL*EMD CRTVNP*NQP
 651 LDGVF*TT*L *D*KPH*QSA HFAFLPHGSV RTH*CHGLFY QRPER*CKLR
 701 CTLFLLPDPW VTSLEPAYLV CLNVVIVMG* DLNLFVILAS LKLCHLSLSE

Figure 2c

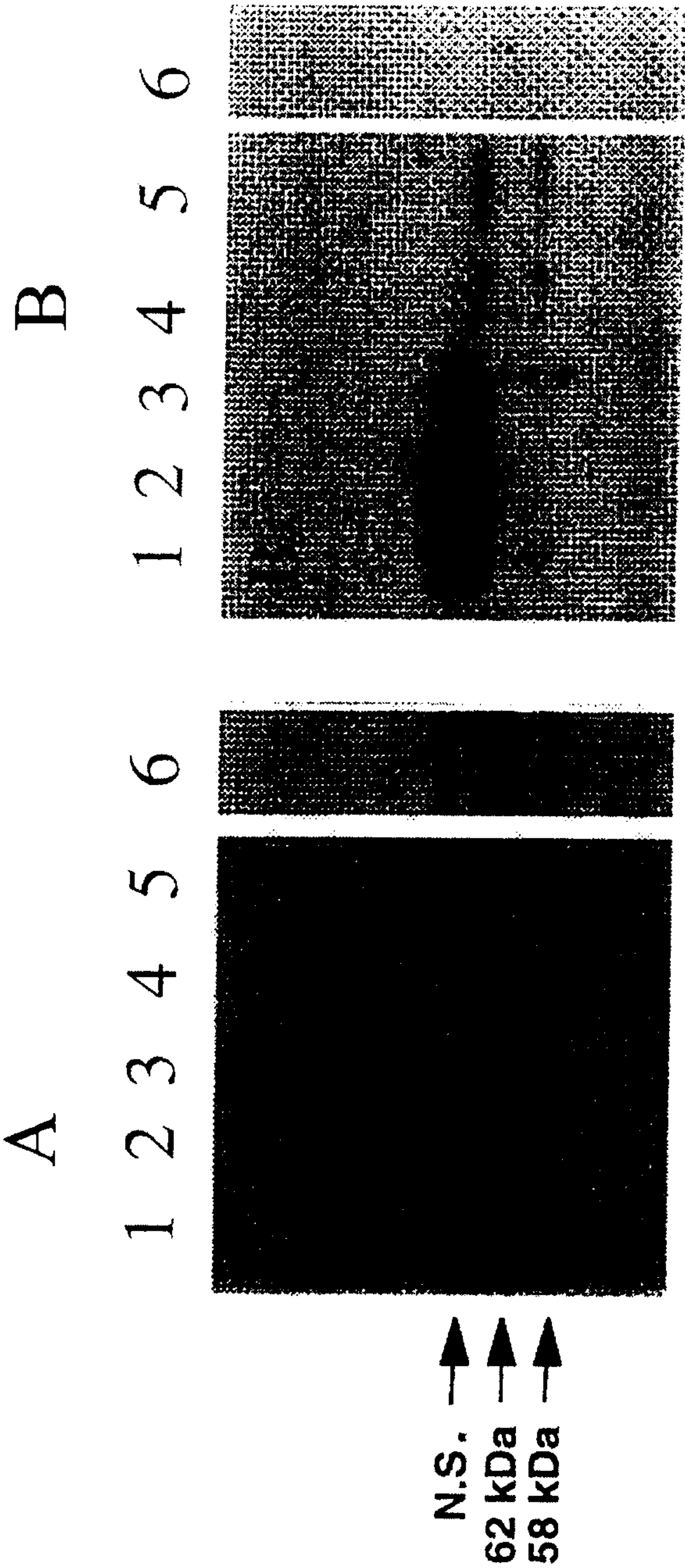


FIGURE 3

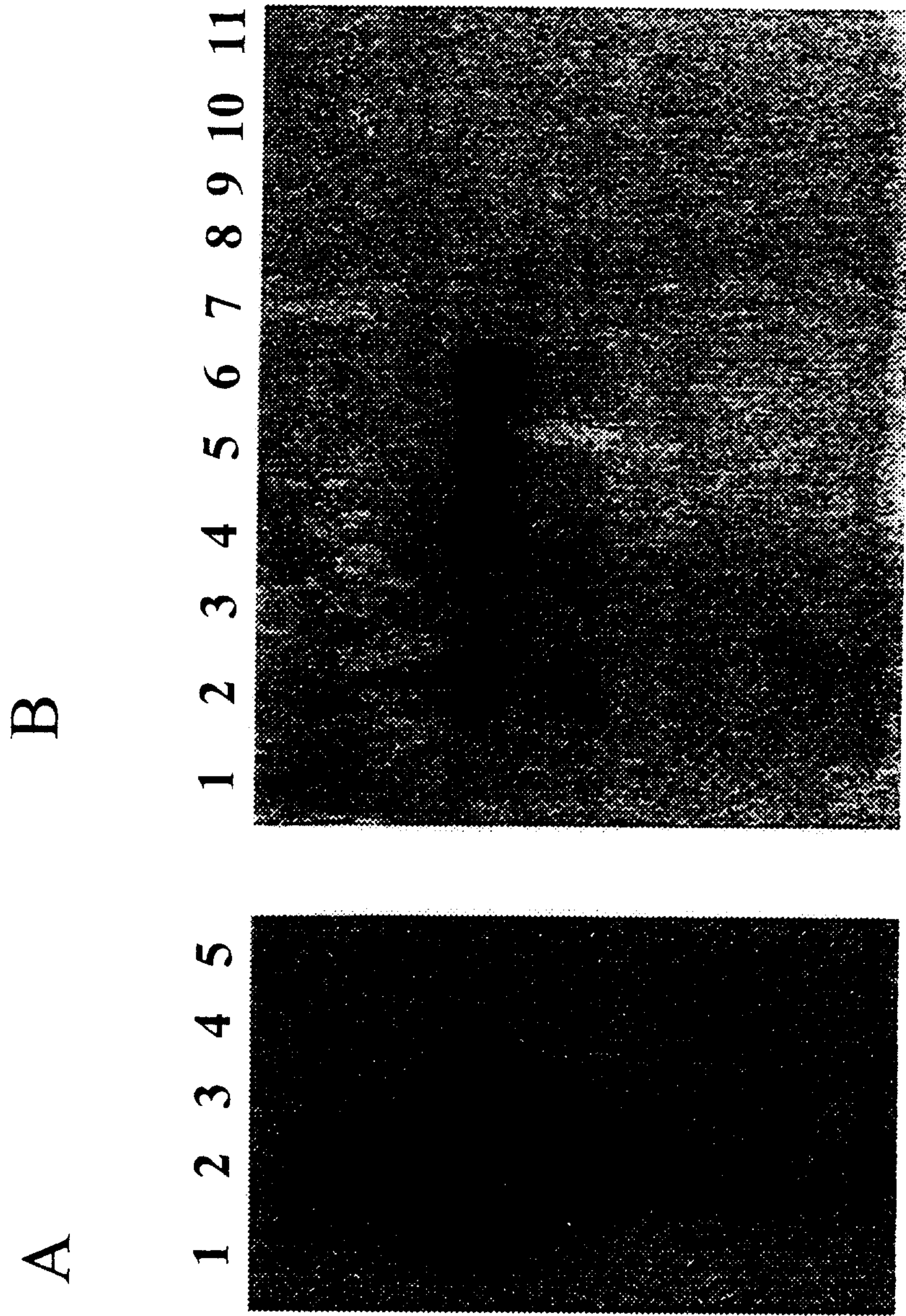


Figure 4

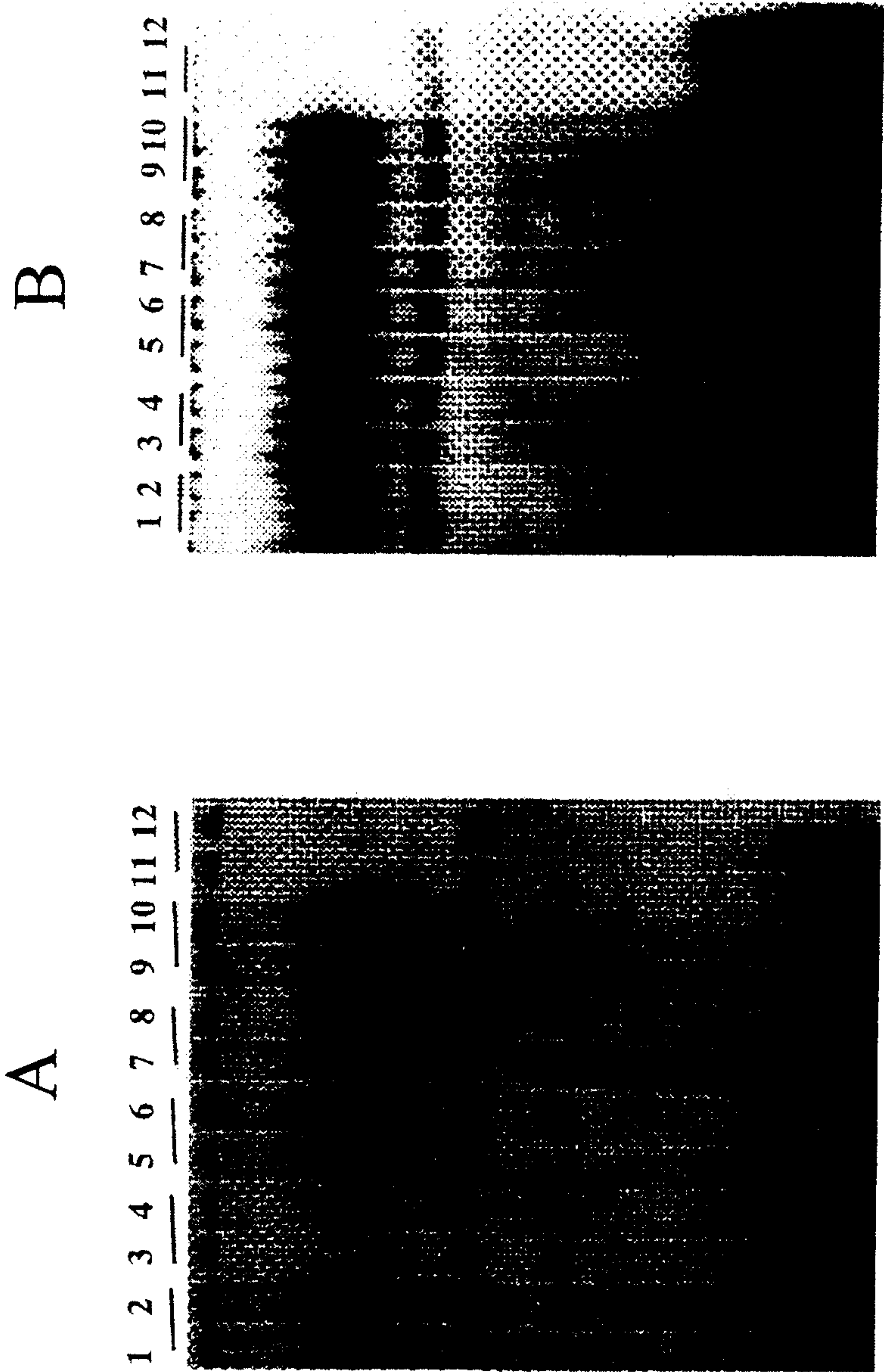


FIGURE 5

A

```

1  ....MSICASSHKDFSQLRPTQDMEIKNSP.....SSLTSPASYNCSQ 39
:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
1  MTMTLHTKASGMALLHQQ.GNELEPLNRPQLKMPMERALGEVYVDNSKP 49

40  SILPLEHGPIYIPSSYVESRHEYSAMTFYSPAVMNYSVPSSTGNLEGGPV 89
:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
50  TVFNYPEGAAYEFNAAAAAAAAAASA.PVYQSGIAYGPGSEAAAFSANS 98

90  .....RQTASPNVLWPTSGHLSP.LATHCQSSLLYAEPQKSPWC.... 127
:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
99  GAFPQLNSVSPSPLMLLHPPPQLSPFLHPHGQQVPYYLENEPSAYAVRDT 148

128  EARSLEHTLPVN.RETLKRKLGGSGCASPVTSAPSAKRDAHFCVCSOYAS 176
:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
149  GPPAFYRSNSDNRRQNGRERLSSSNEKGNMIMESAK.ETRYCAVCNDYAS 197

177  GYHYGVWSCEGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLR 226
:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
198  GYHYGVWSCEGCKAFFKRSIQGHNDYMCPATNQCTIDKNRRKSCQACRLR 247

227  KCYEVGWVKCGSRRERCYRIV..RRQRSASEQVHCLNKAKRTSGHTPRV 274
:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
248  KCYEVGWMMKGGIRKDRRGGRMLKHKRQRDDLEGRNEMGASGDMRAANLWP 297

275  KELLN.....SLSPEQLVLTLLAEPPNVLVS.RPSMPFTEASMM 314
:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
298  SPLVIKHTKKNSPALSLTADQMVSAALLDAEPPMIYSEYDPSRPFSEASMM 347

315  MSLTKLADKELVHMIGWAKKIPGFVELSLLDQVRLLESCWMEVLMVGLMW 364
:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
348  GLLTNLADRELVHMINWAKRVPFGDLNLHDQVHLLCAWLEILMIGLVW 397

365  RSIDHPGKLIFAPDLVLDORDEGKCEGILEIFDMLLATARFRELKLQHK 414
:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
398  RSMEHPGKLLFAPNLLLDNRNQKCEGMEVEIFDMLLATSSRFRMMNLQGE 447

415  EYLCVKAMILLNSMYP.LATASQEAESSRKLTHLLNAVTDALVWVSKS 463
:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
448  EFVCLKSIILLNSGVYTFLSSTLKSLEEKDHIHRVLOKITDTLIHLMKA 497

464  GISSQQQSVRLANLLMLSHVRHISNKGMEHLLSMKCKNVVPVYDLLLEM 513
:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
498  GLTLQQQHRRLAQLLLILSHIRHMSNKGMEHLYNMKCKNVVPLYDLLLEM 547

514  LNAHTLRGYKS..SISGSECCSTE.DSKSKEGSQNLQSQ..... 549
:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
548  LDAHRLHAPASRMGVPPEEPSQTQLATTSSTSAHSLQTYYPPEAEGFPN 597

```

Figure 6

Figure 7

151 ATCCCTTCCTCCTATGTAGAGAGCCGTCACGAATACTCAGCCATGACATT 200
1 TEGAGGTGACATT 19
201 CTACAGTCTGCTGTGATGAACTACAGTGTTCAGCAGCACC GGTAACC 250
20 CTACAGTCTGCTGTGATGAACTACAGTGTTCAGCAGCACC GGTAACC 69
251 TGGAAAGGTGGGCCTGTTCCGAGACTGCAAGCCCAAATGTGCTATGGCCA 300
70 TGGAAAGGTGGGCCTGTTCCGAGACTGCAAGCCCAAATGTGCTATGGCCA 119
301 ACTTCTGGACACCTCTCTCTTTAGCCACCCACTGCCAATCATCGCTTCT 350
120 ACTTCTGGACACCTCTCTCTTTAGCCACCCACTGCCAATCATCGCTTCT 169
351 CTATGCAGAACCTCAAAAGAGTCCTTGGTGTGAAGCAAGATCACTAGAAC 400
170 CTATGCAGAACCTCAAAAGAGTCCTTGGTGTGAAGCAAGATCACTAGAAC 219
401 ACACCTTGCCCTGTAAACAGAGAGACCCTGAAGAGGAAGCTTGGCGGGAGC 450
220 ACACCTTGCCCTGTAAACAGAGAGACCCTGAAGAGGAAGCTTGGCGGGAGC 269
451 GGTTGTGCCAGCCCTGTTACTAGTCCAAGCGCCAAGAGGGATGCTCACTT 500
270 GGTTGTGCCAGCCCTGTTACTAGTCCAAGCGCCAAGAGGGATGCTCACTT 319
501 CTGCGCCGTCTGCAGTGATTATGCATCTGGGTATCATTACGGTGTCTGGT 550
320 CTGCGCCGTCTGCAGTGATTATGCATCTGGGTATCATTACGGTGTCTGGT 369
551 CCTGTGAAGGATGTAAGGCCTTTTAAAGAAGCATTCAAGGACATAAT 600
370 CCTGTGAAGGATGTAAGGCCTTTTAAAGAAGCATTCAAGGACATAAT 419
601 GACTATATCTGTCCAGCCACGAATCAGTGTACCATAGACAAGAACCGGCG 650
420 GACTATATCTGTCCAGCCACGAATCAGTGTACCATAGACAAGAACCGGCG 469
651 TAAAAGCTGCCAGGCCTGCCGACTTCGCAAGTGTTACGAAGTAGGAATGG 700
470 TAAAAGCTGCCAGGCCTGCCGACTTCGCAAGTGTTACGAAGTAGGAATGG 519
701 TCAAGTGTGGATCCAGGAGAGAAAGGTGTGGGTACCGAATAGTACGAAGA 750

Figure 7 (continued)

520 TCAAGTGTGGATCCAGGAGAGAAAGGTGTGGGTACCGAATAGTACGAAGA 569
 751 CAGAGAAGTGCCAGCGAGCAGGTGCATTGCCTGAACAAAGCCAAGAGAAC 800
 ||
 570 CAGAGAAGTGCCAGCGAGCAGGTGCATTGCCTGAACAAAGCCAAGAGAAC 619
 801 CAGTGGGCACACACCCCGGGTGAAGGAGCTACTGCTGAACTCTCTGAGTC 850
 ||
 620 CAGTGGGCACACACCCCGGGTGAAGGAGCTACTGCTGAACTCTCTGAGTC 669
 851 CCGAGCAGCTGGTGTCAACCTGCTGGAAGCTGAGCCACCCAATGTGCTA 900
 ||
 678 CCGAGCAGCTGGTGTCAACCTGCTGGAAGCTGAGCCACCCAATGTGCTA 719
 901 GTGAGCCGTCCAGCATGCCCTTCACCGAGGCCTCCATGATGATGTCCCT 950
 ||
 720 GTGAGCCGTCCAGCATGCCCTTCACCGAGGCCTCCATGATGATGTCCCT 769
 951 CACGAAGCTGGCTGACAAGGAACTGGTGCACATGATTGGCTGGGCCAAGA 1000
 ||
 770 CACGAAGCTGGCTGACAAGGAACTGGTGCACATGATTGGCTGGGCCAAGA 819
 1001 AAATCCCTGGCTTTGTGGAGCTCAGCCTGTTGGACCAAGTCCGCCTCTTG 1050
 ||
 820 AAATCCCTGGCTTTGTGGAGCTCAGCCTGTTGGACCAAGTCCGCCTCTTG 869
 1051 GAAAGCTGCTGGATGGAGGTGCTGATGGTGGGGCTGATGTGGCGCTCCAT 1100
 ||
 870 GAAAGCTGCTGGATGGAGGTGCTGATGGTGGGGCTGATGTGGCGCTCCAT 919
 1101 CGACCACCCCGGCAAGCTCATCTTTGCTCCAGACCTCGTTCTGGACAGGT 1150
 ||
 920 CGACCACCCCGGCAAGCTCATCTTTGCTCCAGACCTCGTTCTGGAC.... 965

1201 AGGGATGAGGGGAAGTGCGTGGAAAGGATTCTGGAAATCTTTGACATGCT 1250
 ||
 968 AGGGATGAGGGGAAGTGCGTGGAAAGGATTCTGGAAATCTTTGACATGCT 1015
 ||
 1251 CCTGGCGACGACGGCAGGTTCCGTGAGTTAAACTGCAGCACAAAGAAT 1300
 ||
 1016 CCTGGCGACGACGGCAGGTTCCGTGAGTTAAACTGCAGCACAAAGAAT 1065
 1301 ATCTGTGTGTGAAGGCCATGATTCTCTCAACTCCAGTATGTACCCCTTG 1350
 ||
 1066 ATCTGTGTGTGAAGGCCATGATTCTCTCAACTCCAGTATGTACCCCTTG 1115
 1351 GCTACCGCAAGCCAGGAAGCAGAGAGTAGCCGGAAGCTGACACACCTATT 1400
 ||
 1118 GCTACCGCAAGCCAGGAAGCAGAGAGTAGCCGGAAGCTGACACACCTATT 1165
 1401 GAACGCAGTGACAGATGCCCTGGTCTGGGTGATTTGGAAGAGTGAATCT 1450
 ||
 1166 GAACGCAGTGACAGATGCCCTGGTCTGGGTGATTTGGAAGAGTGAATCT 1215
 1451 CTTCCCAGCAGCAGTCAGTCCGTCTGGCCAACCTCCTGATGCTTCTTTCT 1500
 ||
 1216 CTTCCCAGCAGCAGTCAGTCCGTCTGGCCAACCTCCTGATGCTTCTTTCT 1265
 1501 CATGTCAGGCACATCAGTAACAAGGGCATGGAACATCTGCTCAGCATGAA 1550
 ||
 1266 CATGTCAGGCACATCAGTAACAAGGGCATGGAACATCTGCTCAGCATGAA 1315
 1551 GTGCAAAAATGTGGTCCCGGTGTACGACCTGCTGCTGGAGATGCTGAATG 1600
 ||
 1316 GTGCAAAAATGTGGTCCCGGTGTACGACCTGCTGCTGGAGATGCTGAATG 1365

D
 Unique
 Nucleotide
 difference.

WO 99/05170

PCT/US98/15539

14/20

```
1601 CTCACACGCTTCGAGGGTACAAGTCCTCAATCTCGGGGTCTGAGTGCTGC 1650
      |||
1366 CTCACACGCTTCGAGGGTACAAGTCCTCAATCTCGGGGTCTGAGTGCTGC 1415
      |||
1651 TCGACAGAGGACAGTAAGAGCAAAGAGGGCTCCAGAACCTCCAGTCACA 1700
      |||
1416 TCGACAGAGGACAGTAAGAGCAAAGAGGGCTCCAGAACCTCCAGTCACA 1465
      |||
1701 GTGA..... 1704
      |||
1466 GTGACGGCCAGGCTGGAGGCGGATGATTACAATGATGGTCAGAAGTGGGA 1515
```

Figure 7 (continued)

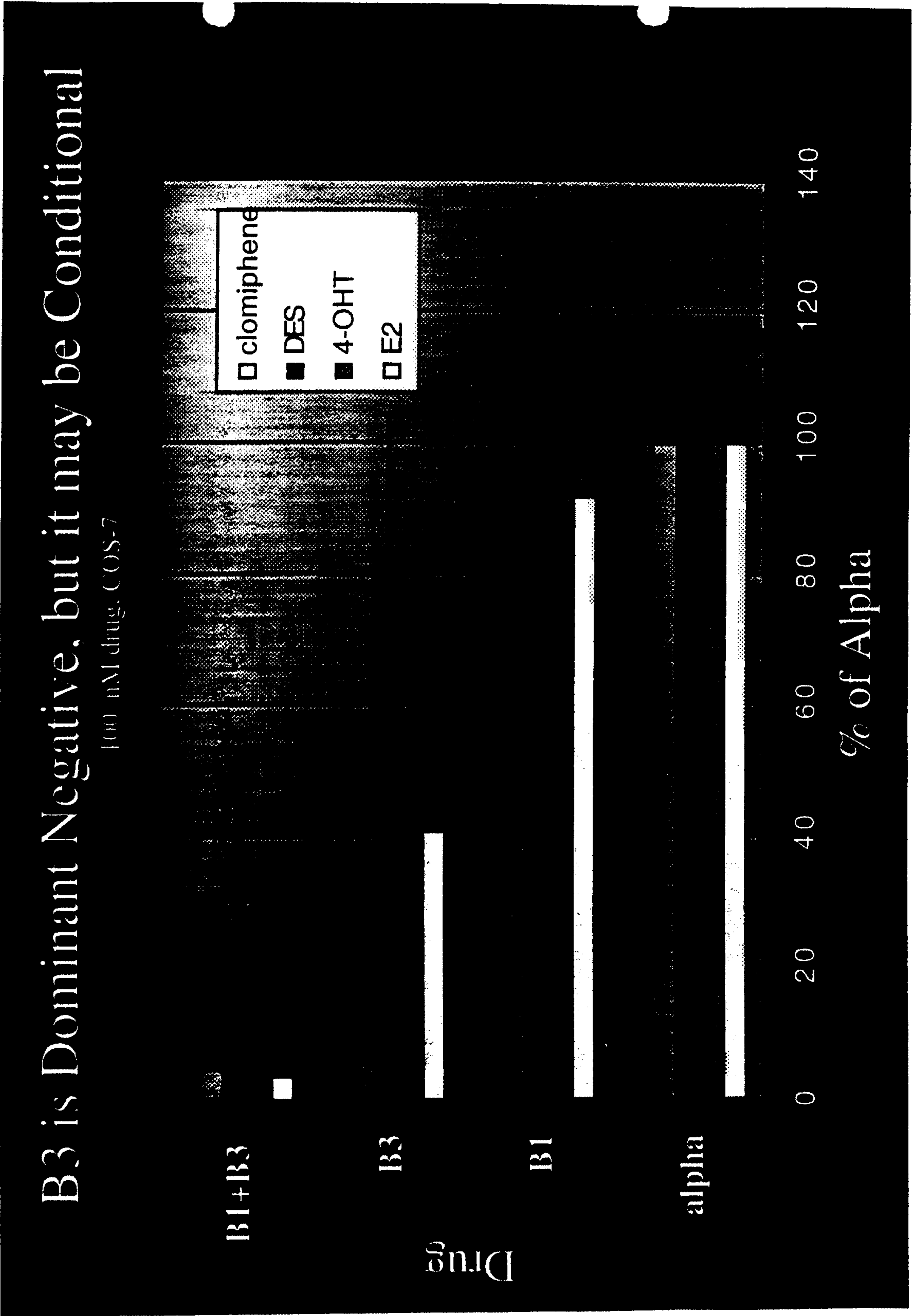


FIGURE 8

Transactivation Profiles-cV2ERE

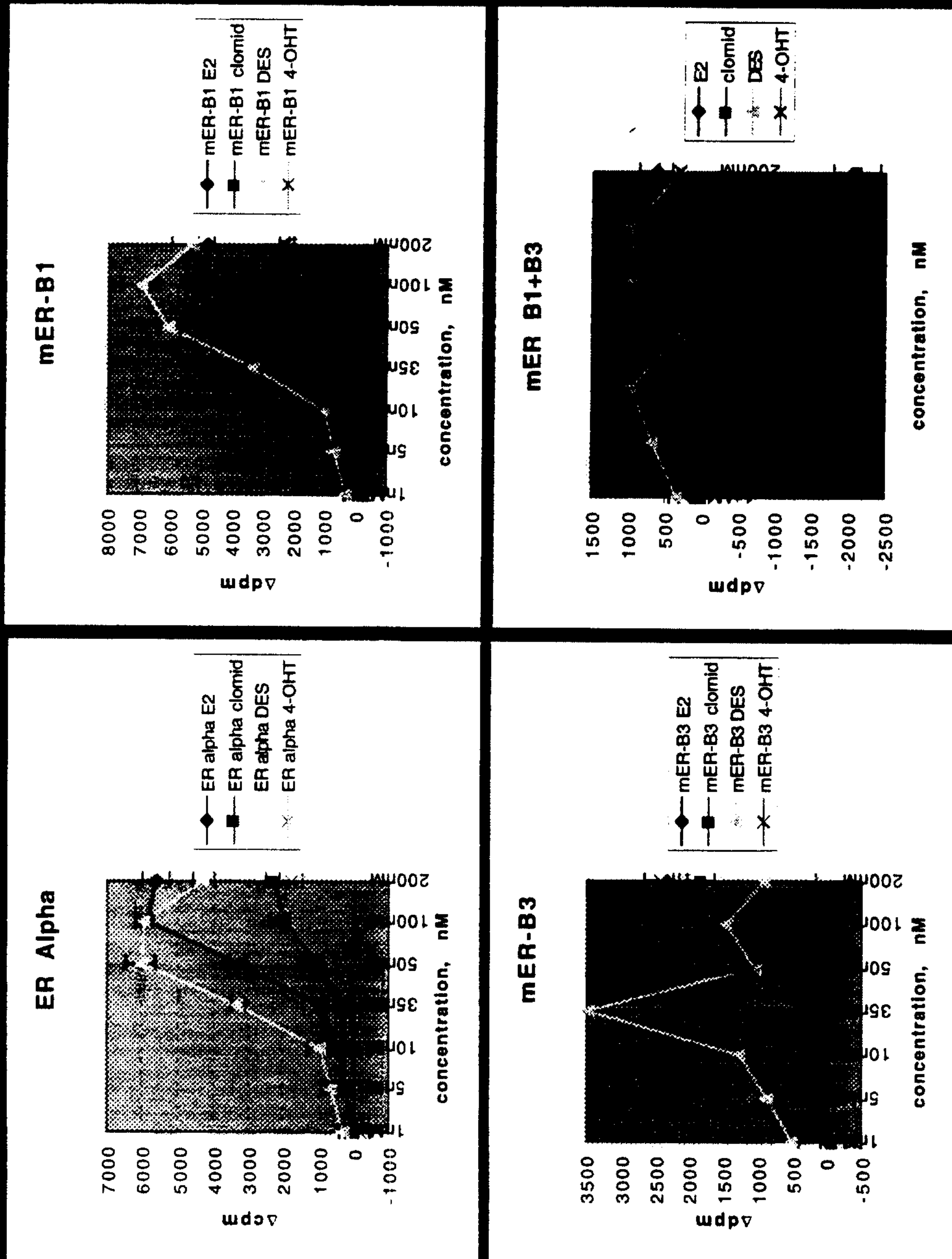


FIGURE 9

WO 99/05170

17/20

PCT/US98/15539

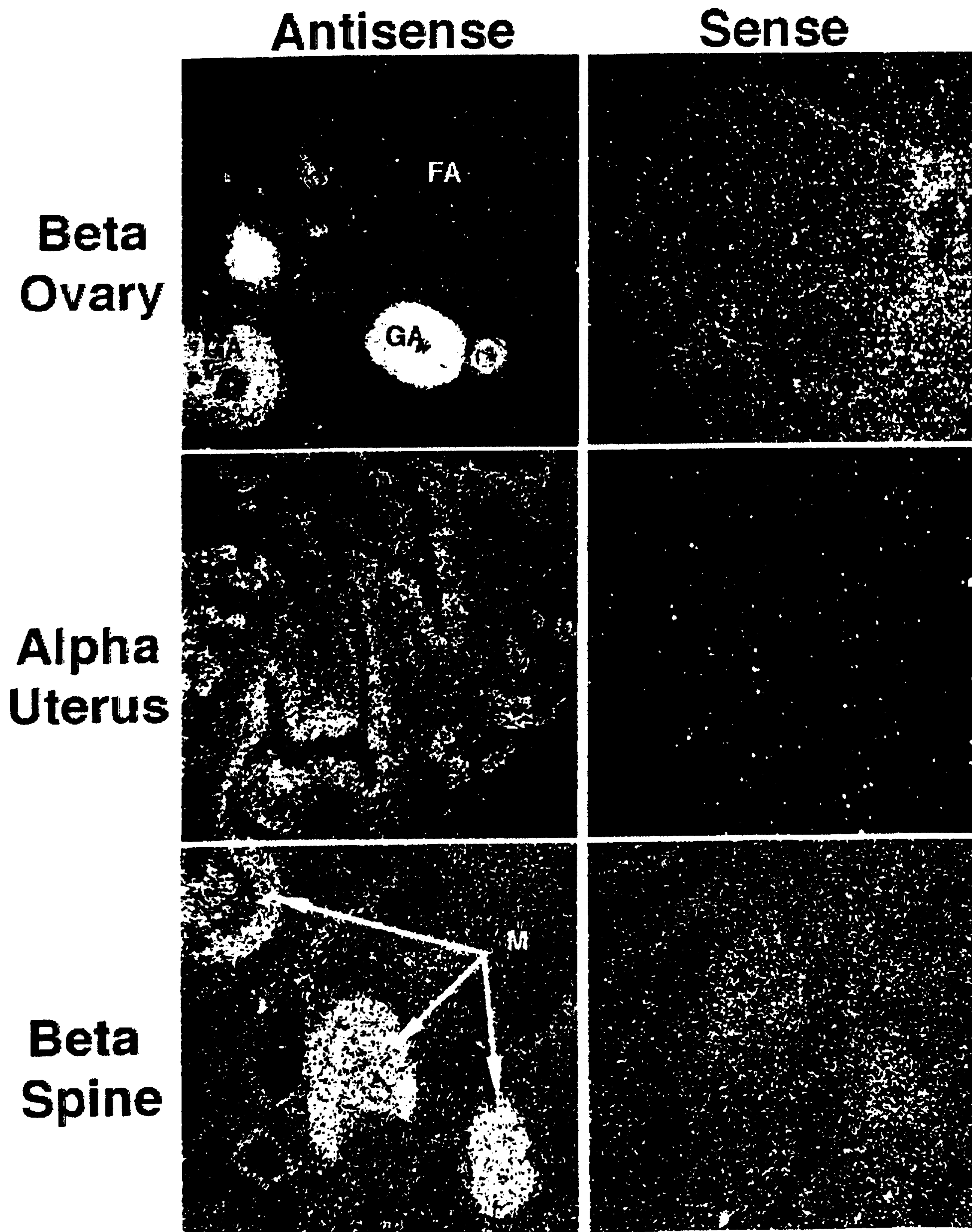


FIGURE 10

1 2 3 4 5 6 7 8



FIGURE 11

FIGURE 12A

112 rat exon 5B	1290	1300	1310	1320	1330	1340	1350	1360
113								
109 murine exon 5B								
114								
111 human intron 5								
	1370	1380	1390	1400	1410	1420	1430	1440
112 rat exon 5B	GTGCTGCCCC AAG							
113								
109 murine exon 5B	GTGCTGCCCC AAG							
114								
111 human intron 5	GTGATGCCCC AAGGTAACAG TCCCATCTCT GCCTTTTGTG GTGGATGACA ATGTTATTAA AAGAGCACTG TTC'INGCATAA							
	1450	1460	1470	1480	1490	1500	1510	1520
112 rat exon 5B	<==							
113	<==							
109 murine exon 5B	<==							
114	<==							
111 human intron 5	TGGTGTGTTTG ATAGAGAACA GATCCTCTGA GAAGAGCTGG AGGACTGATG TGACTTGAAC AGGAGCAAGC CCAGGTGGTA							
	1530	1540	1550	1560	1570	1580	1590	1600
112 rat exon 5B	<==							
113	<==							
109 murine exon 5B	<==							
114	<==							
111 human intron 5	AACCATGGAG GGAGGCTCTG GGAAGACCAG AGAAGTTTCTAG GGCACAGAC CCTTCAGTAA CAAACAAAT AGTTAACCTA							
	1610	1620	1630	1640	1650	1660	1670	1680
112 rat exon 5B	<==							
113	<==							
109 murine exon 5B	<==							
114	<==							
111 human intron 5	TGGCTTCTGA TGTGCTTGGC AGCACCCTAT GCAITTTAACT TATGTCACAA CATTTAACT TCACAAATCT CCTGCCCCCT							
	1690	1700	1710	1720	1730	1740	1750	1760
112 rat exon 5B	<==							
113	<==							
109 murine exon 5B	<==							
114	<==							
111 human intron 5	TTGAGGGAGT AGGATCCATT ATTATCTCTA TCATTTTCAGAT ATTGGAAATG GGAGATTGAG AAACCTGCTT ACAGGTAGGA							

FIGURE 12B

