



Office de la Propriété
Intellectuelle
du Canada

Un organisme
d'Industrie Canada

Canadian
Intellectual Property
Office

An agency of
Industry Canada

CA 2240292 C 2013/05/28

(11)(21) **2 240 292**

(12) **BREVET CANADIEN
CANADIAN PATENT**

(13) **C**

(86) Date de dépôt PCT/PCT Filing Date: 1996/12/19

(87) Date publication PCT/PCT Publication Date: 1997/07/10

(45) Date de délivrance/Issue Date: 2013/05/28

(85) Entrée phase nationale/National Entry: 1998/06/11

(86) N° demande PCT/PCT Application No.: GB 1996/003164

(87) N° publication PCT/PCT Publication No.: 1997/024445

(30) Priorité/Priority: 1995/12/30 (GB9526733.2)

(51) Cl.Int./Int.Cl. *C12N 15/62* (2006.01),
A61K 38/27 (2006.01), *A61K 38/38* (2006.01),
A61K 47/48 (2006.01), *C07K 14/61* (2006.01),
C07K 14/765 (2006.01), *C07K 19/00* (2006.01),
C12N 15/14 (2006.01), *C12N 15/18* (2006.01),
C12P 21/06 (2006.01), *A61K 38/00* (2006.01)

(72) Inventeur/Inventor:
BALLANCE, DAVID JAMES, GB

(73) Propriétaire/Owner:
NOVOZYMES BIOPHARMA DK A/S, DK

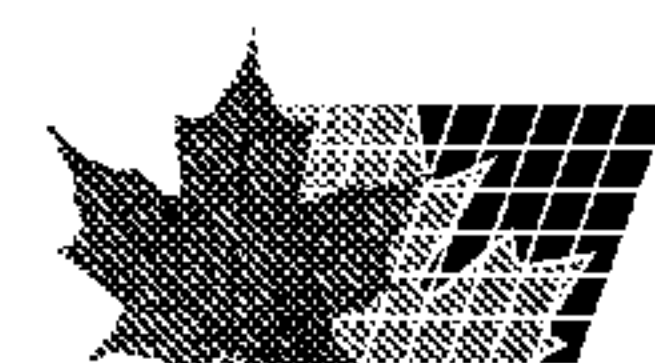
(74) Agent: MBM INTELLECTUAL PROPERTY LAW LLP

(54) Titre : PROTEINES DE FUSION RECOMBINEES D'HORMONE DE CROISSANCE ET D'ALBUMINE SERIQUE

(54) Title: RECOMBINANT FUSION PROTEINS TO GROWTH HORMONE AND SERUM ALBUMIN

(57) Abrégé/Abstract:

Fusion proteins of albumin and growth hormone, or fusions of variants of either, are secreted well in yeast and have increased serum and storage stability.



**PCT**WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C12N 15/62, 15/14, 15/18, C07K 14/765, 14/61, A61K 38/27, 38/38, C07K 19/00	A1	(11) International Publication Number: WO 97/24445 (43) International Publication Date: 10 July 1997 (10.07.97)
(21) International Application Number: PCT/GB96/03164 (22) International Filing Date: 19 December 1996 (19.12.96) (30) Priority Data: 9526733.2 30 December 1995 (30.12.95) GB (71) Applicant (for all designated States except US): DELTA BIOTECHNOLOGY LIMITED [GB/GB]; Castle Court, Castle Boulevard, Nottingham NG7 1FD (GB). (72) Inventor; and (75) Inventor/Applicant (for US only): BALLANCE, David, James [GB/GB]; Delta Biotechnology Limited, Castle Court, Castle Boulevard, Nottingham NG7 1FD (GB). (74) Agent: BASSETT, Richard; Eric Potter Clarkson, St. Mary's Court, St. Mary's Gate, Nottingham NG1 1LE (GB).		(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, ARIPO patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG). 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: RECOMBINANT FUSION PROTEINS TO GROWTH HORMONE AND SERUM ALBUMIN (57) Abstract Fusion proteins of albumin and growth hormone, or fusions of variants of either, are secreted well in yeast and have increased serum and storage stability.		

RECOMBINANT FUSION PROTEINS TO GROWTH HORMONE AND SERUM ALBUMIN

Field of the invention

- 5 The present invention relates to recombinant fusion proteins, to growth hormone (GH), to serum albumin and to production of proteins in yeast.

Background and prior art

- 10 Human serum albumin (HSA), a protein of 585 amino acids, is responsible for a significant proportion of the osmotic pressure of serum and also functions as a carrier of endogenous and exogenous ligands. At present, HSA for clinical use is produced by extraction from human blood. The production of recombinant HA (rHA) in microorganisms has
15 been disclosed in EP 330 451 and EP 361 991.

- The role of albumin as a carrier molecule and its inert nature are desirable properties for use as a stabiliser and transporter of polypeptides. The use of albumin as a component of a fusion protein for stabilising other proteins
20 has been disclosed in WO 93/15199, WO 93/15200, and EP 413 622. The use of N-terminal fragments of HSA for fusions to polypeptides has also been disclosed (EP 399 666). Fusion to the said polypeptide is achieved by genetic manipulation, such that the DNA coding for HSA, or a fragment thereof, is joined to the DNA coding for the said polypeptide.
- 25 A suitable host is then transformed or transfected with the fused nucleotide sequences, so arranged on a suitable plasmid as to express a fusion polypeptide. Nomura *et al* (1995) attempted to express human apolipoprotein E in *S. cerevisiae* as a fusion protein with HSA or fragments of HSA, using the HSA pre-sequence to direct secretion.
- 30 Whilst fusion to full length HSA resulted in the secretion of low levels of

the protein into the medium (maximum yield of 6.3 mg per litre), fusion to HSA (1-198) or HSA (1-390) did not result in secretion into the medium.

5 Human growth hormone (reviewed by Strobl and Thomas, 1994) consists of a single polypeptide of 191 amino acids, internally cross-linked by two disulphide bonds. Two molecules of hGH receptor bind each molecule of hGH to facilitate signal transduction (Cunningham *et al*, 1991; de Vos *et al*, 1992). The C-terminus of the hGH molecule is involved in binding to
10 the first receptor molecule, but the extent to which the N-terminus is involved in receptor binding is not known. The hormone is secreted from the anterior pituitary gland under hypothalamic control, and is responsible for a wide range of growth-promoting effects in the body. Clinically, hGH is used in the treatment of hypopituitary dwarfism, chronic renal
15 insufficiency in childhood, bone fractures and burns. Current methods of production of hGH for therapeutic use are by extraction from human pituitary gland, recombinant expression in *Escherichia coli* as disclosed in EP 127 305 (Genentech) or recombinant expression in mammalian cell culture (Zeisel *et al*, 1992).

20

In addition, hGH has been expressed intracellularly in yeast (Tokunaga *et al*, 1985) and this organism may provide an alternative means of production as disclosed in EP 60 057 (Genentech). Tsiomenko *et al* (1994) reported the role of the yeast MF α -1 prepro leader sequence in the
25 secretion of hGH from yeast. Attachment of the pre-portion of the leader sequence to the hGH gene resulted in hGH accumulation in the periplasm and vacuoles, whilst attachment of the pro-portion to hGH resulted in expression of a non-glycosylated precursor localised inside the cell. Only when both portions of the leader sequence were attached to the hGH gene
30 was hGH secreted into the culture medium. Other secretion signals (pre-

sequences) were also ineffective unless a yeast-derived pro sequence was used, suggesting that such a pro sequence was used is critical to the efficient secretion of hGH in yeast.

5 In humans, hGH is secreted into the blood in pulses, and in the circulation has a half-life of less than 20 minutes (Haffner *et al*, 1994). Elimination of the hormone is primarily via metabolism in the liver and kidneys and is more rapid in adults than in children (Kearns *et al*, 1991). Treatment for hGH deficiency generally lasts for 6 to 24 months, during which hGH
10 is administered either three times a week intramuscularly or on a daily basis subcutaneously. Such a regimen of frequent administration is necessary because of the short half-life of the molecule.

Poznansky *et al* (1988) increased the half-life of porcine growth hormone
15 by conjugation with either porcine or human serum albumin (HSA) to form relatively large conjugates of about 180 kD. Chemical reaction using the cross-linking reagent glutaraldehyde resulted in, on average, two molecules of albumin complexed with six molecules of growth hormone. The resulting 180 kD conjugate was found to have an extended half-life
20 in the circulation of rats of 2 to 3 hours, compared to 5 minutes for unconjugated growth hormone. Activity assays showed that the conjugate retained full, and possibly increased activity *in vitro*, but was inactive *in vivo*.

25 Summary of the invention

The invention relates to proteins formed by the fusion of a molecule of albumin, or variants or fragments thereof, to a molecule of growth hormone or variants or fragments thereof, the fusion proteins having an
30 increased circulatory half-life over unfused growth hormone. For

convenience, we shall refer to human albumin (HA) and human growth hormone (hGH), but the albumin and growth hormones of other vertebrates are included also. Preferably, the fusion protein comprises HA, or a variant or fragment thereof, as the N-terminal portion, with hGH or a variant or fragment thereof as the C-terminal portion, so as to minimise any possible negative effects on receptor binding. Alternatively, a fusion protein comprising HA, or a variant or fragment thereof, as the C-terminal portion, with hGH or a variant or fragment thereof as the N-terminal portion, may also be capable of signal transduction. Generally, the polypeptide has only one HA-derived region and one GH-derived region.

Additionally, the fusion proteins of the invention may include a linker peptide between the two fused portions to provide a greater physical separation between the two moieties and thus maximise the availability of the hGH portion to bind the hGH receptor. The linker peptide may consist of amino acids such that it is flexible or more rigid.

The linker sequence may be cleavable by a protease or chemically to yield the growth hormone related moiety. Preferably, the protease is one which is produced naturally by the host, for example the *S. cerevisiae* protease kex2 or equivalent proteases. Hence, a further aspect of the invention provides a process for preparing growth hormone or a variant or fragment thereof by expressing a polynucleotide which encodes a polypeptide of the invention in a suitable host, cleaving the cleavable linker to yield the GH-type compound and recovering the GH-type compound from the host culture in a more pure form.

We have discovered that the polypeptides of the invention are significantly more stable in solution than hGH. The latter rapidly becomes inactive

when stored in solution at 4°C for over one month. Currently marketed hGH is sold as a freeze-dried powder.

Suitably, the fusion polypeptides are produced as recombinant molecules by secretion from yeast, a microorganism such as a bacterium, human cell line or a yeast. Preferably, the polypeptide is secreted from the host. We have found that, by fusing the hGH coding sequence to the HA coding sequence, either to the 5' end or 3' end, it is possible to secrete the fusion protein from yeast without the requirement for a yeast-derived pro sequence. This was surprising, as other workers have found that a yeast derived pro sequence was needed for efficient secretion of hGH in yeast. For example, Hiramitsu *et al* (1990, 1991) found that the N-terminal portion of the pro sequence in the *Mucor pusillus* rennin pre-pro leader was important. Other authors, using the MF α -1 signal, have always included the MF α -1 pro sequence when secreting hGH. The pro sequences were believed to assist in the folding of the hGH by acting as an intramolecular chaperone. The present invention shows that HA or fragments of HA can perform a similar function.

Hence, a particular embodiment of the invention comprises a DNA construct encoding a signal sequence effective for directing secretion in yeast, particularly a yeast-derived signal sequence (especially one which is homologous to the yeast host), and the fused molecule of the first aspect of the invention, there being no yeast-derived pro sequence between the signal and the mature polypeptide.

The *Saccharomyces cerevisiae* invertase signal is a preferred example of a yeast-derived signal.

Conjugates of the kind prepared by Poznansky *et al* (1988), in which

separately-prepared polypeptides are joined by chemical cross-linking, are not contemplated.

The albumin or hGH may be a variant of normal HSA/rHA (termed hereinafter "HA") or hGH, respectively. By "variants" we include insertions, deletions and substitutions, either conservative or non-conservative, where such changes do not substantially alter one or more of the oncotic, useful ligand-binding and non-immunogenic properties of albumin or, in the case of hGH, its non-immunogenicity and ability to bind and activate the hGH receptor. In particular, we include naturally-occurring polymorphic variants of human albumin and fragments of human albumin, for example those fragments disclosed in EP 322 094 (namely HA (1-n), where n is 369 to 419). The albumin or growth hormone may be from any vertebrate, especially any mammal, for example human, cow, sheep, pig, hen or salmon. The albumin and GH parts of the fusion may be from differing animals.

By "conservative substitutions" is intended swaps within groups such as Gly, Ala; Val, Ile, Leu; Asp, Glu; Asn, Gln; Ser, Thr; Lys, Arg; and Phe, Tyr. The variant will usually have at least 75% (preferably at least 80%, 90%, 95% or 99%) sequence identity with a length of normal HA or hGH which is the same length as the variant and which is more identical thereto than any other length of normal HA or hGH, once the allowance is made for deletions and insertions as is customary in this art. Generally speaking, an HA variant will be at least 100 amino acids long, preferably at least 150 amino acids long. The HA variant may consist of or comprise at least one whole domain of HA, for example domains 1 (1-194), 2 (195-387), 3 (388-585), 1 + 2 (1-387), 2 + 3 (195-585) or 1 + 3 (1-194, + 388-585). Each domain is itself made up of two homologous subdomains namely 1-105, 120-194, 195-291, 316-387, 388-491 and 512-

585, with flexible inter-subdomain linker regions comprising residues Lys106 to Glu199, Glu292 to Val315 and Glu492 to Ala511.. Preferably, the HA part of the fusion comprises at least one subdomain or domain of HA or conservative modifications thereof. If the fusion is based on
5 subdomains, some or all of the adjacent linker is preferably used to link to the hGH moiety. The hGH variant should have GH activity, and will generally have at least 10 amino acids, (although some authors have found activity with only 4 residues), preferably at least 20, preferably at least 50, 100, 150, 180 or 191, amino acids long, and preferably retains its
10 cysteines for both internal disulphide bonds.

The fused molecules of the invention generally have a molecular weight of less than 100 kD, for example less than 90 kD or 70 kD. They are therefore much smaller than the 180 kD conjugates of Poznansky *et al*
15 (referred to above), which were inactive *in vivo*. They will normally have a molecular weight of at least 20 kD, usually at least 30 kD or 50 kD. Most fall within the molecular weight range 60-90 kD.

A second main aspect of the invention provides a yeast transformed to
20 express a fusion protein of the invention.

In addition to the transformed host cells themselves, the present invention also contemplates a culture of those cells, preferably a monoclonal (clonally homogeneous) culture, or a culture derived from a monoclonal
25 culture, in a nutrient medium. Especially if the polypeptide is secreted, the medium will thus contain the polypeptide, with the cells, or without the cells if they have been filtered or centrifuged away.

Many expression systems are known, including bacteria (for example *E. coli* and *Bacillus subtilis*), yeasts (for example *Saccharomyces cerevisiae*,
30

Kluyveromyces lactis and *Pichia pastoris*, filamentous fungi (for example *Aspergillus*), plant cells, animal cells and insect cells.

The desired protein is produced in conventional ways, for example from
5 a coding sequence inserted in the host chromosome or on a free plasmid.

The yeasts are transformed with a coding sequence for the desired protein in any of the usual ways, for example electroporation. Methods for transformation of yeast by electroporation are disclosed in Becker &
10 Guarente (1990) *Methods Enzymol.* **194**, 182.

Successfully transformed cells, ie cells that contain a DNA construct of the present invention, can be identified by well known techniques. For example, cells resulting from the introduction of an expression construct
15 can be grown to produce the desired polypeptide. Cells can be harvested and lysed and their DNA content examined for the presence of the DNA using a method such as that described by Southern (1975) *J. Mol. Biol.* **98**, 503 or Berent *et al* (1985) *Biotech.* **3**, 208. Alternatively, the presence of the protein in the supernatant can be detected using antibodies.

20

Useful yeast plasmid vectors include pRS403-406 and pRS413-416 and are generally available from Stratagene Cloning Systems, La Jolla, CA 92037, USA. Plasmids pRS403, pRS404, pRS405 and pRS406 are Yeast Integrating plasmids (YIps) and incorporate the yeast selectable markers
25 *HIS3*, *TRP1*, *LEU2* and *URA3*. Plasmids pRS413-416 are Yeast Centromere plasmids (YCps).

A variety of methods have been developed to operably link DNA to vectors via complementary cohesive termini. For instance,
30 complementary homopolymer tracts can be added to the DNA segment to

be inserted to the vector DNA. The vector and DNA segment are then joined by hydrogen bonding between the complementary homopolymeric tails to form recombinant DNA molecules.

- 5 Synthetic linkers containing one or more restriction sites provide an alternative method of joining the DNA segment to vectors. The DNA segment, generated by endonuclease restriction digestion, is treated with bacteriophage T4 DNA polymerase or *E. coli* DNA polymerase I, enzymes that remove protruding, 3'-single-stranded termini with their 3'-
10 5'-exonucleolytic activities, and fill in recessed 3'-ends with their polymerizing activities.

The combination of these activities therefore generates blunt-ended DNA segments. The blunt-ended segments are then incubated with a large
15 molar excess of linker molecules in the presence of an enzyme that is able to catalyze the ligation of blunt-ended DNA molecules, such as bacteriophage T4 DNA ligase. Thus, the products of the reaction are DNA segments carrying polymeric linker sequences at their ends. These DNA segments are then cleaved with the appropriate restriction enzyme
20 and ligated to an expression vector that has been cleaved with an enzyme that produces termini compatible with those of the DNA segment.

Synthetic linkers containing a variety of restriction endonuclease sites are commercially available from a number of sources including International
25 Biotechnologies Inc, New Haven, CN, USA.

A desirable way to modify the DNA in accordance with the invention, if, for example, HA variants are to be prepared, is to use the polymerase chain reaction as disclosed by Saiki *et al* (1988) *Science* 239, 487-491.
30 In this method the DNA to be enzymatically amplified is flanked by two

specific oligonucleotide primers which themselves become incorporated into the amplified DNA. The said specific primers may contain restriction endonuclease recognition sites which can be used for cloning into expression vectors using methods known in the art.

5

Exemplary genera of yeast contemplated to be useful in the practice of the present invention as hosts for expressing the fusion proteins are *Pichia* (*Hansenula*), *Saccharomyces*, *Kluyveromyces*, *Candida*, *Torulopsis*, *Torulaspora*, *Schizosaccharomyces*, *Citeromyces*, *Pachysolen*,
10 *Debaromyces*, *Metschnikowia*, *Rhodospiridium*, *Leucosporidium*, *Botryoascus*, *Sporidiobolus*, *Endomycopsis*, and the like. Preferred genera are those selected from the group consisting of *Saccharomyces*, *Schizosaccharomyces*, *Kluyveromyces*, *Pichia* and *Torulaspora*. Examples of *Saccharomyces* spp. are *S. cerevisiae*, *S. italicus* and *S. rouxii*.
15 Examples of *Kluyveromyces* spp. are *K. fragilis*, *K. lactis* and *K. marxianus*. A suitable *Torulaspora* species is *T. delbrueckii*. Examples of *Pichia* (*Hansenula*) spp. are *P. angusta* (formerly *H. polymorpha*), *P. anomala* (formerly *H. anomala*) and *P. pastoris*.

20 Methods for the transformation of *S. cerevisiae* are taught generally in EP 251 744, EP 258 067 and WO 90/01063.

Suitable promoters for *S. cerevisiae* include those associated with the
25 *PGK1* gene, *GAL1* or *GAL10* genes, *CYC1*, *PHO5*, *TRP1*, *ADH1*, *ADH2*, the genes for glyceraldehyde-3-phosphate dehydrogenase, hexokinase, pyruvate decarboxylase, phosphofructokinase, triose phosphate isomerase, phosphoglucose isomerase, glucokinase, α -mating factor pheromone, a-mating factor pheromone, the *PRB1* promoter, the *GUT2* promoter, the
30 *GPD1* promoter, and hybrid promoters involving hybrids of parts of 5'

regulatory regions with parts of 5' regulatory regions of other promoters or with upstream activation sites (eg the promoter of EP-A-258 067).

Convenient regulatable promoters for use in *Schizosaccharomyces pombe* are the thiamine-repressible promoter from the *nmt* gene as described by Maundrell (1990) *J. Biol. Chem.* 265, 10857-10864 and the glucose-repressible *fbp1* gene promoter as described by Hoffman & Winston (1990) *Genetics* 124, 807-816.

- 10 Methods of transforming *Pichia* for expression of foreign genes are taught in, for example, Cregg *et al* (1993), and various Phillips patents (eg US 4 857 467), and *Pichia* expression kits are commercially available from Invitrogen BV, Leek, Netherlands, and Invitrogen Corp., San Diego, California. Suitable promoters include
- 15 *AOX1* and *AOX2*.

- Gleeson *et al* (1986) *J. Gen. Microbiol.* 132, 3459-3465 include information on *Hansenula* vectors and transformation, suitable promoters being *MOX1* and *FMD1*; whilst EP 361 991, Fleer *et al* (1991) and other
- 20 publications from Rhône-Poulenc Rorer teach how to express foreign proteins in *Kluyveromyces* spp., a suitable promoter being *PGK1*.

- The transcription termination signal is preferably the 3' flanking sequence of a eukaryotic gene which contains proper signals for transcription
- 25 termination and polyadenylation. Suitable 3' flanking sequences may, for example, be those of the gene naturally linked to the expression control sequence used, ie may correspond to the promoter. Alternatively, they may be different in which case the termination signal of the *S. cerevisiae* *ADHI* gene is preferred.

The desired fusion protein may be initially expressed with a secretion leader sequence, which may be any leader effective in the yeast chosen. Leaders useful in *S. cerevisiae* include that from the mating factor α polypeptide (MF α -1) and the hybrid leaders of EP-A-387 319. Such
5 leaders (or signals) are cleaved by the yeast before the mature albumin is released into the surrounding medium. Further such leaders include those of *S. cerevisiae* invertase (*SUC2*) disclosed in JP 62-096086 (granted as 91/036516), acid phosphatase (*PHO5*), the pre-sequence of MF α -1, β -glucanase (*BGL2*) and killer toxin; *S. diastaticus* glucoamylase II; *S.*
10 *carlsbergensis* α -galactosidase (*MEL1*); *K. lactis* killer toxin; and *Candida* glucoamylase.

The fusion protein of the invention or a formulation thereof may be administered by any conventional method including parenteral (eg
15 subcutaneous or intramuscular) injection or intravenous infusion. The treatment may consist of a single dose or a plurality of doses over a period of time.

Whilst it is possible for a fusion protein of the invention to be
20 administered alone, it is preferable to present it as a pharmaceutical formulation, together with one or more acceptable carriers. The carrier(s) must be "acceptable" in the sense of being compatible with the fusion protein and not deleterious to the recipients thereof. Typically, the carriers will be water or saline which will be sterile and pyrogen free.
25 The formulation should be non-immunogenic; vaccine-type formulations involving adjuvants are not contemplated.

The formulations may conveniently be presented in unit dosage form and may be prepared by any of the methods well known in the art of
30 pharmacy. Such methods include the step of bringing into association the

fusion protein with the carrier which constitutes one or more accessory ingredients. In general the formulations are prepared by uniformly and intimately bringing into association the active ingredient with liquid carriers or finely divided solid carriers or both, and then, if necessary,
5 shaping the product.

Formulations suitable for parenteral administration include aqueous and non-aqueous sterile injection solutions which may contain anti-oxidants, buffers, bacteriostats and solutes which render the formulation isotonic
10 with the blood of the intended recipient; and aqueous and non-aqueous sterile suspensions which may include suspending agents and thickening agents. The formulations may be presented in unit-dose or multi-dose containers, for example sealed ampoules and vials, and may be stored in a freeze-dried (lyophilised) condition requiring only the addition of the
15 sterile liquid carrier, for example water for injections, immediately prior to use. Extemporaneous injection solutions and suspensions may be prepared from sterile powders.

Preferred unit dosage formulations are those containing a daily dose or
20 unit, daily sub-dose or an appropriate fraction thereof, of an active ingredient.

The fusion proteins of the invention may be used in the treatment of any condition in which growth hormone is indicated, for example isolated
25 growth hormone deficiency, panhypopituitarism, following cranial irradiation (eg in the treatment of leukaemia or brain tumours), Turner's syndrome, Down's syndrome, intrauterine growth retardation, idiopathic growth deficiency, chronic renal failure, achondroplasia, female infertility and various catabolic disorders. They may also be used in the stimulation
30 of growth, and/or enhancement of lean meat proportion, in farm animals

such as cows, sheep, goats and pigs.

The fusion protein may be administered together with insulin-like growth factor I (IGF-I).

5

The dosage can be calculated on the basis of the potency of the fusion protein relative to the potency of hGH, whilst taking into account the prolonged serum half-life of the fusion proteins compared to that of native hGH. Growth hormone is typically administered at 0.3 to 30.0 IU/kg/week, for example 0.9 to 12.0 IU/kg/week, given in three or seven divided doses for a year or more. In a fusion protein consisting of full length HA fused to full length GH, an equivalent dose in terms of units would represent a greater weight of agent but the dosage frequency can be reduced, for example to twice a week, once a week or less.

15

Preferred examples of the invention will now be described by way of example and with reference to the accompanying figures, in which:

Figure 1 shows the human growth hormone cDNA sequence, encoding mature hGH; (SEQ. ID. NOS. 22 & 23)

20

Figure 2 shows a restriction enzyme map of pHGH1;

Figure 3 shows a restriction enzyme map of pBST(+) and the DNA sequence of the polylinker; (SEQ. ID. NO. 24)

25

Figure 4 shows the construction of pHGH12;

Figure 5 shows the construction of pHGH16;

30

Figure 6 shows the HSA cDNA sequence, more particularly the region encoding the mature protein; (SEQ. ID. NOS. 25 & 26)

Figure 7 shows the construction of pHGH14;

5

Figure 8 shows the construction of pHGH38;

Figure 9 shows the construction of pHGH31;

10 Figure 10 shows the construction of pHGH58 or pHGH59 (Example 7);

Figure 11 is a scheme for constructing fusions having spacers (Example 7); and

15 Figure 12 shows the results of a pharmacokinetic study showing the clearance of ^{125}I -labelled rHA-hGH compared to that of hGH following iv injection in rats. Data are from two rats in each group, and include total radioactivity and radioactivity which could be precipitated by TCA, ie that associated with protein rather than as free ^{125}I . The calculated clearance
20 half-life for hGH was approximately 6 minutes, compared to approximately 60 minutes for the rHA-hGH fusion protein. See Example 3.

- - hGH (total counts)
- 25 ■ - hGH (TCA precipitated counts)
- ▼ - rHA-hGH (total counts)
- ▲ - rHA-hGH (TCA precipitated counts).

Detailed description of the invention

All standard recombinant DNA procedures are as described in Sambrook
et al (1989) unless otherwise stated. The DNA sequences encoding HSA
5 were derived from the cDNA disclosed in EP 201 239.

Example 1: Cloning of the hGH cDNA.

The hGH cDNA was obtained from a human pituitary gland cDNA library
10 (catalogue number HL1097v, Clontech Laboratories, Inc) by PCR
amplification. Two oligonucleotides suitable for PCR amplification of the
hGH cDNA, HGH1 and HGH2, were synthesised using an Applied
Biosystems 380B Oligonucleotide Synthesiser.

15 HGH1: 5' - CCCAAGAATTCCCTTATCCAGGC - 3' (SEQ. ID. NO. 1)

HGH2: 5' - GGGAAGCTTAGAAGCCACAGGATCCCTCCACAG - 3' (SEQ. ID. NO. 2)

HGH1 and HGH2 differed from the equivalent portion of the hGH cDNA
sequence (Figure 1, Martial *et al*, 1979) by two and three nucleotides,
20 respectively, such that after PCR amplification an *EcoRI* site would be
introduced to the 5' end of the cDNA and a *BamHI* site would be
introduced into the 3' end of the cDNA. In addition, HGH2 contained a
HindIII site immediately downstream of the hGH sequence.

25 PCR amplification using a Perkin-Elmer-Cetus Thermal Cycler 9600 and
a Perkin-Elmer-Cetus PCR kit, was performed using single-stranded DNA
template isolated from the phage particles of the cDNA library as follows:
10 μ L phage particles were lysed by the addition of 10 μ L phage lysis
buffer (280 μ g/mL proteinase K in TE buffer) and incubation at 55°C for
30 15 min followed by 85°C for 15 min. After a 1 min incubation on ice,

phage debris was pelleted by centrifugation at 14,000 rpm for 3 min. The PCR mixture contained 6 μ L of this DNA template, 0.1 μ M of each primer and 200 μ M of each deoxyribonucleotide. PCR was carried out for 30 cycles, denaturing at 94°C for 30 s, annealing at 65°C for 30 s and
5 extending at 72°C for 30 s, increasing the extension time by 1 s per cycle. Analysis of the reaction by gel electrophoresis showed a single product of the expected size (589 base pairs).

The PCR product was purified using Wizard PCR Preps DNA Purification
10 System (Promega Corp) and then digested with *Eco*RI and *Hind*III. After further purification of the *Eco*RI-*Hind*III fragment by gel electrophoresis, the product was cloned into pUC19 (GIBCO BRL) digested with *Eco*RI and *Hind*III, to give pHGH1 (Figure 2). DNA sequencing of the *Eco*RI-*Hind*III region showed that the PCR product was identical in sequence to
15 the hGH sequence (Martial *et al*, 1979), except at the 5' and 3' ends, where the *Eco*RI and *Bam*HI sites had been introduced, respectively.

Example 2: Expression of the hGH cDNA.

20 The polylinker sequence of the phagemid pBluescribe (+) (Stratagene) was replaced by inserting an oligonucleotide linker, formed by annealing two 75-mer oligonucleotides, between the *Eco*RI and *Hind*III sites to form pBST(+) (Figure 3). The new polylinker included a unique *Not*I site (the full sequence in the region of the polylinker is given in Figure 3).

25

The *Not*I HSA expression cassette of pAYE309 (EP 431 880) comprising the *PRB*I promoter, DNA encoding the HSA/MF α -1 hybrid leader sequence, DNA encoding HSA and the *ADH*I terminator, was transferred to pBST(+) to form pHAl (Figure 4). The HSA coding sequence was
30 removed from this plasmid by digestion with *Hind*III followed by

religation to form pHA2 (Figure 4).

Cloning of the hGH cDNA, as described in Example 1, provided the hGH coding region lacking the pro-hGH sequence and the first 8 base pairs (bp) of the mature hGH sequence. In order to construct an expression plasmid for secretion of hGH from yeast, a yeast promoter, signal peptide and the first 8 bp of the hGH sequence were attached to the 5' end of the cloned hGH sequence as follows:

10 The *HindIII*-*Sfa*NI fragment from pHA1 was attached to the 5' end of the *Eco*RI-*HindIII* fragment from pHGH1 via two synthetic oligonucleotides, HGH3 and HGH4:

HGH3: 5' - GATAAAGATTCCCAAC - 3' (SEQ. ID. NO. 3)
 15 HGH4: 5' - AATTGTTGGGAATCTTT - 3' (SEQ. ID. NO. 4)

The *HindIII* fragment so formed was cloned into *HindIII*-digested pHA2 to make pHGH2 (Figure 4), such that the hGH cDNA was positioned downstream of the *PRB1* promoter and HSA/MF α -1 fusion leader sequence (WO 90/01063). The *NotI* expression cassette contained in pHGH2, which included the *ADHI* terminator downstream of the hGH cDNA, was cloned into *NotI*-digested pSAC35 (Sleep *et al*, 1990) to make pHGH12 (Figure 4). This plasmid comprised the entire 2 μ m plasmid to provide replication functions and the *LEU2* gene for selection of transformants.

pHGH12 was introduced into *S. cerevisiae* DB1 (Sleep *et al*, 1990) by transformation and individual transformants were grown for 3 days at 30°C in 10 mL YEPD (1% w/v yeast extract, 2% w/v peptone, 2% w/v dextrose). After centrifugation of the cells, the supernatants were

examined by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) and were found to contain protein which was of the expected size and which was recognised by anti-hGH antiserum (Sigma, Poole, UK) on Western blots.

5

Example 3: Cloning and expression of an HSA-hGH fusion protein.

In order to fuse the HSA cDNA to the 5' end of the hGH cDNA, the pHA1 *Hind*III-*Bsu*36I fragment (containing most of the HSA cDNA) was
10 joined to the pHGH1 *Eco*RI-*Hind*III fragment (containing most of the hGH cDNA) via two oligonucleotides, HGH7 and HGH8:

HGH7: 5' - TTAGGCTTATTCCCAAC - 3' (SEQ. ID. NO. 5)

HGH8: 5' - AATTGTTGGGAATAAGCC - 3' (SEQ. ID. NO. 6)

15

The *Hind*III fragment so formed was cloned into pHA2 digested with *Hind*III to make pHGH10 (Figure 5), and the *Not*I expression cassette of this plasmid was cloned into *Not*I-digested pSAC35 to make pHGH16 (Figure 5).

20

pHGH16 was used to transform *S. cerevisiae* DB1 and supernatants of cultures were analysed as in Example 2. A predominant band was observed that had a molecular weight of approximately 88 kD, corresponding to the combined masses of HA and hGH. Western blotting
25 using anti-HSA and anti-hGH antisera (Sigma) confirmed the presence of the two constituent parts of the fusion protein.

The fusion protein was purified from culture supernatant by cation exchange chromatography, followed by anion exchange and gel
30 permeation chromatography. Analysis of the N-terminus of the protein by

Example 4: Cloning and expression of an hGH-HSA fusion protein.

In order to fuse the hGH cDNA to the 5' end of the HSA cDNA (Figure 6), the HSA cDNA was first altered by site-directed mutagenesis to introduce an *Eco*NI site near the 5' end of the coding region. This was done by the method of Kunkel *et al* (1987) using single-stranded DNA template prepared from pHA1 and a synthetic oligonucleotide, LEU4:

LEU4 : 5' - GAGATGCACACCTGAGTGAGG - 3' (SEQ. ID. NO. 8)

10

Site-directed mutagenesis using this oligonucleotide changed the coding sequence of the HSA cDNA from Lys4 to Leu4 (K4L). However, this change was repaired when the hGH cDNA was subsequently joined at the 5' end by linking the pHGH2 *Not*I-*Bam*HI fragment to the *Eco*NI-*Not*I fragment of the mutated pHA1, via the two oligonucleotides HGH5 and HGH6:

15

HGH5 : 5' - GATCCTGTGGCTTCGATGCACACAAGA - 3' (SEQ. ID. NO. 9)

HGH6 : 5' - CTCTTGTGTGCATCGAAGCCACAG - 3' (SEQ. ID. NO. 10)

20

The *Not*I fragment so formed was cloned into *Not*I-digested pSAC35 to make pHGH14 (Figure 7). pHGH14 was used to transform *S. cerevisiae* DB1 and supernatants of cultures were analysed as in Example 2. A predominant band was observed that had a molecular weight of approximately 88 kD, corresponding to the combined masses of hGH and HA. Western blotting using anti-HSA and anti-hGH antisera confirmed the presence of the two constituent parts of the fusion protein.

25

The fusion protein was purified from culture supernatant by cation exchange chromatography, followed by anion exchange and gel permeation chromatography. Analysis of the N-terminus of the protein by

30

amino acid sequencing confirmed the presence of the expected hGH sequence.

In vitro studies showed that the fusion protein retained hGH activity, but was significantly less potent than a fusion protein comprising full-length HA (1-585) as the N-terminal portion and hGH as the C-terminal portion, as described in Example 3.

Example 5: Construction of plasmids for the expression of hGH fusions to domains of HSA.

Fusion polypeptides were made in which the hGH molecule was fused to the first two domains of HA (residues 1 to 387). Fusion to the N-terminus of hGH was achieved by joining the pHA1 *HindIII*-*SapI* fragment, which contained most of the coding sequence for domains 1 and 2 of HA, to the pHGH1 *EcoRI*-*HindIII* fragment, via the oligonucleotides HGH11 and HGH12:

HGH11: 5' - TGTGGAAGAGCCTCAGAATTTATTCCCAAC - 3' (SEQ. ID. NO. 11)
HGH12: 5' - AATTGTTGGGAATAAATTCTGAGGCTCTTCC - 3' (SEQ. ID. NO. 12)

The *HindIII* fragment so formed was cloned into *HindIII*-digested pHA2 to make pHGH37 (Fig 8) and the *NotI* expression cassette of this plasmid was cloned into *NotI*-digested pSAC35. The resulting plasmid, pHGH38 (Fig 8), contained an expression cassette that was found to direct secretion of the fusion polypeptide into the supernatant when transformed into *S. cerevisiae* DB1. Western blotting using anti-HSA and anti-hGH antisera confirmed the presence of the two constituent parts of the fusion protein.

The fusion protein was purified from culture supernatant by cation

exchange chromatography followed by gel permeation chromatography. *In vivo* studies with purified protein indicated that the circulatory half-life was longer than that of hGH, and similar to that of a fusion protein comprising full-length HA (1-585) as the N-terminal portion and hGH as the C-terminal portion, as described in Example 3. *In vitro* studies showed that the fusion protein retained hGH activity.

Using a similar strategy as detailed above, a fusion protein comprising the first domain of HA (residues 1-194) as the N-terminal portion and hGH as the C-terminal portion, was cloned and expressed in *S. cerevisiae* DB1. Western blotting of culture supernatant using anti-HSA and anti-hGH antisera confirmed the presence of the two constituent parts of the fusion protein.

15 Example 6: Expression of hGH by introducing a cleavage site between HSA and hGH.

Introduction of a peptide sequence that is recognised by the Kex2 protease, between the HA-hGH fusion protein, allows secretion of hGH. A sequence encoding Ser Leu Asp Lys Arg (SEQ. ID. NO. 13) was introduced using two oligonucleotides, HGH14 and HGH15:

HGH14: 5' - TTAGGCTTAAGCTTGGATAAAAGATTCCCAAC - 3' (SEQ. ID. NO. 14)

HGH15: 5' - AATTGTTGGGAATCTTTTATCCAAGCTTAAGCC - 3' (SEQ. ID. NO. 15)

25 These were used to join the pHA1 *Hind*III-*Bsu*36I fragment to the pHGH1 *Eco*RI-*Hind*III fragment, which were then cloned into *Hind*III-digested pHA2 to make pHGH25 (Fig 9). The *Nor*I expression cassette of this plasmid was cloned into *Nor*I-digested pSAC35 to make pHGH31 (Figure 30 9).

S. cerevisiae DB1 transformed with pHGH31 was found to secrete two major species, as determined by SDS-PAGE analysis of culture supernatants. The two species had molecular weights of approximately 66 kD, corresponding to (full length) HA, and 22 kD, corresponding to (full length) hGH, indicating *in vivo* cleavage of the fusion protein by the Kex2 protease, or an equivalent activity. Western blotting using anti-HSA and anti-hGH antisera confirmed the presence of the two separate species. N-terminal sequence analysis of the hGH moiety confirmed the precise and efficient cleavage from the HA moiety.

10

The hGH moiety was purified from culture supernatant by anion exchange chromatography followed by gel permeation chromatography. *In vitro* studies with the purified hGH showed that the protein was active and fully potent.

15

Using a similar strategy, fusion proteins comprising either the first domain of HA (residues 1-194) or the first two domains of HA (residues 1-387), followed by a sequence recognised by the Kex2p protease, followed by the hGH cDNA, were cloned and expressed in *S. cerevisiae* DB1. Western blotting of culture supernatant using anti-HSA and anti-hGH antisera confirmed the presence of the two separate species.

20

Example 7: Fusion of HSA to hGH using a flexible linker sequence

Flexible linkers, comprising repeating units of [Gly-Gly-Gly-Gly-Ser]_n (SEQ. ID. NOS. 16 & 17) where n was either 2 or 3, were introduced between the HSA and hGH fusion protein by cloning of the oligonucleotides HGH16, HGH17, HGH18 and HGH19.

25

HGH16: 5' - TTAGGCTTAGGTGGCGGTGGATCCGGCGGTGGTGGATCTTTCCCA
AC-3' (SEQ. ID. NO. 18)

HGH17: 5' - AATTGTTGGGAAAGATCCACCACCGCCGGATCCACCGCCACCTAA
GCC-3' (SEQ. ID. NO. 19)

5 HGH18: 5' - TTAGGCTTAGGCGGTGGTGGATCTGGTGGCGCGGATCTGGTGGC
GGTGGATCCTTCCCAAC-3' (SEQ. ID. NO. 20)

HGH19: 5' - AATTGTTGGGAAGGATCCACCACCGCCAGATCCGCCGCCACCA
GATCCACCACCGCCTAAGCC-3' (SEQ. ID. NO. 21)

- 10 Annealing of HGH16 with HGH17 resulted in $n=2$, while HGH18
annealed to HGH19 resulted in $n=3$. After annealing, the double-stranded
oligonucleotides were cloned with the *EcoRI-Bsu36I* fragment isolated
from pHGH1 into *Bsu36I*-digested pHGH10 to make pHGH56 (where
 $n=2$) and pHGH57 (where $n=3$) (Figure 10). The *NorI* expression
15 cassettes from these plasmids were cloned into *NorI*-digested pSAC35 to
make pHGH58 and pHGH59, respectively.

- Cloning of the oligonucleotides to make pHGH56 and pHGH57 introduced
a *BamHI* site in the linker sequences, as shown in Figure 11. It was
20 therefore possible to construct linker sequences in which $n=1$ and $n=4$,
by joining either the *HindIII-BamHI* fragment from pHGH56 to the
BamHI-HindIII fragment from pHGH57 (making $n=1$), or the *HindIII-
BamHI* fragment from pHGH57 to the *BamHI-HindIII* fragment from
pHGH56 (making $n=2$). Cloning of these fragments into the *HindIII* site
25 of pHA2 (described in Example 2), resulted in pHGH60 ($n=1$) and
pHGH61 ($n=4$) (see Figure 11). The *NorI* expression cassettes from
pHGH60 and pHGH61 were cloned into *NorI*-digested pSAC35 to make
pHGH62 and pHGH63, respectively.

- 30 Transformation of *S.cerevisiae* with pHGH58, pHGH59, pHGH62 and
pHGH63 resulted in transformants that secreted the fusion polypeptides

into the supernatant.

Western blotting using anti-HSA and anti-hGH antisera confirmed the presence of the two constituent parts of the fusion proteins.

5

The fusion proteins were purified from culture supernatant by cation exchange chromatography, followed by anion exchange and gel permeation chromatography. Analysis of the N-termini of the proteins by amino acid sequencing confirmed the presence of the expected albumin
10 sequence. Analysis of the purified proteins by electrospray mass spectrometry confirmed an increase in mass of 315 D (n=1), 630 D (n=2), 945 D (n=3) and 1260 D (n=4) compared to the HSA-hGH fusion protein described in Example 3, as expected. The purified protein was found to be active *in vitro*.

References

- Cunningham, B. C. *et al* (1991) *Science* 254, 821-825.
- de Vos, A. M. *et al* (1992) *Science* 255, 306-312.
- 5 Ealey *et al* (1995) *Growth Regulation* 5, 36-44.
- Gleeson *et al* (1986) *J. Gen. Microbiol.* 132, 3459-3465.
- Haffner, D. *et al*, (1994) *J. Clin. Invest.* 93, 1163-1171.
- Hiramitsu *et al* (1990) *App. Env. Microbiol.* 56, 2125-2132.
- Hiramitsu *et al* (1991) *ibid* 57, 2052-2056.
- 10 Hoffman and Winston (1990) *Genetics* 124, 807-816.
- Kearns, G. L. *et al* (1991) *J. Clin. Endocrinol. Metab.* 72, 1148-1156.
- Kunkel, T. A. *et al* (1987) *Methods in Enzymol.* 154, 367-382.
- Martial, J. A. *et al* (1979) *Science* 205, 602-607.
- Maundrell (1990) *J. Biol. Chem.* 265, 10857-10864.
- 15 Nomura, N. *et al* (1995) *Biosci. Biotech. Biochem.* 59, 532-534.
- Poznansky, M. J. *et al* (1988) *FEBS Lett.* 239, 18-22.
- Saiki *et al* (1988) *Science* 239, 487-491.
- Sambrook, J. *et al* (1989) *Molecular Cloning: a Laboratory Manual*, 2nd
 edition. Cold Spring Harbor Laboratory Press, Cold Spring
 20 Harbor, NY.
- Sleep, D. *et al* (1990) *Bio/Technology* 8, 42-46.
- Strobl, J. S. and Thomas, M. J. (1994) *Pharmacol. Rev.* 46, 1-34.
- Tokunaga, T. *et al* (1985) *Gene* 39, 117-120.
- Tsiomenko, A. B. *et al* (1994) *Biochemistry (Moscow)* 59, 1247-1256.
- 25 Zeisel, H. J. *et al* (1992) *Horm. Res.* 37 (Suppl. 2), 5-13.

28

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT:
 - (A) NAME: Delta Biotechnology Limited
 - (B) STREET: Castle Court, Castle Boulevard
 - (C) CITY: Nottingham
 - (E) COUNTRY: UK
 - (F) POSTAL CODE (ZIP): NG7 1FD
- (ii) TITLE OF INVENTION: RECOMBINANT FUSION PROTEINS TO GROWTH HORMONE AND SERUM ALBUMIN
- (iii) NUMBER OF SEQUENCES: 26
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: BERESKIN & PARR
 - (B) STREET: 40 King Street West
 - (C) CITY: Toronto
 - (D) STATE: Ontario
 - (E) COUNTRY: Canada
 - (F) ZIP: M5H 3Y2
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER: 2,240,292
 - (B) FILING DATE: 19-DEC-1996
 - (C) CLASSIFICATION:
- (vii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Rudolph, John R.
 - (B) REGISTRATION NUMBER: 38,003
 - (C) REFERENCE/DOCKET NUMBER: 3167-30
- (viii) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: (416) 364-7311
 - (B) TELEFAX: (416) 361-1398

(2) INFORMATION FOR SEQ ID NO: 1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 23 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid
 - (A) DESCRIPTION: /desc = "PCR PRIMER"
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO

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

CCCAAGAATT CCCTTATCCA GGC

23

29

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "PCR PRIMER"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

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

GGGAAGCTTA GAAGCCACAG GATCCCTCCA CAG

33

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 16 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "OLIGONUCLEOTIDE LINKER"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

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

GATAAAGATT CCCAAC

16

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "OLIGONUCLEOTIDE LINKER"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

AATTGTTGGG AATCTTT

17

(2) INFORMATION FOR SEQ ID NO: 5:

30

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 17 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid
 - (A) DESCRIPTION: /desc = "OLIGONUCLEOTIDE LINKER"
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

TTAGGCTTAT TCCCAAC

17

(2) INFORMATION FOR SEQ ID NO: 6:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 18 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid
 - (A) DESCRIPTION: /desc = "OLIGONUCLEOTIDE LINKER"
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

AATTGTTGGG AATAAGCC

18

(2) INFORMATION FOR SEQ ID NO: 7:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 604 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

Met	Leu	Leu	Gln	Ala	Phe	Leu	Phe	Leu	Leu	Ala	Gly	Phe	Ala	Ala	Lys
1				5					10					15	
Ile	Ser	Ala	Asp	Ala	His	Lys	Ser	Glu	Val	Ala	His	Arg	Phe	Lys	Asp
			20					25					30		

31

Leu	Gly	Glu	Glu	Asn	Phe	Lys	Ala	Leu	Val	Leu	Ile	Ala	Phe	Ala	Gln
		35					40					45			
Tyr	Leu	Gln	Gln	Cys	Pro	Phe	Glu	Asp	His	Val	Lys	Leu	Val	Asn	Glu
	50					55					60				
Val	Thr	Glu	Phe	Ala	Lys	Thr	Cys	Val	Ala	Asp	Glu	Ser	Ala	Glu	Asn
65					70					75					80
Cys	Asp	Lys	Ser	Leu	His	Thr	Leu	Phe	Gly	Asp	Lys	Leu	Cys	Thr	Val
				85					90					95	
Ala	Thr	Leu	Arg	Glu	Thr	Tyr	Gly	Glu	Met	Ala	Asp	Cys	Cys	Ala	Lys
			100					105					110		
Gln	Glu	Pro	Glu	Arg	Asn	Glu	Cys	Phe	Leu	Gln	His	Lys	Asp	Asp	Asn
		115					120					125			
Pro	Asn	Leu	Pro	Arg	Leu	Val	Arg	Pro	Glu	Val	Asp	Val	Met	Cys	Thr
	130					135					140				
Ala	Phe	His	Asp	Asn	Glu	Glu	Thr	Phe	Leu	Lys	Lys	Tyr	Leu	Tyr	Glu
145					150					155					160
Ile	Ala	Arg	Arg	His	Pro	Tyr	Phe	Tyr	Ala	Pro	Glu	Leu	Leu	Phe	Phe
				165					170					175	
Ala	Lys	Arg	Tyr	Lys	Ala	Ala	Phe	Thr	Glu	Cys	Cys	Gln	Ala	Ala	Asp
			180					185					190		
Lys	Ala	Ala	Cys	Leu	Leu	Pro	Lys	Leu	Asp	Glu	Leu	Arg	Asp	Glu	Gly
		195					200					205			
Lys	Ala	Ser	Ser	Ala	Lys	Gln	Arg	Leu	Lys	Cys	Ala	Ser	Leu	Gln	Lys
	210					215					220				
Phe	Gly	Glu	Arg	Ala	Phe	Lys	Ala	Trp	Ala	Val	Ala	Arg	Leu	Ser	Gln
225					230					235					240
Arg	Phe	Pro	Lys	Ala	Glu	Phe	Ala	Glu	Val	Ser	Lys	Leu	Val	Thr	Asp
				245					250					255	
Leu	Thr	Lys	Val	His	Thr	Glu	Cys	Cys	His	Gly	Asp	Leu	Leu	Glu	Cys
			260					265					270		
Ala	Asp	Asp	Arg	Ala	Asp	Leu	Ala	Lys	Tyr	Ile	Cys	Glu	Asn	Gln	Asp
		275					280					285			
Ser	Ile	Ser	Ser	Lys	Leu	Lys	Glu	Cys	Cys	Glu	Lys	Pro	Leu	Leu	Glu
	290					295					300				
Lys	Ser	His	Cys	Ile	Ala	Glu	Val	Glu	Asn	Asp	Glu	Met	Pro	Ala	Asp
305					310					315					320
Leu	Pro	Ser	Leu	Ala	Ala	Asp	Phe	Val	Glu	Ser	Lys	Asp	Val	Cys	Lys
				325					330					335	
Asn	Tyr	Ala	Glu	Ala	Lys	Asp	Val	Phe	Leu	Gly	Met	Phe	Leu	Tyr	Glu
			340					345					350		
Tyr	Ala	Arg	Arg	His	Pro	Asp	Tyr	Ser	Val	Val	Leu	Leu	Leu	Arg	Leu
		355					360					365			
Ala	Lys	Thr	Tyr	Glu	Thr	Thr	Leu	Glu	Lys	Cys	Cys	Ala	Ala	Ala	Asp
	370					375					380				
Pro	His	Glu	Cys	Tyr	Ala	Lys	Val	Phe	Asp	Glu	Phe	Lys	Pro	Leu	Val

32

385		390		395		400									
Glu	Glu	Pro	Gln	Asn	Leu	Ile	Lys	Gln	Asn	Cys	Glu	Leu	Phe	Glu	Gln
				405					410					415	
Leu	Gly	Glu	Tyr	Lys	Phe	Gln	Asn	Ala	Leu	Leu	Val	Arg	Tyr	Thr	Lys
			420					425					430		
Lys	Val	Pro	Gln	Val	Ser	Thr	Pro	Thr	Leu	Val	Glu	Val	Ser	Arg	Asn
		435					440					445			
Leu	Gly	Lys	Val	Gly	Ser	Lys	Cys	Cys	Lys	His	Pro	Glu	Ala	Lys	Arg
	450					455					460				
Met	Pro	Cys	Ala	Glu	Asp	Tyr	Leu	Ser	Val	Val	Leu	Asn	Gln	Leu	Cys
465					470					475					480
Val	Leu	His	Glu	Lys	Thr	Pro	Val	Ser	Asp	Arg	Val	Thr	Lys	Cys	Cys
				485					490					495	
Thr	Glu	Ser	Leu	Val	Asn	Arg	Arg	Pro	Cys	Phe	Ser	Ala	Leu	Glu	Val
			500					505					510		
Asp	Glu	Thr	Tyr	Val	Pro	Lys	Glu	Phe	Asn	Ala	Glu	Thr	Phe	Thr	Phe
		515					520					525			
His	Ala	Asp	Ile	Cys	Thr	Leu	Ser	Glu	Lys	Glu	Arg	Gln	Ile	Lys	Lys
	530					535					540				
Gln	Thr	Ala	Leu	Val	Glu	Leu	Val	Lys	His	Lys	Pro	Lys	Ala	Thr	Lys
545					550					555					560
Glu	Gln	Leu	Lys	Ala	Val	Met	Asp	Asp	Phe	Ala	Ala	Phe	Val	Glu	Lys
				565					570					575	
Cys	Cys	Lys	Ala	Asp	Asp	Lys	Glu	Thr	Cys	Phe	Ala	Glu	Glu	Gly	Lys
			580					585					590		
Lys	Leu	Val	Ala	Ala	Ser	Gln	Ala	Ala	Leu	Gly	Leu				
		595					600								

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "OLIGONUCLEOTIDE FOR USE IN
DNA TEMPLATE PREPARATION"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

GAGATGCACA CCTGAGTGAG G

21

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:

33

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "OLIGONUCLEOTIDE LINKER"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

GATCCTGTGG CTTGATGCA CACAAGA

27

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "OLIGONUCLEOTIDE LINKER"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

CTCTTGTGTG CATCGAAGCC ACAG

24

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "OLIGONUCLEOTIDE LINKER"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

TGTGGAAGAG CCTCAGAATT TATTCCTAAC

30

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 base pairs

34

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
 - (A) DESCRIPTION: /desc = "OLIGONUCLEOTIDE LINKER"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

AATTGTTGGG AATAAATTCT GAGGCTCTTC C

31

- (2) INFORMATION FOR SEQ ID NO: 13:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 5 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

Ser Leu Asp Lys Arg
1 5

- (2) INFORMATION FOR SEQ ID NO: 14:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 32 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "OLIGONUCLEOTIDE LINKER"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

TTAGGCTTAA GCTTGGATAA AAGATTCCCA AC

32

- (2) INFORMATION FOR SEQ ID NO: 15:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 33 base pairs
 - (B) TYPE: nucleic acid

35

- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid
 - (A) DESCRIPTION: /desc = "DNA OLIGONUCLEOTIDE"
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

AATTGTTGGG AATCTTTTAT CCAAGCTTAA GCC

33

(2) INFORMATION FOR SEQ ID NO: 16:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 10 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

Gly	Gly	Gly	Gly	Ser	Gly	Gly	Gly	Gly	Ser
1				5					10

(2) INFORMATION FOR SEQ ID NO: 17:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 15 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

Gly	Gly	Gly	Gly	Ser	Gly	Gly	Gly	Gly	Ser	Gly	Gly	Gly	Gly	Ser
1				5					10					15

(2) INFORMATION FOR SEQ ID NO: 18:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 47 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single

36

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "OLIGONUCLEOTIDE ENCODING FLEXIBLE LINKER"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

TTAGGCTTAG GTGGCGGTGG ATCCGGCGGT GGTGGATCTT TCCCAAC

47

(2) INFORMATION FOR SEQ ID NO: 19:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 48 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "OLIGONUCLEOTIDE ENCODING FLEXIBLE LINKER"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

AATTGTTGGG AAAGATCCAC CACCGCCGGA TCCACCGCCA CCTAAGCC

48

(2) INFORMATION FOR SEQ ID NO: 20:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 62 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "OLIGONUCLEOTIDE ENCODING FLEXIBLE LINKER"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

TTAGGCTTAG GCGGTGGTGG ATCTGGTGGC GCGGATCTG GTGGCGGTGG ATCCTTCCCA

60

AC

62

(2) INFORMATION FOR SEQ ID NO: 21:

37

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 63 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid
 - (A) DESCRIPTION: /desc = "OLIGONUCLEOTIDE ENCODING FLEXIBLE LINKER"
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

AATTGTTGGG AAGGATCCAC CGCCACCAGA TCCGCCGCCA CCAGATCCAC CACCGCCTAA	60
GCC	63

(2) INFORMATION FOR SEQ ID NO: 22:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 576 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

TTCCCAACCA TTCCCTTATC CAGGCTTTTT GACAACGCTA TGCTCCGCGC CCATCGTCTG	60
CACCAGCTGG CCTTTGACAC CTACCAGGAG TTTGAAGAAG CCTATATCCC AAAGGAACAG	120
AAGTATTCAT TCCTGCAGAA CCCCAGACC TCCCTCTGTT TCTCAGAGTC TATTCCGACA	180
CCCTCCAACA GGGAGGAAAC ACAACAGAAA TCCAACCTAG AGCTGCTCCG CATCTCCCTG	240
CTGCTCATCC AGTCGTGGCT GGAGCCCGTG CAGTCCCTCA GGAGTGTCTT CGCCAACAGC	300
CTGGTGTACG GCGCCTCTGA CAGCAACGTC TATGACCTCC TAAAGGACCT AGAGGAAGGC	360
ATCCAAACGC TGATGGGGAG GCTGGAAGAT GGCAGCCCCC GGACTGGGCA GATCTTCAAG	420
CAGACCTACA GCAAGTTCGA CACAACTCA CACAACGATG ACGCACTACT CAAGAACTAC	480
GGGCTGCTCT ACTGCTTCAG GAAGGACATG GACAAGGTCG AGACATTCCT GCGCATCGTG	540
CAGTGCCGCT CTGTGGAGGG CAGCTGTGGC TTCTAG	576

(2) INFORMATION FOR SEQ ID NO: 23:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 191 amino acids
 - (B) TYPE: amino acid

38

(C) STRANDEDNESS:
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
 (iii) HYPOTHETICAL: NO
 (iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

Phe	Pro	Thr	Ile	Pro	Leu	Ser	Arg	Leu	Phe	Asp	Asn	Ala	Met	Leu	Arg	1	5	10	15
Ala	His	Arg	Leu	His	Gln	Leu	Ala	Phe	Asp	Thr	Tyr	Gln	Glu	Phe	Glu	20	25	30	
Glu	Ala	Tyr	Ile	Pro	Lys	Glu	Gln	Lys	Tyr	Ser	Phe	Leu	Gln	Asn	Pro	35	40	45	
Gln	Thr	Ser	Leu	Cys	Phe	Ser	Glu	Ser	Ile	Pro	Thr	Pro	Ser	Asn	Arg	50	55	60	
Glu	Glu	Thr	Gln	Gln	Lys	Ser	Asn	Leu	Glu	Leu	Leu	Arg	Ile	Ser	Leu	65	70	75	80
Leu	Leu	Ile	Gln	Ser	Trp	Leu	Glu	Pro	Val	Gln	Ser	Leu	Arg	Ser	Val	85	90	95	
Phe	Ala	Asn	Ser	Leu	Val	Tyr	Gly	Ala	Ser	Asp	Ser	Asn	Val	Tyr	Asp	100	105	110	
Leu	Leu	Lys	Asp	Leu	Glu	Glu	Gly	Ile	Gln	Thr	Leu	Met	Gly	Arg	Leu	115	120	125	
Glu	Asp	Gly	Ser	Pro	Arg	Thr	Gly	Gln	Ile	Phe	Lys	Gln	Thr	Tyr	Ser	130	135	140	
Lys	Phe	Asp	Thr	Asn	Ser	His	Asn	Asp	Asp	Ala	Leu	Leu	Lys	Asn	Tyr	145	150	155	160
Gly	Leu	Leu	Tyr	Cys	Phe	Arg	Lys	Asp	Met	Asp	Lys	Val	Glu	Thr	Phe	165	170	175	
Leu	Arg	Ile	Val	Gln	Cys	Arg	Ser	Val	Glu	Gly	Ser	Cys	Gly	Phe	180	185	190		

(2) INFORMATION FOR SEQ ID NO: 24:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 81 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
 (A) DESCRIPTION: /desc = "oligonucleotide linker"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

AAGCTCCTAG GCCCGGGCGG CCGCAAGCTT GTCGACGCTA GCTGCAGAAG GATCCAGATC	60
TCGAGGCGCC ATCGATAATT C	81

(2) INFORMATION FOR SEQ ID NO: 25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1758 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

GATGCACACA AGAGTGAGGT TGCTCATCGG TTAAAGATT TGGGAGAAGA AAATTTCAAA	60
GCCTTGGTGT TGATTGCCTT TGCTCAGTAT CTTCAGCAGT GTCCATTTGA AGATCATGTA	120
AAATTAGTGA ATGAAGTAAC TGAATTTGCA AAAACATGTG TTGCTGATGA GTCAGCTGAA	180
AATTGTGACA AATCACTTCA TACCCTTTTT GGAGACAAAT TATGCACAGT TGCAACTCTT	240
CGTGAAACCT ATGGTGAAAT GGCTGACTGC TGTGCAAAAC AAGAACCTGA GAGAAATGAA	300
TGCTTCTTGC AACACAAAGA TGACAACCCA AACCTCCCCC GATTGGTGAG ACCAGAGGTT	360
GATGTGATGT GCACTGCTTT TCATGACAAT GAAGAGACAT TTTTGAAAAA ATACTTATAT	420
GAAATTGCCA GAAGACATCC TTA CTTTTTAT GCCCGGAAC TCCTTTTCTT TGCTAAAAGG	480
TATAAAGCTG CTTTTACAGA ATGTTGCCAA GCTGCTGATA AAGCTGCCTG CCTGTTGCCA	540
AAGCTCGATG AACTTCGGGA TGAAGGGAAG GCTTCGTCTG CCAAACAGAG ACTCAAGTGT	600
GCCAGTCTCC AAAAATTTGG AGAAAGAGCT TTCAAAGCAT GGGCAGTAGC TCGCCTGAGC	660
CAGAGATTTT CCAAAGCTGA GTTTGCAGAA GTTTCCAAGT TAGTGACAGA TCTTACCAA	720
GTCCACACGG AATGCTGCCA TGGAGATCTG CTTGAATGTG CTGATGACAG GGCGGACCTT	780
GCCAAGTATA TCTGTGAAAA TCAAGATTCG ATCTCCAGTA AACTGAAGGA ATGCTGTGAA	840
AAACCTCTGT TGAAAAATC CCACTGCATT GCCGAAGTGG AAAATGATGA GATGCCTGCT	900
GACTTGCCTT CATTAGCTGC TGATTTTGTT GAAAGTAAGG ATGTTTGCAA AAACATGCT	960
GAGGCAAAGG ATGTCTTCCT GGGCATGTTT TTGTATGAAT ATGCAAGAAG GCATCCTGAT	1020
TACTCTGTCG TGCTGCTGCT GAGACTTGCC AAGACATATG AAACCACTCT AGAGAAGTGC	1080
TGTGCCGCTG CAGATCCTCA TGAATGCTAT GCCAAAGTGT TCGATGAATT TAAACCTCTT	1140
GTGGAAGAGC CTCAGAATTT AATCAAACAA AATTGTGAGC TTTTGTGAGCA GCTTGGAGAG	1200
TACAAATTCC AGAATGCGCT ATTAGTTCGT TACACCAAGA AAGTACCCCA AGTGTCAACT	1260

40

```

CCAACCTCTTG TAGAGGTCTC AAGAAACCTA GGAAAAGTGG GCAGCAAATG TTGTAAACAT      1320
CCTGAAGCAA AAAGAATGCC CTGTGCAGAA GACTATCTAT CCGTGGTCCT GAACCAGTTA      1380
TGTGTGTTGC ATGAGAAAAC GCCAGTAAGT GACAGAGTCA CCAAATGCTG CACAGAATCC      1440
TTGGTGAACA GGCGACCATG CTTTTCAGCT CTGGAAGTCG ATGAAACATA CGTTCCCAAA      1500
GAGTTTAATG CTGAAACATT CACCTTCCAT GCAGATATAT GCACACTTTC TGAGAAGGAG      1560
AGACAAATCA AGAAACAAAC TGCACTTGTT GAGCTCGTGA AACACAAGCC CAAGGCAACA      1620
AAAGAGCAAC TGAAAGCTGT TATGGATGAT TTCGCAGCTT TTGTAGAGAA GTGCTGCAAG      1680
GCTGACGATA AGGAGACCTG CTTTGCCGAG GAGGGTAAAA AACTTGTTGC TGCAAGTCAA      1740
GCTGCCTTAG GCTTATAA                                     1758

```

(2) INFORMATION FOR SEQ ID NO: 26:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 585 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

```

Asp Ala His Lys Ser Glu Val Ala His Arg Phe Lys Asp Leu Gly Glu
1          5          10          15
Glu Asn Phe Lys Ala Leu Val Leu Ile Ala Phe Ala Gln Tyr Leu Gln
20          25          30
Gln Cys Pro Phe Glu Asp His Val Lys Leu Val Asn Glu Val Thr Glu
35          40          45
Phe Ala Lys Thr Cys Val Ala Asp Glu Ser Ala Glu Asn Cys Asp Lys
50          55          60
Ser Leu His Thr Leu Phe Gly Asp Lys Leu Cys Thr Val Ala Thr Leu
65          70          75          80
Arg Glu Thr Tyr Gly Glu Met Ala Asp Cys Cys Ala Lys Gln Glu Pro
85          90          95
Glu Arg Asn Glu Cys Phe Leu Gln His Lys Asp Asp Asn Pro Asn Leu
100         105         110
Pro Arg Leu Val Arg Pro Glu Val Asp Val Met Cys Thr Ala Phe His
115         120         125
Asp Asn Glu Glu Thr Phe Leu Lys Lys Tyr Leu Tyr Glu Ile Ala Arg
130         135         140
Arg His Pro Tyr Phe Tyr Ala Pro Glu Leu Leu Phe Phe Ala Lys Arg
145         150         155         160

```

41

Tyr	Lys	Ala	Ala	Phe	Thr	Glu	Cys	Cys	Gln	Ala	Ala	Asp	Lys	Ala	Ala	
				165					170					175		
Cys	Leu	Leu	Pro	Lys	Leu	Asp	Glu	Leu	Arg	Asp	Glu	Gly	Lys	Ala	Ser	
			180					185					190			
Ser	Ala	Lys	Gln	Arg	Leu	Lys	Cys	Ala	Ser	Leu	Gln	Lys	Phe	Gly	Glu	
		195					200					205				
Arg	Ala	Phe	Lys	Ala	Trp	Ala	Val	Ala	Arg	Leu	Ser	Gln	Arg	Phe	Pro	
	210					215					220					
Lys	Ala	Glu	Phe	Ala	Glu	Val	Ser	Lys	Leu	Val	Thr	Asp	Leu	Thr	Lys	
225					230					235					240	
Val	His	Thr	Glu	Cys	Cys	His	Gly	Asp	Leu	Leu	Glu	Cys	Ala	Asp	Asp	
				245					250					255		
Arg	Ala	Asp	Leu	Ala	Lys	Tyr	Ile	Cys	Glu	Asn	Gln	Asp	Ser	Ile	Ser	
			260					265					270			
Ser	Lys	Leu	Lys	Glu	Cys	Cys	Glu	Lys	Pro	Leu	Leu	Glu	Lys	Ser	His	
		275					280					285				
Cys	Ile	Ala	Glu	Val	Glu	Asn	Asp	Glu	Met	Pro	Ala	Asp	Leu	Pro	Ser	
	290					295					300					
Leu	Ala	Ala	Asp	Phe	Val	Glu	Ser	Lys	Asp	Val	Cys	Lys	Asn	Tyr	Ala	
305					310					315					320	
Glu	Ala	Lys	Asp	Val	Phe	Leu	Gly	Met	Phe	Leu	Tyr	Glu	Tyr	Ala	Arg	
				325					330					335		
Arg	His	Pro	Asp	Tyr	Ser	Val	Val	Leu	Leu	Leu	Arg	Leu	Ala	Lys	Thr	
			340					345					350			
Tyr	Glu	Thr	Thr	Leu	Glu	Lys	Cys	Cys	Ala	Ala	Ala	Asp	Pro	His	Glu	
		355					360					365				
Cys	Tyr	Ala	Lys	Val	Phe	Asp	Glu	Phe	Lys	Pro	Leu	Val	Glu	Glu	Pro	
	370					375					380					
Gln	Asn	Leu	Ile	Lys	Gln	Asn	Cys	Glu	Leu	Phe	Glu	Gln	Leu	Gly	Glu	
385					390					395					400	
Tyr	Lys	Phe	Gln	Asn	Ala	Leu	Leu	Val	Arg	Tyr	Thr	Lys	Lys	Val	Pro	
				405					410					415		
Gln	Val	Ser	Thr	Pro	Thr	Leu	Val	Glu	Val	Ser	Arg	Asn	Leu	Gly	Lys	
			420					425					430			
Val	Gly	Ser	Lys	Cys	Cys	Lys	His	Pro	Glu	Ala	Lys	Arg	Met	Pro	Cys	
		435					440					445				
Ala	Glu	Asp	Tyr	Leu	Ser	Val	Val	Leu	Asn	Gln	Leu	Cys	Val	Leu	His	
	450					455					460					
Glu	Lys	Thr	Pro	Val	Ser	Asp	Arg	Val	Thr	Lys	Cys	Cys	Thr	Glu	Ser	
465					470					475					480	
Leu	Val	Asn	Arg	Arg	Pro	Cys	Phe	Ser	Ala	Leu	Glu	Val	Asp	Glu	Thr	
				485					490					495		
Tyr	Val	Pro	Lys	Glu	Phe	Asn	Ala	Glu	Thr	Phe	Thr	Phe	His	Ala	Asp	
			500					505					510			
Ile	Cys	Thr	Leu	Ser	Glu	Lys	Glu	Arg	Gln	Ile	Lys	Lys	Gln	Thr	Ala	

42

515					520					525					
Leu	Val	Glu	Leu	Val	Lys	His	Lys	Pro	Lys	Ala	Thr	Lys	Glu	Gln	Leu
530						535					540				
Lys	Ala	Val	Met	Asp	Asp	Phe	Ala	Ala	Phe	Val	Glu	Lys	Cys	Cys	Lys
545					550					555					560
Ala	Asp	Asp	Lys	Glu	Thr	Cys	Phe	Ala	Glu	Glu	Gly	Lys	Lys	Leu	Val
				565					570					575	
Ala	Ala	Ser	Gln	Ala	Ala	Leu	Gly	Leu							
			580					585							

THE EMBODIMENTS OF THE INVENTION FOR WHICH AN EXCLUSIVE PROPERTY OR PRIVILEGE IS CLAIMED ARE DEFINED AS FOLLOWS:

1. An albumin fusion polypeptide comprising human growth hormone at the C-terminal end of the fusion polypeptide and human serum albumin at the N-terminal end of the fusion polypeptide, wherein said fusion polypeptide has a greater serum half-life and greater aqueous stability than unfused human growth hormone.
2. A polypeptide according to claim 1, wherein the human growth hormone is mature growth hormone.
3. A polypeptide according to claim 1, wherein the human serum albumin is mature serum albumin.
4. A polypeptide according to claim 1, wherein the polypeptide comprises a fusion of the mature form of human serum albumin and the mature form of human growth hormone.
5. A polypeptide according to any one of claims 1 to 4, which further comprises additional amino acids added to either end of the polypeptide.
6. A polypeptide according to any one of claims 1 to 5, wherein said polypeptide has a non-cleavable linker peptide between the human growth hormone and the human serum albumin.
7. A composition comprising the polypeptide according to any one of claims 1 to 6 and a cell culture medium.
8. A polynucleotide encoding the polypeptide according to any one of claims 1 to 6.
9. A host cell comprising the polynucleotide according to claim 8, wherein the

polynucleotide is expressed in the host cell.

10. The host cell of claim 9, wherein said host cell is an animal host cell, a yeast host cell, or a bacterial host cell.

11. The host cell of claim 10, wherein said yeast host cell is *S. cerevisiae*.

12. A host cell according to any one of claims 9 to 11, wherein the polypeptide is secreted from the host cell.

13. A process for obtaining the polypeptide of any one of claims 1 to 6 from the host cell according to any one of claims 9 to 12, comprising culturing the host cell and purifying the polypeptide.

14. A process for obtaining the polypeptide of any one of claims 1 to 6 from the composition of claim 7, comprising purifying the polypeptide from the composition.

15. A pharmaceutical formulation for treating a patient having a condition selected from the group consisting of isolated growth hormone deficiency, panhypopituitarism, post-cranial irradiation, Turner's syndrome, Down's syndrome, intrauterine growth retardation, idiopathic growth deficiency, chronic renal failure, achondroplasia, female infertility and catabolic disorders, the formulation comprising the polypeptide according to any one of claims 1 to 6, or produced by the process of claim 13 or 14, and a pharmaceutically acceptable carrier.

16. The pharmaceutical formulation of claim 15, wherein said post-cranial irradiation condition arises from radiation treatment of leukaemia or brain tumours.

17. Use of a polypeptide according to any one of claims 1 to 6, or produced by the process of claim 13 or 14, in the preparation of a medicament for treating a patient having a condition selected from the group consisting of isolated growth hormone

deficiency, panhypopituitarism, a condition that is treatable with growth hormone following post-cranial irradiation, Turner's syndrome, Down's syndrome, intrauterine growth retardation, idiopathic growth deficiency, chronic renal failure, achondroplasia, female infertility and catabolic disorders.

18. The use according to claim 17, wherein said post-cranial irradiation condition arises from radiation treatment of leukaemia or brain tumours.

19. Use of a polypeptide according to any one of claims 1 to 6, or produced by the process of claim 13 or 14, for treating a patient having a condition selected from the group consisting of isolated growth hormone deficiency, panhypopituitarism, a condition that is treatable with growth hormone following post-cranial irradiation, Turner's syndrome, Down's syndrome, intrauterine growth retardation, idiopathic growth deficiency, chronic renal failure, achondroplasia, female infertility and catabolic disorders.

20. The use according to claim 19, wherein said post-cranial irradiation arises from radiation treatment of leukaemia or brain tumours.

21. A vector comprising the polynucleotide of claim 8.

22. A composition comprising the polypeptide according to any one of claims 1 to 6, and an acceptable carrier.

1/13

```

F   P   T   I   P   L   S   R   L   F   D   N   A   M   L   R   A   H   R
TTC CCA ACC ATT CCC TTA TCC AGG CTT TTT GAC AAC GCT ATG CTC CGC GCC CAT CGT
                                     ^50
L   H   Q   L   A   F   D   T   Y   Q   E   F   E   E   A   Y   I   P   K
CTG CAC CAG CTG GCC TTT GAC ACC TAC CAG GAG TTT GAA GAA GCC TAT ATC CCA AAG
                                     ^100
E   Q   K   Y   S   F   L   Q   N   P   Q   T   S   L   C   F   S   E   S
GAA CAG AAG TAT TCA TTC CTG CAG AAC CCC CAG ACC TCC CTC TGT TTC TCA GAG TCT
                                     ^150
I   P   T   P   S   N   R   E   E   T   Q   Q   K   S   N   L   E   L   L
ATT CCG ACA CCC TCC AAC AGG GAG GAA ACA CAA CAG AAA TCC AAC CTA GAG CTG CTC
                                     ^200
R   I   S   L   L   L   I   Q   S   W   L   E   P   V   Q   S   L   R   S
CGC ATC TCC CTG CTG CTC ATC CAG TCG TGG CTG GAG CCC GTG CAG TCC CTC AGG AGT
                                     ^250
V   F   A   N   S   L   V   Y   G   A   S   D   S   N   V   Y   D   L   L
GTC TTC GCC AAC AGC CTG GTG TAC GGC GCC TCT GAC AGC AAC GTC TAT GAC CTC CTA
                                     ^300
K   D   L   E   E   G   I   Q   T   L   M   G   R   L   E   D   G   S   P
AAG GAC CTA GAG GAA GGC ATC CAA ACG CTG ATG GGG AGG CTG GAA GAT GGC AGC CCC
                                     ^350
R   T   G   Q   I   F   K   Q   T   Y   S   K   F   D   T   N   S   H   N
CGG ACT GGG CAG ATC TTC AAG CAG ACC TAC AGC AAG TTC GAC ACA AAC TCA CAC AAC
^400                                     ^450
D   D   A   L   L   K   N   Y   G   L   L   Y   C   F   R   K   D   M   D
GAT GAC GCA CTA CTC AAG AAC TAC GGG CTG CTC TAC TGC TTC AGG AAG GAC ATG GAC
                                     ^500
K   V   E   T   F   L   R   I   V   Q   C   R   S   V   E   G   S   C   G
AAG GTC GAG ACA TTC CTG CGC ATC GTG CAG TGC CGC TCT GTG GAG GGC AGC TGT GGC
                                     ^550
F   .
TTC TAG

```

Figure 1. Human growth hormone cDNA sequence, encoding mature hGH.

2/13

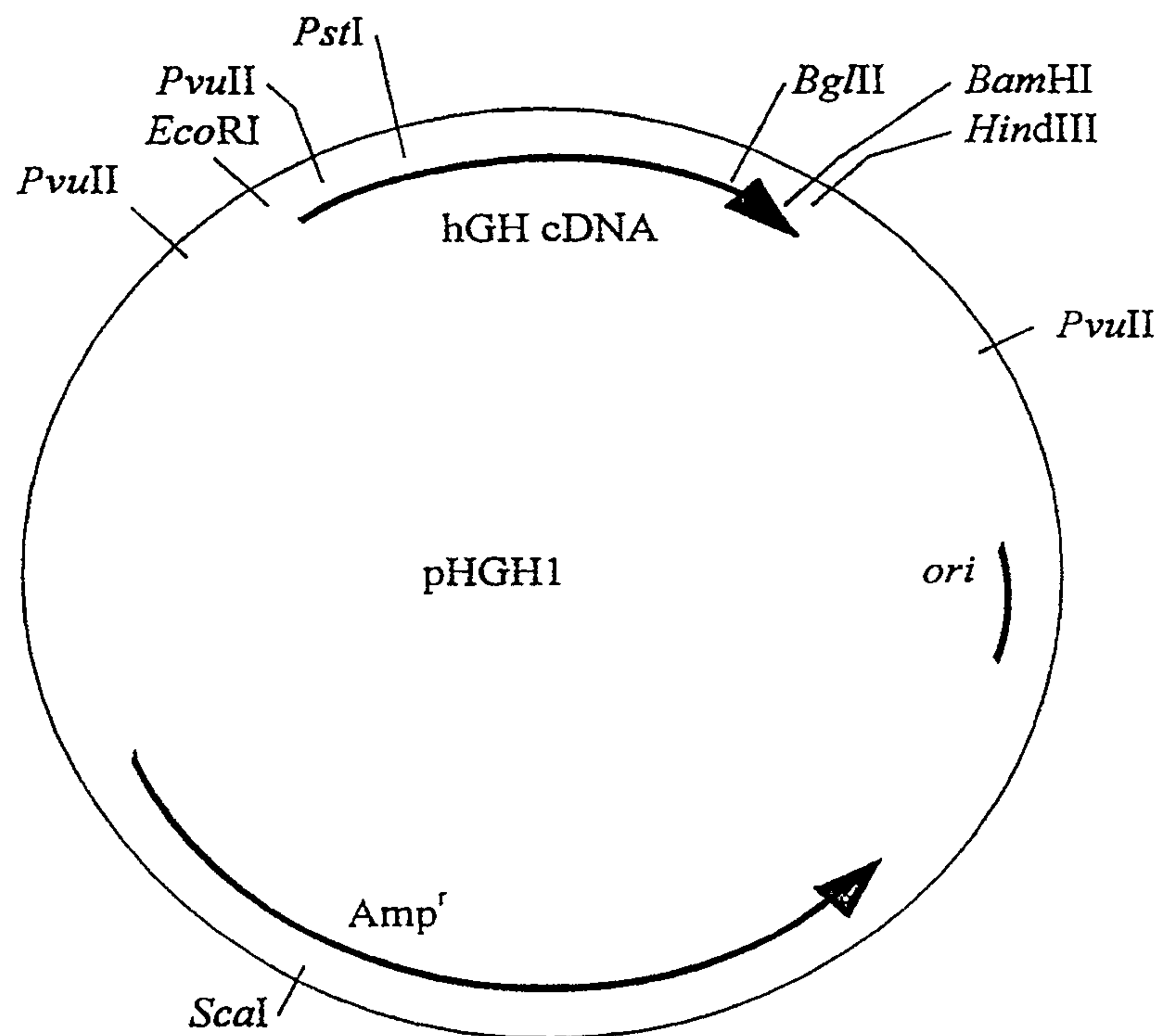


Figure 2

3/13

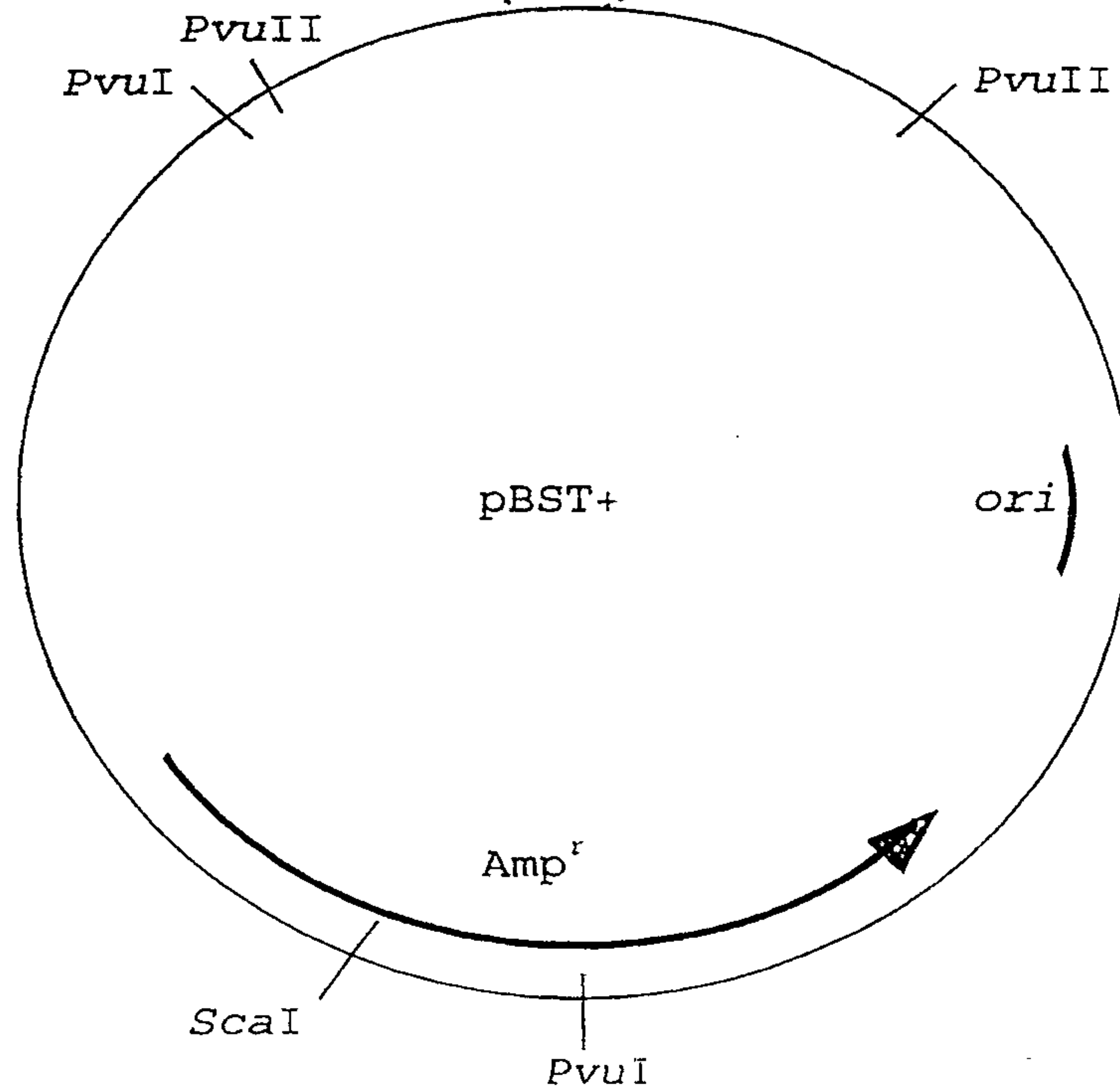
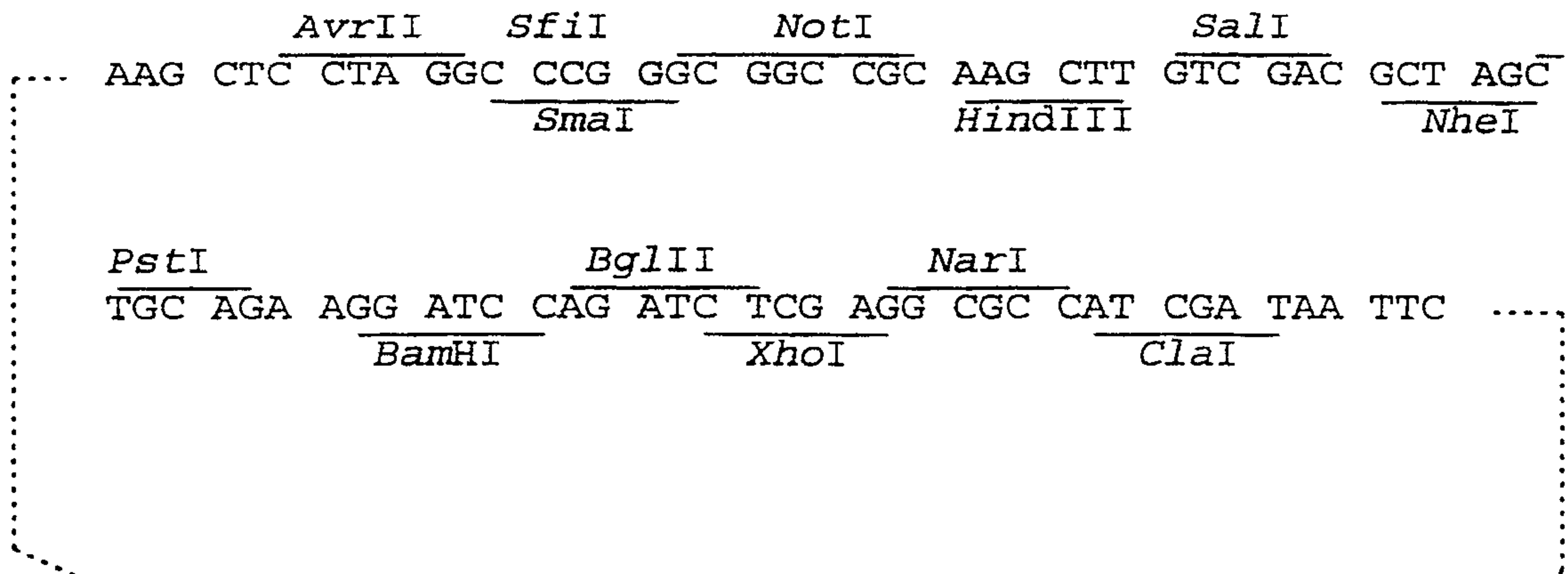


Figure 3

4/13

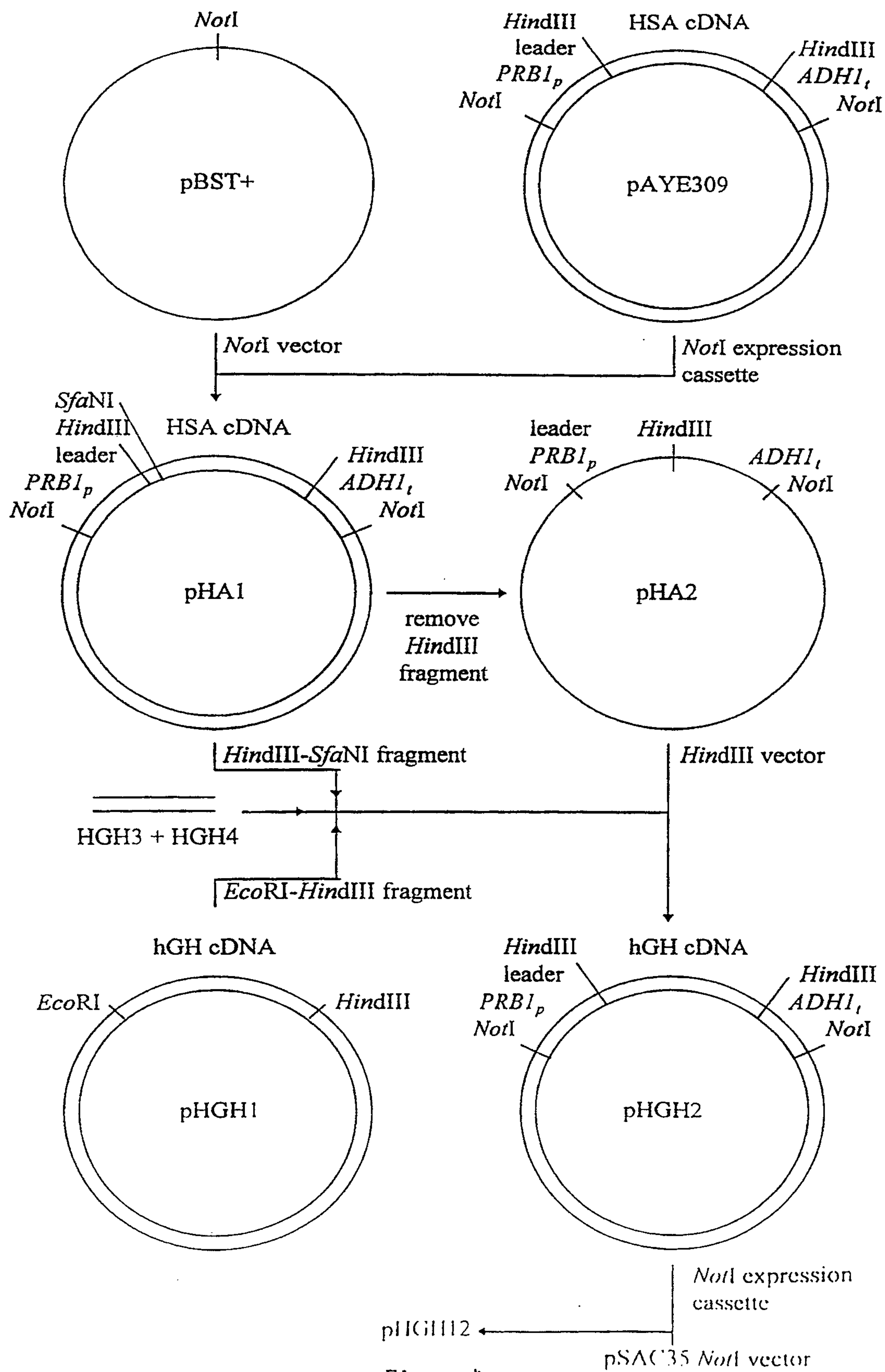


Figure 4

SUBSTITUTE SHEET (RULE 26)

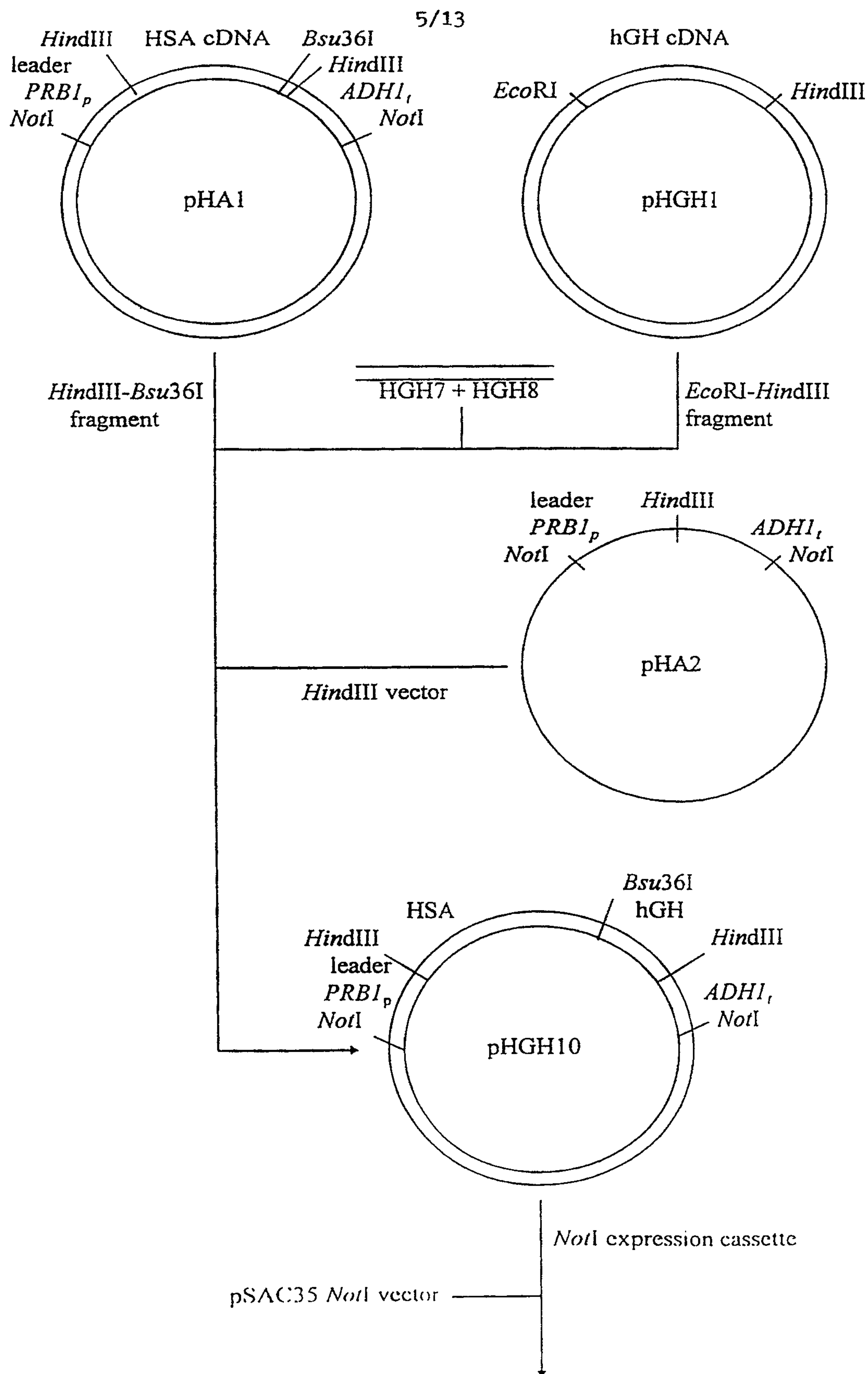


Figure 5 pHGH16

6/13

D A H K S E V A H R F K D L G E E N F
 GAT GCA CAC AAG AGT GAG GTT GCT CAT CGG TTT AAA GAT TTG GGA GAA GAA AAT TTC
 K A L V L I A F A Q Y L Q Q C P F E D
 AAA GCC TTG GTG TTG ATT GCC TTT GCT CAG TAT CTT CAG CAG TGT CCA TTT GAA GAT
 ^100
 H V K L V N E V T E F A K T C V A D E
 CAT GTA AAA TTA GTG AAT GAA GTA ACT GAA TTT GCA AAA ACA TGT GTT GCT GAT GAG
 S A E N C D K S L H T L F G D K L C T
 TCA GCT GAA AAT TGT GAC AAA TCA CTT CAT ACC CTT TTT GGA GAC AAA TTA TGC ACA
 ^200
 V A T L R E T Y G E M A D C C A K Q E
 GTT GCA ACT CTT CGT GAA ACC TAT GGT GAA ATG GCT GAC TGC TGT GCA AAA CAA GAA
 P E R N E C F L Q H K D D N P N L P R
 CCT GAG AGA AAT GAA TGC TTC TTG CAA CAC AAA GAT GAC AAC CCA AAC CTC CCC CGA
 ^300
 L V R P E V D V M C T A F H D N E E T
 TTG GTG AGA CCA GAG GTT GAT GTG ATG TGC ACT GCT TTT CAT GAC AAT GAA GAG ACA
 F L K K Y L Y E I A R R H P Y F Y A P
 TTT TTG AAA AAA TAC TTA TAT GAA ATT GCC AGA AGA CAT CCT TAC TTT TAT GCC CCG
 ^400
 E L L F F A K R Y K A A F T E C C Q A
 GAA CTC CTT TTC TTT GCT AAA AGG TAT AAA GCT GCT TTT ACA GAA TGT TGC CAA GCT
 ^500
 A D K A A C L L P K L D E L R D E G K
 GCT GAT AAA GCT GCC TGC CTG TTG CCA AAG CTC GAT GAA CTT CGG GAT GAA GGG AAG
 A S S A K Q R L K C A S L Q K F G E R
 GCT TCG TCT GCC AAA CAG AGA CTC AAG TGT GCC AGT CTC CAA AAA TTT GGA GAA AGA
 ^600
 A F K A W A V A R L S Q R F P K A E F
 GCT TTC AAA GCA TGG GCA GTA GCT CGC CTG AGC CAG AGA TTT CCC AAA GCT GAG TTT
 A E V S K L V T D L T K V H T E C C H
 GCA GAA GTT TCC AAG TTA GTG ACA GAT CTT ACC AAA GTC CAC ACG GAA TGC TGC CAT
 ^700
 G D L L E C A D D R A D L A K Y I C E
 GGA GAT CTG CTT GAA TGT GCT GAT GAC AGG GCG GAC CTT GCC AAG TAT ATC TGT GAA
 N Q D S I S S K L K E C C E K P L L E
 AAT CAA GAT TCG ATC TCC AGT AAA CTG AAG GAA TGC TGT GAA AAA CCT CTG TTG GAA
 ^800
 K S H C I A E V E N D E M P A D L P S
 AAA TCC CAC TGC ATT GCC GAA GTG GAA AAT GAT GAG ATG CCT GCT GAC TTG CCT TCA
 ^900

Figure 6. HSA cDNA sequence, showing the region encoding the mature protein.

WO 97/24445

PCT/GB96/03164

7/13

L A A D F V E S K D V C K N Y A E A K
 TTA GCT GCT GAT TTT GTT GAA AGT AAG GAT GTT TGC AAA AAC TAT GCT GAG GCA AAG
 D V F L G M F L Y E Y A R R H P D Y S
 GAT GTC TTC CTG GGC ATG TTT TTG TAT GAA TAT GCA AGA AGG CAT CCT GAT TAC TCT
 ^1000
 V V L L L R L A K T Y E T T L E K C C
 GTC GTG CTG CTG CTG AGA CTT GCC AAG ACA TAT GAA ACC ACT CTA GAG AAG TGC TGT
 A A A D P H E C Y A K V F D E F K P L
 GCC GCT GCA GAT CCT CAT GAA TGC TAT GCC AAA GTG TTC GAT GAA TTT AAA CCT CTT
 ^1100
 V E E P Q N L I K Q N C E L F E Q L G
 GTG GAA GAG CCT CAG AAT TTA ATC AAA CAA AAT TGT GAG CTT TTT GAG CAG CTT GGA
 E Y K F Q N A L L V R Y T K K V P Q V
 GAG TAC AAA TTC CAG AAT GCG CTA TTA GTT CGT TAC ACC AAG AAA GTA CCC CAA GTG
 ^1200
 S T P T L V E V S R N L G K V G S K C
 TCA ACT CCA ACT CTT GTA GAG GTC TCA AGA AAC CTA GGA AAA GTG GGC AGC AAA TGT
 ^1300
 C K H P E A K R M P C A E D Y L S V V
 TGT AAA CAT CCT GAA GCA AAA AGA ATG CCC TGT GCA GAA GAC TAT CTA TCC GTG GTC
 L N Q L C V L H E K T P V S D R V T K
 CTG AAC CAG TTA TGT GTG TTG CAT GAG AAA ACG CCA GTA AGT GAC AGA GTC ACC AAA
 ^1400
 C C T E S L V N R R P C F S A L E V D
 TGC TGC ACA GAA TCC TTG GTG AAC AGG CGA CCA TGC TTT TCA GCT CTG GAA GTC GAT
 E T Y V P K E F N A E T F T F H A D I
 GAA ACA TAC GTT CCC AAA GAG TTT AAT GCT GAA ACA TTC ACC TTC CAT GCA GAT ATA
 ^1500
 C T L S E K E R Q I K K Q T A L V E L
 TGC ACA CTT TCT GAG AAG GAG AGA CAA ATC AAG AAA CAA ACT GCA CTT GTT GAG CTC
 V K H K P K A T K E Q L K A V M D D F
 GTG AAA CAC AAG CCC AAG GCA ACA AAA GAG CAA CTG AAA GCT GTT ATG GAT GAT TTC
 ^1600
 A A F V E K C C K A D D K E T C F A E
 GCA GCT TTT GTA GAG AAG TGC TGC AAG GCT GAC GAT AAG GAG ACC TGC TTT GCC GAG
 ^1700
 E G K K L V A A S Q A A L G L
 GAG GGT AAA AAA CTT GTT GCT GCA AGT CAA GCT GCC TTA GGC TTA TAA

Figure 6 (cont'd)

8/13

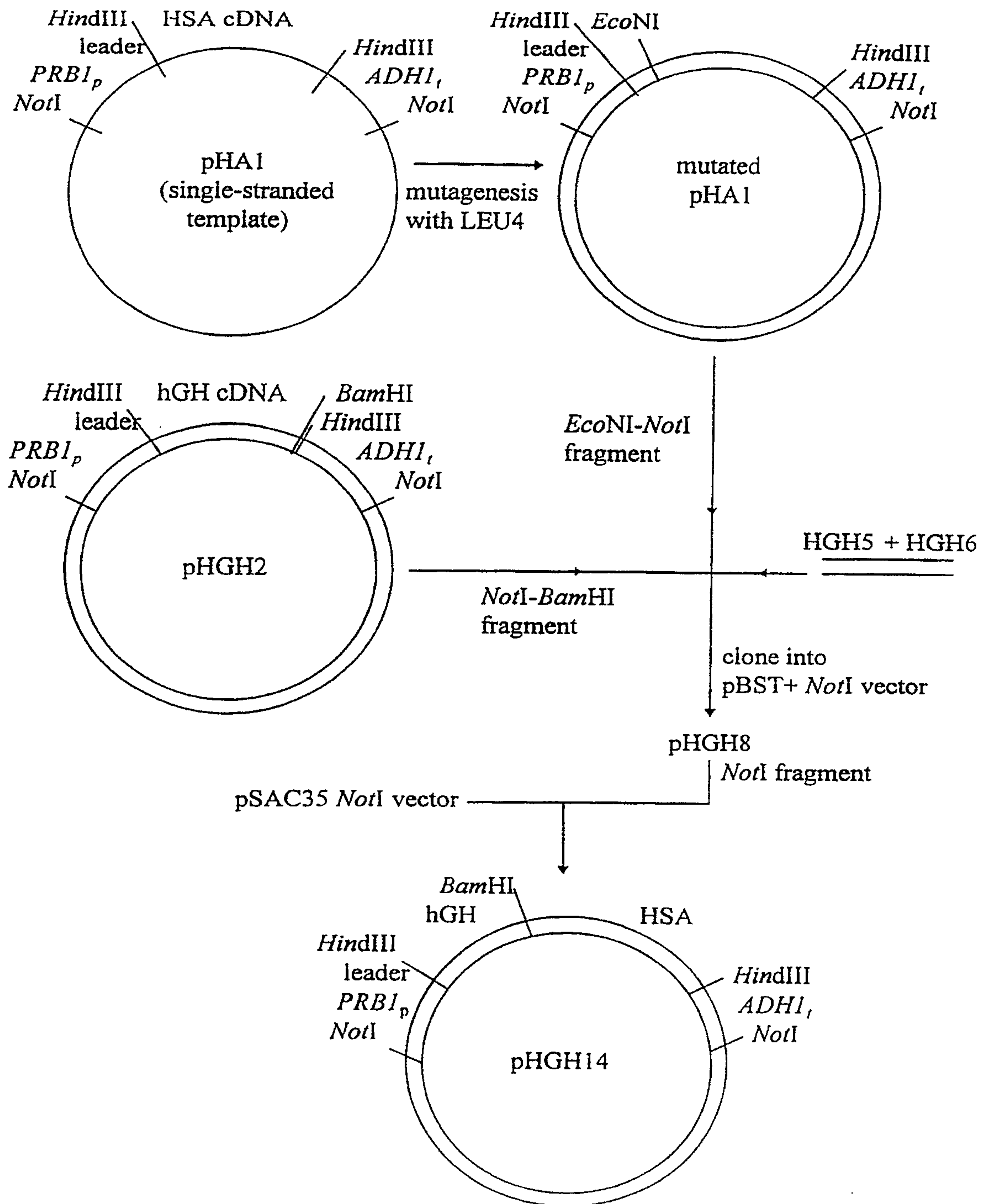


Figure 7

9/13

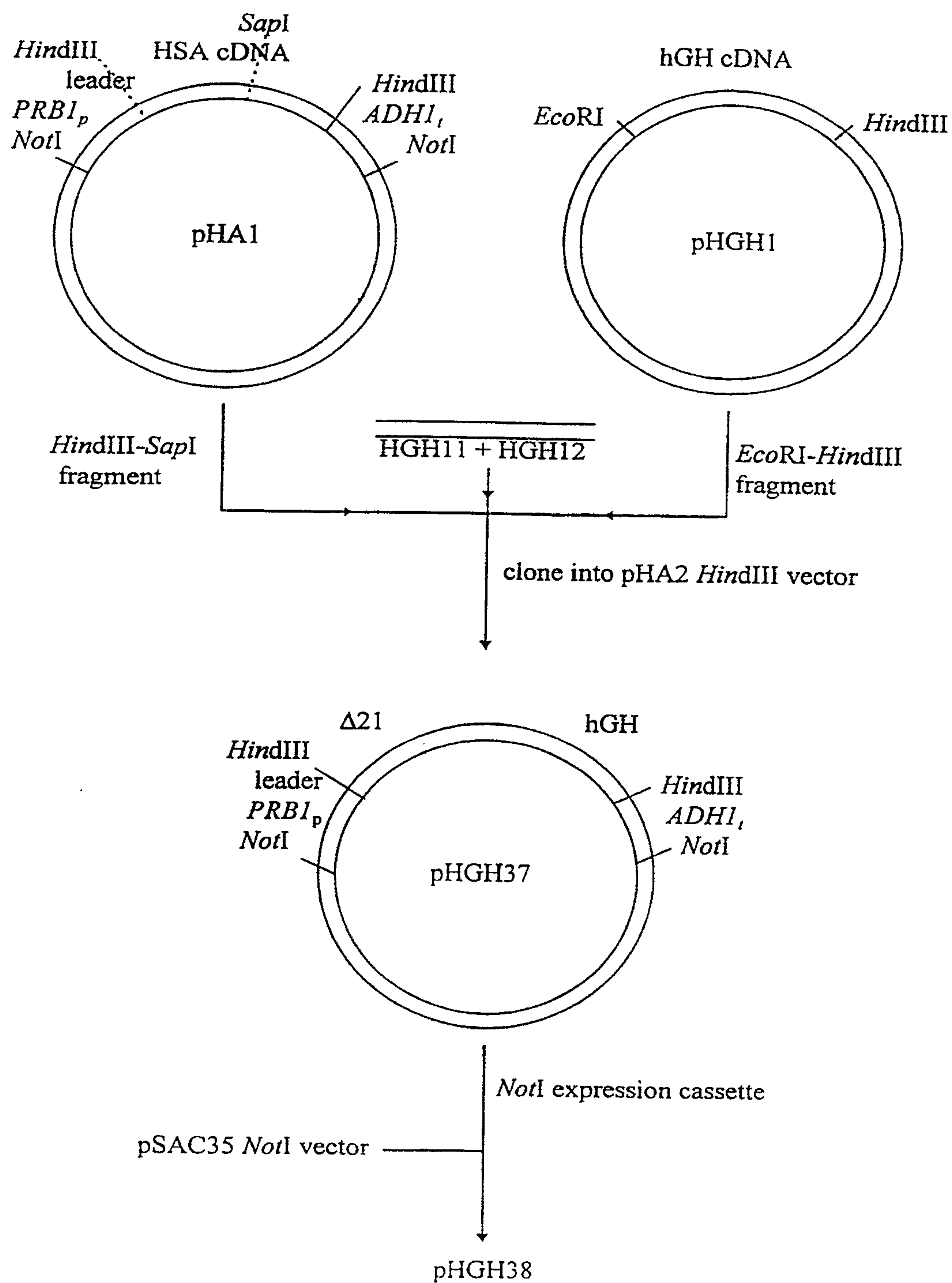


Figure 8

10/13

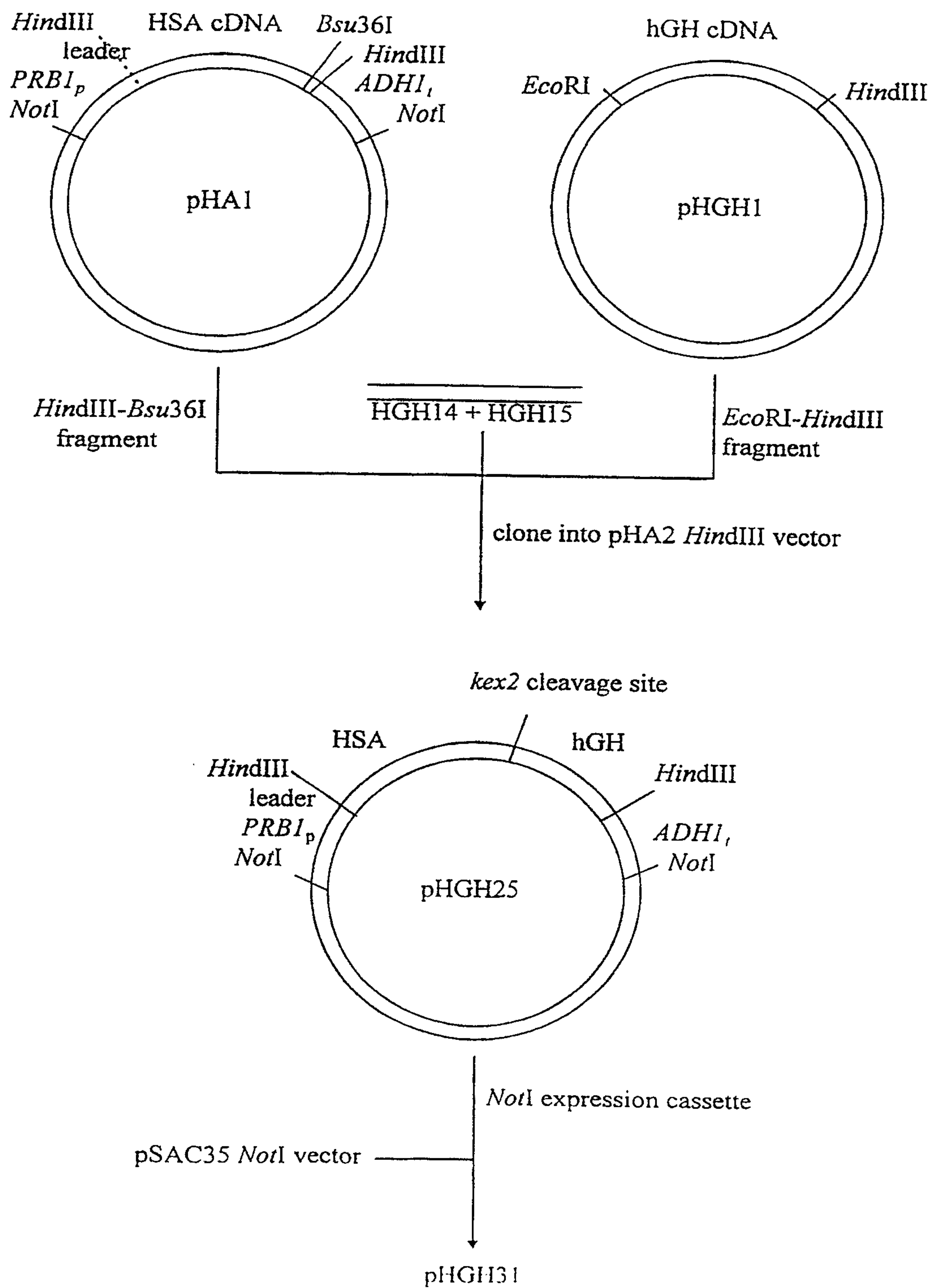


Figure 9

11/13

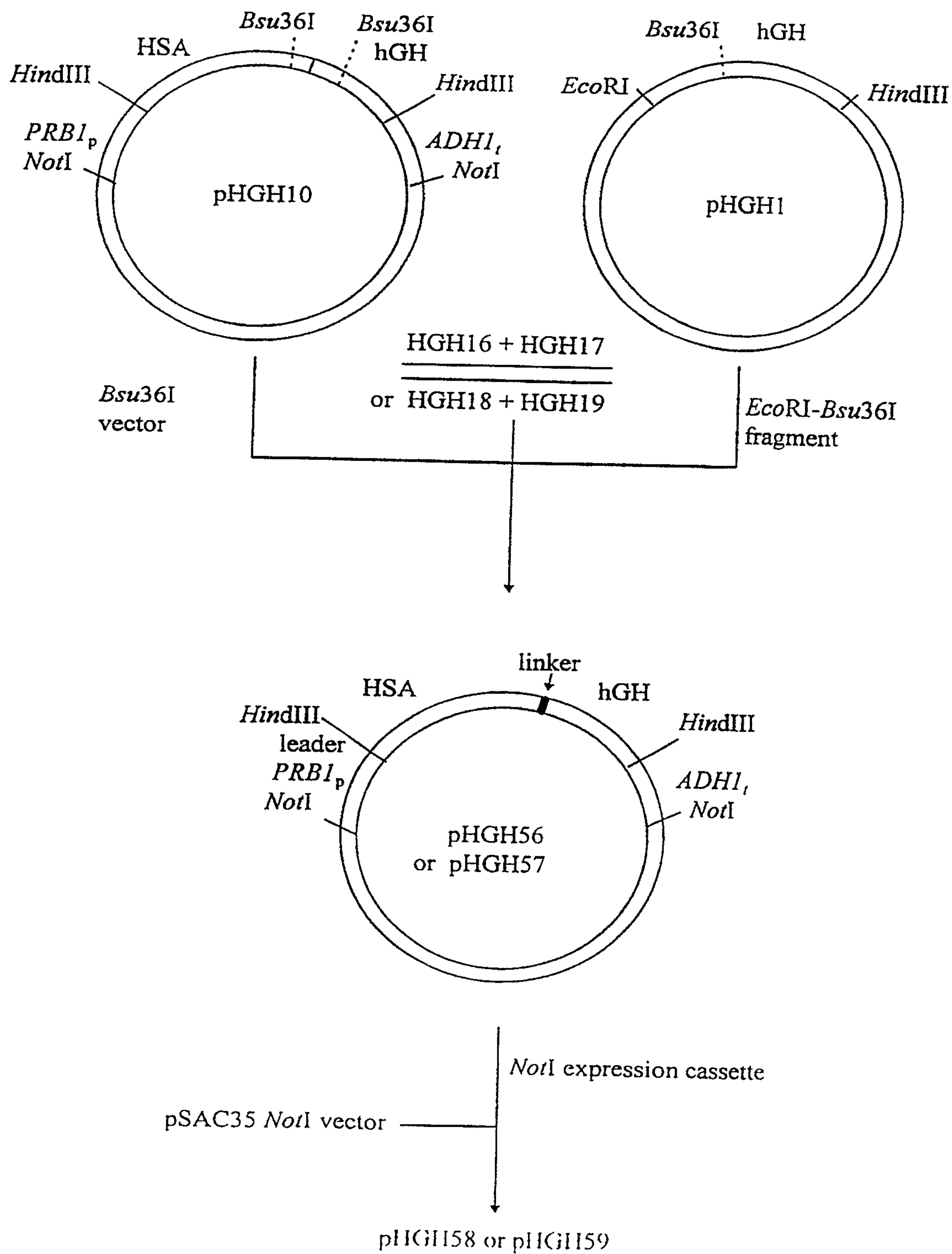


Figure 10

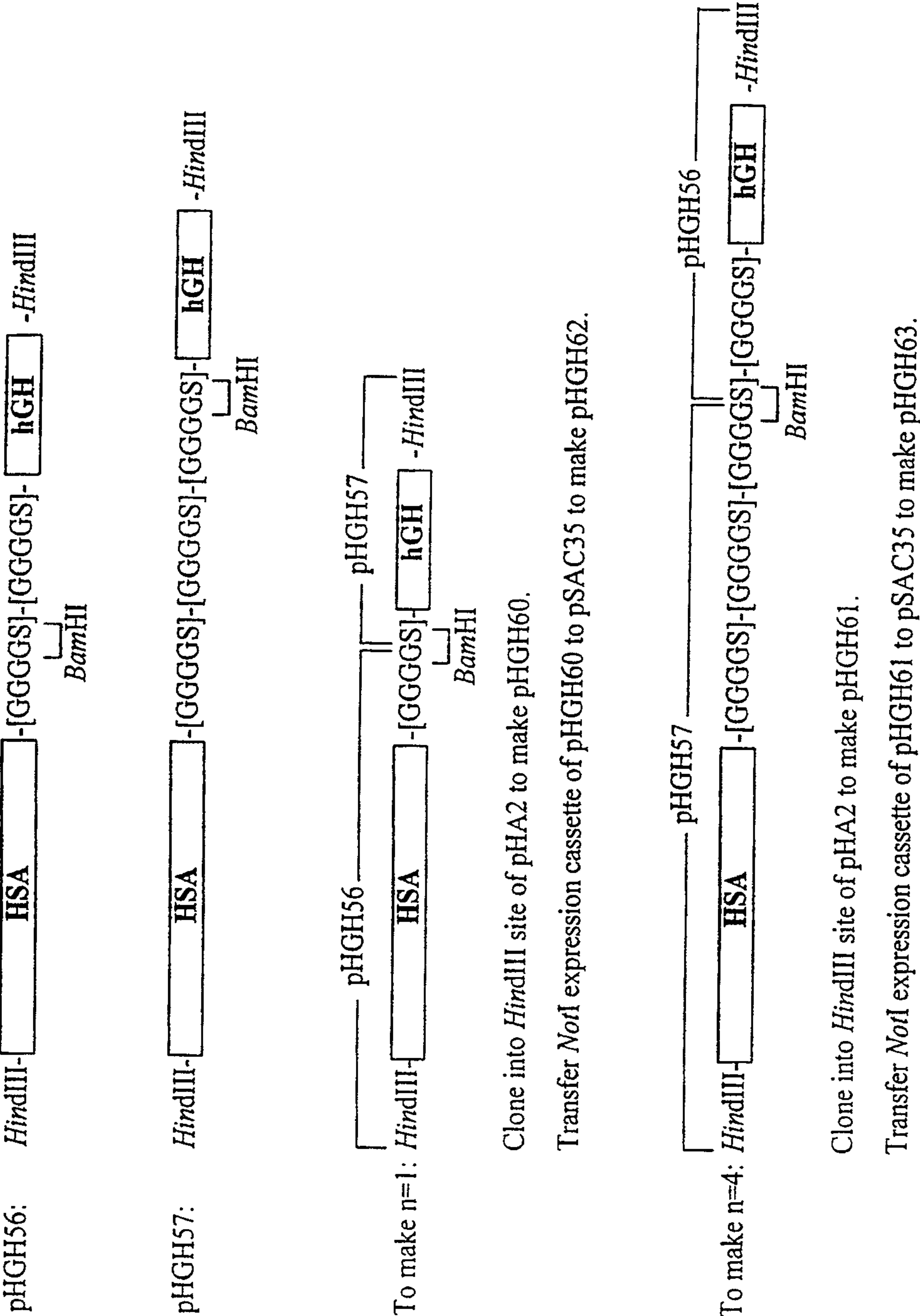


Figure 11

13/13

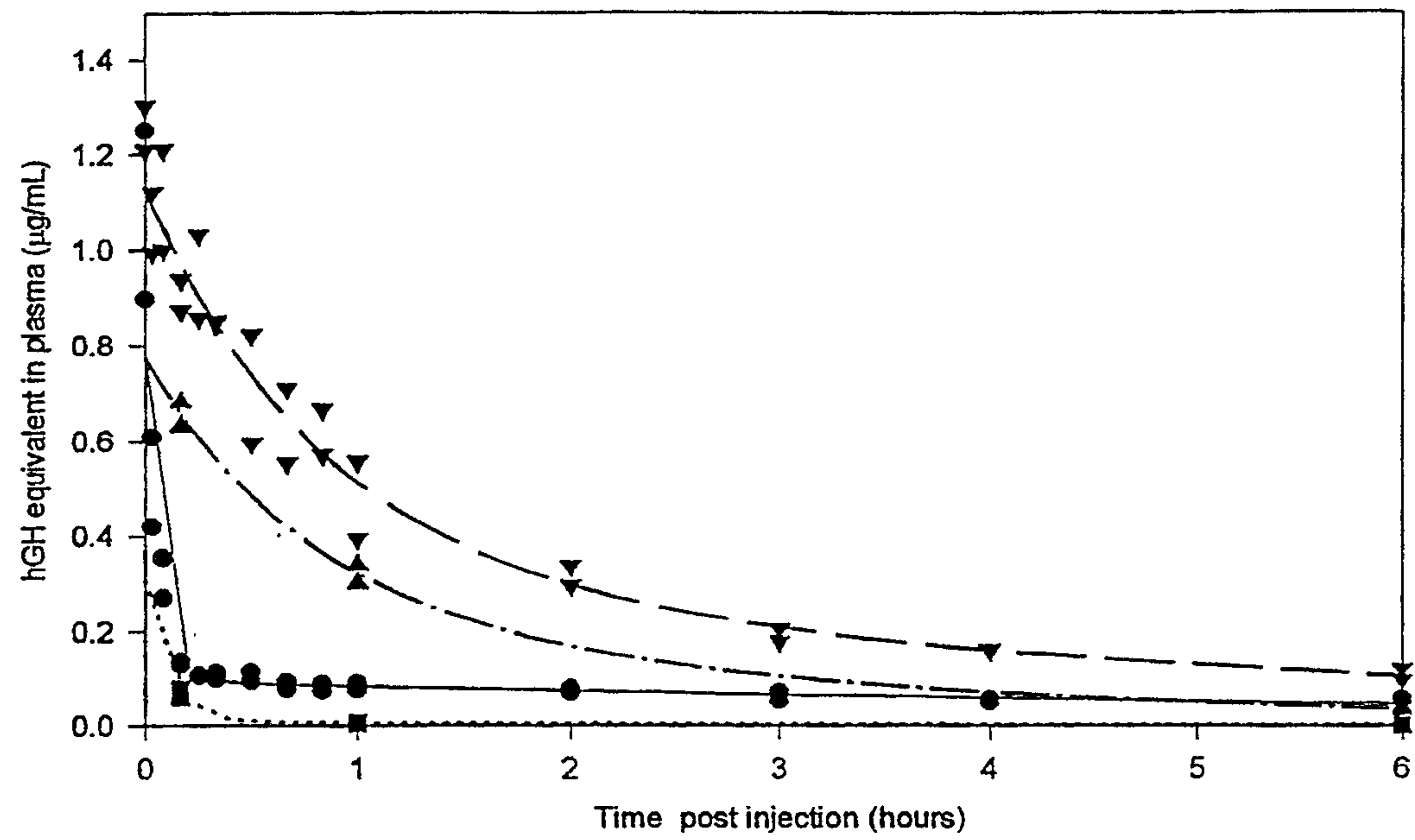


Figure 12