PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/62, 1/15, 1/19, 15/81, 15/12

A2

(11) International Publication Number: **WO 97/12044**

(43) International Publication Date: 3 April 1997 (03.04.97)

(21) International Application Number: PCT/US96/15329

(22) International Filing Date: 25 September 1996 (25.09.96)

(30) Priority Data:

60/004,264	25 September 1995 (25.09.95)	US
60/004,327	26 September 1995 (26.09.95)	US

(60) Parent Application or Grant
(63) Related by Continuation
US 60/004,264 (CIP)
Filed on 25 September 1995 (25.09.95)

(71) Applicant *(for all designated States except US)*: CHIRON CORPORATION [US/US]; 4560 Horton Street, Emeryville, CA 94608 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: CRAWFORD, Kenneth [US/US]; 3214 Encinal Avenue, Alameda, CA 94501 (US). ZAVOR, Isabel [US/US]; 6 Valley View Drive, Orinda, CA 94563 (US). BISHOP, Robert, J. [US/US]; 1501 Lincoln Way #404, San Francisco, CA 94122 (US). INNS, Michael, A. [US/US]; 315 Constance Place, Moraga, CA 94556 (US).

(74) Agents: CHUNG, Ling, Fong et al.; Chiron Corporation, Intellectual Property - R440, P.O. Box 8097, Emeryville, CA 94662-8097 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, 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
Without international search report and to be republished upon receipt of that report.

(54) Title: PICHIA SECRETORY LEADER FOR PROTEIN EXPRESSION

(57) Abstract

Polynucleotides, vectors and host cells comprising a polynucleotide having a fragment of a leader sequence and a second nucleotide sequence that encodes a polypeptide heterologous to the leader sequence, wherein the leader sequence fragment is sufficient for secretion and comprises an amino acid sequence that comprises at least about 70 % sequence identity to the leader sequence of *Pichia acaciae* killer toxin, wherein the heterologous polypeptide is not naturally contiguous to the leader sequence, and wherein upon expression of the polynucleotide molecule in a host cell suitable for expression thereof, the heterologous polypeptide is produced that is free of additional N-terminal amino acids.

wherein the leader sequence fragment is sufficient for secretion and comprises an amino acid sequence that comprises at least about 70% sequence identity to the leader sequence of *Picahia acaciae* killer toxin,

wherein the heterologous polypeptide is not naturally contiguous to the leader sequence, and

wherein upon expression of the polynucleotide molecule in a host cell suitable for expression thereof, the heterologous polypeptide is produced that is free of additional N-terminal amino acids.

The polynucleotide of the invention can be used to construct expression vectors and host cells capable of producing the polynucleotide or expressing the desired polypeptide.

Yet another object of the invention is to provide a method of producing a polypeptide encoded by a polynucleotide comprising

(a) transforming a host cell with the polynucleotide,

(b) allowing the expression thereof to produce the polypeptide and

(c) obtaining the polypeptide therefrom,

wherein the polynucleotide molecule comprises a first nucleotide sequence that encodes at least a fragment of a leader sequence and a second nucleotide sequence that encodes a polypeptide heterologous to the leader sequence,

wherein the leader sequence fragment is sufficient for secretion and comprises an amino acid sequence that comprises at least about 70% sequence identity to the leader sequence of *Picahia acaciae* killer toxin,

wherein the heterologous polypeptide is not naturally contiguous to the leader sequence, and

wherein upon expression of the polynucleotide molecule in a host cell suitable for expression thereof, the heterologous polypeptide is produced that is free of additional N-terminal amino acids.

A specific embodiment of the invention is where the heterologous polypeptide is human insulin-like growth factor 1 (IGF-1).

Pichia Secretory Leader for Protein Expression

Inventors: Kenneth Crawford, Isabel Zaror, Bob Bishop and Michael Innis

Description

Background of the Invention

Recombinant DNA technology has revolutionized the ability to produce polypeptides economically. Yeast host cells and expression systems are useful for such production. Examples of yeast expression systems are Brake, US Pat No. 4,870,008; Cregg, US Pat No 4,837,148; Stroman *et al.*, US Pat No 4,855,231; Stroman *et al.* US Pat No 4,879,231; Brierley *et al.*, US Pat No 5,324,639; Prevatt *et al.*, US Pat No 5,330,901; Tschopp, EP 256 421; Sreekrishna *et al.*, J. Basic Microbiol. 28(1988): 4 265-278; Tschopp *et al.*, Bio/Technology 5(1987): 1305-1308; Cregg *et al.*, Bio/Technology 5(1987): 479-485; Sreekrishna *et al.* Biochemistry 28(1989): 4117-4125; and Bolen *et al.*, Yeast 10: 403-414 (1994).

General recombinant DNA methods can be found, for example, in Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual* (2nd ed., 1989).

All publications, patents and patent applications cited herein, whether supra or infra, are hereby incorporated by reference in their entirety.

Summary of the Invention

It is an object of the invention to provide a polynucleotide molecule comprising a first nucleotide sequence that encodes at least a fragment of a leader sequence and a second nucleotide sequence that encodes a polypeptide heterologous to the leader sequence,

generally hydrophobic and exhibit a three dimensional helical structure. Also, a cleavage site can be incorporated in the fragment to facilitate removal of the leader fragment from the heterologous polypeptide. Examples are peptidase cleavage sites, which include KEX2 as an example. Preferably, the cleavage site comprises a dibasic dipeptide such as, lys-lys, arg-arg, more preferably lys-arg.

The leader sequence can be altered for convenience or to optimize expression. For example, the amino acid sequence of *Picahia acaciae* signal peptide can be mutated. The following are examples of conservative substitutions: Gly⇔Ala; Val⇔Ile⇔Leu; Asp⇔Glu; Lys⇔Arg; Asn⇔Gln; and Phe⇔Trp⇔Tyr. A subset of mutants, called muteins, is a group of polypeptides with the non-disulfide bond participating cysteines substituted with a neutral amino acid, generally, with serines..

The amino acid sequence of the *Picahia acaciae* killer toxin leader sequence, SEQ ID NO:1 can be aligned with the leader sequence of other yeast killer toxin genes to detemine the positions of variable and conserved amino acid residues.

Full length and fragments of *Picahia acaciae* killer toxin leader sequences as well as mutants thereof, can be fused with additional amino acid residues. For example, the consensus sequence of pro-regions from other leader sequences can be determined and incorporated into the leader sequence. Such pro-region sequences can be helpful to optimize expression in a particular host cell.

Polynucleotide sequence encoding the leader sequence can be based on the sequence found in genomic DNA or be made by using codons preferred by the host cell. In both cases, the polynucleotides can be synthesized using the methods described in Urdea *et al.*, Proc. Natl. Acad. Sci. USA 80: 7461 (1983), for example. Alternatively, the polynucleotides from nucleic acid libraries using probes based on the nucleic acid sequence shown in SEQ ID NO:1. Techniques for producing and probing nucleic acid sequence libraries are described, for example, in Sambrook *et al.*, "Molecular Cloning: A Laboratory Manual" (New York, Cold Spring Harbor Laboratory, 1989). Other recombinant techniques, such as site specific mutagenesis, PCR, enzymatic digestion and ligation, can also be used to clone or modify the sequences found from natural sources.

Similarly, the polynucleotides encoding the desired polypeptide can also be constructed using synthetic or recombinant means. Amino acid sequence of polypeptides to be expressed can also be found in publically available databases.

Brief Description of the Drawings

Figure 1 is a plasmid map of pHIL-A1.

Detailed Description

Definitions

"Heterologous" means not naturally contiguous. For example, a yeast leader and a human protein are heterologous because the two are not naturally contiguous.

A host cell suitable of "expression of a polynucleotide" is capable of effecting transcription and translation of the polynucleotide to produce the encoded heterolgous polypeptide free of additional N-terminal amino acids.

General Methods and Detailed Description

Preferably, polynucleotides of the instant invention are produced by recombinant DNA techniques. The polynucleotide encoding at least a fragment of a leader sequence can be either synthesized or cloned.

The amino acid sequence of the leader sequence comprises at least 70% sequence identity to the leader sequence of the *Picahia acaciae* killer toxin, described in Bolen *et al.*, Yeast 10: 403-414 (1994) and shown in SEQ ID NO:1. More preferably, the leader sequence comprises at least 80%; even more preferably, at least 90%; more preferably, at least 95% sequence identity to SEQ ID NO:1; most preferably, 100% sequence identity to SEQ ID NO:1.

A full length leader sequence begins at the initiating methionine and ends at the last amino acid residue before the beginning of the encoded mature polypeptide. Amino acid residues can be removed from full length leader to construct leader fragments. These fragments can be tested to determine if they are sufficient for secretion.

Empirical data can be used, for example, to determine if a fragment of a leader sequence is sufficient for secretion. Host cells with the polynucleotide of the instant invention exhibit increased expression levels as compared to a negative control. See below for assays to detect polypeptide expression.

A full length leader sequence from a native gene, such a *Picahia acaciae* killer toxin, can be divided into a signal peptide region and a pro-region. Typically, a fragment sufficient for secretion comprises a signal peptide. Signal peptides are

transcription. These sequences can overlap the sequences that initiate expression. Most host cell systems include regulatory sequences within the promoter sequences. For example, when a repressor protein binds to the lac operon, an *E. coli* regulatory promoter sequence, transcription of the downstream gene is inhibited. Another example is the yeast alcohol dehydrogenase promoter, which has an upstream activator sequence (UAS) that modulates expression in the absence of glucose. Additionally, some viral enhancers not only amplify but also regulate expression in mammalian cells. These enhancers can be incorporated into mammalian promoter sequences, and the promoter will become active only in the presence of an inducer, such as a hormone or enzyme substrate (Sassone-Corsi and Borelli (1986) Trends Genet. 2:215; Maniatis *et al.* (1987) Science 236:1237).

Functional non-natural promoters may also be used, for example, synthetic promoters based on a consensus sequence of different promoters. Also, effective promoters can contain a regulatory region linked with a heterologous expression initiation region. Examples of hybrid promoters are the *E. coli* lac operator linked to the *E. coli* tac transcription activation region; the yeast alcohol dehydrogenase (ADH) regulatory sequence linked to the yeast glyceraldehyde-3-phosphate-dehydrogenase (GAPDH) transcription activation region (U.S. Patent Nos. 4,876,197 and 4,880,734, incorporated herein by reference); and the cytomegalovirus (CMV) enhancer linked to the SV40 (simian virus) promoter.

Typically, terminators are regulatory sequences, such as polyadenylation and transcription termination sequences, located 3' or downstream of the stop codon of the coding sequences. Usually, the terminator of native host cell proteins are operable when attached 3' of the polynucleotide of the instant invention. Examples are the *Saccharomyces cerevisiae* alpha-factor terminator and the baculovirus terminator. Further, viral terminators are also operable in certain host cells; for instance, the SV40 terminator is functional in CHO cells.

For convenience, selectable markers, an origin of replication, and homologous host cell sequences may optionally be included in an expression vector. A selectable marker can be used to screen for host cells that potentially contain the expression vector. Such markers may render the host cell immune to drugs such as ampicillin, chloramphenicol, erythromycin, neomycin, and tetracycline. Also, markers

Useful polypeptides to be expressed include, for example, hormones, growth factors, cytokines, haematopoietic factors, immunoglobulins, enzymes, repressors, cell differentiation factors, binding proteins, or transcription factors. Specific examples are: growth hormone, luteinizing hormone, thyroid stimulating hormone, oxytocin, insulin, vasopresin, renin, calcitonin, follicle stimulating hormone, prolactin, insulin-like growth factor (IGF-I, IGF-II), an IGF-binding protein, epidermal growth factor (EGF), platelet derived growth factor (PDGF), keratinocyte growth factor (KGF), fibroblast growth factor (FGF), nerve growth factor (NGF), TGF-beta, vascular endothelial cell growth factor (VEGF), erythropoietin (EPO), colony stimulating factor (CSF), interferon, endorphin, enkaphalin, dynorphin and an active fragment thereof.

The two polynucleotides, encoding at least a fragment of a leader sequence and the heterologous polypeptide, are linked together to produce the polynucleotide of the instant invention. Preferably, the polynucleotides are linked together in proper reading frame.

Polynucleotides encoding at least a fragment of a leader sequence and encoding polypeptides can be expressed by a variety of host cells. Although the leader sequence may be yeast derived and linked to a human protein, for example, host cells as diverse as yeast, insect, and mammalian host cells can express the polypeptide.

Typically, the polynucleotide of the instant invention, leader sequence and polypeptide, can be incorporated into an expression vector, which is in turn inserted into the desired host cell for expression.

At the minimum, an expression vector will contain a promoter which is operable in the host cell and operably linked to polynucleotide of the instant invention. Expression vectors may also include signal sequences, terminators, selectable markers, origins of replication, and sequences homologous to host cell sequences. These additional elements are optional but can be included to optimize expression.

A promoter is a DNA sequence upstream or 5' to the polynucleotide of the instant invention to be expressed. The promoter will initiate and regulate expression of the coding sequence in the desired host cell. To initiate expression, promoter sequences bind RNA polymerase and initiate the downstream (3') transcription of a coding sequence (*e.g.* structural gene) into mRNA. A promoter may also have DNA sequences that regulate the rate of expression by enhancing or specifically inducing or repressing

Wright (1986) Nature 321:718; Smith *et al.*, (1983) Mol. Cell. Biol. 3:2156; and see generally, Fraser, *et al.* (1989) In Vitro Cell. Dev. Biol. 25:225).

Transformation

After vector construction, the expression vector is inserted into the host cell. Many transformation techniques exist for inserting expression vectors into bacterial, yeast, insect, and mammalian cells. The transformation procedure to introduce the expression vector depends upon the host to be transformed.

Methods of introducing exogenous DNA into bacterial hosts are well-known in the art, and typically protocol includes either treating the bacteria with $CaCl_2$ or other agents, such as divalent cations and DMSO. DNA can also be introduced into bacterial cells by electroporation or viral infection. Transformation procedures usually vary with the bacterial species to be transformed. See *e.g.*, (Masson *et al.* (1989) FEMS Microbiol. Lett. 60:273; Palva *et al.* (1982) Proc. Natl. Acad. Sci. USA 79:5582; EP Publ. Nos. 036 259 and 063 953; PCT WO 84/04541, *Bacillus*), (Miller *et al.* (1988) Proc. Natl. Acad. Sci. 85:856; Wang *et al.* (1990) J. Bacteriol. 172:949, *Campylobacter*), (Cohen *et al.* (1973) Proc. Natl. Acad. Sci. 69:2110; Dower *et al.* (1988) Nucleic Acids Res. 16:6127; Kushner (1978) "An improved method for transformation of *Escherichia coli* with ColE1-derived plasmids in Genetic Engineering: Proceedings of the International Symposium on Genetic Engineering (eds. H.W. Boyer and S. Nicosia); Mandel *et al.* (1970) J. Mol. Biol. 53:159; Taketo (1988) Biochim. Biophys. Acta 949:318; Escherichia), (Chassy *et al.* (1987) FEMS Microbiol. Lett. 44:173 *Lactobacillus*); (Fiedler *et al.* (1988) Anal. Biochem 170:38, *Pseudomonas*); (Augustin *et al.* (1990) FEMS Microbiol. Lett. 66:203, *Staphylococcus*), (Barany *et al.* (1980) J. Bacteriol. 144:698; Harlander (1987) "Transformation of *Streptococcus lactis* by electroporation," in Streptococcal Genetics (ed. J. Ferretti and R. Curtiss III); Perry *et al.* (1981) Infec. Immun. 32:1295; Powell *et al.* (1988) Appl. Environ. Microbiol. 54:655; Somkuti *et al.* (1987) Proc. 4th Evr. Cong. Biotechnology 1:412, *Streptococcus*).

Transformation methods for yeast hosts are well-known in the art, and typically include either the transformation of spheroplasts or of intact yeast cells treated with alkali cations. Electroporation is another means for transforming yeast hosts. See for example, Methods in Enzymology, Volume 194, 1991, "Guide to Yeast Genetics and

may be biosynthetic genes, such as those in the histidine, tryptophan, and leucine biosynthetic pathways. Thus, when leucine is absent from the media, for example, only the cells with a biosynthetic gene in the leucine pathway will survive.

An origin of replication may be needed for the expression vector to replicate in the host cell. Certain origins of replication enable an expression vector to be reproduced at a high copy number in the presence of the appropriate proteins within the cell. Examples of origins are the 2m and autonomously replicating sequences, which are effective in yeast; and the viral T-antigen, effective in COS-7 cells.

Expression vectors may be integrated into the host cell genome or remain autonomous within the cell. Polynucleotide sequences homologous to sequences within the host cell genome may be needed to integrate the expression cassette. The homologous sequences do not always need to be linked to the expression vector to be effective. For example, expression vectors can integrate into the CHO genome via an unattached dihydrofolate reductase gene. In yeast, it is more advantageous if the homologous sequences flank the expression cassette. Particularly useful homologous yeast genome sequences are those disclosed in PCT WO90/01800, and the HIS4 gene sequences, described in Genbank, accession no. J01331.

The choice of promoter, terminator, and other optional elements of an expression vector will also depend on the host cell chosen. The invention is not dependent on the host cell selected. Convenience and the level of protein expression will dictate the optimal host cell. A variety of hosts for expression are known in the art and available from the American Type Culture Collection (ATCC). Bacterial hosts suitable for expression include, without limitation: *Campylobacter*, *Bacillus*, *Escherichia*, *Lactobacillus*, *Pseudomonas*, *Staphylococcus*, and *Streptococcus*. Yeast hosts from the following genera may be utilized: *Candida*, *Hansenula*, *Kluyveromyces*, *Pichia*, *Saccharomyces*, *Schizosaccharomyces*, and *Yarrowia*. Immortalized mammalian host cells include but are not limited to CHO cells, HeLa cells, baby hamster kidney (BHK) cells, monkey kidney cells (COS), human hepatocellular carcinoma cells (*e.g.*, Hep G2), and other cell lines. A number of insect cell hosts are also available for expression of heterologous proteins: *Aedes aegypti*, *Bombyx mori*, *Drosophila melanogaster*, and *Spodoptera frugiperda* (PCT WO 89/046699; Carbonell *et al.*, (1985) J. Virol. 56:153;

be expressed if operably linked to a suitable promoter. A variety of suitable insect cells and viruses are known and include following without limitation.

Insect cells from any order of the Class Insecta can be grown in the media of this invention. The orders Diptera and Lepidoptera are preferred. Example of insect species are listed in Weiss *et al.*, "Cell Culture Methods for Large-Scale Propagation of Baculoviruses," in Granados *et al.* (eds.), The Biology of Baculoviruses: Vol. II Practical Application for Insect Control, pp. 63-87 at p. 64 (1987). Insect cell lines derived from the following insects are exemplary: *Carpocapsa pomeonella* (preferably, cell line CP-128); *Trichoplusia ni* (preferably, cell line TN-368); *Autograph californica*; *Spodoptera frugiperda* (preferably, cell line Sf9); *Lymantria dispar*; *Mamestra brassicae*; *Aedes albopictus*; *Orgyia pseudotsugata*; *Neodiprio sertifer*; *Aedes aegypti*; *Antheraea eucalypti*; *Gnorimoschema operceullela*; *Galleria mellonella*; *Spodoptera littolaris*; *Blatella germanic*; *Drosophila melanogaster*; *Heliothis zea*; *Spodoptera exigua*; *Rachiplusia ou*; *Plodia interpunctella*; *Amsaeta moorei*; *Agrotis c-nigrum*, *Adoxophyes orana*; *Agrotis segetum*; *Bombyx mori*; *Hyponomeuta malinellu;*, *Colias eurytheme*; *Anticarsia germmetalia*; *Apanteles melanoscelu*; *Arctia caja*; and *Porthetria dispar*. Preferred insect cell lines are from *Spodoptera frugiperda*, and especially preferred is cell line Sf9. The Sf9 cell line used in the examples herein was obtained from Max D. Summers (Texas A & M University, College Station, Texas, 77843, U.S.A.) Other *S. frugiperda* cell lines, such as IPL-Sf-21AE III, are described in Vaughn *et al.*, In Vitro 13: 213-217 (1977).

The insect cell lines of this invention are suitable for the reproduction of numerous insect-pathogenic viruses such as parvoviruses, pox viruses, baculoviruses and rhabdcoviruses, of which nucleopolyhedrosis viruses (NPV) and granulosis viruses (GV) from the group of baculoviruses are preferred. Further preferred are NPV viruses such as those from Autographa spp., Spodoptera spp., Trichoplusia spp., Rachiplusia spp., Gallerai spp., and Lymantria spp. More preferred are baculovirus strain *Autographa californica* NPV (AcNPV), *Rachiplusia ou* NPV, *Galleria mellonella* NPV, and any plaque purified strains of AcNPV, such as E2, R9, S1, M3, characterized and described by Smith *et al.*, J Virol 30: 828-838 (1979); Smith *et al.*, J Virol 33: 311-319 (1980); and Smith *et al.*, Virol 89: 517-527 (1978).

Molecular Biology." Transformation procedures usually vary with the yeast species to be transformed. See *e.g.*, (Kurtz *et al.* (1986) Mol. Cell. Biol. 6:142; Kunze *et al.* (1985) J. Basic Microbiol. 25:141; *Candida*); (Gleeson *et al.* (1986) J. Gen. Microbiol. 132:3459; Roggenkamp *et al.* (1986) Mol. Gen. Genet. 202:302; *Hansenula*); (Das *et al.* (1984) J. Bacteriol. 158:1165; De Louvencourt *et al.* (1983) J. Bacteriol. 154:1165; Van den Berg *et al.* (1990) Bio/Technology 8:135; *Kluyveromyces*); (Cregg *et al.* (1985) Mol. Cell. Biol. 5:3376; Kunze *et al.* (1985) J. Basic Microbiol. 25:141; U.S. Patent Nos. 4,837,148 and 4,929,555; *Pichia*); (Hinnen *et al.* (1978) Proc. Natl. Acad. Sci. USA 75;1929; Ito *et al.* (1983) J. Bacteriol. 153:163 *Saccharomyces*); (Beach and Nurse (1981) Nature 300:706; *Schizosaccharomyces*); (Davidow *et al.* (1985) Curr. Genet. 10:39; Gaillardin *et al.* (1985) Curr. Genet. 10:49; *Yarrowia*).

Methods for introducing heterologous polynucleotides into mammalian cells are known in the art and include viral infection, dextran-mediated transfection, calcium phosphate precipitation, polybrene mediated transfection, protoplast fusion, electroporation, encapsulation of the polynucleotide(s) in liposomes, and direct microinjection of the DNA into nuclei.

The method for construction of an expression vector for transformation of insect cells for expression of recombinant protein herein is slightly different than that generally applicable to the construction of a bacterial expression vector, a yeast expression vector, or a mammalian expression vector. In an embodiment of the present invention, a baculovirus vector is constructed in accordance with techniques that are known in the art, for example, as described in Kitts *et al.*, BioTechniques 14: 810-817 (1993), Smith *et al.*, Mol. Cell. Biol. 3: 2156 (1983), and Luckow and Summer, Virol. 17: 31 (1989). In one embodiment of the present invention, a baculovirus expression vector is constructed substantially in accordance to Summers and Smith, Texas Agricultural Experiment Station Bulletin No. 1555 (1987). Moreover, materials and methods for baculovirus/insect cell expression systems are commercially available in kit form, for example, the MaxBac® kit from Invitrogen (San Diego, CA).

Also, methods for introducing heterologous DNA into an insect host cell are known in the art. For example, an insect cell can be infected with a virus containing a coding sequence. When the virus is replicating in the infected cell, the polypeptide will

Antibodies to the desired proteins can be used in Western blots to determine with greater sensitivity if protein was expressed.

Examples

The examples presented below are provided as a further guide to the practitioner of ordinary skill in the art, and are not to be construed as limiting the invention in any way.

Example 1

Construction of *Pichia pastoris* autonomously replicating vector containing *P. pastoris* HIS4 gene as a selectable marker and an expression cassette containing a *P. acaciae* killer toxin leader and IGF-1 gene.

A. CLONING

I. Killer Toxin Leader Fragment

Construction of fragment by annealing of synthetic oligomers.

Synthesis of oligomers with a phosphate group attached or kinase.

The sequence of the oligomers, KAC 34, KAC 37, KAC 39, KAC 59, KAC 60, and KAC 61 are attached.

Ligation of fragment and base vector for sequencing and ease of handling

Fragment: as described above

Base vector: pLITMUS28 available from New England Biolabs (Beverly, Massachusetts, USA)

II. IGF-1 Fragment

Isolation: from a yeast strain with an integrated vector. Sequence of gene attached.

III. Overlapping PCR

Construction of a single fragment containing the leader sequence and IGF-1 gene.

PCR #1:

Reaction Mix:

4 µL of IGF-1 gene fragment for a total of 10 ng

Typically, insect cells *Spodoptera frugiperda* type 9 (SF9) are infected with baculovirus strain *Autographa californica* NPV (AcNPV) containing a coding sequence. Such a baculovirus is produced by homologous recombination between a transfer vector containing the coding sequence and baculovirus sequences and a genomic baculovirus DNA. Preferably, the genomic baculovirus DNA is linearized and contains a disfunctional essential gene. The transfer vector, preferably, contains the nucleotide sequences needed to restore the disfunctional gene and a baculovirus polyhedrin promoter and terminator operably linked to the polynucleotides of the instant invention. (See Kitts *et al.*, BioTechniques 14(5): 810-817 (1993).

The transfer vector and linearized baculovirus genome are transfected into SF9 insect cells, and the resulting viruses probably containing the desired coding sequence. Without a functional essential gene the baculovirus genome cannot produce a viable virus. Thus, the viable viruses from the transfection most likely contain the coding sequence and the needed essential gene sequences from the transfer vector. Further, lack of occlusion bodies in the infected cells are another verification that the coding sequence was incorporated into the baculovirus genome.

The essential gene and the polyhedrin gene flank each other in the baculovirus genome. The coding sequence in the transfer vector is flanked at its 5' with the essential gene sequences and the polyhedrin promoter and at its 3' with the polyhedrin terminator. Thus, when the desired recombination event occurs the coding sequence displaces the baculovirus polyhedrin gene. Such baculoviruses without a polyhedrin gene will not produce occlusion bodies in the infected cells. Of course, another means for determining if coding sequence was incorporated into the baculovirus genome is to sequence the recombinant baculovirus genomic DNA. Alternatively, expression of the desired polypeptide by cells infected with the recombinant baculovirus is another verification means.

Once transformed the host cells can be used to produce either polynucleotides of the instant invention or express the desired polypeptide.

Simple gel electrophoresis techniques can be used to detect expression of the desired polypeptide. For example, media from a host cell without an expression vector can be compared to media from host cell with the desired vector. Polyacrylamide gel electrophoresis ("PAGE") can be used to determine if any proteins were expressed.

19 cycles: 97°C for 1 minute and 72°C for 1 minute

PCR #3

Reaction Mix:

5 µL of result PCR#2

5 µL of 1:100 dilution of result of PCR#1

10 µL of 10X Pfu DNA Polymerase buffer available from Stratagene (La Jolla, California, USA)

4 µL of 2 mM dNTP

1 µL of 2.5 units/µL of Pfu DNA Polymerase available from Stratagene(La Jolla, California)

2 µL of oligomer KAC74 for a total of 2 picomoles

2 µL of oligomer KAC57 for a total of 2 picomoles

71 µL of water.

Temperature Cycle:

5 cycles: 97°C for 1 minute, 58°C for 1 minute, and 72°C for 1 minute

24 cycles: 97°C for 1 minute and 72°C for 1 minute.

PCR#4

Reaction Mix:

1 µL of results of PCR#3

10 µL of KAC 74 for a total of 10 picomoles

30 µL of KAC 57 for a total of 10 picomoles

10 µL of 10X PCR buffer (same as used in PCR#2)

2 µL of 2 mM dNTP

5 µL of 0.5 units/µL of taq DNA Polymerase available from Boehringer Mannheim catalog number 1 146 173 (Indianapolis, Indiana, USA)

42 µL of water.

Temperature Cycle:

10 µL of Pfu DNA Polymerase buffer available from Stratagene (La Jolla, California, USA)

4 µL of a 2 mM dNTP

20 µL of oligomer KAC58 for a total of 20 picomoles

20 µL of oligomer KAC57 for a total of 20 picomoles

1 µL of 2.5 units/µL Pfu DNA Polymerase available from Stratagene (La Jolla, California, USA)

41 µL of water

Temperature cycle:

5 cycles: 97°C for 1 minute, 43°C for 1 minute, and 72°C for 1 minute

24 cycles: 97°C for 1 minute and 72°C for 1 minute

PCR#2

Reaction Mix

1 µL of Killer toxin fragment in pLITMUS28 for a total of 10 ng

10 µL of 10X PCR buffer

2 µL of 2 mM dNTP

10 µL of oligomer KAC74 for a total of 10 picomoles

10 µL of oligomer KAC75 for a total of 10 picomoles

0.5 µL of 5 units/µL taq DNA Polymerase available from Boehringer Mannheim catalog number 1 146 173 (Indianapolis, Indiana, USA)

66.5 µL of H_2O

10X PCR buffer

0.25 M Tris-HCl, pH 8.3

0.015 M $MgCl_2$ in 0.0015 M EDTA

0.25 M KCl

0.5% Tween 20

Temperature cycle:

5 cycles: 97°C for 1 minute, 63°C for 1 minute, and 72°C for 1 minute

[redacted] 63-MER

[redacted]CC

MW AMMONIUM SALT	20303.6
MOLAR EXTINCTION AT 260nm	586200
MICROGRAMS PER OD260nm	34.6359
PICOMOLES PER OD260nm	1705.9
BASE COMPOSITION: ACGT	1312929
Td (blot) .1M Na+	71
Tm @ .1M Na+, .000001M probe	84

7.2

11.3 pm/l

KAC39 35-MER

AGTCAAGCATTAACAGCGGTTAAAAATATTAAACG

MW AMMONIUM SALT	11366.2
MOLAR EXTINCTION AT 260nm	366500
MICROGRAMS PER OD260nm	31.0129
PICOMOLES PER OD260nm	2728.51
BASE COMPOSITION: ACGT	16568
Td (blot) .1M Na+	64
Tm @ .1M Na+, .000001M probe	73

6.2

17.4 pm/l

Enter extinction of Z. <RET> for 0: ?

KAC59 38-MER

ZTAAAGTAGCTAAAAATAATAAGACTATAATTAACATG

MW AMMONIUM SALT	12024.7
MOLAR EXTINCTION AT 260nm	463200
MICROGRAMS PER OD260nm	25.9601
PICOMOLES PER OD260nm	2158.89
BASE COMPOSITION: ACGT	203410
MIXED BASES: YRNMKSWHBVDXZ	0000000000001
Z=phosphate	
Td (blot) .1M Na+	59
Tm @ .1M Na+, .000001M probe	63

12.3 OD/mL

26.6 pmol/l

In KAC60

DEFINE Z

Enter extinction of Z. <RET> for 0: ?

KAC60 49-MER

ZTCTAGTATATCCAAAAAATACATCTCCGCTACAATCGAGGGAATTAGC

MW AMMONIUM SALT	15495.2
MOLAR EXTINCTION AT 260nm	543500
MICROGRAMS PER OD260nm	28.51
PICOMOLES PER OD260nm	1839.93
BASE COMPOSITION: ACGT	1811712
MIXED BASES: YRNMKSWHBVDXZ	0000000000001
Z=phosphate	
Td (blot) .1M Na+	70
Tm @ .1M Na+, .000001M probe	80

9.7 OD/mL

17.9 pmol/l

In KAC61

DEFINE Z

Enter extinction of Z. <RET> for 0: ?

24 cycles: 97°C for 1 minute and 72°C for 1 minute

Ligation of PCR#4 fragment to a shuttle vector for sequencing

Fragment: 1 µL of result of PCR#4

Base vector: 2 µL of pCRII from Invitrogen (San Diego, California, USA)

Ligase: 1 µL from Invitrogen (San Diego, California, USA) kit # 45-0046

10X Ligase buffer: 1 µL from Invitrogen (San Diego, California, USA) kit # 45-0046

Water: 5 µL

KAC34 65-MER
AATTCATGTTAATTATAGTCTTATTATTTTTAGCTACTTTAGCTAATTCCCTCGATTGTAGCGGA
MW AMMONIUM SALT 21012.1
MOLAR EXTINCTION AT 260nm 627400
MICROGRAMS PER OD260nm 33.4908
PICOMOLES PER OD260nm 1593.88
BASE COMPOSITION: ACGT 1710929
Td (blot) .1M Na+ 69
Tm @ .1M Na+. .000001M probe 80

~~KAC36~~ 37-MER
[redacted]CATGAAAT
MW AMMONIUM SALT 12024.7
MOLAR EXTINCTION AT 260nm 400800
MICROGRAMS PER OD260nm 30.0018
PICOMOLES PER OD260nm 2495.01
BASE COMPOSITION: ACGT 203410
Td (blot) .1M Na+ 59
Tm @ .1M Na+. .000001M probe 67

~~KAC38~~ 48-MER
[redacted]GT
MW AMMONIUM SALT 15495.2
MOLAR EXTINCTION AT 260nm 477200
MICROGRAMS PER OD260nm 32.471
PICOMOLES PER OD260nm 2095.56
BASE COMPOSITION: ACGT 1811712
Td (blot) .1M Na+ 70
Tm @ .1M Na+. .000001M probe 79

~~KAC37~~ KAC37 48-MER
GATGTATTTTTTGGATATACTAGAGGAGACAAAACAGATGTTCATAAA
MW AMMONIUM SALT 15684.3
MOLAR EXTINCTION AT 260nm 500800
MICROGRAMS PER OD260nm 31.3185
PICOMOLES PER OD260nm 1996.81
BASE COMPOSITION: ACGT 1941015
Td (blot) .1M Na+ 66
Tm @ .1M Na+. .000001M probe 76

```
Enter probe conc. for Tm in uM. <RET> for default 1uM:  ?
KAC58          39-MER
GTTAAAAATATTAAAAGAGGACCGGAGACGCTCTGCGGG
MW AMMONIUM SALT                12768.2
MOLAR EXTINCTION AT 260nm       399300
MICROGRAMS PER OD260nm          31.9764
PICOMOLES PER OD260nm           2504.38
BASE COMPOSITION: ACGT          146127
Td (blot) .1M Na+               71
Tm @ .1M Na+. .000001M probe                    81
```

KAC61 64-MER
ZCATGCGTTTAATATTTTTAACCGCTGTTAATGCTTGACTTTTATGAACATCTGTTTTGTCTCC
MW AMMONIUM SALT 20303.6
MOLAR EXTINCTION AT 260nm 644800
MICROGRAMS PER OD260nm 31.4882
PICOMOLES PER OD260nm 1550.87
BASE COMPOSITION: ACGT 1312929
MIXED BASES: YRNMKSWHBVDXZ 0000000000001
Z=phosphate
Td (blot) .1M Na+ 71
Tm @ .1M Na+. .000001M probe 85

14.3 OD/mL 222 pmol/λ

Enter probe conc. for Tm ; uM. <RET> for default 1uM ?

KAC74 44-MER
CGCGAATTCGACAGAATGTTAATTATAGTCTTATTATTTTTAGC
MW AMMONIUM SALT 14240.3
MOLAR EXTINCTION AT 260nm 432500
MICROGRAMS PER OD260nm 32.9255
PICOMOLES PER OD260nm 2312.14
BASE COMPOSITION: ACGT 136718
Td (blot) .1M Na+ 66
Tm @ .1M Na+. .000001M probe 75

735 OD/mL 164 pm/λ

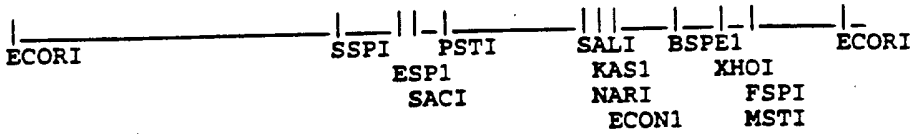
KAC75 42-MER
CGTCTCCGGTCCTCTTTTAATATTTTTAACCGCTGTTAATGC
MW AMMONIUM SALT 13443.7
MOLAR EXTINCTION AT 260nm 378000
MICROGRAMS PER OD260nm 35.5653
PICOMOLES PER OD260nm 2645.5
BASE COMPOSITION: ACGT 711618
Td (blot) .1M Na+ 70
Tm @ .1M Na+. .000001M probe 80

700 OD/mL 164 pm/λ

KAC57 31-MER
CGCGAATTCGGTTCCTTATCAAGCTGACTTG
MW AMMONIUM SALT 9971.18
MOLAR EXTINCTION AT 260nm 286200
MICROGRAMS PER OD260nm 34.8399
PICOMOLES PER OD260nm 3494.06
BASE COMPOSITION: ACGT 68710
Td (blot) .1M Na+ 69
Tm @ .1M Na+. .000001M probe 79

204 pm/λ 1/20

Argument Map in DNA Strand paKT.IGF
from the '/arp/lib/6mers' file.
Translation shown at frame 1.




```
                    MetLeuIleIleValLeuLeuPheLeuAlaThrLeuAlaAsnSer
  1 CGCGAATTCGACAGAATGTTAATTATAGTCTTATTATTTTTAGCTACTTTAGCTAATTCC
    GCGCTTAAGCTGTCTTACAATTAATATCAGAATAATAAAAATCGATGAAATCGATTAAGG
    ^
    4 ECORI,

    LeuAspCysSerGlyAspValPhePheGlyTyrThrArgGlyAspLysThrAspValHis
 61 CTCGATTGTAGCGGAGATGTATTTTTTGGATATACTAGAGGAGACAAAACAGATGTTCAT
    GAGCTAACATCGCCTCTACATAAAAAACCTATATGATCTCCTCTGTTTTGTCTACAAGTA

    LysSerGlnAlaLeuThrAlaValLysAsnIleLysArgGlyProGluThrLeuCysGly
121 AAAAGTCAAGCATTAACAGCGGTTAAAAATATTAAAAGAGGACCGGAGACGCTCTGCGGG
    TTTTCAGTTCGTAATTGTCGCCAATTTTTATAATTTTCTCCTGGCCTCTGCGAGACGCCC
                                  ^
    148 SSPI,

    AlaGluLeuValAspAlaLeuGlnPheValCysGlyAspArgGlyPheTyrPheAsnLys
181 GCTGAGCTCGTGGATGCTCTGCAGTTCGTGTGTGGAGACAGGGGCTTTTATTTCAACAAG
    CGACTCGAGCACCTACGAGACGTCAAGCACACACCTCTGTCCCCGAAAATAAAGTTGTTC
    ^  ^           ^
    181 ESP1, 184 SACI, 199 PSTI,

    ProThrGlyTyrGlySerSerSerArgArgAlaProGlnThrGlyIleValAspGluCys
241 CCCACAGGGTATGGCTCCAGCAGTCGACGGGCGCCTCAGACAGGCATCGTGGATGAGTGC
    GGGTGTCCCATACCGAGGTCGTCAGCTGCCCGCGGAGTCTGTCCGTAGCACCTACTCACG
                          ^     ^  ^
    263 SALI, 270 KAS1 NARI, 274 ECON1,

    CysPheArgSerCysAspLeuArgArgLeuGluMetTyrCysAlaProLeuLysProAla
301 TGCTTCCGGAGCTGTGATCTAAGGAGGCTCGAGATGTATTGCGCACCCCTCAAGCCTGCC
    ACGAAGGCCTCGACACTAGATTCCTCCGAGCTCTACATAACGCGTGGGGAGTTCGGACGG
    ^                     ^              ^
    305 BSPE1, 328 XHOI, 340 FSPI MSTI,

    LysSerAlaOP OC GlyThrGluPheArg
361 AAGTCAGCTTGATAAGGAACCGAATTCCGC
    TTCAGTCGAACTATTCCTTGGCTTAAGGCG
                       ^
    382 ECORI,
```

expression cassette =
a killer toxin leader (P. pastoris)
and IGF-I gene

Ligation into Expression

Base Vector: 2 μL of pHIL-A1, linear with EcoRI ends and dephosphorylated

Fragment: 2 μL of EcoRI from pCRII with expression cassette containing a killer toxin leader fragment with IGF-1 gene

Ligase: 1μL of T4 DNA ligase available from Boerhinger Mannheim

10X Ligase buffer: 1 μL available from Boerhinger Mannheim

Water: 4 μL

Verification that expression cassette in correct orientation.

Nucleotide sequence of the 3'AOX1 transcriptional termination region

Eco RI
GAA TTC CCC TTA GAC ATG ACT GTT CCT CAG TTC AAG TTG GGC ACT TAC GAG AAG

ACC GGT CTT GCT AGA TTC *TAA* TCA AGA GGA TGT CAG AAT GCC ATT TGC CTG AGA

GAT GCA GGC TTC ATT TTT GAT ACT TTT TTA TTT GTA ACC TAT ATA GTA TAG GAT

3' end of AOX1 mRNA
TTT TTT TGT **CA**T TTT GTT TCT TCT CGT ACG AGC TTG CTC CTG ATC AGC CTA TCT

CGC AGC TGA TGA ATA TCT TGT GGT AGG GGT TTG GGA AAA TCA TTC GAG TTT GAT

GTT TTT CTT GGT ATT TCC CAC TCC TCT TCA GAG TAC AGA AGA TTA AGT GAG ACG

TTC GTT TGT GCA AGC TT
Hind III

NOTE: The 3'AOX1 contains a small stretch (22 amino acids long) of carboxy terminal alcohol oxidase coding sequences upto translational stop codon TAA (italicized and underlined). The 3' end of AOX1 mRNA is in bold and also underlined (**A**).

Nucleotide Sequence of the PARS1 (164 bp) Taq I fragment in pHIL-A1

Nru I
TCG AGA TAA GCT GGG GGA ACA TTC GCG AAA ATG AAA CAA GTC GGC TGT TAT

Bgl II
AGT ATA TTT ATT ATA ATA TTG AAA GAT CTC AAA AGA CTA CTT ATT TTT GAA

Hinc II
TGA ACC AAG TAT GAA ATC AAC CTA TTT GGG GTT GAC CAA AAT AAG TAA ATA

TTA ATT GTC GA

Description of pHIL-A1

Plasmid pHIL-A1 is and E. coli - P. pastoris shuttle vector, with sequences for selection and autonomous replication in each host. One component of the plasmid is a modified portion of plasmid pBR322 containing the ampicillin resistance gene and the origin of replication (ori). The regions between nucleotides 1,100 and 2,485 of pBR322 and between NaeI sites 404 and 932 were deleted to eliminate "poison sequences" and the Sal I site, respectively.

The DNA elements comprising the rest of the plasmid are derived from the genome of P. pastoris, except for short regions of pBR322 used to the link the yeast elements. The yeast elements are as follows: proceeding clockwise:

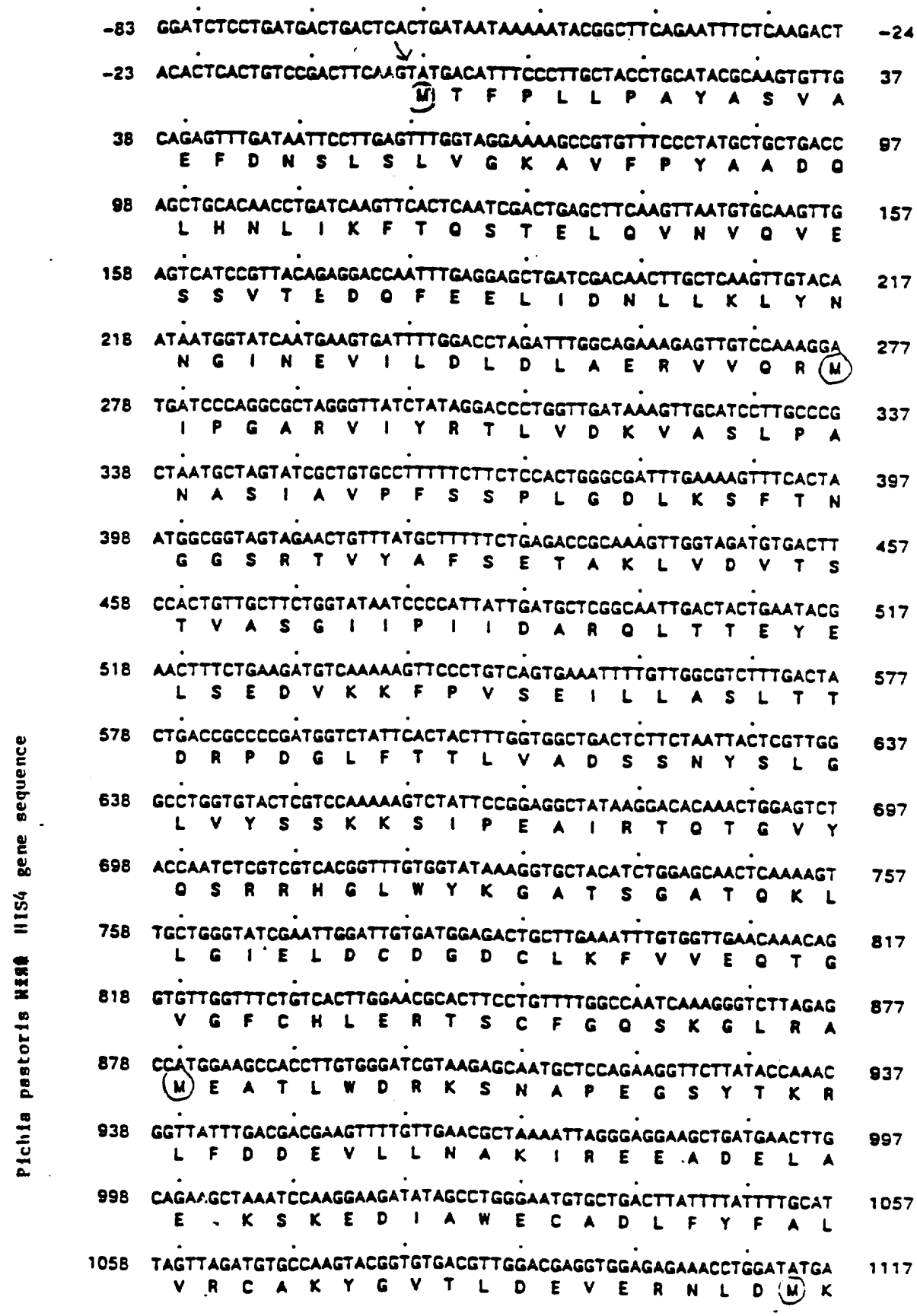
1. 3' AOX1, alcohol oxidase, approximately 300 bp segment of the AO terminating sequence. Sequence attached.

2. 5' AOX1, approximately 750 bp segment of the alcohol oxidase promoter. The alcohol oxidase coding sequences following the A of the ATG initiating methionine codon have been removed, and a synthetic linker used to generate a unique EcoRI site, as described for pHIL-D1 (available from Invitrogen, San Diego, California, USA). Sequence attached.

3. PARS1, approximately 190 bp segment of P. pastoris autonomous replication sequence. Sequence attached.

4. HIS4, approximately 2.8 kb segment of *P. pastoris* histidinol dehydrogenase gene to complement the defective his4 gene in *P. pastoris*, strain GS115. Sequence attached.

Pichia pastoris HIS4 gene sequence



Nucleotide sequence of 5'AOX1 (1018 Nucleotides)

[5'AOX1 begins at (Hind III/Hinc II) in pHIL-Ds and pHIL-S1]
(Hind III/Hinc II) junction
AAG CTG ACT CAT GTT GGT ATT GTG AAA TAG ACG CAG ATC GGG AAC ACT GAA AAA
Bgl II
TAA CAG TTA TTA TTC GAG ATC TAA CAT CCA AAG ACG AAA GGT TGA ATG AAA CCT

TTT TGC CAT CCG ACA TCC ACA GGT CCA TTC TCA CAC ATA AGT GCC AAA CGC AAC

AGG AGG GGA TAC ACT AGC AGC AGA CCG TTG CAA ACG CAG GAC CTC CAC TCC TCT

TCT CCT CAA CAC CCA CTT TTG CCA TCG AAA AAC CAG CCC AGT TAT TGG GCT TGA
(5' AOX1 of pHIL-A1 begins from Sst I)
SstI
TTG GAG CTC GCT CAT TCC AAT TCC TTC TAT TAG GCT ACT AAC ACC ATG ACT TTA

TTA GCC TGT CTA TCC TGG CCC CCC TGG CGA GGT TCA TGT TTG TTT ATT TCC GAA

TGC AAC AAG CTC CGC ATT ACA CCC GAA CAT CAC TCC AGA TGA GGG CTT TCT GAG

TGT GGG GTC AAA TAG TTT CAT GTT CCC CAA ATG GCC CAA AAC TGA CAG TTT AAA

CGC TGT CTT GGA ACC TAA TAT GAC AAA AGC GTG ATC TCA TCC AAG ATG AAC TAA

GTT TGG TTC GTT GAA ATG CTA ACG GCC AGT TGG TCA AAA AGA AAC TTC CAA AAG

TCG GCA TAC CGT TTG TCT TGT TTG GTA TTG ATT GAC GAA TGC TCA AAA ATA ATC

TCA TTA ATG CTT AGC GCA GTC TCT CTA TCG CTT CTG AAC CCC GGT GCA CCT GTG

CCG AAA CGC AAA TGG GGA AAC ACC CGC TTT TTG GAT GAT TAT GCA TTG TCT CCA

CAT TGT ATG CTT CCA AGA TTC TGG TGG GAA TAC TGC TGA TAG CCT AAC GTT CAT

GAT CAA AAT TTA ACT GTT CTA ACC CCT ACT TGA CAG CAA TAT ATA AAC AGA AGG
^ * ^

AAG CTG CCC TGT CTT AAA CCT TTT TTT TTA TCA TCA TTA TTA GCT TAC TTT CAT

AAT TGC GAC TGG TTC CAA TTG ACA AGC TTT TGA TTT TAA CGA CTT TTA ACG ACA
AsuII/FspII EcoRI
ACT TGA GAA GAT CAA AAA ACA ACT AAT TAT TCG AAA CGA *GGA ATT C*

Note: Nucleotides added immediately following the "A" of the translation initiation codon to create Eco RI site is italicized. The 5' end of the alcohol oxidase mRNA have been denoted as a major species(*) or minor species(^) of mRNA transcripts.

```
2378  ACACTGCAACCTTCCAGAAGTTCATCACTTCACAAGACGTAACTCCTGAGGGACTGAAAC  2437
        T  A  T  F  Q  K  F  I  T  S  Q  D  V  T  P  E  G  L  K  H

2438  ATATTGGCCAAGCAGTGATGGATCTGGCTGCTGTTGAAGGTCTAGATGCTCACCGCAATG  2497
        I  G  Q  A  V  M  D  L  A  A  V  E  G  L  D  A  H  R  N  A

2498  CTGTTAAGGTTCGTATGGAGAAACTGGGACTTATTTAATTATTTAGAGATTTTAACTTAC  2557
        V  K  V  R  M  E  K  L  G  L  I  *

2558  ATTTAGATTCGATAGATCC  2576
```

```
1118  AGTCCCTAAAGGTCACTAGAAGGAAAGGAGATGCCAAGCCAGGATACACCAAGGAACAAC  1177
        S  L  K  V  T  R  R  K  G  D  A  K  P  G  Y  T  K  E  Q  P

1178  CTAAAGAAGAATCCAAACCTAAAGAAGTCCCTTCTGAAGGTCGTATTGAATTGTGCAAAA  1237
        K  E  E  S  K  P  K  E  V  P  S  E  G  R  I  E  L  C  K  I

                 StuI
1238  TTGACGTTTCTAAGGCCTCCTCACAAGAAATTGAAGATGCCCTTCGTCGTCCTATCCAGA  1297
        D  V  S  K  A  S  S  Q  E  I  E  D  A  L  R  R  P  I  Q  K

                                        SalI
1298  AAACGGAACAGATTATGGAATTAGTCAAACCAATTGTCGACAATGTTCGTCAAAATGGTG  1357
        T  E  Q  I  M  E  L  V  K  P  I  V  D  N  V  R  Q  N  G  D

1358  ACAAAGCCCTTTTAGAACTAACTGCCAAGTTTGATGGAGTCGCTTTGAAGACACCTGTGT  1417
        K  A  L  L  E  L  T  A  K  F  D  G  V  A  L  K  T  P  V  L

1418  TAGAAGCTCCTTTCCCAGAGGAACTTATGCAATTGCCAGATAACGTTAAGAGAGCCATTG  1477
        E  A  P  F  P  E  E  L  M  Q  L  P  D  N  V  K  R  A  I  D

1478  ATCTCTCTATAGATAACGTCAGGAAATTCCATGAAGCTCAACTAACGGAGACGTTGCAAG  1537
        L  S  I  D  N  V  R  K  F  H  E  A  Q  L  T  E  T  L  Q  V

1538  TTGAGACTTGCCCTGGTGTAGTCTGCTCTCGTTTTGCAAGACCTATTGAGAAAGTTGGCC  1597
        E  T  C  P  G  V  V  C  S  R  F  A  R  P  I  E  K  V  G  L

1598  TCTATATTCCTGGTGGAACCGCAATTCTGCCTTCCACTTCCCTGATGCTGGGTGTTCCTG  1657
        Y  I  P  G  G  T  A  I  L  P  S  T  S  L  M  L  G  V  P  A

                                                          KpnI
1658  CCAAAGTTGCTGGTTGCAAAGAAATTGTTTTTGCATCTCCACCTAAGAAGGATGGTACCC  1717
        K  V  A  G  C  K  E  I  V  F  A  S  P  P  K  K  D  G  T  L

1718  TTACCCCAGAAGTCATCTACGTTGCCCACAAGGTTGGTGCTAAGTGTATCGTGCTAGCAG  1777
        T  P  E  V  I  Y  V  A  H  K  V  G  A  K  C  I  V  L  A  G

1778  GAGGCGCCCAGGCAGTAGCTGCTATGGCTTACGGAACAGAAACTGTTCCTAAGTGTGACA  1837
        G  A  Q  A  V  A  A  M  A  Y  G  T  E  T  V  P  K  C  D  K

              SspI
1838  AAATATTTGGTCCAGGAAACCAGTTCGTTACTGCTGCCAAGATGATGGTTCAAAATGACA  1897
        I  F  G  P  G  N  Q  F  V  T  A  A  K  M  M  V  Q  N  D  T

1898  CATCAGCCCTGTGTAGTATTGACATGCCTGCTGGGCCTTCTGAAGTTCTAGTTATTGCTG  1957
        S  A  L  C  S  I  D  M  P  A  G  P  S  E  V  L  V  I  A  D

1958  ATAAATACGCTGATCCAGATTTCGTTGCCTCAGACCTTCTGTCTCAAGCTGAACATGGTA  2017
        K  Y  A  D  P  D  F  V  A  S  D  L  L  S  Q  A  E  H  G  I

2018  TTGATTCCCAGGTGATTCTGTTGGCTGTCGATATGACAGACAAGGAGCTTGCCAGAATTG  2077
        D  S  Q  V  I  L  L  A  V  D  M  T  D  K  E  L  A  R  I  E

2078  AAGATGCTGTTCACAACCAAGCTGTGCAGTTGCCAAGGGTTGAAATTGTACGCAAGTGTA  2137
        D  A  V  H  N  Q  A  V  Q  L  P  R  V  E  I  V  R  K  C  I

2138  TTGCACACTCTACAACCCTATCGGTTGCAACCTACGAGCAGGCTTTGGAAATGTCCAATC  2197
        A  H  S  T  T  L  S  V  A  T  Y  E  Q  A  L  E  M  S  N  Q

2198  AGTACGCTCCTGAACACTTGATCCTGCAAATCGAGAATGCTTCTTCTTATGTTGATCAAG  2257
        Y  A  P  E  H  L  I  L  Q  I  E  N  A  S  S  Y  V  D  Q  V

2258  TACAACACGCTGGATCTGTGTTTGTTGGTGCCTACTCTCCAGAGAGTTGTGGAGATTACT  2317
        Q  H  A  G  S  V  F  V  G  A  Y  S  P  E  S  C  G  D  Y  S

2318  CCTCCGGTACCAACCACACTTTGCCAACGTACGGATATGCCCGTCAATACAGCGGAGTTA  2377
        S  G  T  N  H  T  L  P  T  Y  G  Y  A  R  Q  Y  S  G  V  N
```

Media: 25 mL of MGY

MGY=

13.g/L of Yeast Nitrogen Base without amino acids, available from Difco (Michigan, Detroit, USA)

400 µg/L biotin

1% (v/v) glycerol

0.1% leucine

0.1% lysine

0.1% tryptophan

0.1% adenine

0.1% uracil

Inoculum:

250 µL of the preculture

Temperature

30°C

Aeration:

275 rpm

Time:

Approximately 48 hours or 5-10 OD_{600}

Harvest:

4000 rpm for 10 minutes

Wash, Resuspension, and Dilution of cells:

Use MM media for all.

MM=

13.g/L of Yeast Nitrogen Base without amino acids, available from Difco (Michigan, Detroit, USA)

400 µg/L biotin

0.5% (v/v) methanol

B. TRANSFORMATION

I. YEAST STRAIN

P. pastoris, GS115 available from Invitrogen (San Diego, California, USA), also available from the USDA, Northern Regional Research Center in Peoria, Illinois, under the accession number NRRL Y-15851

or

P. pastoris SMD1163

II. ELECTROPORATION

Cells: Cells from preculture at approximately 16 OD_{600}. 1:20 dilution into 10% glycerol with water. 50µL of cells in 10% glycerol with water for electroporation.

Equipment:

BioLab Pulse Controller and BioLab Gene Pulser

Pulse:

2.0 Kilovolts

25 µFD

200 ohms

Time Constant:

5 Milliseconds

Selection:

Cells on minimal medium in minus histidine with glucose

C. EXPRESSION

I. Precultures

Media:

Minimal his minus media plus glucose

Inoculum:

One transformed colony

Temperature:

30°C

Time: until culture is saturated

II. Expression Cultures

STAGE 2 CLONING:

Fragment:

BglII-BamHI fragment from Resulting vector 1.

Base vector:

The entire resulting vector 1, linear with BamHI ends

Resulting vector 2:

pALIGF1-2 with two expression cassettes each with

One AO1 gene promoter

One *P. acaciae* killer toxin leader

One IGF-1 gene

One AO1 gene terminator.

STAGE 3:

Fragment:

BglII-BamHI fragment from Resulting vector 2, pALIGF1-2.

Base Vector:

The entire pALIGF1-2, linear with BamHI ends

Resulting Vector:

pALIGF1-3 with four expression cassettes with

One AO1 gene promoter

One *P.acaciae* killer toxin leader

One IGF-1 gene

One AO1 gene terminator.

STAGE 4:

Fragment:

BglII-BamHI fragment from Resulting vector 2, pALIGF1-2.

Base Vector:

The entire pALIGF1-3, linear with BamHI ends

Resulting Vector:

pALIGF1-4 with six expression cassettes with

0.1% leucine

0.1% lysine

0.1% tryptophan

0.1% adenine

0.1% uracil

Resuspension: with approximately 5 mL

Dilution: to approximately 3 OD_{600}.

Temperature:

30°C

Aeration:

275 rpm

Time:

Approximately 96 hours

Example 2

Construction of *Pichia pastoris* integrating vector containing *P. pastoris* HIS4 gene as a selectable marker and multiple copies of an expression cassette containing the *P. acaciae* leader and IGF1 gene.

STAGE 1 CLONING:

Starting vector:

pA0815 as described by Brierley *et al.*, U.S. Patent No. 5,324,639 and available from Invitrogen (San Diego, California, USA). The vector contains a unique EcoRI restriction site flanked by the *P. pastoris* alcohol oxidase 1 ("AO1") gene promoter and terminator.

Insert Fragment:

Described above in Example 1 comprising EcoRI restriction ends.

Resulting vector 1:

One AO1 gene promoter

One *P. acaciae* killer toxin leader

One IGF-1 gene

One AO1 gene terminator.

```
Enter extinction of X. <RET> for 0:        ?

KAC117         69-MER
XAATTCGACAGAATGTTAATTATAGTCTTATTATTTTTAGCTACTTTAGCTAATTCCCTCGATTGTAGC
MW AMMONIUM SALT                  21978.8
MOLAR EXTINCTION AT 260nm         729800
MICROGRAMS PER OD260nm            30.1162
PICOMOLES PER OD260nm             1370.24
BASE COMPOSITION: ACGT            1911929
MIXED BASES: YRNMKSWHBVDXZ        0000000000010
X=PHOSPHATE
Td (blot) .1M Na+                 70
Tm @ .1M Na+. .000001M probe                  80
In  KAC118
DEFINE X
Enter extinction of X. <RET> for 0:        ?

KAC118         55-MER
XGGAGATGTATTTTTTGGATATACTAGAGGAGACAAAACAGATGTTCATAAAAGT
MW AMMONIUM SALT                  17704.7
MOLAR EXTINCTION AT 260nm         641700
MICROGRAMS PER OD260nm            27.5903
PICOMOLES PER OD260nm             1558.36
BASE COMPOSITION: ACGT            2141316
MIXED BASES: YRNMKSWHBVDXZ        0000000000010
X=PHOSPHATE
Td (blot) .1M Na+                 69
Tm @ .1M Na+. .000001M probe                  78
In  KAC119
DEFINE X
Enter extinction of X. <RET> for 0:        ?

KAC119         49-MER
XGGTCCAGAAACCTTGTGTGGTGCTGAATTGGTCGATGCTTTGCAATTC
MW AMMONIUM SALT                  15619.2
MOLAR EXTINCTION AT 260nm         505400
MICROGRAMS PER OD260nm            30.9046
PICOMOLES PER OD260nm             1978.63
BASE COMPOSITION: ACGT            991416
MIXED BASES: YRNMKSWHBVDXZ        0000000000010
X=PHOSPHATE
Td (blot) .1M Na+                 74
Tm @ .1M Na+. .000001M probe                  90
In  KAC120
DEFINE X
Enter extinction of X. <RET> for 0:        ?

KAC120         56-MER
XGTTTGTGGTGACAGAGGTTTCTACTTCAACAAGCCAACCGGTTACGGTTCTTCTT
MW AMMONIUM SALT                  17815.8
MOLAR EXTINCTION AT 260nm         573000
MICROGRAMS PER OD260nm            31.0921
PICOMOLES PER OD260nm             1745.2
BASE COMPOSITION: ACGT            11121319
MIXED BASES: YRNMKSWHBVDXZ        0000000000010
X=PHOSPHATE
Td (blot) .1M Na+                 74
Tm @ .1M Na+. .000001M probe                  88
```

One AO1 gene promoter

One *P.acaciae* killer toxin leader

One IGF-1 gene

One AO1 gene terminator.

TRANSFORMATION:

Yeast:

P. pastoris, GS115, available from Invitrogen (San Diego, California, USA) or *P. pastoris*, SMD1163.

Electroporation: Same as Example 1.

EXPRESSION: Same as Example 1.

Example 3

Construction of three vectors, pKK, pKG, and pKGK.

These vectors comprise the IGF-1 coding sequence. Further, the vectors comprise killer toxin leader sequences as described below:

(The asterisks indicate the amino acid positions that are different from the native killer toxin sequence.)

pKG = killer toxin leader with glycosylation site, sequence below:

Met-Leu-Ile-Ile-Val-Leu-leu-Phe-Leu-Ala-Thr-Leu-Ala-Asn-Ser-Leu-Asp-Cys-Ser-Gly-Asp-Val-Phe-Phe-Gly-Tyr-Thr-Arg-Gly-Asp-Lys-Thr-Asp-Val-His-Lys-Ser-Gln-Asn*-Leu-Thr-Ala-Val-Lys-Asn-Ile-Lys-Arg-

pKK = killer toxin with KEX2 site, sequence below:

Met-Leu-Ile-Ile-Val-Leu-leu-Phe-Leu-Ala-Thr-Leu-Ala-Asn-Ser-Leu-Asp-Cys-Ser-Gly-Asp-Val-Phe-Phe-Gly-Tyr-Thr-Arg-Gly-Asp-Lys-Thr-Asp-Val-His-Lys-Ser-Gln-Ala-Leu-Thr-Ala-Val-Pro*-Met*-Tyr*-Lys-Arg

pKGK = killer toxin with glycosylation site and KEX2 site, sequence below:

Met-Leu-Ile-Ile-Val-Leu-leu-Phe-Leu-Ala-Thr-Leu-Ala-Asn-Ser-Leu-Asp-Cys-Ser-Gly-Asp-Val-Phe-Phe-Gly-Tyr-Thr-Arg-Gly-Asp-Lys-Thr-Asp-Val-His-Lys-Ser-Gln-Asn*-Leu-Thr-Ala-Val-Pro*-Met*-Tyr*-Lys-Arg

A. ANNEALING OLIGOMERS

Construction of killer toxin fragments by annealing of synthetic oligomers. The DNA oligomers comprises a 5' phosphate group.

```
Enter extinction of X. <RET> for 0:        ?

KAC129          58-MER
XGGTTTCTGGACCTCTTTTAATATTTTTAACCGCTGTTAAGTTTTGACTTTTATGAAC
MW AMMONIUM SALT                  18412.2
MOLAR EXTINCTION AT 260nm         589800
MICROGRAMS PER OD260nm            31.2178
PICOMOLES PER OD260nm             1695.49
BASE COMPOSITION: ACGT            129927
MIXED BASES: YRNMKSWHBVDXZ        0000000000010
X=PHOSPHATE
Td (blot) .1M Na+                 69
Tm @ .1M Na+. .000001M probe                    81
In  KAC130
DEFINE X
Enter extinction of X. <RET> for 0:        ?

KAC130          34-MER
XCAAGCATTAACAGCGGTTCCAATGTACAAAAGA
MW AMMONIUM SALT                  10675.7
MOLAR EXTINCTION AT 260nm         388600
MICROGRAMS PER OD260nm            27.4723
PICOMOLES PER OD260nm             2573.34
BASE COMPOSITION: ACGT            14766
MIXED BASES: YRNMKSWHBVDXZ        0000000000010
X=PHOSPHATE
Td (blot) .1M Na+                 66
Tm @ .1M Na+. .000001M probe                    77
In  KAC131
DEFINE X
Enter extinction of X. <RET> for 0:        ?

KAC131          58-MER
XGGTTTCTGGACCTCTTTTGTACATTGGAACCGCTGTTAATGCTTGACTTTTATGAAC
MW AMMONIUM SALT                  18448.2
MOLAR EXTINCTION AT 260nm         588900
MICROGRAMS PER OD260nm            31.3266
PICOMOLES PER OD260nm             1698.08
BASE COMPOSITION: ACGT            11111223
MIXED BASES: YRNMKSWHBVDXZ        0000000000010
X=PHOSPHATE
Td (blot) .1M Na+                 73
Tm @ .1M Na+. .000001M probe                    86
In  KAC132
DEFINE X
Enter extinction of X. <RET> for 0:        ?

KAC132          34-MER
XCAAAACTTAACAGCGGTTCCAATGTACAAAAGA
MW AMMONIUM SALT                  10659.7
MOLAR EXTINCTION AT 260nm         392500
MICROGRAMS PER OD260nm            27.1585
PICOMOLES PER OD260nm             2547.77
BASE COMPOSITION: ACGT            15756
MIXED BASES: YRNMKSWHBVDXZ        0000000000010
X=PHOSPHATE
Td (blot) .1M Na+                 65
Tm @ .1M Na+. .000001M probe                    75
```

```
Enter extinction of X. <RET> for 0:        ?

KAC121          46-MER
XCTAGAAGAGCTCCACAAACCGGTATCGTTGACGAATGTTGTTTCA
MW AMMONIUM SALT                14576.5
MOLAR EXTINCTION AT 260nm       493600
MICROGRAMS PER OD260nm          29.531
PICOMOLES PER OD260nm           2025.93
BASE COMPOSITION: ACGT          13101012
MIXED BASES: YRNMKSWHBVDXZ      0000000000010
X=PHOSPHATE
Td (blot) .1M Na+               72
Tm @ .1M Na+. .000001M probe                    85
In  KAC122
DEFINE X
Enter extinction of X. <RET> for 0:        ?

KAC122          51-MER
XAATTAGCTAAAGTAGCTAAAAATAATAAGACTATAATTAACATTCTGTCG
MW AMMONIUM SALT                16232.7
MOLAR EXTINCTION AT 260nm       598100
MICROGRAMS PER OD260nm          27.1405
PICOMOLES PER OD260nm           1671.96
BASE COMPOSITION: ACGT          236615
MIXED BASES: YRNMKSWHBVDXZ      0000000000010
X=PHOSPHATE
Td (blot) .1M Na+               65
Tm @ .1M Na+. .000001M probe                    71
In  KAC123
DEFINE X
Enter extinction of X. <RET> for 0:        ?

KAC123          57-MER
XATCTGTTTTGTCTCCTCTAGTATATCCAAAAAATACATCTCCGCTACAATCGAGGG
MW AMMONIUM SALT                18027
MOLAR EXTINCTION AT 260nm       598600
MICROGRAMS PER OD260nm          30.1152
PICOMOLES PER OD260nm           1670.56
BASE COMPOSITION: ACGT          1614818
MIXED BASES: YRNMKSWHBVDXZ      0000000000010
X=PHOSPHATE
Td (blot) .1M Na+               72
Tm @ .1M Na+. .000001M probe                    83
In  KAC124
DEFINE X
Enter extinction of X. <RET> for 0:        ?

KAC124          49-MER
XGTCACCACAAACGAATTGCAAAGCATCGACCAATTCAGCACCACACAA
MW AMMONIUM SALT                15413.1
MOLAR EXTINCTION AT 260nm       547600
MICROGRAMS PER OD260nm          28.1467
PICOMOLES PER OD260nm           1826.15
BASE COMPOSITION: ACGT          201666
MIXED BASES: YRNMKSWHBVDXZ      0000000000010
X=PHOSPHATE
Td (blot) .1M Na+               73
Tm @ .1M Na+. .000001M probe                    91
```

```
Enter extinction of X. <RET> for 0:        ?

KAC133          58-MER
XGGTTTCTGGACCTCTTTTGTACATTGGAACCGCTGTTAAGTTTTGACTTTTATGAAC
MW AMMONIUM SALT                    18463.2
MOLAR EXTINCTION AT 260nm           590200
MICROGRAMS PER OD260nm              31.283
PICOMOLES PER OD260nm               1694.34
BASE COMPOSITION: ACGT              11101224
MIXED BASES: YRNMKSWHBVDXZ          0000000000010
X=PHOSPHATE
Td (blot) .1M Na+                   72
Tm @ .1M Na+. .000001M probe                    85
```

```
Enter extinction of X. <RET> for 0:        ?

KAC125         56-MER
XGAGCTCTTCTAGAAGAAGAACCGTAACCGGTTGGCTTGTTGAAGTAGAAACCTCT
MW AMMONIUM SALT                 17900.8
MOLAR EXTINCTION AT 260nm        610600
MICROGRAMS PER OD260nm           29.3168
PICOMOLES PER OD260nm            1637.73
BASE COMPOSITION: ACGT           16111414
MIXED BASES: YRNMKSWHBVDXZ       0000000000010
X=PHOSPHATE
Td (blot) .1M Na+                74
Tm @ .1M Na+. .000001M probe                  85
In  KAC126
DEFINE X
Enter extinction of X. <RET> for 0:        ?

KAC126         38-MER
XGATCTGAAACAACATTCGTCAACGATACCGGTTTGTG
MW AMMONIUM SALT                 11968.6
MOLAR EXTINCTION AT 260nm        407600
MICROGRAMS PER OD260nm           29.3636
PICOMOLES PER OD260nm            2453.39
BASE COMPOSITION: ACGT           118810
MIXED BASES: YRNMKSWHBVDXZ       0000000000010
X=PHOSPHATE
Td (blot) .1M Na+                69
Tm @ .1M Na+. .000001M probe                  81
In  KAC127
DEFINE X
Enter extinction of X. <RET> for 0:        ?

KAC127         70-MER
XAATTCTTATCAAGCAGACTTAGCTGGCTTCAATGGAGCACAGTACATTTCCAATCTTCTCAAGTCACAG
MW AMMONIUM SALT                 22278
MOLAR EXTINCTION AT 260nm        749700
MICROGRAMS PER OD260nm           29.7159
PICOMOLES PER OD260nm            1333.87
BASE COMPOSITION: ACGT           2117112C
MIXED BASES: YRNMKSWHBVDXZ       0000000000010
X=PHOSPHATE
Td (blot) .1M Na+                74
Tm @ .1M Na+. .000001M probe                  89
In  KAC128
DEFINE X
Enter extinction of X. <RET> for 0:        ?

KAC128         34-MER
XCAAAACTTAACAGCGGTTAAAAATATTAAAAGA
MW AMMONIUM SALT                 10706.8
MOLAR EXTINCTION AT 260nm        413700
MICROGRAMS PER OD260nm           25.8805
PICOMOLES PER OD260nm            2417.21
BASE COMPOSITION: ACGT           18447
MIXED BASES: YRNMKSWHBVDXZ       0000000000010
X=PHOSPHATE
Td (blot) .1M Na+                60
Tm @ .1M Na+. .000001M probe                  68
```

2μL of a fragment from pHIL-A1 vector digested with EcoRI and phosphotased for a total of 30ng (plasmid described above)

1μL of T4 DNA ligase for a total of 1 one unit

q.s. to final volume of 20 μL with water.

Either 1μL or 5μL of the above three oligomer mixtures were used for the ligation. Incubated overnight at 4°C.

TRANSFORMATION INTO YEAST HOST

The vectors were transformed into *Pichia pastoris* yeast host, SMD1163, available from Invitrogen (San Diego, California, United States).

Before transformation, 3 mL of YEPD was inoculated with *P. pastoris* SMD1163. This culture was incubated overnight. Ten microliters of this overnight culture was used to inoculate 100 mL of YEPD.

These cells were grown to an OD_{650} of 0.78. Then, the cells were entrifuged for 5 minutes at 3.5K. Cell pellets were resuspended in 100 mL sterile water. The cells were centrifuged for 5 minutes at 3.5K. The cell pellets were resuspended in 8 mL of 0.1 M lithium acetate.

The cells were incubated in the lithium acetate for 30 minutes at 30°C while shaking. Next, the cells were centrifuged again for 5 minutes in a table top centrifuge and the cell pellets were resuspended in 8 mL of 0.1 M lithium acetate.

Ten microliters of either pKK, pKG, or pKGK, ** pg, was added to 100 μL of the cells in 0.8 M lithium acetate. The cells and DNA were incubated for 30 minutes at 30°C.

Next, 0.6 mL of 40% PEG 3550, was added to the cells and DNA. The mixture was vortexed, and the mixture was incubated for 60 minutes at 30°C.

Then, the cells were centrifuged for 30 minutes and the cell pellets were resuspended in 60 μL of water. The mixture was plated on histidine minus, yeast minimal media.

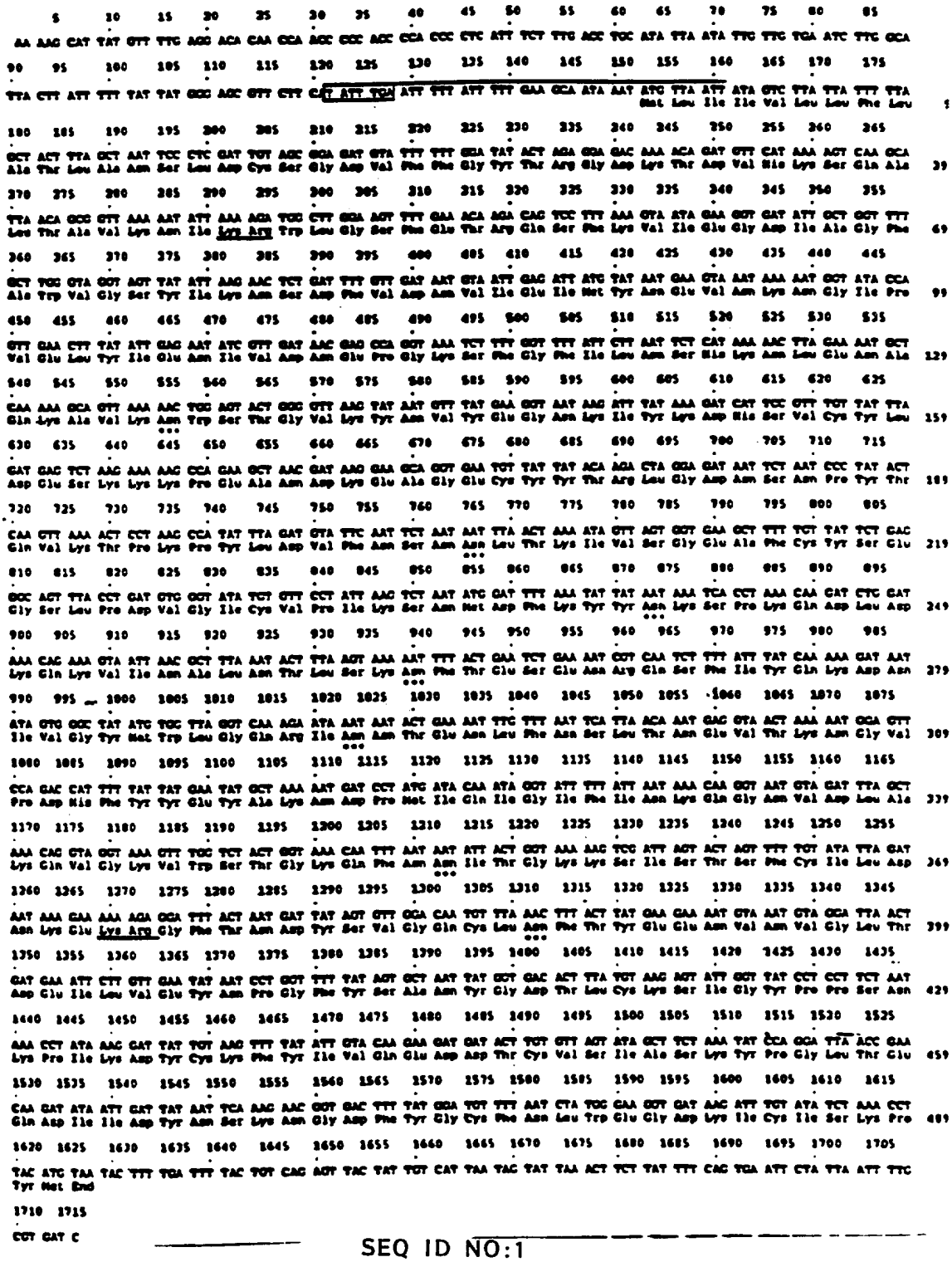
Oligomers were diluted to a concentration of 100 picomoles in final volume of 500µl with 5µl polyA (1mg/mL) and 50 µl of 10X ligase buffer. Ligase buffer purchased from New England Biolabs, Beverly, Massachusetts, United States.

	pKK	pKG.	pKGK	pmoles/µl
KAC117	4.8 µL	4.8	4.8	20.7
KAC122	2.9	2.9	2.9	34.3
KAC118	4.5	4.5	4.5	22.1
KAC123	5	5	5.	20.0
KAC119	3.8	3.8	3.8	26.3
KAC124	4.6	4.6	4.6	21.5
KAC120	3.5	3.5	3.5	28.4
KAC125	4	4	4	24.9
KAC121	5.4	5.4	5.4	18.4
KAC126	2.1	2.1	2.1	46.6
KAC109	1	1	1	
KAC127	9	9	9	11.1
KAC128		3.6		27.8
KAC129		3.6		27.6
KAC130	3.5			28.8
KAC131	4.1			24.4
KAC132			2.2	44.3
KAC133			4.4	22.9

Oligomer mixtures were incubated for two minutes in boiling water. The mixture was cooled to room temperature (~3 hours) with a little ice in bath, which was removed from the heat source.

LIGATION INTO YEAST VECTOR:

The following is the ligation mixture used to construct the leader/coding sequences: 2µL of 10X ligation solution with ATP



SEQ ID NO:1

Deposit Information:

The following materials were deposited with the American Type Culture Collection:

Name	Deposit Date	Accession No.
Escherichia coli XL1 Blue pHIL-A1 paKT	26 Sept 1995	69903

The above materials have been deposited with the American Type Culture Collection, Rockville, Maryland, under the accession numbers indicated. This deposit will be maintained under the terms of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for purposes of Patent Procedure. The deposits will be maintained for a period of 30 years following issuance of this patent, or for the enforceable life of the patent, whichever is greater. Upon issuance of the patent, the deposits will be available to the public from the ATCC without restriction.

These deposits are provided merely as convenience to those of skill in the art, and are not an admission that a deposit is required under 35 U.S.C. §112. The sequence of the polynucleotides contained within the deposited materials, as well as the amino acid sequence of the polypeptides encoded thereby, are incorporated herein by reference and are controlling in the event of any conflict with the written description of sequences herein. A license may be required to make, use, or sell the deposited materials, and no such license is granted hereby.

5. The polynucleotide of claim 3, wherein the yeast cell is selected from the group consisting of *Pichia pastoris*, *Saccharomyces cerevisiae*, *Kluyveromyces lactis*, and *Hansenula polymorpha*.

6. The polynucleotide molecule of claim 1, wherein the host cell is a protease A deficient cell.

7. The polynucleotide molecule of claim 1, wherein the host cell is a protease B deficient cell.

8. The polynucleotide molecule of claim 1, wherein the host cell is a protease A and protease B deficient cell.

9. The polynucleotide molecule of claim 1, wherein the leader sequence comprises a signal peptide sequence and a peptidase cleavage site that comprises dibasic amino acid residues.

10. The polynucleotide molecule of claim 1, wherein the amino acid sequence comprises at least about 80% sequence identity to the leader sequence of *Picahia acaciae* killer toxin.

11. The polynucleotide molecule of claim 1, wherein the amino acid sequence comprises at least about 90% sequence identity to the leader sequence of *Picahia acaciae* killer toxin.

What is claimed:

1. A polynucleotide molecule comprising a first nucleotide sequence that encodes at least a fragment of a leader sequence and a second nucleotide sequence that encodes a polypeptide heterologous to the leader sequence,

wherein the leader sequence fragment is sufficient for secretion and comprises an amino acid sequence that comprises at least about 70% sequence identity to the leader sequence of *Picahia acaciae* killer toxin,

wherein the heterologous polypeptide is not naturally contiguous to the leader sequence, and

wherein upon expression of the polynucleotide molecule in a host cell suitable for expression thereof, the heterologous polypeptide is produced that is free of additional N-terminal amino acids.

2. The polynucleotide molecule of claim 1, wherein the host cell is an eukaryotic cell.

3. The polynucleotide molecule of claim 2, wherein the eukaryotic cell is a yeast cell.

4. The polynucleotide molecule of claim 3, wherein the yeast cell belongs to a genus that is selected from the genera consisting of Pichia, Saccharomyces, Kluyveromyces, and Hansenula.

19. A method of producing the polynucleotide molecule of claim 1, comprising linking together in proper reading frame the first nucleotide sequence and the second nucleotide sequence.

20. A method of producing the vector of claim 15, wherein the vector is capable of independent replication, comprising linking together in proper reading frame a replicon and a polynucleotide molecule,

wherein the polynucleotide molecule comprises a first nucleotide sequence that encodes at least a fragment of a leader sequence and a second nucleotide sequence that encodes a polypeptide heterologous to the leader sequence,

wherein the leader sequence fragment is sufficient for secretion and comprises an amino acid sequence that comprises at least about 70% sequence identity to the leader sequence of *Picahia acaciae* killer toxin,

wherein the heterologous polypeptide is not naturally contiguous to the leader sequence and wherein upon expression of the polynucleotide molecule in a host cell suitable for expression thereof, the heterologous polypeptide is produced that is free of additional N-terminal amino acids.

22. The holst cell of claim 16, wherein the cell is selected from the group consisting of a prokaryotic cell and an eukarytoic cell.

23. The host cell of claim 22, wherein the host cell is an eukaryotic cell and the eukaryotic cell is selected from the group consisting of a yeast cell, an avian cell, an insect cell, and a mammalian cell.

24. The host cell of claim 23, wherein the cell is a yeast cell, and the yeast cell is selected from the genera consisting of Pichia, Saccharomyces, and Kluyveromyces.

12. The polynucleotide molecule of claim 1, wherein the amino acid sequence comprises at least about 95% sequence identity to the leader sequence of *Picahia acaciae* killer toxin.

13. The polynucleotide of claim 1, wherein the polynucleotide is DNA.

14. The polynucleotide of claim 1, wherein the polynucleotide is RNA.

15. An expression vector comprising the polynucleotide of claim 1, wherein the vector is capable of independent replication or integration into a host genome.

16. A host cell comprising the polynucleotide of claim 1, wherein the host cell is capable of effecting transcription and translation of the polynucleotide to produce the heterologous polypeptide.

17. A host cell comprising the vector of claim 15, wherein the host cell is capable of effecting transcription and translation of the polynucleotide to produce the heterologous polypeptide.

18. A method of producing a polypeptide culturing the host cell of claim 16 and obtaining the polypeptide molecule therefrom.

30. A method of producing a polypeptide encoded by a polynucleotide comprising

(a) transforming a host cell with the polynucleotide,

(b) allowing the expression thereof to produce the polypeptide and

(c) obtaining the polypeptide therefrom,

wherein the polynucleotide molecule comprises a first nucleotide sequence that encodes at least a fragment of a leader sequence and a second nucleotide sequence that encodes a polypeptide heterologous to the leader sequence,

wherein the leader sequence fragment is sufficient for secretion and comprises an amino acid sequence that comprises at least about 70% sequence identity to the leader sequence of *Picahia acaciae* killer toxin,

wherein the heterologous polypeptide is not naturally contiguous to the leader sequence, and

wherein upon expression of the polynucleotide molecule in a host cell suitable for expression thereof, the heterologous polypeptide is produced that is free of additional N-terminal amino acids.

25. The host cell of claim24, wherein the yeast cell is selected from the group consisting of *Pichia pastoris*, Saccharomyces cerevisiae, and Kluyveromyces lactis.

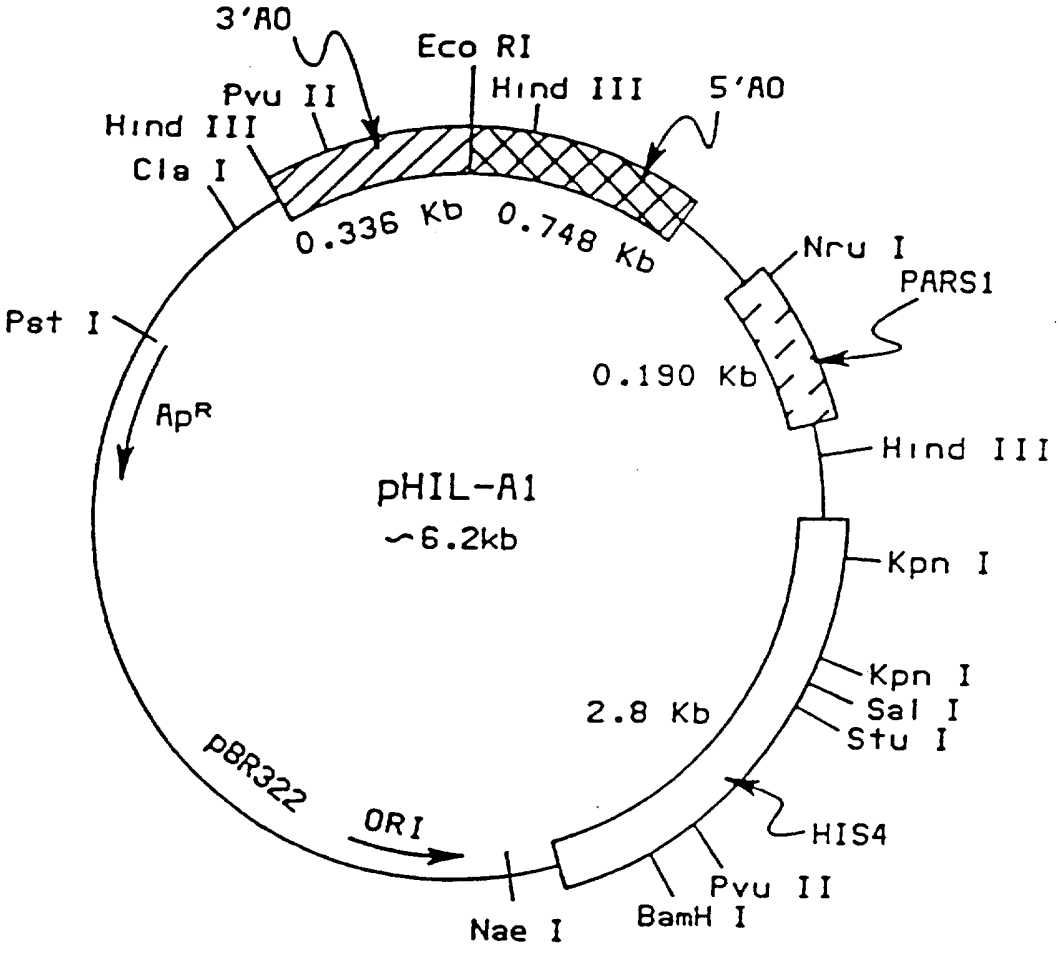
26. The polynucleotide of claim 1, wherein the heterologous polypeptide is a mammalian polypeptide.

27. The polynucleotide of claim 26, wherein the mammalian polypeptide is a human polypeptide.

28. The polynucleotide of claim 1, wherein the polypeptide is one selected from the group consisting of a hormone, a growth factor, a cytokine, a haematopoietic factor, an immunoglobulin, an enzyme, a repressor, a cell differentiation factor, a binding protein, and a transcription factor.

29. The polynucleotide of claim 1, wherein the polypeptide is one selected from the group consisting of growth hormone, luteinizing hormone, thyroid stimulating hormone, oxytocin, insulin, vasopresin, renin, calcitonin, follicle stimulating hormone, prolactin, insulin-like growth factor (IGF-I, IGF-II), an IGF-binding protein, epidermal growth factor (EGF), platelet derived growth factor (PDGF), keratinocyte growth factor (KGF), fibroblast growth factor (FGF), nerve growth factor (NGF), TGF-beta, vascular endothelial cell growth factor (VEGF), erythropoietin (EPO), colony stimulating factor (CSF), interferon, endorphin, enkaphalin, dynorphin and an active fragment thereof.

FIGURE 1


3'AO
Eco RI
Pvu II
Hind III
5'AO
Hind III
Cla I
0.336 Kb
0.748 Kb
Nru I
PARS1
Pst I
0.190 Kb
Ap^R
Hind III
pHIL-A1
~6.2kb
Kpn I
Kpn I
Sal I
Stu I
2.8 Kb
pBR322
ORI
HIS4
Pvu II
BamH I
Nae I