(12) STANDARD PATENT (11) Application No. **AU 2006303870 B2**
(19) AUSTRALIAN PATENT OFFICE

(54) Title
Polyoleosins

(51) International Patent Classification(s)
C12N 15/29 (2006.01) ***C12N 15/62*** (2006.01)
C12N 15/54 (2006.01)

(21) Application No: **2006303870** (22) Date of Filing: **2006.10.16**

(87) WIPO No: **WO07/045019**

(30) Priority Data

(31) Number	(32) Date	(33) Country
2005906364	**2005.11.16**	**AU**
2005905787	**2005.10.19**	**AU**

(43) Publication Date: **2007.04.26**
(44) Accepted Journal Date: **2012.12.06**

(71) Applicant(s)
Agriculture Victoria Services Pty Ltd;AgResearch Limited

(72) Inventor(s)
Scott, Richard William;Arcus, Vickery Laurence;Roberts, Nicholas John

(74) Agent / Attorney
Jones Tulloch, 484 Glenferrie Road, Hawthorn, VIC, 3122

(56) Related Art
US 2003/0167524 A1
WO 2004/113376 A1

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
26 April 2007 (26.04.2007)

PCT

(10) International Publication Number
WO 2007/045019 A3

(51) International Patent Classification:
C12N 15/29 (2006.01) *C12N 15/62* (2006.01)
C12N 15/54 (2006.01)

(21) International Application Number:
PCT/AU2006/001528

(22) International Filing Date: 16 October 2006 (16.10.2006)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
2005905787 19 October 2005 (19.10.2005) AU
2005906364 16 November 2005 (16.11.2005) AU

(71) Applicants *(for all designated States except US)*: **AGRICULTURE VICTORIA SERVICES PTY LTD** [AU/AU]; 475 Mickleham Road, Attwood, Victoria 3049 (AU). **AGRESEARCH LIMITED** [NZ/NZ]; 5th Floor, Tower Block Ruakura Research, Centre East Street, Hamilton 3240 (NZ).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **SCOTT, Richard William** [NZ/NZ]; 25 Colombo Street, Palmerston North 4410 (NZ). **ARCUS, Vickery Laurence** [NZ/NZ]; 38 Mcnicol Street, Fairfield, Hamilton 3214 (NZ). **ROBERTS, Nicholas John** [NZ/NZ]; 97 Mckay Line, Rd 7, Feilding 4777 (NZ).

(74) Agent: **FREEHILLS PATENT & TRADE MARK ATTORNEYS**; 101 Collins Street, Melbourne, Victoria 3000 (AU).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declaration under Rule 4.17:
— *of inventorship (Rule 4.17(iv))*

Published:
— *with international search report*

(88) Date of publication of the international search report:
12 July 2007

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: POLYOLEOSINS

(57) Abstract: The present invention relates to constructs including one or more nucleic acids encoding two or more oleosin repeat units, and methods of use thereof. The present invention also relates to recombinant polypeptides including two or more oleosin repeat units, and methods of use thereof.

WO 2007/045019 A3

POLYOLEOSINS

The present invention relates to constructs including one or more nucleic acids encoding two or more oleosin repeat units, and methods of use thereof. The present invention also relates to recombinant polypeptides including two or more oleosin repeat units, and methods of use thereof.

Triacylglycerol

Most plants synthesise and store significant amounts of triacylglycerol (TAG) only in developing embryos and pollen cells where it is subsequently utilised to provide catabolizable energy during germination and pollen tube growth. Dicotyledonous plants can accumulate up to approximately 60% of their seed weight as TAG. Ordinarily, this level is considerably lower in the monocotyledonous seeds where the main form of energy storage is carbohydrates (e.g., starch).

The only committed step in TAG biosynthesis is the last one, i.e., the addition of a third fatty acid to an existing diacylglycerol, thus generating TAG. In plants this step is performed by one of three enzymes including: acyl CoA: diacylglycerol acyltransferase (DGAT1); an unrelated acyl CoA: diacylglycerol acyl transferase (DGAT2); and phospholipid: diacylglycerol acyltransferase (PDAT) (Bouvier-Nave et al., 2000; Dahlqvist et al., 2000; Lardizabal et al., 2001; Zou et al., 1999).

Oleosin

Oleosins are specific plant proteins usually found only in seeds and pollen. Their function is to stop oil bodies coalescing during seed and pollen dehiscence. In nature, TAG produced in seeds and pollen form micelles encapsulated by a spherical phospholipid monolayer and one or several species of oleosin proteins. Oil bodies in fruit tissues (such as olives and avocados) do not contain oleosins.

The physiochemical properties of the major oleosins is relatively conserved between plants and is characterised by the following:

- 15-25kDa protein corresponding to approximately 140-230 amino acid residues.
- The protein sequence can be divided almost equally along its length into 4 parts that correspond to a N-terminal hydrophilic region; two centre hydrophobic regions (joined by a proline knot or knob) and a C-terminal hydrophilic region.
- The topology of oleosin is attributed to its physiochemical properties, which includes a folded hydrophobic core flanked by hydrophilic domains (Figure 1). This arrangement confers an amphipathic nature to oleosin resulting in the hydrophobic domain being embedded in the phospholipid monolayer (Tzen et al., 1992) while the flanking hydrophilic domains are exposed to the aqueous environment of the cytoplasm (Figure 2).

Oil Bodies

An oil body that is produced in seed or pollen consists of a droplet of TAG surrounded by a monolayer of phospholipid where the hydrophobic acyl moieties of the phospholipids interact with the encapsulated TAG and the hydrophilic head groups face the cytoplasm. Oil bodies are naturally produced in the seeds and pollen of many plants. Oil bodies can also be generated artificially by combining oleosins, triacylglycerides and phospholipids (Peng et al., 2004).

The outside of the oil body is coated with oleosins which are orientated with their central hydrophobic amino acid domains protruding through the phospholipid monolayer and into the TAG core of the oil body (Figures 2-6).

The size of the oil body may be regulated by oleosin imparting a defined curvature; the curvature is dependent on the oleosin::oil ratio as well as the type of oleosin and oil.

Biohydrogenation

It has been demonstrated that the lipid profile of ruminant animal feed in turn influences the lipid profile of meat and dairy products (Demeyer and Doreau, 1999). Different plants have different lipid profiles; by selectively feeding animals only plants with the desired lipid profile it is possible to positively influence the lipid profile of downstream

2006303870 16 Oct 2012

meat and dairy products. In ruminants the final lipid make up of the meat and milk is not only influenced by the dietary lipids but is also heavily influenced by biohydrogenation. Biohydrogenation is the hydrogenation of non-reduced compounds (such as unsaturated fats) by the biota present in the rumen.

Emulsions

Emulsions are produced when one or more liquids that are immiscible (usually due to different polarities and thus different hydrophobicities) in another liquid are uniformly suspended within that liquid, for example when oil droplets are dispersed uniformly in water or water droplets dispersed uniformly in oil. Generation of a relatively stable emulsion requires the use of an emulsifier, which lowers the interfacial tension between the liquids. The stability of an emulsion is generally a measure of what conditions and for how long the uniform dispersion persists. Emulsifiers are commonly used in the food and cosmetic industry; as such the emulsifiers need to have high emulsion stability as well as be safe for consumption and topical application.

It is an object of the present invention to overcome, or at least alleviate, one or more of the difficulties or deficiencies associated with the prior art.

In one aspect the present invention provides a construct including one or more nucleic acids encoding two or more oleosin repeat units. For expression in vegetative plant tissue preferably the construct further includes a nucleic acid encoding a diacylglycerol acryltransferase.

In a preferred embodiment, the present invention provides a construct including one or more nucleic acids encoding three or more oleosin repeat units.

Preferably, the nucleic acids encoding oleosin repeat units include a nucleotide sequence selected from the group consisting of Seq ID Nos. 29-43, 50, 50-57, 64-73 and 84, and functionally active fragments and variants thereof.

2006303870 16 Oct 2012

The term “oleosin” as used herein includes any functionally active plant oleosin, including functionally active fragments and variants thereof (e.g., L-oleosin, H-oleosin, steroleosin and caoleosin).

In a preferred embodiment, the oleosin repeat unit may be from white clover or sesame seed, or may be a synthetic or recombinant version of an oleosin repeat unit from white clover or sesame seed.

While Applicants have exemplified the invention using white clover and sesame seed oleosins, the invention is not limited thereto and any functionally active plant oleosin sequence may be used in the constructs of the present invention. The oleosin sequence may be naturally occurring, recombinant or synthetic.

By "repeat units" is meant multiple copies of nucleotide sequences encoding oleosin within a single polynucleotide, or multiple copies of amino acid sequences encoding oleosin within a single polypeptide, which may or may not contain intervening nucleotide or amino acid sequences. The repeat units may be tandem repeats. The repeat units may be homo- or hetero- repeats (homo or heteromeric).

The oleosin repeat units may be linked either directly, for example by direct fusion of the repeats, or by using linking sequences.

In a preferred embodiment of this aspect of the invention linker sequences may be included between the oleosin repeats to facilitate rotation of the oleosin hydrophilic domains relative to each other to form the correct topology. Such linker sequences also facilitate recombinatorial subcloning of sequences encoding desired peptides in between the oleosin repeats, and also facilitate chemical or enzymatic cleavage, for example to destroy the multimeric oleosin chain and/or release the desired peptide from the multimeric oleosin chain.

Thus, the construct may further include nucleotide sequences encoding linking sequences between two or more of the oleosin repeat units. The linking sequences may be short sequences, for example sequences that allow flexibility between the repeat units (act as a flexible hinge) or induce a directional change (act as a directional induction hinge) between the repeats, enable degradation between the oleosin repeat units, for example by peptidases, unrelated peptide sequences that may have bioactive properties, or sites for enzymatic cleavage and or subsequent fusion.

More particularly, the linking nucleotide sequence(s) between the oleosin repeat units may encode the native N- and C-termini of the respective repeats, sequences that code for comparatively short peptides that allow flexibility between the repeats,

comparatively short peptides with amino acid residues that induce a directional change in the chain (for example proline), sequences that enable the future targeted cloning of additional sequences between the repeats, sequences that encode for specific targeted peptidase degradation, sequences that encode for bioactive peptides, and/or sequences encoding sites for specific enzymatic cleavage and subsequent fusion, e.g., modified self splicing intein and polymerisation cyclisation (Williams et al., 2005).

The multiple oleosin nucleotide sequences may code for the same oleosin peptide sequence or code for different oleosin peptide sequences. The multiple oleosin nucleotide sequences may code for the same oleosin peptide sequence but use alternate codons in the nucleotide sequence where applicable. This enables the construct to also be used in prokaryotic expression systems that are not non-recombinant minus (rec$^-$). It also enables the use of oleosins that contain different affinities for the oil body.

In a particularly preferred embodiment of the present invention, the nucleic acid encoding an oleosin repeat unit includes a nucleotide sequence selected from the group consisting of sequences shown in Figures 13, 15, 17, 19, 21, 23, 25, 27, 31, 34, 70, 72-77 and 97-106 hereto, and functionally active fragments and variants thereof.

The construct of the present invention includes one or more nucleic acids encoding two or more oleosin repeats, preferably between two or three and twenty oleosin repeats, more preferably between two or three and ten oleosin repeats, most preferably between two or three and five or six oleosin repeats.

The diacylglycerol acyltransferase may be of any suitable type, including functionally active fragments and variants thereof. In a preferred form, the diacylglycerol acyltransferase is selected from the group consisting of DGAT1, DGAT2 and PDAT.

Nucleic acids according to the invention may be full-length genes or part thereof, and are also referred to as “nucleic acid fragments” and “nucleotide sequences” in this specification.

The nucleic acid may be of any suitable type and includes DNA (such as cDNA or genomic DNA) and RNA (such as mRNA) that is single- or double-stranded, optionally containing synthetic, non-natural or altered nucleotide bases, and combinations thereof.

By "functionally active" in respect of a nucleotide sequence is meant that the fragment or variant (such as an analogue, derivative or mutant) is capable of modifying lipids in a plant. Such variants include naturally occurring allelic variants and non-naturally occurring variants. Additions, deletions, substitutions and derivatizations of one or more of the nucleotides are contemplated so long as the modifications do not result in loss of functional activity of the fragment or variant. Preferably the functionally active fragment or variant has at least approximately 80% identity to the relevant part of the oleosin repeat sequences exemplified herein, more preferably at least approximately 90% identity, most preferably at least approximately 95% identity. Such functionally active variants and fragments include, for example, those having nucleic acid changes that result in conservative amino acid substitutions of one or more residues in the corresponding amino acid sequence. For example, the nucleic acid sequence may be modified to enhance expression without altering the amino acid sequence. Preferably the fragment has a size of at least 30 nucleotides, more preferably at least 45 nucleotides, most preferably at least 60 nucleotides.

By "functionally active" in the context of a polypeptide is meant that the fragment or variant has one or more of the biological properties of oleosin polypeptides. Additions, deletions, substitutions and derivatizations of one or more of the amino acids are contemplated so long as the modifications do not result in loss of functional activity of the fragment or variant. Preferably the functionally active fragment or variant has at least approximately 60% identity to the relevant part of the oleosin polypeptides exemplified herein, more preferably at least approximately 80% identity, most preferably at least approximately 90% identity. Such functionally active variants and fragments include, for example, those having conservative amino acid substitutions of one or more residues in the corresponding amino acid sequence. Preferably the fragment has a size of at least 10 amino acids, more preferably at least 15 amino acids, most preferably at least 20 amino acids.

By "operatively linked" is meant that a regulatory element is capable of causing expression of said nucleic acid in a cell and/or a terminator is capable of terminating expression of said nucleic acid in a cell. Preferably, said regulatory element is upstream of said nucleic acid and said terminator is downstream of said nucleic acid.

By "an effective amount" is meant an amount sufficient to result in an identifiable phenotypic trait in said plant, or a plant, plant seed or other plant part derived therefrom. Such amounts can be readily determined by an appropriately skilled person, taking into account the type of plant, the route of administration and other relevant factors. Such a person will readily be able to determine a suitable amount and method of administration. See, for example, Maniatis et al, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, the relevant disclosure of which is incorporated herein by reference.

It will also be understood that the term "comprises" (or its grammatical variants) as used in this specification is equivalent to the term "includes" and should not be taken as excluding the presence of other elements or features.

The construct of the present invention may be a vector. In a preferred embodiment of this aspect of the invention, the vector may include at least one regulatory element, such as a promoter, operatively linked to the nucleic acid. The vector may also include an operatively linked terminator.

The vector may be of any suitable type and may be viral or non-viral. The vector may be an expression vector. Such vectors include chromosomal, non-chromosomal and synthetic nucleic acid sequences, eg. derivatives of plant viruses; bacterial plasmids; derivatives of the Ti plasmid from *Agrobacterium tumefaciens*, derivatives of the Ri plasmid from *Agrobacterium rhizogenes*; phage DNA; yeast artificial chromosomes; bacterial artificial chromosomes; binary bacterial artificial chromosomes; vectors derived from combinations of plasmids and phage DNA. However, any other vector may be used as long as it is replicable, or integrative or viable in the relevant cell.

The regulatory element and terminator may be of any suitable type and may be endogenous to the target cell or may be exogenous, provided that they are functional in the target cell.

Preferably one of the regulatory elements is a promoter. A variety of promoters which may be employed in the vectors of the present invention are well known to those skilled in the art. Factors influencing the choice of promoter include the desired tissue specificity of the vector, and whether constitutive or inducible expression is desired and the nature of the cell to be transformed. Particularly suitable constitutive promoters for use in plants include the Cauliflower Mosaic Virus 35S (CaMV 35S) promoter, the maize Ubiquitin promoter, and the rice Actin promoter. In a preferred embodiment the promoter may be chosen to enable the expression of oleosin in the desired organ, tissue and stage of development.

A variety of terminators which may be employed in the vectors of the present invention are also well known to those skilled in the art. The terminator may be from the same gene as the promoter sequence or a different gene. Particularly suitable terminators for use in plants are polyadenylation signals, such as the CaMV 35S polyA and other terminators from the nopaline synthase (*nos*) and the octopine synthase (*ocs*) genes.

The vector, in addition to the regulatory element, the nucleic acid and the terminator, may include further elements necessary for expression of the nucleic acid, in different combinations, for example vector backbone, origin of replication (ori), multiple cloning sites, spacer sequences, enhancers, introns (such as the maize Ubiquitin Ubi intron), antibiotic resistance genes and other selectable marker genes (such as the neomycin phosphotransferase (*npt2*) gene, the hygromycin phosphotransferase (*hph*) gene, the phosphinothricin acetyltransferase (*bar* or *pat*) gene), and reporter genes (such as green fluorescence protein (GFP), beta-glucuronidase (GUS) gene (*gusA*)). The vector may also contain a ribosome binding site for translation initiation. The vector may also include appropriate sequences for facilitating correct transcription, amplifying expression, increasing mRNA stability, enhancing protein translation or facilitating accurate translation.

As an alternative to use of a selectable marker gene to provide a phenotypic trait for selection of transformed host cells, the presence of the construct or vector in transformed cells may be determined by other techniques well known in the art, such as PCR (polymerase chain reaction), Southern blot hybridisation analysis, histochemical GUS assays, northern and Western blot hybridisation analyses.

Those skilled in the art will appreciate that the various components of the construct or vector are operatively linked; so as to result in expression of said nucleic acid. Techniques for operatively linking the components of the construct or vector of the present invention are well known to those skilled in the art. Such techniques include the use of linkers, such as synthetic linkers, for example including one or more restriction enzyme sites.

The constructs and vectors of the present invention may be incorporated into a variety of plants, including monocotyledons (such as grasses from the genera *Lolium*, *Festuca*, *Paspalum*, *Pennisetum*, *Panicum* and other forage and turfgrasses, corn, rice, sugarcane, oat, wheat and barley) dicotyledons (such as arabidopsis, tobacco, soybean, canola, cotton, potato, chickpea, medics, white clover, red clover, subterranean clover, alfalfa, eucalyptus, poplar, hybrid aspen, and gymnosperms (pine tree)). In a preferred embodiment, the constructs and vectors are used to transform commercial crops that are fed directly to animals.

The constructs and vectors of the present invention may also be incorporated into other eukaryotic expression systems, including yeast, insect and mammalian cells. Thus, repeat oleosins, such as multimeric tandem repeat oleosins, (either homo or hetero repeats) may be generated by recombinant protein expression in eukaryotic expression systems and subsequently purified as functional recombinant oil bodies having the additional properties afforded by the presence of repeat oleosins.

The constructs and vectors of the present invention may also be incorporated into prokaryotic expression systems. Thus, repeat oleosins, such as multimeric tandem repeat oleosins, (either homo or hetero repeats) may be generated by recombinant protein expression in bacteria such as *Escherichia coli* and subsequently purified and

recombined with for example, phospholipids and triacyl glyceride to generate functional recombinant oil bodies with the additional properties afforded by the presence of repeat oleosins.

The constructs and vectors of the present invention may be constructed by any suitable method, including:

- Amplification and subsequent cloning of oleosin sequences, for example using PCR with primers designed to publicly available oleosin sequences. For plant based expression clones these may include oleosins amplified from genomic DNA that may also contain introns depending on the clone. For prokaryotic and non-plant eukaryotic expression systems the source of the template for PCR preferably comes from cDNA which will not contain plant introns.
- Restriction digestion and ligation of oleosin clones that contain suitable inframe restriction sites at both N and C termini. Restriction sites may be added by techniques such as PCR or ligation of sticky ends.
- Engineering different restriction sites (that correspond to a cloning polylinker sequence) onto the ends of individual oleosin clones, then building up the multimer by a series of digestions and ligations by firstly ligating the polylinker to the 3' end of the first oleosin clone.
- Chemically synthesising the complete construct.
- A combination of the above.

Techniques for incorporating the constructs and vectors of the present invention into cells (for example by transduction, transfection or transformation) are well known to those skilled in the art. For plant cells, techniques include *Agrobacterium* mediated introduction, electroporation to tissues, cells and protoplasts, protoplast fusion, injection into reproductive organs, injection into immature embryos and high velocity projectile introduction to cells, tissues, calli, immature and mature embryos. The choice of technique will depend largely on the type of cell to be transformed.

Cells incorporating the constructs and vectors of the present invention may be selected, as described above, and then cultured in an appropriate medium to regenerate transformed plants, using techniques well known in the art. The culture conditions, such

as temperature, pH and the like, will be apparent to the person skilled in the art. The resulting plants may be reproduced, either sexually or asexually, using methods well known in the art, to produce successive generations of transformed plants.

In a further aspect of the present invention there is provided a plant cell, plant, plant seed or other plant part, including, e.g. transformed with, a construct or vector of the present invention.

The plant cell, plant, plant seed or other plant part may be from any suitable species, including monocotyledons, dicotyledons and gymnosperms. In a preferred embodiment the plant cell, plant, plant seed or other plant part is from a commercial plant that is normally fed directly to animals.

The present invention also provides a plant, plant seed or other plant part derived from a plant cell of the present invention and including a construct or vector of the present invention. The present invention also provides a plant, plant seed or other plant part derived from a plant of the present invention and including a construct or vector of the present invention.

The present invention also provides a eukaryotic cell, such as a yeast, insect or mammalian cell, including, eg. transformed with, a construct or vector of the present invention.

The present invention also provides a prokaryotic cell, eg. a bacteria such *E. coli*, including, eg. transformed with, a construct or vector of the present invention.

The present invention also provides a method of producing repeat oleosins in a plant, said method including introducing into said plant a construct or vector of the present invention. By "repeat oleosin" is meant a recombinant polypeptide including two or more oleosin repeat units.

The present invention also provides a method of producing repeat oleosins in a eukaryotic cell, said method including introducing into said eukaryotic cell a construct or vector of the present invention. The method may include the further step of partially or

2006303870 16 Oct 2012

substantially purifying said repeat oleosin from said cell. In a preferred form the repeat oleosin may be partially or substantially purified by the generation of oil bodies. The oil bodies may be produced within the cell, e.g. plant seed.

The present invention also provides a method of producing repeat oleosins in a prokaryotic cell, said method including introducing into said prokaryotic cell a construct or vector of the present invention. The method may include the further steps of partially or substantially purifying said repeat oleosin from said cell. In a preferred form the repeat oleosin may be partially or substantially purified by the generation of oil bodies. The oil bodies may be produced artificially using recombinant oleosin repeats from other expression systems such as *E. coli*.

The oleosin protein repeats may be purified using affinity chromatography, such as a fused Histidine tag and Ni^{2+} resin purification.

The present invention also provides a partially or substantially purified and/or recombinant polypeptide including two or more oleosin repeat units.

In a preferred embodiment, the present invention provides a partially or substantially purified and/or recombinant polypeptide including three or more oleosin repeat units.

Preferably, the polypeptide includes an amino acid sequence selected from the group consisting of Seq ID Nos: 47, 48, 51, 58-63 and 74-83, and functionally active fragments and variants thereof.

The polypeptide may be produced by expression of a construct or vector according to the present invention.

The present invention also provides a lipid encapsulated by a polypeptide according to the present invention.

The oleosin repeat units may be of any suitable type, including functionally active fragments and variants thereof. For example, the oleosin repeat units may be selected from the group consisting of L-oleosin, H-oleosin, steroleosin and caoleosin and

2006303870 16 Oct 2012

functionally active fragments and variants thereof. The oleosin repeats may be tandem repeats. The oleosin repeats may be multimeric tandem repeats. The oleosin repeats may be homo- or hetero- repeats.

In a preferred embodiment, the oleosin repeat unit may be from white clover or sesame seed, or may be a synthetic or recombinant version of an oleosin repeat unit from white clover or sesame seed. However, the invention is not limited thereto and the.

polypeptide of the present invention may include any functionally active plant oleosin sequence.

The oleosin repeat units may be linked either directly or by linking sequences. Thus, the polypeptide may further include linking sequences between two or more of the oleosin repeat units. The linking sequences may allow flexibility, induce a directional change or enable degradation between the oleosin repeat units.

In a particularly preferred embodiment of the present invention, the oleosin repeat unit includes an amino acid sequence selected from the group consisting of the sequences shown in Figures 69, 70, 78-83 and 107-116 hereto, and functionally active fragments and variants thereof.

In another particularly preferred embodiment of the present invention, the oleosin repeat unit is encoded by a nucleotide sequence selected from the group consisting of sequences shown in Figures 13, 15, 17, 19, 21, 23, 25, 27, 31, 34, 70, 72-77 and 97-106 hereto, and functionally active fragments and variants thereof.

The recombinant polypeptide of the present invention includes two or more oleosin repeats, preferably between two or three and twenty oleosin repeats, more preferably between two or three and ten oleosin repeats, most preferably between two or three and five or six oleosin repeats.

Applicants have found that that repeat oleosins, such as recombinant multimeric tandem repeat oleosins, (either homo or hetero repeats) may be expressed in the seeds of plants which would mean that the extract would contain a mixture of native oleosins along with the repeat oleosins. If a single species of oleosin multiples are required the repeat oleosins may be expressed in plants in which native oleosin expression has been suppressed, e.g., by mutagenesis, gene silencing or natural selection.

In addition, the repeat oleosins may be coexpressed with a diacylglyerol acyltransferase such as DGAT1 or DGAT2 or PDAT, in plant vegetative organs allowing the generation of emulsion complexes containing substantially only the desired species of oleosin (since oleosins are not normally expressed in the vegetative portions of plants). Thus,

co-expression of DGAT1, DGAT2 or PDAT with polyoleosin in the vegetative portions of plants produces oil bodies encapsulated by the polyoleosin.

Furthermore, the repeat oleosins may be expressed and purified from bacteria, such as *E. coli,* since oleosins are not naturally present in *E. coli*. Expression of repeat oleosins in some bacteria may require modification of the nucleotide sequence to avoid recombination events occurring in rec$^-$ strains.

Applicants have found that use of a series of oleosin repeats generates a recombinant protein with exploitable properties. Linking oleosin units to give homo or hetero multimeric repeats reduces the number of N-termini available for the initiation amino peptidase degradation in the rumen and/or stomach as well as altering the physiochemical characteristics (e.g., the hydrophobic interactions) of the protein and thus broadening the range of emulsification properties. In particular, this has a number of exploitable benefits, including:

- Generating a simple method for producing foodstuffs such as meat and milk with health promoting lipid profiles by reducing the degree of biohydrogenation.

- Allowing for the delivery of fused bioactive peptides to be ingested and delivered intact to the site of absorption.

- Providing a rapid mechanism to alter the emulsion properties of reconstituted feeds, pharmaceuticals, beauty products etc.

- Allowing for the delivery of lipid soluble active compounds to be delivered and released in a controlled fashion by topical application.

- Allowing for the delivery of lipid soluble bioactive compounds to be ingested and delivered intact to the site of absorption.

Applicants have found that by the use of repeat oleosins, such as recombinant multimeric tandem repeat oleosins (either homo or hetero repeats) the stability of an oil body may be regulated. The incorporation of amino acid residues between the oleosin

repeats may be desirable to allow sufficient flex between the repeats, to influence the curvature of the oil body, to control the direction of the repeats relative to each other, to incorporate bioactive peptides, to provide sites for directed proteolytic degradation, or to provide sites for specific enzymatic cleavage and subsequent fusion, e.g., modified self splicing intein and polymerisation cyclisation (Williams et al., 2005).

Thus, the present invention also provides a mechanism for the delivery of bioactive peptides that may be fused between the repeats or at the end of the repeats.

The present invention also provides a mechanism for allowing free rotation between the repeats or for directing the orientation of the repeats relative to each other by incorporating linking sequences between the repeats.

The size of the suspended particles contributes to the stability of the suspension and the amount of material that may be suspended and this allows the emulsification properties to be tailored for the application. By altering the number of oleosin repeats in a peptide sequence (preferably using recombinant technologies) oleosins with different emulsification properties may be generated. These may confer enhanced stabilities in terms of, for example, temperature stability, pH stability and/or altered particle size. In turn this broadens the number of compounds that may be emulsified, as well as expanding the applications of the emulsifications, for example, by extending their stability and the amount of compounds that may be emulsified.

In a further aspect the present invention provides a method of manipulating lipids in a plant, said method including introducing into said plant a construct including one or more nucleic acids encoding two or more oleosin repeat units.

Manipulation of lipids includes, but is not limited to, alteration of emulsification properties, including stability of suspension and amount of material that may be suspended, alteration of physiochemical properties, including hydrophobic interactions, and altering the degree of biohydrogenation.

In a preferred embodiment, the present invention provides a method of altering stability of an oil body in a plant, said method including modifying an oleosin in said plant to

include two or more oleosin repeat units. This may in turn alter the ratio of oleosin to oil in said oil body.

Reduction of Biohydrogenation

In ruminants, biohydrogenation is the hydrogenation of non-reduced compounds (such as unsaturated fats) with hydrogen from rumen biota. Applicants have engineered oleosin to generate oil bodies containing unsaturated fats in the TAG (surrounded by a phospholipid monolayer) encapsulated by the oleosin repeats of the present invention. This may be less susceptible to the process of biohydrogenation in the rumen before it passes into the intestine for absorption. While applicants do not wish to be restricted by theory, it is thought that a chain of oleosin units would have fewer N-termini per unit of oleosin available for amino peptidase degradation (the main form of protein degradation in the rumen). In turn this would reduce the degree of oleosin degradation and therefore reduce the loss of oil body integrity and subsequently reduce the degree of biohydrogenation of the unsaturated fats in the TAG. In turn this would lead to an increase in the level of unsaturated fats reaching the site of adsorption in the intestine. In turn this would lead to a change in the fatty acid profile of foodstuffs such as milk and meat products from the animal eating such a product.

Thus, the present invention provides a method of altering biohydrogenation of a lipid, said method including encapsulating said lipid in a recombinant polypeptide including two or more oleosin repeat units.

The present invention also provides a method of protecting unsaturated lipids from biohydrogenation, said method including incorporating said unsaturated lipids into an oil body including a recombinant polypeptide including two or more oleosin repeat units.

Delivery of Fused Bioactive Peptides Orally and Topically

Applicants have engineered restriction sites between the oleosin repeats to enable insertion of a frame coding sequence (such as those encoding bioactive peptides that would normally be susceptible to degradation in the intestines). While applicants do not wish to be restricted by theory it is thought that a chain of oleosin units would have

enhanced stability in the digestive acid conditions of the stomach as well as fewer N-termini per unit of oleosin available for amino peptidase degradation (the main form of protein degradation in the rumen). Thus a peptide inserted between multimeric oleosin tandem repeats is afforded a degree of protection from degradation by digestive acidic conditions and aminopeptidases. In turn this would reduce the degree of active peptide degradation in the stomach and rumen. In turn this would lead to higher levels of active peptides reaching the site of adsorption in the intestine. In turn this would lead to an increased absorption of bioactive peptides by the organism. The degree of protection may be altered by the position of the bioactive peptide in the oleosin chain, the nature of the oleosins in the chain, and/or the amino acid sequences joining the oleosins and bioactive peptides.

Thus, the present invention provides a method of delivering a bioactive peptide, to animals including humans, said method including inserting said peptide at the N- or C-terminus of a series of two or more oleosin repeat units or between two or more oleosin repeat units to produce a recombinant polypeptide and administering said recombinant polypeptide to said animal.

Preferably, the bioactive peptides are delivered orally or topically. The bioactive peptides may be delivered to the intestine. The bioactive peptides may be delivered by timed release, eg. sustained release or delayed release.

The present invention also provides a partially or substantially purified or recombinant polypeptide including two or more oleosin repeat units and further including one or more bioactive peptides.

The bioactive peptide may be inserted at the N- or C-terminus of a series of two or more oleosin repeat units or between two or more oleosin repeat units.

The recombinant polypeptide may be produced by expression of a construct or vector according to the present invention.

Thus, the present invention also provides a construct including one or more nucleic acids encoding two or more oleosin repeat units and further including one or more nucleic acids encoding bioactive peptides.

Delivery of Encapsulated Bioactive Peptides and Other Compounds Orally and Topically

Polyoleosin provides a mechanism for the development and delivery of compounds such as therapeutic and prophylactic drugs, including drugs for internal parasites in humans and animals and bioactive peptides; and organisms such as health promoting bacteria (e.g., lactobacillus) by encapsulation of the compound or organism within the oil body. In particular, it provides a mechanism for delivering bioactive peptides through the rumen and into the digestive system, substantially without loss of bioactivity. This issue currently represents a major hurdle for the development of bioactives for internal parasites in rumens. Polyoleosins thus facilitate development of bioactive drug development and delivery.

In addition, polyoleosin has a similar application in the delivery of encapsulated bioactives in cosmetics, eg. creams and may be applied epidermally, for example to wounds or skin problems. Polyolesin may also have applications for controlled release in dermal applications.

Thus, the present invention also provides a method of delivering compounds and/or organisms to animals including humans, said method including encapsulation said compound or organism in an oil body including two or more oleosin repeat units and administering said oil body to said animal.

Preferably the compound is a therapeutic or prophylactic drug or a bioactive peptide or protein.

Preferably the organism is a bacterium, more preferably a health promoting bacterium such as lactobacillus.

Preferably the compound or organism is delivered orally or topically. The compound or organism may be delivered to the intestine. The compound or organism may be delivered by timed release eg. sustained release or delayed release.

Emulsification and Encapsulation of Bioactives

Repeat oleosins, such as recombinant multimeric tandem repeat oleosins (either homo or hetero repeats), may be used to tailor emulsion complexes and to encapsulate bioactive compounds that may exist in the emulsion. By the use of repeat oleosins, the stability of an oil body may be tailored by oleosin. The size of the suspended particles contributes both to the stability of the suspension and the amount of material that may be suspended. By altering the number of oleosin repeats in a peptide sequence (preferably using recombinant technologies) oleosins with different emulsification properties are generated. These confer enhanced stabilities in terms of, for example, temperature and pH stability and altered particle size. In turn this broadens the number of compounds that may be emulsified as well as expanding the applications of the emulsifications.

Polyoleosin allows the manufacturer to tailor the emulsification properties by altering the number of oleosin repeats. The majority of processed foods utilise emulsifiers of some form or another. In addition, polyoleosins may be utilised in the cosmetics industry. The oil based encapsulation mechanism provides an ideal delivery mechanism for any compound, notably bioactives for dermal application.

Accordingly, in a further aspect the present invention provides a method of altering the emulsification properties of an oleosin, said method including recombinantly producing the oleosin with two or more oleosin repeat units.

Altering the emulsification properties of the oleosin may include altering temperature and/or pH stability and/or size of oil bodies including the oleosin.

The present invention also provides a recominbant oleosin with altered emulsification properties, said oleosin including two or more oleosin repeat units.

The recombinant polypeptide may be produced by expression of a construct or vector according to the present invention.

Thus, in the polypeptides of the present invention tandem oleosins may be fused directly or by small linking (eg. hinge) sequences. The polypeptides may contain fused bioactive peptides either at the terminal ends of the repeats or located between the repeats.

Such polyoleosins may be used to deliver encapsulated products/compounds/proteins that do not have to be fused to the oleosin peptides but rather are generated during the process of creating oil bodies or AOBs.

Polyoleosins may also be used to deliver bioactive peptides either encapsulated within oil bodies, AOBs or fused to the ends of oleosin repeats or fused between the repeats.

Polyoleosins used for emulsification or delivery or protection of compounds may be tailored by changing the number of oleosin tandem repeats rather than simply being the fusion of a peptide of insert fused between two oleosin repeats.

The present invention will now be more fully described with reference to the accompanying Examples and drawings. It should be understood, however, that the description following is illustrative only and should not be taken in any way as a restriction on the generality of the invention described above.

In the Figures

Figure 1. Kyte and Doolittle hydrophobic plot (window = 17) of a typical oleosin sequence.

Figure 2. Oil body showing topology of oleosin single peptide chains.

Figure 3. Oil body showing topology of dimeric oleosin repeat peptide chains joined by unique peptide linkers.

Figure 4. Oil body showing topology of trimeric oleosin repeat peptide chains joined by unique peptide linkers.

Figure 5. Oil body showing topology of tetrameric oleosin repeat peptide chains joined by unique peptide linkers.

Figure 6. Oil body showing topology of pentameric oleosin repeat peptide chains joined by unique peptide linkers.

Figure 7. Typical Kyte and Doolittle hydrophobicity plot (window = 17) of homo or hetero pentameric oleosin tandem repeat containing unique inter oleosin peptide liners.

Figure 8. pBLUESCRIPT II SK MCS sequence and translated amino acid sequence. With the exception of ***Eco*RV** all other restriction sites are not present in either the white clover oleosin clone or the Gateway PCR entry vector pENTR/D _•_ and XXX = codon not part of a restriction site.

Figure 9. Standard layout of primers and PCR products to generate white clover polyoleosin constructs.

Figure 10. Map of pENTR/D-TOPO.

Figure 11. Map of pCR2.1-TOPO.

Figure 12. Map of pOLE01-MC.

Figure 13. Sequence of the oleosin coding region in pOLE01-MC (Seq ID No. 29). **CACC** = GATEWAY™ adapter, XXXX = oleosin sequence, *XXXXXX* = restriction enzyme site - included in primer, TGA = stop codon.

Figure 14. Map of pOLE02-MC.

Figure 15. Sequence of the oleosin coding region in pOLE02-MC (Seq ID No. 30). XXXX = oleosin sequence, *XXXXXX* = restriction enzyme site - included in primer, ***XXXXXX*** = codons for extra amino acids – included in primer.

Figure 16. Map of pOLE03-MC.

Figure 17. Sequence of the oleosin coding region in pOLE03-MC (Seq ID No. 31). XXXX = oleosin sequence, *XXXXXX* = restriction enzyme site - included in primer, ***XXXXXX*** = codons for extra amino acids – included in primer.

Figure 18. Map of pOLE04-MC.

Figure 19. Sequence of the oleosin coding region in pOLE04-MC (Seq ID No. 32). XXXX = oleosin sequence, *XXXXXX* = restriction enzyme site - included in primer, ***XXXXXX*** = codons for extra amino acids – included in primer.

Figure 20. Map of pOLE05-MC.

Figure 21. Sequence of the oleosin coding region in pOLE05-MC (Seq ID No. 33). XXXX = oleosin sequence, *XXXXXX* = restriction enzyme site - included in primer, ***XXXXXX*** = codons for extra amino acids – included in primer.

Figure 22. Map of pOLE06-MC.

Figure 23. Sequence of the oleosin coding region in pOLE06-MC (Seq ID No. 34). **XXXX**= GATEWAY™ adapter sequence, XXXX= oleosin sequence, *XXXXXX*= base pairs encoding for additional amino acids- included in primer, XXXX= polylinker sequence, TGA = stop codon.

Figure 24. Map of pOLE07-MC.

Figure 25. Sequence of the oleosin coding region in pOLE07-MC (Seq ID No. 35). **XXXX**= GATEWAY™ adapter sequence, XXXX= oleosin sequence, *XXXXXX*= base pairs encoding for additional amino acids- included in primer, XXXX= polylinker sequence, TGA = stop codon.

Figure 26. Map of pOLE08-MC.

Figure 27. Sequence of the oleosin coding region in pOLE08-MC (Seq ID No. 36). **XXXX**= GATEWAY™ adapter sequence, XXXX= oleosin sequence, *XXXXXX*= base pairs encoding for additional amino acids- included in primer, XXXX= polylinker sequence, TGA = stop codon.

Figure 28. Agarose gel of *Hind*III/*Cla*I digested pOLE04-MC and pOLE08-MC. Dotted boxes outline the regions excised for extraction from gel.

Figure 29. Map of pOLE09-MC.

Figure 30. Agarose gel results from checking pOLE09-MC.

Expected sizes:

Clone 7

NcoI/XhoI - 2.6kb and 1.3kb

NcoI/SpeI/XhoI - 2.6kb, 1.3kb & 57bp

NotI/XhoI - 2.6kb, 697bp, & 643bp

XhoI - 3.9kb

Clone 9

NcoI/XhoI - two bands at ~2.6kb clones C9.1-C9.3 correct

NcoI/SpeI/XhoI - 2.6kb & two bands at ~1.3kb clones C9.1-C9.3 correct

NotI/XhoI - 2.6kb, 1.9kb & 643bp clones C9.1-C9.3 correct

XhoI - 5.2kb clones C9.1-C9.3 correct

Clones 9.1 - 9.3 all showed the correct/expected sizes for pOLE09-MC.

The region in the pOLE09-MC cassette containing the newly inserted oleosin was then sequenced.

Figure 31. Sequence of pOLE09-MC, confirming the addition of the oleosin from pOLE04-MC into pOLE08-MC. Sequence was obtained by sequencing across the last oleosin in pOLE09-MC into the original pOLE08-MC vector. All mismatches were checked back to the original electropherogram to confirm errors were due to software misreads.

Figure 32. Map of pOLE10-MC.

Figure 33. Agarose gel analysis of a restriction enzyme digested pOLE10-MC clone, compared to pOLE09-MC.

Expected sizes:

Clone 9	*Xho*I/*Pst*I	-4.5kb and 650bp	correct
	*Pst*I	-5.1kb correct	
	*Sal*I/*Pst*I	-4.5kb and 650bp	correct

Expected sizes:

Clone 10	*Xho*I/*Pst*I	-4.5kb and 1.2kb	correct
	*Pst*I	-5.7kb correct	
	*Sal*I/*Pst*I	-5.1kb and 650bp	correct

Both pOLE09-MC and pOLE10-MC clones were found to be correct when analysed by restriction digests.

The region in the pOLE10-MC cassette containing the newly inserted oleosin was then sequenced using the M13-Reverese primer (five clones sequenced and found to match the predicted sequence.

Figure 34. Sequence of pOLE10-MC, confirming the addition of the oleosin from pOLE05-MC, into pOLE09-MC. Sequence was obtained via sequencing into the fifth pOLE10-MC and back across into pOLE09-MC.Figure 35. Map of pRSh1.

Figure 36. Map of pRS12.

Figure 37. *Eco*RI digest of C6 and C7 clones in pRSH1.

Expected sizes:

Clone 6-pRSH1 10,349bp, 3084bp, 725bp and 566bp clones 1,3, & 4 correct

Clone 7-pRSH1 10,349bp, 3084bp, 1352bp and 566bp clones 1,2, & 3 correct

pRSH1 11,525bp, 3084bp, 566bp and 463bp.

Figure 38. *Eco*RI digest of C8 clones in pRSH1.

Expected sizes:

Clone 8-pRSH1 10,349bp, 3084bp, 1973bp and 566bp clones 2-8 correct

pRSH1 11,525bp, 3084bp, 566bp and 463bp.

Figure 39. *Eco*RI digest of C9 and C10 clones in pRSH1.

Expected sizes:

Clone 9-pRSH1 10,970bp, 3084bp, 1973bp and 566bp clones 1,2, & 4 correct

Clone 10-pRSH1 11,594bp, 3084bp, 1973bp and 566bp clones 1,2, & 4 correct.

Figure 40. *Bam*HI digest of C9 and C10 clones in pRSH1.

Expected sizes:

Clone 9-pRSH1 10826bp, 4400bp and 1367bp clone correct

Clone 10-pRSH1 10826bp, 4400bp and 1991bp clone correct

pRSH1 10826bp, 3279bp, 830bp and 703bp.

Figure 41. *Eco*RI digest of DGAT1 clones in pRSH1.

Expected sizes:

DGAT1-pRSH1 10700bp, 3084bp, 1232bp and 566bp clones 1-3 correct

pRSH1 **11525bp, 3084bp, 566bp and 463bp.**Figure 42. *Eco*RV digest of C6 clones in pRS12.

Expected sizes:

C6-pRS12 9131bp, 7424bp and 3762bp clones 2,3,5, & 8 correct

pRS12 13810bp and 7424bp.

Figure 43. Restriction digest of DGAT1 and C7 clones in pRS12.

Expected sizes:

DGAT1-pRS12	*Spe*I	14437bp, 3833bp and 2908bp	clones 1 & 2 correct
C7-pRS12	*Spe*I	14437bp, 2908bp, 2724bp and 878bp	clones 1 & 2 correct
pRS12	*Spe*I	14437bp, 3889bp and 2908bp	
DGAT1-pRS12	*Bam*HI	17039bp and 4139bp	clones 1 & 2 correct
C7-pRS12	*Bam*HI	16683bp, 4139 bp and 125bp	clones 1 & 2 correct
pRS12	*Bam*HI	15562bp, 4139bp, 830bp and 703bp.	

Figure 44. Restriction digest of C8, C9 and C10 clones in pRS12.

Expected sizes:

C8-pRS12	*Spe*I	14437bp, 2908bp, 2724bp and 1499bp	clones 1 & 2 correct
C9-pRS12	*Spe*I	14437bp, 2908bp, 2724bp and 2120bp	clones 2 & 3 correct
C10-pRS12	*Spe*I	14437bp, 2908bp, 2744bp and 2724bp	clones 1-3 correct
pRS12	*Spe*I	14437bp, 3889bp and 2908bp	
C8-pRS12	*Bam*HI	16683bp, 4139bp and 746bp	clones 1 & 2 correct
C9-pRS12	*Bam*HI	16683bp, 4139bp and 1367bp	clones 2 & 3 correct
C10-pRS12	*Bam*HI	16683bp, 4139bp and 1991bp	clone 1... correct
pRS12	*Bam*HI	15562bp, 4139bp, 830 bp and 703bp.	

Figure 45. Restriction digest of C8, C9 and C10 clones in pRS12.

Expected sizes:

C10-pRS12	*Bam*HI	16683bp, 4139bp and 1991bp	clones ...2 & 3 correct
C8-pRS12	*Eco*RV	10382bp, 7424bp and 3762bp	clones 1 & 2 correct
C9-pRS12	*Eco*RV	10382bp, 7424bp and 4383bp	clones 2 & 3 correct
C10-pRS12	*Eco*RV	10382bp, 7424bp and 5007bp	clones 1-3 correct
pRS12	*Spe*I	14437bp, 3889bp and 2908bp	
pRS12	*Bam*HI	15562bp, 4139bp, 830 bp and 703bp	
pRS12	*Eco*RV	13810bp and 7424bp.	

Figure 46. A representative diagram of the pRSh1 binary vector containing oleosin multimers.

Figure 47. A representative diagram of the pRS12 binary vector containing oleosin multimers.

Figure 48. Plasmid map for the expression vector p17NON (N terminal oleosin, in pDEST17).

Figure 49. Plasmid maps for the expression vectors p42CON (N terminal oleosin in pET DEST42).

Figure 50. Characterisation of p17NON by PCR analysis.
Expected sizes: NT - 505bp
All plasmids appear to contain an insert of the expected size
This amplification confirms the presence of the N terminal oleosin gene in the pDEST17 plasmid → p17NON

Figure 51. Characterisation of p42CON by PCR analysis.
Expected sizes: NT - 561bp
NT#1 appears to contain a slightly larger insert..
All others appear to contain an insert of the expected size
This amplification confirms the presence of the N-terminal oleosin gene in the pET DEST42 plasmid → p42CON

Figure 52. Expression analysis by SDS PAGE of 6HON. N = Non-induced I = Induced < = putative expressed 6HON expected size 6HON = 14.7kD.

Figure 53. Inclusion body analysis by SDS PAGE of the 6HONpeptide. N = Non induced, I = Induced, < = putative expressed 6HON, expected size 6HON = 14.7kD
* = putative expressed 6HOC expected size 6HOC = 12.8kD

Figure 54. Trial scale Nickel column purification analysis by SDS PAGE of 6HON. < = putative expressed 6HON expected size 6HON = 14.7kD.
Samples loaded as equivalent volumes.

Figure 55. Large scale Nickel purification analysis by SDS PAGE of 6HON. Expected size of 6HON = 14.7kD.

Figure 56. Concentration of 6HON. Volume of each eluted fraction was approximately 1mL.

Figure 57. SDS-PAGE/immunoblot screened against anti oleosin N terminal antibody. Run on 18% SDS PAGE gel, 90min 150V. 1° and 2° antibody diluted in 20mL PBS. 1° antibody = 1:2000 (10µL). 2° antibody = 1:5000 (4µL) anti rabbit Ig, Horseradish Peroxidase linked whole antibody from donkey.

Lanes 4, 5, and 6 contained 1.8, 6.9, and 4.7ng purified 6HON, respectively. Antibody appears to bind to 17OF (Lane 2), 17OC (Lane 3) and a number of peptides present in the extract from clover seed. Reason for binding to 17OC unclear, but may be due to similarities in 6xHis tag and neighbouring regions.

Figure 58. Coomassie and subsequent immunoblot analysis of oil bodies extracted from clover seed. Arrow indicates expected size of clover seed oleosin (20kDa).
L = Ladder; A & B = duplicate oil body extracts from clover seed.

Figure 59. PCR analysis of Agrobacterium colonies transformed with oleosin clone plasmids.
Lane A = AgR6; B = AgR 7; C = AgR 8;
D = AgR 9; E & F = AgR 10; G, H, I = AgR 7;
J, K, L = AgR 8; M, N, O = AgR 9; P = untransformed LBA4404; Q and R = AgR 7 plasmid DNA; S = 1Kb plus ladder

Figure 60. Effect of Basta on regeneration from hypocotyl (top) and cotyledon explants (bottom).

Figure 61. Growth of Basta resistant and sensitive shoots from DGAT co cultivated explants on 5 mg/L Basta.

Figure 62. PCR for the bar gene on putative transgenic shoots.
Lanes A and N 1Kb plus ladder; lanes B and C DNA from AgR 6 shoots;
lanes D and E DNA from AgR 7 shoots; lanes F and G DNA from AgR 8 shoots;
lanes H and I DNA from AgR 9 shoots; lanes J and K DNA from AgR 10 shoots;
lane L AgR 7 plasmid DNA; lane M blank

Figure 63. PCR for the Bar and oleosin gene.
Top panel = PCR for bar gene, Bottom panel = PCR with the same DNA using oleosin clone specific primers

Lanes	A = 1Kb plus ladder;	B F = pRSH1 C6 plants;
	G = pRSH1 C6 plasmid;	H = blank;
	I L = pRSH1 C7 plants;	M = pRSH1 C7 plasmid;
	N = blank;	O R = pRSH1 C8 plants;
	S = pRSH1 C6 plasmid;	T= 1Kb plus ladder

Figure 64. RT PCR for oleosin from pRSH1 C6 transgenic shoots.
Lanes A and N = 1Kb plus ladder;
B F = reverse transcriptase treated RNA from pRSH1 C6 plants;
G K RNASE treated RNA from pRSH1 C6 plants;
L= pRSH1 C6 plasmid DNA;
M = blank

Figure 65. Northern blot analysis of 5µg total RNA from Lotus japonicus hairy roots transformed with 1,2,3,4 or 5 oleosin repeats.

Probed with 25ng random primed 5µCi 32P labelled oleosin cDNA.
Positive controls: + = 4pg; + = 20pg (arrows in ladder lane); oleosin gene (PCR product).
Line in single oleosin lanes indicates main transcript size detected (approximately 1.4Kb);
Line in tandem oleosin construct lanes indicates main transcript size detected (approximately 2Kb);

Line in trimeric oleosin construct lanes indicates main transcript size detected (approximately 2.7Kb);
Line in tetrameric oleosin construct lanes indicates main transcript size detected(approximately 3.4Kb);
Line in pentameric construct lanes indicates main transcript size detected (approximately 4Kb);
Small black arrows indicate oleosin hybridising transcripts of aberrant size.

Figure 66. Schematic diagram of the synthetic sesame seed polyoleosin construct. The sequence was generated using optimised codons selected for expression and elevated GC content, randomised between the repeats with elevated mRNA stability, cryptic splice sites removed, UBQ10 intron, Kozak, tetranucleotide stop codon and restriction sites added.

Figure 67. Arabidopsis thaliana UBQ10 intron modifications.

A. Modified *Arabidopsis thaliana* UBQ10 intron (3' splice site modified to be *Pst*I as per Rose and Beliakoff, 2000) (Seq ID No. 44).

B. Comparison of the modified UBQ10 intron (Seq ID No. 44) with the original sequence from Norris et al., (1993) (Seq ID No. 45). Consensus = Seq ID No. 46.

Figure 68. NetGene2 graphical output prediction of splicing of the synthetic sesame seed polyoleosin construct. The GC content had been elevated, cryptic splice sites had been manually removed, and the Arabidopsis thaliana UBQ10 intron had been added.

Figure 69. Alignment of the translated sequences of 6 tandem repeats of the original sesame seed oleosin (without linkers) (Seq ID No. 47) with the 6 tandem repeats containing random assignment of the appropriate degenerate codons, GC optimised and modified linkers (with the intron removed) (Seq ID No. 48). Consensus = Seq ID No. 49.

The alignment shows the oleosin peptide sequences of each repeat are identical, also there is no change between the randomised codons sequences with and without intron.

Figure 70. Nucleic acid (Seq ID No. 50) and amino acid (Seq ID No. 51) Sequence and feature map of the synthetic sesame seed hexameric polyoleosin with randomised optimal degenerate codons, enriched GC content, enhanced mRNA stability, engineered restriction sites (also optimised for codon usage) and modified UBQ10 intron.

Key: attL sites (bold); engineered restriction sites and linkers (italics); oleosin repeats (grey box, white letters); double stop codon to ensure no read through due to amber codon (bold, underline); tetra-nucleotide stop codon (black box, white letter); UBQ10 intron (grey box, black letters); -20 to -11 from original cDNA (underline); -10 to -1 modified Joshi et al (1997) consensus sequence (bold, italics, underline).

Figure 71. Vector map of the synthetic sesame seed polyoleosin construct in PCR Script.

Figure 72. Nucleotide sequence of clone synthesised by GENEART AG containing 6 tandem Sesame seed oleosin repeats with randomised degerate codons (Seq ID No. 52).

Figure 73. Feature map of the nucleotide sequence of the identical triple oleosin repeats synthesised by GENEART AG (Seq ID No. 53).

Figure 74. Feature map of nucleotide sequence of pET29 + Ole6-6xHis (Seq ID No. 54).

Figure 75. Feature map of nucleotide sequence of pET29 + Ole2-6xHis (Seq ID No. 55).

Figure 76. Feature map of nucleotide sequence of pET29 + Ole4-6xHis (Seq ID No. 56).

Figure 77. Feature map of nucleotide sequence of pET29 + Ole5-6xHis (Seq ID No. 57).

Figure 78. Peptide sequence of: pET29 + Ole6-6xHis (Seq ID No. 58).

Figure 79. Peptide sequence of: pET29 + Ole2-6xHis (Seq ID No. 59).

Figure 80. Peptide sequence of: pET29 + Ole4-6xHis (Seq ID No. 60).

Figure 81. Peptide sequence of: pET29 + Ole5-6xHis (Seq ID No. 61).

Figure 82. Peptide sequence of: pUCOle3+ (Seq ID No. 62).

Figure 83. Peptide sequence of: p29Ole (Seq ID No. 63).

Figure 84. Whole gel elution results from prokaryotic expression and purification of sesame seed oleosin.

A. Preliminary analysis of 1× - 6× AOBs by SDS-PAGE separation.

Small arrows indicate expected sizes of 4×, 5× and 6× oleosin (rectangular boxes). Large arrows indicate bands of putative polyoleosin protein corresponding to 1×, 2, and 3× oleosins (oval boxes). Expected molecular weights are shown in brackets.

B. Coomassie stained SDS-PAGE gel after whole gel elution (39C_sesame_oleosin_AOBs.doc).

L_K = BioRad Kaleidoscope Standards (cat# 161-0324)
Pre = Coomassie stained slice of acrylamide gel prior to elution
← = expected size of oleosin.

Figure 85. SDS-PAGE analysis of fractions from Ni-affinity purification of prokaryotically expressed denatured sesame seed oleosin (48C_sesame_oleosin_His_pur.doc).

PI = Post Induction S_{PL} = Supernatant, Post Lysis

L = Lysate FT = Flow Through

L_S = BioRad Prestained SDS-PAGE Standards, Low Range (cat# 161-0305)

$\#_A$ & $\#_B$ = Fraction number from either column A or B.

← = expected size of oleosin.

Figure 86. Analysis of rabbit anti-sesame seed oleosin antiserum by immunoblot and silver staining. Analysis of antibody titre and specificity.

LK = BioRad Kaleidoscope Standards (cat# 161 0324)
11 = 3ng of 1A 21 = 13ng 1A 31 = 27ng 1A 41 = 132ng 1A 51 = 663ng 1A
S1 & S2 = sesame seed extract C = clover seed extract
12 = 3ng of 1B 22 = 14ng 1B 32 = 29ng 1B 42 = 143ng 1B
I = crude post-induction sample LP = BioRad Precision Plus Standard (cat# 161 0363).
□ = expected size of oleosin. Overflow of the standards can be seen in lanes 11 and I.

The rabbit raised antibodies show a high affinity for the affinity purified recombinant oleosin, detecting down to at least 3ng. The antibodies also show high specificity, with no cross-reactivity against soluble seed proteins and low cross-reactivity to bacterial proteins (lane I).

Figure 87. Visualisation of emulsification layer containing artificial oil bodies (AOBs) after first (top) and second (bottom) rounds of sonication.

Figure 88. SDS-PAGE/immunoblot analysis of AOB 1×, 2×, 3×, 4× and 6× oleosin.

For analysis of the 6× polyoleosin two samples were loaded per lane. $_N$ = non-induced negative control.

Figure 89. Analysis of AOBs containing 1× - 6× Sesame seed polyoleosin proteins by SDS/urea-Gradient PAGE/SafeStain.

Figure 90. SDS, urea-PAGE/immunoblot of AOB Analysis of 1× - 6× AOBs by SDS/urea-PAGE/SafeStain.

Small arrow indicates expected position of 6× oleosin (rectangular box). Large arrows indicate bands of putative polyoleosin protein corresponding to 1×, 2, 3×, 4×, and 5× oleosins (oval boxes). Expected molecular weights are shown in brackets.

Figure 91. Microscopic analysis of AOB containing a different number of oleosin repeats. Size of AOBs when prepared with polyoleosins containing increasing numbers of repeating oleosin units.

Figure 92. Effect of polyoleosin on the stability over time of AOB containing a different number of oleosin repeats after 7 days at 4°C. AOBs were prepared with different polyoleosins containing increasing numbers of repeating oleosin units. Vector control AOBs were prepared using inclusion bodies from E. coli containing a vector control..

Figure 93. Microscopic analysis of size of AOB containing a different number of oleosin repeats after heat treatment at 90°C for 15min.

Figure 94. Heat stability of emulsification layers (containing AOB) in relation to the number of repeat oleosin units.

Figure 95. Microscopic analysis of AOB containing a different number of oleosin repeats after incubation in different pH buffers.

Figure 96. SDS-PAGE/immunoblot (A) and Coomassie (B) analysis of prokaryotically produced recombinant polyoleosin purified by Artificial Oil Body or Ni^{2+} affinity column.

Figure 97. Feature map of nucleotide sequence of attB flanking regions, single oleosin clone and UBQ10 in pRSh1-PSP1 (CaMV35S driving 1 oleosin tandem repeats with the UBQ10 intron in the first repeat) (Seq ID No. 64).

Figure 98. Feature map of nucleotide sequence of attB flanking regions, single oleosin clone and UBQ10 in pRSh1-PSP3 (CaMV35S driving 3 oleosin tandem repeats with the UBQ10 intron in the first repeat) (Seq ID No. 65).

Figure 99. Feature map of nucleotide sequence of attB flanking regions, single oleosin clone and UBQ10 in pRSh1-PSP4 (CaMV35S driving 4 oleosin tandem repeats with the UBQ10 intron in the first repeat) (Seq ID No. 66).

Figure 100. Feature map of nucleotide sequence of attB flanking regions, single oleosin clone and UBQ10 in pRSh1-PSP6 (CaMV35S driving 6 oleosin tandem repeats with the UBQ10 intron in the first repeat) (Seq ID No. 67).

Figure 101. Feature map of nucleotide sequence of attB flanking regions, single oleosin clone and UBQ10 in pRSh3-PSP1 (Arabidopsis oleosin seed promoter driving 1 oleosin repeat with the UBQ10 intron in the first repeat) (Seq ID No. 68).

Figure 102. Feature map of nucleotide sequence of attB flanking regions, single oleosin clone and UBQ10 in pRSh3-PSP3 (Arabidopsis oleosin seed promoter driving 3 oleosin tandem repeats with the UBQ10 intron in the first repeat) (Seq ID No. 69).

Figure 103. Feature map of nucleotide sequence of attB flanking regions, single oleosin clone and UBQ10 in pRSh3-PSP4 (Arabidopsis oleosin seed promoter driving 4 oleosin tandem repeats with the UBQ10 intron in the first repeat) (Seq ID No. 70).

Figure 104. Feature map of nucleotide sequence of attB flanking regions, single oleosin clone and UBQ10 in pRSh3-PSP6 (Arabidopsis oleosin seed promoter driving 6 oleosin tandem repeats with the UBQ10 intron in the first repeat) (Seq ID No. 71).

Figure 105. Feature map of nucleotide sequence of attB flanking regions, identical oleosin clones in pRSh1-Ole3+ (CaMV35s promoter driving 3 itentical oleosin tandem repeats; no intron) (Seq ID No. 72).

Figure 106. Feature map of nucleotide sequence of attB flanking regions, identical oleosin clones in pRSh1-Ole3+ (Arabidopsis oleosin seed promoter driving 3 itentical oleosin tandem repeats; no intron) (Seq ID No. 73).

Figure 107. Peptide sequence of: pRSh1-PSP1 (CaMV35S driving 1 oleosin tandem repeats with the UBQ10 intron in the first repeat) (Seq ID No. 74).

Figure 108. Peptide sequence of: pRSh1-PSP3 (CaMV35S driving 3 oleosin tandem repeats with the UBQ10 intron in the first repeat) (Seq ID No. 75).

Figure 109. Peptide sequence of: pRSh1-PSP4 (CaMV35S driving 4 oleosin tandem repeats with the UBQ10 intron in the first repeat) (Seq ID No. 76).

Figure 110. Peptide sequence of: pRSh1-PSP6 (CaMV35S driving 6 oleosin tandem repeats with the UBQ10 intron in the first repeat) (Seq ID No. 77).

Figure 111. Peptide sequence of: pRSh3-PSP1 (Arabidopsis oleosin seed promoter driving 1 oleosin repeat with the UBQ10 intron in the first repeat) (Seq ID No. 78).

Figure 112. Peptide sequence of: pRSh3-PSP3 (Arabidopsis oleosin seed promoter driving 3 oleosin tandem repeats with the UBQ10 intron in the first repeat) (Seq ID No. 79).

Figure 113. Peptide sequence of: pRSh3-PSP4 (Arabidopsis oleosin seed promoter driving 4 oleosin tandem repeats with the UBQ10 intron in the first repeat) (Seq ID No. 80).

Figure 114. Peptide sequence of: pRSh3-PSP6 (Arabidopsis oleosin seed promoter driving 6 oleosin tandem repeats with the UBQ10 intron in the first repeat) (Seq ID No. 81).

Figure 115. Peptide sequence of: pRSh1-Ole3+ (CaMV35s promoter driving 3 identical oleosin tandem repeats; no intron) (Seq ID No. 82).

Figure 116. Peptide sequence of: pRSh3-Ole3+ (Arabidopsis oleosin seed promoter driving 3 itentical oleosin tandem repeats; no intron) (Seq ID No. 83).

Figure 117. SDS-PAGE/immunoblot showing production of Sesame seed polyoleosin in Arabidopsis thaliana seeds. Analysis of oil bodies from transgenic *Arabidopsis thaliana* seeds over expressing varying tandem repeats of the sesame seed oleosin.

Figure 118. SDS-PAGE/Coomassie showing crude protein extracts from Arabidopsis thaliana wild type seeds and transgenic seeds expressing Sesame seed polyoleosin.

Figure 119. Picture of relative recombinant protein levels in Arabidopsis oil bodies and effects of different Sesame seed polyoleosin lengths on emulsifcation thicknesses after heating. SDS-PAGE/immunoblot analysis and emulsification stability analysis of oil bodies from transgenic *Arabidopsis thaliana* over expressing varying tandem repeats of the sesame seed oleosin and the increase in heat stability of the emulsification layers extracted from these transformants. The reduction in oleosin repeat numbers correlates with a decrease in the emulsification layer remaining after 24hr at 90°C, where the thickest emulsification can be seen from plants expressing the hexameric polyoleosin construct.

Figure 120. Picture of heating effect on emulsifcation thicknesses of transgenic Arabidopsis oil bodies expressing different polyoleosins. Heat stability analysis of the emulsification layers from transformants expressing 1, 3 or 6 oleosin repeats. The reduction in oleosin repeat numbers correlates with a decrease in the emulsification layer remaining after 24hr at 90°C, where the thickest emulsification can be seen from plants expressing the hexameric polyoleosin construct. At 0hrs through to 24hrs at 90°C the higher the number of oleosin repeats correlates with a thicker emulsification layer. After 58hrs at 90°C the oil bodies from the wild type arabidopsis has been reduced to a very thin layer with a small ring of emulsion deposited on the tube above the remaining aqueous phase. In comparison, the ring of emulsion remaining in the transformant samples is much greater.

EXAMPLES

EXAMPLE 1: WHITE CLOVER POLYOLEOSIN CONSTRUCTION

We have generated a dimeric, trimeric, tetrameric and a pentameric homo white clover oleosin tandem repeat construct and placed these into plant binary vectors for plant transformation. The following is a description of the methods used to generate the constructs. It should be noted that the oleosin sequence used is for example only. Any oleosin sequence or combinations of oleosins, steroleosins and caoleosins and oleosin linking sequences could be used.

Overview of experimental approach to generation of white clover oleosin repeats

Five oleosin clones (Pole01-MC to pOLE05-MC), containing suitable restriction sites, were prepared for the subsequent generation of constructs containing oleosin repeats. In this description clones pOLE01-MC to pOLE05-MC were used to prepare monomer, dimer, trimer, tetramer and pentamer oleosin repeat constructs, pOLE06-MC to pOLE10-MC respectively. The pBLUESCRIPT polylinker (Figure 8) was used in the preparation of pOLE01-MC to provide the multiple restriction sites required to build the oleosin repeat constructs.

Standard method of generating restriction enzyme/white clover oleosin cassette

PCR products were produced using a proof-reading polymerase to amplify the existing oleosin open reading frame (609 bases) of an oleosin clone containing no *Sac*I, *Xba*I, *Spe*I, *Bam*HI, *Sma*I, *Pst*I, *Eco*RI, *Hind*III *Bsp*106I, *Sal*I or *Xho*I sites. For each cassette unique primers were designed to amplify the oleosin sequence and add specific flanking restriction sites (Figure 9). Additional base pairs coding for specific amino acids were also included between the restriction site and oleosin priming sequence. In the polyoleosin peptide these amino acids could be expected to act as spacer arms enabling the individual oleosins to be separated from each other, reducing the chances of any misfolding occurring.

The oleosin PCR product to generate pOLE01-MC was then TOPO cloned into the pENTR/D-TOPO vector with the primary purpose of maintaining the flanked oleosin sequence. The oleosin PCR product to generate pOLE02, 03, 04, 05-MC was then TOPO cloned into the pCR2.1-TOPO vector with the primary purpose of maintaining the flanked oleosin sequence. Maps of the vectors used for cloning, pENTR/D-TOPO and pCR2.1-TOPO are shown in Figures 10 and 11, respectively.

The new construct was transformed into competent *E.coli* cells and plated out for single colonies. Individual colonies were cultured and the purified plasmid from these cultures was analysed by digestion with a range of suitable restriction enzymes. If no suitable colonies were identified then this process was repeated. Those colonies/plasmid preps that produced predicted patterns by restriction enzyme digestion were then sequenced across the cassette containing the oleosin insert and the resulting data compared against the predicted sequence for the construct/plasmid.

The construction of constructs pOLE01-MC to pOLE05-MC is summarised below, based on the method outlined above. Each cassette was fully sequenced after construction, with all sequences being confirmed as correct.

We have generated polyclonal antibodies to a number of white clover oleosin fragments, the generation and characterisation of these antisera is summarised below.

pOLE01-MC

(1) Primer design

Forward primer CACCXXXXXXXXXXXXXXXXX (Seq ID No. 1)

Reverse primer TCACTCGAGGAGCTCXXXXXXXXXXXXXXXXXX (Seq ID No. 2)

In the forward primer XXXX represents an oleosin specific primer that has a 5' end beginning with the first ATG site in the open reading frame.

In the reverse primer XXXX represents an oleosin specific primer that has a 5' end complementary to the sequence for last codon in the open reading frame.

(2) Outline of PCR product

CACC■■■■■■*SacI*/*XhoI*/TGA cloned into pENTR/D-TOPO to generate pOLE01-MC

■■■■■■ Oleosin ORF sequence

CACC pENTR/D-TOPO adapter sequence

SacI/*XhoI* Engineered restriction sites, each site adds 6 base pairs

TGA/TCA Stop codon

A map of pOLE01-MC is shown in Figure 12. The sequence of pOLE01-MC is shown in Figure 13.

pOLE02-MC

(3) Primer design

Forward primer TCTAGAGGTACTXXXXXXXXXXXXXXXXXX (Seq ID No. 3)

Reverse primer ACTAGTAGTACCXXXXXXXXXXXXXXXXXX (Seq ID No. 4)

In the forward primer XXXX represents an oleosin specific primer that has a 5' end beginning with the first ATG site in the open reading frame.

In the reverse primer XXXX represents an oleosin specific primer that has a 5' end complementary to the sequence for last codon in the open reading frame.

Outline of PCR product

*Xba*I/GGTACT■■■■■■GGTACT/*Spe*I cloned into pCR2.1-TOPO to generate pOLE02-MC

■■■■■■ Oleosin ORF sequence

GGTACT Extra bases to code for glycine and threonine.

A map of pOLE02-MC is shown in Figure 14. The sequence of pOLE02-MC is shown in Figure 15.

pOLE03-MC

Primer design

Forward primer CCCGGGGGTACTXXXXXXXXXXXXXXXXXX (Seq ID No. 5)

Reverse primer CTGCAGAGTACCXXXXXXXXXXXXXXXXXX (Seq ID No. 6)

In the forward primer XXXX represents an oleosin specific primer that has a 5' end beginning with the first ATG site in the open reading frame.

In the reverse primer XXXX represents an oleosin specific primer that has a 5' end complementary to the sequence for last codon in the open reading frame.

Outline of PCR product

*Sma*I/GGTACT■■■■■■GGTACT/*Pst*I cloned into pCR2.1-TOPO to generate pOLE03-MC

■■■■■■ Oleosin ORF sequence

GGTACT Extra bases to code for glycine and threonine

A map of pOLE03-MC is shown in Figure 16. The sequence of pOLE03-MC is shown in Figure 17.

pOLE04-MC

(4) Primer design

Forward primer AAGCTTGGTACTXXXXXXXXXXXXXXXXXX (Seq ID No. 7)

Reverse primer ATCGATAGTACCXXXXXXXXXXXXXXXXXX (Seq ID No. 8)

In the forward primer XXXX represents an oleosin specific primer that has a 5' end beginning with the first ATG site in the open reading frame.

In the reverse primer XXXX represents an oleosin specific primer that has a 5' end complementary to the sequence for last codon in the open reading frame.

Outline of PCR product

*Hin*dIII/GGTACT■■■■■■GGTACT/*Bsp*106I cloned into pCR2.1-TOPO to generate pOLE04-MC

■■■■■■ Oleosin ORF sequence

GGTACT Extra bases to code for glycine and threonine

A map of pOLE04-MC is shown in Figure 18. The sequence of pOLE04-MC is shown in Figure 19.

pOLE05-MC

(5) Primer design

Forward primer GTCGACGGTACTTCTXXXXXXXXXXXXXXXXXX (Seq ID No. 9)

Reverse primer CTCGAGXXXXXXXXXXXXXXXXXX (Seq ID No. 10)

In the forward primer XXXX represents an oleosin specific primer that has a 5' end beginning with the first ATG site in the open reading frame.

In the reverse primer XXXX represents an oleosin specific primer that has a 5' end complementary to the sequence for last codon in the open reading frame.

Outline of PCR product

*Sal*I/GGTACTTCT■■■■■■*Xho*I cloned into pCR2.1-TOPO to generate pOLE05-MC

■■■■■■ Oleosin ORF sequence

GGTACTTCT Extra bases to code for glycine, threonine and serine

A map of pOLE05-MC is shown in Figure 20. The sequence of pOLE05-MC is shown in Figure 21.

pOLE06-MC

pOLE01-MC and pBLUESCRIPT were digested using *Sac*I and *Xho*I. The pBLUESCRIPT fragment was ligated into the open pOLE01-MC vector to form

pOLE06-MC (Figure 22). The sequence of pOLE06-MC is shown in Figure 23.

CACC■*Sac*I/••••••/*Xba*I/*Spe*I/*Bam*HI/*Sma*I/*Pst*I/*Eco*RI/*Eco*RV/*Hin*dIII/*Bsp*106I/•/*Sal*I/*Xho*I/TGA

CACC pENTR/D-TOPO adapter sequence

■ Oleosin ORF sequence

*Sac*I to *Xho*I Restriction sites within pBLUESCRIPT

• amino acids encoded within the MCS but not corresponding to a restriction site

TGA Stop codon

Confirmation of the addition of the polylinker was carried out via digestion with *Sac*I, *Xho*I, *Xba*I, *Bam*HI, *Sma*I, *Pst*I, *Eco*RI, *Hin*dIII and *Sal*I. The pOLE06-MC cassette was then sequenced.

pOLE07-MC

pOLE06-MC was digested using *Xba*I and pOLE02-MC was digested with *Xba*I and *Spe*I, the resulting fragments were ligated together to form pOLE07-MC (Figure 24). The sequence of pOLE07-MC is shown in Figure 25.

CACC■*Sac*I/••••••/*Xba*I/••/■/••/*Spe*I/*Bam*HI/*Sma*I/*Pst*I/*Eco*RI/*Eco*RV/*Hin*dIII/*Bsp*106I/•/*Sal*I/*Xho*I/TGA

CACC pENTR/D-TOPO adapter sequence

■ Oleosin ORF sequences

*Spe*I to *Xho*I Restriction sites within pBLUESCRIPT

• amino acids encoded within the MCS but not corresponding to a restriction site

TGA Stop codon

Confirmation of the generation of pOLE07-MC was carried via *Xba*I/*Eco*RI, *Pst*I, and *Hind*III/*Pst*I digests. In order to gain a complete sequence of the clone, an LR reaction was performed to transfer the pOLE07-MC cassette into pET-DEST42. The pOLE07-MC cassette was then sequenced.

pOLE08-MC

pOLE07-MC and pOLE03-MC were digested using *Sma*I and *Pst*I, then the required fragments were ligated together to form pOLE08-MC (Figure 26). The sequence of pOLE08-MC is shown in Figure 27.

CACC■*Sac*I/••••••/*Xba*I/••/■/••/*Spe*I/*Bam*HI/*Sma*I/••/■/••/*Pst*I/*Eco*RI/*Eco*RV/*Hin*dIII/*Bsp*106I/•/*Sal*I/*Xho*I/TGA

CACC pENTR/D-TOPO adapter sequence

■ Oleosin ORF sequences

*Pst*I to *Xho*I Restriction sites within pBLUESCRIPT

• amino acids encoded within the MCS but not corresponding to a restriction site

TGA Stop codon

Confirmation of the generation of pOLE08-MC was carried out via *Pst*I/*Nco*I, *Not*I/*Xho*I and *Bam*HI digests. In order to gain a complete sequence of the clone, the pOLE08-MC oleosin cassette was transferred into pET-DEST42 by an LR reaction. The cassette was then sequenced and the sequence of the polylinker, and the pOLE03-MC oleosin fragment ligated into pOLE07-MC.

pOLE09-MC

pOLE08-MC containing the oleosin trimer, was digested using *Hind*III and *Cla*I (*Cla*I is an isoschizomer of *Bsp*106I), and pOLE04-MC was digested with *Hind*III and *Cla*I (Table 1), the resulting fragments were gel purified (Figure 28). The digests were purified from the agarose gel using the QIAGEN QIAquick Gel Extraction Kit. The *Hind*III/*Cla*I digested oleosin isolated from pOLE04-MC was then cloned into the *Hind*III/*Cla*I linearised pOLE08-MC vector (table 2) to generate pOLE09-MC (Figure 29).

CACC■*Sac*I/••••••/*Xba*I/••/■/••/*Spe*I/*Bam*HI/*Sma*I/••/■/••*Pst*I/*Eco*RI/*Eco*RV/*Hind*III/••/■/••*Cla*I/•/*Sal*I/*Xho*I/TGA

CACC pENTR/D-TOPO adapter sequence

■ Oleosin ORF sequences

*Cla*I to *Xho*I Restriction sites within pBLUESCRIPT

• amino acids encoded within the MCS but not corresponding to a restriction site

TGA Stop codon

4	8		
10		µL	pOLE04-MC
	10	µL	pOLE08-MC
1	1	µL	*Hind*III (Roche)
1	1	µL	*Cla*I (Roche)
2	2	µL	Buffer B (Roche)
6	6	µL	sH_2O
20	20	µL	Total volume
After 2.5h at 37°C added			
1	1	µL	Phosphatase (Roche, CAP, 20U/uL)
Incubated for a further 30min at 37°C.			

Table 1. Table describing the components used in the restriction enzyme digest to produce the linearised pOLE08-MC vector and the oleosin insert/fragment from pOLE04-MC

Ct8	Ct4	1:1	1:3	1:6		
3.5		3.5	3.5	3.5	µL	pOLE08-MC *Hind*III/*Cla*I linearised vector
	8	2	6	12	µL	Oleosin *Hind*III/*Cla*I fragment from
4	4	4	4	4	µL	Ligase buffer
1	1	1	1	1	µL	T4 DNA Ligase (Roche)
11.5	7	9.5	5.5		µL	sH_20
20	20	20	20	20.5	µL	Total volume
Incubated at 11°C overnight.						

Table 2. Table describing the components used in the ligation of the linearised pOLE08-MC vector and the oleosin insert/fragment from pOLE04-MC to produce pOLE09-MC.

5µL of each of the above ligation reactions were then transformed into *E. coli* DH5α to allow screening for correct pOLE09-MC plasmids.

Confirmation of the generation of pOLE09-MC was carried via agarose gel/ethidium bromide analysis of the *Nco*I/*Xho*I, *Nco*I/*Spe*I/*Xho*I, *Xho*I, and *Not*I/*Xho*I digests (Figure 30).

In order to gain a complete sequence of the clone, the pOLE09-MC oleosin cassette was transferred into pET-DEST42 by an LR reaction. The cassette was then sequenced (Figure 31).

pOLE10-MC

Due to poor restriction enzyme digestion efficiency when digesting the oleosin tetramer pOLE09-MC with *Sal*I, pOLE09-M was digested using only *Xho*I (Table 3). *Sal*I generated ends that are compatible with *Xho*I generated ends and thereby the *Sal*I/*Xho*I generated fragment was ligated into the *Xho*I linearised vector. pOLE05-MC was digested with *Sal*I and *Xho*I and the resulting oleosin fragment was ligated with pOLE09-MC (table 3) to form pOLE10-MC (Figure 32).

CACC■*Sac*I/••••••/*Xba*I/••/■/••/*Spe*I/*Bam*HI/*Sma*I/••/■/••*Pst*I/*Eco*RI/*Eco*RV/*Hin*dIII/••/■/••*Cla*I/•/**SalI**/••/■**XhoI**/TGA

CACC pENTR/D-TOPO adapter sequence

■ Oleosin ORF sequences

*Cla*I to *Xho*I Restriction sites within pBLUESCRIPT

• amino acids encoded within the MCS but not corresponding to a restriction site

TGA **Stop codon**

5	9		
x3	x2		
6		µL	pOLE05-MC (midipreped DNA)
	6	µL	pOLE09-MC (midipreped DNA)
1	2	µL	*Xho*I (Roche)
1	x	µL	*Sal*I (Roche)
2	2	µL	Buffer H (Roche)
10	10	µL	sH_2O
20	20	µL	Total volume
After 2.5h at 37°C added			
-	1	µL	Phosphatase (Roche, CAP, 20U/uL)
Incubated for a further 30min at 37°C.			

Table 3. Table describing the components used in the restriction enzyme digest to produce the linearised oleosin tetramer pOLE09-MC vector and the oleosin insert/fragment from pOLE05-MC.

The digests were purified from the agarose gel using the QIAGEN QIAquick Gel Extraction Kit. The *Xho*I/*Sal*I digested oleosin isolated from pOLE05-MC was then cloned into the *Xho*I linearised pOLE09-MC vector (Table 4) to generate pOLE10-MC.

Ct9	Ct5	1:3	1:9		
3	6	3	3	µL	pOLE09-MC *Xho*I/*Sal*I linearised vector
		3	9	µL	Oleosin *Xho*I/*Sal*I fragment from pOLE05-MC
1	1	1	1	µL	Ligase buffer
4	4	4	4	µL	T4 DNA Ligase (Roche)
12	9	9	3	µL	sH_20
20	20	20	20.5	µL	Total volume
Incubated at 11°C overnight.					

Table 4. Table describing the components used in the ligation of the linearised pOLE09-MC vector and the oleosin insert/fragment from pOLE05-MC to produce pOLE10-MC.

5µL of each of the above ligation reactions were then transformed into *E. coli* DH5α to allow screening for correct pOLE10-MC plasmids. Confirmation of the generation of pOLE10-MC was carried via *PstI*/*XhoI*, *PstI* and *SalI*/*PstI* restriction enzyme digests. One clone was selected for detailed agarose/ethidium bromide gel analysis, and compared to pOLE09-MC digested with the same enzymes (Figure 33).

In order to gain a complete sequence of the clone, the pOLE10-MC oleosin cassette was transferred into pET-DEST42 by an LR reaction. The cassette was then sequenced (Figure 34).

Generation of binary constructs

Clones pOLE06-MC through to pOLE10-MC were generated as GATEWAY™ entry vectors. The following is a description of subsequent LR reactions performed to transfer the oleosin constructs into two plant binary vectors, pRSh1 (Table 5 and Figure 35) and pRS12 (Table 5 and Figure 36).

Construct Name	Plasmid Type	Cassette Design	Bacterial Selectable Marker	Plant Selectable Marker
pRSh1	Plant binary	CaMV35s::attR1::GW::attR2::OCS3'	Spectinomycin^r	BASTA^r
pRS12	Plant binary	CaMV35s::attR1::GW::attR2::OCS3'	Spectinomycin^r	Kanamycin^r & GFP expression

Table 5. Table describing the components between the left and right borders of two plant binary vectors used to transform the multimeric oleosin repeats into plants.

Key:

attR1 GATEWAY™ recombination site

attR2 GATEWAY™ recombination site

CaMV35s **Ca**uliflower **m**osaic **v**irus promoter sequence

GFP Jelly Fish **G**reen **F**luoresence **P**rotein sequence

GW GATEWAY™ destination cassette

OCS3' **Oc**tapine **S**ynthase Terminator sequence

Overview of experimental approaches used to clone white clover oleosin repeats and DGAT1 into pRSH1.

Linerasing the entry clones

When cloning into pRSH1, linearising the entry clone seemed to make the reaction more efficient. 2µg of each of the six constructs were digested in 20µL volumes with the enzyme Alw441 (Roche) which cuts once in the back bone of pENTRD™. Digests were cleaned up using the QIAGEN QIAquick PCR Purification Kit.™.

LR reactions were set up between pRSH1 with pOLE06-MC to pOLE10-MC as well as with DGAT1 as per Table 6.

Component	Rxn 1 C6	Rxn 2 C7	Rxn 3 C8	Rxn 4 C9	Rxn 5 C10	Rxn 6 DGAT1
Linear entry clone	5.55	5.55	5.55	5.55	5.55	5.55
pRSH1 (1.5µg/µL)	0.2	0.2	0.2	0.2	0.2	0.2
LR Clonase™ (µL)	2	2	2	2	2	2
LR Rxn mix™ (µL)	2	2	2	2	2	2
Topoisomerase (µL)	0.25	0.25	0.25	0.25	0.25	0.25
H_2O (µL)	-	-	-	-	-	-
Total (µL):	10	10	10	10	10	10

Table 6. LR reactions for cloning oleosin repeats and DGAT1 into pRSH1.

C6 = pOLE06-MC etc.

The LR reactions were incubated overnight at 25°C, and were then transformed into *E. coli* DH5α. Four colonies were picked for each of the constructs C6-pRSH1, C7-pRSH1, C9-pRSH1 & C10-pRSH1, eight colonies were picked for C8-pRSH1 and three colonies for DGAT1-pRSH1.

Restriction enzyme digests to screen colonies

Plasmid DNA was extracted from clones, and purified plasmid was analysed by restriction digest using the following protocol:

µL	Component
2	Plasmid DNA (construct in pRSH1)
1	*Eco*RI (Roche)
2	Buffer H (Roche)
15	sH_2O
20	Total volume

Incubated at 37°C for 2 hours.

Digests were analysed by agarose/ethidium bromide gel electrophoresis and are shown in Figures 37-39.

Due to difficulties distinguishing between the two constructs C9-pRSH1 and C10-pRSH1 with the *Eco*RI digest clone1 of C9-pRSH1 and clone2 of C10-pRSH1 were digested with *Bam*HI to determine if both constructs were correct. Digests were analysed by agarose/ethidum bromide gel electrophoresis and are shown in Figures 40-41.

Overview of experimental approaches used to clone white clover oleosin repeats and DGAT1 into pRS12.

Linearising the binary plasmid

When cloning into pRS12, linearising the binary plasmid rather than the entry clone seemed to make the reaction more efficient. 2µg of pRS12 was digested in a 20µL reaction volume with the enzyme *Sma*I (Roche), which makes a single cut between the GATEWAY™ att recombination sites in pRS12. The digest was cleaned up using phenol:chloroform:isoamyl alcohol (25:24:1 v/v) and then chloroform. 100% recovery of DNA was assumed from this cleanup method.

LR reactions were set up between pRS12 with pOLE06-MC to pOLE10-MC as well as with DGAT1 as per Table 7.

Component	Rxn 1 C6	Rxn 2 C7	Rxn 3 C8	Rxn 4 C9	Rxn 5 C10	Rxn 6 DGAT1
Linear entry clone	1	1	1	1	1	1
pRS12 (2µg/µL)	0.2	0.2	0.2	0.2	0.2	0.2
LR Clonase™ (µL)	2	2	2	2	2	2
LR Rxn mix™ (µL)	2	2	2	2	2	2
Topoisomerase	0.25	0.25	0.25	0.25	0.25	0.25
H_2O (µL)	5.3	5.3	5.3	5.3	5.3	5.3
Total (µL):	10	10	10	10	10	10

Table 7. For LR cloning oleosin repeats and DGAT1 into pRS12.

C6 = pOLE06-MC etc.

The LR reactions were incubated overnight at 25°C, and were then transformed into *E. coli* DH5α. Twelve colonies were picked for analysis from the C6-pRS12 construct,

two colonies from C7-pRS12, two from C8-pRS12, three from C9-pRS12, three from C10-pRS12, and two from DGAT1-pRS12.

Restriction enzyme digests to screen colonies

Plasmid DNA was extracted from clones, and purified plasmid was analysed by restriction digest using the following protocol:

µL	Component
2	Plasmid DNA (construct in pRS12)
1	Restriction enzyme (Roche)
2	Buffer (Roche)
15	sH_2O
20	Total volume

Incubated at 37°C for 2 hours.

Digests were analysed by agarose/ethidium bromide gel electrophoresis and are shown in Figures 42-45.

Representative maps of a pRSH1 binary vector containing white clover oleosin multimers is shown in Figure 46.

Representative maps of a pRS12 binary vector containing white clover oleosin multimers is shown in Figure 47.

Generation of polyclonal antibodies to white clover oleosin

Cloning of N- terminal region of white clover oleosin into the pDEST17 and pET-DEST42 expression vectors

The N- terminal region of oleosin was cloned into the pENTR-D vector, to produce the construct pEON. The N- terminal region of oleosin was transferred into the pDEST17

and pET-DEST42 vectors (Invitrogen) by the LR reaction to produce the constructs p17NON (Figure 48) and p42CON (Figure 49), which were then transformed into *E. coli* DH5α. p17NON was designed to express the N- terminal oleosin peptide with an N-terminal 6xHis tag (6HON). The p42CON was designed to express the N- terminal oleosin peptides with a C-terminal 6xHis tag (ON6H).

Plasmid DNA purified from the putative constructs was analysed by PCR using T7 and T7t primers (as marked on the maps in Figure 48 and Figure 49), to determine if the terminal regions were present (Figure 50 and Figure 51).

Plasmid DNA from lines p42CON-2 and p17NON-1 were transformed into *E. coli* strains BL21 and BL21-AI for protein expression.

(b) Sequence analysis of putative p17NON and p42CON

The p17NON and p42CON plasmid lines were sequenced from the 5' and 3' ends using T7 and T7t primers, respectively. The sequences obtained showed that the N-terminal regions of oleosin had been cloned in frame into both the pDEST17 and pET-DEST42 vectors.

Expression of tagged peptides

Example of standard culture and induction protocol

10mL LB Amp100 broths were inoculated with expression construct in *E. coli* BL21-AI and incubated overnight at 37°C 220rpm. The next day duplicate 10mL LB Amp100 broths were inoculated with 500µL of the overnight cultures (to give OD_{600} = 0.05 - 0.1). All cultures were incubated for approximately three hours at 37°C 220rpm, until OD_{600} = 0.5 - 0.7 was reached. One set of the duplicate cultures was induced by the addition of 20% Arabinose to a final concentration of 0.2% and IPTG (isopropyl β-D-thiogalactopyranoside) to a final concentration of 1mM. The second set of duplicate cultures was used as the non-induced negative controls, and therefore nothing was added to these cultures. All cultures were then incubated overnight at 37°C

220rpm. 1mL aliquots were removed from each culture and prepared for SDS-PAGE analysis (Table 8).

1° broth	10mL LB Amp100 broths
1° inoculation	Colony picked from plate by toothpick
1° incubation conditions	Overnight at 37°C 220rpm
2° broth	10mL LB Amp100 broths (in duplicate)
2° inoculation	500µL of the overnight cultures (to give OD600 = 0.05 - 0.1)
2° incubation conditions	Approximately three hours at 37°C 220rpm, until OD600 = 0.5 - 0.7
Induction	**+ve** - 20% Arabinose to a final concentration of 0.2% and IPTG (isopropyl β-D-thiogalactopyranoside) to a final concentration of 1mM **-ve** - nothing added to non-induced negative control cultures
Induction incubation conditions	Overnight at 37°C 220rpm
Additional preparation	This may include harvesting and lysing cells
Analysis	1mL aliquots were removed from each culture and prepared for SDS-PAGE analysis

Table 8. Standard prokaryotic induction protocol.

Purification of tagged white clover oleosin peptides

HIS-Select Nickel Affinity Gel (Sigma) was used to purify the oleosin 6xHIS tagged proteins under denaturing conditions (Table 9 shows the buffers used).

Denaturing buffers

				Final vol. (mL) 100	100	100
		Stock (M)	Final conc (M)	EQUIL'N BUFFER	WASH BUFFER	ELUTION BUFFER
PO	mL	0.25	0.1	40	40	40
Urea (Mr=60.06)	g		6	36 g	36 g	36 g
NaCl	mL	5	0.3	6	6	6
Imidazole	mL	2	0.005	0.25		
Imidazole	mL	2	0.02		1	
Imidazole	mL	2	0.25			12.5
pH		5M HCl		8.0	8.0	8.0

Table 9. Buffers used to purify tagged white clover oleosin.

Example of standard large scale purification protocol

1.5mL of HIS-Select Nickel Affinity Gel was transferred to a chromatography column. The affinity gel was washed with two volumes of deionized water and then equilibrated with three volumes of equilibration buffer. The clarified crude extract was loaded onto the column at a flow rate of 2 to 10 column volumes/hour. After the extract was loaded, the column was washed with wash buffer at a flow rate of about 10 to 20 column volumes/hour. The 6xHis tagged protein was eluted from the column using 3 to 10 column volumes of elution buffer at a flow rate of 2 to 10 column volumes/hour. Fractions were collected continuously and assayed for the target protein using Bradford's protein assay.

N-terminal fragment of white clover oleosin

Induction of p17NON and the expression of 6HON (oleosin N-terminal with N-6xHis tag)

1° broth	10mL LB Amp100 broths
1° inoculation	p17NON in *E. coli* BL21-AI, colony picked from plate by toothpick
1° incubation conditions	Overnight at 37°C 220rpm
2° broth	10mL LB Amp100 broths
2° inoculation	500µL of the overnight cultures (to give OD600 = 0.05 - 0.1)
2° incubation conditions	Approximately three hours at 37°C 220rpm, until OD600 = 0.5 - 0.7
Induction	20% Arabinose to a final concentration of 0.2% and IPTG (isopropyl β-D-thiogalactopyranoside) to a final concentration of 1mM
Induction incubation conditions	Overnight at 37°C 220rpm
Analysis	1mL aliquots were removed from each culture and prepared for SDS-PAGE analysis (Figure 52)

Table 10. Large scale induction protocol.

Inclusion body preparation of 6HON for purification

1° broth	5mL LB Amp100 broths
1° inoculation	p17NON in *E. coli* BL21-AI
1° incubation conditions	Overnight at 37°C 220rpm
2° broth	50mL LB Amp100 broths
2° inoculation	3.5mL of the overnight cultures (to give OD600 = 0.05 - 0.1)
2° incubation conditions	Approximately three hours at 37°C 220rpm, until OD600 = 0.5 - 0.7
Induction	20% Arabinose to a final concentration of 0.2% and IPTG (isopropyl β-D-thiogalactopyranoside) to a final concentration of 1mM
Induction incubation conditions	Overnight at 37°C 220rpm
Additional preparation	The induced cells were collected by centrifugation and resuspended in 1/10 culture volume of buffer (50mM Tris-HCl pH8.0, 2mM EDTA). Lysozyme was added to a concentration of 100µg/mL, 1/10 volume of 1% Triton X-100 was then added, mixed gently and incubated at 30°C for 15 minutes. DNA was sheared by sonication and the insoluble protein fraction was separated from the soluble fraction by centrifugation.
Analysis	1mL aliquots were removed from each culture and prepared for SDS-PAGE analysis (Figure 53)

Table 11. Inclusion body preparation for *E. coli* expressed white clover oleosin.

Nickel Affinity Chromatography to purify 6HON

HIS-Select Nickel Affinity Gel (Sigma) was used to purify the 6HON under denaturing conditions (as described above).

A trial scale purification of 6HON was carried out using a 2mL sample from the insoluble protein fraction (as described above). The isolated 6HON was visualised using SDS-PAGE (Figure 54).

Large scale purification of 6HON

1° broth	50mL LB Amp100 broths
1° inoculation	p17NON in *E. coli* BL21-AI
1° incubation conditions	Overnight at 37°C 220rpm
2° broth	500mL LB Amp100 broths
2° inoculation	50mL of the overnight cultures (to give OD600 = 0.05 - 0.1)
2° incubation conditions	Approximately three hours at 37°C 220rpm, until OD600 = 0.5 - 0.7
Induction	20% Arabinose to a final concentration of 0.2% and IPTG (isopropyl β-D-thiogalactopyranoside) to a final concentration of 1mM
Induction incubation conditions	Overnight at 37°C 220rpm
Additional preparation	The induced cells were collected by centrifugation and resuspended in 1/10 culture volume of buffer (50mM Tris-HCl pH8.0, 2mM EDTA). Lysozyme was added to a concentration of 100µg/mL, 1/10 volume of 1% Triton X-100 was then added, mixed gently and incubated at 30°C for 15 minutes. DNA was sheared by sonication and the insoluble protein fraction was separated from the soluble fraction by centrifugation.
Analysis	1mL aliquots were removed from each culture and prepared for SDS-PAGE analysis

Table 12. Large scale preparation for prokaryotically produced white clover oleosin for antibody generation.

HIS-Select Nickel Affinity Gel was used to purify 6HON under denaturing conditions (as described above). A large scale purification of 6HON was carried out using the 500mL sample (induced as described above). The purified 6HON was visualised using SDS-PAGE (Figure 55) and the concentration of the eluted protein was estimated using a Bradford's protein assay, and found to be approximately 0.72, 2.74 and 1.89mg/mL, for Elutions 1, 2 and 3 respectively (Figure 56).

Generation of rabbit anti-white clover oleosin antisera

Immunisation of rabbit

66µL of Elution 2 (see Large scale purification of 6HON, above), containing 250µg purified protein, was mixed with 433µL PBS. This was passed to the Massey University Small Animals Unit for immunisation of a rabbit (male, ID# I24, Massey University Animal Ethics Committee (MUAEC) approval number 04/28). 25 days after the primary immunisation the first boost was administered (66µL of Elution 2, containing 250µg purified protein, mixed with 433µL PBS), and the second boost was given 46 days after the primary immunisation (128µL of Elution 1, containing 250µg purified protein, mixed with 372µL PBS).

Preparation of rabbit anti-white clover oleosin antisera

53 days after the primary immunisation approximately 3mL of blood was taken from the rabbit and placed at 4°C for approximately 16h. To separate the serum the blood was centrifuged at 1500xg 20min 4°C. The serum (1.5mL) was transferred to a fresh tube and stabilised with 0.25% Phenol in PBS and 0.01% Methiolate.

Assaying titre and specificity of rabbit anti-white clover oleosin antisera by SDS-PAGE immunoblot analysis

The samples outlined in Table 13 were analysed by SDS-PAGE immunoblot analysis using the anti oleosin N terminal antibody (Figure 57).

Lane	ID	Load	Oleosin	Induced	Expressed
1	FLA-BL21_AI, *E. coli*	1µL	6HOF	-	-
2	FLA-BL21_AI, *E. coli*	1µL	6HOF	+	+
3	17C-2-BL21_AI, *E. coli*	1µL	6HOC	+	+
4	Elu1, 1:200, pure	1µL	6HON, 1.8ng	+	+
5	Elu2, 1:200, pure	1µL	6HON, 6.9ng	+	+
6	Elu3, 1:200, pure	1µL	6HON, 4.7ng	+	+
8	clover seed extract	0.1µL	1.6µg total protein		
9	clover seed extract	1µL	16µg total protein		
10	clover seed extract	5µL	80µg total protein		

Table 13. Samples used to determine titre and specificity of polyclonal rabbit anti-white clover oleosin antisera.

Using rabbit anti-white clover oleosin antisera to detect oleosin in clover seed oil bodies

We checked to see if the antibodies were capable of detecting full-length oleosin. Early analysis of the antibodies on crude clover seed protein extracts had shown that the antibodies were able to be used to detect a protein of the expected size when clover seed oil body extracts were analysed by Immunoblot analysis (Figure 58).

Results from the immunoblot showed that the anti-oleosin N-terminal antibodies could be used to specifically detect a ~20kDa protein - the expected size of the clover oleosin protein.

Overview of experimental approaches used to transform plants with white clover polyoleosin clones.

Transformation of *Brassica oleracea* with white clover polyoleosin and DGAT1 constructs

A total of 12 constructs were transformed into *Brassica oleracea* (Table 14. These were all transferred into *Agrobacterium tumefaciens* strain LBA4404 via the freeze-thaw method described in Christey and Braun (2004). Colonies were selected on LB medium containing 100ml/l streptomycin and 100ml/l spectinomycin. PCR analysis for the BAR gene was used to confirm plasmid presence in *Agrobacterium* colonies.

Construct	C&F code	Genes	Marker
pRSH1-C6	AgR 6	35S-oleosin monomer	35S-BAR
pRSH1-C7	AgR 7	35S-oleosin dimer	35S-BAR
pRSH1-C8	AgR 8	35S-oleosin trimer	35S-BAR
pRSH1-C9	AgR 9	35S-oleosin tetramer	35S-BAR
pRSH1-C10	AgR 10	35S-oleosin pentamer	35S-BAR
pRSH1-At DGAT	DGAT	35S-Arab. DGAT	35S-BAR
pRSh4		Arab. oleosin seed promoter + GUSi	35S-BAR
pRSh6		oleosin seed prom.+ oleosin monomer	35S-BAR
pRSh7		oleosin seed prom.+ oleosin dimer	35S-BAR
pRSh8		oleosin seed prom.+ oleosin trimer	35S-BAR
pRSh9		oleosin seed prom.+ oleosin tetramer	35S-BAR
pRSh10		oleosin seed prom.+ oleosin pentamer	35S-BAR

Table 14. Summary of constructs placed into *Brassica oleracea*.

Plant material and transformation

For these experiments a rapid cycling (RC) *B. oleracea* line, DH1012, selected for high regeneration and transformation ability by Sparrow *et al.* (2004) was used. Seeds were germinated *in vitro* as described in Christey *et al.* (1997). Hypocotyl and cotyledonary petiole explants from 4-5 day-old seedlings were co-cultivated briefly with a culture of *Agrobacterium* grown overnight in LB medium containing antibiotics prior to 1:10 dilution in antibiotic-free minimal medium (7.6mM $(NH_4)_2SO_4$, 1.7mM sodium citrate, 78.7mM K_2HPO_4, 0.33M KH_2PO_4, 1mM $MgSO_4$, 0.2% sucrose) with growth for a further 4 hrs. Explants were cultured on Murashige-Skoog (MS; Murashige and Skoog, 1962) based medium with B5 vitamins and 2.5mg/L BA. After 3 days co-cultivation, explants were transferred to the same medium with 300mg/L Timentin (SmithKline Beecham) and subsequently placed on selection medium with 2.5mg/L Basta. Green shoots were transferred to hormone-free Linsmaier-Skoog based medium (LS; Linsmaier and Skoog, 1965) containing 5mg/L Basta and solidified with 10gm/L Danisco Standard Agar. Explants were cultured in tall Petri dishes (9cm diameter, 2cm tall) sealed with Micropore (3M) surgical tape. Shoots were cultured in clear plastic tubs (98mm, 250ml, Vertex). All culture manipulations were conducted at 25°C with a 16h/day photoperiod, provided by Cool White fluorescent lights (20 $\mu E/m^2/s$).

Determination of selective levels of Basta

Hypocotyl and cotyledon explants from 4 day-old seedlings were cultured on LSN medium containing 2.5mg/L BAP and 300mg/L Timentin and 5 different levels of Basta (0, 2.5, 5, 10 and 25mg/L) to determine the appropriate level to use for selection in co-cultivation experiments.

To determine the selective level of Basta for maintenance of *in vitro* shoots, small healthy apical cuttings were transferred to tall pots of hormone-free LSn containing either 2.5 or 5mg/L Basta.

Conformation of *Brassica oleracea* transformation by PCR

DNA was isolated from leaves of *in vitro* shoots using the rapid method described in Christey and Braun (2004). PCR was conducted to test for the presence of the bar

gene using the primers baran (5'-tcagatctcggtgacgggcagg-3') (Seq ID No. 11) and barse (5'-atgagccaagaacgacgcccgg-3') (Seq ID No. 12) to produce a 551bp product. PCR conditions were: 94°C for 30sec, 65°C for 30sec and 72°C for 30sec over 40 cycles.

PCR was conducted on Bar positive samples to check for the presence of the oleosin clones using the primers outlined in Table15. PCR conditions were: 94°C for 30 s, 45°C for 30 s and 68°C for 30 s for 30 cycles. A 650bp product was expected.

Primer name	Sequence	Sequence IDs
AgR 6 forward	CACCATGGCACAACCTCAAGTT	SEQUENCE ID NO. 13
AgR7 forward	GCTCCACCGCG	SEQUENCE ID NO. 14
AgR 6/7 reverse	TCCACTAGTTCTAG	SEQUENCE ID NO. 15
AgR 8 forward	CTAGAACTAGTGGAT	SEQUENCE ID NO. 16
AgR 8 reverse	TCCTGCAGAGTAC	SEQUENCE ID NO. 17
AgR 9 forward	GTACTCTGCAGGA	SEQUENCE ID NO. 18
AgR 9 reverse	ATACGATAGTAC	SEQUENCE ID NO. 19
AgR 10 forward	CTATCGATACCGTCG	SEQUENCE ID NO. 20
AgR 10 reverse	CTCGAGTGTTGATCTCTTAGCTTC	SEQUENCE ID NO. 21

Table 15. Primers used for oleosin PCR.

Conformation of white clover polyoleosin construct gene expression in *Brassica oleracea* transformants by RT-PCR

RNA was isolated from plants using the Concert™ small scale isolation reagent (Invitrogen). DNA was digested using Turbo DNase (Ambion). For RT-PCR the primers described in Table 15 were used to amplify the transcribed region. Reactions were performed on RNA samples using the Superscript™ III (Invitrogen) Reverse Transcriptase, with the RT step performed at 50°C followed by PCR.

Agrobacterium transformation

PCR for the Bar gene was used to confirm the presence of the oleosin-containing plasmids in *Agrobacterium* colonies (Figure 59). Two BAR positive colonies from each transformation event were selected for storage at –80°C in glyercols and used in subsequent explant co-cultivation experiments.

Determination of selective levels of Basta

The results of the explant selection experiment confirmed that use of 2.5mg/L Basta was suitable for use as the selection level for subsequent co-cultivation experiments. After 2-4 weeks on this level rare shoot regeneration was noted (Figure 60). In contrast, *in vitro* shoot cuttings showed some development on 2.5mg/L Basta with root development apparent on some. In contrast, after 4-5 weeks on 5mg/L Basta cuttings turned brown with no root development. This level was used in subsequent experiments for growth of putative transgenic shoots.

Selection of transgenic shoots

At least 150 explants were co-cultivated with each construct (Table 16). Three to four weeks after co-cultivation rare small green shoots were apparent. Putative Basta resistant shoots were excised and transferred to hormone-free media containing 5mg/L Basta. *In vitro* shoot and root growth on Basta-containing medium confirmed plants were transgenic. Transgenic plants were healthy with good shoot and root development and normal phenotype. In contrast, Basta sensitive cuttings showed little development (Figure 61). Transgenic shoots were obtained mainly from hypocotyl explants with only rare transgenic shoots obtained from cotyledon explants. While the overall transformation rate was only 1.2%, there was a lot of variation between both experiments and explants, with some combinations producing over 5% transformation in some experiments.

Construct	No. explants cocultivated		No. PCR +ve for BAR		Total
	cotyledon	hypocotyl	cotyledon	hypocotyl	
pRSH1-C6	70	90	0	6	6
pRSH1-C7	170	110	1	3	4
pRSH1-C8	390	190	2	2	4
pRSH1-C9	390	190	0	4	4
pRSH1-C10	250	120	3	7	10
At DGAT	416	195	0	8	8
pRSh4	200	120	2	2	4
pRSh6	100	50	0	0	0
pRSh7	110	30	0	1	1
pRSh8	110	50	1	2	3
pRSh9	110	50	0	0	0
pRSh10	110	50	1	0	1
Total	**2426**	**1245**	**10**	**35**	**45**

Table 16. Summary of PCR positive shoots.

Conformation of transformation by PCR

Shoots growing on 5mg/L Basta that were still green and showing good root development after at least two subcultures had DNA extracted and PCR performed for the bar gene. Over 250 shoots were analysed with 45 confirmed as having the bar gene (20%) (Table 10). PCR analysis confirmed the presence of the Bar gene in all plants that had shown good growth on selective levels of Basta in vitro. PCR for the *Bar* gene produced a 551bp fragment (Figure 62). PCR analysis of Bar positive shoots for the oleosin gene indicated that not all plants contain the oleosin gene as well (Figure 63).

Transgenic shoots were obtained for 10 of the 12 constructs in Table 14. Most constructs had at least 4 independent transgenic shoots but no transgenic shoots were obtained for pRSh6 and pRSh9. Negative shoots were retained on 5mg/L Basta in case

they were false negatives and re-tested if they remained green with roots. Generally PCR negative shoots showed Basta sensitivity symptoms in subsequent subculture rounds suggesting negative PCR results were genuine. For all transgenic lines 1-2 pots of clonal cuttings remain *in vitro*. All lines appeared healthy with good root and shoot development and normal phenotype.

Conformation of expression by RT-PCR

RT-PCR analysis confirmed expression of the oleosin construct in 4 out of 5 lines tested (Figure 64).

Transformation of Lotus japonicus with Agrobacterium rhizogenes

Northern analysis of RNA extracted from hairy roots

Northern blot analysis was performed on transformed *Lotus japonicus* hairy roots (see below) to confirm the expression of the polyoleosin constructs. The blot was probed with ^{32}P labled, random primed, white clover oleosin cDNA (Figure 65). The Northern blot analysis shows that the constructs are transcribed. In each case it appears that transcription has finished close to the 3' end of the OCS terminator where a single oleosin transcript and the terminator are roughly 1.5kb in length. It is clear that the transcript from the single oleosin is highly expressed in the majority of the single oleosin transformants. Each additional oleosin repeat resulted in the increase in the transcript size by the appropriate amount (~650b). As such that the sizes are as follows 1.4kb, 2kb, 2.7kb, 3.4kb and 4kb for the oleosin monomer, dimer, trimer, tetramer and pentamer, respectively. The level of transcript accumulated appears to be inversely proportional to the number of oleosin repeats. Furthermore, the transformants containing either the oleosin monomer or dimer did not have any transcripts of unexpected size hybridising to the oleosin probe, whereas the transformants containing the oleosin trimer, tetramer or pentamer had a number of additional hybridising bands. In the majority of cases the bands were considerably larger than the expected size, frequently >4.4kb, and in only one case did the hybridising band appear to be smaller than the predicted transcript. Since the extra bands only appear in lines transformed

with the trimeric or larger oleosin repeats, we have assumed that these bands represent aberrantly processed oleosin RNA, and in many cases the processing appears to have failed to terminate transcription.

Transformation of *Lotus japonicus* with *Agrobacterium rhizogenes* (A4T)

Protocol

Day 1

Scarify lotus seeds using p220 wet/dry sand paper

Sterilise seeds by rotating for 20 min in 10mL sterilisation soln:

7 mL 100% ethanol
1 mL 30% H_2O_2
2 mL H_2O

Wash 3 times in sterile H_2O

Place seeds on 1% water agar plates

Wrap plates in tinfoil (dark) and germinate at 25°C for 2 days

Streak TY agar plate with Agrobacterium rhizogenes (A4T) glycerol stock and grow overnight @ 28°C

Day 2

Inoculate 50mL YEB culture broth with colony from A4T plate and grow overnight @ 28°C shaking (220rpm)

Day 3

Make Agrobacterium competent cells and transform with binary plasmid containing gene of interest, plate on TY agar plates and grow for 2 days at 28°C. (refer: Transformation of Agrobacterium)

Transfer germinated seeds to ½ B5 media, approx 10 across each plate, roots pointing down.

Tape plates together, grow vertically on lab bench.

Day 5

Pick colonies from Agrobacterium plates into 10mL TY-spectomycin broths and grow at 28°C shaking (220rpm) for 2 days.

Day 6

Perform PCR on Agrobacterium broths to check for desired gene.

Day 7

Inoculate Lotus japonicus plants by dipping a sterile scalpel into the Agrobacterium broth and cutting off the root. After inoculation tape plates together, wrap in tinfoil and leave overnight on lab bench

Day 8

Unwrap plates and grow for 2 days vertically on lab bench

Day 9

Transfer plants to MS(CRO) media containing the antibiotic cephotaximine, 10 across a plate. Grow vertically on lab bench.

Roots can be viewed (for GFP) under a Microscope 10-20 days later.

Media

½ B5 media (No sucrose)

0.0425	g	$NaH_2PO_4.2H_20$
0.625	g	$KN0_3$
0.0335	g	$(NH_4)_2S0_4$
0.0625	g	$MgS0_4.2H_20$
0.01	g	Ferric EDTA
0.025	g	Myo-Inositol
0.25	mL	Stock A
0.25	mL	Stock B
0.25	mL	Stock C
0.25	mL	Stock D

Adjust pH to 5.5 with 0.2M KOH or 0.2M HCl

6	g	Agar

Make up to 500mL with sterile H20

MS(CRO) media

50	mL	MS Macro Stock
5	mL	MS Macro Stock
5	mL	MS Fe (EDTA) Stock
1	mL	B5B Vitamins stock
30	g	Sucrose
100	mg	Myo-Inositol
8	g	Phytagel agar

pH to 5.7 with NaOH
Final volume = 1000mL

MS Macro Stock

33	g	NH_4NO_3
38	g	KNO_3
8.8	g	$CaCl_2.2H_2O$
3.4	g	KH_2PO_4
7.4	g	$MgSO_4.7H_2O$

Final volume = 1000mL

MS Fe (EDTA) Stock

4	g	Ferric EDTA (Fe Na Ethylene diaminetetra acetic)

Final volume = 500mL

MS Micro Stock

1.24	g	H_3BO_3
4.46	g	$MnSO_4.4H_20$
1.72	g	$ZnSO_4.7H_20$
0.166	g	KI
0.05	g	$Na_2MoO_4.2H_20$
0.005	g	$CuSO_4.5H_20$
0.005	g	$CoCl_2.6H_20$

Final volume = 1000mL

B5 Vitamin Stock

0.1	g	Nicotinic Acid
1.0	g	Thiamine HCl
0.1	g	Pyridoxine HCl

Final volume = 100mL

1 mL aliquots into eppendorfs

Store at -20°C

Notes

Transformation of Lotus:

This protocol includes the transformation of Agrobacterium which sometimes is just being grown from glycerols. Also is written as for GFP - if for PPT resistant plants then on day 9 (or just when roots are visible) we transfer to selective plates. Usually 4 plants (roots only) per plate. Once selection has taken place (either GFP or PPT) individual roots are transferred to individual plates. Supposedly we transfer all or some of the tissue to a new plate each month. Once we are growing root portions only the plates are kept horizontal, sealed and in darkness, usually in the 25 C room.

Transformation of *Agrobacterium rhizogenes* (A4T)

1. Streak a TY agar plate with *Agrobacterium rhizogenes* (A4T) glycerol stock and grow 28°C overnight.
2. Inoculate 50mL of YEB broth with a colony from Agrobacterium plate and grow at 28°C, shaking (220rpm) until OD_{600} is approx 0.5 (16h)
3. Centrifuge cells for 15 min @ 4000rpm, discard supernatant and resuspend in 10mL of **0.15M NaCl**
4. Centrifuge cells for 10 min @ 4000rpm, discard supernatant, and resuspend in 1mL of ice-cold **20mM $CaCl_2$**
5. Aliquot 200µL of cells into an eppendorf tube, add 5µg of DNA and incubate on ice for 30 min.
6. With what is left of the 1mL aliquot 186µL of cells and 14µL of DMSO into eppendorf tubes and freeze in liquid N_2 then store at -70°C.
7. After incubation on ice for 30mins freeze the DNA/cells in liquid N_2 for 1 min.
8. Thaw in a 37°C waterbath
9. Repeat steps 7 & 8

10. Add 1mL of YEB broth and incubate cells for 4h @ 28°C with gentle shaking
11. Plate cells on TY agar containing spectomycin and grow for 2 days @ 28°C.
12. Pick colonies from the *Agrobacterium* plates into 10mL TY broths containing spectomycin and grow for 2 days @ 28°C, shaking at 220rpm.

Agrobacterium transformed with a binary vector are prepared for infiltration by plating or spreading a bacterial lawn on the appropriate plate with appropriate antibiotics (although Rifampicin is not normally included). The *Agrobacterium* should be grown at 28°C for 2-3 days.

Clover seed oilbody purification

Buffers for oilbody extraction

Buffer A

4	mL	500mM NaPhosphate Buffer, pH 7.5
41.1	g	sucrose

Final volume = 200mL (with dH_2O)

Buffer D

4	mL	500mM NaPhosphate Buffer, pH 7.5
17.1	g	sucrose
23.4	g	NaCl

Final volume = 200mL (with dH_2O)

Detergent washing solution

0.4	mL	10% Tween 20 (0.1%)
20	mL	Floating Buffer (½x Buffer D)

Final volume = 40mL (with dH_2O)

Procedure

1. Weigh 2g of clover seed.
2. Homogenate the seed in ~20mL of **Buffer A** with sand in a mortar and pestle.
3. Transfer to two 15mL Falcon tubes and centrifuge 3200xg 10min RT.
4. Transfer upper aqueous phase and any floating material to two 15mL Falcon tubes and add 1mL **Buffer D**.
5. Resuspend the oil bodies thoroughly by sonication Power-1 Duty Cycle-100% 5x30sec.
6. Centrifuge 10,000rpm 10min RT.
7. Transfer upper/floating oil-body layer at the top with a bended spatula to 2mL microfuge tubes containing 1.5mL of **Detergent Washing Solution**.
8. Resuspend the oil bodies thoroughly by sonication Power-1 Duty Cycle-100% 5x30sec.
9. Centrifuge 10,000rpm 10min RT.
10. Transfer upper/floating oil-body layer at the top with a bended spatula to 2mL microfuge tubes containing 1.5mL of **Buffer A**.
11. Resuspend the oil bodies thoroughly by sonication Power-1 Duty Cycle-100% 5x30sec.
12. Centrifuge 10,000rpm 10min RT.
13. Transfer upper/floating oil-body layer at the top with a bended spatula to 2mL microfuge tubes containing 500µL of **Buffer A**.
14. Resuspend the oil bodies thoroughly by vortexing and prepare samples for SDS-PAGE analysis.

RNA isolation from plants using the QIAGEN RNeasy kit

RNA was isolated from *Agrobacterium tumefaciens* infected *N. benthamiana* leaves and *Agrobacterium rhizogenes* infected *L. japonicus* roots using the RNeasy Plant Mini Kit (QIAGEN, Hilden, Germany). Approximately 0.1g of tissue was removed from the plant, placed into a 1.5mL microcentrifuge tube, frozen in liquid nitrogen and ground to a fine powder using a stainless steel rod. 450µL RLT Buffer (guanidinium isothiocyanate) and 4.5µL ß-mercaptoethanol were added and the mixture vortexed vigorously until the

sample thawed then for a further 30sec. The lysate was transferred onto a QIAshredder spin column (lilac), fitted in a 2mL collection tube (supplied) using a 1mL pipette with the end of the tip cut off to prevent blockages. The mixture was spun in microfuge for 2min at 13000rpm. The filtrate was transferred to a 1.5mL centrifuge tube taking care not to disturb the debris in collection tube. 225µL Absolute Ethanol (0.5vols.) was added and mixed immediately by pipetting. The filtrate/ethanol mix, and any precipitate, was decanted onto the RNeasy mini spin column (pink) which had been fitted in a 2mL collection tube (supplied). The mixture was spun in a microfuge for 15sec at 13000rpm. The filtrate was then discarded, but the column and collection tube were retained for the next step. To column was washed by loading on 700µL RW1 Buffer then spun in a microfuge for 15sec at 13000rpm. The filtrate was discarded, and 500µL RPE Buffer was loaded onto the column, which was then spun in a microfuge for 15sec at 13000rpm. Again, the filtrate was discarded and another 500µL RPE Buffer was loaded onto the column and spun for 15sec at 13000rpm. The filtrate was discarded, and the column was spun for an additional min at 13000rpm to thoroughly dry the column. Next the column was carefully removed from the collection tube, avoiding the carry-over of ethanol/wash buffer. The RNeasy column was then placed into a fresh 1.5mL collection tube (supplied) and 30µL RNase–free H_2O was added onto the column. The column was spun in a microfuge for 60sec at 13000rpm to elute the RNA. The RNA was stored at -20°C until required.

RT-PCR

cDNA synthesis.

Expand Reverse Transcriptase, Roche, cat#1785826, supplied with 5 x Buffer and DTT.

	8.5-9.5	µL	Total RNA (*1µg recommended*)
either	2	µL	10µM gene specific reverse primer
or	1	µL	Oligo(dT)$_{12-18}$ @ 100 pmoles/ µL
Total volume	10.5 µL		(with RNase-free sterile H_2O)

Denatured RNA and primer for 10min at 65°C.

Immediately cooled on ice.

Added respective reagents separately to each tube in the following order (no cocktail):

4	µL	5 x Expand reverse transcriptase buffer (first-strand) (1x)
2	µL	100mM DTT (10mM)
2	µL	10mM dNTP Mix (dATP+dCTP+dGTP+dTTP, Roche cat#1581295; 1mM)
0.5	µL	RNase Inhibitor (Roche, cat#799017, 40U/µL; 20U)
1	µL	Expand Reverse Transcriptase (Roche cat#11 785 834 001; 50U/µL; 50U)

Total volume 20 µL(with RNase-free sterile H_2O)

Mixed and pulse spun.

Incubated at 43°C for 45-60min.

Quenched on ice.

Used immediately.

First Round PCR reaction.

5	µL	cDNA
1	µL	10mM dNTP Mix (dATP+dCTP+dGTP+dTTP, Roche cat#1581295; 200 µM)
0.3	µL	25µM Forward primer (150 nM)
0.3	µL	25µM Reverse primer (150 nM)
5	µL	10x PCR reaction buffer with $MgCl_2$ (Roche cat#1 647 679; 1x)
1	µL	Taq DNA Polymerase (Roche cat#1 647 679; 1U/ µL;1U)

Total volume 50µL (with sH_2O)

First Round PCR amplification.

95°C	2:00		1 cycle
95°C	0:30	]	
58°C	0:30	]	30 cycles
72°C	2:30	]	
72°C	7:00		1 cycle

Northern analysis of RNA

4pg pMC6 (Clone 6) used as template.

Amplified using pMC1 (Clone 1) forward and reverse primers and standard Taq reaction.

Extracted bands purified using QIAGEN Gel Purification kit.

Product quantified to 55ng/µL.

Formaldehyde Gels

***All water should be sourced fresh from the still.

10x MOPS/EDTA Buffer

(A) 52.3 g MOPS (free acid, Mr=209, 500mM)
Adjust pH to 7.0 with 10N NaOH.
Final vol 250mL.

(B) 1.86 g EDTA (Na salt, Mr=372, 10mM, or 1mL 500mM EDTA in 500mL)
Adjust pH to 7.5 with 10N NaOH.
Final vol 250mL

Add 250mL MOPS to 250mL EDTA to obtain correct final concentrations.

Store in the dark at room temperature.

The buffer may yellow with age or autoclaving but this will not affect buffering capacity.

Buffer A

300 µL 10x MOPS/EDTA Buffer
Final vol 1mL.

Formaldehyde/Formamide

89 µL AR Formaldehyde (37-40%)
250 µL Formamide (TOXIC. Fumehood and Gloves)

Gel Loading Buffer

400 mg Sucrose
320 µL Buffer A
Mix to dissolve then add

5 mg Xylene cyanol
5 mg Bromophenol blue
Mix to dissolve. Add

178 µL AR Formaldehyde (37-40%) (TOXIC. Fumehood and Gloves)
500 µL Formamide (TOXIC. Fumehood and Gloves)
Store in 100 µL aliquots at -20oC

Running Buffer (= 1x MOPS/EDTA)

225 mL 10x MOPS/EDTA Buffer
Total vol = 2.25L

Sample Preparation

30 µg RNA (as a dry pellet)
5.5 µL Buffer A
Resuspend RNA. Add

12 µL Formamide/Formaldehyde (TOXIC. Fumehood and Gloves)
Heat to 70°C (heating block) for 10 min then quench on ice.

Add 4.5 µL Gel Loading Buffer
Load onto submerged gel.

Pouring Formaldehyde Gel

Prepare gel tray by wiping out with ethanol and taping both ends with a double layer of masking tape.

250	200	125	mL	Final vol.
2.5	2.0	1.25	g	Agarose
25	20	12.5	mL	10x MOPS/EDTA Buffer
180	144	90mL	H_2O	

For today's work 2 gels 150x100mm prepared with 30-tooth combs in each.

Mix in a 500mL flask and dissolve in microwave. Stir while cooling.

All of following steps should be done in the fumehood:

Add	45mL	36mL	22.5	mL AR Formaldehyde (37-40%)

Mix and pour into prepared gel tray, with comb in place.

Allow to set for 30-45min.

Cover with Running Buffer, load and run initially at 50V for 15min to get samples into gel, then turn voltage down to 12V and run overnight.

Northern Blot

20x SSC

175.3 g AR NaCl (=3M)
88.2 g AR Sodium citrate (=0.3M)
Adjust pH to 7.0

Total volume 1L

This Northern Blot method uses the flow of buffer **DOWN** through a gel then a nylon membrane to transfer denatured RNA from the gel onto the membrane. Based on the design of the Schleicher & Schuell TurboBlotter.

1. Two 3cm stacks of interfolded paper hand towels (Hygenex Royale, 245x270mm, cat#2170360) were placed side-by-side in a plastic tray
2. Two sheet of 3MM paper, cut to 220x160mm were placed on top, followed by one sheet pre-wetted in 10xSSC.
3. The nylon membrane (Hybond-XL (Amersham; cat no: RPN303S) was cut to 220x160mm, pre-wetted with dH_2O, rinsed for 1min in 10xSSC, and finally placed onto the stack.
4. The gel was removed from the gel plate and washed for 20min in 10xSSC.
5. The washed gel was then placed onto the membrane ensuring that no bubbles were trapped between the gel and the membrane. **Note** care was taken to ensure that the gel was positioned correctly as transfer may be initiated as soon as the gel comes into contact with the membrane.
6. Three sheets of pre-wetted 3MM paper, cut to 230x170mm, were layered onto the stack. Again ensuring that no bubbles were trapped between the layers.
7. Finally the pre-wetted 3MM wick, cut to ~800x180mm, was placed across the stack, with each end folded over and submerged in ~100mL 20xSSC, which was in containers at the side of the blot/stack.

   NOTE

   1. The wick was kept flat at all times (*i.e.* the containers were long enough so that the wick did not have to be crammed in).
   2. The sides of the containers were only 1-2cm high so that the capillary action could easily draw the 20xSSC up and out of the reservoirs.

Next day

1. The blot was carefully taken apart down to the gel. A soft leaded pencil was pushed through the wells, and used to mark the position of each well on the membrane.
2. The RNA crosslinked to the membrane using a
3. The membrane was then placed between two sheets of dry 3MM paper and allowed to dry.
4. Finally the membrane was wrapped in Gladwrap and stored at 4°C overnight.

Labelling the probes

RadPrime DNA Labeling System (Invitrogen, Cat. No.: 18428-011)

1. Denature 25ng DNA dissolved in 19µL of sterile distilled water or TE in a microcentrifuge tube by heating for 5min in a boiling water bath; then immediately cool on ice
2. Perform the following additions on ice:

1	µL	500 µM dATP
1	µL	500 µM dGTP
1	µL	500 µM dTTP
20	µL	2.5X Random Primers Solution
5	µL	[α-32P]dCTP (3000Ci/mmol, 10 mCi/mL, approximately 50µCi)
To 49	µL	with dH_2O

3. Pulse spin
4. Mix briefly, add

1	µL	Klenow Fragment

5. Mix gently but thoroughly
6. Pulse spin
7. Incubate at 37°C for 10min
8. Add 5 µL Stop Buffer

Purifying the probe.

ProbeQuant G-50 Micro Columns, Amersham Biosciences, cat#27-5335-01.

1. Loosen the lid of the column and snap off the seal at the bottom.
2. Place the column into a 1.5mL microcentrifuge tube.
3. Spin at 735xg for 1min to remove the liquid.
4. Place the column into a fresh 1.5mL microcentrifuge tube.
5. Layer the labelled probe onto the column.
6. Spin at 735xg for 2min (unincorporated nucleotides, dyes, and salts will remain on the column).

7. The purified labelled probe will collect in the microfuge tube.
8. **Denature the labelled DNA by heating to 95-100°C for 5min and snap-cool on ice.**
9. Pulse spin.
10. Immediately add into Pre/Hybridisation Solution in tubes (do not place directly onto blot).

Probing blots

20x SSC

175.3	g	AR NaCl (=3M)
88.2	g	AR Sodium citrate (=0.3M)

Adjust pH to 7.0

Total volume 1L

Pre/Hybridisation Solution.

(from: *Church GM & Gilbert W, 1984, Genomic Sequencing, Proc. Natl. Acad. Sci. USA*).

12.5	mL	1M $NaHPO_4$ Buffer pH 7.2 (250mM)
20	mL	H_2O
100	µL	0.5M EDTA (1mM)
17.5	mL	20% SDS (7%)
0.5	g	BSA (1%)

Final volume ~50mL.

Pre-wet blots in 2 x SSC before placing in hybridisation oven tubes.

Pre-hybridise in 45mL Pre/Hybridisation Solution at 65°C for 2h.

Add **denatured** probe (see above) to the Pre/Hybridisation Solution and hybridise at 65°C overnight.

Washing blots

2x Wash Buffer.

50 mL 20 x SSC (2 x SSC)
2.5 mL 20% SDS (0.1 % SDS)
Final volume = 500mL.

1x Wash Buffer.

25 mL 20 x SSC (1 x SSC)
2.5 mL 20% SDS (0.1 % SDS)
Final volume = 500mL.

0.2x Wash Buffer.

5 mL 20 x SSC (0.2 x SSC)
2.5 mL 20% SDS (0.1 % SDS)
Final volume = 500mL.

0.1x Wash Buffer.

2.5 mL 20 x SSC (0.1 x SSC)
2.5 mL 20% SDS (0.1 % SDS)
Final volume = 500mL.

Wash filters

2x Wash Buffer 2x 15min at 65°C (~200mL per hybridisation tube).

1x Wash Buffer 2x 15min at 65°C (~200mL per hybridisation tube).

0.2x Wash Buffer2x 15min at 65°C (~200mL per hybridisation tube).

0.1x Wash Buffer2x 15min at 65°C (~200mL per hybridisation tube).

Seal membranes in thin plastic bag, check counts using Geiger counter, and expose to X-ray film with intensifying screen at -70°C, overnight (if ~20 counts per second) or longer depending on signal intensity after washing.

Blots were exposed for 3 days.

EXAMPLE 2: SYNTHETIC SESAME SEED POLYOLEOSIN CONSTRUCTION

Background

We designed a synthetic sesame seed oleosin with tandem repeats for expression in both *E. coli* as well as plants (e.g., Arabidopsis and Lotus). It should be noted that the oleosin sequence used is for example only. Any oleosin sequence or combinations of oleosins, steroleosins and caoleosins and oleosin linking sequences could be used. The original sesame seed oleosin nucleotide sequence and translated peptide sequence are from a sesame seed oleosin, GenBank clone AF091840 (Tai et al., 2002). The codons were optimised for both *E. coli* and Arabidopsis expression. Each repeat used randomised degenerate codons to code for the specific amino acid sequence thus ensuring that the repeats will not be rearranged by non rec$^-$ bacteria such as *Agrobacterium tumefaciens* or *Agrobacterium rhizogenes*. The construct was designed so that it can be relatively simply sublconed from the original backbone (pUC57) into both pET29a and various plant binary vectors. In order to allow simple restriction digestion and re-ligation to reduce the number of repeats as well as to enable us to paste in future peptides between the repeats we engineered restriction sites between them.

The design allows various numbers of tandem repeats to be easily transferred into pET29a and to perform simple digestions on the original clone to remove different numbers of inserts then transfer to binary vectors. This included a NcoI site on each end of the sesame oleosin to place it into pET29a which gives the peptide an N-terminal S•Tag and thrombin cleavage site and a C-terminal His•Tag. For transfer to plant binary vectors we designed an attL1 site to the 5'end and an attL2 site to the 3'end, these are compatible for with our GATEWAY plant binary constructs built in house.

Unique restriction sites (*Eco*47 III, *Dra*I, *Mlu*NI, *Sac*I, *Sal*I, *Eco*RI, *Hind*III, *Sca*I, *Hpa*I, *Alw*44I) were also engineered between the repeats to allow future additions of peptides between the repeats. A single oleosin repeat can be generated by *Xho*I digestion and re-ligation. Similarly, constructs for a dimeric, trimeric and tetrameric oleosin repeats can be generated by digestion and re-ligation with *Cla*I, *Bst*XI and *Nde*I respectively. These can then be transferred to chosen binary vectors via the LR reaction. If the intron is not required then this can be removed after transfer to the binary. *Not*I sites flanked the ORF of the complete clone to allow sub-cloning into pART binary vectors if necessary.

We have mainly concentrated on constructs using cauliflower mosaic virus 35S promoter (CaMV35S), a well-characterized over-expression promoter. It is expected that polyoleosin expression will only be seen in the seeds of transformed plants where triglyceride will be present (de Boer and Somerville, 2001). We have also created a binary vector containing the *Arabidopsis* oleosin promoter; this can also be utilised for discreet polyoleosin expression in the seed..

In addition we created several alternative trimeric oleosin-repeat constructs, In *E. coli*, we have had success expressing a trimer of direct repeats of the sesame oleosin that was successfully expressed by Dr. Bocky Chi-Chung Peng (Peng et al., 2004). We created two binary vectors (CaMV35S and Oleosin promoters) containing this direct repeat trimer. These constructs have been transformed into Arabidopsis.

Codon analysis

We compared the codon usage by both *E. coli* and Arabidopsis (Table 17). From this we were able to identify codons that were not suitable for use in this construct; these included the following:

Arg	AGG	Ile	ATA
Arg	AGA	Leu	CTA
Arg	CGG	stop	TAG
Arg	CGA		
Arg	CGC		

These codons were removed from the codon Table, the remaining codons were placed in a spread sheet which randomised the possibilities still available for each amino acid. Thus, while the codon usage was randomised the peptide sequence for the sesame seed oleosin was conserved. The randomisation was repeated 6 times, one for each oleosin repeat. An alignment of these sequences showed that the homology dropped to approximately 75% between each repeat and the drop in homology was generally distributed evenly across the whole sequence.

Amino Acid	Acid	Codon	*E.coli* frequency	*A.thaliana* frequency
Gly		GGG	0.16	0.15
Gly		GGA	0.12	0.37
Gly		GGT	0.34	0.34
Gly		GGC	0.38	0.14
Glu		GAG	0.32	0.48
Glu		GAA	0.68	0.52
Asp		GAT	0.63	0.68
Asp		GAC	0.37	0.32
Val		GTG	0.36	0.26
Val		GTA	0.16	0.15

Val	GTT	0.26	0.41
Val	GTC	0.22	0.19
Ala	GCG	0.34	0.14
Ala	GCA	0.22	0.27
Ala	GCT	0.17	0.44
Ala	GCC	0.27	0.16
Arg	AGG	0.03	0.20
Arg	AGA	0.05	0.35
Ser	AGT	0.16	0.16
Ser	AGC	0.27	0.13
Lys	AAG	0.24	0.52
Lys	AAA	0.76	0.48
Asn	AAT	0.47	0.52
Asn	AAC	0.53	0.48
Met	ATG	1.00	1.00
Ile	ATA	0.09	0.24
Ile	ATT	0.50	0.41
Ile	ATC	0.41	0.35
Thr	ACG	0.27	0.15
Thr	ACA	0.15	0.30
Thr	ACT	0.17	0.34
Thr	ACC	0.41	0.20
Trp	TGG	1.00	1.00
End	TGA	0.32	0.43
Cys	TGT	0.44	0.60
Cys	TGC	0.56	0.40
End	TAG	0.08	0.20
End	TAA	0.60	0.36
Tyr	TAT	0.58	0.52
Tyr	TAC	0.42	0.48
Leu	TTG	0.13	0.22
Leu	TTA	0.13	0.14
Phe	TTT	0.58	0.51
Phe	TTC	0.42	0.49
Ser	TCG	0.15	0.10
Ser	TCA	0.13	0.20
Ser	TCT	0.15	0.28
Ser	TCC	0.15	0.13
Arg	CGG	0.11	0.09
Arg	CGA	0.07	0.12
Arg	CGT	0.36	0.17
Arg	CGC	0.37	0.07
Gln	CAG	0.66	0.44
Gln	CAA	0.34	0.56
His	CAT	0.58	0.61
His	CAC	0.42	0.39
Leu	CTG	0.48	0.11
Leu	CTA	0.04	0.11
Leu	CTT	0.11	0.26
Leu	CTC	0.10	0.17
Pro	CCG	0.50	0.18
Pro	CCA	0.20	0.33
Pro	CCT	0.17	0.38
Pro	CCC	0.13	0.11

Table 17: Comparison of *E. coli* and *A. thaliana* codon usage.

Each number represents the proportion that codon is used to code for the respective amino acid. Codons in grey were not used in the polyoleosin construct since they coded for the respective amino acid less than 10% of the time in either organism. Codons in bold and underlined were removed to raise the GC content and to remove cryptic splice sites and mRNA degradation signals (ATTTA).

Selection and location of restriction sites between oleosin repeats

Restriction sites were inserted as linkers between the repeats. The sites were chosen to allow the subcloning of different combinations of oleosin repeats; they also allowed for the generation of 8 amino acid linkers between each repeat to allow for free rotation etc. Linkers with undesirable codons were not used.

The multiple cloning sites of both pUC57 and pET29a allowed the design of a sub-cloning strategy using multiple placements of the following restriction sites within the polyoleosin construct.

Bst XI Cla I Pst I Nco I Nde I Not I Xho I

The randomised oleosin repeats were checked for these sites and alternative codons were then used to eliminate the sites when discovered.

Unique restriction sites

We also engineered unique Eco47 III, Dra I, Mlu NI, Sac I, Sal I, Sca I, Hpa I, Alw44I sites between different repeats. These have been included to allow future additions of peptides between the repeats.

Not I sites

Not I sites flank the ORF of the complete clone. This is to allow sub-cloning into pART binary vectors if necessary.

A schematic diagram of the construct is shown in Figure 66.

Optimisation for: Improving Translation Efficiency; Increasing RNA stability; Correct Splicing.

Tetranucleotide stop codon.

Brown et al., (1990) reported that there was an increasing number of reports where the tri-nucleotide stop codons do not signal the termination of protein synthesis; they found that the signals UAA(A/G) and UGA(A/G) are the preferred stop codons in eukaryotes. Hence we have added an A to the 3' end of the second stop codon (TGA) in our construct.

mRNA degradation signal

Beelman and Parker (1995) reported the degradation signal ATTTA (AUUUA) can destabilize transcripts in plants as well as animals. The proposed construct ORF originally had 7 ATTTA sites. These were predominantly caused by the sequence coding for isoleucine followed by a tyrosine residue. The ATTTA sites were removed by changing the relevant isoleucine codons to ATC. Re-analysis of the splice sites after the removal of the ATTTA sites showed that fewer regions were predicted to be introns (partially determined by the GC content).

poly T

The original proposed construct ORF would have had 27 TTTT sites and 12 TTTTT sites.

To reduce the number of these regions the phenylalanine codon TTT was removed and replaced by TTC; in one case the site was eliminated by moving the engineered DraI site to the 5' end of the MluNI site. Combined these changes reduced the number of TTTT sites to 14 and the number of TTTTT sites to 1.

Plant Intron Insertion

The insertion of a recognised plant intron into an expression construct frequently results in a significantly enhanced expression of the construct *in planta*; this is termed Intron Mediated Enhancement (Rose 2004 and references therein). The sequence and position of the intron is important in terms of expression enhancement with the highest gains obtained by placing the *Arabidopsis thaliana* ubiquitin10 (UBQ10) intron within the first 250 bases or so of the 5' end of the transcript (Rose, 2002; 2004 and references therein). Rose and Beliakoff (2000) found that utilising a PstI site was a useful way to insert introns. This was achieved by engineering a PstI site to the 5' end of the intron and by modifying the existing 3' end of the intron to contain a PstI site, from this they were able to add or delete functional introns wherever a PstI site existed in the gene or cDNA.

Vector NTI identified approximately 4 PstI sites within the proposed polyoleosin construct with the closest to the 5' end occurring approximately 500 bases downstream. All these sites were eliminated using various combinations of degenerate codons and a new PstI site was engineered at position 300 using the degenerate codons. This places the intron in the first oleosin repeat and therefore enables the generation of all truncated versions with the intron (see **Sub-Cloning Strategy** below). Using the UBQ10 intron sequence (Norris et al., 1993) we engineered the 3' end to include a Pst1 site; the comparison with the original sequence is shown in Figure 67.

The new polyoleosin construct (containing the intron) was then analysed by NetGene2 (Hebsgaard et al., 1996; Brunak et al., 1991, web site address http://www.cbs.dtu.dk/services/NetGene2/) to confirm that the engineered intron would be predicted to be spliced correctly. This analysis revealed that not only would the UBQ10 be spliced out correctly but there were also high confidence cryptic donor and acceptor sites that would likely result in aberrant splicing. The putative cryptic splice sites were either eliminated wherever possible or reduced in confidence by using alternative redundant codons. The analysis was repeated and showed that the only high confidence donor and acceptor sites remaining were flanking the engineered UBQ10 intron (NetGene 2 results of both the polyoleosin prior to cryptic splice site

removal and polyoleosin with intron and cryptic splice sites removed are shown below); it was predicted that the splicing would remove only the intron and would leave the construct in frame. The analysis showed that a number of regions did not appear to be coding regions and as such may be susceptible to some aberrant splicing. To further reduce the possibility of cryptic splicing we then modified the GC content of the construct (see **GC content** below).

GC content.

Oleosins with optimised and randomised codons; no ATTTA sites or TTT were still found to have relatively low GC content compared to the original sequence. To increase the GC content the additional codons were removed: ATT, AAT, TTA, CTT; this raised the GC content to close to the original content (Table 18).

The repeats were linked using the previously engineered linking regions. These sequences were modified to remove all but the first Pst I sites in the first oleosin repeat and the removal of an Xho site in the second oleosin repeat. In our construct the ORF has no ATTTA or TTTT sites. Furthermore, when the sequence was re-analysed by NetGene2 the only predicted intron splice site in the ORF was the UBQ10 intron engineered into the PstI site and the % identity of the repeats increased from an average of 74.8% identical to 79.1% identical.

NetGene2 was used to predict the splicing of the proposed construct. The results indicated that the RNA should only be spliced at the acceptor and donor sites of the UBQ10 intron (Table 19 and Figure 68).

Conformation of sequence

The coding sequence of the complete ORF (after splicing) was then checked against a heptameric repeat of the original oleosin translated sequence and found to be identical over the oleosin coding regions (Figure 69).

The sequence and feature map of the proposed construct is shown in Figure 70 and the vector map is shown in Figure 71.

Sub-Cloning Strategy.

Plant Intron Removal

The intron is designed to be removed by Pst I digestion, fragment removal and re-ligation. However, due to the presence of a Pst I site in the multiple cloning site of pUC57 it would be necessary to perform a partial digestion to remove this fragment. It would be preferable to attempt this at least once in order to generate full length clones with and without the intron in the initial cloning vector. Alternatively the full length clones can be transferred to expression vectors prior to Pst I digestion as detailed above.

Transfer to *E.coli* expression vector pET29a and generation of different number of repeats.

Transfer the complete clone via partial NcoI digestion then re-ligation into pET29a. The intron can then be removed by Pst I digestion and re-ligation. Full Nco I digestion of the original clone then transfer to pET29a will also yield two trimeric oleosin clones, one will contain the intron. All NcoI digests going into pET29a will have to also be checked for orientation, as this step is not directional. Following intron removal in pET29a a single oleosin repeat can be generated by Xho I digestion and re-ligation. Similarly, constructs for a dimeric, trimeric and tetrameric oleosin repeats can be generated by digestion and re-ligation with Cla I, Bst XI and Nde I respectively.

Transfer to pRS series of GATEWAY adapted binary vectors and generation of different number of repeats.

A single oleosin repeat can be generated by Xho I digestion and religation. Similarly, constructs for a dimeric, trimeric and tetrameric oleosin repeats can be generated by digestion and re-ligation with Cla I, Bst XI and Nde I respectively. These can then be transferred to chosen binary vectors via the LR reaction. If the intron is not required then this can be removed after transfer to the binary vector or by partial digestion in pUC57.

Construct Design	Oleosin Repeat Number						
	Original Oleosin	1	2	3	4	5	6
Oleosins with initial selection of optimised and randomised codons; no ATTTA sites or TTT codons	60.9%	54%	54%	53%	57%	56.8%	53.8%
Oleosins with initial selection of optimised and randomised codons; no ATTTA sites or TTT codons and additionally no ATT, AAT, TTA or CTT codons	60.9%	58.7%	59.1%	59.8%	60.5%	59.5%	57.7%

Table 18. Comparison of GC content of each synthetic oleosin repeat and the original oleosin sequence.

With Intron, optimised codons, enriched GC content, RNA stability modifications

Donor splice sites, direct strand

pos	5'->3'	phase	strand	confidence	5' exon	intron 3'
332	2	+	1.00	GTACCTGCAG^GTAAATTTCT		H

Acceptor splice sites, direct strand

pos	5'->3'	phase	strand	confidence	5' intron	exon 3'
97	1	+	0.90	AAAAAAGCAG^GCTCCGCGGC		
635	1	+	1.00	TGATCTGCAG^TCATCACAAT		H

Branch points, direct strand

acceptor pos	5'->3'	branch pos	5'->3'	strand	score	5'	A 3'
97		52		+	-2.41	CAACAAATTGATAAGCAATG	
635		617		+	-3.71	GATTAATCTGAGTTTTTCTG	

Table 19. NetGene2 tabulated output prediction of splicing in polyoleosin after manual removal of cryptic splice sites plus the addition of the UBQ10 intron and enriched GC

content. The predicted cryptic splice site at position 97 was ignored since this is in the attL1 site and is not part of the transcribed region.

Prokaryotic Expression Constructs.

Three sesame seed oleosin direct repeat constructs were used to generate the prokaryotic expression constructs (Figures 66, 72 and 73). .One clone provided by Dr Bocky Peng (Graduate Institute of Biotechnology, National Chung Hsing University, Taichung, Taiwan 40227) contained a single oleosin clone using the cDNA nucleotide sequence obtained from the sesame seed oleosin clone as reported by Peng et al., (2004). A second clone generated synthetically by GENEART AG contained 3 identical tandem repeats of the cDNA nucleotide sequence obtained from the sesame seed oleosin clone (Figure 73); each repeats was linked with a nucleotide sequence encoding glycine-glycine-glycine-glycine-serine-glycine-glycine-glycine-glycine-serine (GGGGSGGGGS) (Seq ID No. 22). The third clone was the synthetic sesame seed oleosin hexameric repeat generated by GENEART AG (Figure 66). Combinations of the GENEART AG synthesised clones including the prokaryotic specific constructs and the plant optimised construct were used to generate prokaryotic expression vectors in the Novagen pET29 backbone containing between 1 to 6 repeats of the oleosin transcript (Figures 74-83).

The first new synthetic oleosin clone was designed around the sesame seed oleosin minus the six C terminal amino acids but a different C terminal linker to the 6xHis tag. Successful expression of this clone in the pET29a vector would indicate that a specific C terminal linker amino acid sequence is not essential for successful expression. The design of the linker also allows for many different oleosin sequences/repeats to be cloned into the 3' region of the clone, including components of the existing synthetic oleosin sequence with altered codons.

In the synthetic sesame seed oleosin identical triplicate nucleotide repeat clone (Figures 73 and 83), the linker sequence between each unit in the repeat has a repetitive amino acid motif of GGGGSGGGGS, which is designed to allow free movement of each unit in relation to the other units. The C terminal sequence is also be the same as C terminal

sequence from the protein expressed by p29Ole, i.e., the single sesame seed oleosin construct. A summary of the constructs and their design is listed in Table 20.

Construct name	Purpose	Number of sesame seed oleosin repeats	Source of repeats
p29Ole	prokaryotic expression	1	original cDNA (Figure 83)
p29Ole2-6xHis	prokaryotic expression	2	triplet identical repeats from GENEART AG clone (Figure 73)
p29PS3a	prokaryotic expression	3	first 3 randomised codon repeats from GENEART AG (Figure 70)
p29PS3b	prokaryotic expression	3	second 3 random repeats from GENEART AG (Figure 70)
p29Ole3+	prokaryotic expression	3	Triplet identical repeats from GENEART AG (Figure 73)
p29Ole4-6xHis	prokaryotic expression	4	Original cDNA and triplicate identical repeats from GENEART AG (Figure 73)
p29Ole5-6xHis	prokaryotic expression	5	Triplicate identical repeats from GENEART AG (Figure 73)
p29Ole6-6xHis	prokaryotic expression	6	randomised codon repeats from GENEART AG (Figure 70)

Table 20. Constructs in the pET29 (Novagen) backbone for sesame seed polyoleosin expression in *E. coli.*

Generation of polyclonal antibodies to sesame seed oleosin

Culture and induction of expression system

A 50mL LB broth, supplemented with 50µg/L kanamycin (Kan50), was inoculated with a loop of pET29 (containing the nucleotide sequence encoding a single sesame seed oleosin with a C-terminal His tag) culture from a plate and incubated 37°C 200rpm overnight (16h). The following day 6mL aliquots of the overnight culture were used to inoculate 2x 125mL LB Kan50 broths in 1L conical flasks (initial OD600 = 0.16). The cultures were incubated 37°C 200rpm for ~2h (OD_{600} = 0.6) and IPTG was added to a

final concentration of 0.24mg/mL (1mM). The induced cultures were returned to incubate at 37°C 200rpm for a further 4h. The cultures were transferred to 250mL centrifuge pots and spun at 4000×g 4°C for 10min (5000rpm, Sorvall, SS34 rotor).

The supernatant was discarded and the remaining cells were lysed by adding 10mL of B PER® Reagent (Novagen) to the pellet and vortexing/pipetting up and down until the cell suspension was homogeneous. The lysed culture was then incubated with gentle mixing for 20min. Soluble proteins were separated from insoluble proteins by centrifugation at 12,000×g 4°C for 25min (13,400rpm, microfuge). The supernatant was discarded (i.e. retained insoluble proteins in pellet). To increase the purity of the inclusion bodies, 5mL B PER® Reagent + 200µg/mL lysozyme was added to the pellet, which was then vortexed/pipetted up and down to resuspend. The mixture was then incubated at RT for 5min, then spun at 12,000×g 4°C for 25min (13,400rpm, microfuge). The supernatant was discarded (i.e. retained insoluble proteins in pellet, pellet very diffuse) and pellet/dissolved inclusion bodies was resuspended in 4mL Binding Buffer (100mM $NaPO_4$, pH8.0, 500mM NaCl, 6M urea and 10mM imidazole). To remove debris the mixture was spun at 12,000×g 4°C for 15min (13,400rpm, microfuge).

Ni-agarose purification and concentration of His-tagged recombinant protein

2mL aliquots of Invitrogen Ni agarose slurry was placed into two empty columns (A and B) and spun in a hand-operated centrifuge to remove storage buffer. To equilibrate the columns 3x 4mL Binding Buffer was passed through the columns, spinning in the hand-operated centrifuge between each equilibration. The lysate was added to the column and the Ni agarose was gently resuspended into the lysate. The columns were incubated at RT for 10min with gentle end-over-end mixing then spun in a hand-operated centrifuge to remove post bind filtrate. To remove non-bound proteins the columns were washed 3x 4mL Wash Buffer (100mM NaPO4, pH8.0, 500mM NaCl, 6M urea and 50mM imidazole), spinning in a hand-operated centrifuge between each wash. Note that the Wash Buffer previously contained 25mM imidazole the higher concentration (50mM) was found to increase the purity of the His-tagged target protein in the eluted fractions. Finally, fractions were eluted in the following sequential volumes of Elution Buffer (100mM $NaPO_4$, pH8.0, 500mM NaCl, 6M urea and 250mM

imidazole), with spinning in a hand-operated centrifuge between each fraction: 1mL (fraction 1), 2mL (fraction 2), 1mL (fraction 3), 1mL overnight (fraction 4).

200µL of fraction 3A and 3B was removed and put aside, fractions 1A, 2A, and the remainder of fraction 3A were pooled as were fractions 1B, 2B, and the remainder of fraction 3B giving approximately 2.5mL each. Each pool was placed into a spin concentrator with a 10kD molecular weight cut off. The concentrators were spun 3181×g 4°C for 60min. The concentrate was transferred to a fresh tube and the membrane washed with 2x 100uL of the respective retained Fraction 3, the washes were then added to the concentrate.

Whole gel elution/purification of target protein

Whole gel elution was performed using the Bio Rad Mini Whole Gel Eluter as per manufacturers instructions. A denaturing elution buffer was used (250mM Tris, 125mM boric acid, 0.1% SDS, check final pH ≈ 8.7) and the gel was eluted at 85mA (constant) for 25min.

The whole gel elution technique was used to increase the purity of a sample of Ni^{2+} agarose purified recombinant sesame oleosin. Table 21 outlines the loss of recombinant protein that occurred at each stage of the purification process.

	Total Protein	Volume	Remaining
Following...	(µg)	(µL)	(%)
Ni-agarose column	4106	830	100
Concentration	978	60	24
Whole gel elution	272	3000	7
Concentration	83	70	2

Table 21. Protein losses during the whole gel elution procedure.

Following the whole gel elution the gel (Figure 84) and cellophane (not shown) were Coomassie stained to try to identify where the loss had occurred.

An intense band of the same molecular weight as the oleosin could be seen both in the gel and on the cellophane. This indicated that the region of the gel containing the oleosin had not been completely over one of the slots in the whole gel eluter, and only an estimated 40% of the target protein was actually eluted from the gel. The presence of oleosin on the cellophane membrane suggested that the oleosin was either precipitating out at high concentrations or the oleosin was binding to the cellophane, which is possible given the hydropathic clusters within the protein.

Ni agarose chromatography affinity purification

Following the failure of the whole gel elution another culture was prepared and passed through two Ni2+ agarose affinity columns (A and B) and washed at a high stringency (50mM imidazole). Eluted fractions were analysed by Coomassie stained SDS PAGE (Figure 85) and the protein concentration of fractions 1A and 1B was measured using the Bradford's Assay (Table 22).

Sample	Read 1	Read 2	Average	ug	uL in rxn	ug/uL	Final vol. (uL)	Final amt. (ug)
1A	0.584	0.583	0.584	2.65	1	2.65	400	1060
1B	0.608	0.605	0.608	2.86	1	2.86	400	1146

Table 22. Protein concentration of selected fractions from Ni^{2+} agarose purification.

Raising antibodies in rabbits

The first injection was prepared by mixing equal amount of Freunds Complete Adjuvant and the solution containing 265µg of the target protein, to a maximum final volume of 0.5mL. The injection was then administered into multiple sites across the back of the neck and shoulder area of the rabbit. Booster shots containing 77µg of the target protein were delivered 3 4 weeks after the primary. Then 10-14 days later a test bleed of ~3mL was removed for preliminary analysis.

Bleeds were normally stored overnight at 4°C. The following day this was spun at 1500×g 4°C for 5min. Clear serum was removed from the top of the clot and initially stored as 200µL aliquots at 20°C. After thawing, phenol and methiolate was added to the serum (to a final concentration of 0.25% and 0.01% respectively)

Analysis of rabbit anti-sesame seed oleosin antiserum

Four SDS PAGE gels were prepared and loaded identically with samples containing varying amounts of affinity purified oleosin (first fractions from columns A and B). The gels were run out and three were used for immunoblot analysis and one was silver stained.

Serum from the test sample was prepared as described above for the rabbit anti-white clover oleosin antisera and used at 1:200 and 1:1000 dilutions (in TBS-Tween) to screen two of the immunoblots. The remaining immunoblot was screened with a 1:200 dilution (in TBS-Tween) of a chicken anti oleosin antibody raised by Professor Jason Tzen (Graduate Institute of Biotechnology, National Chung Hsing University, Taichung, Taiwan 40227). Results from the immunoblot analysis and silver staining are shown in Figure 86.

Expression and analysis of 1× - 6× polyoleosin in *E. coli*

Preparation of *E. coli* expression lines

100µL competent *E. coli* Rosetta cells were transformed with 2.5µL plasmid DNA. Samples were then incubated on ice for 20min; heat shocked 42°C 1min; cooled

rapidly on ice 2min; and incubated with the addition of 900µL LB broth at 37°C 60min 1400rpm (Thermomixer). Pelleted cells were spread on LB Kan50 plates and incubated at 37°C overnight.

.Preparation of artificial oil bodies (AOBs)

The primary methods for investigating the properties of the putative triple sesame polyoleosin protein was to compare with the single sesame recombinant protein in terms of:

- AOB size
- AOB stability over time and at various temperatures and pH values
- Stability of AOBs in the rumen

Day 1

10mL LB Kan50 broth toothpick inoculated with original colony from p29Ole3+(8b) and p29Ole, then incubated at 37°C overnight.

Day 2

5mL overnight culture inoculated into a 100mL LB Kan50 broth (500mL flask). Incubated 37°C ~150min. Induced with 400µL 250mM IPTG, incubated 37°C 3h.

Each culture was transferred to 2x 50mL Falcon tubes and cells pelleted by centrifugation 10min 4°C 3200×g. Pellet resuspended in 4mL B Per Reagent (Pierce) and incubated to lyse (RT 20min, gentle end-over-end mixing).

Insoluble protein pelleted by centrifugation 20min 4°C 3200×g. Supernatant discarded and pellets resuspended in 2mL Oilbody Buffer (total vol.; 50mM $NaPO_4$ pH8.0, 50mM NaCl) and stored at -20°C.

Day 3

Samples removed from -20°C and thawed. 2x 500µL each prep mixed with 4.5mL Oilbody Buffer and either 200µL(A) or 500µL(B) of purified sesame oil (remaining 1mL of each prep returned to -20°C).

Samples were then sonicated (Sonics & Materials Vibra~Cell VC600, 600W, 20kHz; 1/8" tapered micro-tip probe) on ice at Power 4, 80% pulse, 1x 90sec (probe heated up at this setting); followed by Power 4, 50% pulse, 2x 180sec. After incubation on ice for 90min the samples were again sonicated on ice at Power 4, 50% pulse, 2x 180sec.

To concentrate the AOBs the samples were spun 10min 4°C 3200×g and the lower phase decanted off from under upper oilbody phase. Oilbodies were then mixed in 5mL Oilbody Buffer and completely resuspended by sonication at Power 4, 50% pulse, 1x 180sec. The AOBs were then stored at 4°C until required for subsequent analysis.

Alternatively, pellets from 20mL induced cultures of 1× 6× polyoleosin lines were resuspended in 1mL Oilbody Buffer and sonicated (Sonics & Materials Vibra~Cell VC600, 600W, 20kHz; 1/8" tapered micro tip probe) off/on ice at:

Pulse	Time	Power
100	20	3
100	20	3
100	20	3
100	20	3

The mixture was then spun in a microfuge 4°C 10min 14,000×g (14,500rpm), the supernatant discarded and 1mL Oilbody Buffer added to the pellets. 10µL Purified Sesame Oil added to 240606 samples and 50µL Purified Sesame Oil added to 210606 samples. Sonicated:

Pulse	Time	Power
100	20	3
100	20	3
100	20	3
100	10	5

Spun in a microfuge 4°C 15min 14,000×g (14,500rpm). None of the samples had formed AOBs. Supernatant discarded and 1mL Oilbody Buffer added to pellets. 50µL Purified Sesame Oil added to all 240606 and 210606 samples, including those that had already formed AOBs (white layer). Sonicated:

Pulse	Time	Power
100	15	5
100	15	5
100	15	5
100	15	5

Spun in a microfuge 4°C 15min 14,000×g (14,500rpm). The supernatant was transferred to a fresh 1.5mL tube, and the supernatant and AOBs were stored at 4°C overnight.

It can be seen by the formation of two layers that (AOB layer = upper), all samples form AOBs including the negative controls (Figure 87), which consist of induced *E. coli* Rosetta strain containing the pET22 vector. However, the AOBs formed using non-oleosin *E. coli* proteins are unstable and rapidly breakdown and coalesce (Figure 92).

Alternative methods to prepare artificial oil bodies can also be used including varying the oil/buffer/protein emulsion via different proportions or different oils and buffers. It is also possible to vary the ultrasonic energy required or to dispense with the requirement for ultrasonication via use of other disruptive methods such as vortexing and the use of organic solvents to purify the polyoleosin prior to use.

Analysis of the purity of polyoleosins from AOBs

750µL Oilbody Buffer added to AOBs, then sonicated:

Pulse	Time	Power
100	15	5

30µL of each sample was mixed with an equal volume of SDS GLB [@ 2× SDS], and the remaining AOBs were stored at 4°C. Protein denatured in boiling water bath for 5min. Samples loaded onto 12% SDS PAGE [@ 2× SDS] gel at 15µL per lane for the supernatant (S) and 2.5µL per lane for the AOBs (A). Run at 150V 75min. After electrophoresis gels were transferred to PVDF membrane for subsequent immunoblot analysis with rabbit anti-oleosin antibodies (Figure 88). Alternatively samples were loaded onto a 4-15% gradient SDS-PAGE (with no stacking gel) and after electrophoresis gels were stained with SafeStain (Invitrogen) (Figure 89). Earlier attempts to analyse artificial oil bodies from bacteria by SDS-PAGE showed that varying portions of the 3×, 4×, 5× and 6× samples do not run into the gel, and stop at the

stacking/separating gel interface. As it was possible that the polyoleosin proteins with higher numbers of repeats were forming high molecular weight aggregates. This problem is partially alleviated by using SDS/urea denaturing PAGE (Table 23) and visualised using SafeStain (Figure 90).

SDS/urea Separating Gel				SDS/urea Stacking Gel		
2.5	mL	4× Tris-SDS Separating Buffer		5	mL	2× Tris-SDS Stacking Buffer
3	mL	40% Acrylamide		1.25	mL	40% Acrylamide
4.8	g	Urea (Mr=60, 8M final conc.)		4.8	g	Urea (Mr=60, 8M final conc.)
Made to 10mL with H_2O				Made to 10mL with H_2O		
pH adjusted to 8.8 with conc. HCl				pH adjusted to 6.8 with 1M HCl		
10	µL	TEMED		10	µL	TEMED
100	µL	10% APS		Washed top of polymerised separating gel with 200µL aliquots.		
				100	µL	10% APS

Table 23. Recipe for SDS/urea denaturing PAGE.
* Rinsed wells immediately before loading samples *

AOB properties

5µL aliquots of AOBs were placed on a microscope slide and observed at 1000× magnification (Figure 91). When observed under the microscope the 1× oleosin AOBs appeared to be the smallest, and as the number of oleosin repeats increased the average size of the AOBs increased; up to the 3× polyoleosin, when AOB size appeared to remain similar for the 3×, 4×, 5× and 6× polyoleosin.

It is likely that significantly more single oleosin was used to make AOB than any of the multimeric tandem repeat proteins. This would explain the comparatively small size of the 1x oleosin AOB size.

AOB stability over time

A 50µL aliquot of AOBs was transferred to a 250µL PCR tube and incubated at 4°C for 168h. 5µL aliquots of the AOBs were then placed on a microscope slide and observed at 1000× (Figure 92). AOB generated from vector control protein extracts showed

almost complete coalescence of the oil. After 168h the size of the AOB in the samples containing recombinant oleosins was inversley related to the number of oleosin repeats. In other words the AOB containing single chain oleosin units were on average larger than those containing 3 linked oleosin units that in turn were on average larger than those containing 6 linked oleosin units (Figure 92). Thus the long-term stability of the emulsification can be tailored by altering the number of oleosin repeats used to generate the AOB.

AOB heat stability

A 50µL aliquot of AOBs was transferred to a 250µL PCR tube and incubated either at 70°C for 4h, or at 90°C for 15min in a PCR machine. To clearly define the amount of intact AOBs remaining, the tubes were spun 10min 4°C 3200×g (4000rpm, Eppendorf 5810R centrifuge, A 4 62 swing out rotor). 5µL aliquots of the AOBs were then placed on a microscope slide and observed at 1000× (Figure 93).

Although there did not appear to be much difference in the size of the different AOBs after heat treatment, a large proportion of the AOBs formed with the 1× and 2× polyoleosins had coalesced (Figure 93) to form large pools of oil. Those AOBs that were observed from samples formed with the 3× polyoleosin were of relatively the same size as before the heat treatment, but had formed a very thick emulsion. Some coalescence was observed in the 4× and 5× AOBs, but no large pools of oil were observed (Figure 93). With the 6× polyoleosin there was no change in the size of the AOBs, nor was there any evidence of coalescence (Figure 93). Thus the heat stability of the emulsification can be tailored by altering the number of oleosin repeats used to generate the AOB.

In addition, thickness of the emulsion layer remaining after heat treatment was correlated with an increase in oleosin repeat number (Figure 94). Thus the stability of the emulsification at elevated temperatures can be tailored by altering the number of oleosin repeats used to generate the AOB.

Stability of AOB with different polyoleosins at pH 3.5, 8.0 and 10.5

A large number of proteins have been found experimentally to have different optimum pH of maximal stability where pH influences the folding and the net charge of the proteins. We tested for stability of AOBs generated with different polyoleosins at pH 3.5, 8.0 and10.5.

Buffers: 50 mM Sodium Phosphate pH 8, 100 mM NaCl

50 mM Glycine-HCl pH 3.5

50 mm Glycine-NaOH pH 10.5

AOBs were generated using 1, 2, 3 or 6 polyoleosins; the negative control consisted of AOBs generated using inclusion bodies generated from E. coli containing an empty expression vector. Each preparation was sonicated to evenly suspend them in buffer containing 50 mM sodium phosphate pH 8, 100 mM NaCl. A 25 µl suspension of each polyoleosin-oil body was aliquot into microtubes containing 75 µL of the buffers at different pH. The tubes were incubated at room temperature (~ 22 °C) with every 5 minutes interval of 15 seconds vigorously shaking (~1,400 rpm) using a Thermomixer. After approximately 4 hours, samples were taken (~4 µL) and dropped on glass slice for microscoping at 1000x magnification.

After 4 hours at room temperature the negative control AOBs were beginning to coalesce at both pH3.5 and 10.5; while no coalescence of the negative control was seen at pH8 the AOBs were no longer spherical (Figure 95). The AOBs containing 1 oleosin repeat appeared to be unstable at both pH3.5 and 10.5 but were stable at pH8.0. In comparison, AOBs containing 2 or more oleosin repeats were relatively stable at both pH 8.0 and 3.5. At pH 10.5 the AOBs containing 2 or more oleosin repeats were still visible but appeared to be smaller than those in lower pH buffers

(Figure 95). Some precipitation/aggregation of the emulsification was noted in the preparations containing 2 or more oleosin repeats at pH 10.5.

Comparison of purification of oleosin by AOB generation versus affinity purification

Prokaryotically produced polyoleosins were purified by either AOB generation or Ni^{2+} affinity column. These samples were analysed by both SDS-PAGE/immunoblot or Coomassie stain or SafeStain (Figure 96). This demonstrates that both methods can be used for polyoleosin purification.

Transformation of *Arabidopsis thaliana* with polyoleosin constructs under the control of the CaMV35s construct or the Arabidopsis oleosin seed promoter.

A range of plant binary vectors containing from 1 to 6 oleosin repeats were created using the synthetic sequences generated by GENEART AG (Figures 66, 70, 71 & 72) and two plant binary vectors containing either the CaMV35s promoter or the Arabidopsis oleosin seed promoter (Table 24 and Figures 97-116).

Construct name	Plant promoter	Terminator	Oleosin cassette	UBQ10 intron in first repeat	Selectable marker
pRSh1-PSP1	CaMV35s	ocs3'	1 oleosin, randomised degenerarte codons	yes	Spectinomycin & Basta
pRSh1-PSP3	CaMV35s	ocs3'	3 oleosins, randomised degenerate codons	yes	Spectinomycin & Basta
pRSh1-PSP4	CaMV35s	ocs3'	4 oleosins, randomised degenerate codons	yes	Spectinomycin & Basta
pRSh1-PSP6	CaMV35s	ocs3'	6 oleosins, randomised degenerate codons	yes	Spectinomycin & Basta
pRSh3-PSP1	Oleosin	ocs3'	1 oleosin,	yes	Spectinomycin & Basta

	seed		randomised degenearte codons		
pRSh3-PSP3	Oleosin seed	ocs3'	3 oleosins, randomised degenerate codons	yes	Spectinomycin & Basta
pRSh3-PSP4	Oleosin seed	ocs3'	4 oleosins, randomised degenerate codons	yes	Spectinomycin & Basta
pRSh3-PSP6	Oleosin seed	ocs3'	6 oleosins, randomised degenerate codons	yes	Spectinomycin & Basta
pRSh1-Ole3+	CaMV35s	ocs3'	3 oleosins, identical nucleotides codons	no	Spectinomycin & Basta

Table 24. Summary of constructs generated for expression of polyoleosin in plants.

Analysis of Agrobacterium Plasmid preps

Agrobacterium tumefaciens (strain GV3101) was transformed using the freeze thaw method. 5 µg of plasmid was added to 250 µL of competent GV3101 cells and incubated on ice for 30 min. The cells were then frozen in liquid nitrogen for 1 min and thawed in a 37°C water bath for 1 min. This process was repeated once then 1 mL LB broth was added to the cell mix. Following incubation at 28°C for 4–6 h, the cells were pelleted, resuspended in 100 µL LB and plated (20 µL and 50 µL) onto LB Spec.

Transformed colonies were visualised on LB spec plates after approximately 48 h and single colonies were re streaked on LB spec to ensure the use of single colonies. From these new plates, single colonies were selected and plasmid preps obtained by using 8–10 mL of overnight culture in a modified QIAGEN Mini Prep protocol. Quantification of the preps was by NanoDrop spectrophotometer, and as yields were typically low, the presence of the construct in Agrobacterium was detected by PCR.

Primers

pRSh1-PSP4:

35S(3'end)Fwd 5' GAC ACG CTC GAG GAA TTC GGT ACC (Seq ID No. 23)

UBQ10IntRev 5' GAT GGT GAT GAC TGC AGA TCA GAA (Seq ID No. 24)

Product size = 609 bp

pRSh1-PSP6:

UBQ10IntFwd 5' CGA TTA ATC TGA GTT TTT CTG ATC TGC AGT CA (Seq ID No. 25)

PolySes3R 5' CGA TCA CCG TTC CGG CCA ATG TC (Seq ID No. 26)

Product size = 889 bp

The pRSh1-PSP4 primers could be used on any of the 35S promoter polyoleosin constructs.

Cycle

94°C / 2 min

(94°C / 30 sec; 63°C / 30 sec; 72°C / 1 min) x 30

72°C / 7 min

Primers

pRSh3-PSP4 and pRSh3-PSP6:

OlePro 5' GAC ACG TGA CTT CTC GTC TCC TT (Seq ID No. 27)

UBQ10IntRev 5' GAT GGT GAT GAC TGC AGA TCA GAA (Seq ID No. 28)

Product size = 722 bp

In theory, these primers will detect any of the plant polyoleosin constructs containing the oleosin promoter. The constructs were thoroughly checked by restriction digestion and sequencing when purified from *E. coli*. The positive controls were the respective pRSh3-PSP4 and pRSh3-PSP6 plasmid preps from *E. coli*.

Cycle

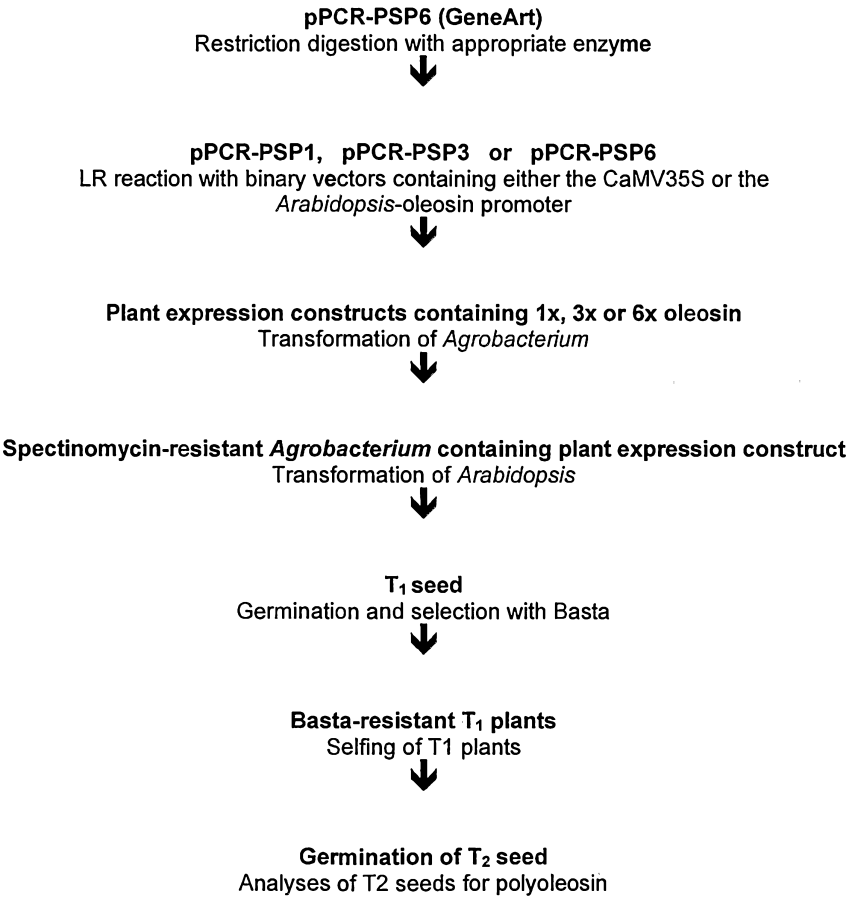
94°C / 2 min

(94°C / 30 sec; 63°C / 30 sec; 72°C / 45 sec) x 30

72°C / 7 min

Transformation of *Arabidopsis thaliana* var Columbia with Polyoleosin binary constructs

Flow Chart of Project - Polyoleosin Expression in Plants



Columbia plants were transformed using the floral dip method. The efficiency of this process can vary depending on a number of variables, including the construct itself, the health and floral development stage of the plant and the strain of Agrobacterium. For some constructs, two variations on the floral dip method were employed in order to try and optimise infiltration effectiveness.

One method involved dipping the entire plant or pot of plants (Full-dip) into Agrobacterium culture suspended in a solution of 5 % (w/v) sucrose and 10 mM magnesium chloride, pH 5.7. Prior to dipping, Silwet L77 (a silicone polyether

copolymer) was added to the culture solution to aid infiltration. Plants could be dipped up to three times, at a frequency of no more than once per week. We now have some idea of the efficiency of our full dip method using the polyoleosin constructs, which looks to be about 0.05 0.1%. This indicates that 1g of T1 seed (approx 50,000 seed) should yield 25 herbicide resistant plants. One batch of 200 dipped plants should yield between 2g and 5g of seed.

The alternative method involved dropping Agrobacterium culture, suspended in the floral dip solution described above, onto individual flowers using a sterile transfer pipette (Floral drop). The rationale here was to avoid the entire plant being covered in Agrobacterium, thus aiding plant recovery. Florets were infiltrated as often as every second day and up to four times in total.

The seed collected from these transformation events represent T1 seed, which were germinated and sprayed with Basta herbicide in order to select transformed T1 plants. The total number of plants subjected to either of these methods with the different constructs is summarised in Table 25.

Batch Number	Construct	Promoter	Terminator	Resistance	Approx number of plants treated	Method of transformation
GH 22083	pRSh1-PSP1	CaMV 35S	ocs3'	Spec / Basta	200	Full-dip
GH 22242 GH 22243	pRSh1-PSP1	CaMV 35S	ocs3'	Spec / Basta	60	Floral drop
GH 22082	pRSh1-PSP3	CaMV 35S	ocs3'	Spec / Basta	200	Full-dip
GH 22244 GH 22188	pRSh1-PSP3	CaMV 35S	ocs3'	Spec / Basta	60	Floral drop
GH 22446	pRSh1-PSP4	CaMV 35S	ocs3'	Spec / Basta	200	Full dip
GH 22017 GH 22150	pRSh1-PSP6	CaMV 35S	ocs3'	Spec / Basta	300	Full dip
GH 22171-2 GH 22209-10	pRSh1-PSP6	CaMV 35S	ocs3'	Spec / Basta	60	Floral drop
GH 22447	pRSh3-PSP4	Arabidopsis oleosin	ocs3'	Spec / Basta	200	Full dip
GH 22286 GH 22430	pRSh3-PSP6	Arabidopsis oleosin	ocs3'	Spec / Basta	300	Full dip
GH 22196	pRSh3-PSP6	Arabidopsis oleosin	ocs3'	Spec / Basta	60	Floral drop
GH 22646	pRSh1-Ole3+	CaMV 35S	ocs3'	Spec / Basta	200	Full dip
GH 22485	pRSh3-Ole3+	Arabidopsis oleosin	ocs3'	Spec / Basta	200	Full dip

Table 25. Summary of Arabidopsis transformations completed with the polyoleosin monomeric, trimeric, tetrameric and hexameric polyoleosin constructs.

PCR testing of herbicide resistant T1 plants

T1 seed from batch GH 22017 (Table 25) was collected and 1g of seed germinated. Spraying with Basta herbicide resulted in approximately 20 herbicide resistant plants. These plants were also tested by PCR on genomic DNA. Genomic DNA was extracted from the rosette leaves of selected plants once they were of a suitable size. Extraction was carried out using the QIAGEN DNeasy mini kit protocol and 100 ng genomic DNA was used as PCR template. Two primer pairs were used to detect different areas of the polyoleosin insertion, thus giving information as to whether any gross rearrangements occurred during integration of the construct into the plant genome.

Plants transformed with tandem repeats of identical nucleotide Sesame seed oleosin transcripts.

The concern with using a construct containing repeating units of exactly the same sequence, as is the case with the Ole3+ constructs, is that non *rec*⁻ bacteria such as *Agrobacterium* may utilise their natural recombination mechanism to rearrange the sequence each time a new generation of bacteria is grown. Most laboratory strains of *E. coli* have the *rec* mutation to prevent this phenomenon. The PCR results go some way towards investigating this possibility, by using 2 sets of primers at each end of the repeating unit structure in the constructs. This suggests the constructs are intact after one round of sub-culturing. An additional check is to digest the plasmid preps from *Agrobacterium* with restriction enzymes and check the banding pattern obtained. The plasmid prep yield from *Agrobacterium* is typically too low to be able to perform digests, so the prep of clone #1 from each construct was used to transform *E. coli* TOP10 cells by heatshock. Plasmid preps were prepared by overnight culture of three single colonies and extraction using the Qiagen mini-prep kit. The preps were analysed by three restriction enzymes the correct banding pattern is observed for all the clones tested. This result indicates that the Ole3+ constructs are intact in the *Agrobacterium* with no rearrangement. In the white clover constructs (discussed above), aberrent RNA species were only found to be expressed in *Lotus japonicus* roots when the number of identical white clover oleosin repeats was greater than 3. This could be due to rearrangements occurring in the Agrobacterium prior to plant transformation or during the process of stable integration into the plant geneome.

An additional check of integrity is to sequence through the repeating units of the *E. coli* preps from TOP10 cells transformed with plasmid obtained from *Agrobacterium*. We have sequenced through 90% of the repeating units and for the sequence completed, the construct sequence matches with the expected database sequence.

The possibility of rearrangement is potentially greater as the *Agrobacterium* integrates its T-DNA into the plant genome. We have no control over this event, so can only check for rearrangement events in the genomic DNA of herbicide resistant plants once they have been identified.

Analysis of *Arabidopis thaliana* seeds expressing sesame seed oleosin.

Protein expression

The herbicide resistant Arabidopsis seeds (T1) from the floral dip and full dip plants (T0) were allowed to self and set seed (T2). Seeds were collected from plants as the sliques matured (i.e., turned brown and became dry). Seeds were separated from the sliques and all seeds from 1 plant were pooled. Given the method of transformation the T2 seeds consist of segregating populations. It could be expected that for a single insertion event the T2 seed will demonstrate a 3:1 segregation pattern for a dominant trait (as would be expected for protein expression under the CaMV35s promoter). This would consist of 25% homozygous, 50% hemizygous and 25% wild type (untransformed).

Immunoblot analysis of *Arabidopsis thaliana* oil bodies containing sesame seed oleosin.

T2 seeds from approximately 10 individual transformation events for each construct were analysed for protein expression. Seed lots for analysis were chosen based on the total amount of seed collected per plant.

Weigh out 25mg of seed collected from 1 plant; combine with approximately 100mg of dried, clean sand and 500µL buffer A (600mM Sucrose in 10mM Sodium Phosphate buffer, pH 7.5).

Grind to a paste using mortar and pestel.

Recover maximum volume using pipette.

Rinse mortar and pestel 2x with 500µL buffer A.

Combine extract and rinse in eppendorf tube, spin at 13k rpm for 5 minutes.

Quickly recover the majority of the aqueous layer (bottom layer) using a piptette and transfer to a fresh tube.

Recover overlying oil layer by piptette after tilting the tube horizontally; transfer to a fresh tube.

To both fractions (aqueous and oil layer) add buffer A to a final volume of 500µL; mix by pipetting.

Mix with equal volumes of loading buffer, vortex, boil for 5 minutes and analyse by SDS-PAGE /immunoblot using rabbit anti-sesame seed oleosin antiserum as the primary antibody (Figure 117).

The rabbit anti-sesame seed oleosin antiserum showed no binding with extracts from wild type plants. Immnoreactive protein of the expected size was detected in the plants transformed with the polyoleosin constructs including the monomeric, trimeric and tetrameric constructs (dimeric and tetrameric transformants were not included) (Figure 117). The SDS-PAGE/immunoblot analysis demonstrates that the constructs can be expressed and translated and that the protein is of the expected molecular weight, it accumulates in the seed and is targeted to the oil bodies in the correct manner. The T2 seed populations screened here contained segregating populations, hence it could be reasonably anticipated that even higher levels of expression would exist in homozygous plants as well perhaps in transgenic plants yet to be analysed. The presences of the immunoreactive band at the molecular weight of the monomeric sesame seed oleosin in all extracts from plants transformed with the polyoleosin trimer and polyoleosin hexamer suggests that either some protein cleavage is occurring or that some of the transcript is only translated as far as the first oleosin repeat before being terminated. The ratio of monomer to either trimer or hexamer is relatively low indicating that the level of accumulation of the monomer is also low.

SDS-PAGE/Coomassie analysis of the crude protein extracts from seeds of the same samples shows that the level of recombinant protein in the hemizygous transgenic

plants is too low to be detected by Coomassie stain in the background of native proteins (Figure 118).

Properties of sesame seed polyoleosin expressed in *Arabidopsis thaliana* seeds.

Heat stability

Oil bodies from two wild type, pRSh1-PSP1, pRSh1-PSP3 and pRSh1-PSP6 Arabidopsis transformants were compared for heat stability. 25mg of seed from each plant was ground separately in a mortar and pestle containing xxμL oil body extraction buffer (10mM Sodium Phosphate buffer containing 600mM sucrose)). This was transferred to a microfuge tube, the mortar and pestle was rinsed a further two times with 500μL of oil body extraction buffer, combined with the original extract and the volumes were made up to 500μL with oil body extraction buffer. The tubes were spun at 13k rpm for 10 minutes and phographed (Figure 119). Each sample showed a thin layer (containing the oil bodies) floating on the surface of extraction buffer. 400μL of the oil body extraction buffer was removed using a protein gel loading tip attached to an eppendorf. Small aliquots of oil body extraction buffer were used to resuspend the oil layer and removed to a new 2ml eppendorf tube. The oil body layers were made up to the same final volume (approximately 500μL with oil body extraction buffer and resuspended fully by pipetting. 50μL of the suspension was loaded into PCR tubes and capped; 50μL was stored at 4°C. The PCR tubes were heated at 90°C for 2, 4, 24 and 58hrs. At each time point the tubes were removed from the heating block, spun for 10 minutes at 13k rpm to separate the suspensions and photographed (Figure 120). The reduction in oleosin repeat numbers correlates with a decrease in the emulsification layer remaining after 24hr at 90°C, where the thickest emulsification can be seen from plants expressing the hexameric polyoleosin construct. The reduction in oleosin repeat numbers correlates with a decrease in the emulsification layer remaining after 24hr at 90°C, where the thickest emulsification can be seen from plants expressing the hexameric polyoleosin construct. At 0hrs through to 24hrs at 90°C the higher the number of oleosin repeats correlates with a thicker emulsification layer. After 58hrs at 90°C the oil bodies from the wild type arabidopsis has been reduced to a very thin layer with a small ring of emulsion deposited on the tube above the remaining aqueous

phase. In comparison, the ring of emulsion remaining in the transformant samples is much greater (Figure 120). Thus the stability of the emulsification derived from seed extracts at elevated temperatures can be tailored by altering the number of oleosin repeats used to transform the plant.

5

References:

Beelman CA and Parker R. (1995). Degradation of mRNA ineukaryotes. Cell, 81(2) 179-183.

Bouvier-Nave P, Benveniste P, Oelkers P, Sturley SL, Schaller H. (2000) Expression in yeast and tobacco of plant cDNAs encoding acyl CoA:diacylglycerol acyltransferase. *Eur. J. Biochem.* 267, 85-96.

Brown CM, Stockwell PA, Trotman CN, Tate WP. (1990). Sequence analysis suggests that tetra-nucleotides signal the termination of protein synthesis in eukaryotes. Nucleic Acids Research. 18(21): 6339-6349.

Brunak, S., Engelbrecht, J., and Knudsen, S. (1991). Prediction of Human mRNA Donor and Acceptor Sites from the DNA Sequence, Journal of Molecular Biology, 220, 49-65)

Christey, M.C. and Braun R.H. 2004. Production of Hairy Root Cultures and Transgenic Plants by *Agrobacterium rhizogenes*-mediated Transformation p.47-60. In: L. Pena (ed.), Methods in Molecular Biology, Vol. 286: Transgenic Plants. Humana Press Inc., Totowa, NJ.

Christey, M.C., Sinclair, B.K., Braun, R.H. and Wyke, L. 1997. Regeneration of transgenic vegetable brassicas (*Brassica oleracea* and *B. campestris*) via Ri-mediated transformation. Plant Cell Reports 16: 587-593.

Dahlqvist A, Stahl U, Lenman M, Banas A, Lee M, Sandager L, Ronne H, Stymne S. (2000) Phospholipid:diacylglycerol acyltransferase: an enzyme that catalyzes the acyl-CoA-independent formation of triacylglycerol in yeast and plants. Proc Natl Acad Sci U S A. 97, 6487-6492.

de Boer G-J and Somerville C. (2001). The many changing colours of fatty acid biosynthesis. Proc. Nat. Plant Lipid Cooperative Supplement. Stanford Sierra Summer Camp, Fallen Leaf Lake, CA)

Demeyer, D., Doreau, M. 1999. Targets and procedures for altering ruminant meat and milk lipids. Proceedings of the Nutrition Society 58:593-607.

Hebsgaard, SM. Korning, P.G. Tolstrup, N. Engelbrecht, J. Rouze, P. Brunak. S. (1996): Splice site prediction in Arabidopsis thaliana DNA by combining local and global sequence information, Nucleic Acids Research,24(17)3439-3452.

Lardizabal KD, Mai JT, Wagner NW, Wyrick A, Voelker T, Hawkins DJ. (2001) DGAT2 Is a new diacylglycerol acyltransferase gene family. J.B.C. 276, 38862-38869.

Linsmaier, E.M. and Skoog, F. 1965. Organic growth factor requirements of tobacco tissue cultures. Physiol. Plant. 18:100-127.

Murashige, T. and Skoog, F. 1962. A revised medium for rapid growth and bioassay with tobacco tissue cultures. Physiol. Plant. 15: 473-497.

Norris SR, Meyer SE and Callis J (1993). The intron of Arabidopsis thaliana polyubiquitin genes is conserved in location and is a quantitative determinant of chimeric gene expression. Plant Molecular Biology, 21, 895-906

Peng C-C, Chen JCF, Shyu DJH, Chen M-J, Tzen JTC. (2004) A system for purification of recombinant proteins in Escherichia coli via artificial oil bodies constituted with their oleosin-fused polypeptides. *J. Biotech.* 111, 51-57

Rose AB (2002). Requirements for intron-mediated enhancement of gene expression in Arabidopsis. RNA, 8, 1444-1453

Rose AB (2004). The effect of intron location on intron-mediated enhancement of gene expression in Arabidopsis. The Plant Journal, 40, 744-751

Rose AB and Beliakoff JA (2000). Intron-mediated enhancement of gene expression independent of unique intron sequences and splicing. Plant Physiology, 122, 535-542

Sparrow, P.A.C., Dale, P.J. and Irwin, J.A. 2004. The use of phenotypic markers to identify *Brassica oleracea* genotypes for routine high throughput *Agrobacterium*-mediated transformation. Plant Cell Reports 23: 64-70.

Tai, SSK, Chen, MCM, Peng, CC, Tzen JTC. (2002). Gene family of oleosin isoforms and their structural stabilization in sesame seed oil bodies. Biosci. Biotechnol. Biochem. 66(10): 2146-2153.

Tzen JT, Lie GC, Huang AH. (1992). Characterization of the charged components and their topology on the surface of plant seed oil bodies. J Biol Chem. 267(22):15626-34.

Williams NK, Liepinsh E, Watt SJ, Prosselkov P, Matthews JM, Attard P, BeckJL, Dixon NE, Otting G. (2005). Stabilization of native protein fold by intein-mediated covalent cyclization. J. Mol. Biol., 346(4):1095-1108.

Zou J, Wei Y, Jako C, Kumar A, Selvaraj G, Taylor DC. (1999) The Arabidopsis thaliana TAG1 mutant has a mutation in a diacylglycerol acyltransferase gene. *Plant J.* 19, 645-653.

Reference to any prior art in the specification is not, and should not be taken as, an acknowledgement or any form of suggestion that this prior art forms part of the common general knowledge in Australia or any other jurisdiction.

Finally, it is to be understood that various alterations, modifications and/or additions may be made without departing from the spirit of the present invention as outlined herein.

07 Nov 2012

2006303870

CLAIMS

1. A construct including one or more nucleic acids encoding three or more oleosin repeat units.

2. A construct according to claim 1, wherein the oleosin repeat units are tandem repeats.

3. A construct according to claim 1, wherein the construct further includes one or more nucleotide sequences encoding peptide linking sequences between two or more of the oleosin repeat units.

4. A construct according to any one of claims 1 to 3, wherein the construct includes nucleic acids encoding between three and ten oleosin repeat units.

5. A construct according to claim 3 wherein said peptide linking sequences enable degradation, allow flexibility and/or induce a directional change between the oleosin repeat units.

6. A construct according to claim 3 wherein the nucleic acids encoding peptide linking sequences include sites for enzymatic cleavage, and/or subsequent fusion.

7. A construct according to any one of claims 1 to 6 further including one or more nucleic acids encoding bioactive peptides.

8. A construct according to any one of claims 1 to 7, wherein the nucleic acids encoding oleosin repeat units are from white clover or sesame seed or are a recombinant or synthetic version thereof.

9. A construct according to claim 8, wherein said nucleic acids are modified to enhance expression of said construct.

2006303870 07 Nov 2012

10. A construct according to any one of claims 1 to 9 further including a nucleic acid encoding a diacylglycerol acyltransferase or a functionally active fragment or variant thereof.

11. A construct according to claim 1, wherein the nucleic acids encoding oleosin repeat units include a nucleotide sequence selected from the group consisting of Seq ID Nos. 29-43, 50, 52-57 and 64-73, and functionally active fragments and variants thereof.

12. A plant cell, plant, plant seed or other plant part including a construct according to any one of claims 1 to 11.

13. An eukaryotic or prokaryotic cell including a construct according to any one of claims 1 to 11.

14. A method of producing repeat oleosins in:

(i) a plant, or

(ii) a eukaryotic cell, or

(iii) a prokaryotic cell,

said method including introducing into said plant, eukaryotic cell, or prokaryotic cell a construct according to any one of claims 1 to 11.

15. A partially or substantially purified and/or recombinant polypeptide including three or more oleosin repeat units.

16. A polypeptide according to claim 15, wherein the oleosin repeat units are tandem repeats.

17. A polypeptide according to claim 16, wherein the polypeptide further includes peptide linking sequences between two or more of the oleosin repeat units.

2006303870 07 Nov 2012

18. A polypeptide according to any one of claims 15 to 17, wherein the polypeptide includes between three and ten oleosin repeat units.

19. A polypeptide according to any one of claims 15 to 18 further including one or more bioactive peptides.

20. A polypeptide according to claim 19, wherein said bioactive peptide is inserted at the N or C terminus of said oleosin repeat units or between two or more oleosin repeat units.

21. A polypeptide according to any one of claims 15 to 20, wherein the oleosin repeat units are from white clover or sesame seed or are a synthetic or recombinant version thereof.

22. A polypeptide according to claim 15 including an amino acid sequence selected from the group consisting of Seq ID Nos: 47, 48, 51, 58-63 and 74-83, and functionally active fragments and variants thereof.

23. A polypeptide according to claim 15 produced by expression of a construct according to claim 1.

24. A lipid encapsulated by a polypeptide according to any one of claims 15 to 23.

25. A method of manipulating lipids in a plant, said method including introducing into said plant a construct according to any one of claims 1 to 11.

26. A method according to claim 25, wherein manipulation of lipids includes alteration of emulsification properties, physiochemical properties or degree of biohydrogenation of said lipids.

27. A method of:

(i) altering stability of an oil body in a plant, said method including modifying an oleosin in said plant to include three or more oleosin repeat units, or

2006303870 07 Nov 2012

(ii) delivering compounds and/or organisms to animals including humans, said method including encapsulation said compound or organism in an oil body including three or more oleosin repeat units and administering said oil body to said animal, or

(iii) altering the emulsification properties of an oleosin, said method including recombinantly producing the oleosin with three or more oleosin repeat units.

28. A method of:

(i) altering biohydrogenation of a lipid, said method including encapsulating said lipid in a recombinant polypeptide including three or more oleosin repeat units, or

(ii) protecting unsaturated lipids from biohydrogenation, said method including incorporating said unsaturated lipids into an oil body including a recombinant polypeptide including three or more oleosin repeat units, or

(iii) delivering a bioactive peptide, to animals including humans, said method including inserting said peptide at the N- or C- terminus of a series of three or more oleosin repeat units or between two or more oleosin repeat units to produce a recombinant polypeptide.

29. A recominbant oleosin with altered emulsification properties, said oleosin including three or more oleosin repeat units.

30. A recombinant oleosin according to claim 29, wherein said oleosin includes between three and ten oleosin repeat units.

31. A construct according to claim 1 or a polypeptide according to claim 15 substantially as hereinbefore described with reference to any one of the applicable examples.

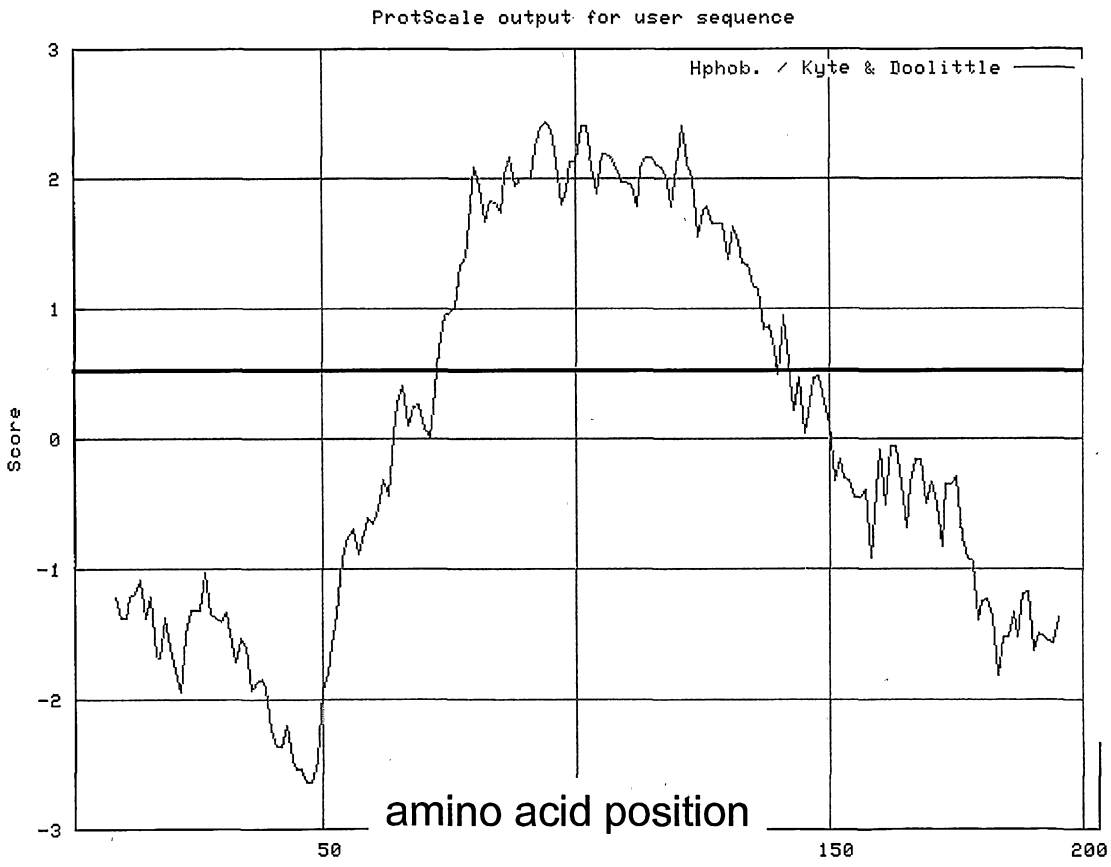

ProtScale output for user sequence
Hphob. / Kyte & Doolittle
Score
3
2
1
0
-1
-2
-3
amino acid position
50
150
200

Figure 1

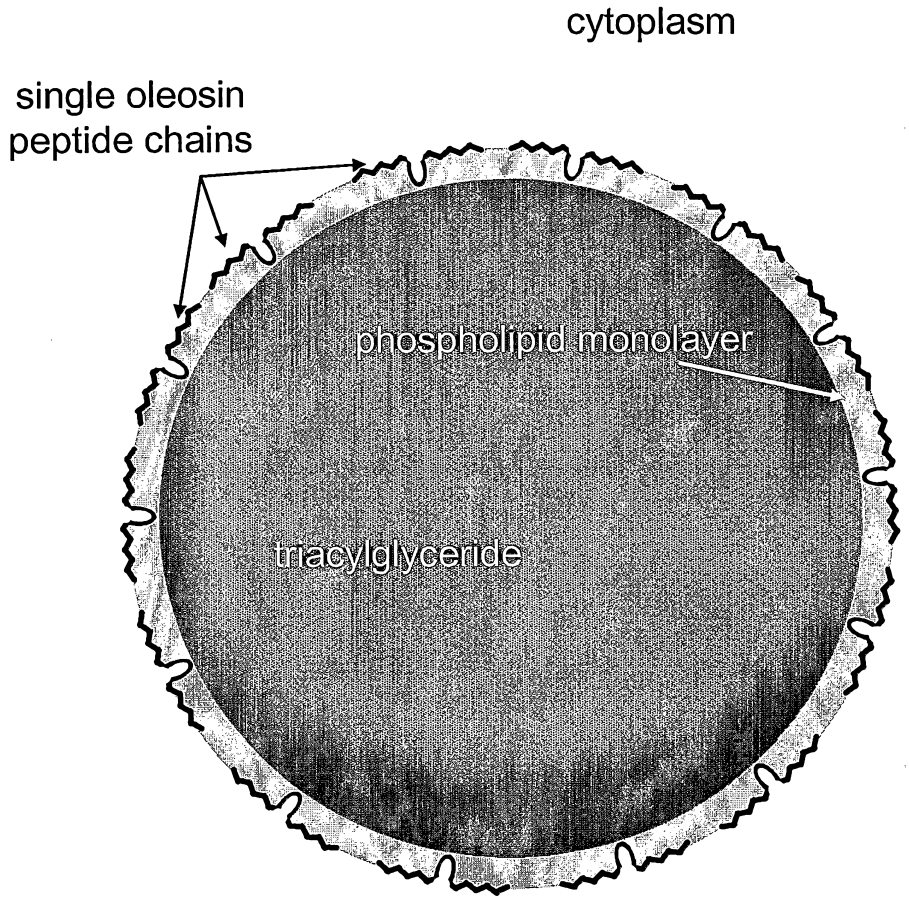

cytoplasm
single oleosin
peptide chains
phospholipid monolayer
triacylglyceride

Figure 2

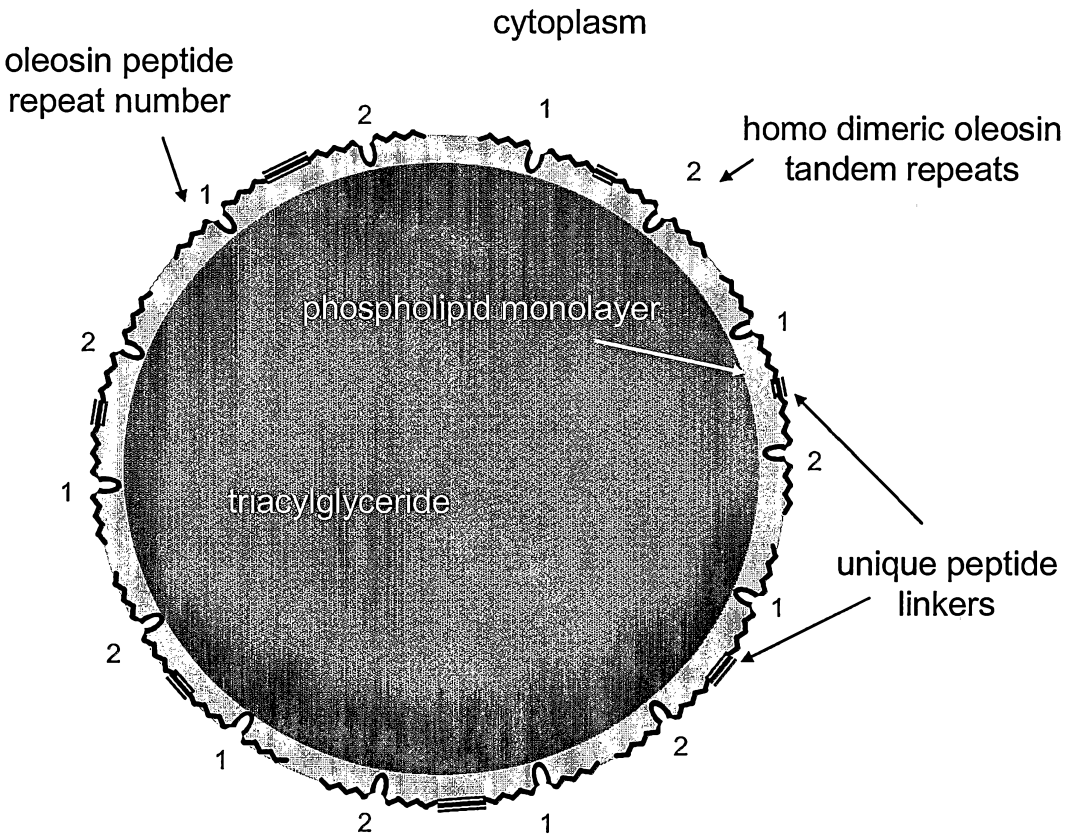

cytoplasm
oleosin peptide repeat number
homo dimeric oleosin tandem repeats
phospholipid monolayer
triacylglyceride
unique peptide linkers
1
2

Figure 3

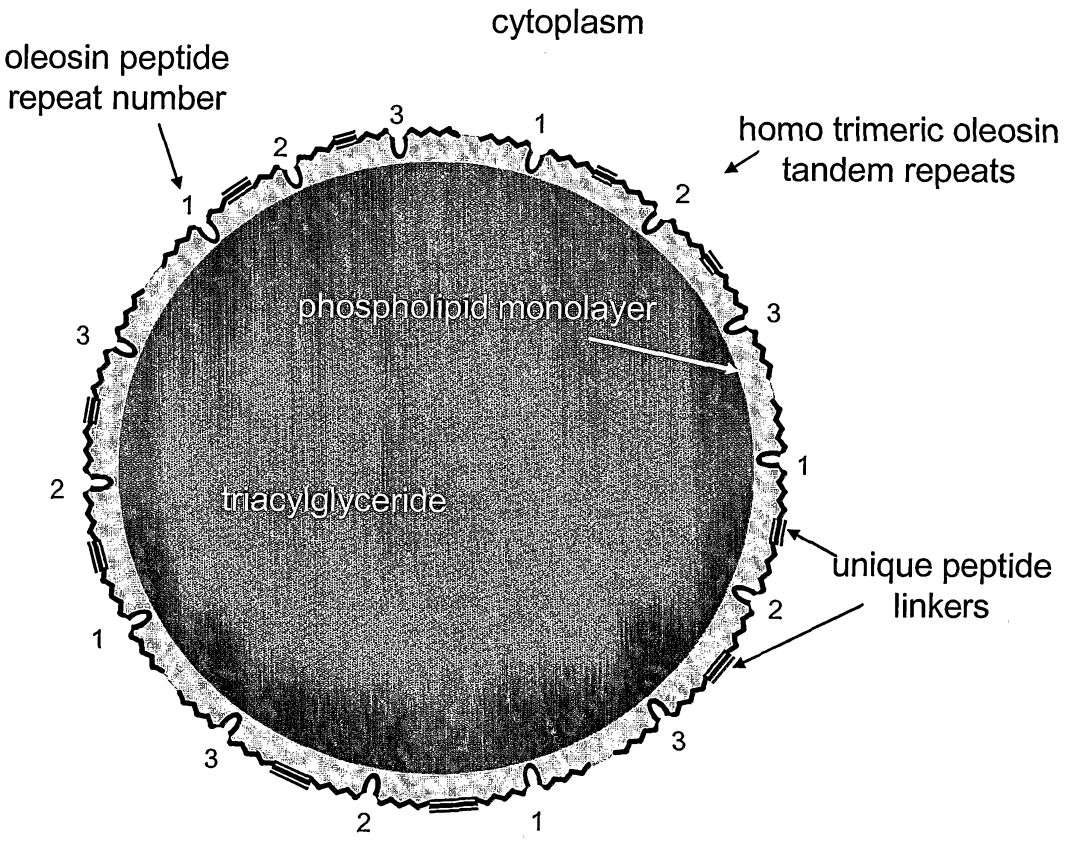

cytoplasm
oleosin peptide repeat number
homo trimeric oleosin tandem repeats
phospholipid monolayer
triacylglyceride
unique peptide linkers
1
2
3

Figure 4

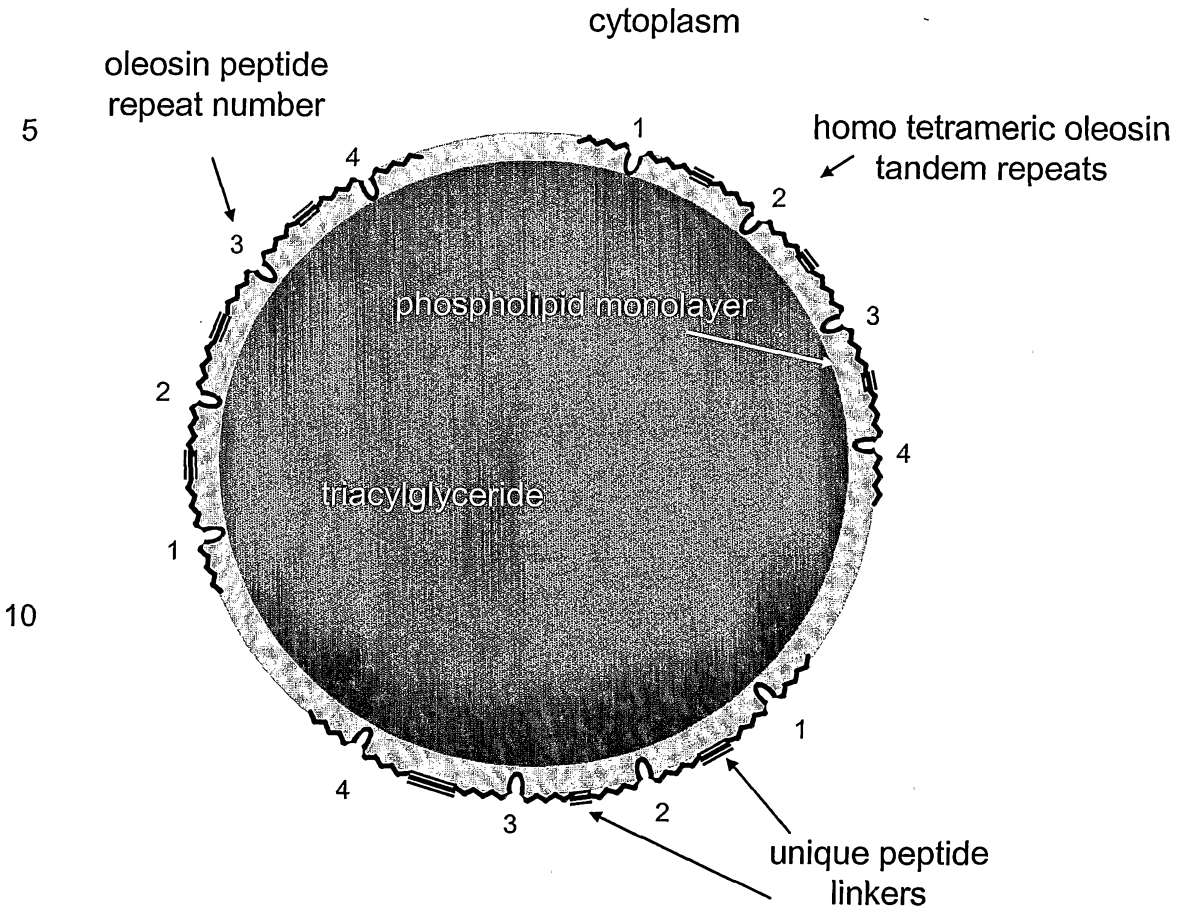

cytoplasm
oleosin peptide repeat number
homo tetrameric oleosin tandem repeats
phospholipid monolayer
triacylglyceride
unique peptide linkers
1
2
3
4

Figure 5

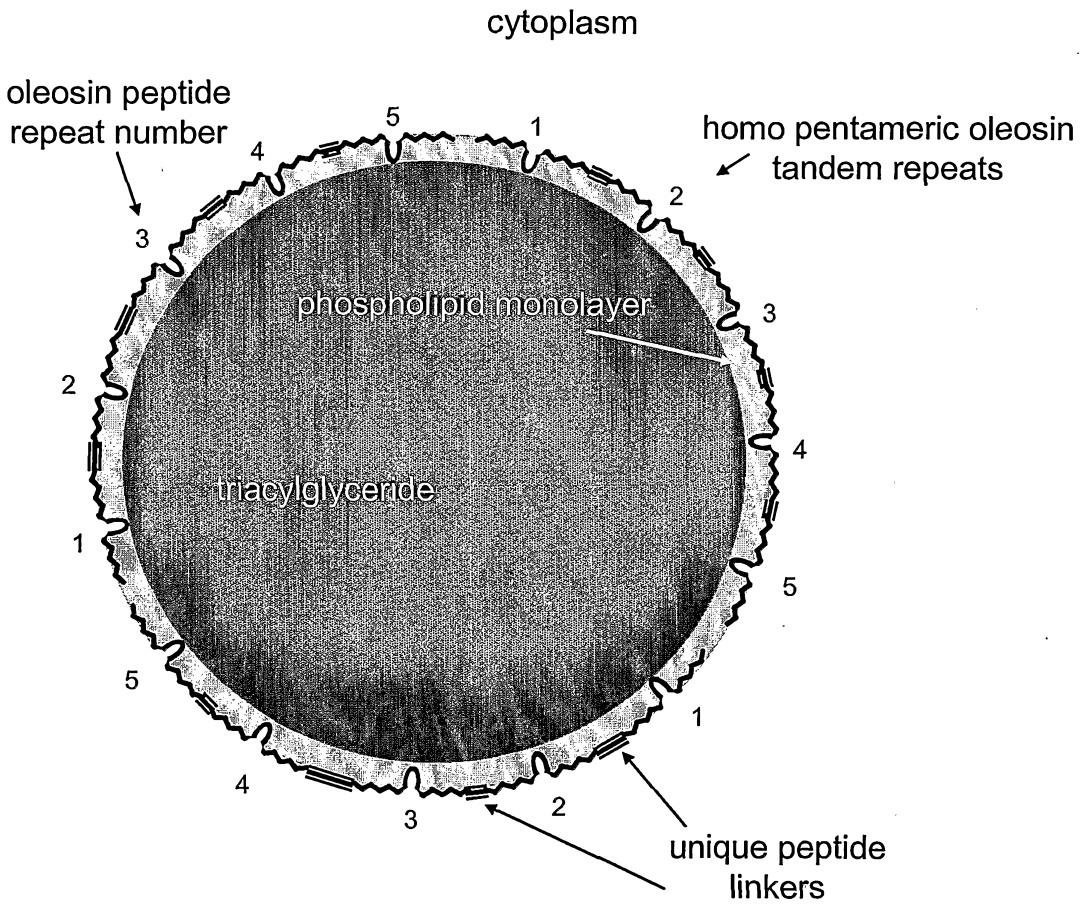

cytoplasm
oleosin peptide repeat number
homo pentameric oleosin tandem repeats
phospholipid monolayer
triacylglyceride
unique peptide linkers
1
2
3
4
5

Figure 6

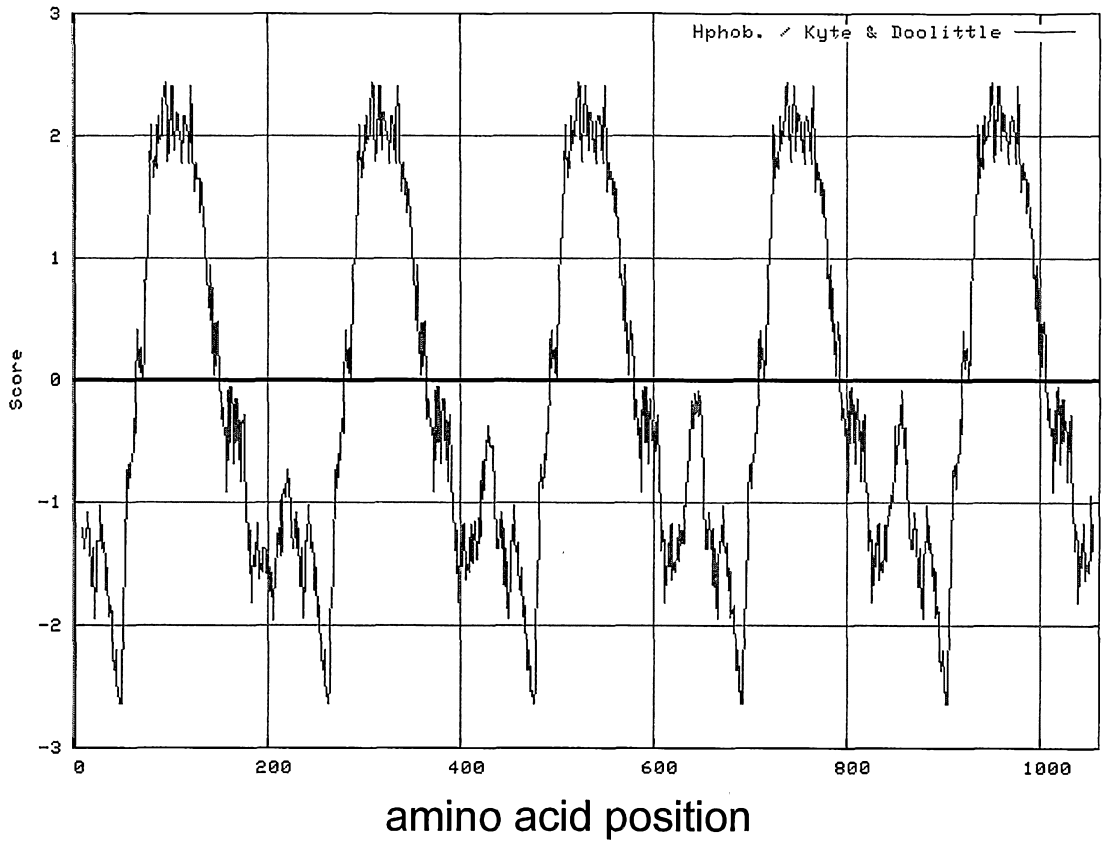

Hphob. / Kyte & Doolittle
Score
3
2
1
0
-1
-2
-3
0
200
400
600
800
1000
amino acid position

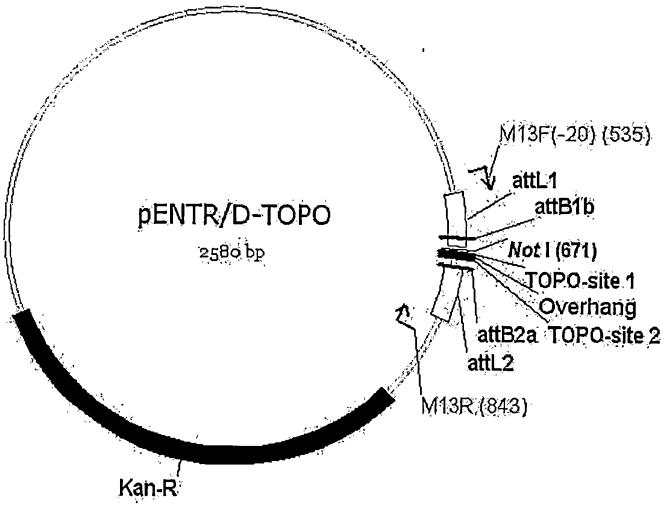
Figure 7


SacI • • • • • • XbaI SpeI BamHI SmaI
GAG CTC CAC CGC GGT GGC GGC CGC TCT AGA ACT AGT □G ATCC CCC GGG
E L H R G G G R S R T S G S P G
PstI EcoRI EcoRV HindIII Bsp106I • SalI XhoI
CTG CAG GAA TTC GAT ATC AAG CTT ATC GAT ACC GTC GAC CTC GAG
L Q E F D I K L I D T V D L E

Figure 8

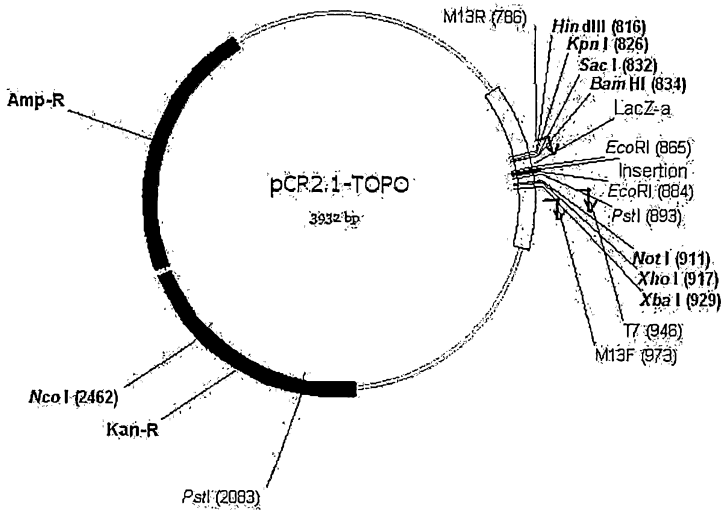
	restriction sites	additional amino acids	~18bp of oleosin sequence as initiator of
Primer	RSRSRS	AAAAAA	OLESOSIN SEQUENCE
PCR Product	***Enz*I/aacids/OLEOSIN/aacids/*Enz*II**		
TOPO Cloned	**generates pOLE01,02,03,04,05-MC**		

Figure 9


pENTR/D-TOPO
2580 bp
M13F(-20) (535)
attL1
attB1b
Not I (671)
TOPO-site 1
Overhang
attB2a TOPO-site 2
attL2
M13R (843)
Kan-R

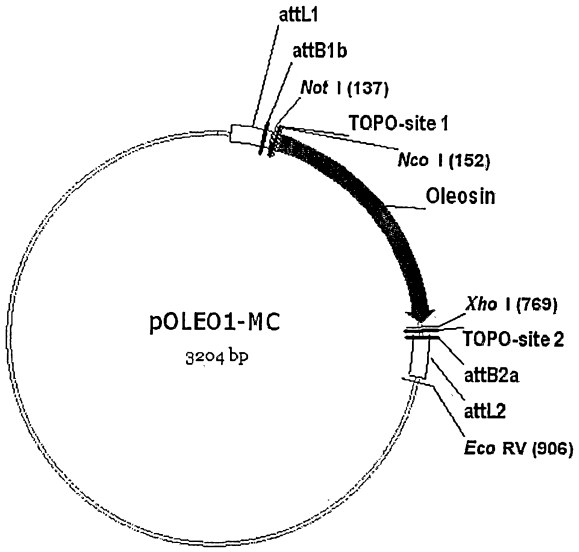
5

Figure 10


M13R (786)
Hin dIII (816)
Kpn I (826)
Sac I (832)
Bam HI (834)
LacZ-a
EcoRI (865)
Insertion
EcoRI (884)
PstI (893)
Not I (911)
Xho I (917)
Xba I (929)
T7 (946)
M13F (973)
Amp-R
pCR2.1-TOPO
3932 bp
Nco I (2462)
Kan-R
PstI (2083)

5

Figure 11


attL1
attB1b
Not I (137)
TOPO-site 1
Nco I (152)
Oleosin
pOLEO1-MC
3204 bp
Xho I (769)
TOPO-site 2
attB2a
attL2
Eco RV (906)

5

Figure 12

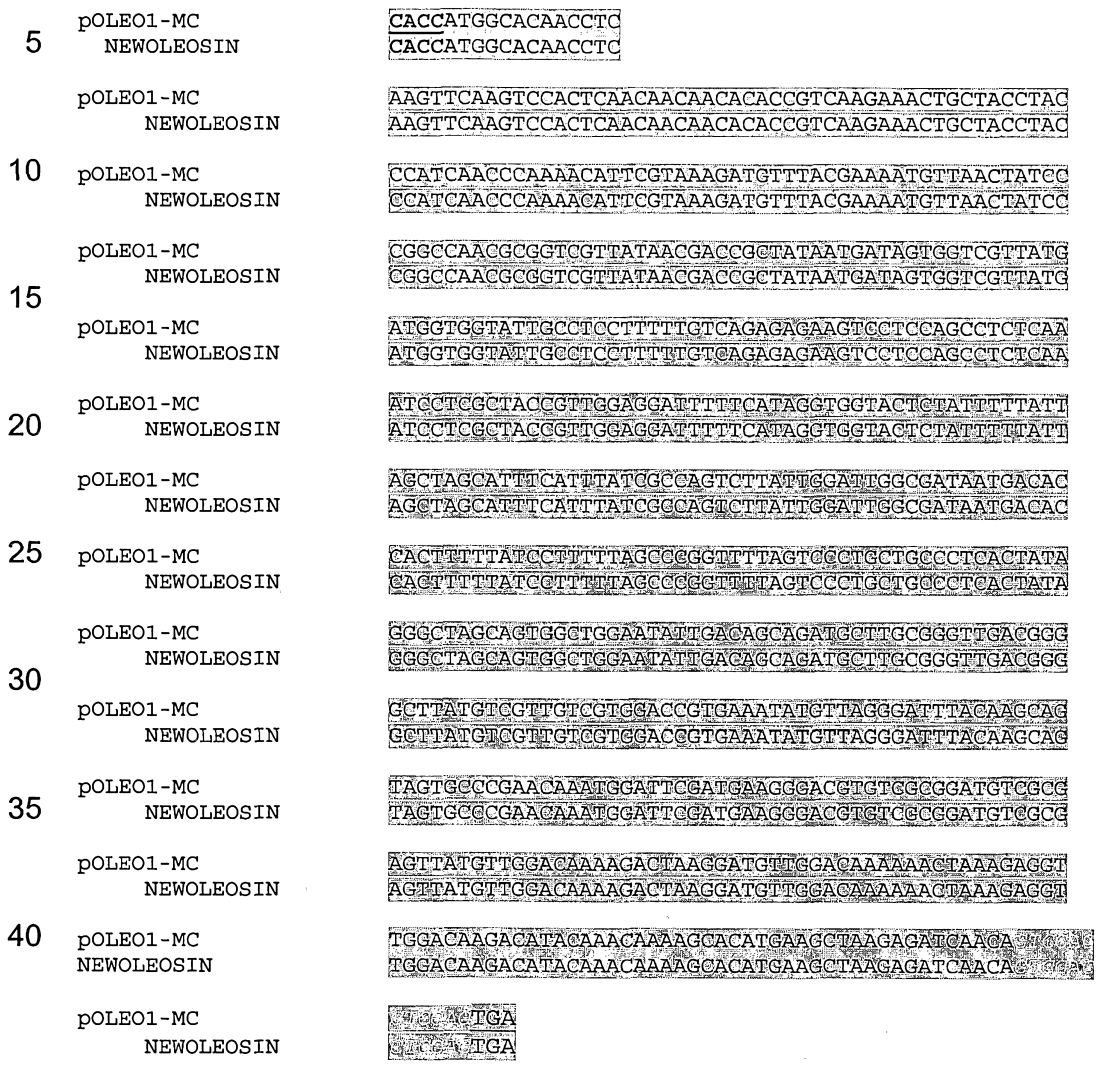

pOLEO1-MC
NEWOLEOSIN
CACCATGGCACAACCTC
CACCATGGCACAACCTC
AAGTTCAAGTCCACTCAACAACAACACACCGTCAAGAAACTGCTACCTAC
AAGTTCAAGTCCACTCAACAACAACACACCGTCAAGAAACTGCTACCTAC
CCATCAACCCAAAACATTCGTAAAGATGTTTACGAAAATGTTAACTATCC
CCATCAACCCAAAACATTCGTAAAGATGTTTACGAAAATGTTAACTATCC
CGGCCAACGCGGTCGTTATAACGACCGCTATAATGATAGTGGTCGTTATG
CGGCCAACGCGGTCGTTATAACGACCGCTATAATGATAGTGGTCGTTATG
ATGGTGGTATTGCCTCCTTTTTGTCAGAGAGAAGTCCTCCAGCCTCTCAA
ATGGTGGTATTGCCTCCTTTTTGTCAGAGAGAAGTCCTCCAGCCTCTCAA
ATCCTCGCTACCGTTGGAGGATTTTTCATAGGTGGTACTCTATTTTTATT
ATCCTCGCTACCGTTGGAGGATTTTTCATAGGTGGTACTCTATTTTTATT
AGCTAGCATTTCATTTATCGGCCAGTCTTATTGGATTGGCGATAATGACAC
AGCTAGCATTTCATTTATCGGCAGTCTTATTGGATTGGCGATAATGACAC
CACTTTTTATCCTTTTTAGCCCGGTTTTAGTCCCTGCTGCCCTCACTATA
CACTTTTTATCGTTTTTAGCCCGGTTTTAGTCCCTGCTGCCTCACTATA
GGGCTAGCAGTGGGTGGAATATTGACAGCAGATGCTTGCGGGTTGACGGG
GGGCTAGCAGTGGCTGGAATATTGACAGCAGATGCTTGCGGGTTGACGGG
GCTTATGTCGTTGTCGTGGACCGTGAAATATGTTAGGGATTTACAAGCAG
GCTTATGTCGTTGTCGTGGACCGTGAAATATGTTAGGGATTTACAAGCAG
TAGTGCCCGAACAAATGGATTCGATGAAGGGACGTGTGGCGGATGTCGCG
TAGTGCCCGAACAAATGGATTCGATGAAGGGACGTGTCGCCGATGTCGCG
AGTTATGTTGGACAAAAGACTAAGGATGTTGGACAAAAAACTAAAGAGGT
AGTTATGTTGGACAAAAGACTAAGGATGTTGGACAAAAAACTAAAGAGGT
TGGACAAGACATACAAACAAAAGCACATGAAGCTAAGAGATCAACA
TGGACAAGACATACAAACAAAAGCACATGAAGCTAAGAGATCAACA
TGA
TGA

Figure 13

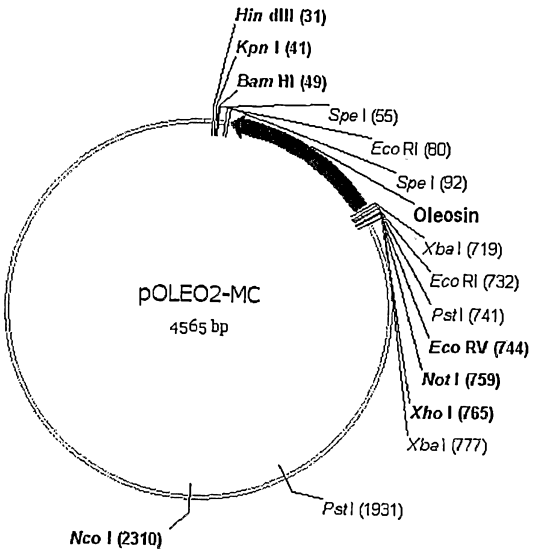

Hin dIII (31)
Kpn I (41)
Bam HI (49)
Spe I (55)
Eco RI (80)
Spe I (92)
Oleosin
Xba I (719)
Eco RI (732)
Pst I (741)
Eco RV (744)
Not I (759)
Xho I (765)
Xba I (777)
pOLEO2-MC
4565 bp
Pst I (1931)
Nco I (2310)

Figure 14

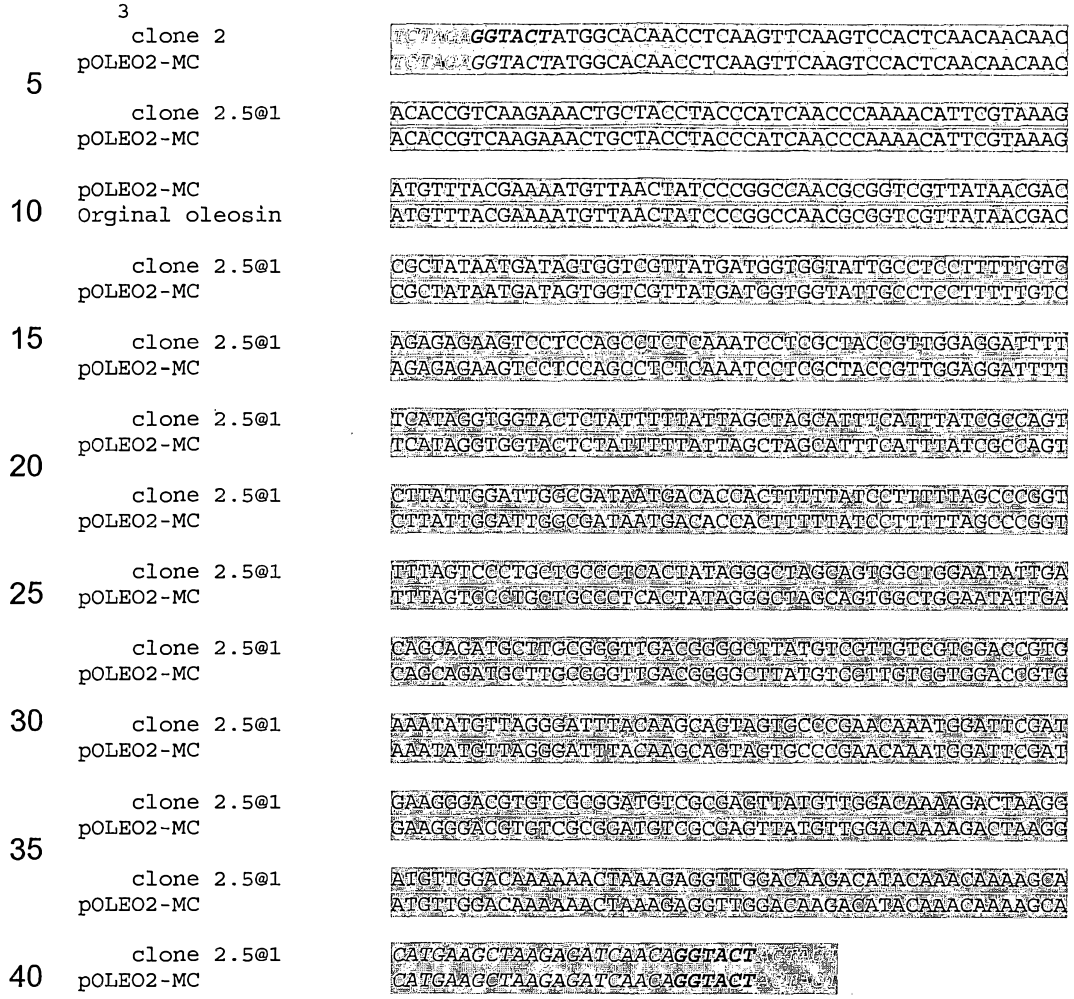

3
clone 2
pOLEO2-MC
clone 2.5@1
pOLEO2-MC
pOLEO2-MC
Orginal oleosin
clone 2.5@1
pOLEO2-MC
clone 2.5@1
pOLEO2-MC
clone 2.5@1
pOLEO2-MC
clone 2.5@1
pOLEO2-MC
clone 2.5@1
pOLEO2-MC
clone 2.5@1
pOLEO2-MC
clone 2.5@1
pOLEO2-MC
clone 2.5@1
pOLEO2-MC
clone 2.5@1
pOLEO2-MC
clone 2.5@1
pOLEO2-MC

Figure 15

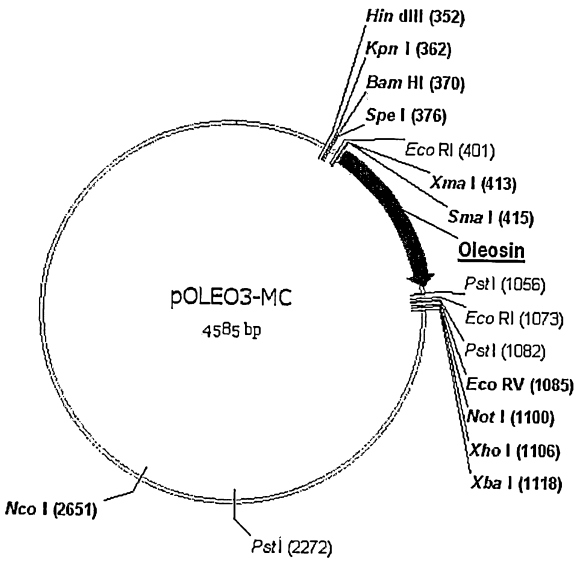

Hin dIII (352)
Kpn I (362)
Bam HI (370)
Spe I (376)
Eco RI (401)
Xma I (413)
Sma I (415)
Oleosin
pOLEO3-MC
4585 bp
PstI (1056)
Eco RI (1073)
PstI (1082)
Eco RV (1085)
Not I (1100)
Xho I (1106)
Xba I (1118)
Nco I (2651)
PstI (2272)

Figure 16

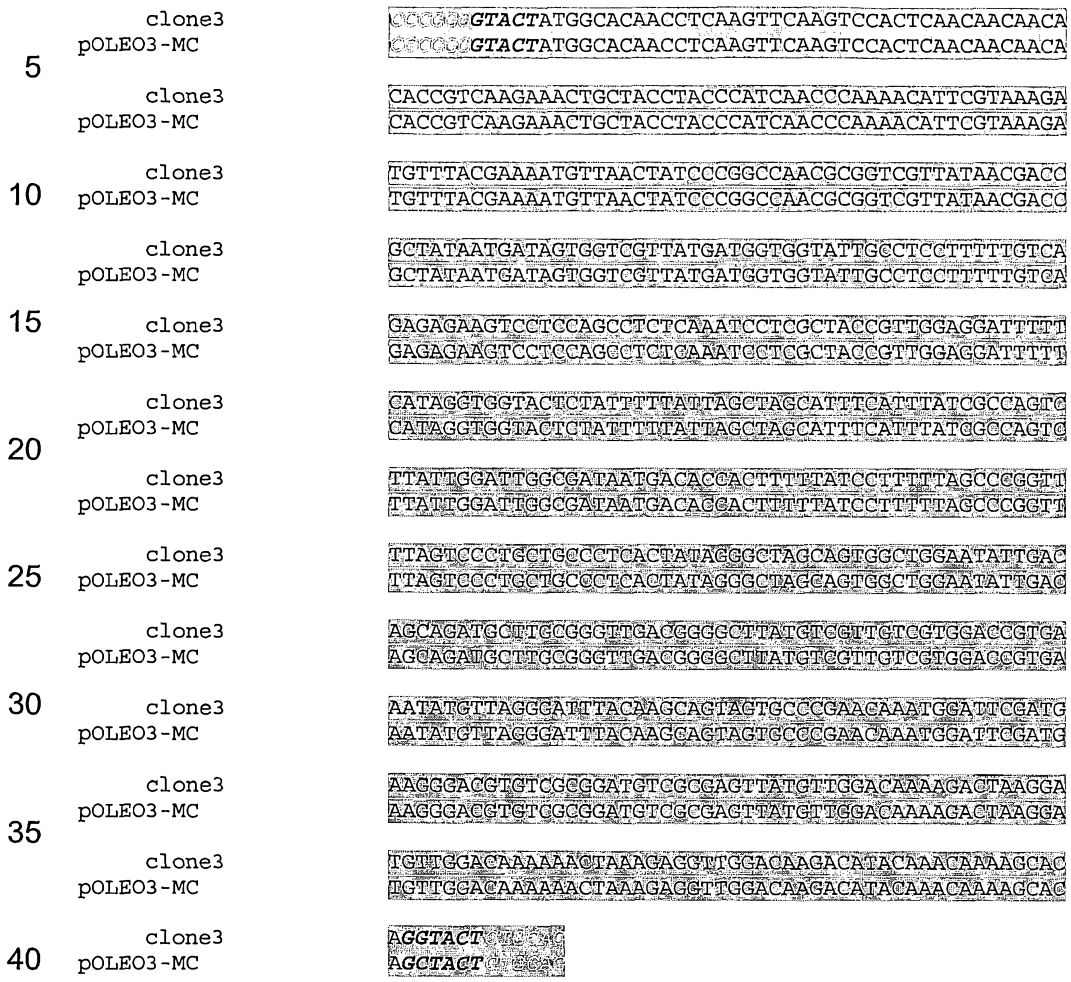

clone3
pOLEO3-MC
GTACTATGGCACAACCTCAAGTTCAAGTCCACTCAACAACAACA
GTACTATGGCACAACCTCAAGTTCAAGTCCACTCAACAACAACA
CACCGTCAAGAAACTGCTACCTACCCATCAACCCAAAACATTCGTAAAGA
CACCGTCAAGAAACTGCTACCTACCCATCAACCCAAAACATTCGTAAAGA
TGTTTACGAAAATGTTAACTATCCCGGCCAACGCGGTCGTTATAACGACC
TGTTTACGAAAATGTTAACTATCCCGGCCAACGCGGTCGTTATAACGACC
GCTATAATGATAGTGGTCGTTATGATGGTGGTATTGCCTCCTTTTTGTCA
GCTATAATGATAGTGGTCGTTATGATGGTGGTATTGCCTCCTTTTTGTCA
GAGAGAAGTCCTCCAGCCTCTCAAATCCTCGCTACCGTTGGAGGATTTTT
GAGAGAAGTCCTCCAGCCTCTCAAATGCTCGCTACCGTTGGAGGATTTTT
CATAGGTGGTACTCTATTTTTATTAGCTAGCATTTCATTTATCGCCAGTC
CATAGGTGGTACTCTATTTTTATTAGCTAGCATTTCATTTATCGCCAGTC
TTATTGGATTGGCGATAATGACACCACTTTTTATCCTTTTTAGCCCGGTT
TTATTGGATTGGCGATAATGACACCACTTTTTATCCTTTTTAGCCCGGTT
TTAGTCCCTGCTGCCCTCACTATAGCGCTAGCAGTGGCTGGAATATTGAC
TTAGTCCCTGCTGCCCTCACTATAGCGCTAGCAGTGGCTGGAATATTGAC
AGCAGATGCTTGCGGGTTGACGGGGCTTATGTCGTTGTCGTGGACCGTGA
AGCAGATGCTTGCGGGTTGACGGGGCTTATGTCGTTGTCGTGGACCGTGA
AATATGTTAGGGATTTACAAGCAGTAGTGCCCGAACAAATGGATTCGATG
AATATGTTAGGGATTTACAAGCAGTAGTGCCCGAACAAATGGATTCGATG
AAGGGACGTGTCGCGGATGTCGCGAGTTATGTTGGACAAAAGACTAAGGA
AAGGGACGTGTCGCGGATGTCGCGAGTTATGTTGGACAAAAGACTAAGGA
TGTTGGACAAAAAACTAAACAGCGTTGGACAAGACATACAAACAAAAGCAC
TGTTGGACAAAAAACTAAAGAGGTTGGACAAGACATACAAACAAAAGCAC
AGGTACT
AGCTACT

Figure 17

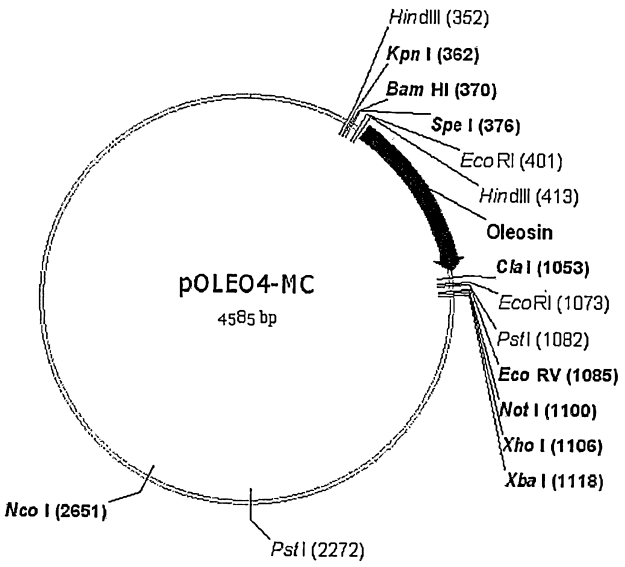

HindIII (352)
Kpn I (362)
Bam HI (370)
Spe I (376)
EcoRI (401)
HindIII (413)
Oleosin
Cla I (1053)
EcoRI (1073)
PstI (1082)
Eco RV (1085)
Not I (1100)
Xho I (1106)
Xba I (1118)
pOLEO4-MC
4585 bp
Nco I (2651)
PstI (2272)

Figure 18

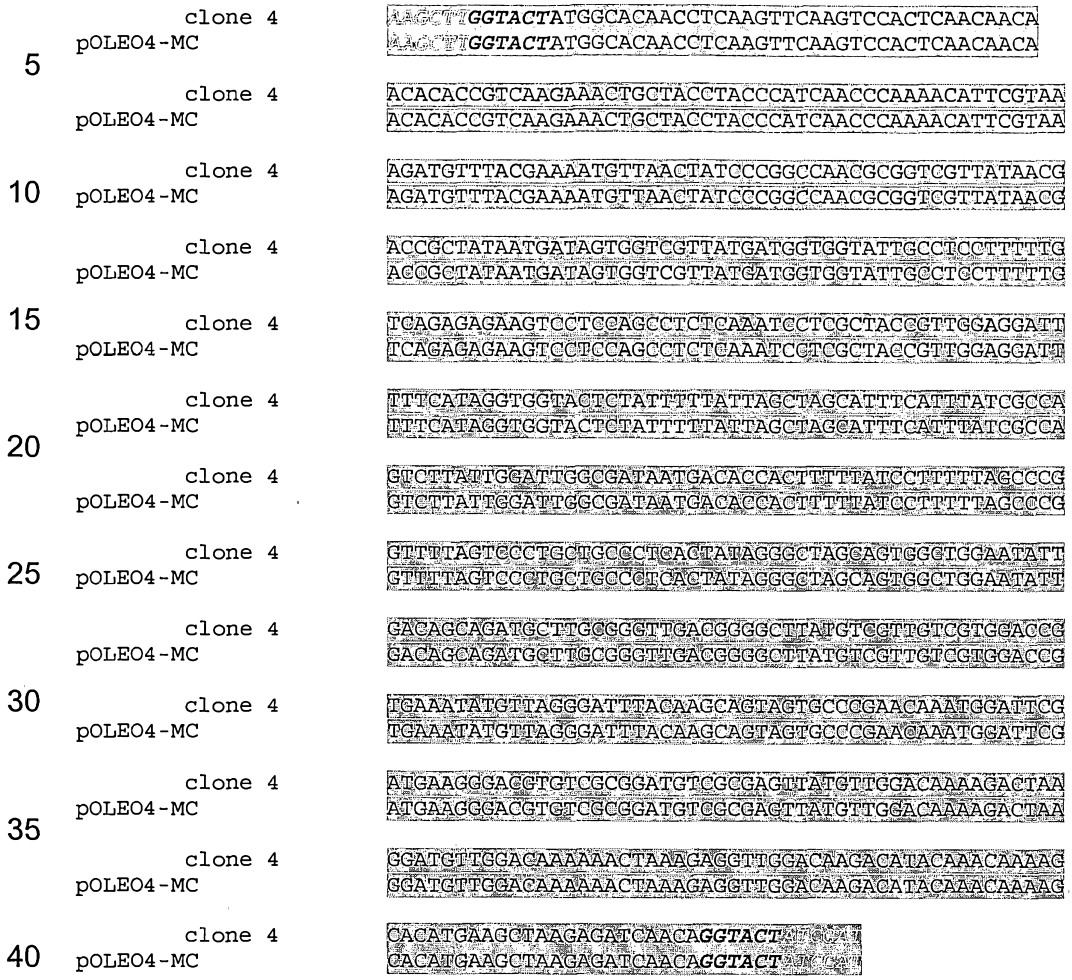

clone 4
pOLEO4-MC
AAGCTTGGTACTATGGCACAACCTCAAGTTCAAGTCCACTCAACAACA
ACACACCGTCAAGAAACTGCTACCTACCCATCAACCCAAAACATTCGTAA
AGATGTTTACGAAAATGTTAACTATCCCGGCCAACGCGGTCGTTATAACG
ACCGCTATAATGATAGTGGTCGTTATGATGGTGGTATTGCCTCCTTTTTG
TCAGAGAGAAGTCCTCCAGCCTCTCAAATCCTCGCTACCGTTGGAGGATT
TTTCATAGGTGGTACTCTATTTTTATTAGCTAGCATTTCATTTATCGCCA
GTCTTATTGGATTGGCGATAATGACACCACTTTTTATCCTTTTTAGCCCG
GTTTTAGTCCCTGCTGCCCTCACTATAGGGCTAGCAGTGGCTGGAATATT
GACAGCAGATGCTTGCGGGTTGACGGGGCTTATGTCGTTGTCGTGGACCG
TGAAATATGTTAGGGATTTACAAGCAGTAGTGCCCGAACAAATGGATTCG
ATGAAGGGACGTGTCGCGGATGTCGCGAGTTATGTTGGACAAAAGACTAA
GGATGTTGGACAAAAAACTAAAGAGGTTGGACAAGACATACAAACAAAAG
CACATGAAGCTAAGAGATCAACAGGTACT

Figure 19

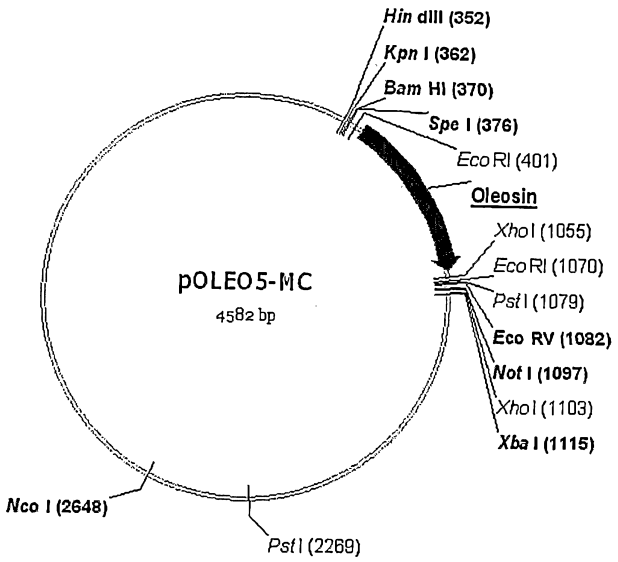

Hin dIII (352)
Kpn I (362)
Bam HI (370)
Spe I (376)
EcoRI (401)
Oleosin
XhoI (1055)
EcoRI (1070)
PstI (1079)
Eco RV (1082)
Not I (1097)
XhoI (1103)
Xba I (1115)
pOLEO5-MC
4582 bp
Nco I (2648)
PstI (2269)

Figure 20

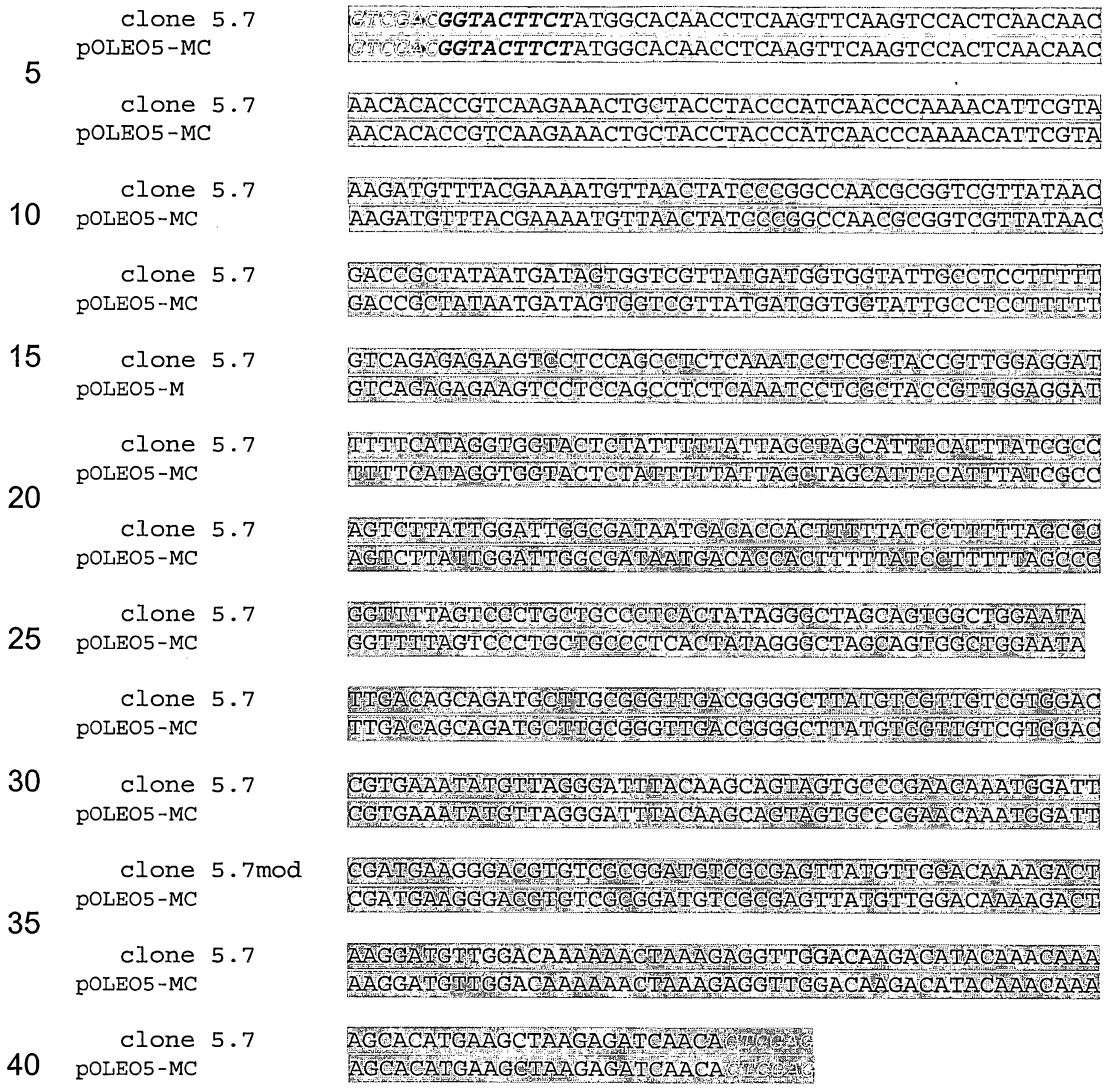

clone 5.7
pOLEO5-MC
GTCGACGGTACTTCTATGGCACAACCTCAAGTTCAAGTCCACTCAACAAC
GTCGACGGTACTTCTATGGCACAACCTCAAGTTCAAGTCCACTCAACAAC
clone 5.7
pOLEO5-MC
AACACACCGTCAAGAAACTGCTACCTACCCATCAACCCAAAACATTCGTA
AACACACCGTCAAGAAACTGCTACCTACCCATCAACCCAAAACATTCGTA
clone 5.7
pOLEO5-MC
AAGATGTTTACGAAAATGTTAACTATCCCGGCCAACGCGGTCGTTATAAC
AAGATGTTTACGAAAATGTTAACTATCCCGGCCAACGCGGTCGTTATAAC
clone 5.7
pOLEO5-MC
GACCGCTATAATGATAGTGGTCGTTATGATGGTGGTATTGCCTCCTTTTT
GACCGCTATAATGATAGTGGTCGTTATGATGGTGGTATTGCCTCCTTTTT
clone 5.7
pOLEO5-M
GTCAGAGAGAAGTGCTCCAGCCTCTCAAATCCTCGCTACCGTTGGAGGAT
GTCAGAGAGAAGTCCTCCAGCCTCTCAAATCCTCGCTACCGTTGGAGGAT
clone 5.7
pOLEO5-MC
TTTTCATAGGTGGTACTCTATTTTTATTAGCTAGCATTTCATTTATCGCC
TTTTCATAGGTGGTACTCTATTTTTATTAGCTAGCATTTCATTTATCGCC
clone 5.7
pOLEO5-MC
AGTCTTATTGGATTGGCGATAATGACACCACTTTTTATCCTTTTTAGCGC
AGTCTTATTGGATTGGCGATAATGACACCACTTTTTATCCTTTTTAGCCC
clone 5.7
pOLEO5-MC
GGTTTTAGTCCCTGCTGCCCTCACTATAGGGCTAGCAGTGGCTGGAATA
GGTTTTAGTCCCTGCTGCCCTCACTATAGGGCTAGCAGTGGCTGGAATA
clone 5.7
pOLEO5-MC
TTGACAGCAGATGCTTGCGGGTTGACGGGGCTTATGTCGTTGTCGTGGAC
TTGACAGCAGATGCTTGCGGGTTGACGGGGCTTATGTCGTTGTCGTGGAC
clone 5.7
pOLEO5-MC
CGTGAAATATGTTAGGGATTTACAAGCAGTAGTGCCCGAACAAATGGATT
CGTGAAATATGTTAGGGATTTACAAGCAGTAGTGCCCGAACAAATGGATT
clone 5.7mod
pOLEO5-MC
CGATGAAGGGACGTGTCGCGGATGTCGCGAGTTATGTTGGACAAAAGACT
CGATGAAGGGACGTGTCGCGGATGTCGCGAGTTATGTTGGACAAAAGACT
clone 5.7
pOLEO5-MC
AAGGATGTTGGACAAAAAACTAAAGAGGTTGGACAAGACATACAAACAAA
AAGGATGTTGGACAAAAAACTAAAGAGGTTGGACAAGACATACAAACAAA
clone 5.7
pOLEO5-MC
AGCACATGAAGCTAAGAGATCAACA
AGCACATGAAGCTAAGAGATCAACA

Figure 21

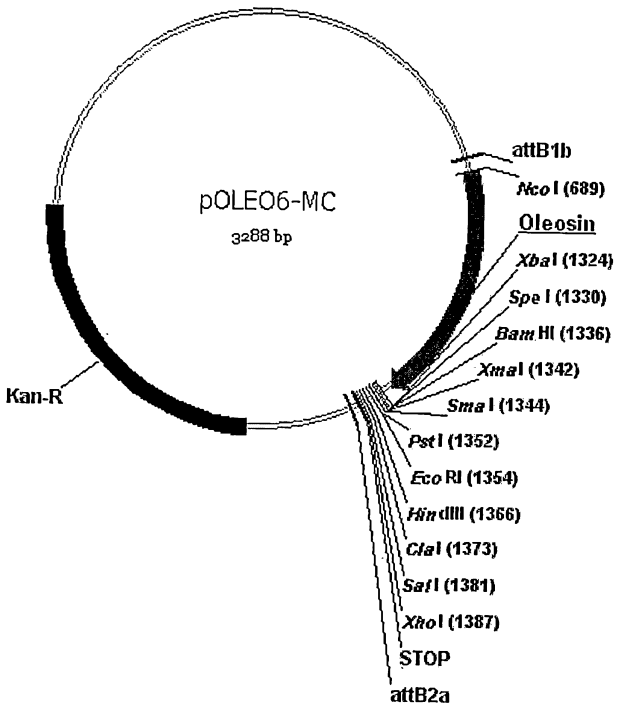

pOLEO6-MC
3288 bp
attB1b
Nco I (689)
Oleosin
Xba I (1324)
Spe I (1330)
Bam HI (1336)
Xma I (1342)
Sma I (1344)
Pst I (1352)
Eco RI (1354)
Hin dIII (1366)
Cla I (1373)
Sal I (1381)
Xho I (1387)
STOP
attB2a
Kan-R

Figure 22

5

```
pOLEO6-MC          CACCATGGCACAACC
   clone 6 fwd     CACCATGGCACAACC

pOLEO6-MC          TCAAGTTCAAGTCCACTCAACAACAACACACCGTCAAGAAACTGCTACCT
   clone 6 fwd     TCAAGTTCAAGTCCACTCAACAACAACACACCGTCAAGAAACTGCTACCT

pOLEO6-MC          ACCCATCAACCCAAAACATTCGTAAAGATGTTTACGAAAATGTTAACTAT
   clone 6 fwd     ACCCATCAACCCAAAACATTCGTAAAGATGTTTACGAAAATGTTAACTAT

pOLEO6-MC          CCCGGCCAACGCGGTCGTTATAACGACCGCTATAATGATAGTGGTCGTTA
   clone 6 fwd     CCCGGCCAACGCGGTCGTTATAACGACCGCTATAATGATAGTGGTCGTTA

pOLEO6-MC          TGATGGTGGTATTGCCTCCTTTTTGTCAGAGAGAAGTCCTCCAGCCTCTC
   clone 6 fwd     TGATGGTGGTATTGCCTCCTTTTTGTCAGAGAGAAGTCCTCCAGCCTCTC

pOLEO6-MC          AAATCCTCGCTACCGTTGGAGGATTTTTCATAGGTGGTACTCTATTTTTA
   clone 6 fwd     AAATCCTCGCTACCGTTGGAGGATTTTTCATAGGTGGTACTCTATTTTTA

pOLEO6-MC          TTAGCTAGCATTTCATTTATCGCCAGTCTTATTGGATTGGCGATAATGAC
   clone 6 fwd     TTAGCTAGCATTTCATTTATCGCCAGTCTTATTGGATTGGCGATAATGAC

pOLEO6-MC          ACCACTTTTTATCCTTTTTAGCCCGGTTTTAGTCCCTGCTGCCCTCACTA
   clone 6 rvs     ACCACTTTTTATCCTTTTTAGCCCGGTTTTAGTCCCTGCTGCCCTCACTA

pOLEO6-MC          TAGGGCTAGCAGTGGCTGGAATATTGACAGCAGATGCTTGCGGGTTGACG
   clone 6 rvs     TAGGGCTAGCAGTGGCTGGAATATTGACAGCAGATGCTTGCGGGTTGACG

pOLEO6-MC          GGGCTTATGTCGTTGTCGTGGACCGTGAAATATGTTAGGGATTTACAAGC
   clone 6 rvs     GGGCTTATGTCGTTGTCGTGGACCGTGAAATATGTTAGGGATTTACAAGC

pOLEO6-MC          AGTAGTGCCCGAACAAATGGATTCGATGAAGGGACGTGTCGCGGATGTCG
   clone 6 rvs     AGTAGTGCCCGAACAAATGGATTCGATGAAGGGACGTGTCGCGGATGTCG

pOLEO6-MC          CGAGTTATGTTGGACAAAAGACTAAGGATGTTGGACAAAAAACTAAAGAG
   clone 6 rvs     CGAGTTATGTTGGACAAAAGACTAAGGATGTTGGACAAAAAACTAAAGAG

pOLEO6-MC          GTTGGACAAGACATACAAACAAAAGCACATGAAGCTAAGAGATCAACAGA
   clone 6 rvs     GTTGGACAAGACATACAAACAAAAGCACATGAAGCTAAGAGATCAACAGA

pOLEO6-MC          GCTCCACCGCGGTGGCGGCCGCTCTAGAACTAGTGGATCCCCCGGGCTGC
   clone 6 rvs     GCTCCACCGCGGTGGCGGCCGCTCTAGAACTAGTGGATCCCCCGGGCTGC

pOLEO6-MC          AGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGTGA
   clone 6 rvs     AGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGTGA
```

Figure 23

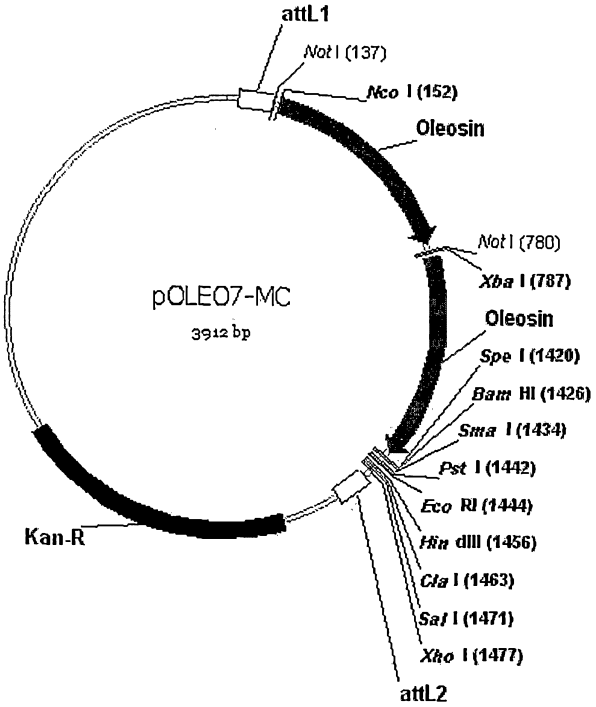

attL1
Not I (137)
Nco I (152)
Oleosin
Not I (780)
Xba I (787)
Oleosin
Spe I (1420)
Bam HI (1426)
Sma I (1434)
Pst I (1442)
Eco RI (1444)
Hin dIII (1456)
Cla I (1463)
Sal I (1471)
Xho I (1477)
attL2
Kan-R
pOLEO7-MC
3912 bp

Figure 24

```
pOLEO7-MC        CACCATGGCACAACCTCAAGTTCAAGTCCACTCAACAACAACACACCGTCAA
 clone 7.3fwd    CACCATGGCACAACCTCAAGTTCAAGTCCACTCAACAACAACACACCGTCAA

pOLEO7-MC        GAAACTGCTACCTACCCATCAACCCAAAACATTCGTAAAGATGTTTACGA
 clone 7.3fwd    GAAACTGCTACCTACCCATCAACCCAAAACATTCGTAAAGATGTTTACGA

pOLEO7-MC        AAATGTTAACTATCCCGGCCAACGCGGTCGTTATAACGACCGCTATAATG
 clone 7.3fwd    AAATGTTAACTATCCCGGCCAACGCGGTCGTTATAACGACCGCTATAATG

pOLEO7-MC        ATAGTGGTCGTTATGATGGTGGTATTGCCTCCTTTTTGTCAGAGAGAAGT
 clone 7.3fwd    ATAGTGGTCGTTATGATGGTGGTATTGCCTCCTTTTTGTCAGAGAGAAGT

pOLEO7-MC        CCTCCAGCCTCTCAAATCCTCGCTACCGTTGGAGGATTTTTCATAGGTGG
 clone 7.3fwd    CCTCCAGCCTCTCAAATCCTCGCTACCGTTGGAGGATTTTTCATAGGTGG

pOLEO7-MC        TACTCTATTTTTATTAGCTAGCATTTCATTTATCGCCAGTCTTATTGGAT
 clone 7.3fwd    TACTCTATTTTTATTAGCTAGCATTTCATTTATCGCCAGTCTTATTGGAT

pOLEO7-MC        TGGCGATAATGACACCACTTTTTATCCTTTTTAGCCCGGTTTTAGTCCCT
 clone 7.3fwd    TGGCGATAATGACACCACTTTTTATCCTTTTTAGCCCGGTTTTAGTCCCT

pOLEO7-MC        GCTGCCCTCACTATAGGGCTAGCAGTGGCTGGAATATTGACAGCAGATGC
 clone 7.3fwd    GCTGCCCTCACTATAGGGCTAGCAGTGGCTGGAATATTGACAGCAGATGC

pOLEO7-MC        TTGCGGGTTGACGGGGCTTATGTCGTTGTCGTGGACCGTGAAATATGTT
 clone 7.3fwd    TTGCGGGTTGACGGGGCTTATGTCGTTGTCGTGCACCGTGAAATATGTT

pOLEO7-MC        AGGGATTTACAAGCAGTAGTGCCCGAACAAATGGATTCGATGAAGGGACG
 clone 7.3fwd    AGGGATTTACAAGCAGTAGTGCCCGAACAAATGGATTCGATGAAGGGACG

pOLEO7-MC        TGTCGCGGATGTCGCGAGTTATGTTGGACAAAAGACTAAGGATGTTGGAC
 clone 7.3fwd    TGTCGCGGATGTCGCGAGTTATGTTGGACAAAAGACTAAGGATGTTGGAC

pOLEO7-MC        AAAAAACTAAAGAGGTTGGACAAGACATACAAACAAAAGCACATGAAGCT
 clone 7.3fwd    AAAAAACTAAAGAGGTTGGACAAGACATACAAACAAAAGCACATGAAGCT

pOLEO7-MC        AAGAGATCAACAGAGCTCCACCGCGGTGGCGGCCGCTCTAGAGGTACTAT
 clone 7.3fwd    AAGAGATCAACAGAGCTCCACCGCGGTGGCGGCCGCTCTAGAGGTACTAT

pOLEO7-MC        GGCACAACCTCAAGTTCAAGTCCACTCAACAACAACACACCGTCAAGAAA
 clone 7.3fwd    GGCACAACCTCAAGTTCAAGTCCACTCAACAACAACACACCGTCAAGAAA

pOLEO7-MC        CTGCTACCTACCCATCAACCCAAAACATTCGTAAAGATGTTTACCAAAA
 clone 7.3fwd    CTGCTACCTACCCATCAACCCAAAACATTCGTAAAGATGTTTACGAAAA

pOLEO7-MC        TGTTAACTATCCCGGCCAACGCGGTCGTTATAACGACCGCTATAATGAT
      7.3revs    TGTTAACTATCCCGGCCAACGCGGTCGTTATAACGACCGCTATAATGAT

pOLEO7-MC        AGTGGTCGTTATGATGGTGGTATTGCCTCCTTTTTGTCAGAGAGAAGTCC
      7.3revs    AGTGGTCGTTATGATGGTGGTATTGCCTCCTTTTTGTCAGAGAGAAGTCC

pOLEO7-MC        TCCAGCCTCTCAAATCCTCGCTACCGTTGGAGGATTTTTCATAGGTGGTA
      7.3revs    TCCAGCCTCTCAAATCCTCGCTACCGTTGGAGGATTTTTCATAGGTGGTA

pOLEO7-MC        CTCTATTTTTATTAGCTAGCATTTCATTTATCGCCAGTCTTATTGGATTG
      7.3revs    CTCTATTTTTATTAGCTAGCATTTCATTTATCGCCAGTCTTATTGGATTG

pOLEO7-MC        GCGATAATGACACCACTTTTTATCCTTTTTAGCCCGGTTTTAGTCCCTGC
      7.3revs    GCGATAATGACACCACTTTTTATCCTTTTTAGCCCGGTTTTAGTCCCTGC

pOLEO7-MC        TGCCCTCACTATAGGGCTAGCAGTGGCTGGAATATTGACAGCAGATGCTT
      7.3revs    TGCCCTCACTATAGGGCTAGCAGTGGCTGGAATATTGACAGCAGATGCTT
```

Figure 25

```
pOLEO7-MC       GCGGGTTCACGGGGCTTATGTCGTTGTCCTGGACCGTGAAATATGTTAGG
7.3revs         GCGGGTTGACGGGGCTTATGTCGTTGTCGTGGACCGTGAAATATGTTAGC

pOLEO7-MC       GATTTACAAGCAGTAGTGCCCGAACAAATGGATTCGATGAAGGGACGTGT
7.3revs         GATTTACAAGCAGTAGTGCCCGAACAAATGGATTCGATGAAGGGACGTGT

pOLEO7-MC       CGCGGATGTCGCGAGTTATGTTGGACAAAAGACTAAGGATGTTGGACAAA
7.3revs         CGCGGATGTCGCGAGTTATGTTGGACAAAAGACTAAGGATGTTGGACAAA

pOLEO7-MC       AAACTAAAGAGGTTCGACAAGACATACAAACAAAAGCACATGAAGCTAAG
7.3revs         AAACTAAAGAGGTTGGACAAGACATACAAACAAAAGCACATGAAGCTAAG

pOLEO7-MC       AGATCAACAGGTACTACTAGAACTAGTGGATCCCCCGGGCTGCAGGAATT
7.3revs         AGATCAACAGGTACTACTAGAACTAGTGGATCCCCCGGGCTGCAGGAATT

pOLEO7-MC       CGATATCAAGCTTATCGATACCGTCGACCTCGAGTGA
7.3revs         CGATATCAAGCTTATCGATACCGTCGACCTCGAGTGA
```

Figure 25 continued

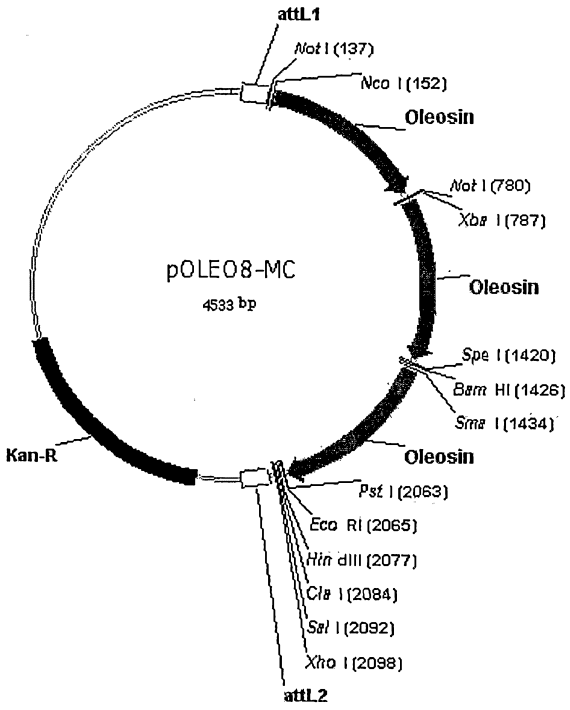

attL1
Not I (137)
Nco I (152)
Oleosin
Not I (780)
Xba I (787)
pOLEO8-MC
4533 bp
Oleosin
Spe I (1420)
Bam HI (1426)
Sma I (1434)
Kan-R
Oleosin
Pst I (2063)
Eco RI (2065)
Hin dIII (2077)
Cla I (2084)
Sal I (2092)
Xho I (2098)
attL2

Figure 26

```
pOLEO8-MC    CACCATGGCACAACCTCAAGTTCAAGTCCACTCAACAACAACACACCGTCAA
Clone8fwd    CACCATGGCACAACCTCAAGTTCAAGTCCACTCAACAACAACACACCGTCAA

pOLEO8-MC    GAAACTGCTACCTACCCATCAACCCAAAACATTCGTAAAGATGTTTACGGA
Clone8fwd    GAAACTGCTACCTACCCATCAACCCAAAACATTCGTAAAGATGTTTACGGA

pOLEO8-MC    AAATGTTAACTATCCCGGCCAACGCGGTCGTTATAACGACCGCTATAATG
Clone8fwd    AAATGTTAACTATCCCGGCCAACGCGGTCGTTATAACGACCGCTATAATG

pOLEO8-MC    ATAGTGGTCGTTATGATGGTGGTATTGCCTCCTTTTTGTCAGAGAGAAGT
Clone8fwd    ATAGTGGTCGTTATGATGGTGGTATTGCCTCCTTTTTGTCAGAGAGAAGT

pOLEO8-MC    CCTCCAGCCTCTCAAATCCTCGCTACCGTTGGAGGATTTTTCATAGGTGG
Clone8fwd    CCTCCAGCCTCTCAAATCCTCGCTACCGTTGGAGGATTTTTCATAGGTGG

pOLEO8-MC    TACTCTATTTTTATTAGCTAGCATTTCATTTATCGCCAGTCTTATTGGAT
Clone8fwd    TACTCTATTTTTATTAGCTAGCATTTCATTTATCGCCAGTCTTATTGGAT

pOLEO8-MC    TGGCGATAATGACACCACTTTTTATCCTTTTTAGCCCGGTTTTAGTCCCT
Clone8fwd    TGGCGATAATGACACCACTTTTTATCCTTTTTAGCCCGGTTTTAGTCCCT

pOLEO8-MC    GCTGCCCTCACTATAGGGCTAGCAGTGGCTGGAATATTGACAGCAGATGC
Clone8fwd    GCTGCCCTCACTATAGGGCTAGCAGTGGCTGGAATATTGACAGCAGATGC

pOLEO8-MC    TTGCGGGTTGACGGGGCTTATGTCGTTGTCGTGGACCGTGAAATATGTTA
Clone8fwd    TTGCGGGTTGACGGGGCTTATGTCGTTGTCGTGGACCGTGAAATATGTTA

pOLEO8-MC    GGGATTTACAAGCAGTAGTGCCCGAACAAATGGATTCGATGAAGGGACGT
Clone8rev    GGGATTTACAAGCAGTAGTGCCCGAACAAATGGATTCGATGAAGGGACGT

pOLEO8-MC    GTCGCGGATGTCGCGAGTTATGTTGGACAAAAGACTAAGGATGTTGGACA
Clone8rev    GTCGCGGATGTCGCGAGTTATGTTGGACAAAAGACTAAGGATGTTGGACA

pOLEO8-MC    AAAAACTAAAGAGGTTGGACAAGACATACAAACAAAAGCACATGAAGCTA
Clone8rev    AAAAACTAAAGAGGTTGGACAAGACATACAAACAAAAGCACATGAAGCTA

pOLEO8-MC    AGAGATCAACAGAGCTCCACCGCGGTGGCGGCCGCTCTAGAGGTACTAT
Clone8rev    AGAGATCAACAGAGCTCCACCGCGGTGGCGGCCGCTCTAGAGGTACTAT

pOLEO8-MC    GGCACAACCTCAAGTTCAAGTCCACTCAACAACAACACACCGTCAAGAAA
Clone8rev    GGCACAACCTCAAGTTCAAGTCCACTCAACAACAACACACCGTCAAGAAA

pOLEO8-MC    CTGCTACCTACCCATCAACCCAAAACATTCGTAAAGATGTTTACGAAAAT
Clone8rev    CTGCTACCTACCCATCAACCCAAAACATTCGTAAAGATGTTTACGAAAAT

pOLEO8-MC    GTTAACTATCCCGGCCAACGCGGTCGTTATAACGACCGCTATAATGATAG
Clone8rev    GTTAACTATCCCGGCCAACGCGGTCGTTATAACGACCGCTATAATGATAG

pOLEO8-MC    TGGTCGTTATGATGGTGGTATTGCCTCCTTTTTGTCAGAGAGAAGTCCTC
Clone8rev    TGGTCGTTATGATGGTGGTATTGCCTCCTTTTTGTCAGAGAGAAGTCCTC
```

Figure 27

```
pOLEO8-MC    CAGCCTCTCAAATCCTCGCTACCGTTGGAGGATTTTTCATAGGTGGTACT
Clone8rev    CAGCCTCTCAAATCCTCGCTACCGTTGGAGGATTTTTCATAGGTGGTACT

pOLEO8-MC    CTATTTTTATTAGCTAGCATTTCATTTATCGCCAGTCTTATTGGATTGGC
Clone8rev    CTATTTTTATTAGCTAGCATTTCATTTATCGCCAGTCTTATTGGATTGGC

pOLEO8-MC    GATAATGACACCACTTTTTATCCTTTTTAGCCCGGTTTTAGTCCCTGCTG
Clone8rev    GATAATGACACCACTTTTTATCCTTTTTAGCCCGGTTTTAGTCCCTGCTG

pOLEO8-MC    CCCTCACTATAGGGCTAGCAGTGGCTGGAATATTGACAGCAGATGCTTGC
Clone8rev    CCCTCACTATAGGGCTAGCAGTGGCTGGAATATTGACAGCAGATGCTTGC

pOLEO8-MC    GGGTTGACGGGGCTTATGTCGTTGTCGTGGACCGTGAAATATGTTAGGGA
Clone8rev    GGGTTGACGGGGCTTATGTCGTTGTCGTGGACCGTGAAATATGTTAGGGA

pOLEO8-MC    TTTACAAGCAGTAGTGCCCGAACAAATGGATTCGATGAAGGGACGTGTCG
Clone8rev    TTTACAAGCAGTAGTGCCCGAACAAATGGATTCGATGAAGGGACGTGTCG

pOLEO8-MC    CGGATGTCGCGAGTTATGTTGGACAAAAGACTAAGGATGTTGGACAAAAA
Clone8rev    CGGATGTCGCGAGTTATGTTGGACAAAAGACTAAGGATGTTGGACAAAAA

pOLEO8-MC    ACTAAAGAGGTTGGACAAGACATACAAACAAAAGCACATGAAGCTAAGAG
Clone8rev    ACTAAAGAGGTTGGACAAGACATACAAACAAAAGCACATGAAGCTAAGAG

pOLEO8-MC    ATCAACAGGTACTACTAGAACTAGTGGATCCCCCGGGGGTACT
Clone8rev    ATCAACAGGTACTACTAGAACTAGTGGATCCCCCGGGGGTACT

pOLEO8-MC       ATGGCACAACCTCAAGTTCAAGTCCACTCAACAACAACACACCGTCA
Clone8rev       ATGGCACAACCTCAAGTTCAAGTCCACTCAACAACAACACACCGTCA

pOLEO8-MC    AGAAACTGCTACCTACCCATCAACCCAAAACATTCGTAAAGATGTTTACG
Clone8rev    AGAAACTGCTACCTACCCATCAACCCAAAACATTCGTAAAGATGTTTACG

pOLEO8-MC    AAAATGTTAACTATCCCGGCCAACGCGGTCGTTATAACGACCGCTATAAT
Clone8rev    AAAATGTTAACTATCCCGGCCAACGCGGTCGTTATAACGACCGCTATAAT

pOLEO8-MC    GATAGTGGTCGTTATGATGGTGGTATTGCCTCCTTTTTGTCAGAGAGAAG
Clone8rev    GATAGTGGTCGTTATGATGGTGGTATTGCCTCCTTTTTGTCAGAGAGAAG

pOLEO8-MC    TCCTCCAGCCTCTCAAATCCTCGCTACCGTTGGAGGATTTTTCATAGGTG
Clone8rev    TCCTCCAGCCTCTCAAATCCTCGCTACCGTTGGAGGATTTTTCATAGGTG

pOLEO8-MC    GTACTCTATTTTTATTAGCTAGCATTTCATTTATCGCCAGTCTTATTGGA
Clone8rev    GTACTCTATTTTTATTAGCTAGCATTTCATTTATCGCCAGTCTTATTGGA

pOLEO8-MC    TTGGCGATAATGACACCACTTTTTATCCTTTTTAGCCCGGTTTTAGTCCC
Clone8rev    TTGGCGATAATGACACCACTTTTTATCCTTTTTAGCCCGGTTTTAGTCCC

pOLEO8-MC    TGCTGCCCTCACTATAGGGCTAGCAGTGGCTGGAATATTGACAGCAGATG
Clone8rev    TGCTGCCCTCACTATAGGGCTAGCAGTGGCTGGAATATTGACAGCAGATG
```

Figure 27 Continued

```
pOLEO8-MC    CTTGCGGGTTGACGGGGCTTATGTCGTTGTCGTGGACCGTGAAATATGTT
Clone8rev    CTTGCGGGTTGACGGGGCTTATGTCGTTGTCGTGGACCGTGAAATATGTT

pOLEO8-MC    AGGGATTTACAAGCAGTAGTGCCCGAACAAATGGATTCGATGAAGGGACG
Clone8rev    AGGGATTTACAAGCAGTAGTGCCCGAACAAATGGATTCGATGAAGGGACG

pOLEO8-MC    TGTCGCGGATGTCGCGAGTTATGTTGGACAAAAGACTAAGGATGTTGGAC
Clone8rev    TGTCGCGGATGTCGCGAGTTATGTTGGACAAAAGACTAAGGATGTTGGAC

pOLEO8-MC    AAAAAACTAAAGAGGTTGGACAAGACATACAAACAAAAGCACATGAAGCT
Clone8rev    AAAAAACTAAAGAGGTTGGACAAGACATACAAACAAAAGCACATGAAGCT

pOLEO8-MC    AAGAGATCAACAGGTACTCTGCAGGAATTCGATATCAAGCTTATCGATAC
Clone8rev    AAGAGATCAACAGGTACTCTGCAGGAATTCGATATCAAGCTTATCGATAC

pOLEO8-MC    CGTCGACCTCGAGTGA
Clone8rev    CGTCGACCTCGAGTGA
```

Figure 27 continued

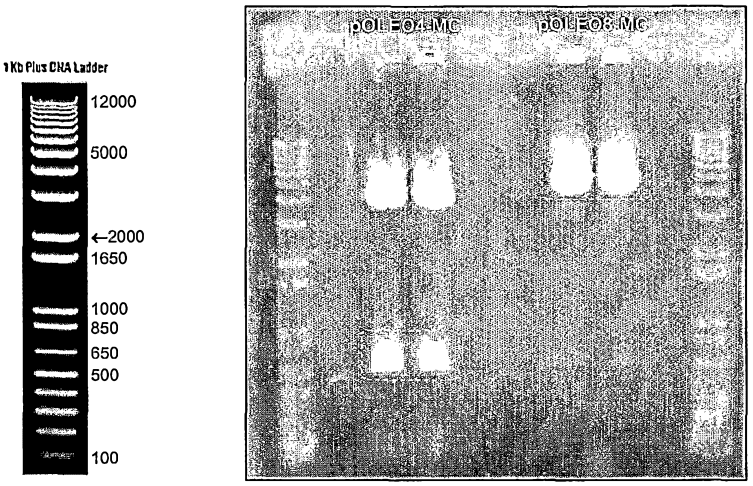

1Kb Plus DNA Ladder
12000
5000
←2000
1650
1000
850
650
500
100
pOL E04 MC
pOL E08 MC

Figure 28

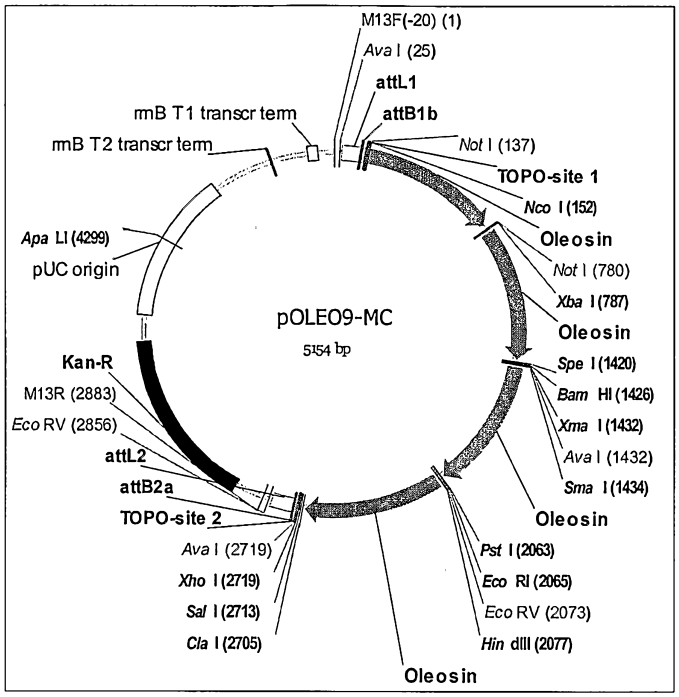

M13F(-20) (1)
Ava I (25)
attL1
attB1b
rrnB T1 transcr term
rrnB T2 transcr term
Not I (137)
TOPO-site 1
Nco I (152)
Oleosin
Not I (780)
Xba I (787)
Oleosin
Spe I (1420)
Bam HI (1426)
Xma I (1432)
Ava I (1432)
Sma I (1434)
Oleosin
Pst I (2063)
Eco RI (2065)
Eco RV (2073)
Hin dIII (2077)
Oleosin
Apa LI (4299)
pUC origin
pOLEO9-MC
5154 bp
Kan-R
M13R (2883)
Eco RV (2856)
attL2
attB2a
TOPO-site 2
Ava I (2719)
Xho I (2719)
Sal I (2713)
Cla I (2705)

Figure 29

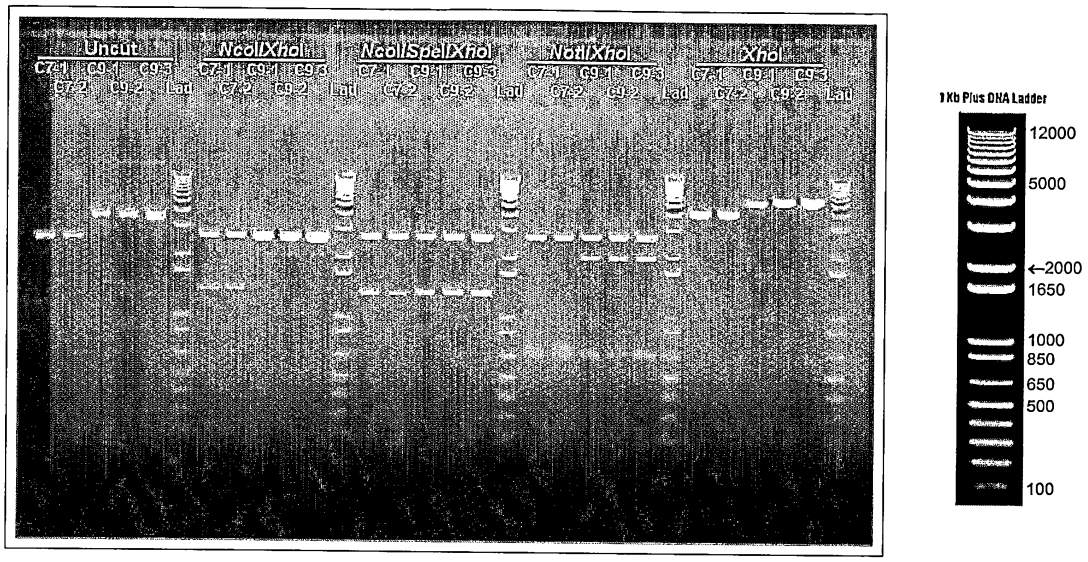

Uncut
NcoI/XhoI
NcoI/SpeI/XhoI
NotI/XhoI
XhoI
C7-1
C9-1
C9-3
C7-2
C9-2
Ladd
1 Kb Plus DNA Ladder
12000
5000
←2000
1650
1000
850
650
500
100

Figure 30

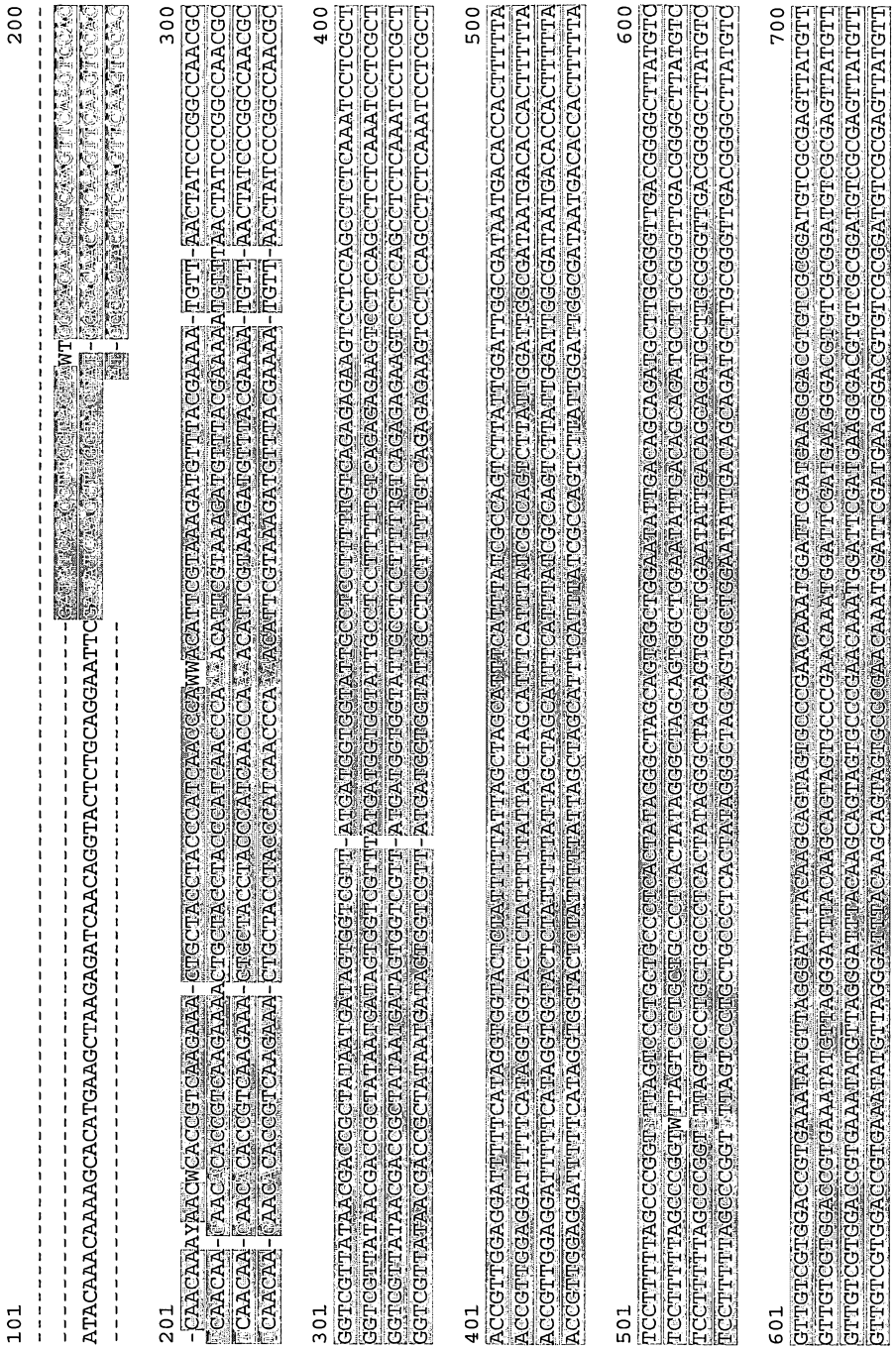


Figure 31

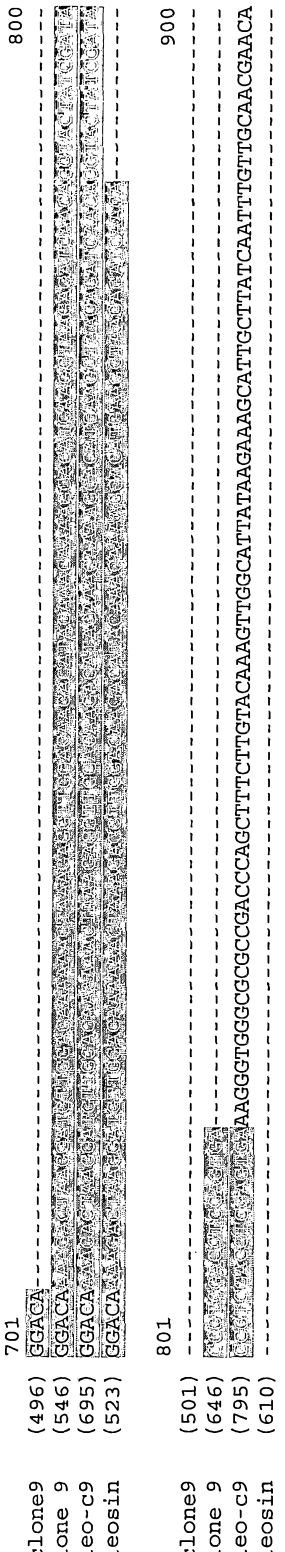


Figure 31 continued

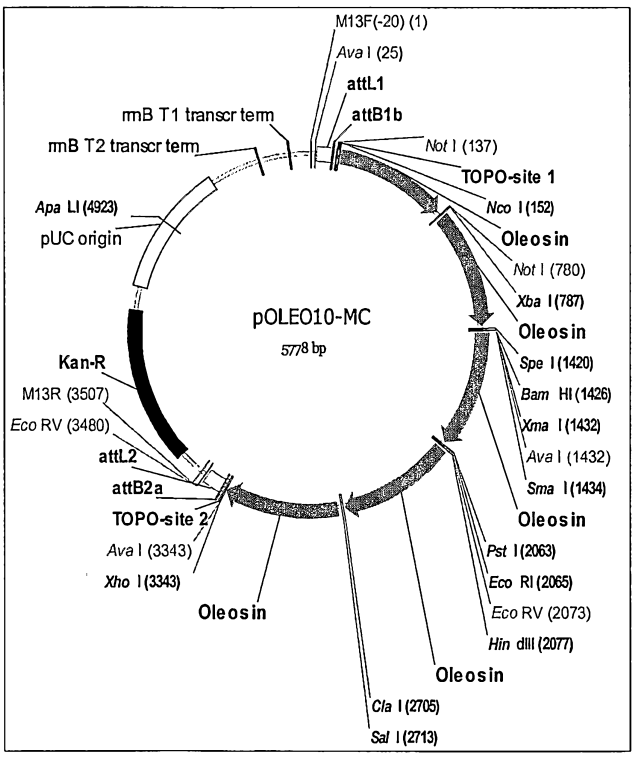
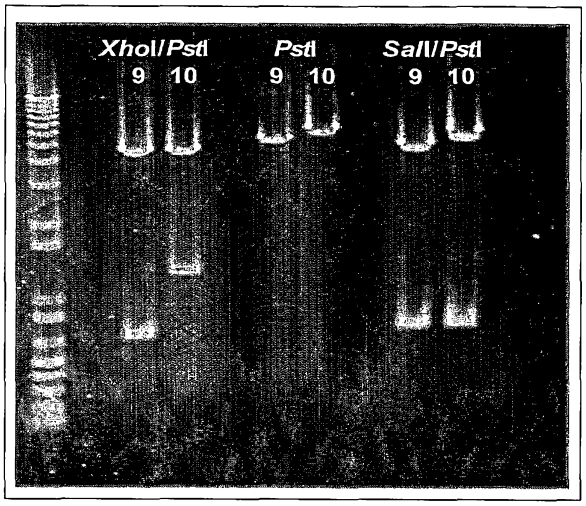

M13F(-20) (1)
Ava I (25)
attL1
attB1b
Not I (137)
TOPO-site 1
Nco I (152)
Oleosin
Not I (780)
Xba I (787)
Oleosin
Spe I (1420)
Bam HI (1426)
Xma I (1432)
Ava I (1432)
Sma I (1434)
Oleosin
Pst I (2063)
Eco RI (2065)
Eco RV (2073)
Hin dIII (2077)
Oleosin
Cla I (2705)
Sal I (2713)
Oleosin
Ava I (3343)
Xho I (3343)
TOPO-site 2
attB2a
attL2
Eco RV (3480)
M13R (3507)
Kan-R
pUC origin
Apa LI (4923)
rrnB T2 transcr term
rrnB T1 transcr term
pOLE010-MC
5778 bp

Figure 32


XhoI/PstI
9 10
PstI
9 10
SalI/PstI
9 10

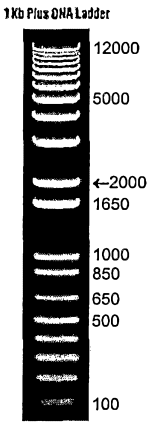

1Kb Plus DNA Ladder
12000
5000
←2000
1650
1000
850
650
500
100

Figure 33

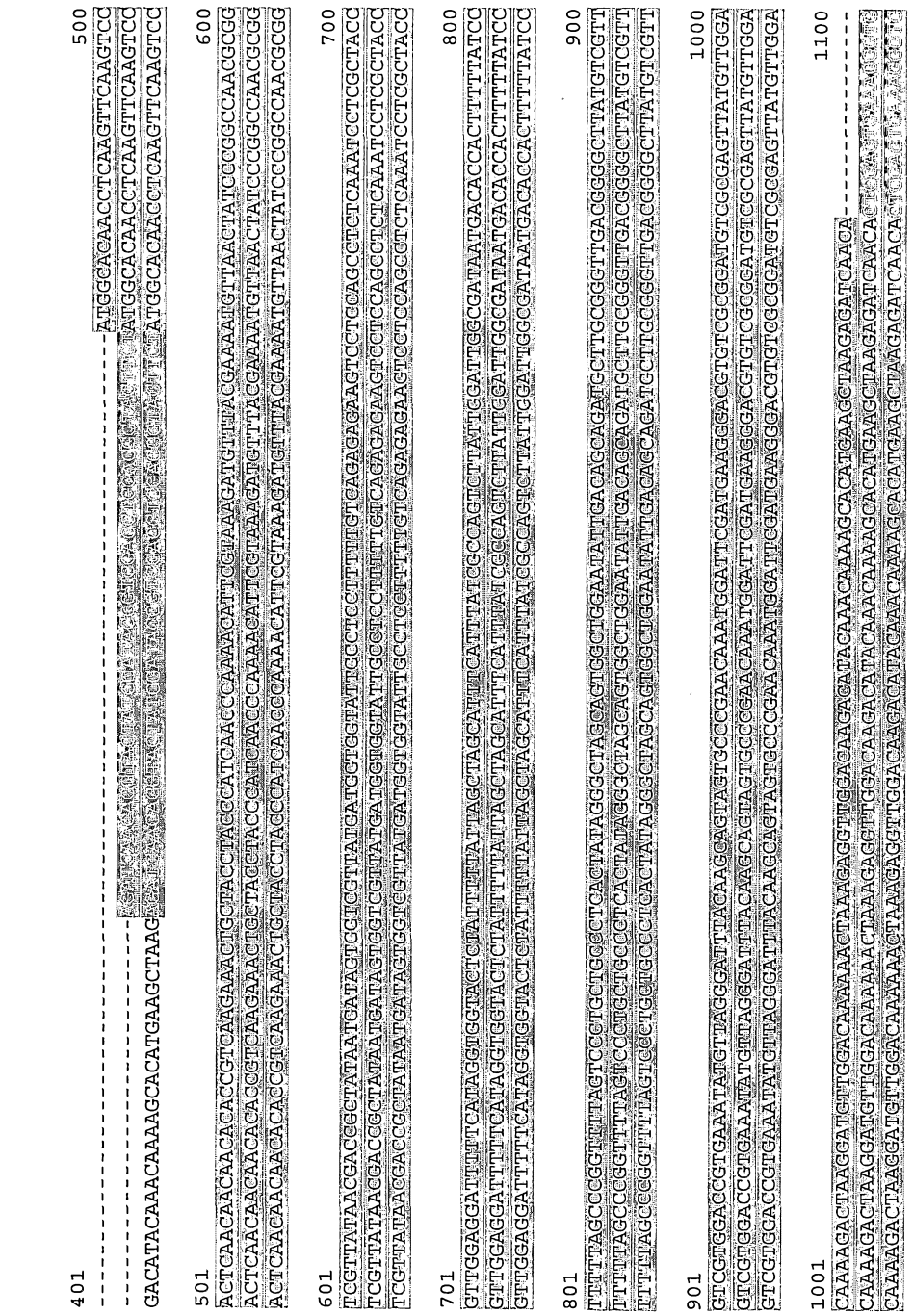


Figure 34

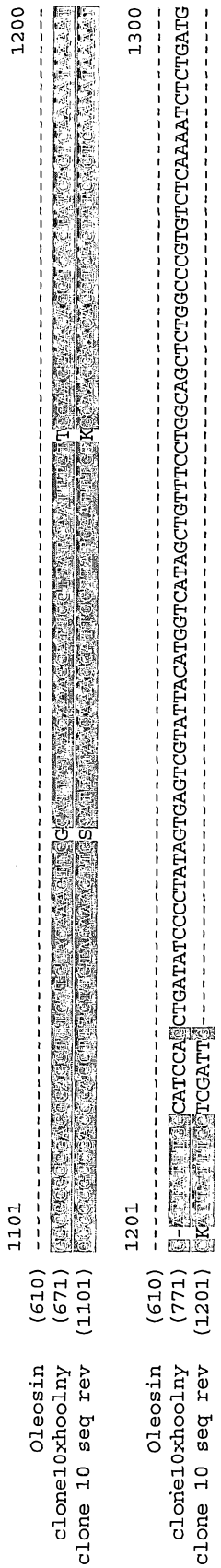


Figure 34 Continued

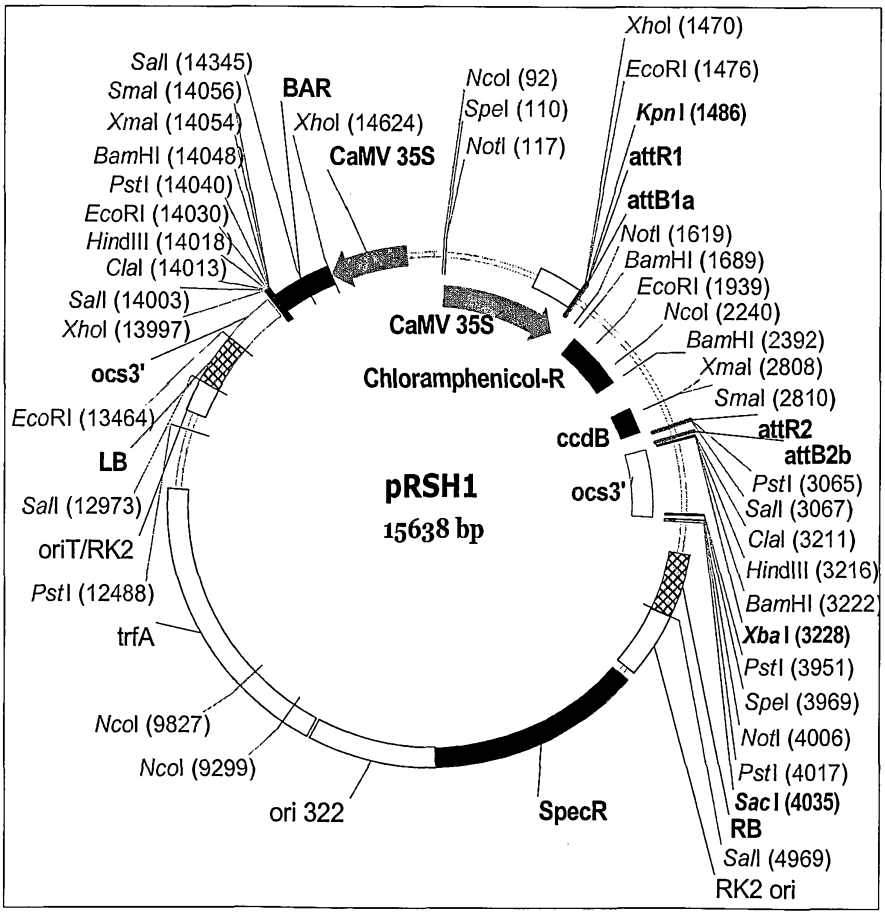

XhoI (1470)
EcoRI (1476)
KpnI (1486)
attR1
attB1a
NotI (1619)
BamHI (1689)
EcoRI (1939)
NcoI (2240)
BamHI (2392)
XmaI (2808)
SmaI (2810)
attR2
attB2b
PstI (3065)
SalI (3067)
ClaI (3211)
HindIII (3216)
BamHI (3222)
XbaI (3228)
PstI (3951)
SpeI (3969)
NotI (4006)
PstI (4017)
SacI (4035)
RB
SalI (4969)
RK2 ori
SalI (14345)
SmaI (14056)
XmaI (14054)
BamHI (14048)
PstI (14040)
EcoRI (14030)
HindIII (14018)
ClaI (14013)
SalI (14003)
XhoI (13997)
ocs3'
EcoRI (13464)
LB
SalI (12973)
oriT/RK2
PstI (12488)
trfA
NcoI (9827)
NcoI (9299)
ori 322
SpecR
BAR
XhoI (14624)
CaMV 35S
NcoI (92)
SpeI (110)
NotI (117)
CaMV 35S
Chloramphenicol-R
ccdB
ocs3'
pRSH1
15638 bp

Figure 35

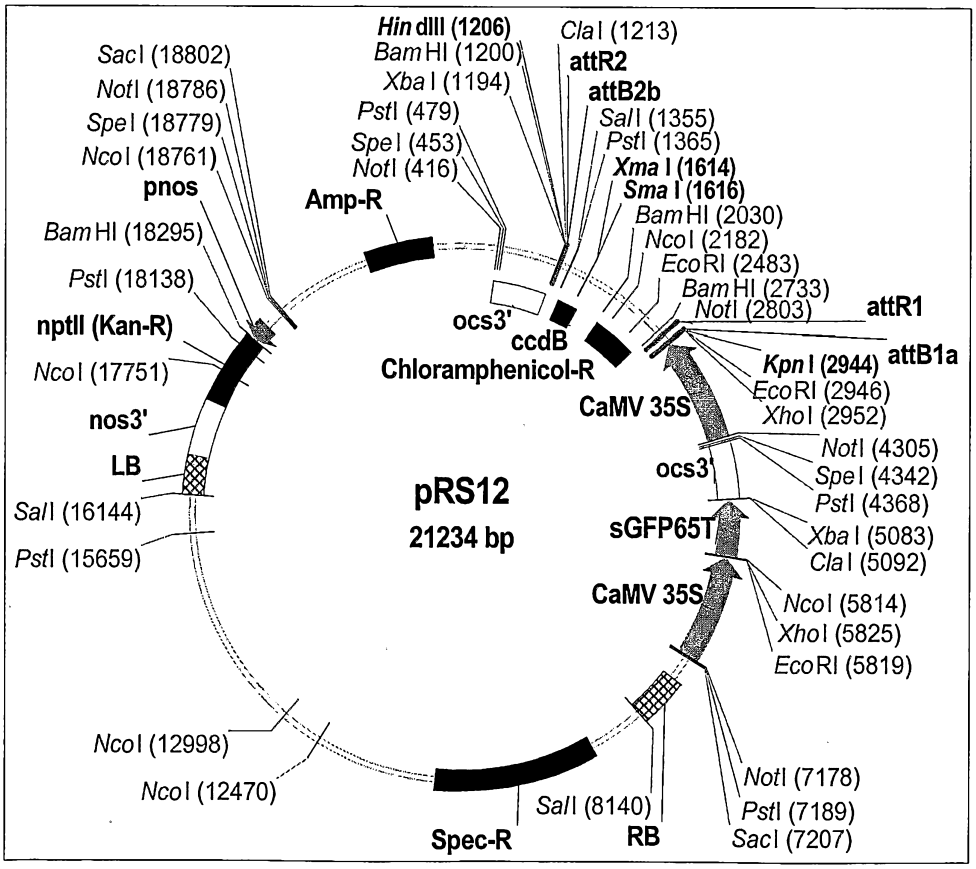

HindIII (1206)
ClaI (1213)
BamHI (1200)
attR2
XbaI (1194)
attB2b
PstI (479)
SalI (1355)
SpeI (453)
PstI (1365)
NotI (416)
XmaI (1614)
SmaI (1616)
SacI (18802)
BamHI (2030)
NotI (18786)
NcoI (2182)
SpeI (18779)
EcoRI (2483)
NcoI (18761)
BamHI (2733)
pnos
NotI (2803)
attR1
Amp-R
BamHI (18295)
ocs3'
PstI (18138)
ccdB
nptII (Kan-R)
Chloramphenicol-R
KpnI (2944)
attB1a
NcoI (17751)
EcoRI (2946)
CaMV 35S
nos3'
XhoI (2952)
NotI (4305)
LB
ocs3'
SpeI (4342)
pRS12
PstI (4368)
SalI (16144)
21234 bp
sGFP65T
XbaI (5083)
PstI (15659)
ClaI (5092)
CaMV 35S
NcoI (5814)
XhoI (5825)
EcoRI (5819)
NcoI (12998)
NotI (7178)
NcoI (12470)
SalI (8140)
PstI (7189)
Spec-R
RB
SacI (7207)

Figure 36

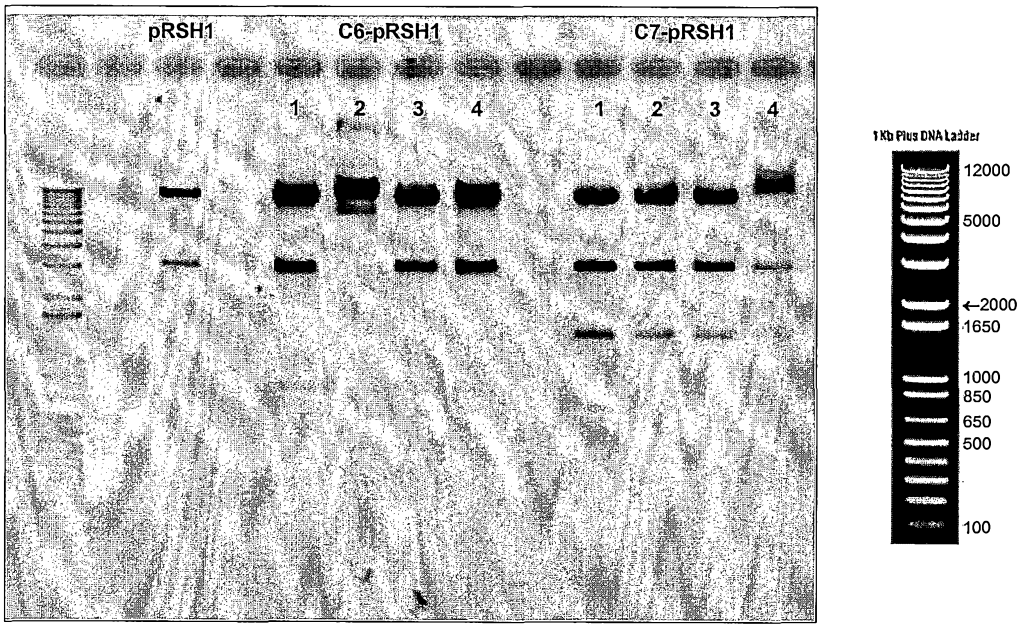

pRSH1
C6-pRSH1
C7-pRSH1
1
2
3
4
1
2
3
4
1Kb Plus DNA Ladder
12000
5000
2000
1650
1000
850
650
500
100

Figure 37

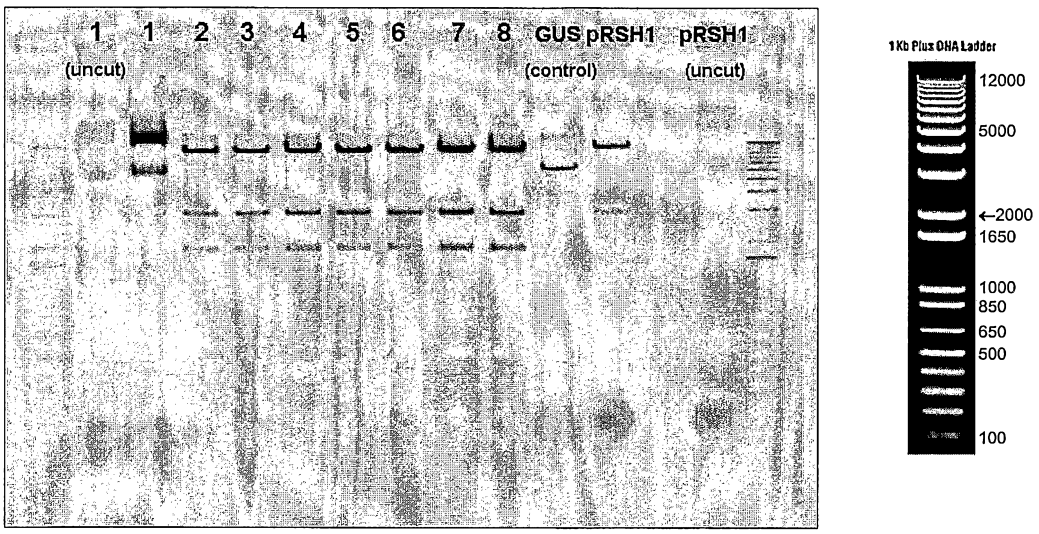

1 (uncut)
1
2
3
4
5
6
7
8
GUS (control)
pRSH1
pRSH1 (uncut)
1Kb Plus DNA Ladder
12000
5000
←2000
1650
1000
850
650
500
100

Figure 38

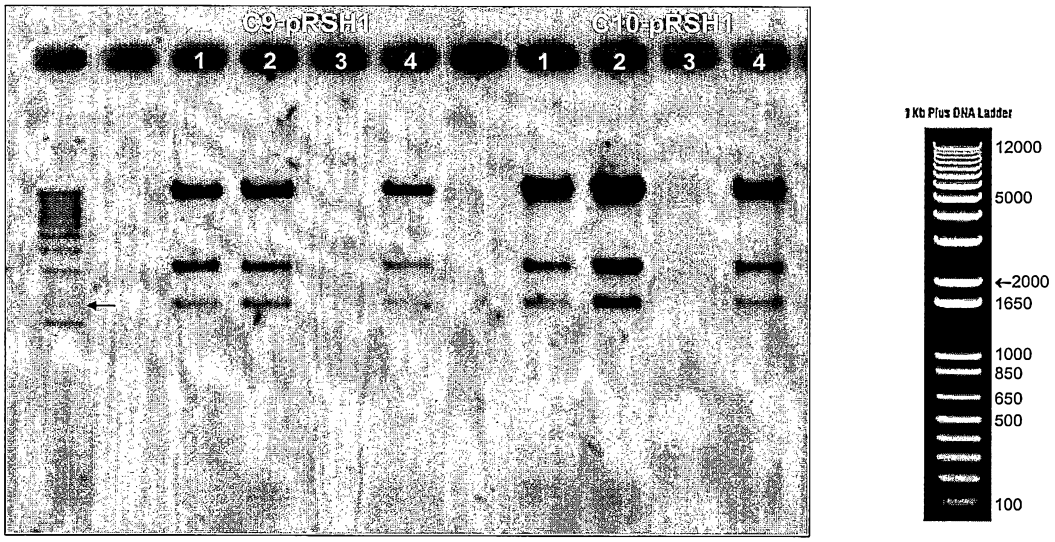

C9-pRSH1
C10-pRSH1
1
2
3
4
1
2
3
4
1 Kb Plus DNA Ladder
12000
5000
←2000
1650
1000
850
650
500
100

Figure 39

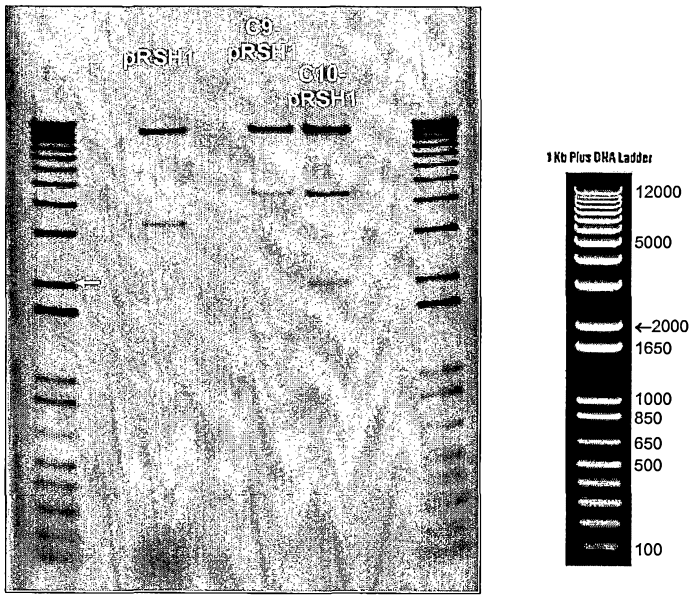

C9-
pRSH1
C10-
pRSH1
1 Kb Plus DNA Ladder
12000
5000
←2000
1650
1000
850
650
500
100

Figure 40

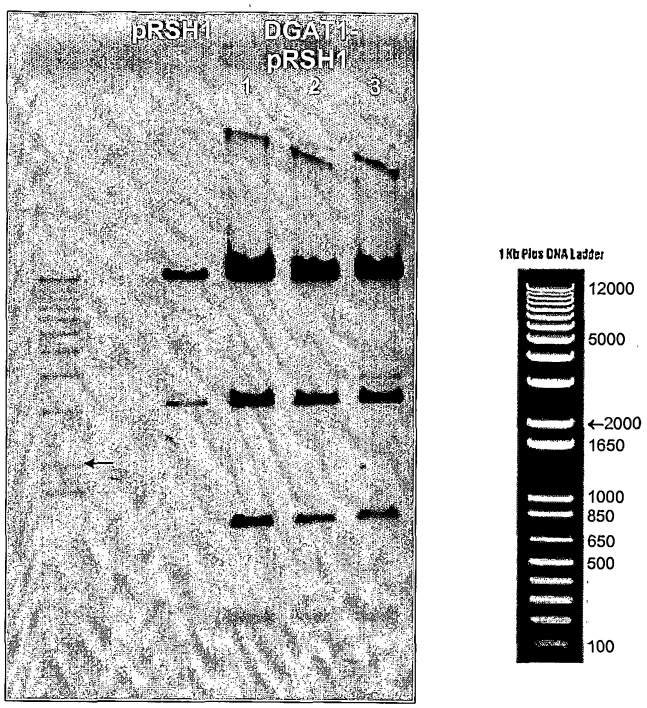

pRSH1
DGAT1-pRSH1
1
2
3
1 Kb Plus DNA Ladder
12000
5000
←2000
1650
1000
850
650
500
100

Figure 41

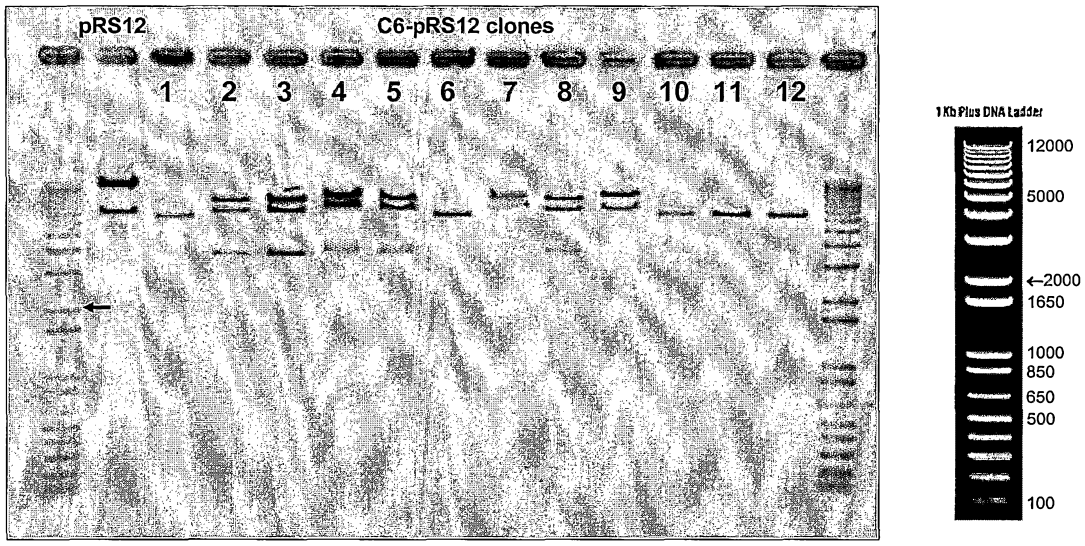

pRS12
C6-pRS12 clones
1
2
3
4
5
6
7
8
9
10
11
12
1 Kb Plus DNA Ladder
12000
5000
←2000
1650
1000
850
650
500
100

Figure 42

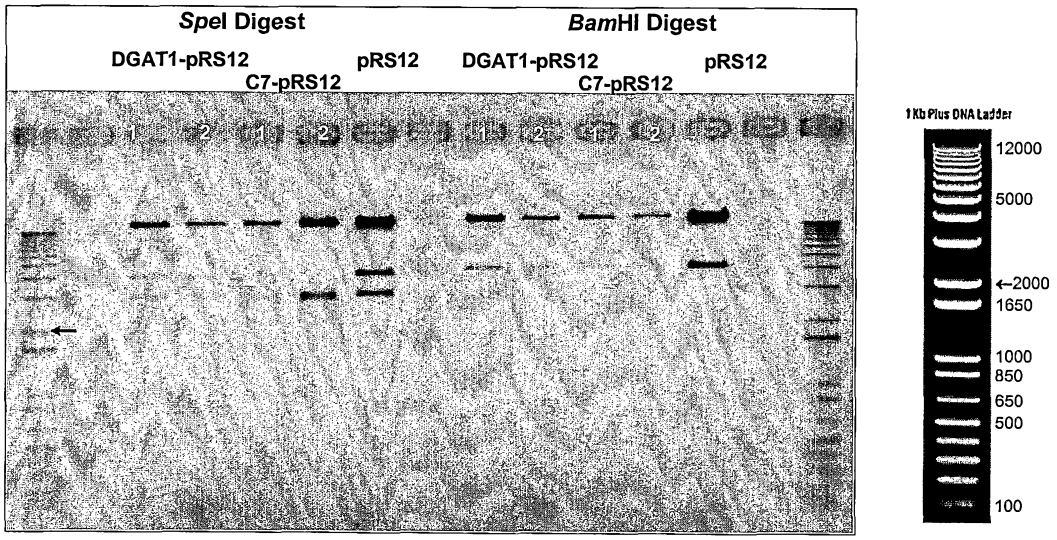

SpeI Digest
BamHI Digest
DGAT1-pRS12
C7-pRS12
pRS12
DGAT1-pRS12
C7-pRS12
pRS12
1 2 1 2 3 1 2 1 2 3
1 Kb Plus DNA Ladder
12000
5000
←2000
1650
1000
850
650
500
100

Figure 43

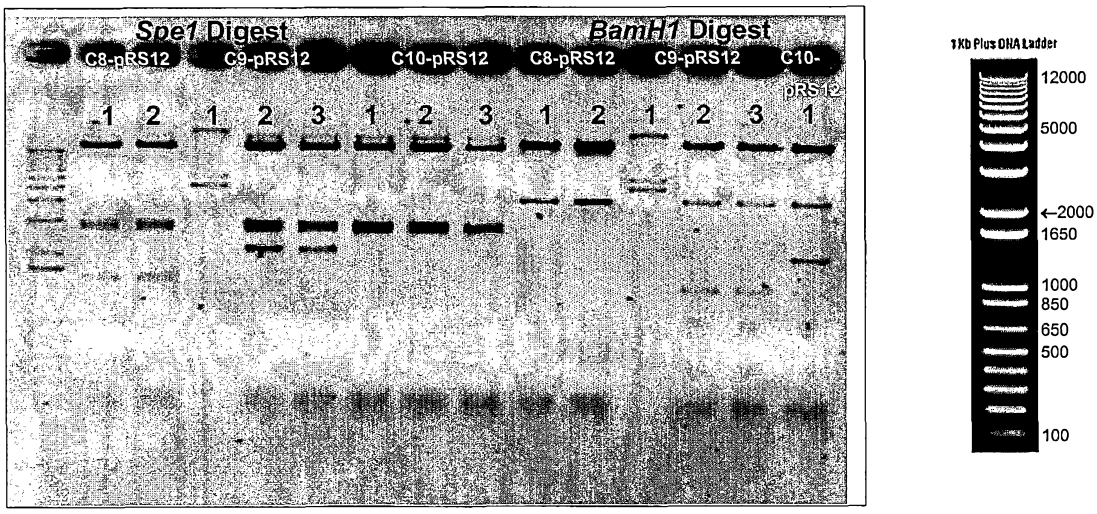

SpeI Digest
BamH1 Digest
C8-pRS12
C9-pRS12
C10-pRS12
C8-pRS12
C9-pRS12
C10-pRS12
1 2 1 2 3 1 2 3 1 2 1 2 3 1
1Kb Plus DNA Ladder
12000
5000
←2000
1650
1000
850
650
500
100

Figure 44

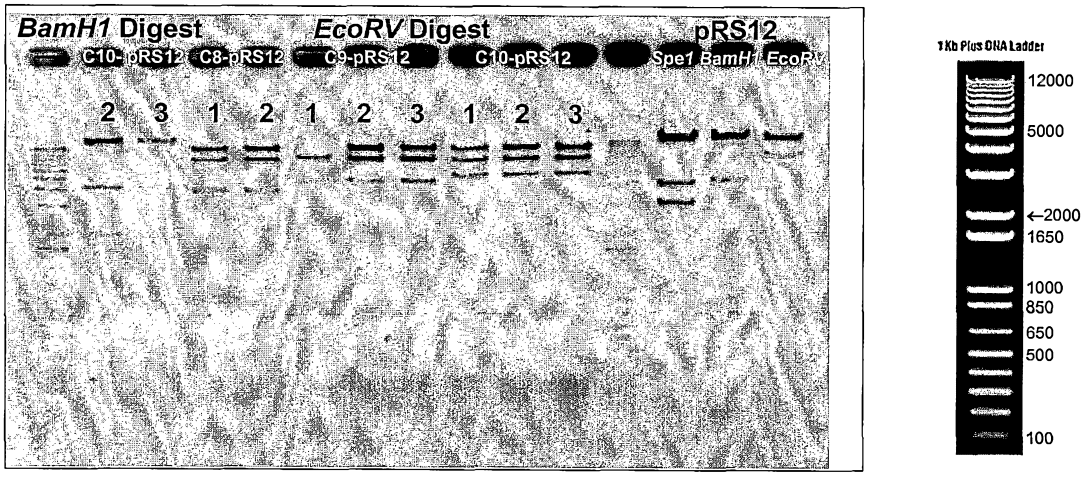

BamH1 Digest
EcoRV Digest
pRS12
C10-pRS12
C8-pRS12
C9-pRS12
C10-pRS12
Spe1 BamH1 EcoR1
2 3 1 2 1 2 3 1 2 3
1Kb Plus DNA Ladder
12000
5000
←2000
1650
1000
850
650
500
100

Figure 45

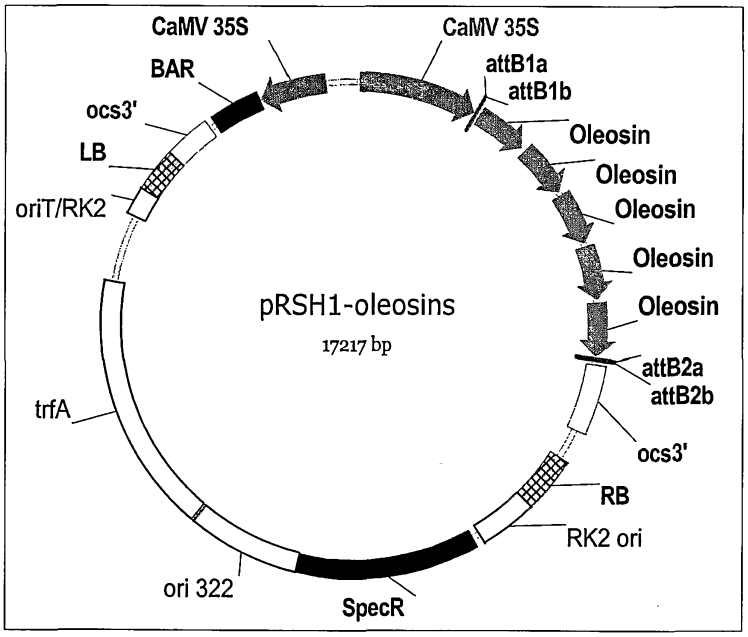

CaMV 35S
CaMV 35S
BAR
attB1a
attB1b
ocs3'
Oleosin
LB
Oleosin
oriT/RK2
Oleosin
Oleosin
pRSH1-oleosins
17217 bp
Oleosin
attB2a
attB2b
trfA
ocs3'
RB
RK2 ori
ori 322
SpecR

Figure 46

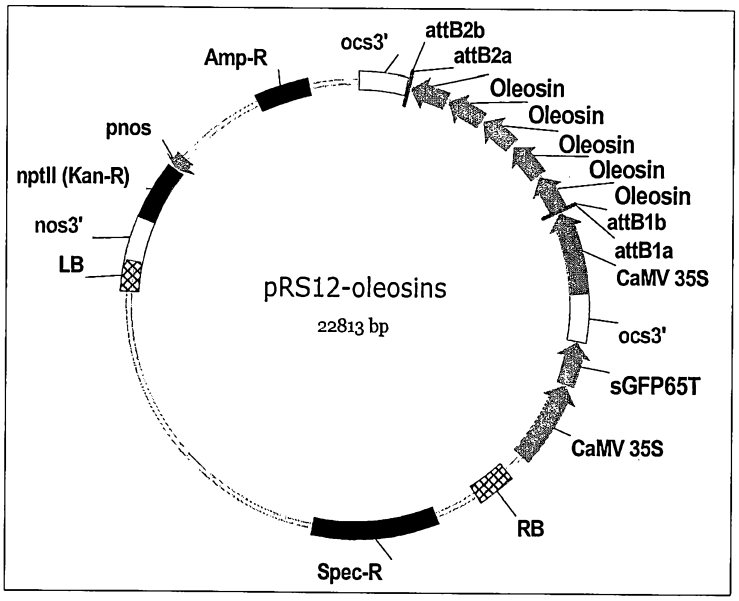

Amp-R
ocs3'
attB2b
attB2a
Oleosin
Oleosin
Oleosin
Oleosin
Oleosin
attB1b
attB1a
CaMV 35S
ocs3'
sGFP65T
CaMV 35S
RB
Spec-R
LB
nos3'
nptII (Kan-R)
pnos
pRS12-oleosins
22813 bp

Figure 47

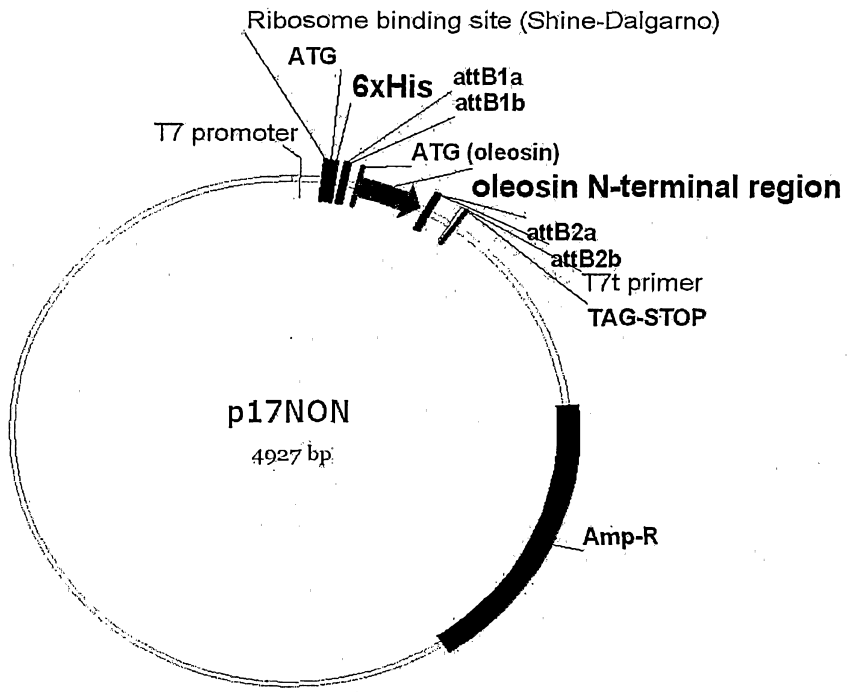

Ribosome binding site (Shine-Dalgarno)
ATG
6xHis
attB1a
attB1b
T7 promoter
ATG (oleosin)
oleosin N-terminal region
attB2a
attB2b
T7t primer
TAG-STOP
p17NON
4927 bp
Amp-R

Figure 48

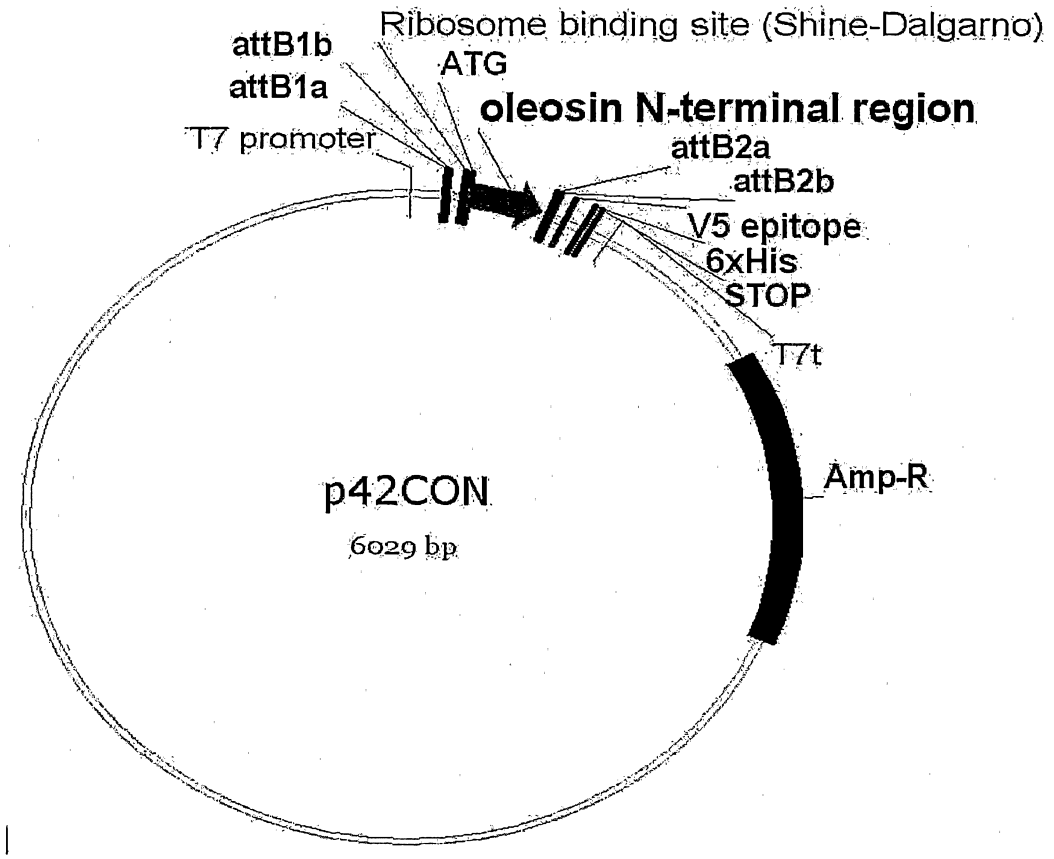

Ribosome binding site (Shine-Dalgarno)
attB1b
attB1a
ATG
oleosin N-terminal region
T7 promoter
attB2a
attB2b
V5 epitope
6xHis
STOP
T7t
Amp-R
p42CON
6029 bp

Figure 49

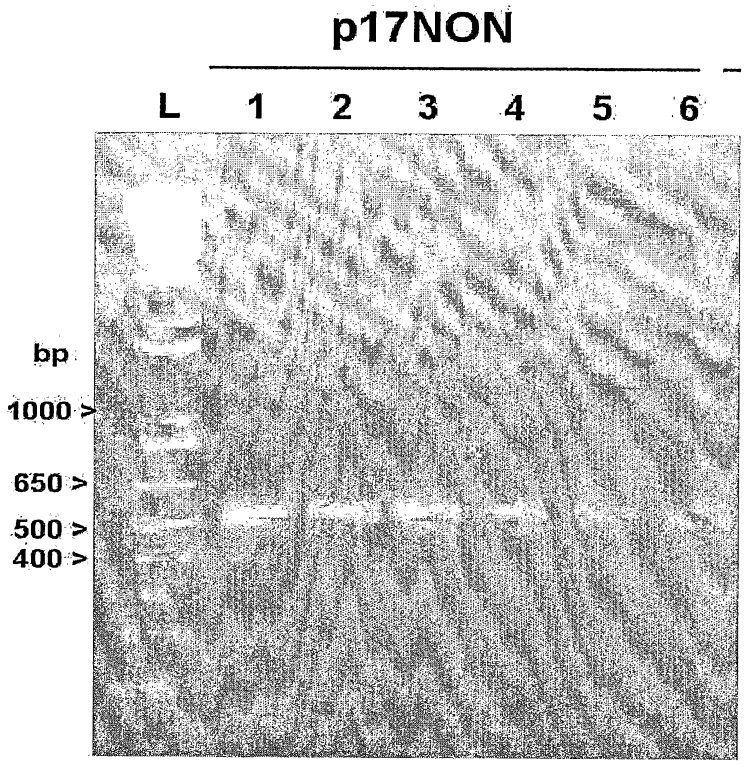


Figure 50

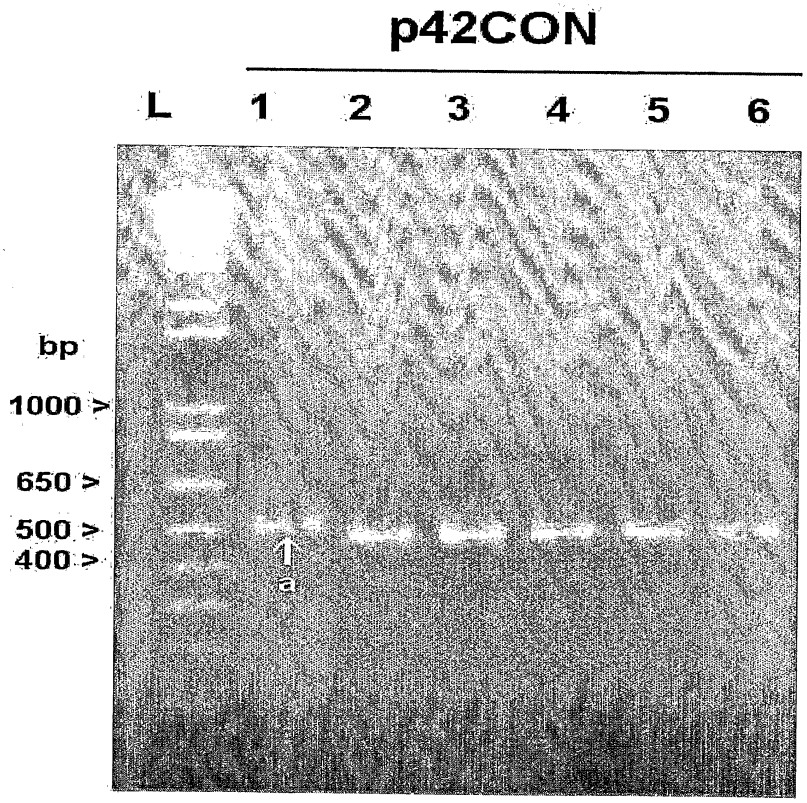


Figure 51

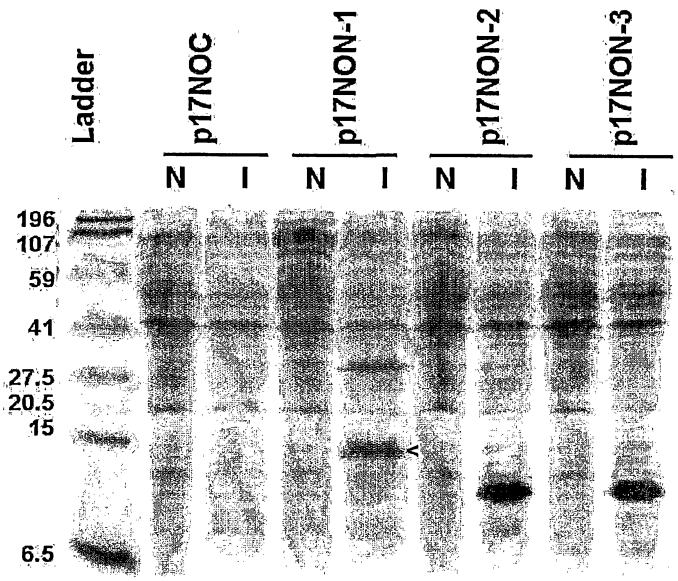

Ladder
p17NOC
p17NON-1
p17NON-2
p17NON-3
N
I
N
I
N
I
N
I
196
107
59
41
27.5
20.5
15
6.5

Figure 52

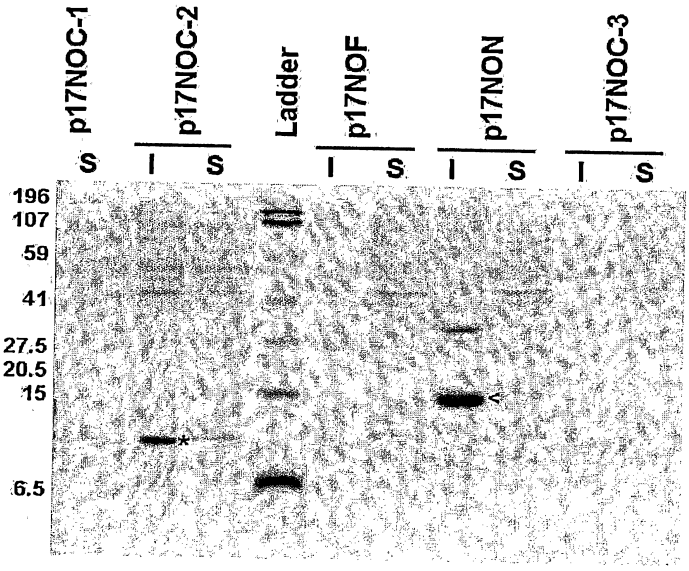

p17NOC-1
p17NOC-2
Ladder
p17NOF
p17NON
p17NOC-3
S
I
S
I
S
I
S
I
S
196
107
59
41
27.5
20.5
15
6.5

Figure 53

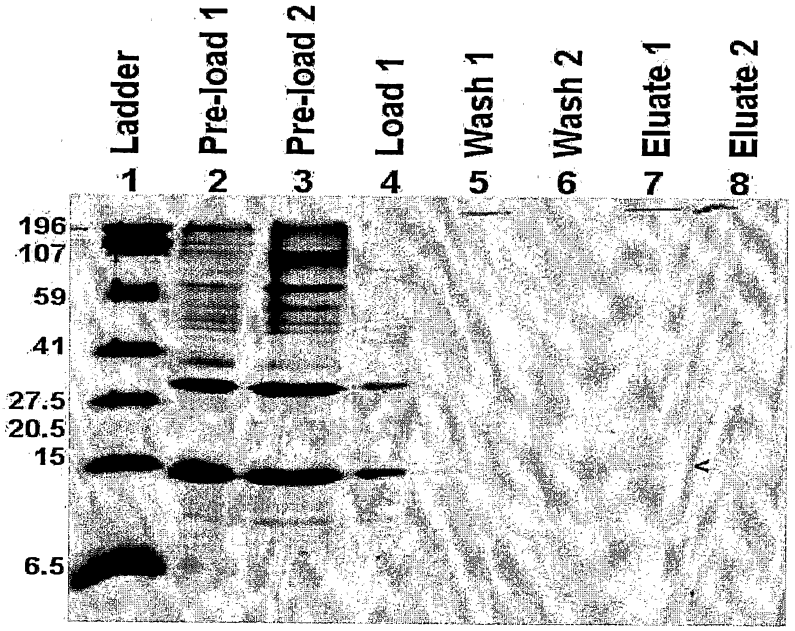

Ladder
Pre-load 1
Pre-load 2
Load 1
Wash 1
Wash 2
Eluate 1
Eluate 2
1
2
3
4
5
6
7
8
196
107
59
41
27.5
20.5
15
6.5

Figure 54

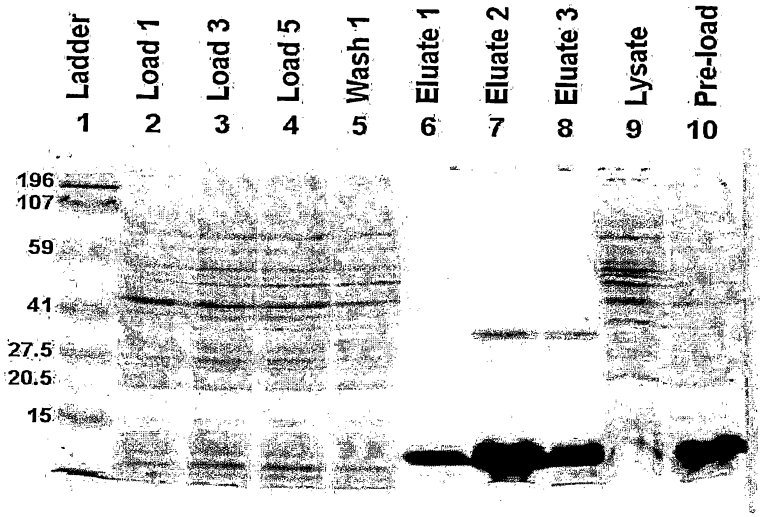

Ladder
Load 1
Load 3
Load 5
Wash 1
Eluate 1
Eluate 2
Eluate 3
Lysate
Pre-load
1
2
3
4
5
6
7
8
9
10
196
107
59
41
27.5
20.5
15

Figure 55

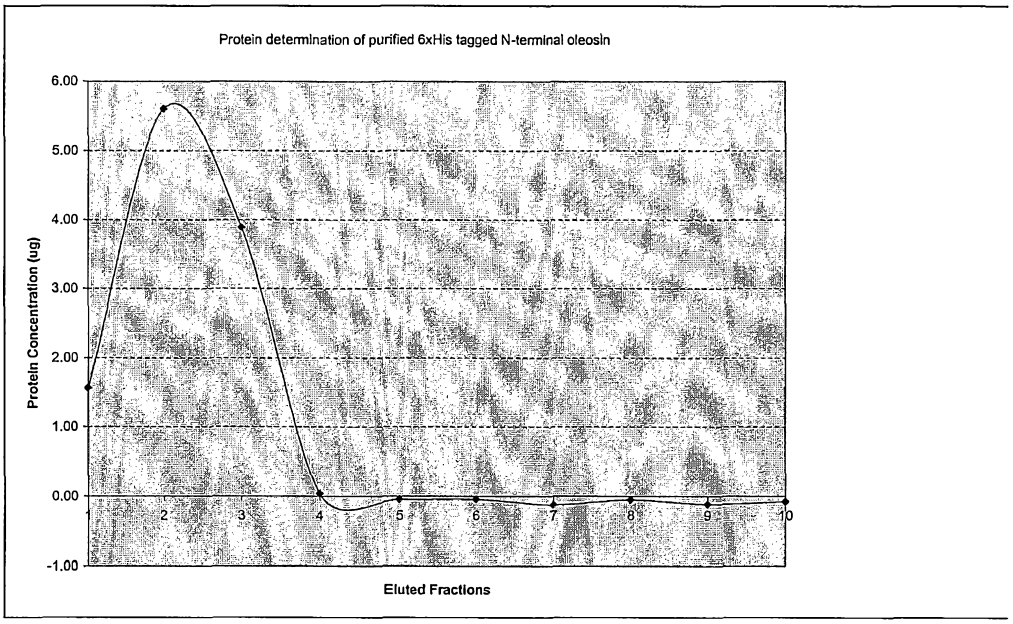

Protein determination of purified 6xHis tagged N-terminal oleosin
Protein Concentration (ug)
6.00
5.00
4.00
3.00
2.00
1.00
0.00
-1.00
1
2
3
4
5
6
7
8
9
10
Eluted Fractions

Figure 56

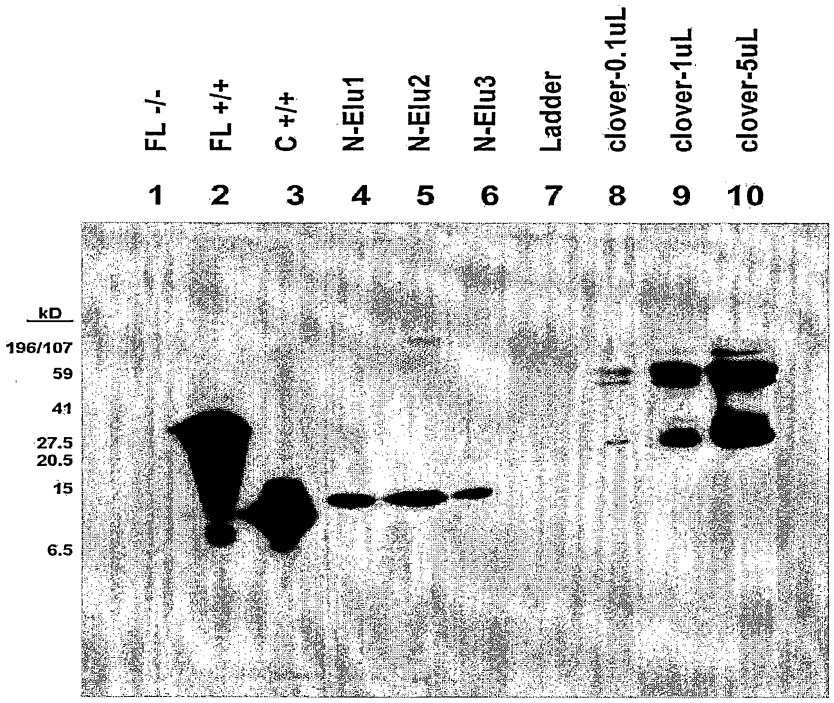

FL -/-
FL +/+
C +/+
N-Elu1
N-Elu2
N-Elu3
Ladder
clover-0.1uL
clover-1uL
clover-5uL
1 2 3 4 5 6 7 8 9 10
kD
196/107
59
41
27.5
20.5
15
6.5

Figure 57

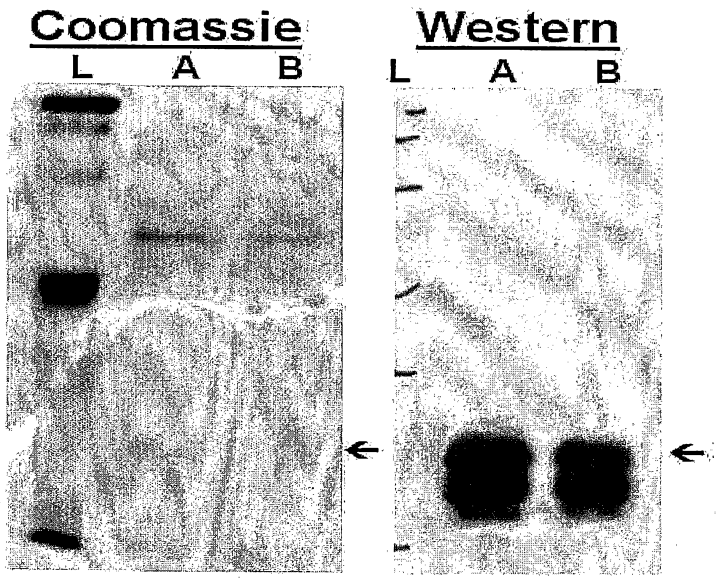


Figure 58

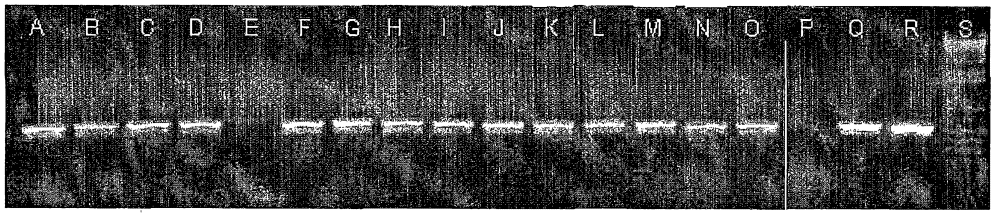


Figure 59

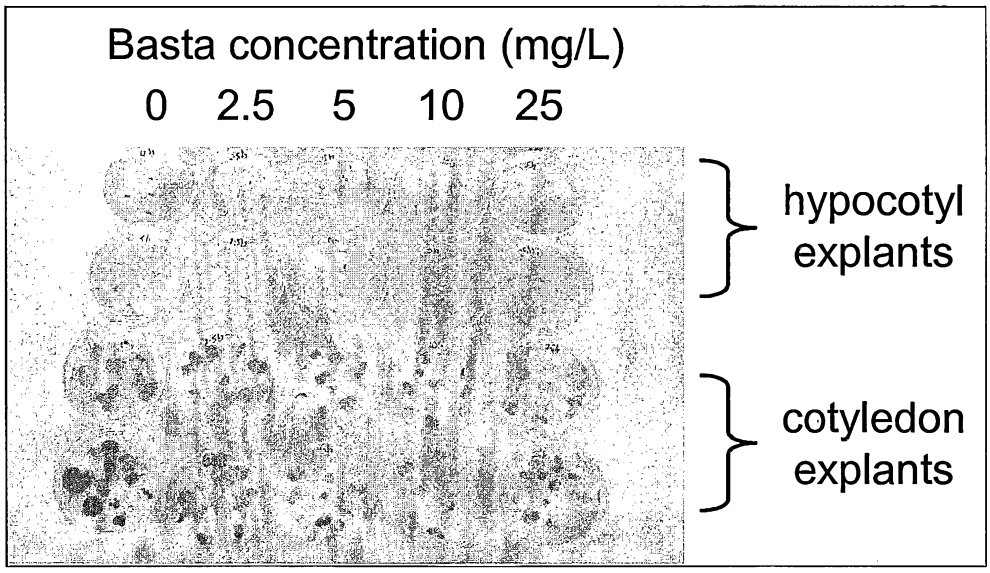

Basta concentration (mg/L)
0 2.5 5 10 25
hypocotyl explants
cotyledon explants

Figure 60

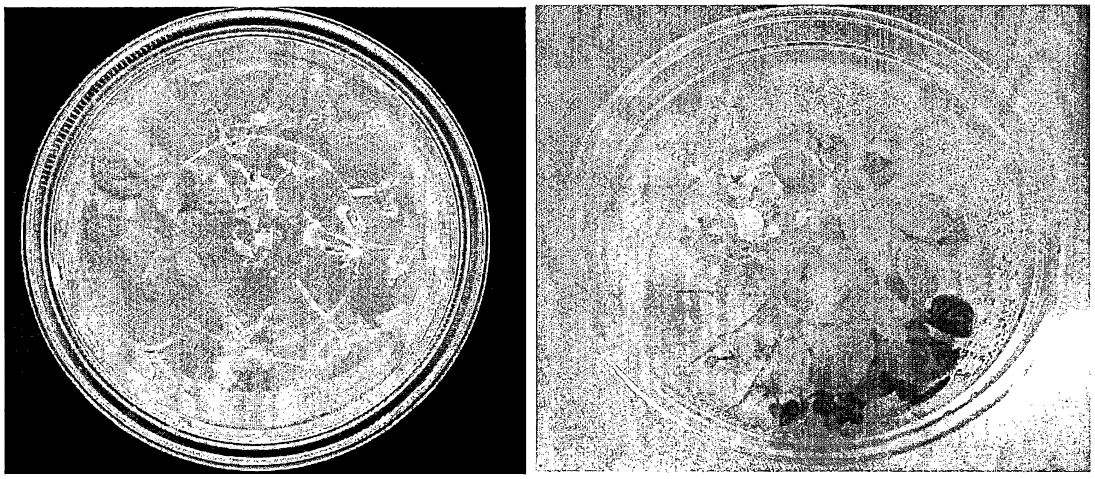


Figure 61

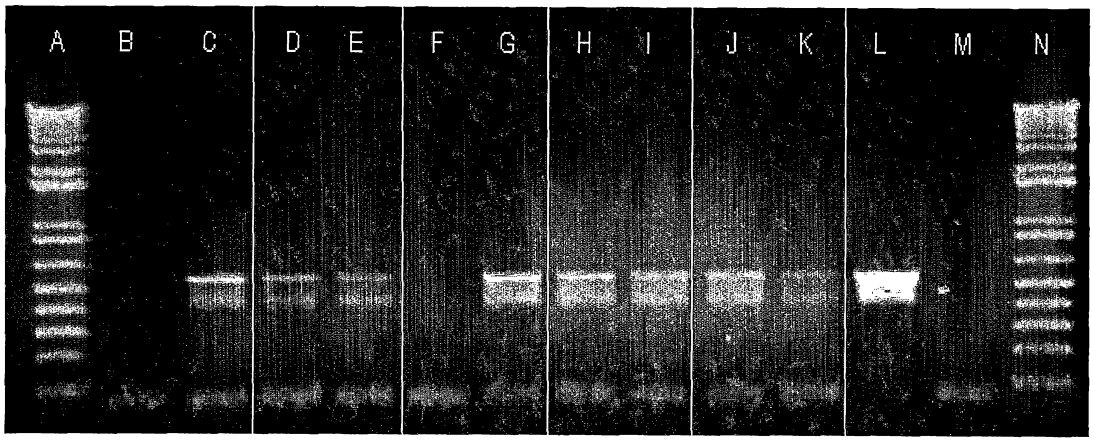


Figure 62



Figure 63

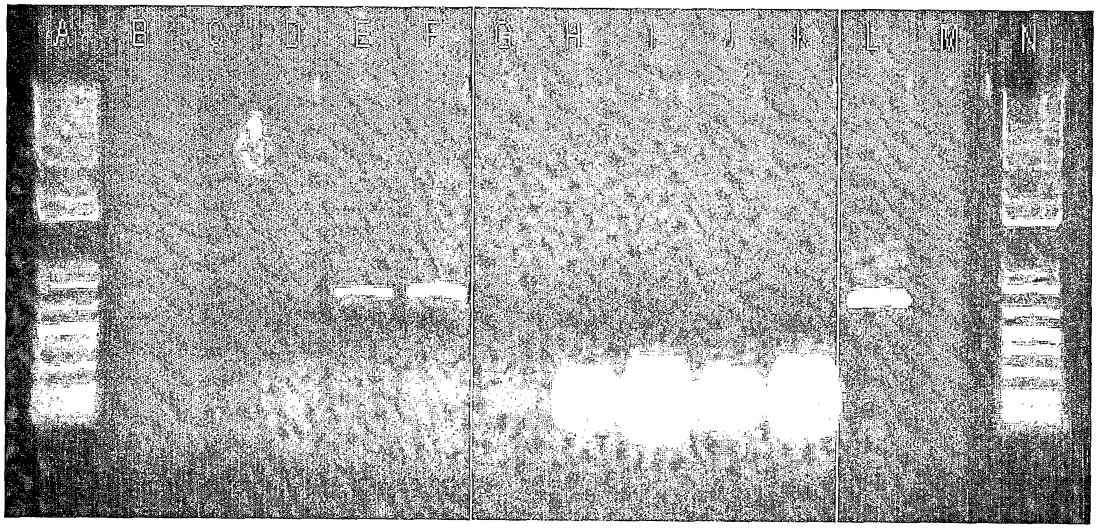


Figure 64

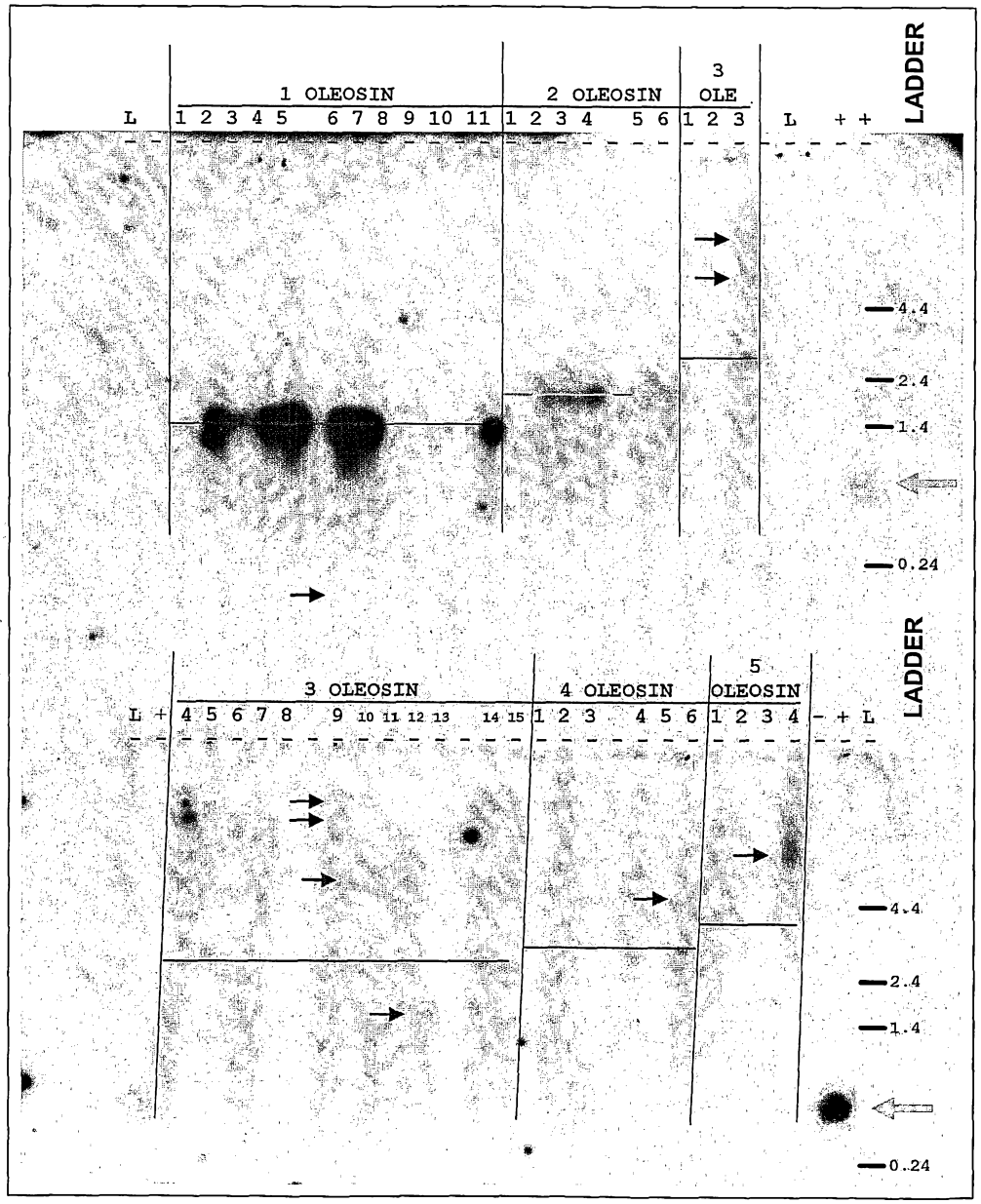

L
1 OLEOSIN
1 2 3 4 5 6 7 8 9 10 11
2 OLEOSIN
1 2 3 4 5 6
3 OLE
1 2 3
L + + LADDER
4.4
2.4
1.4
0.24
3 OLEOSIN
L + 4 5 6 7 8 9 10 11 12 13 14 15
4 OLEOSIN
1 2 3 4 5 6
5 OLEOSIN
1 2 3 4
- + L LADDER
4.4
2.4
1.4
0.24

Figure 65

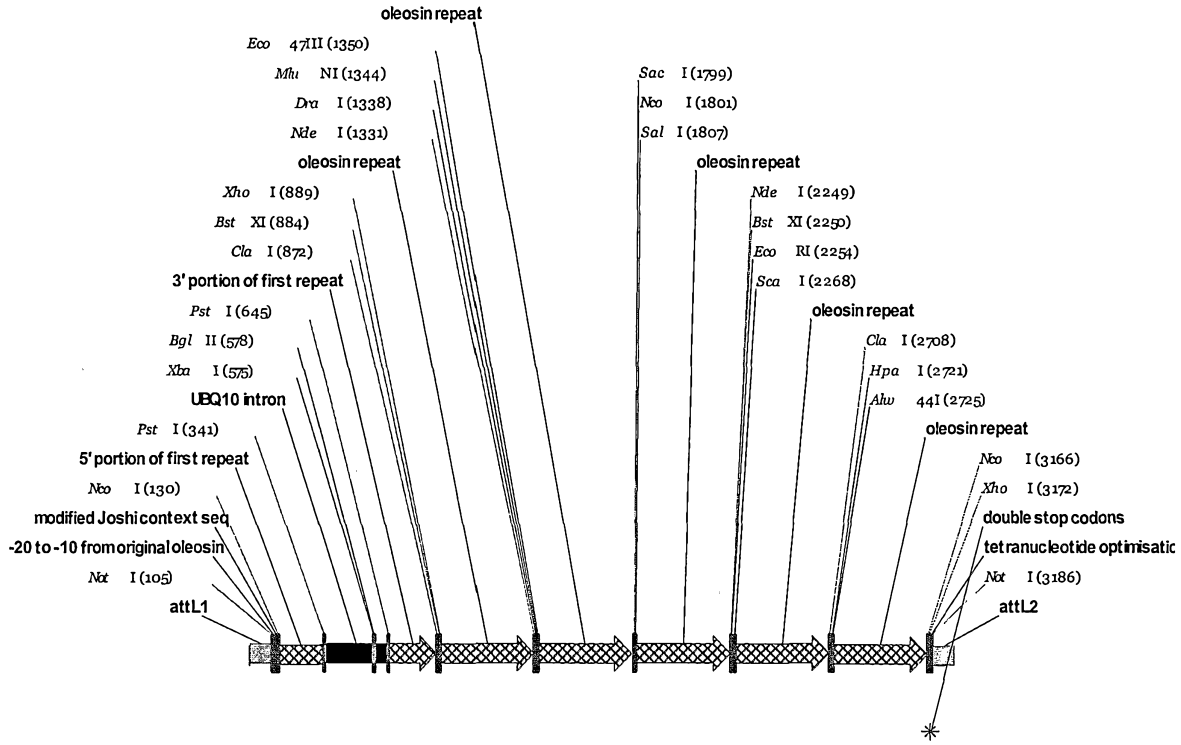

oleosin repeat
Eco 47II (1350)
Mlu NI (1344)
Dra I (1338)
Nde I (1331)
oleosin repeat
Xho I (889)
Bst XI (884)
Cla I (872)
3' portion of first repeat
Pst I (645)
Bgl II (578)
Xba I (575)
UBQ10 intron
Pst I (341)
5' portion of first repeat
Nco I (130)
modified Joshi context seq
-20 to -10 from original oleosin
Not I (105)
attL1
Sac I (1799)
Nco I (1801)
Sal I (1807)
oleosin repeat
Nde I (2249)
Bst XI (2250)
Eco RI (2254)
Sca I (2268)
oleosin repeat
Cla I (2708)
Hpa I (2721)
Alw 44I (2725)
oleosin repeat
Nco I (3166)
Xho I (3172)
double stop codons
tetranucleotide optimisatic
Not I (3186)
attL2

Figure 66

A

```
   1  GTA AAT TTC TGT GTT CCT TAT TCT CTC AAA ATC TTC GAT TTT GTT TTC GTT CGA TCC CAA TTT CGT ATA TGT TCT TTG GTT TAG ATT CTG
  91  TTA ATC TTA GAT CGA AGA CGA TTT TCT GGG TTT GAT CGT TAG ATA TCA TCT TAA TTC TCG ATT AGG GTT TCA TAG ATA TCA TCC GAT TTG

                                                    XbaI
                                                   BglII
 181  TTC AAA TAA TTT GAG TTT TGT CGA ATA ATT ACT CTT CGA TTT GTG ATT TCT ATC TAG ATC TGG TGT TAG TTT CTA GTT TGT GCG ATC GAA

                                          PstI
 271  TTT GTC GAT TAA TCT GAG TTT TTC TGA TCT GCA G
```

B

```
                                                                                                                    1
Norris et al UBQ10 intron                                                                                     (1)  GTAAATTTCTGTGTTCCTTATTCTCTCAAAATCTTCGATTTTGTTTTCGTTCGATCCCAATTTCGTATGTTCTTTGGTTTAGATTCTGTTAATCTTTAG
  UBQ10 for polyoleosin                                                                                       (1)  GTAAATTTCTGTGTTCCTTATTCTCTCAAAATCTTCGATTTTGTTTTCGTTCGATCCCAATTTCGTATGTTCTTTGGTTTAGATTCTGTTAATCTTTAG
        Consensus                                                                                             (1)  GTAAATTTCTGTGTTCCTTATTCTCTCAAAATCTTCGATTTTGTTTTCGTTCGATCCCAATTTCGTATGTTCTTTGGTTTAGATTCTGTTAATCTTTAG

                                                                                                                  101
Norris et al UBQ10 intron                                                                                     (101)  ATCGAAGACGATTTTCTGGGTTTGATCGTTAGATATCATCTTAATTCTCGATTAGGGTTTCATAGATATCATCCGATTTGTTCAAATAATTTGAGTTTTG
  UBQ10 for polyoleosin                                                                                       (101)  ATCGAAGACGATTTTCTGGGTTTGATCGTTAGATATCATCTTAATTCTCGATTAGGGTTTCATAGATATCATCCGATTTGTTCAAATAATTTGAGTTTTG
        Consensus                                                                                             (101)  ATCGAAGACGATTTTCTGGGTTTGATCGTTAGATATCATCTTAATTCTCGATTAGGGTTTCATAGATATCATCCGATTTGTTCAAATAATTTGAGTTTTG

                                                                                                                  201
Norris et al UBQ10 intron                                                                                     (201)  TCGAATAATTACTCTTCGATTTGTGATTTCTATCTAGATCTGGTGTTAGTTTCTAGTTTGTGCGATCGAATTTGTCGATTAATCTGAGTTTTTCTGATTA
  UBQ10 for polyoleosin                                                                                       (201)  TCGAATAATTACTCTTCGATTTGTGATTTCTATCTAGATCTGGTGTTAGTTTCTAGTTTGTGCGATCGAATTTGTCGATTAATCTGAGTTTTTCTGATCT
        Consensus                                                                                             (201)  TCGAATAATTACTCTTCGATTTGTGATTTCTATCTAGATCTGGTGTTAGTTTCTAGTTTGTGCGATCGAATTTGTCGATTAATCTGAGTTTTTCTGAT

                                                                                                                  301
Norris et al UBQ10 intron                                                                                     (301)  ACAG
  UBQ10 for polyoleosin                                                                                       (301)  GCAG
        Consensus                                                                                             (301)  CAG
```

Figure 67

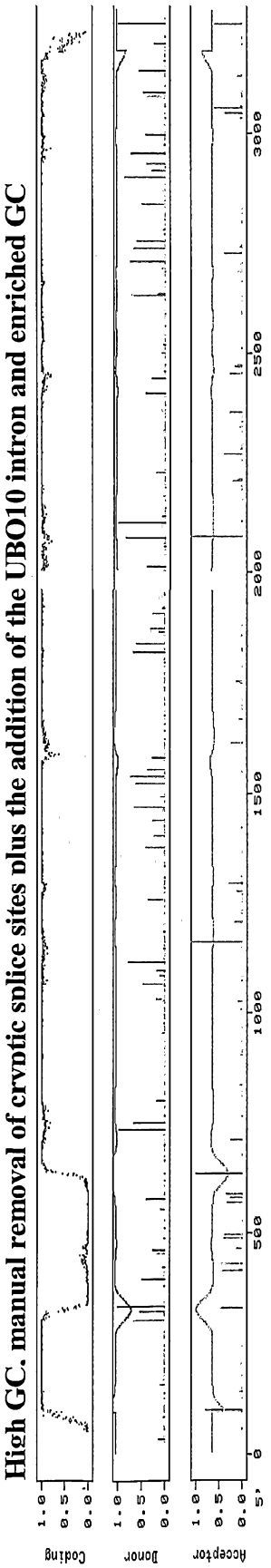

High GC, manual removal of cryptic splice sites plus the addition of the UBO10 intron and enriched GC
Coding
Donor
Acceptor
1.0
0.5
0.0
0
500
1000
1500
2000
2500
3000
5'

Figure 68

```
  MAEHYGQQQQTRAPHLQLQPRAQRVVKAATAVTAGGSLLVLSGLTLAGTVIALTIATPLLVIFSPVLVPAVITIF
Translation of High GC oleosins spliced (1)
MAEHYGQQQQTRAPHLQLQPRAQRVVKAATAVTAGGSLLVLSGLTLAGTVIALTIATPLLVIFSPVLVPAVITIF
                                Consensus (1)
MAEHYGQQQQTRAPHLQLQPRAQRVVKAATAVTAGGSLLVLSGLTLAGTVIALTIATPLLVIFSPVLVPAVITIF
                                          76
150
          6x OLEOSIN AF091840 ORF only (76)
LLGAGFLASGGFGVAALSVLSWIYRYLTGKHPPGADQLESAKTKLASKAREMKDRAEQFSQQPVAGSQTS-----
Translation of High GC oleosins spliced (76)
LLGAGFLASGGFGVAALSVLSWIYRYLTGKHPPGADQLESAKTKLASKAREMKDRAEQFSQQPVAGSQTSIDPSS
                                Consensus (76)
LLGAGFLASGGFGVAALSVLSWIYRYLTGKHPPGADQLESAKTKLASKAREMKDRAEQFSQQPVAGSQTS
                                          151
225
          6x OLEOSIN AF091840 ORF only (145)
---MAEHYGQQQQTRAPHLQLQPRAQRVVKAATAVTAGGSLLVLSGLTLAGTVIALTIATPLLVIFSPVLVPAVI
Translation of High GC oleosins spliced (151)
WLEMAEHYGQQQQTRAPHLQLQPRAQRVVKAATAVTAGGSLLVLSGLTLAGTVIALTIATPLLVIFSPVLVPAVI
                                Consensus (151)
MAEHYGQQQQTRAPHLQLQPRAQRVVKAATAVTAGGSLLVLSGLTLAGTVIALTIATPLLVIFSPVLVPAVI
                                          226
300
          6x OLEOSIN AF091840 ORF only (218)
TIFLLGAGFLASGGFGVAALSVLSWIYRYLTGKHPPGADQLESAKTKLASKAREMKDRAEQFSQQPVAGSQTS--
Translation of High GC oleosins spliced (226)
TIFLLGAGFLASGGFGVAALSVLSWIYRYLTGKHPPGADQLESAKTKLASKAREMKDRAEQFSQQPVAGSQTSHM
                                Consensus (226)
TIFLLGAGFLASGGFGVAALSVLSWIYRYLTGKHPPGADQLESAKTKLASKAREMKDRAEQFSQQPVAGSQTS
                                          301
375
          6x OLEOSIN AF091840 ORF only (291)
------MAEHYGQQQQTRAPHLQLQPRAQRVVKAATAVTAGGSLLVLSGLTLAGTVIALTIATPLLVIFSPVLVP
Translation of High GC oleosins spliced (301)
FKWPSAMAEHYGQQQQTRAPHLQLQPRAQRVVKAATAVTAGGSLLVLSGLTLAGTVIALTIATPLLVIFSPVLVP
                                Consensus (301)
MAEHYGQQQQTRAPHLQLQPRAQRVVKAATAVTAGGSLLVLSGLTLAGTVIALTIATPLLVIFSPVLVP
                                          376
450
          6x OLEOSIN AF091840 ORF only (360)
AVITIFLLGAGFLASGGFGVAALSVLSWIYRYLTGKHPPGADQLESAKTKLASKAREMKDRAEQFSQQPVAGSQT
Translation of High GC oleosins spliced (376)
AVITIFLLGAGFLASGGFGVAALSVLSWIYRYLTGKHPPGADQLESAKTKLASKAREMKDRAEQFSQQPVAGSQT
                                Consensus (376)
AVITIFLLGAGFLASGGFGVAALSVLSWIYRYLTGKHPPGADQLESAKTKLASKAREMKDRAEQFSQQPVAGSQT
                                          451
525
          6x OLEOSIN AF091840 ORF only (435)
S--------MAEHYGQQQQTRAPHLQLQPRAQRVVKAATAVTAGGSLLVLSGLTLAGTVIALTIATPLLVIFSPV
Translation of High GC oleosins spliced (451)
SSSELPWVDMAEHYGQQQQTRAPHLQLQPRAQRVVKAATAVTAGGSLLVLSGLTLAGTVIALTIATPLLVIFSPV
                                Consensus (451)
S       MAEHYGQQQQTRAPHLQLQPRAQRVVKAATAVTAGGSLLVLSGLTLAGTVIALTIATPLLVIFSPV
                                          526
600
          6x OLEOSIN AF091840 ORF only (502)
LVPAVITIFLLGAGFLASGGFGVAALSVLSWIYRYLTGKHPPGADQLESAKTKLASKAREMKDRAEQFSQQPVAG
Translation of High GC oleosins spliced (526)
LVPAVITIFLLGAGFLASGGFGVAALSVLSWIYRYLTGKHPPGADQLESAKTKLASKAREMKDRAEQFSQQPVAG
                                Consensus (526)
LVPAVITIFLLGAGFLASGGFGVAALSVLSWIYRYLTGKHPPGADQLESAKTKLASKAREMKDRAEQFSQQPVAG
                                          601
675
          6x OLEOSIN AF091840 ORF only (577)
SQTS--------MAEHYGQQQQTRAPHLQLQPRAQRVVKAATAVTAGGSLLVLSGLTLAGTVIALTIATPLLVIF
Translation of High GC oleosins spliced (601)
SQTSHMEFKLSTMAEHYGQQQQTRAPHLQLQPRAQRVVKAATAVTAGGSLLVLSGLTLAGTVIALTIATPLLVIF
                                Consensus (601) SQTS
MAEHYGQQQQTRAPHLQLQPRAQRVVKAATAVTAGGSLLVLSGLTLAGTVIALTIATPLLVIF
                                          676
```

Figure 69

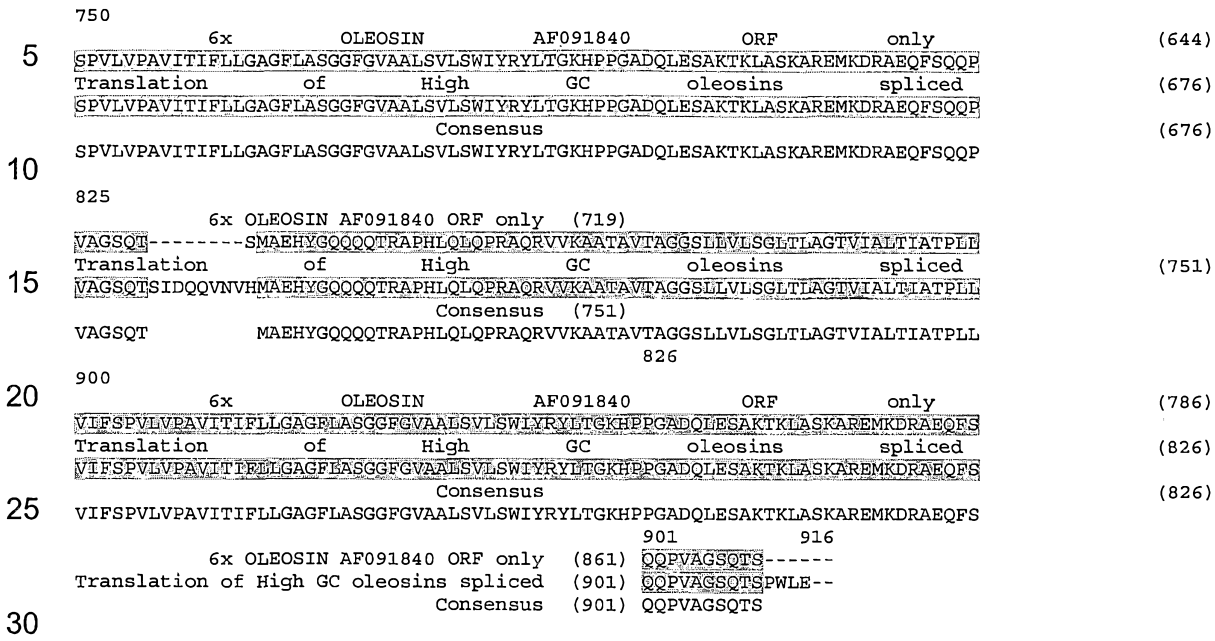


Figure 69 continued
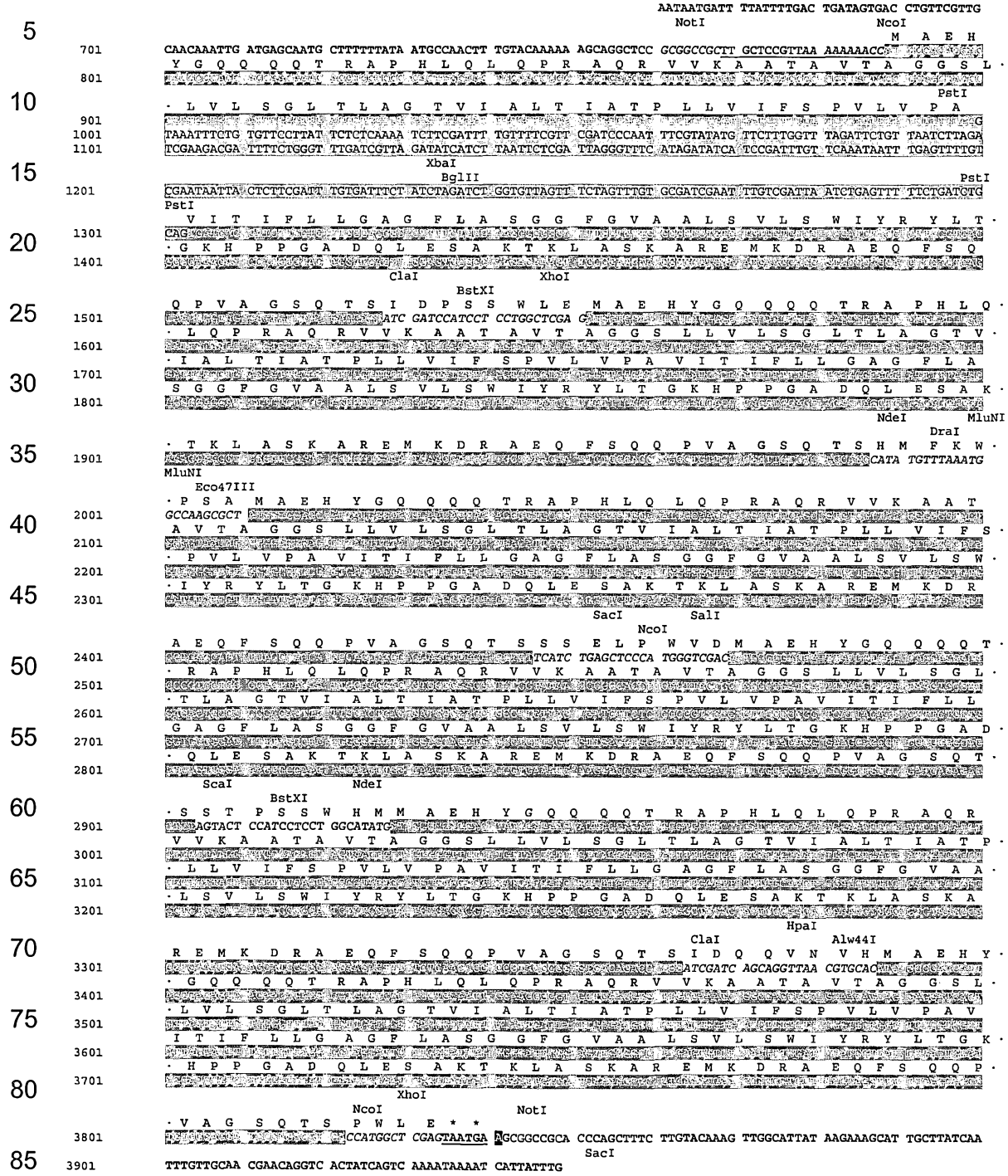


Figure 70

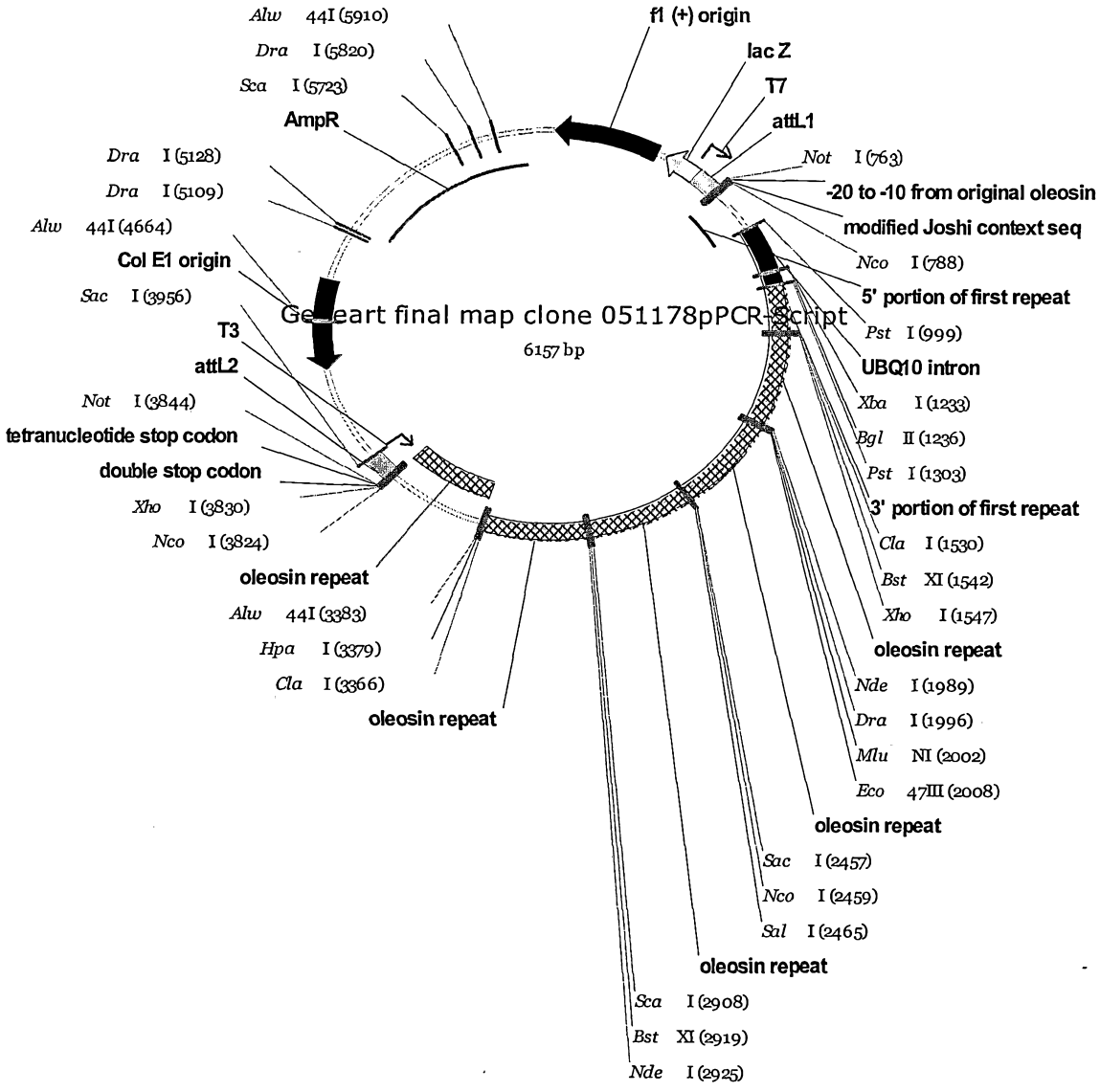

Alw 44I (5910)
Dra I (5820)
Sca I (5723)
AmpR
Dra I (5128)
Dra I (5109)
Alw 44I (4664)
Col E1 origin
Sac I (3956)
T3
attL2
Not I (3844)
tetranucleotide stop codon
double stop codon
Xho I (3830)
Nco I (3824)
oleosin repeat
Alw 44I (3383)
Hpa I (3379)
Cla I (3366)
oleosin repeat
f1 (+) origin
lac Z
T7
attL1
Not I (763)
-20 to -10 from original oleosin
modified Joshi context seq
Nco I (788)
5' portion of first repeat
Pst I (999)
UBQ10 intron
Xba I (1233)
Bgl II (1236)
Pst I (1303)
3' portion of first repeat
Cla I (1530)
Bst XI (1542)
Xho I (1547)
oleosin repeat
Nde I (1989)
Dra I (1996)
Mlu NI (2002)
Eco 47III (2008)
oleosin repeat
Sac I (2457)
Nco I (2459)
Sal I (2465)
oleosin repeat
Sca I (2908)
Bst XI (2919)
Nde I (2925)
Geneart final map clone 051178pPCR-Script
6157 bp

Figure 71

CAAATAATGATTTTATTTTGACTGATAGTGACCTGTTCGTTGCAACAAATTGATGAGCAATGCTTTTTTATAATGCCAACTTTGTACAAAAAAG
CAGGCTCCGCGGCCGCTTGCTCCGTTAAAAAAAACCATGGCAGAGCACTACGGGCAGCAGCAACAAACACGTGCCCCCCACTTGCAACTGCAAC
CGCGTGCTCAGCGTGTTGTAAAGGCAGCGACCGCGGTTACTGCCGGAGGTAGCCTGTTGGTGTTATCCGGGTTGACCCTGGCTGGAACGGTCAT
TGCGCTGACAATCGCCACACCTCTCCTGGTGATCTTCTCGCCTGTACTGGTCCCTGCAGGTAAATTTCTGTGTTCCTTATTCTCTCAAAATCTT
CGATTTTGTTTTCGTTCGATCCCAATTTCGTATATGTTCTTTGGTTTAGATTCTGTTAATCTTAGATCGAAGACGATTTTCTGGGTTTGATCGT
TAGATATCATCTTAATTCTCGATTAGGGTTTCATAGATATCATCCGATTTGTTCAAATAATTTGAGTTTTGTCGAATAATTACTCTTCGATTTG
TGATTTCTATCTAGATCTGGTGTTAGTTTCTAGTTTGTGCGATCGAATTTGTCGATTAATCTGAGTTTTTCTGATCTGCAGTCATCACCATCTT
CCTGTTGGGAGCGGGTTTTCTGGCAAGCGGGGGATTTGGTGTTGCCGCTCTGTCTGTGCTGTCCTGGATCTACCGTTACCTGACGGGGAAACAT
CCACCCGGAGCGGATCAGTTGGAGTCGGCCAAGACAAAGTTGGCGAGCAAGGCCCGTGAAATGAAGGACCGTGCCGAGCAGTTCAGTCAGCAAC
CGGTAGCAGGGTCTCAGACCAGCATCGATCCATCCTCCTGGCTCGAGATGGCGGAACACTACGGGCAACAACAGCAGACTCGTGCTCCCCACCT
GCAATTACAACCCCGTGCCCAACGTGTTGTGAAAGCGGCAACAGCAGTAACGGCAGGGGGAAGTTTGCTGGTCTTATCGGGGTTGACCTTAGCG
GGAACCGTGATTGCCCTGACAATTGCGACTCCGCTGCTGGTTATCTTCAGCCCCGTATTGGTTCCGGCCGTGATCACGATTTTTTTGCTGGGGG
CAGGATTTTTAGCCAGCGGAGGATTTGGGGTCGCAGCGTTGTCTGTGCTGAGTTGGATCTATCGTTATTTGACCGGGAAGCACCCACCTGGAGC
AGACCAGCTGGAGAGCGCGAAAACGAAGCTGGCATCGAAGGCGCGTGAAATGAAGGATCGTGCTGAACAATTCTCCCAGCAGCCTGTTGCCGGT
TCTCAGACCAGCCATATGTTTAAATGGCCAAGCGCTATGGCCGAGCATTATGGGCAGCAACAGCAAACCCGTGCCCCGCATCTGCAATTGCAAC
CTCGTGCCCAGCGTGTCGTTAAGGCGGCTACTGCGGTAACAGCGGGAGGGAGCTTACTGGTATTAAGCGGGCTGACATTGGCCGGAACGGTGAT
CGCCTTAACAATCGCGACACCCTTGCTGGTCATCTTCAGTCCGGTTCTGGTGCCCGCGGTGATTACGATTTTCCTGCTGGGAGCCGGTTTCTTA
GCATCGGGGGGTTTTGGGGTAGCAGCCTTGAGTGTCCTGTCGTGGATCTATCGTTACTTAACTGGAAAACACCCGCCAGGAGCTGACCAGTTGG
AGTCTGCAAAAACTAAGCTGGCGTCCAAAGCCCGTGAAATGAAGGATCGTGCTGAGCAGTTTAGCCAGCAGCCAGTTGCGGGAAGTCAGACCTC
TTCATCTGAGCTCCCATGGGTCGACATGGCGGAGCATTACGGTCAACAGCAACAGACCCGTGCTCCGCACTTACAATTGCAACCACGTGCTCAA
CGTGTCGTAAAAGCCGCCACGGCAGTTACTGCGGGGGGATCATTGCTGGTGTTAAGTGGGTTGACACTGGCGGGGACAGTTATTGCACTGACGA
TCGCGACCCCCTTGTTAGTGATCTTCTCCCCCGTTCTGGTTCCGGCGGTCATTACAATCTTTCTGTTGGGTGCCGGATTTTTAGCCTCTGGGGG
ATTTGGAGTAGCTGCCCTGTCAGTGTTGAGCTGGATCTACCGTTACTTAACAGGGAAGCACCCTCCCGGGGCAGATCAGTTGGAAAGCGCCAAG
ACCAAGCTGGCAAGTAAAGCGCGTGAAATGAAGGACCGTGCCGAACAATTTTCGCAGCAACCGGTTGCGGGATCACAGACCTCTAGTACTCCAT
CCTCCTGGCATATGATGGCCGAGCACTATGGACAACAGCAGCAGACGCGTGCCCCTCATCTGCAACTGCAACCCCGTGCTCAACGTGTCGTTAA
GGCTGCGACAGCAGTAACCGCTGGGGGTTCTCTGTTAGTGTTGTCAGGGCTGACTTTGGCGGGGACGGTAATTGCGTTGACCATTGCCACCCCG
CTGTTAGTGATTTTCAGCCCGGTACTGGTGCCAGCAGTTATCACGATCTTCTTGCTGGGTGCCGGATTCTTGGCAAGTGGAGGTTTTGGAGTTG
CGGCGCTGTCAGTTTTATCCTGGATCTATCGTTATCTGACAGGAAAACATCCCCCAGGTGCCGATCAGCTGGAGAGTGCCAAGACAAAACTGGC
GTCTAAGGCACGTGAAATGAAGGATCGTGCCGAACAGTTTTCTCAACAGCCCGTAGCGGGGTCACAGACCTCGATCGATCAGCAGGTTAACGTG
CACATGGCCGAACATTACGGACAGCAACAACAGACGCGTGCTCCACACCTGCAATTGCAACCGCGTGCTCAACGTGTTGTCAAAGCGGCGACCG
CCGTAACAGCAGGAGGATCACTGTTAGTGCTGTCGGGTTTAACCTTGGCCGGGACCGTCATTGCATTGACTATTGCGACGCCCTTACTGGTGAT
CTTTTCTCCGGTGCTGGTTCCCGCCGTTATTACCATCTTCTTGTTAGGGGCAGGATTCCTGGCATCAGGGGGATTCGGAGTTGCGGCGTTGAGT
GTCTTAAGTTGGATCTACCGTTATCTGACTGGAAAGCACCCGCCTGGGGCCGATCAACTGGAGTCAGCCAAAACGAAATTGGCGTCAAAAGCGC
GTGAAATGAAGGACCGTGCTGAGCAGTTTTCTCAGCAGCCTGTGGCAGGATCCCAGACATCACCATGGCTCGAGTAATGAAGCGGCCGCACCCA
GCTTTCTTGTACAAAGTTGGCATTATAAGAAAGCATTGCTTATCAATTTGTTGCAACGAACAGGTCACTATCAGTCAAAATAAAATCATTATTT
G

Figure 72

ATG
-20 to -10 from original oleosin
sesame oleosin1
modified Joshi context seq
NcoI
PstI
PstI
BamHI
PstI
sesame oleosin1
sesame oleosin2
101 TTGC TCCGTTAAAA AAAACCATGG CTGAGCATTA TGGTCAACAA CAGCAGACCA GGGCGCCTCA CCTGCAGCTG CAGCCGCGCG
201 CCCAGCGGGT AGTGAAGGCG GCCACCGCCG TGACAGCCGG CGGCTCGCTT CTCGTCCTCT CTGGCCTCAC TTTAGCCGGA ACTGTTATTG CGCTCACCAT
301 CGCCACTCCG CTGCTTGTGA TCTTTAGCCC CGTTCTGGTG CCGGCGGTCA TAACCATTTT CTTGCTGGGT GCGGGTTTTC TGGCATCCGG AGGCTTCGGC
401 GTGGCGGCGC TGAGTGTGCT GTCGTGGATT TACAGATATC TGACAGGGAA ACACCCGCCG GGGGCGGATC AGCTGGAATC GGCAAAGACG AAGCTGGCGA
501 GCAAGGCGCG AGAGATGAAG GATAGGGCAG AGCAGTTCTC GCAGCAGCCT GTTGGAGGCG GTGGATCCGG AGGCGGTGGT AGTATGGCTG AGCATTATGG
601 TCAACAACAG CAGACCAGGG CGCCTCACCT GCAGCTGCAG CCGCGCGCCC AGCGGGTAGT GAAGGCGGCC ACCGCCGTGA CAGCCGGCGG CTCGCTTCTC
701 GTCCTCTCTG GCCTCACTTT AGCCGGAACT GTTATTGCGC TCACCATCGC CACTCCGCTG CTTGTGATCT TTAGCCCCGT TCTGGTGCCG GCGGTCATAA
801 CCATTTTCTT GCTGGGTGCG GGTTTTCTGG CATCCGGAGG CTTCGGCGTG GCGGCGCTGA GTGTGCTGTC GTGGATTTAC AGATATCTGA CAGGGAAACA

Figure 73

```
                                                
```

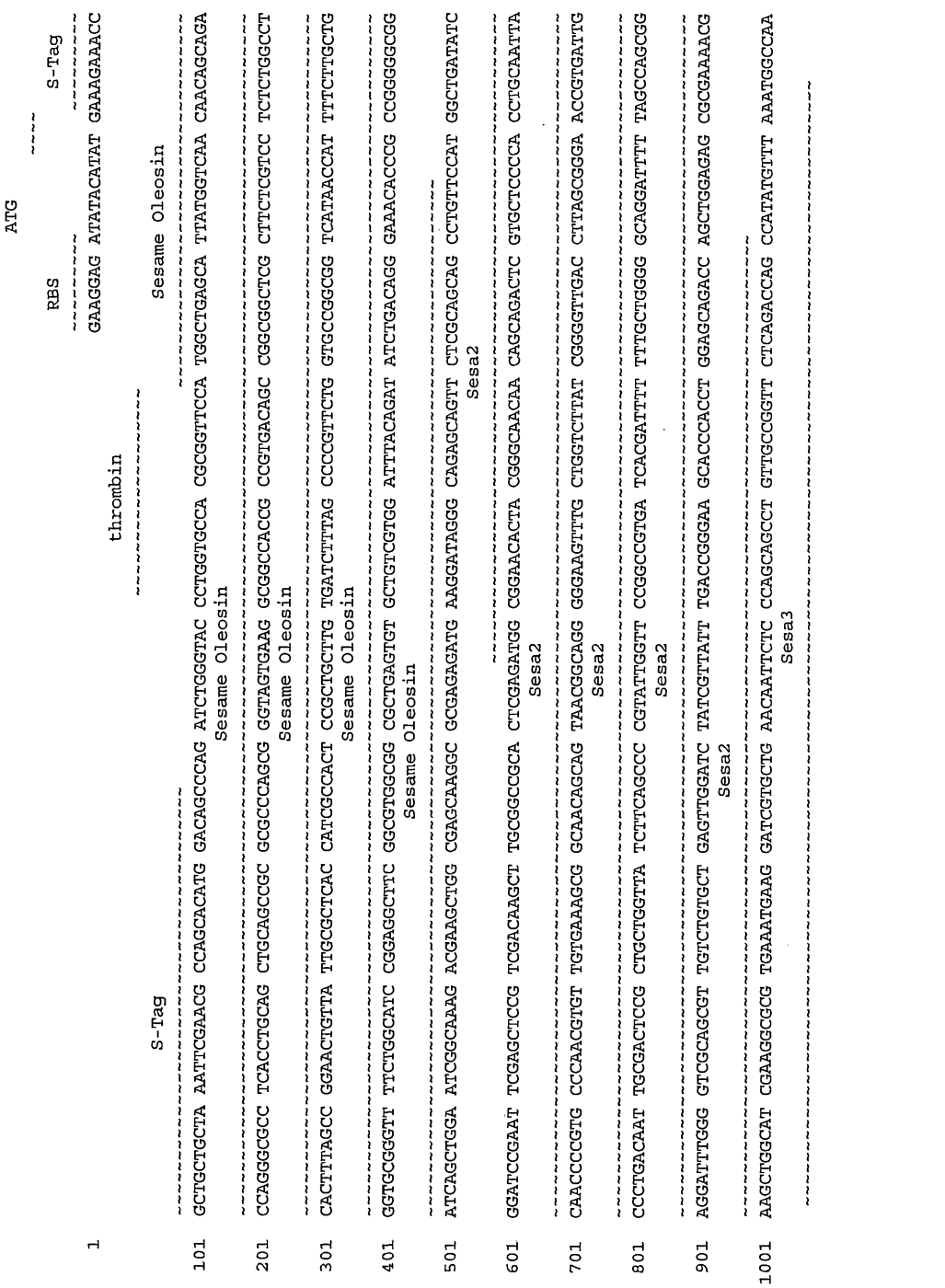

ATG
S-Tag
RBS
S-Tag
GAAGGAG ATATACATAT GAAAGAACC
thrombin
Sesame Oleosin
S-Tag
1
GCTGCTGCTA AATTCGAACG CCAGCACATG GACAGCCCAG ATCTGGGTAC CCTGGTGCCA CGCGGTTCCA TGGCTGAGCA TTATGGTCAA CAACAGCAGA
Sesame Oleosin
101
CCAGGGCGCC TCACCTGCAG CTGCAGCCGC GCGCCCAGCG GGTAGTGAAG GCGGCCACCG CCGTGACAGC CGGCGGCTCG CTTCTCGTCC TCTCTGGCCT
Sesame Oleosin
201
CACTTTAGCC GGAACTGTTA TTGCGCTCAC CATCGCCACT CCGCTGCTTG TGATCTTTAG CCCCGTTCTG GTGCCGGCGG TCATAACCAT TTTCTTGCTG
Sesame Oleosin
301
GGTGCGGGTT TTCTGGCATC CGGAGGCTTC GGCGTGGCGG CGCTGAGTGT GCTGTCGTGG ATTTACAGAT ATCTGACAGG GAAACACCCG CCGGGGGCGG
Sesame Oleosin
401
ATCAGCTGGA ATCGGCAAAG ACGAAGCTGG CGAGCAAGGC GCGAGAGATG AAGGATAGGG CAGAGCAGTT CTCGCAGCAG CCTGTTCCAT GGCTGATATC
Sesa2
501
GGATCCGAAT TCGAGCTCCG TCGACAAGCT TGCGGCCGCA CTCGAGATGG CGGAACACTA CGGGCAACAA CAGCAGACTC GTGCTCCCCA CCTGCAATTA
Sesa2
601
CAACCCCGTG CCCAACGTGT TGTGAAAGCG GCAACAGCAG TAACGGCAGG GGGAAGTTTG CTGGTCTTAT CGGGGTTGAC CTTAGCGGGA ACCGTGATTG
Sesa2
701
CCCTGACAAT TGCGACTCCG CTGCTGGTTA TCTTCAGCCC CGTATTGGTT CCGGCCGTGA TCACGATTTT TTTGCTGGGG GCAGGATTTT TAGCCAGCGG
Sesa2
801
AGGATTTGGG GTCGCAGCGT TGTCTGTGCT GAGTTGGATC TATCGTTATT TGACCGGGAA GCACCCACCT GGAGCAGACC AGCTGGAGAG CGCGAAAACG
Sesa2
901
AAGCTGGCAT CGAAGGCGCG TGAAATGAAG GATCGTGCTG AACAATTCTC CCAGCAGCCT GTTGCCGGTT CTCAGACCAG CCATATGTTT AAATGGCCAA
Sesa3
1001

Figure 74

```
1101  GCGCTATGGC CGAGCATTAT GGGCAGCAAC AGCAAACCCG TGCCCCGCAT CTGCAATTGC AACCTCGTGC CCAGCGTGTC GTTAAGGCGG CTACTGCGGT
                                                        Sesa3
      ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

1201  AACAGCGGGA GGGAGCTTAC TGGTATTAAG CGGGCTGACA TTGGCCGGAA CGGTGATCGC CTTAACAATC GCGACACCCT TGCTGGTCAT CTTCAGTCCG
                                                  Sesa3
      ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

1301  GTTCTGGTGC CCGCGGTGAT TACGATTTTC CTGCTGGGAG CCGGTTTCTT AGCATCGGGG GGTTTTGGGG TAGCAGCCTT GAGTGTCCTG TCGTGGATCT
                                        Sesa3
      ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

1401  ATCGTTACTT AACTGGAAAA CACCCGCCAG GAGCTGACCA GTTGGAGTCT GCAAAAACTA AGCTGGCGTC CAAAGCCCGT GAAATGAAGG ATCGTGCTGA
                Sesa3                                                                 Sesa4
      ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

1501  GCAGTTTAGC CAGCAGCCAG TTGCGGGAAG TCAGACCTCT TCATCTGAGC TCCCATGGGT CGACATGGCG GAGCATTACG GTCAACAGCA ACAGACCCGT
                                        Sesa4
      ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

1601  GCTCCGCACT TACAATTGCA ACCACGTGCT CAACGTGTCG TAAAAGCCGC CACGGCAGTT ACTGCGGGGG GATCATTGCT GGTGTTAAGT GGGTTGACAC
                                              Sesa4
      ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

1701  TGGCGGGGAC AGTTATTGCA CTGACGATCG CGACCCCCTT GTTAGTGATC TTCTCCCCCG TTCTGGTTCC GGCGGTCATT ACAATCTTTC TGTTGGGTGC
                                        Sesa4
      ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

1801  CGGATTTTTA GCCTCTGGGG GATTTGGAGT AGCTGCCCTG TCAGTGTTGA GCTGGATCTA CCGTTACTTA ACAGGGAAGC ACCCTCCCGG GGCAGATCAG
                                        Sesa4
      ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

1901  TTGGAAAGCG CCAAGACCAA GCTGGCAAGT AAAGCGCGTG AAATGAAGGA CCGTGCCGAA CAATTTTCGC AGCAACCGGT TGCGGGATCA CAGACCTCTA
                                              Sesa5
      ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

2001  GTACTCCATC CTCCTGGCAT ATGATGGCCG AGCACTATGG ACAACAGCAG CAGACGCGTG CCCCTCATCT GCAACTGCAA CCCCGTGCTC AACGTGTCGT
                                        Sesa5
      ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

2101  TAAGGCTGCG ACAGCAGTAA CCGCTGGGGG TTCTCTGTTA GTGTTGTCAG GGCTGACTTT GGCGGGGACG GTAATTGCGT TGACCATTGC CACCCCGCTG
                                              Sesa5
      ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

2201  TTAGTGATTT TCAGCCCGGT ACTGGTGCCA GCAGTTATCA CGATCTTCTT GCTGGGTGCC GGATTCTTGG CAAGTGGAGG TTTTGGAGTT GCGGCGCTGT
                                        Sesa5
      ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

2301  CAGTTTTATC CTGGATCTAT CGTTATCTGA CAGGAAAACA TCCCCCAGGT GCCGATCAGC TGGAGAGTGC CAAGACAAAA CTGGCGTCTA AGGCACGTGA
                Sesa5                                                                 Sesa6
      ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

2401  AATGAAGGAT CGTGCCGAAC AGTTTTCTCA ACAGCCCGTA GCGGGGTCAC AGACCTCGAT CGATCAGCAG GTTAACGTGC ACATGGCCGA ACATTACGGA
                                        Sesa6
      ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
```

Figure 74 continued

```
2501  CAGCAACAAC AGACGCGTGC TCCACACCTG CAATTGCAAC CGCGTGCTCA ACGTGTTGTC AAAGCGGCGA CCGCCGTAAC AGCAGGAGGA TCACTGTTAG
                                                 Sesa6
      ~~~
```


rbs
S-Tag
GAAGGAG ATATACATAT GAAAGAAACC
Ole3.1
S-Tag
thrombin
GCTGCTGCTA AATTCGAACG CCAGCACATG GACAGCCCAG ATCTGGGTAC CCTGGTGCCA CGCGGTTCCA TGGCTGAGCA TTATGGTCAA CAACAGCAGA
Ole3.1
CCAGGGCGCC TCACCTGCAG CTGCAGCCGC GCGCCCAGCG GGTAGTGAAG GCGGCCACCG CCGTGACAGC CGGCGGCTCG CTTCTCGTCC TCTCTGGCCT
Ole3.1
CACTTTAGCC GGAACTGTTA TTGCGCTCAC CATCGCCACT CCGCTGCTTG TGATCTTTAG CCCCGTTCTG GTGCCGGCGG TCATAACCAT TTTCTTGCTG
Ole3.1
GGTGCGGGTT TTCTGGCATC CGGAGGCTTC GGCGTGGCGG CGCTGAGTGT GCTGTCGTGG ATTTACAGAT ATCTGACAGG GAAACACCCG CCGGGGCGG
Ole3.1
Hinge-1_to_2
ATCAGCTGGA ATCGGCAAAG ACGAAGCTGG CGAGCAAGGC GCGAGAGATG AAGGATAGGG CAGAGCAGTT CTCGCAGCAG CCTGTTGGAG GCGGTGGATC
Ole3.2
Hinge-1_to_2
CGGAGGCGGT GGTAGTATGG CTGAGCATTA TGGTCAACAA CAGCAGACCA GGGCGCCTCA CCTGCAGCTG CAGCCGCGCG CCCAGCGGGT AGTGAAGGCG
Ole3.2
GCCACCGCCG TGACAGCCGG CGGCTCGCTT CTCGTCCTCT CTGGCCTCAC TTTAGCCGGA ACTGTTATTG CGCTCACCAT CGCCACTCCG CTGCTTGTGA
Ole3.2
TCTTTAGCCC CGTTCTGGTG CCGGCGGTCA TAACCATTTT CTTGCTGGGT GCGGGTTTTC TGGCATCCGG AGGCTTCGGC GTGGCGGCGC TGAGTGTGCT
Ole3.2
GTCGTGGATT TACAGATATC TGACAGGGAA ACACCCGCCG GGGGCGGATC AGCTGGAATC GGCAAAGACG AAGCTGGCGA GCAAGGCGCG AGAGATGAAG
Hinge-2_to_3
Ole3.2
GATAGGGCAG AGCAGTTCTC GCAGCAGCCT GTTGGGGGCG GTGGATCCGG TGGAGGGGGA TCCATGGCGA TATCGGATCC GAATTCGAGC TCCGTCGACA
stop
His-Tag
AGCTTGCGGC CGCACTCGAG CACCACCACC ACCACCACTG A

Figure 75

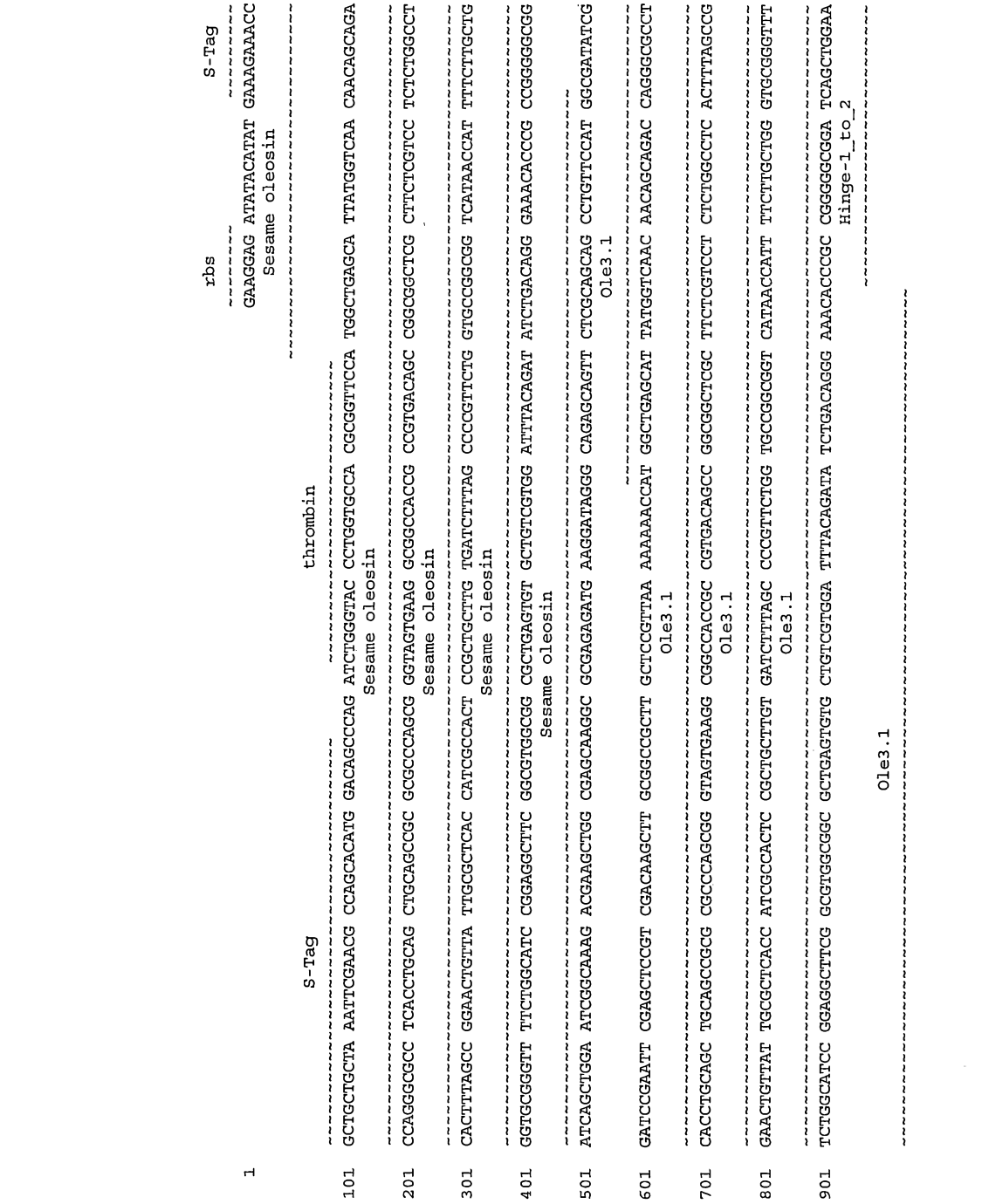

rbs
S-Tag
S-Tag
GAAGGAG ATATACATAT GAAAGAAACC
Sesame oleosin
thrombin
S-Tag
GCTGCTGCTA AATTCGAACG CCAGCACATG GACAGCCCAG ATCTGGGTAC CCTGGTGCCA CGCGGTTCCA TGGCTGAGCA TTATGGTCAA CAACAGCAGA
Sesame oleosin
CCAGGGCGCC TCACCTGCAG CTGCAGCCGC GCGCCCAGCG GGTAGTGAAG GCGGCCACCG CCGTGACAGC CGGCGGCTCG CTTCTCGTCC TCTCTGGCCT
Sesame oleosin
CACTTTAGCC GGAACTGTTA TTGCGCTCAC CATCGCCACT CCGCTGCTTG TGATCTTTAG CCCCGTTCTG GTGCCGGCGG TCATAACCAT TTTCTTGCTG
Sesame oleosin
GGTGCGGGTT TTCTGGCATC CGGAGGCTTC GGCGTGGCGG CGCTGAGTGT GCTGTCGTGG ATTTACAGAT ATCTGACAGG GAAACACCCG CCGGGGGCGG
Sesame oleosin
ATCAGCTGGA ATCGGCAAAG ACGAAGCTGG CGAGCAAGGC GCGAGAGATG AAGGATAGGG CAGAGCAGTT CTCGCAGCAG CCTGTTCCAT GGCGATATCG
Ole3.1
GATCCGAATT CGAGCTCCGT CGACAAGCTT GCGGCCGCTT GCTCCGTTAA AAAAAACCAT GGCTGAGCAT TATGGTCAAC AACAGCAGAC CAGGGCGCCT
Ole3.1
CACCTGCAGC TGCAGCCGCG CGCCCAGCGG GTAGTGAAGG CGGCCACCGC CGTGACAGCC GGCGGCTCGC TTCTCGTCCT CTCTGGCCTC ACTTTAGCCG
Ole3.1
GAACTGTTAT TGCGCTCACC ATCGCCACTC CGCTGCTTGT GATCTTTAGC CCCGTTCTGG TGCCGGCGGT CATAACCATT TTCTTGCTGG GTGCGGGTTT
Ole3.1
TCTGGCATCC GGAGGCTTCG GCGTGGCGGC GCTGAGTGTG CTGTCGTGGA TTTACAGATA TCTGACAGGG AAACACCCGC CGGGGGCGGA TCAGCTGGAA
Hinge-1_to_2
Ole3.1

Figure 76

```
1001  TCGGCAAAGA CGAAGCTGGC GAGCAAGGCG CGAGAGATGA AGGATAGGGC AGAGCAGTTC TCGCAGCAGC CTGTTGGAGG CGGTGGATCC GGAGGCGGTG
                                                                  Ole3.2
      ~~~
```

```
                                                
```

```
901  GTCGTGGATT TACAGATATC TGACAGGGAA ACACCCGCCG GGGGCGGATC AGCTGGAATC GGCAAAGACG AAGCTGGCGA GCAAGGCGCG AGAGATGAAG
                                                 Hinge-2_to_3
                                                 ~~~~~~~~~~~~
                                                 ~~~~~~~~~~~~
                                                 ~~~~~~~~~~~~
```

```
                                    Ole3.3
                                    ~~~
```

```
   1 MKETAAAKFE RQHMDSPDLG TLVPRGSMAE HYGQQQQTRA
  PHLQLQPRAQ
  51 RVVKAATAVT AGGSLLVLSG LTLAGTVIAL TIATPLLVIF
  SPVLVPAVIT
 101 IFLLGAGFLA SGGFGVAALS VLSWIYRYLT GKHPPGADQL
  ESAKTKLASK
 151 AREMKDRAEQ FSQQPVPWLI SDPNSSSVDK LAAALEMAEH
  YGQQQQTRAP
 201 HLQLQPRAQR VVKAATAVTA GGSLLVLSGL TLAGTVIALT
  IATPLLVIFS
 251 PVLVPAVITI FLLGAGFLAS GGFGVAALSV LSWIYRYLTG
  KHPPGADQLE
 301 SAKTKLASKA REMKDRAEQF SQQPVAGSQT SHMFKWPSAM
  AEHYGQQQQT
 351 RAPHLQLQPR AQRVVKAATA VTAGGSLLVL SGLTLAGTVI
  ALTIATPLLV
 401 IFSPVLVPAV ITIFLLGAGF LASGGFGVAA LSVLSWIYRY
  LTGKHPPGAD
 451 QLESAKTKLA SKAREMKDRA EQFSQQPVAG SQTSSSELPW
  VDMAEHYGQQ
 501 QQTRAPHLQL QPRAQRVVKA ATAVTAGGSL LVLSGLTLAG
  TVIALTIATP
 551 LLVIFSPVLV PAVITIFLLG AGFLASGGFG VAALSVLSWI
  YRYLTGKHPP
 601 GADQLESAKT KLASKAREMK DRAEQFSQQP VAGSQTSSTP
  SSWHMMAEHY
 651 GQQQQTRAPH LQLQPRAQRV VKAATAVTAG GSLLVLSGLT
  LAGTVIALTI
 701 ATPLLVIFSP VLVPAVITIF LLGAGFLASG GFGVAALSVL
  SWIYRYLTGK
```

Figure 78

751 HPPGADQLES AKTKLASKAR EMKDRAEQFS QQPVAGSQTS
IDQQVNVHMA
801 EHYGQQQQTR APHLQLQPRA QRVVKAATAV TAGGSLLVLS
GLTLAGTVIA
851 LTIATPLLVI FSPVLVPAVI TIFLLGAGFL ASGGFGVAAL
SVLSWIYRYL
901 TGKHPPGADQ LESAKTKLAS KAREMKDRAE QFSQQPVAGS
QTSPWLEHHH
951 HHH*

Figure 78 continued

```
   1 MKETAAAKFE RQHMDSPDLG TLVPRGSMAE HYGQQQQTRA
  PHLQLQPRAQ
  51 RVVKAATAVT AGGSLLVLSG LTLAGTVIAL TIATPLLVIF
  SPVLVPAVIT
 101 IFLLGAGFLA SGGFGVAALS VLSWIYRYLT GKHPPGADQL
  ESAKTKLASK
 151 AREMKDRAEQ FSQQPVGGGG SGGGGSMAEH YGQQQQTRAP
  HLQLQPRAQR
 201 VVKAATAVTA GGSLLVLSGL TLAGTVIALT IATPLLVIFS
  PVLVPAVITI
 251 FLLGAGFLAS GGFGVAALSV LSWIYRYLTG KHPPGADQLE
  SAKTKLASKA
 301 REMKDRAEQF SQQPVGGGGS GGGGSMAISD PNSSSVDKLA
  AALEHHHHHH
 351 *
```

Figure 79

```
  1 MKETAAAKFE RQHMDSPDLG TLVPRGSMAE HYGQQQQTRA
PHLQLQPRAQ
 51 RVVKAATAVT AGGSLLVLSG LTLAGTVIAL TIATPLLVIF
SPVLVPAVIT
101 IFLLGAGFLA SGGFGVAALS VLSWIYRYLT GKHPPGADQL
ESAKTKLASK
151 AREMKDRAEQ FSQQPVPWRY RIRIRAPSTS LRPLAPLKKT
MAEHYGQQQQ
201 TRAPHLQLQP RAQRVVKAAT AVTAGGSLLV LSGLTLAGTV
IALTIATPLL
251 VIFSPVLVPA VITIFLLGAG FLASGGFGVA ALSVLSWIYR
YLTGKHPPGA
301 DQLESAKTKL ASKAREMKDR AEQFSQQPVG GGGSGGGGSM
AEHYGQQQQT
351 RAPHLQLQPR AQRVVKAATA VTAGGSLLVL SGLTLAGTVI
ALTIATPLLV
401 IFSPVLVPAV ITIFLLGAGF LASGGFGVAA LSVLSWIYRY
LTGKHPPGAD
451 QLESAKTKLA SKAREMKDRA EQFSQQPVGG GGSGGGGSMA
EHYGQQQQTR
501 APHLQLQPRA QRVVKAATAV TAGGSLLVLS GLTLAGTVIA
LTIATPLLVI
551 FSPVLVPAVI TIFLLGAGFL ASGGFGVAAL SVLSWIYRYL
TGKHPPGADQ
601 LESAKTKLAS KAREMKDRAE QFSQQPVPWL EHHHHHH*
```

Figure 80

1 MKETAAAKFE RQHMDSPDLG TLVPRGSMAE HYGQQQQTRA PHLQLQPRAQ

51 RVVKAATAVT AGGSLLVLSG LTLAGTVIAL TIATPLLVIF SPVLVPAVIT

101 IFLLGAGFLA SGGFGVAALS VLSWIYRYLT GKHPPGADQL ESAKTKLASK

151 AREMKDRAEQ FSQQPVGGGG SGGGGSMAEH YGQQQQTRAP HLQLQPRAQR

201 VVKAATAVTA GGSLLVLSGL TLAGTVIALT IATPLLVIFS PVLVPAVITI

251 FLLGAGFLAS GGFGVAALSV LSWIYRYLTG KHPPGADQLE SAKTKLASKA

301 REMKDRAEQF SQQPVGGGGS GGGGSMGYRI RIRAPSTSLR PLAPLKKTMA

351 EHYGQQQQTR APHLQLQPRA QRVVKAATAV TAGGSLLVLS GLTLAGTVIA

401 LTIATPLLVI FSPVLVPAVI TIFLLGAGFL ASGGFGVAAL SVLSWIYRYL

451 TGKHPPGADQ LESAKTKLAS KAREMKDRAE QFSQQPVGGG GSGGGGSMAE

501 HYGQQQQTRA PHLQLQPRAQ RVVKAATAVT AGGSLLVLSG LTLAGTVIAL

551 TIATPLLVIF SPVLVPAVIT IFLLGAGFLA SGGFGVAALS VLSWIYRYLT

601 GKHPPGADQL ESAKTKLASK AREMKDRAEQ FSQQPVGGGG SGGGGSMAEH

651 YGQQQQTRAP HLQLQPRAQR VVKAATAVTA GGSLLVLSGL TLAGTVIALT

Figure 81

701 IATPLLVIFS PVLVPAVITI FLLGAGFLAS GGFGVAALSV
LSWIYRYLTG

751 KHPPGADQLE SAKTKLASKA REMKDRAEQF SQQPVPWLEH HHHHH*

Figure 81 continued

```
  1 MAEHYGQQQQ TRAPHLQLQP RAQRVVKAAT AVTAGGSLLV
LSGLTLAGTV
 51 IALTIATPLL VIFSPVLVPA VITIFLLGAG FLASGGFGVA
ALSVLSWIYR
101 YLTGKHPPGA DQLESAKTKL ASKAREMKDR AEQFSQQPVG
GGGSGGGGSM
151 AEHYGQQQQT RAPHLQLQPR AQRVVKAATA VTAGGSLLVL
SGLTLAGTVI
201 ALTIATPLLV IFSPVLVPAV ITIFLLGAGF LASGGFGVAA
LSVLSWIYRY
251 LTGKHPPGAD QLESAKTKLA SKAREMKDRA EQFSQQPVGG
GGSGGGGSMA
301 EHYGQQQQTR APHLQLQPRA QRVVKAATAV TAGGSLLVLS
GLTLAGTVIA
351 LTIATPLLVI FSPVLVPAVI TIFLLGAGFL ASGGFGVAAL
SVLSWIYRYL
401 TGKHPPGADQ LESAKTKLAS KAREMKDRAE QFSQQPVPWL E*
```

Figure 82

```
301  MAEHY GQQQQTRAPH LQLQPRAQRV
351  VKAATAVTAG     GSLLVLSGLT     LAGTVIALTI     ATPLLVIFSP
  VLVPAVITIF
401  LLGAGFLASG     GFGVAALSVL     SWIYRYLTGK     HPPGADQLES
  AKTKLASKAR
451  EMKDRAEQFS QQPVPWLEHH HHHH*
```

Figure 83

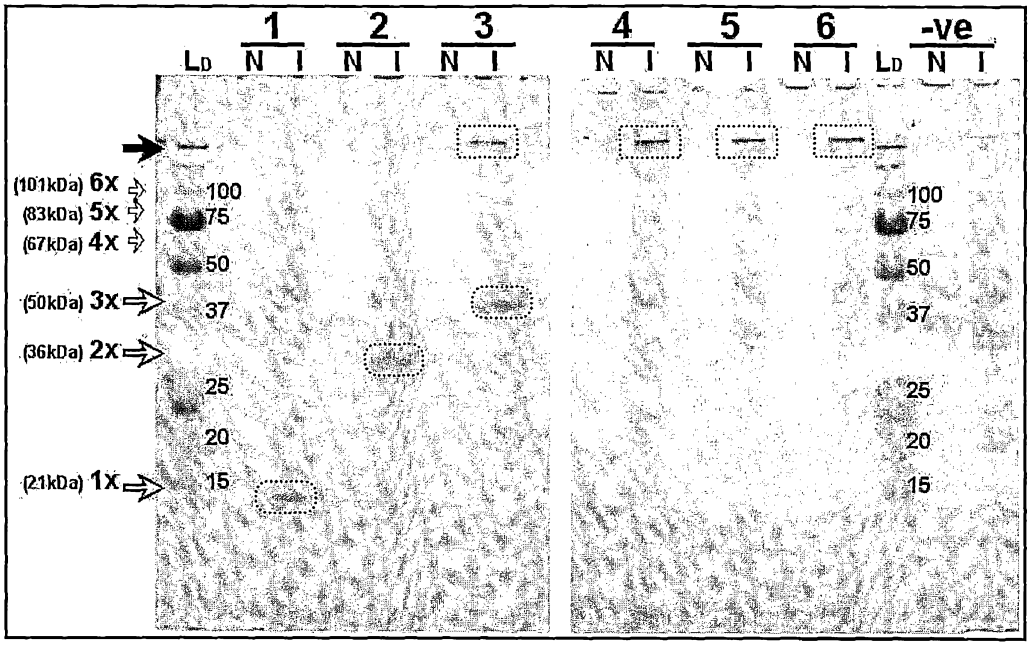


Figure 84A

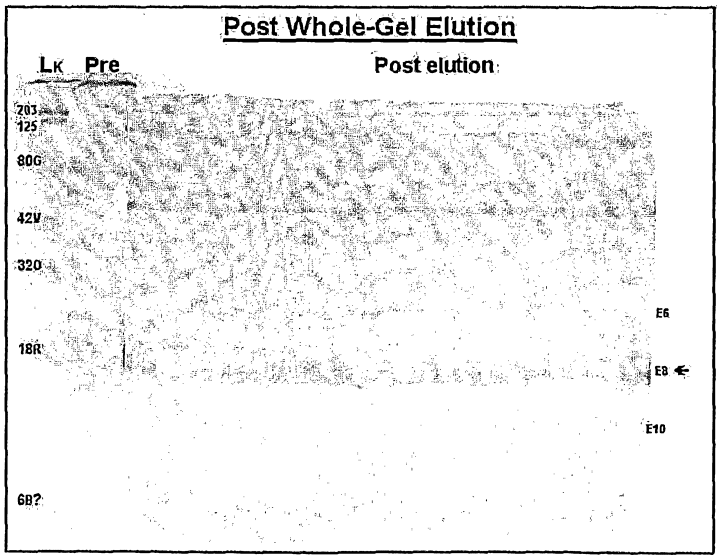


Figure 84B

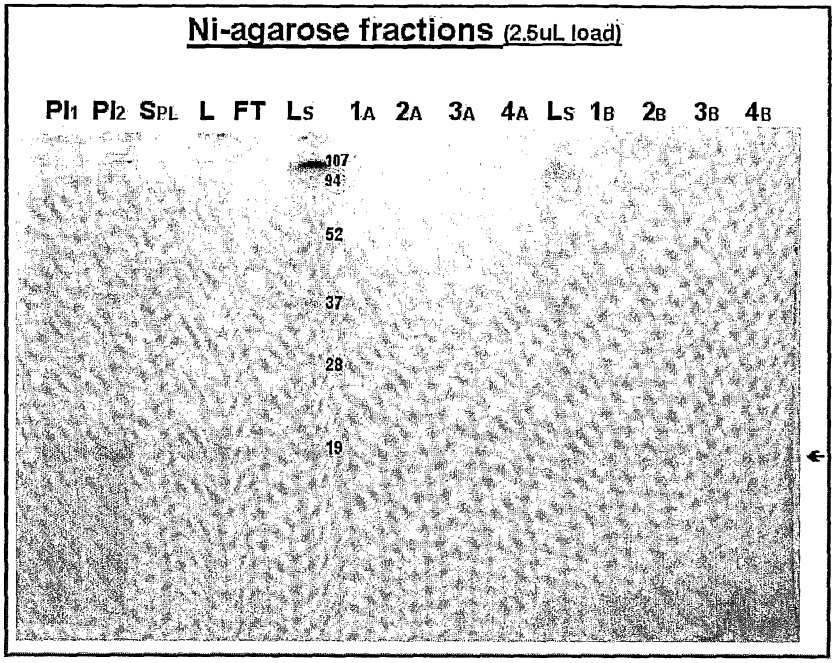

Ni-agarose fractions (2.5uL load)
PI1 PI2 SPL L FT LS 1A 2A 3A 4A LS 1B 2B 3B 4B
107
94
52
37
28
19

Figure 85

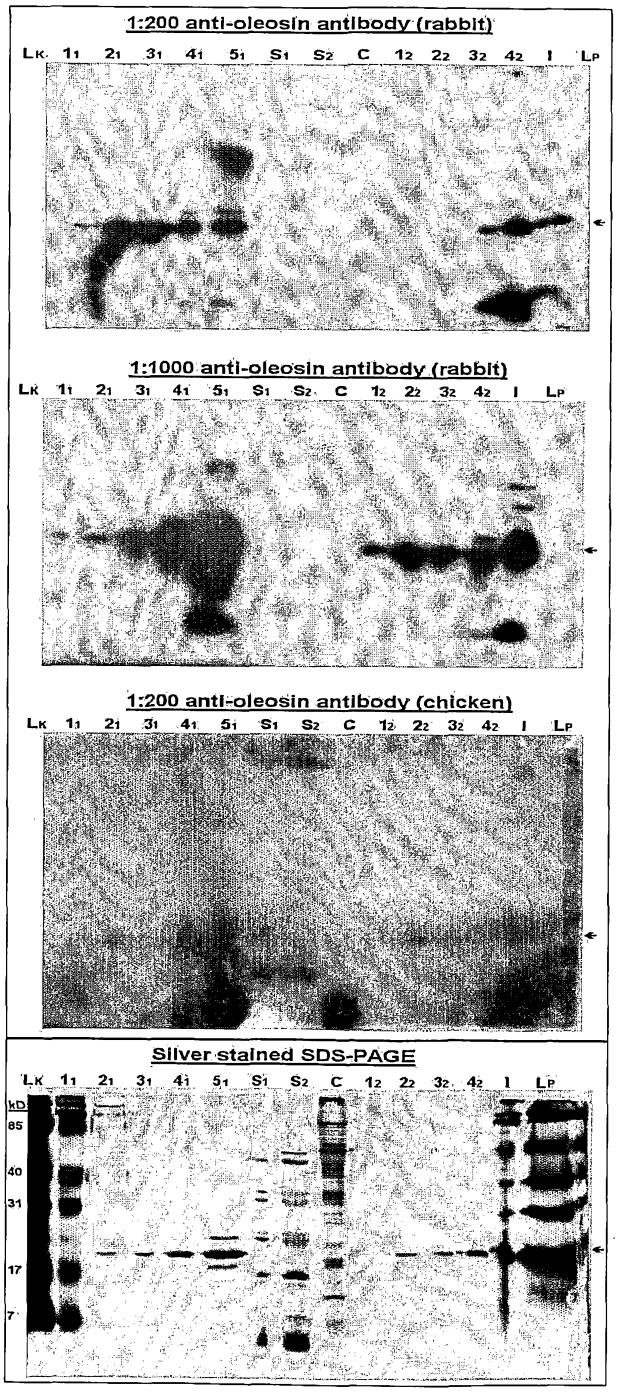

1:200 anti-oleosin antibody (rabbit)
LK 11 21 31 41 51 S1 S2 C 12 22 32 42 I LP
1:1000 anti-oleosin antibody (rabbit)
LK 11 21 31 41 51 S1 S2 C 12 22 32 42 I LP
1:200 anti-oleosin antibody (chicken)
LK 11 21 31 41 51 S1 S2 C 12 22 32 42 I LP
Silver stained SDS-PAGE
LK 11 21 31 41 51 S1 S2 C 12 22 32 42 I LP
kD
85
40
31
17
7

Figure 86

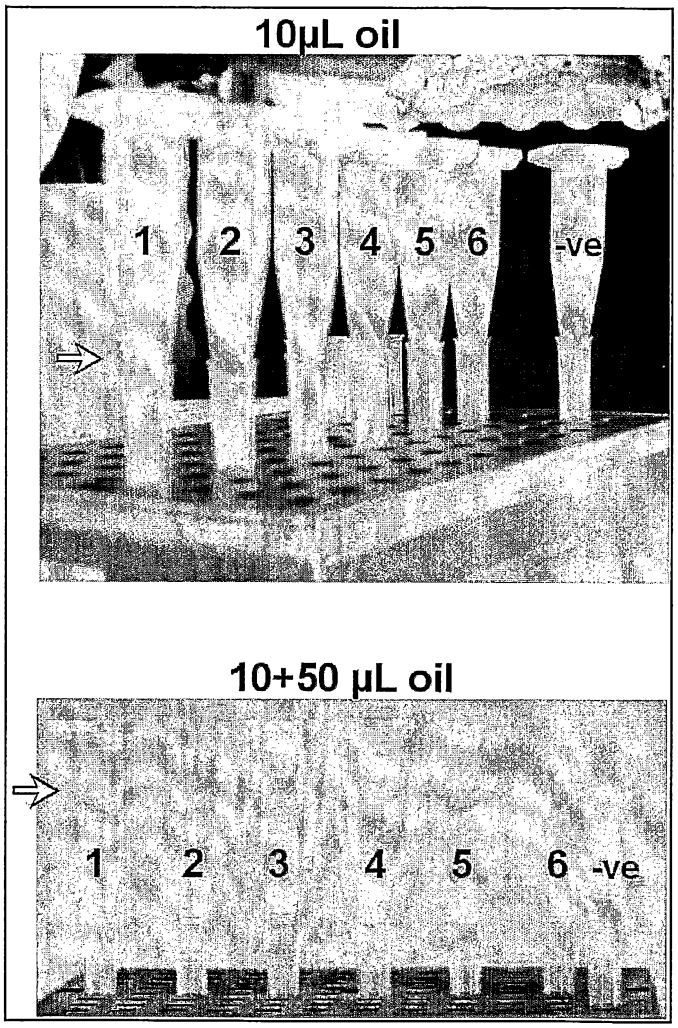


Figure 87

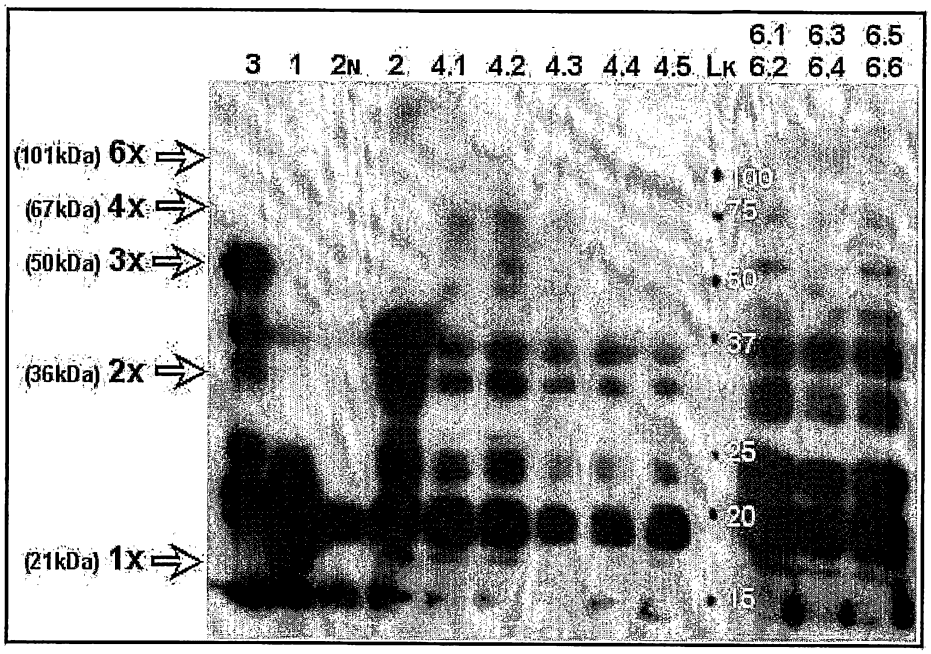

3 1 2N 2 4.1 4.2 4.3 4.4 4.5 LK 6.1 6.2 6.3 6.4 6.5 6.6
(101kDa) 6x
(67kDa) 4x
(50kDa) 3x
(36kDa) 2x
(21kDa) 1x
100
75
50
37
25
20
15

Figure 88

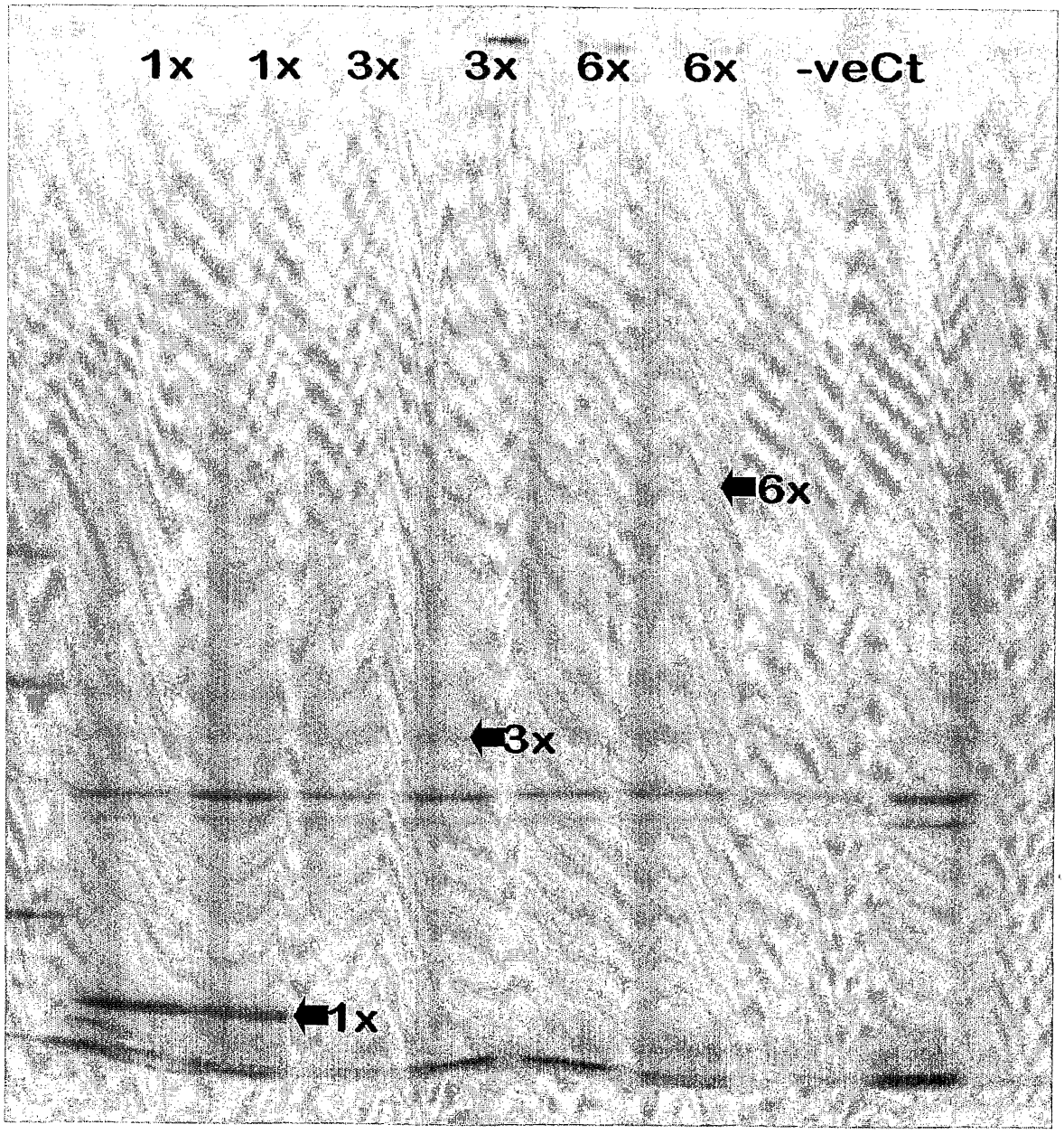


Figure 89

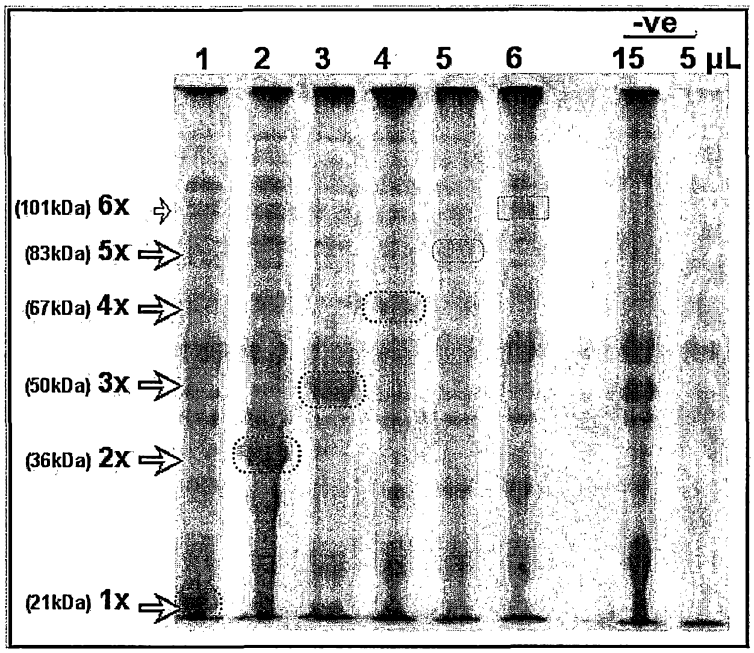

-ve
1
2
3
4
5
6
15
5 µL
(101kDa) 6x
(83kDa) 5x
(67kDa) 4x
(50kDa) 3x
(36kDa) 2x
(21kDa) 1x

Figure 90

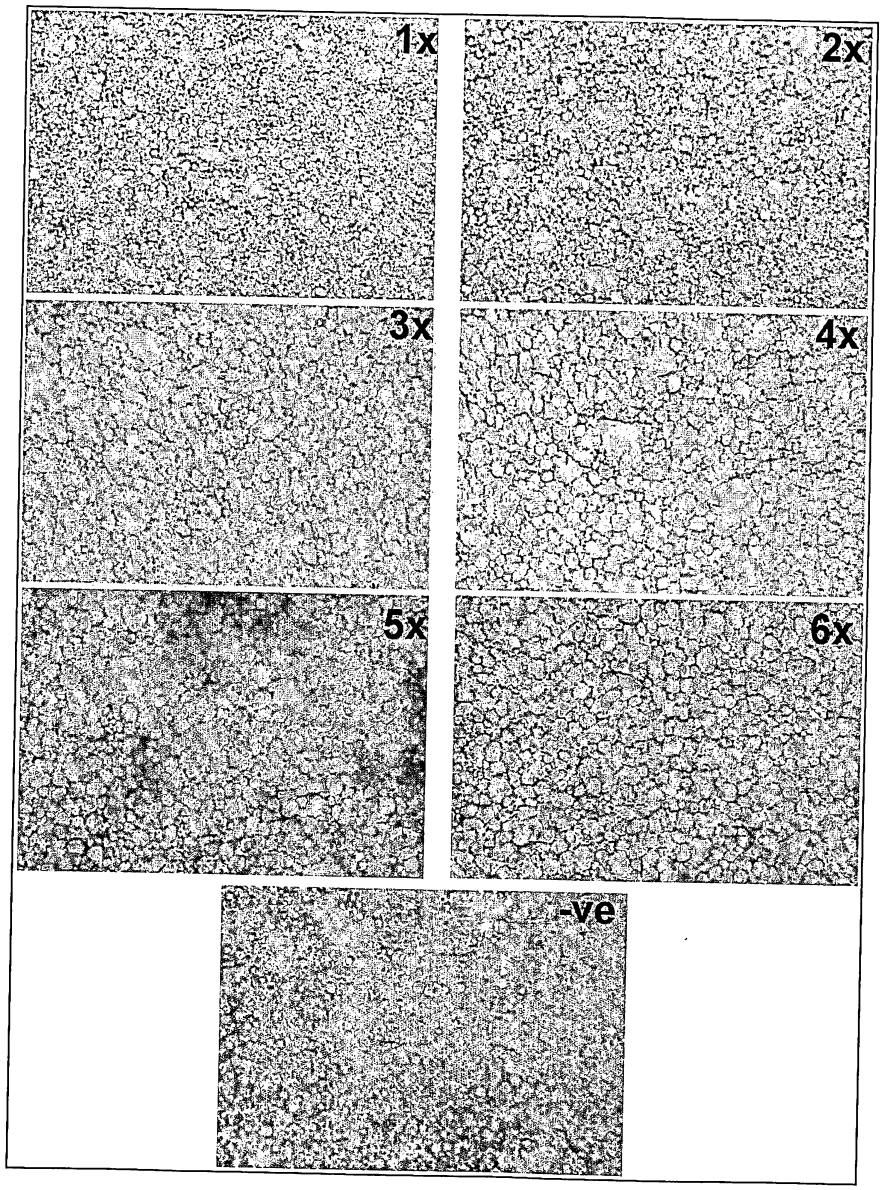


Figure 91

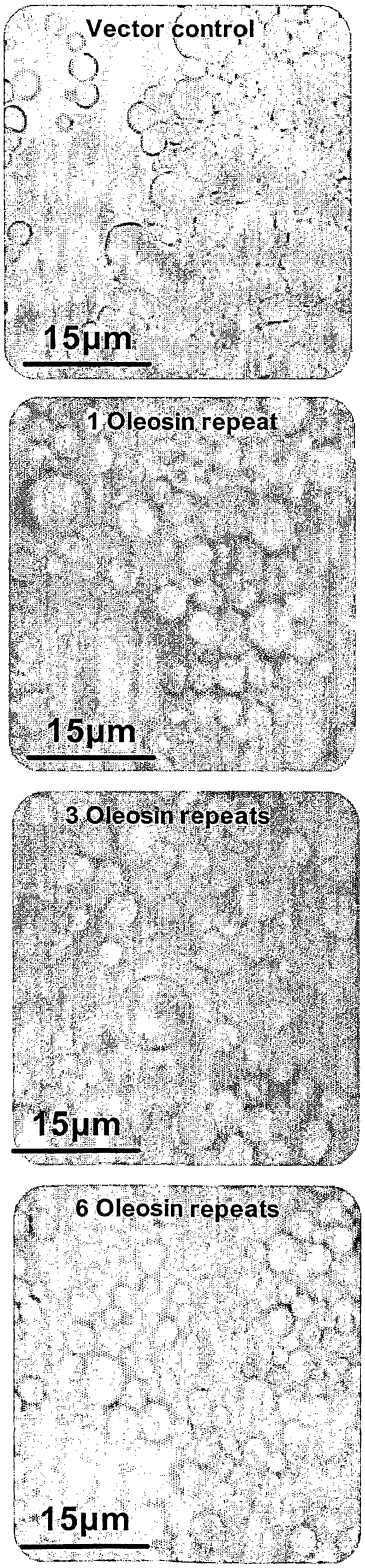

Vector control
15µm
1 Oleosin repeat
15µm
3 Oleosin repeats
15µm
6 Oleosin repeats
15µm

Figure 92

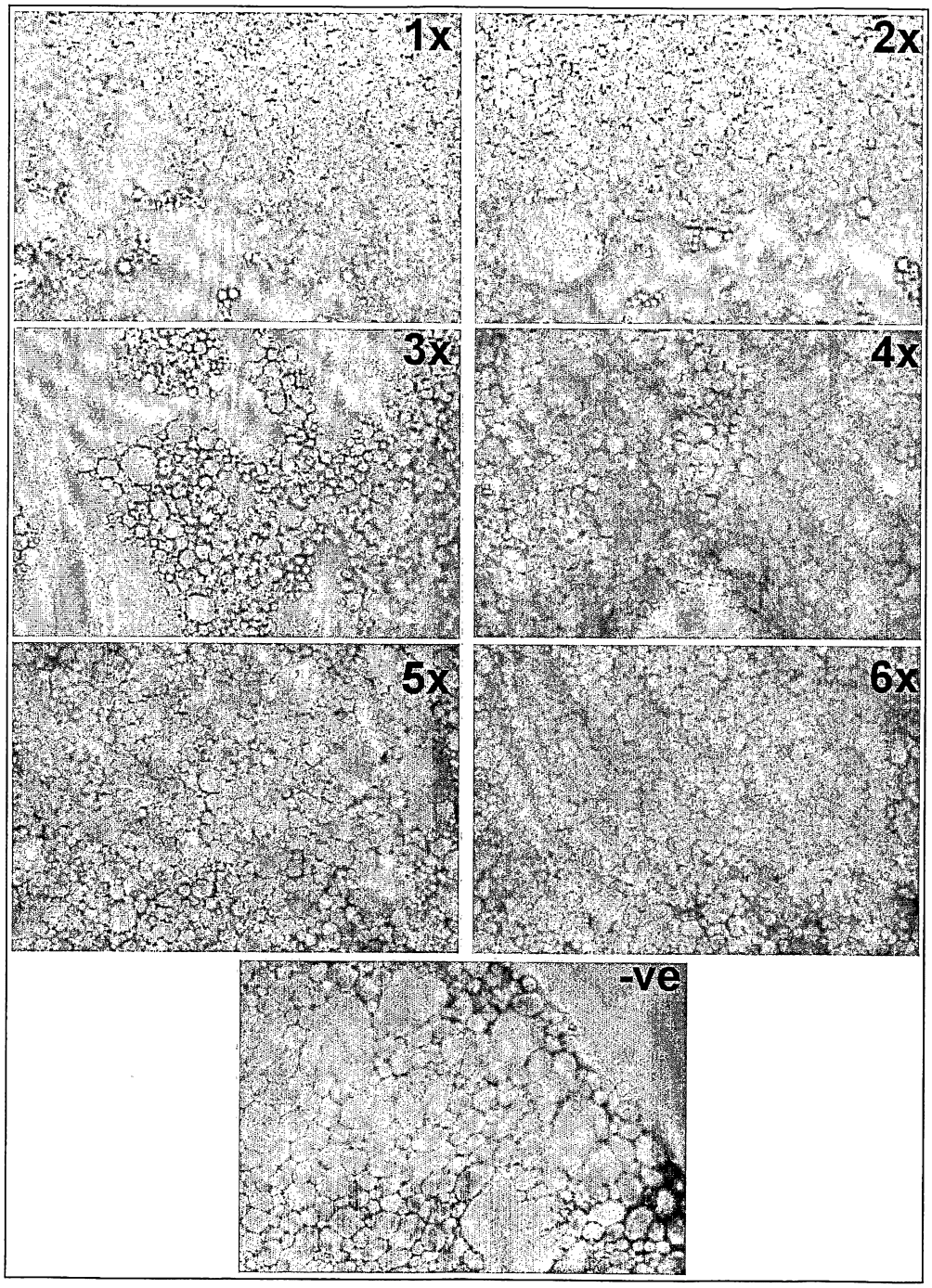


Figure 93

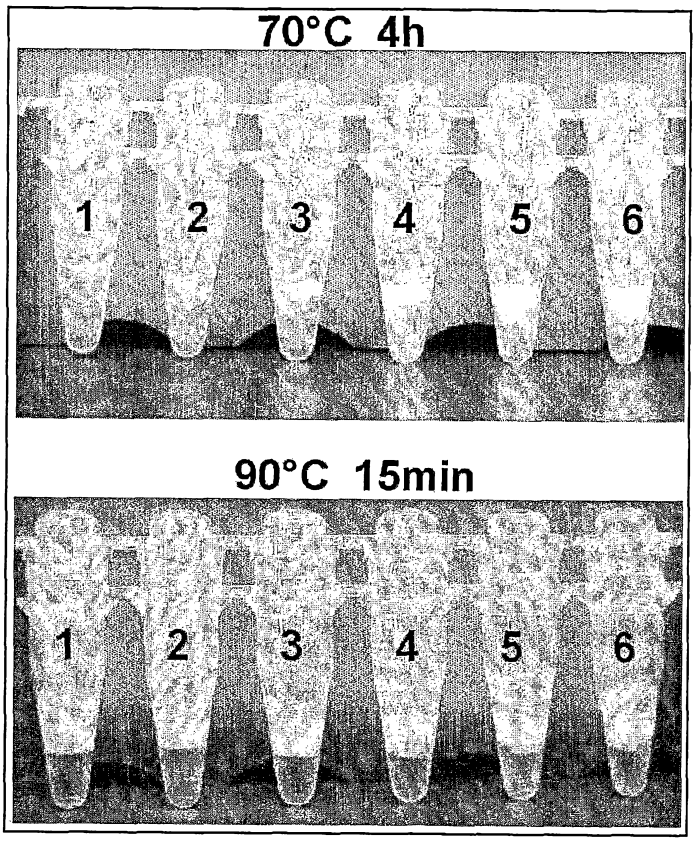
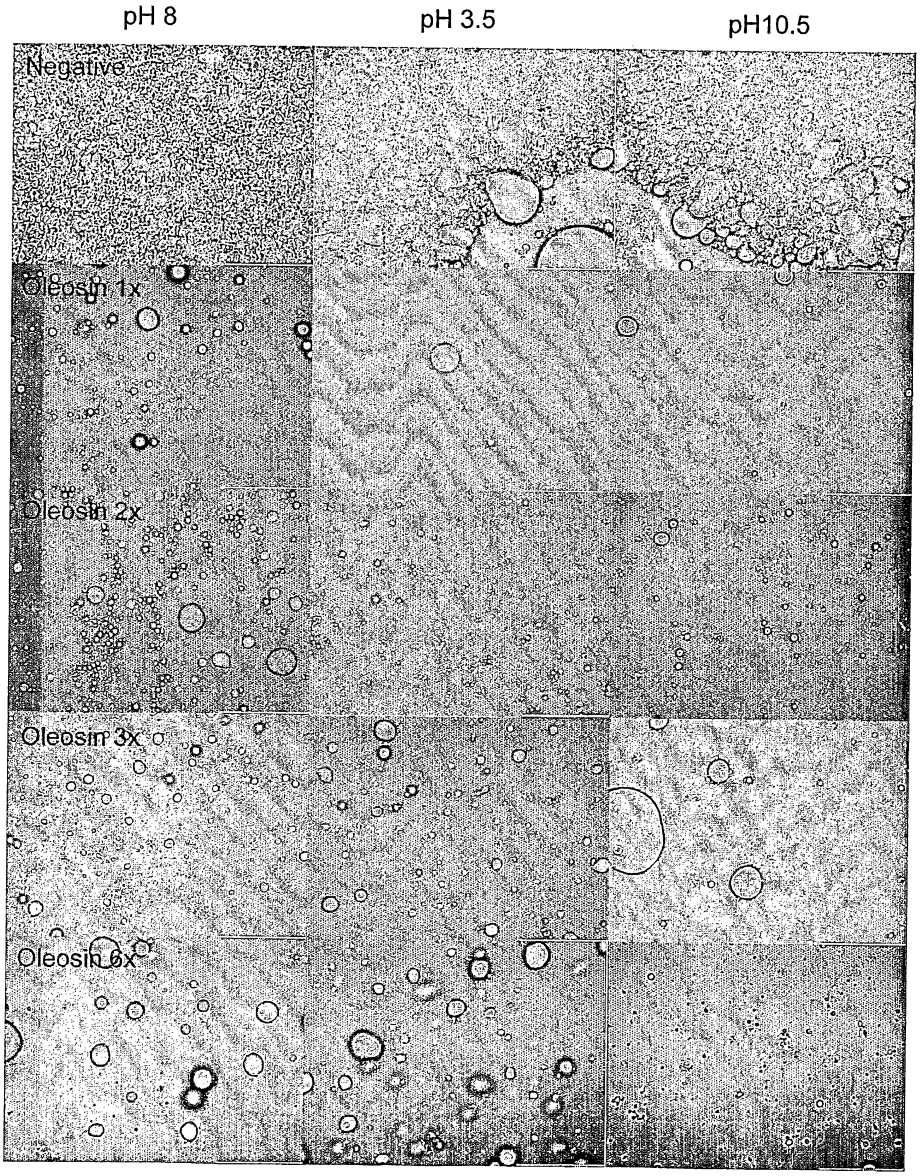


Figure 94


pH 8
pH 3.5
pH10.5
Negative
Oleosin 1x
Oleosin 2x
Oleosin 3x
Oleosin 6x

5

Figure 95

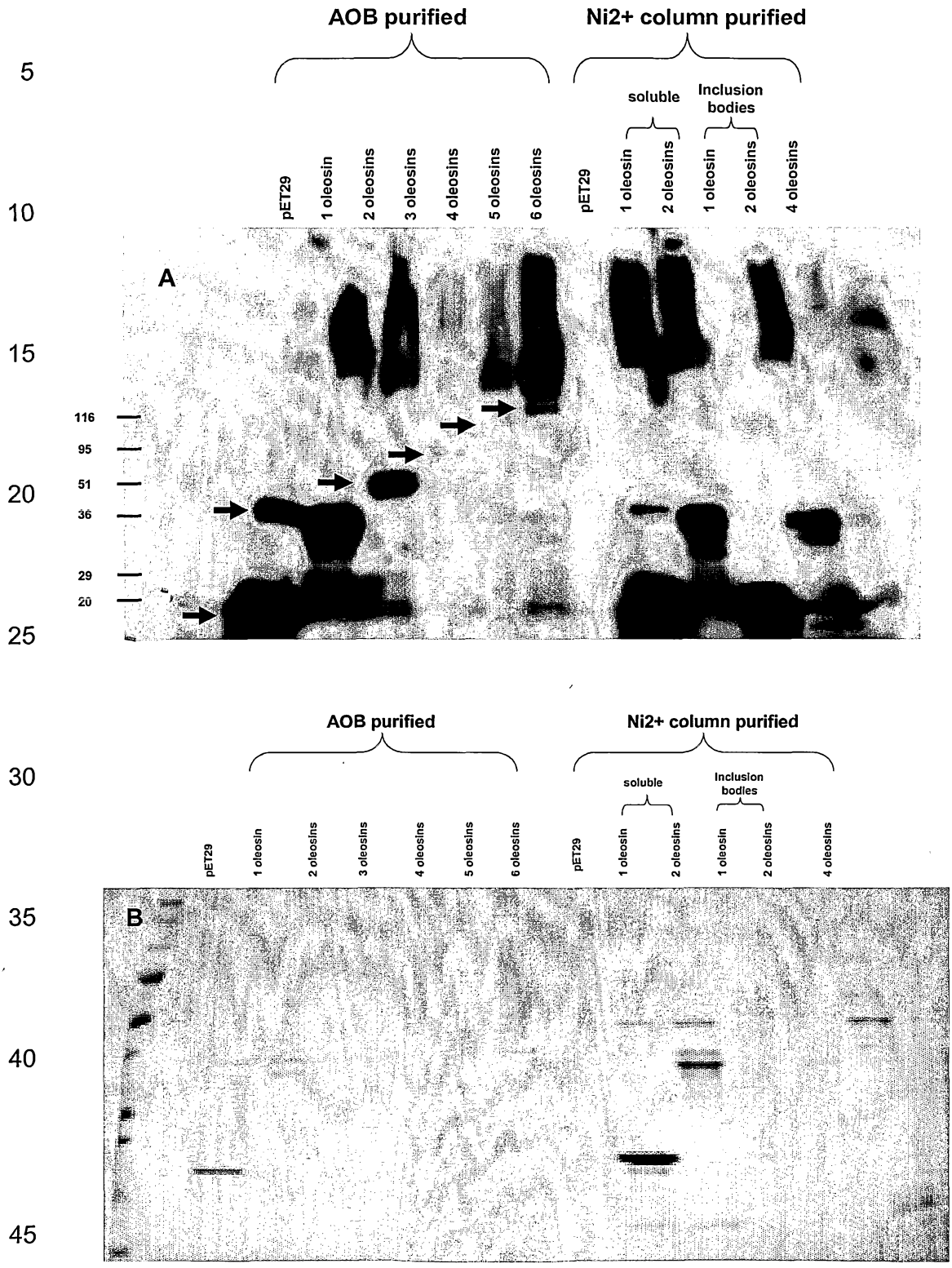

AOB purified
Ni2+ column purified
soluble
Inclusion bodies
pET29
1 oleosin
2 oleosins
3 oleosins
4 oleosins
5 oleosins
6 oleosins
A
116
95
51
36
29
20
B

Figure 96

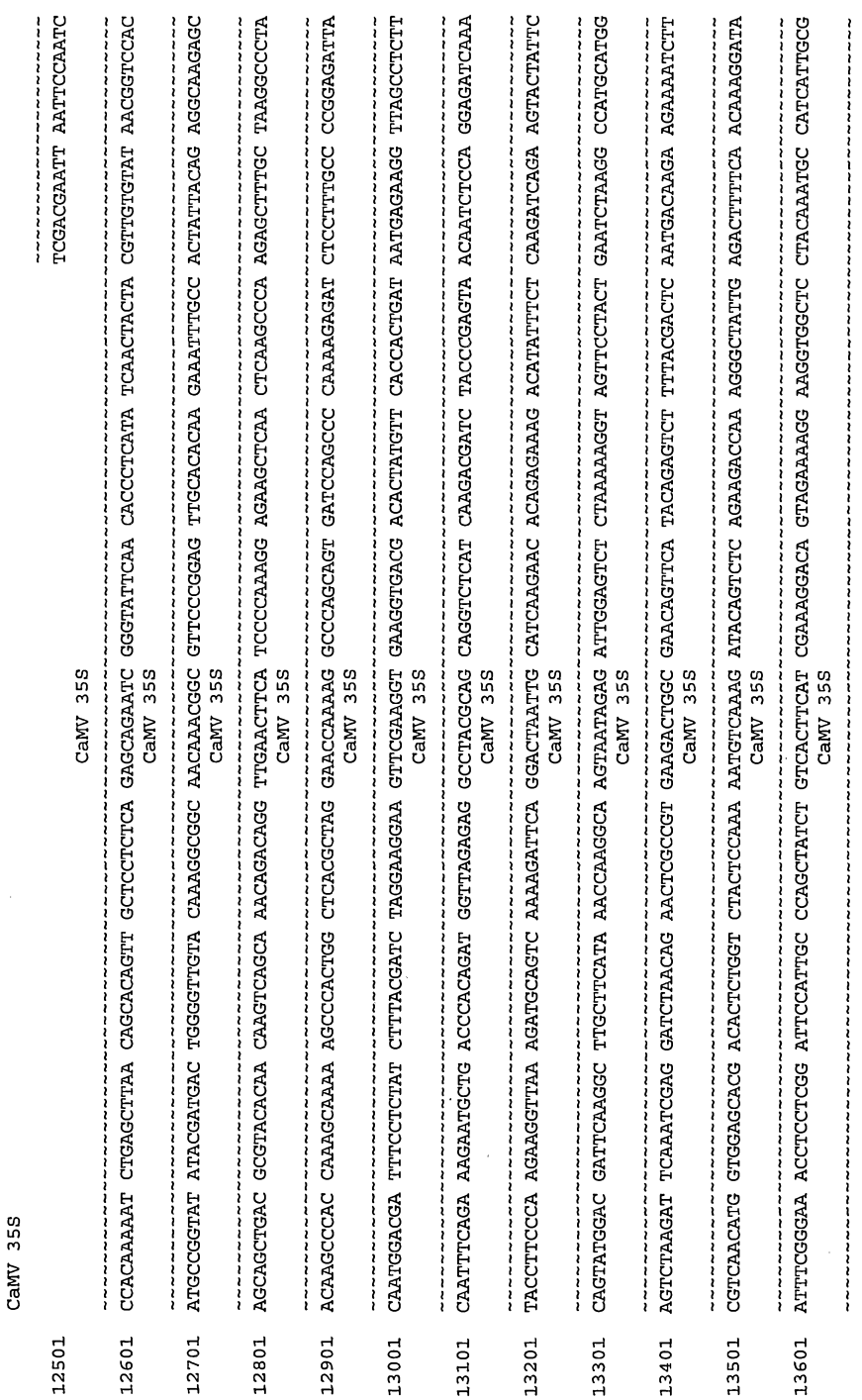

CaMV 35S
12501
CaMV 35S
12601
CaMV 35S
12701
CaMV 35S
12801
CaMV 35S
12901
CaMV 35S
13001
CaMV 35S
13101
CaMV 35S
13201
CaMV 35S
13301
CaMV 35S
13401
CaMV 35S
13501
CaMV 35S
13601
CaMV 35S

Figure 97

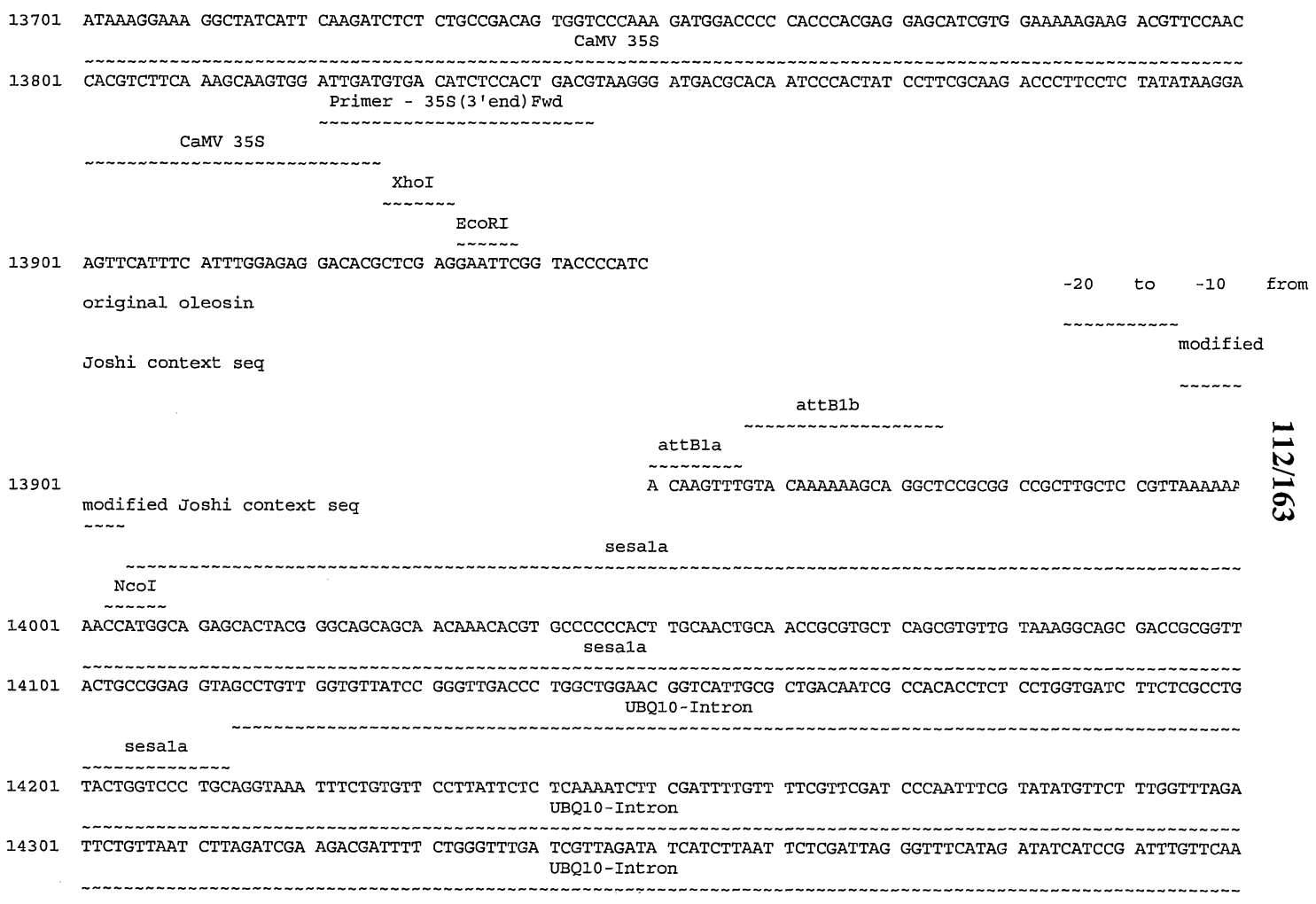


Figure 97 continued

```
14401  ATAATTTGAG TTTTGTCGAA TAATTACTCT TCGATTTGTG ATTTCTATCT AGATCTGGTG TTAGTTTCTA GTTTGTGCGA TCGAATTTGT CGATTAATCT
          UBQ10-Intron
       ~~~~~~~~~~~~~~~~~~~~~~
                                    sesa1b
                                ~~~~~~~~~~~~~~~~~~~
```

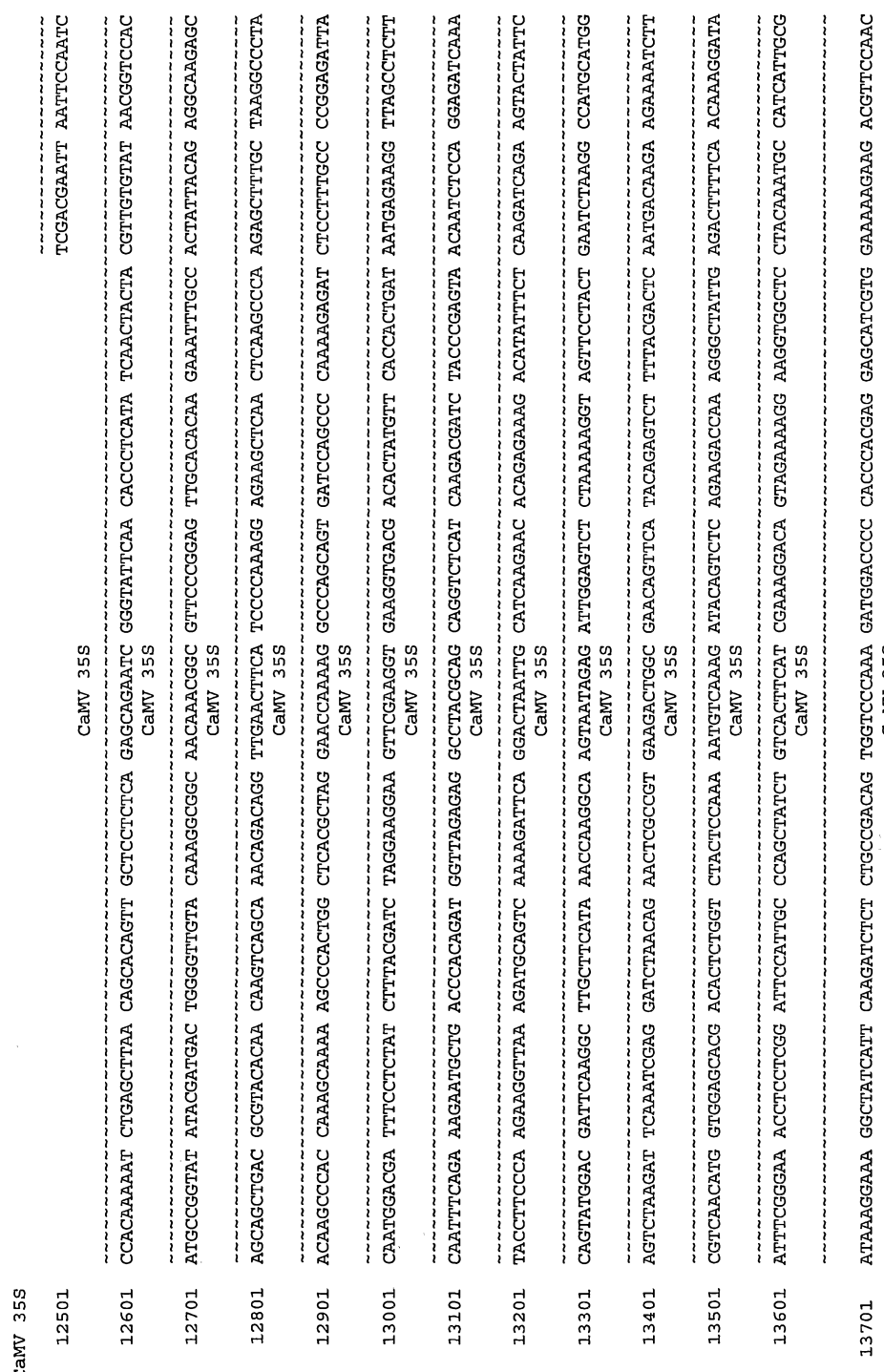

CaMV 35S
12501
TCGACGAATT AATTCCAATC
CaMV 35S
12601
CCACAAAAAT CTGAGCTTAA CAGCACAGTT GCTCCTCTCA GAGCAGAATC GGGTATTCAA CACCCTCATA TCAACTACTA CGTTGTGTAT AACGGTCCAC
CaMV 35S
12701
ATGCCGGTAT ATACGATGAC TGGGGTTGTA CAAAGGCGGC AACAAACGGC GTTCCCGGAG TTGCACACAA GAAATTTGCC ACTATTACAG AGGCAAGAGC
CaMV 35S
12801
AGCAGCTGAC GCGTACACAA CAAGTCAGCA AACAGACAGG TTGAACTTCA TCCCCAAAGG AGAAGCTCAA CTCAAGCCCA AGAGCTTTGC TAAGGCCCTA
CaMV 35S
12901
ACAAGCCCAC CAAAGCAAAA AGCCCACTGG CTCACGCTAG GAACCAAAAG GCCCAGCAGT GATCCAGCCC CAAAAGAGAT CTCCTTTGCC CCGGAGATTA
CaMV 35S
13001
CAATGGACGA TTTCCTCTAT CTTTACGATC TAGGAAGGAA GTTCGAAGGT GAAGGTGACG ACACTATGTT CACCACTGAT AATGAGAAGG TTAGCCTCTT
CaMV 35S
13101
CAATTTCAGA AAGAATGCTG ACCCACAGAT GGTTAGAGAG GCCTACGCAG CAGGTCTCAT CAAGACGATC TACCCGAGTA ACAATCTCCA GGAGATCAAA
CaMV 35S
13201
TACCTTCCCA AGAAGGTTAA AGATGCAGTC AAAAGATTCA GGACTAATTG CATCAAGAAC ACAGAGAAAG ACATATTTCT CAAGATCAGA AGTACTATTC
CaMV 35S
13301
CAGTATGGAC GATTCAAGGC TTGCTTCATA AACCAAGGCA AGTAATAGAG ATTGGAGTCT CTAAAAAGGT AGTTCCTACT GAATCTAAGG CCATGCATGG
CaMV 35S
13401
AGTCTAAGAT TCAAATCGAG GATCTAACAG AACTCGCCGT GAAGACTGGC GAACAGTTCA TACAGAGTCT TTTACGACTC AATGACAAGA AGAAAATCTT
CaMV 35S
13501
CGTCAACATG GTGGAGCACG ACACTCTGGT CTACTCCAAA AATGTCAAAG ATACAGTCTC AGAAGACCAA AGGGCTATTG AGACTTTTCA ACAAAGGATA
CaMV 35S
13601
ATTTCGGGAA ACCTCCTCGG ATTCCATTGC CCAGCTATCT GTCACTTCAT CGAAAGGACA GTAGAAAAGG AAGGTGGCTC CTACAAATGC CATCATTGCG
CaMV 35S
13701
ATAAAGGAAA GGCTATCATT CAAGATCTCT CTGCCGACAG TGGTCCCAAA GATGGACCCC CACCCACGAG GAGCATCGTG GAAAAAGAAG ACGTTCCAAC
CaMV 35S

Figure 98

```
                                                
```

```
14501  GAGTTTTTCT GATCTGCAGT CATCACCATC TTCCTGTTGG GAGCGGGTTT TCTGGCAAGC GGGGGATTTG GTGTTGCCGC TCTGTCTGTG CTGTCCTGGA
                                                  sesa1b

14601  TCTACCGTTA CCTGACGGGG AAACATCCAC CCGGAGCGGA TCAGTTGGAG TCGGCCAAGA CAAAGTTGGC GAGCAAGGCC CGTGAAATGA AGGACCGTGC
            sesa1b                                                                          sesa5

14701  CGAGCAGTTC AGTCAGCAAC CGGTAGCAGG GTCTCAGACC AGCATCGATC CATCCTCCTG GCATATGATG GCCGAGCACT ATGGACAACA GCAGCAGACG
                                                  sesa5

14801  CGTGCCCCTC ATCTGCAACT GCAACCCCGT GCTCAACGTG TCGTTAAGGC TGCGACAGCA GTAACCGCTG GGGGTTCTCT GTTAGTGTTG TCAGGGCTGA
                                                  sesa5

14901  CTTTGGCGGG GACGGTAATT GCGTTGACCA TTGCCCACCCC GCTGTTAGTG ATTTTCAGCC CGGTACTGGT GCCAGCAGTT ATCACGATCT TCTTGCTGGG
                                                  sesa5

15001  TGCCGGATTC TTGGCAAGTG GAGGTTTTGG AGTTGCGGCG CTGTCAGTTT TATCCTGGAT CTATCGTTAT CTGACAGGAA AACATCCCCC AGGTGCCGAT
                                        sesa5

15101  CAGCTGGAGA GTGCCAAGAC AAAACTGGCG TCTAAGGCAC GTGAAATGAA GGATCGTGCC GAACAGTTTT CTCAACAGCC CGTAGCGGGG TCACAGACCT
                                                  sesa6

15201  CGATCGATCA GCAGGTTAAC GTGCACATGG CCGAACATTA CGGACAGCAA CAACAGACGC GTGCTCCACA CCTGCAATTG CAACCGCGTG CTCAACGTGT
                                                  sesa6

15301  TGTCAAAGCG GCGACCGCCG TAACAGCAGG AGGATCACTG TTAGTGCTGT CGGGTTTAAC CTTGGCCGGG ACCGTCATTG CATTGACTAT TGCGACGCCC
                                                  sesa6

15401  TTACTGGTGA TCTTTTCTCC GGTGCTGGTT CCCGCCGTTA TTACCATCTT CTTGTTAGGG GCAGGATTCC TGGCATCAGG GGGATTCGGA GTTGCGGCGT
                                                  sesa6

15501  TGAGTGTCTT AAGTTGGATC TACCGTTATC TGACTGGAAA GCACCCGCCT GGGGCCGATC AACTGGAGTC AGCCAAAACG AAATTGGCGT CAAAAGCGCG
                  sesa6                                                                   4xSTOP

15601  TGAAATGAAG GACCGTGCTG AGCAGTTTTC TCAGCAGCCT GTGGCAGGAT CCCAGACATC ACCATGGCTC GAGTAATGAA GCGGCCGCAC CCAGCTTT
         attB2b                                                                                                attB2a

     1  CTTGTACAAA GTGGT
```

Figure 98 continued

```
CaMV 35S

                                                
```

```
13701  ATAAAGGAAA GGCTATCATT CAAGATCTCT CTGCCGACAG TGGTCCCAAA GATGGACCCC CACCCACGAG GAGCATCGTG GAAAAAGAAG ACGTTCCAAC
                                                    CaMV 35S
       ~~~
```

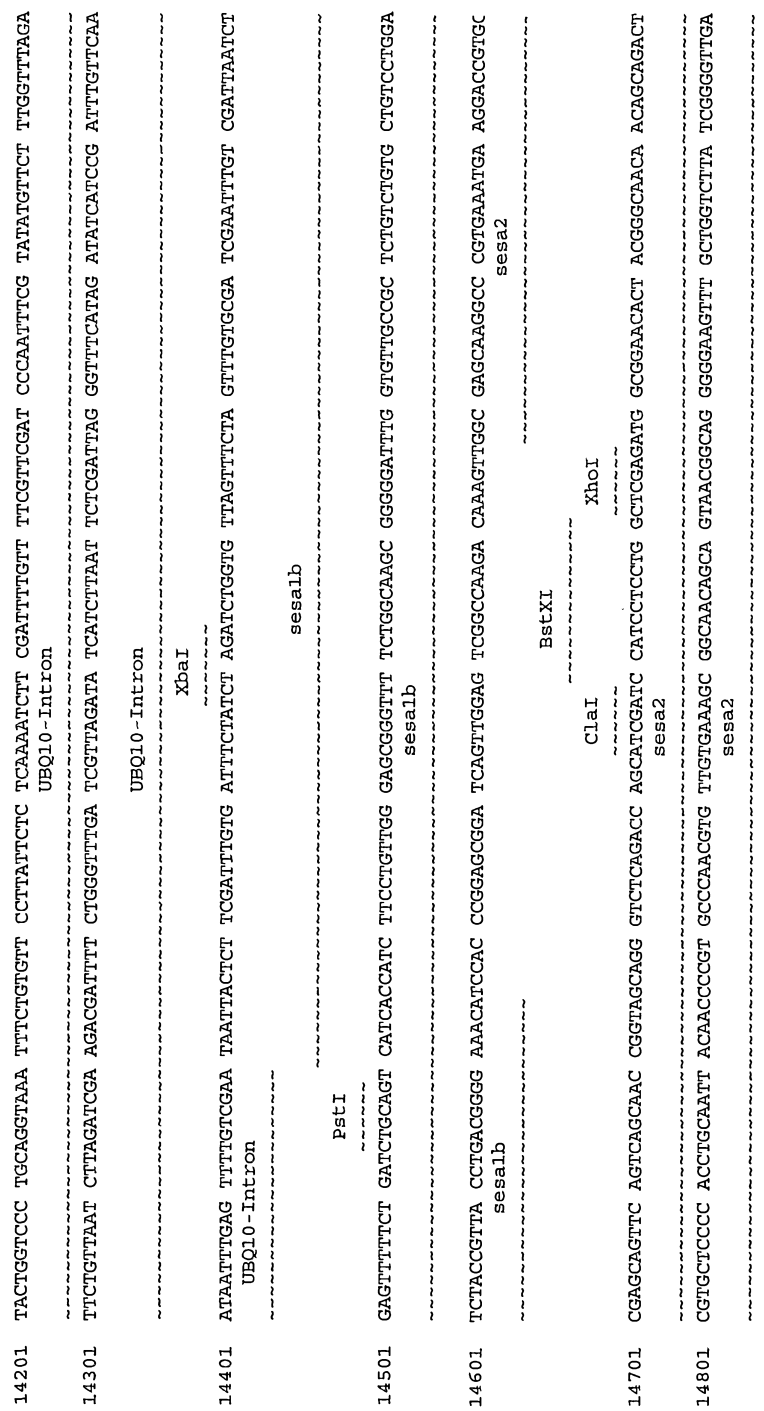


Figure 99 continued

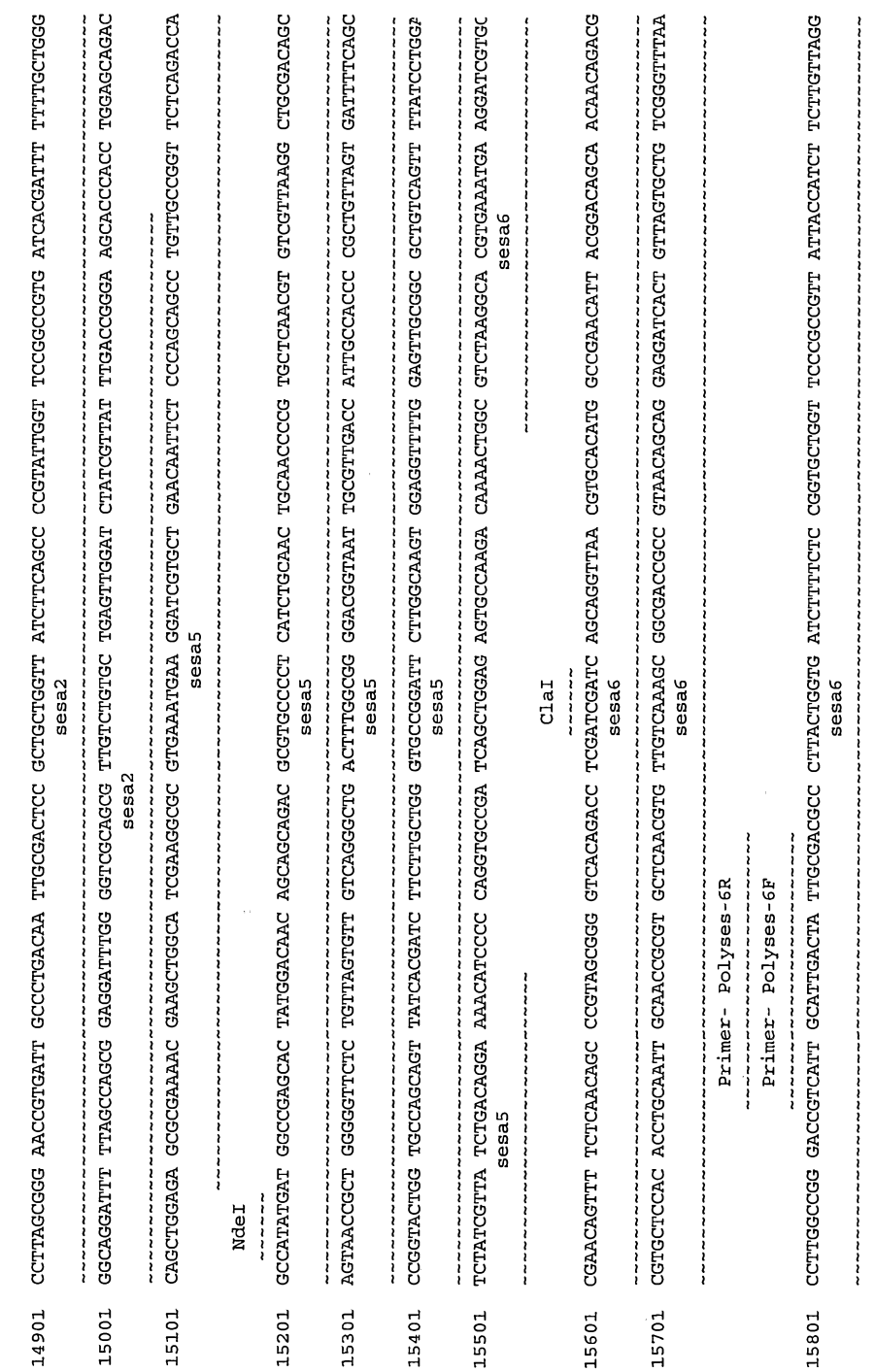


Figure 99 continued

```
15901  GGCAGGATTC CTGGCATCAG GGGGATTCGG AGTTGCGGCG TTGAGTGTCT TAAGTTGGAT CTACCGTTAT CTGACTGGAA AGCACCCGCC TGGGGCCGAT
                                    sesa6
       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
                                                                                                  BamHI
                                                                                                  ~~~~~~~
16001  CAACTGGAGT CAGCCAAAAC GAAATTGGCG TCAAAAGCGC GTGAAAATGAA GGACCGTGCT GAGCAGTTTT CTCAGCAGCC TGTGGCAGGA TCCCAGACAT
                                4xSTOP
                                ~~~~~
                          XhoI       NotI
                          ~~~~~~~    ~~~~~~~~
                    NcoI             attB2a
                    ~~~~~~           ~~~~~~~~~~~
16101  CACCATGGCT CGAGTAATGA AGCGGCCGCA CCCAGCTTT          XbaI
                                        HindIII       ~~~~~~~
                                        ~~~~~~
                                  attB2b      ClaI      BamHI
                                  ~~~~~~~~~~~~~~~~~~    ~~~~~~~
                                1 CTTGTACAAA GTGGTGATGG GTTCGAAATC GATAAGCTTG GATCCTCTAG A
```

Figure 99 continued

```
                                                
```

Figure 100 continued

```
13801  CACGTCTTCA AAGCAAGTGG ATTGATGTGA CATCTCCACT GACGTAAGGG ATGACGCACA ATCCCACTAT CCTTCGCAAG ACCCTTCCTC TATATAAGGA
                          Primer - 35S(3'end)Fwd
                         ~~~~~~~~~~~~~~~~~~~~~~~~~~
               CaMV 35S
       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~
                                 XhoI
                                ~~~~~~~
                                       EcoRI
                                       ~~~~~~
13901  AGTTCATTTC ATTTGGAGAG GACACGCTCG AGGAATTCGG TACCCCATC
                                                                                                  -20    to    -10    from
       original oleosin
                                                                                                  ~~~~~~~~~~~~
                                                                                                             modified
       Joshi context seq
                                                                                                             ~~~~~~
                                                                                attB1b
                                                                               ~~~~~~~~~~~~~~~~~~~~
                                                                        attB1a                          NotI
                                                                       ~~~~~~~~~                       ~~~~~~~~~
13901                                                                  A CAAGTTTGTA CAAAAAAGCA GGCTCCGCGG CCGCTTGCTC CGTTAAAAAA
       modified Joshi context seq
       ~~~~
                                                       sesa1a
           ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
14001  AACCATGGCA GAGCACTACG GGCAGCAGCA ACAAACACGT GCCCCCCACT TGCAACTGCA ACCGCGTGCT CAGCGTGTTG TAAAGGCAGC GACCGCGGTT
                                                  sesa1a
       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
14101  ACTGCCGGAG GTAGCCTGTT GGTGTTATCC GGGTTGACCC TGGCTGGAAC GGTCATTGCG CTGACAATCG CCACACCTCT CCTGGTGATC TTCTCGCCTG
                                                        UBQ10-Intron
                  ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
           sesa1a
       ~~~~~~~~~~~~~~~
14201  TACTGGTCCC TGCAGGTAAA TTTCTGTGTT CCTTATTCTC TCAAAATCTT CGATTTTGTT TTCGTTCGAT CCCAATTTCG TATATGTTCT TTGGTTTAGA
                                          UBQ10-Intron
       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
14301  TTCTGTTAAT CTTAGATCGA AGACGATTTT CTGGGTTTGA TCGTTAGATA TCATCTTAAT TCTCGATTAG GGTTTCATAG ATATCATCCG ATTTGTTCAA
                                          UBQ10-Intron
       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
14401  ATAATTTGAG TTTTGTCGAA TAATTACTCT TCGATTTGTG ATTTCTATCT AGATCTGGTG TTAGTTTCTA GTTTGTGCGA TCGAATTTGT CGATTAATCT
          UBQ10-Intron
       ~~~~~~~~~~~~~~~~~~~~~
                                                    sesa1b
                              ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
```

```
14501  GAGTTTTTCT GATCTGCAGT CATCACCATC TTCCTGTTGG GAGCGGGTTT TCTGGCAAGC GGGGGATTTG GTGTTGCCGC TCTGTCTGTG CTGTCCTGGA
                                                  sesa1b
14601  TCTACCGTTA CCTGACGGGG AAACATCCAC CCGGAGCGGA TCAGTTGGAG TCGGCCAAGA CAAAGTTGGC GAGCAAGGCC CGTGAAATGA AGGACCGTGC
             sesa1b                                                                    sesa2
14701  CGAGCAGTTC AGTCAGCAAC CGGTAGCAGG GTCTCAGACC AGCATCGATC CATCCTCCTG GCTCGAGATG GCGGAACACT ACGGCAACA ACAGCAGACT
                                                  sesa2
14801  CGTGCTCCCC ACCTGCAATT ACAACCCCGT GCCCAACGTG TTGTGAAAGC GGCAACAGCA GTAACGGCAG GGGGAAGTTT GCTGGTCTTA TCGGGGTTGA
                                                  sesa2
14901  CCTTAGCGGG AACCGTGATT GCCCTGACAA TTGCGACTCC GCTGCTGGTT ATCTTCAGCC CCGTATTGGT TCCGGCCGTG ATCACGATTT TTTTGCTGGG
                                                  sesa2
15001  GGCAGGATTT TTAGCCAGCG GAGGATTTGG GGTCGCAGCG TTGTCTGTGC TGAGTTGGAT CTATCGTTAT TTGACCGGGA AGCACCCACC TGGAGCAGG
                                    sesa2
15101  CAGCTGGAGA GCGCGAAAAC GAAGCTGGCA TCGAAGGCGC GTGAAATGAA GGATCGTGCT GAACAATTCT CCCAGCAGCC TGTTGCCGGT TCTCAGACC
                                                  sesa3
15201  GCCATATGTT TAAATGGCCA AGCGCTATGG CCGAGCATTA TGGGCAGCAA CAGCAAACCC GTGCCCCGCA TCTGCAATTG CAACCTCGTG CCCAGCGTGT
                                                  sesa3
15301  CGTTAAGGCG GCTACTGCGG TAACAGCGGG AGGGAGCTTA CTGGTATTAA GCGGGCTGAC ATTGGCCGGA ACGGTGATCG CCTTAACAAT CGCGACACCC
                                                  sesa3
15401  TTGCTGGTCA TCTTCAGTCC GGTTCTGGTG CCCGCGGTGA TTACGATTTT CCTGCTGGGA GCCGGTTTCT TAGCATCGGG GGGTTTTGGG GTAGCAGCCT
                                                  sesa3
15501  TGAGTGTCCT GTCGTGGATC TATCGTTACT TAACTGGAAA ACACCCGCCA GGAGCTGACC AGTTGGAGTC TGCAAAAACT AAGCTGGCGT CCAAAGCCCG
                                    sesa3                                                                 sesa4
15601  TGAAATGAAG GATCGTGCTG AGCAGTTTAG CCAGCAGCCA GTTGCGGGAA GTCAGACCTC TTCATCTGAG CTCCCATGGG TCGACATGGC GGAGCATTAC
                                                  sesa4
15701  GGTCAACAGC AACAGACCCG TGCTCCGCAC TTACAATTGC AACCACGTGC TCAACGTGTC GTAAAAGCCG CCACGGCAGT TACTGCGGGG GGATCATTGC
                                                  sesa4
15801  TGGTGTTAAG TGGGTTGACA CTGGCGGGGA CAGTTATTGC ACTGACGATC GCGACCCCCT TGTTAGTGAT CTTCTCCCCC GTTCTGGTTC CGGCGGTCAT
                                                  sesa4
```

Figure 100 continued

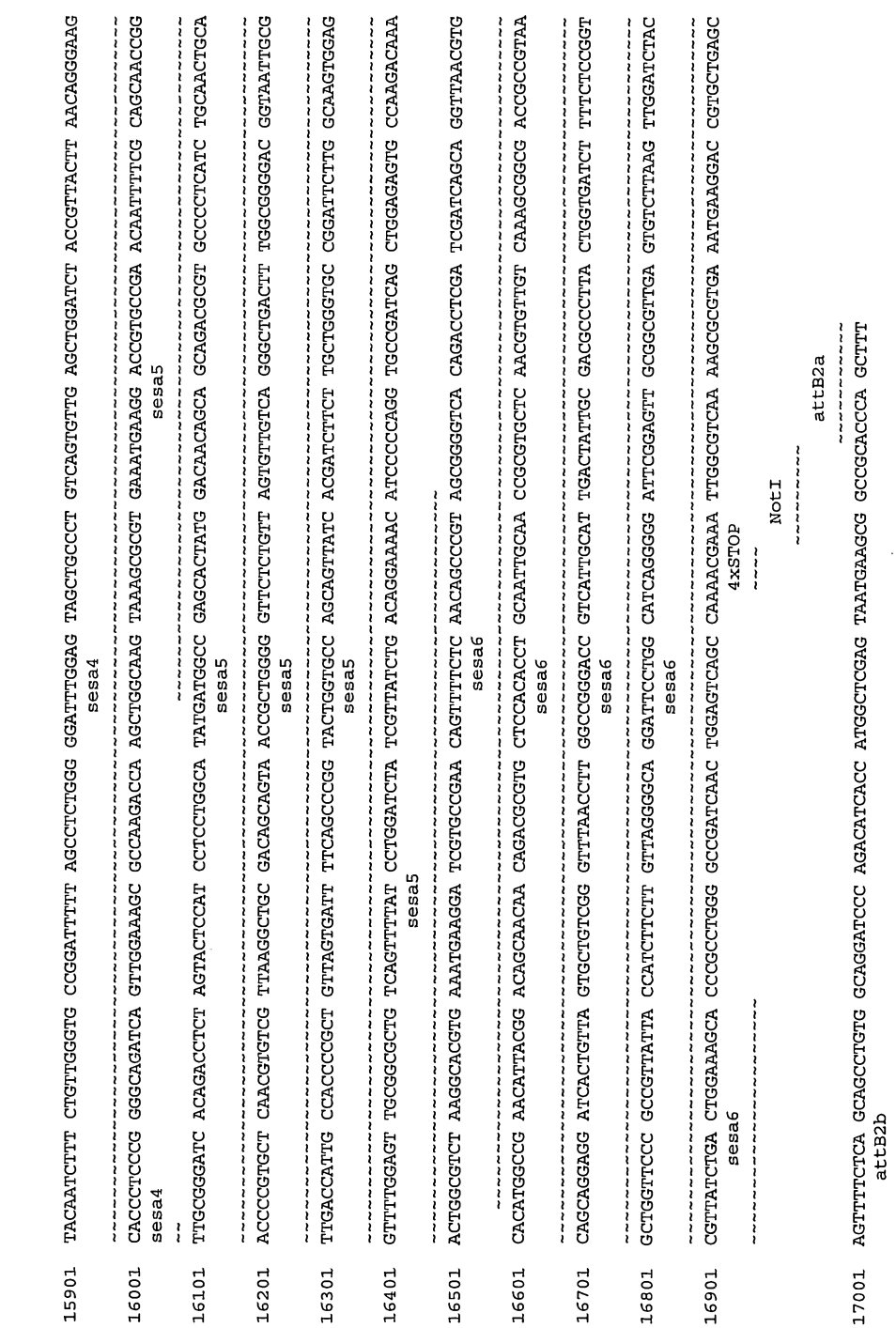


Figure 100 continued

```
                                                
```

```
                                                
```

```
14601  TCTACCGTTA CCTGACGGGG AAACATCCAC CCGGAGCGGA TCAGTTGGAG TCGGCCAAGA CAAAGTTGGC GAGCAAGGCC CGTGAAATGA AGGACCGTGC
           sesa1b                                                          4xSTOP
       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~                                      ~~~~
                                                                                 attB2a
                                                                                 ~~~~~~~~~~~
14701  CGAGCAGTTC AGTCAGCAAC CGGTAGCAGG GTCTCAGACC AGCATCGATC CATCCTCCTG GCTCGAGTAA TGAAGCGGCC GCACCCAGCT TT
       attB2b
       ~~~~~~~~~~~~~~~~~
           1  CTTGTACAAA GTGGT
```

Figure 101 continued

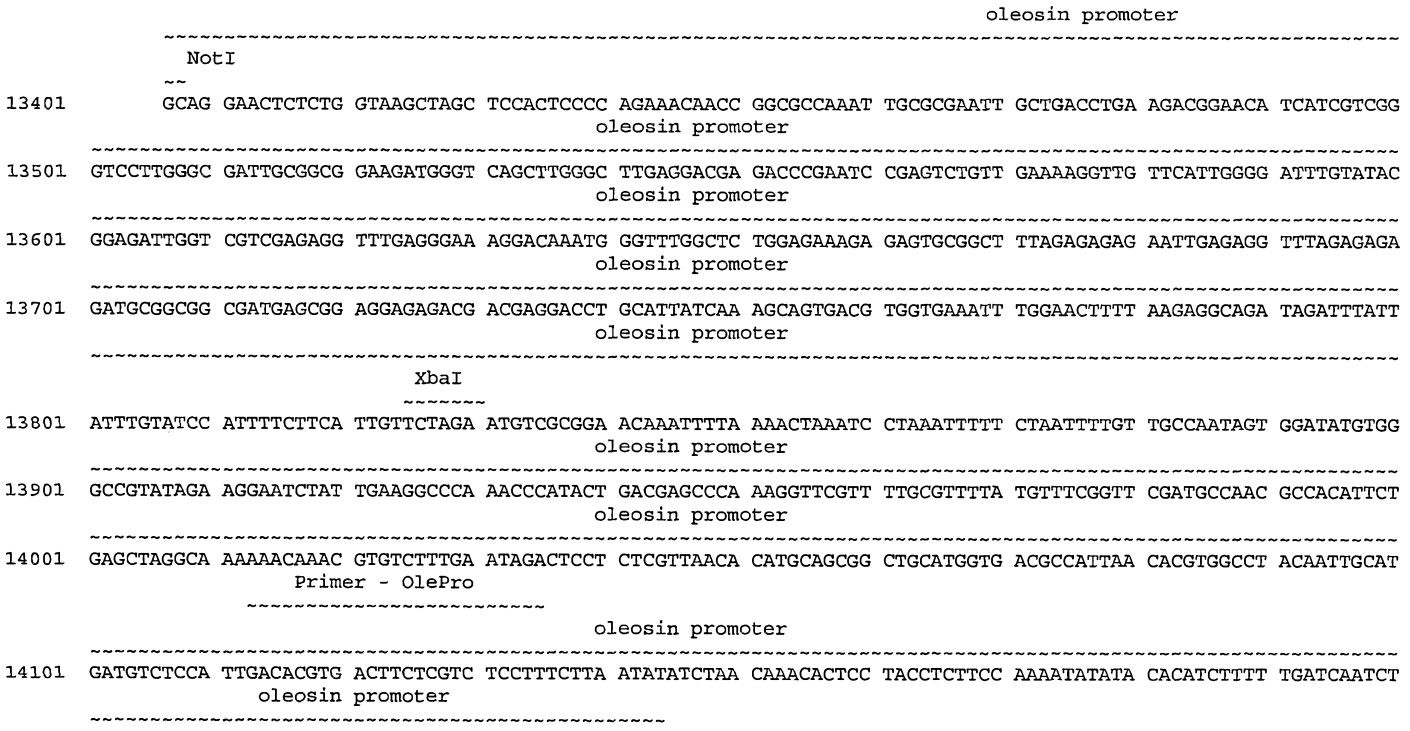

oleosin promoter
NotI
13401 GCAG GAACTCTCTG GTAAGCTAGC TCCACTCCCC AGAAACAACC GGCGCCAAAT TGCGCGAATT GCTGACCTGA AGACGGAACA TCATCGTCGG
oleosin promoter
13501 GTCCTTGGGC GATTGCGGCG GAAGATGGGT CAGCTTGGGC TTGAGGACGA GACCCGAATC CGAGTCTGTT GAAAAGGTTG TTCATTGGGG ATTTGTATAC
oleosin promoter
13601 GGAGATTGGT CGTCGAGAGG TTTGAGGGAA AGGACAAATG GGTTTGGCTC TGGAGAAAGA GAGTGCGGCT TTAGAGAGAG AATTGAGAGG TTTAGAGAGA
oleosin promoter
13701 GATGCGGCGG CGATGAGCGG AGGAGAGACG ACGAGGACCT GCATTATCAA AGCAGTGACG TGGTGAAATT TGGAACTTTT AAGAGGCAGA TAGATTTATT
oleosin promoter
XbaI
13801 ATTTGTATCC ATTTTCTTCA TTGTTCTAGA ATGTCGCGGA ACAAATTTTA AAACTAAATC CTAAATTTTT CTAATTTTGT TGCCAATAGT GGATATGTGG
oleosin promoter
13901 GCCGTATAGA AGGAATCTAT TGAAGGCCCA AACCCATACT GACGAGCCCA AAGGTTCGTT TTGCGTTTTA TGTTTCGGTT CGATGCCAAC GCCACATTCT
oleosin promoter
14001 GAGCTAGGCA AAAAACAAAC GTGTCTTTGA ATAGACTCCT CTCGTTAACA CATGCAGCGG CTGCATGGTG ACGCCATTAA CACGTGGCCT ACAATTGCAT
Primer - OlePro
oleosin promoter
14101 GATGTCTCCA TTGACACGTG ACTTCTCGTC TCCTTTCTTA ATATATCTAA CAAACACTCC TACCTCTTCC AAAATATATA CACATCTTTT TGATCAATCT
oleosin promoter

Figure 102

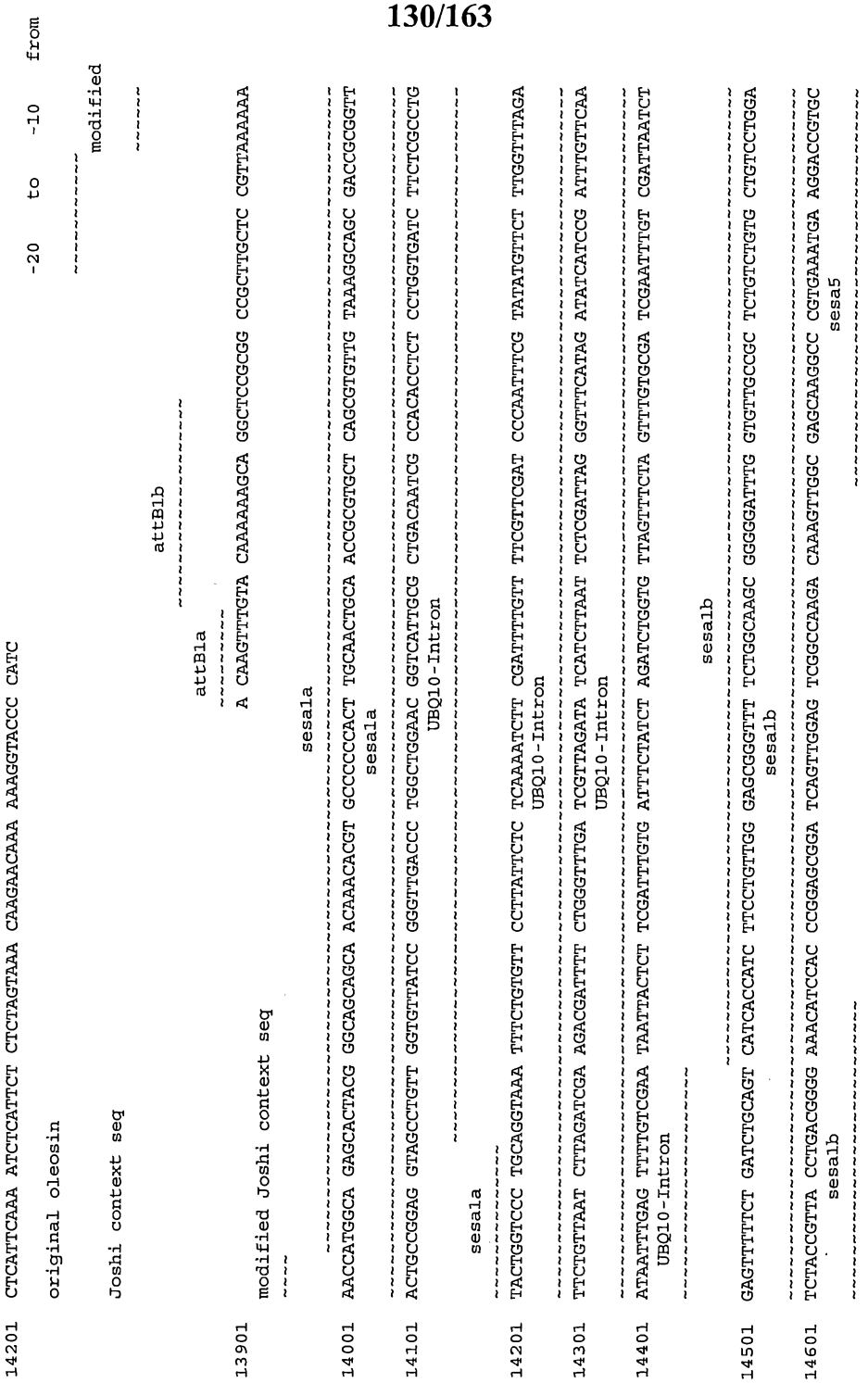


Figure 102 continued

```
14701  CGAGCAGTTC AGTCAGCAAC CGGTAGCAGG GTCTCAGACC AGCATCGATC CATCCTCCTG GCATATGATG GCCGAGCACT ATGGACAACA GCAGCAGACG
                                          sesa5
       ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
14801  CGTGCCCCTC ATCTGCAACT GCAACCCCGT GCTCAACGTG TCGTTAAGGC TGCGACAGCA GTAACCGCTG GGGGTTCTCT GTTAGTGTTG TCAGGGCTGA
                                          sesa5
       ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
14901  CTTTGGCGGG GACGGTAATT GCGTTGACCA TTGCCACCCC GCTGTTAGTG ATTTTCAGCC CGGTACTGGT GCCAGCAGTT ATCACGATCT TCTTGCTGGG
                                          sesa5
       ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
15001  TGCCGGATTC TTGGCAAGTG GAGGTTTTGG AGTTGCGGCG CTGTCAGTTT TATCCTGGAT CTATCGTTAT CTGACAGGAA AACATCCCCC AGGTGCCGAT
                                    sesa5
       ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
15101  CAGCTGGAGA GTGCCAAGAC AAAACTGGCG TCTAAGGCAC GTGAAATGAA GGATCGTGCC GAACAGTTTT CTCAACAGCC CGTAGCGGGG TCACAGACCT
                                                sesa6
       ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
15201  CGATCGATCA GCAGGTTAAC GTGCACATGG CCGAACATTA CGGACAGCAA CAACAGACGC GTGCTCCACA CCTGCAATTG CAACCGCGTG CTCAACGTGT
                                          sesa6
       ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
15301  TGTCAAAGCG GCGACCGCCG TAACAGCAGG AGGATCACTG TTAGTGCTGT CGGGTTTAAC CTTGGCCGGG ACCGTCATTG CATTGACTAT TGCGACGCCC
                                          sesa6
       ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
15401  TTACTGGTGA TCTTTTCTCC GGTGCTGGTT CCCGCCGTTA TTACCATCTT CTTGTTAGGG GCAGGATTCC TGGCATCAGG GGGATTCGGA GTTGCGGCGT
                                          sesa6
       ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
15501  TGAGTGTCTT AAGTTGGATC TACCGTTATC TGACTGGAAA GCACCCGCCT GGGGCCGATC AACTGGAGTC AGCCAAAACG AAATTGGCGT CAAAAGCGCG
                                    sesa6                                                 4xSTOP
       ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~                              ~~~~~
                                                                                                  attB2a
                                                                                                  ~~~~~~~~~~
15601  TGAAATGAAG GACCGTGCTG AGCAGTTTTC TCAGCAGCCT GTGGCAGGAT CCCAGACATC ACCATGGCTC GAGTAATGAA GCGGCCGCAC CCAGCTTT
         attB2b
       ~~~~~~~~~~ ~~~~~~~~~~
           1  CTTGTACAAA GTGGT
```

Figure 102 continued

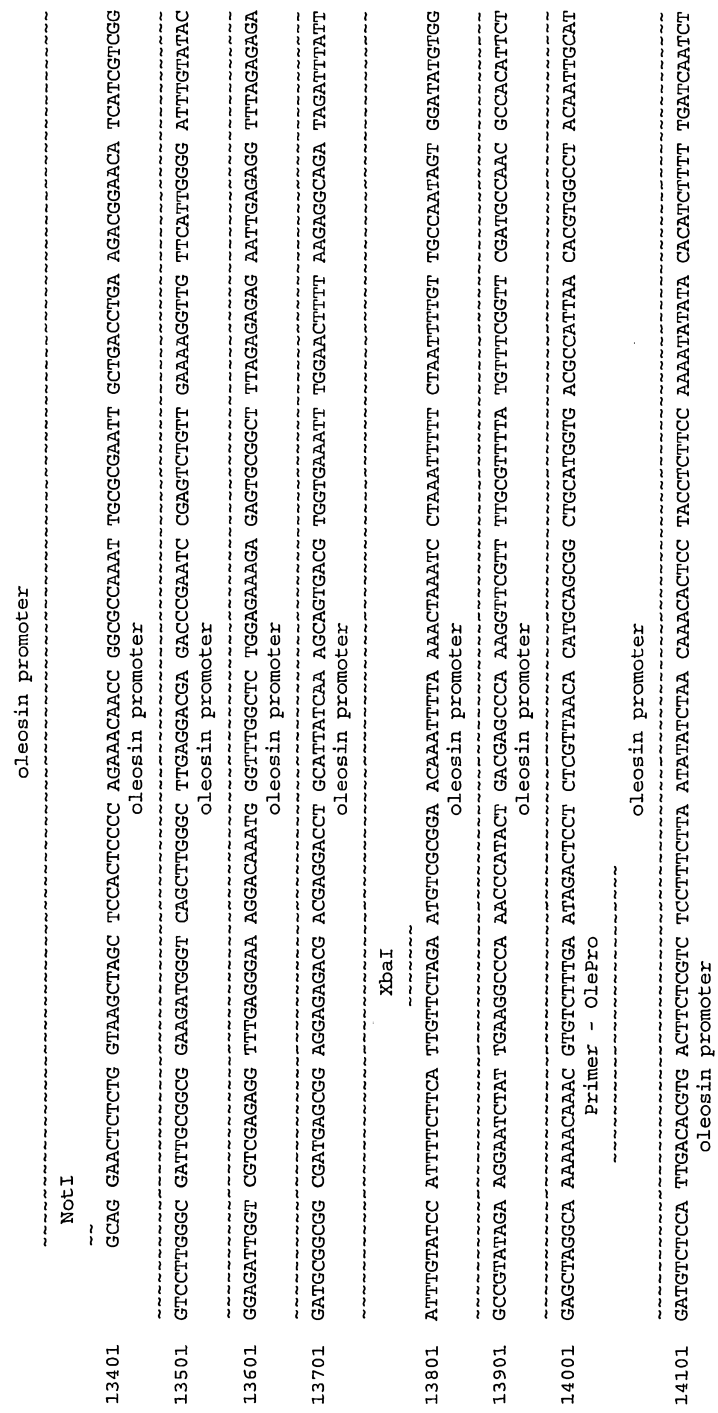

oleosin promoter
NotI
13401 GCAG GAACTCTCTG GTAAGCTAGC TCCACTCCCC AGAAACAACC GGCGCCAAAT TGCGCGAATT GCTGACCTGA AGACGGAACA TCATCGTCGG
oleosin promoter
13501 GTCCTTGGGC GATTGCGGCG GAAGATGGGT CAGCTTGGGC TTGAGGACGA GACCCGAATC CGAGTCTGTT GAAAAGGTTG TTCATTGGGG ATTTGTATAC
oleosin promoter
13601 GGAGATTGGT CGTCGACAGG TTTGAGGGAA AGGACAAATG GGTTTGGCTC TGGAGAAAGA GAGTGCGGCT TTAGAGAGAG AATTGACAGG TTTAGAGAGA
oleosin promoter
13701 GATGCGGCGG CGATGAGCGG AGGAGAGACG ACGAGGACCT GCATTATCAA AGCAGTGACG TGGTGAAATT TGGAACTTTT AAGAGGCAGA TAGATTTATT
oleosin promoter
XbaI
13801 ATTTGTATCC ATTTTCTTCA TTGTTCTAGA ATGTCGCGGA ACAAATTTTA AAACTAAATC CTAAATTTTT CTAATTTTGT TGCCAATAGT GGATATGTGG
oleosin promoter
13901 GCCGTATAGA AGGAATCTAT TGAAGGCCCA AACCCATACT GACGAGCCCA AAGGTTCGTT TTGCGTTTTA TGTTTCGGTT CGATGCCAAC GCCACATTCT
oleosin promoter
14001 GAGCTAGGCA AAAAACAAAC GTGTCTTTGA ATAGACTCCT CTCGTTAACA CATGCAGCGG CTGCATGGTG ACGCCATTAA CACGTGGCCT ACAATTGCAT
Primer - OlePro
oleosin promoter
14101 GATGTCTCCA TTGACACGTG ACTTCTCGTC TCCTTTCTTA ATATATCTAA CAAACACTCC TACCTCTTCC AAAATATATA CACATCTTTT TGATCAATCT
oleosin promoter

Figure 103

```
                                                
```



Figure 103 continued

```
15401  CCGGTACTGG TGCCAGCAGT TATCACGATC TTCTTGCTGG GTGCCGGATT CTTGGCAAGT GGAGGTTTTG GAGTTGCGGC GCTGTCAGTT TTATCCTGGA
                                          sesa5
       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
15501  TCTATCGTTA TCTGACAGGA AAACATCCCC CAGGTGCCGA TCAGCTGGAG AGTGCCAAGA CAAAACTGGC GTCTAAGGCA CGTGAAATGA AGGATCGTGC
            sesa5                                                                                 sesa6
       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
                                                ClaI
                                                ~~~~~~
15601  CGAACAGTTT TCTCAACAGC CCGTAGCGGG GTCACAGACC TCGATCGATC AGCAGGTTAA CGTGCACATG GCCGAACATT ACGGACAGCA ACAACAGACG
                                                  sesa6
       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
15701  CGTGCTCCAC ACCTGCAATT GCAACCGCGT GCTCAACGTG TTGTCAAAGC GGCGACCGCC GTAACAGCAG GAGGATCACT GTTAGTGCTG TCGGGTTTAA
                                          sesa6
       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
                                Primer- Polyses-6R
                                ~~~~~~~~~~~~~~~~~~~~~~~~~~
                                  Primer- Polyses-6F
                                  ~~~~~~~~~~~~~~~~~~~~~~~~~
15801  CCTTGGCCGG GACCGTCATT GCATTGACTA TTGCGACGCC CTTACTGGTG ATCTTTTCTC CGGTGCTGGT TCCCGCCGTT ATTACCATCT TCTTGTTAGC
                                                  sesa6
       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
15901  GGCAGGATTC CTGGCATCAG GGGGATTCGG AGTTGCGGCG TTGAGTGTCT TAAGTTGGAT CTACCGTTAT CTGACTGGAA AGCACCCGCC TGGGGCCGAT
                                          sesa6
       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
                                                                                        BamHI
                                                                                        ~~~~~~~
16001  CAACTGGAGT CAGCCAAAAC GAAATTGGCG TCAAAAGCGC GTGAAATGAA GGACCGTGCT GAGCAGTTTT CTCAGCAGCC TGTGGCAGGA TCCCAGACAT
                                4xSTOP
                                ~~~~~
                          XhoI       NotI
                          ~~~~~~~    ~~~~~~~~
                     NcoI                 attB2a
                     ~~~~~~               ~~~~~~~~~~
16101  CACCATGGCT CGAGTAATGA AGCGGCCGCA CCCAGCTTT
                                          HindIII      XbaI
                                          ~~~~~~       ~~~~~~~
                          attB2b                  ClaI       BamHI
                          ~~~~~~~~~~~~~~~~~~~~    ~~~~~~~    ~~~~~~~
     1    CTTGTACAAA GTGGTGATGG GTTCGAAATC GATAAGCTTG GATCCTCTAG A
```

Figure 103 continued

```
                                                oleosin promoter
                                                ~~~~~~~~~~~~~~~~
                                                
```

```
                                                
```

```
14801  CGTGCTCCCC ACCTGCAATT ACAACCCCGT GCCCAACGTG TTGTGAAAGC GGCAACAGCA GTAACGGCAG GGGGAAGTTT GCTGGTCTTA TCGGGGTTGA
                                                    sesa2
14901  CCTTAGCGGG AACCGTGATT GCCCTGACAA TTGCGACTCC GCTGCTGGTT ATCTTCAGCC CCGTATTGGT TCCGGCCGTG ATCACGATTT TTTGCTGGG
                                                    sesa2
15001  GGCAGGATTT TTAGCCAGCG GAGGATTTGG GGTCGCAGCG TTGTCTGTGC TGAGTTGGAT CTATCGTTAT TTGACCGGGA AGCACCCACC TGGAGCAGAC
                                                    sesa2
15101  CAGCTGGAGA GCGCGAAAAC GAAGCTGGCA TCGAAGGCGC GTGAAATGAA GGATCGTGCT GAACAATTCT CCCAGCAGCC TGTTGCCGGT TCTCAGACCA
                                                    sesa3
15201  GCCATATGTT TAAATGGCCA AGCGCTATGG CCGAGCATTA TGGGCAGCAA CAGCAAACCC GTGCCCCGCA TCTGCAATTG CAACCTCGTG CCCAGCGTGT
                                                    sesa3
15301  CGTTAAGGCG GCTACTGCGG TAACAGCGGG AGGGAGCTTA CTGGTATTAA GCGGGCTGAC ATTGGCCGGA ACGGTGATCG CCTTAACAAT CGCGACACCC
                                                    sesa3
15401  TTGCTGGTCA TCTTCAGTCC GGTTCTGGTG CCCGCGGTGA TTACGATTTT CCTGCTGGGA GCCGGTTTCT TAGCATCGGG GGGTTTTGGG GTAGCAGCCT
                                                    sesa3
15501  TGAGTGTCCT GTCGTGGATC TATCGTTACT TAACTGGAAA ACACCCGCCA GGAGCTGACC AGTTGGAGTC TGCAAAAACT AAGCTGGCGT CCAAAGCCCG
                                                    sesa3                                                sesa4
15601  TGAAATGAAG GATCGTGCTG AGCAGTTTAG CCAGCAGCCA GTTGCGGGAA GTCAGACCTC TTCATCTGAG CTCCCATGGG TCGACATGGC GGAGCATTAC
                                                    sesa4
15701  GGTCAACAGC AACAGACCCG TGCTCCGCAC TTACAATTGC AACCACGTGC TCAACGTGTC GTAAAAGCCG CCACGGCAGT TACTGCGGGG GGATCATTGC
                                                    sesa4
15801  TGGTGTTAAG TGGGTTGACA CTGGCGGGGA CAGTTATTGC ACTGACGATC GCGACCCCCT TGTTAGTGAT CTTCTCCCCC GTTCTGGTTC CGGCGGTCAT
                                                    sesa4
15901  TACAATCTTT CTGTTGGGTG CCGGATTTTT AGCCTCTGGG GGATTTGGAG TAGCTGCCCT GTCAGTGTTG AGCTGGATCT ACCGTTACTT AACAGGGAAG
                                                    sesa4
16001  CACCCTCCCG GGGCAGATCA GTTGGAAAGC GCCAAGACCA AGCTGGCAAG TAAAGCGCGT GAAATGAAGG ACCGTGCCGA ACAATTTTCG CAGCAACCGG
sesa4                                                                                    sesa5
16101  TTGCGGGATC ACAGACCTCT AGTACTCCAT CCTCCTGGCA TATGATGGCC GAGCACTATG GACAACAGCA GCAGACGCGT GCCCCTCATC TGCAACTGCA
```

Figure 104 continued

```
                                                
```

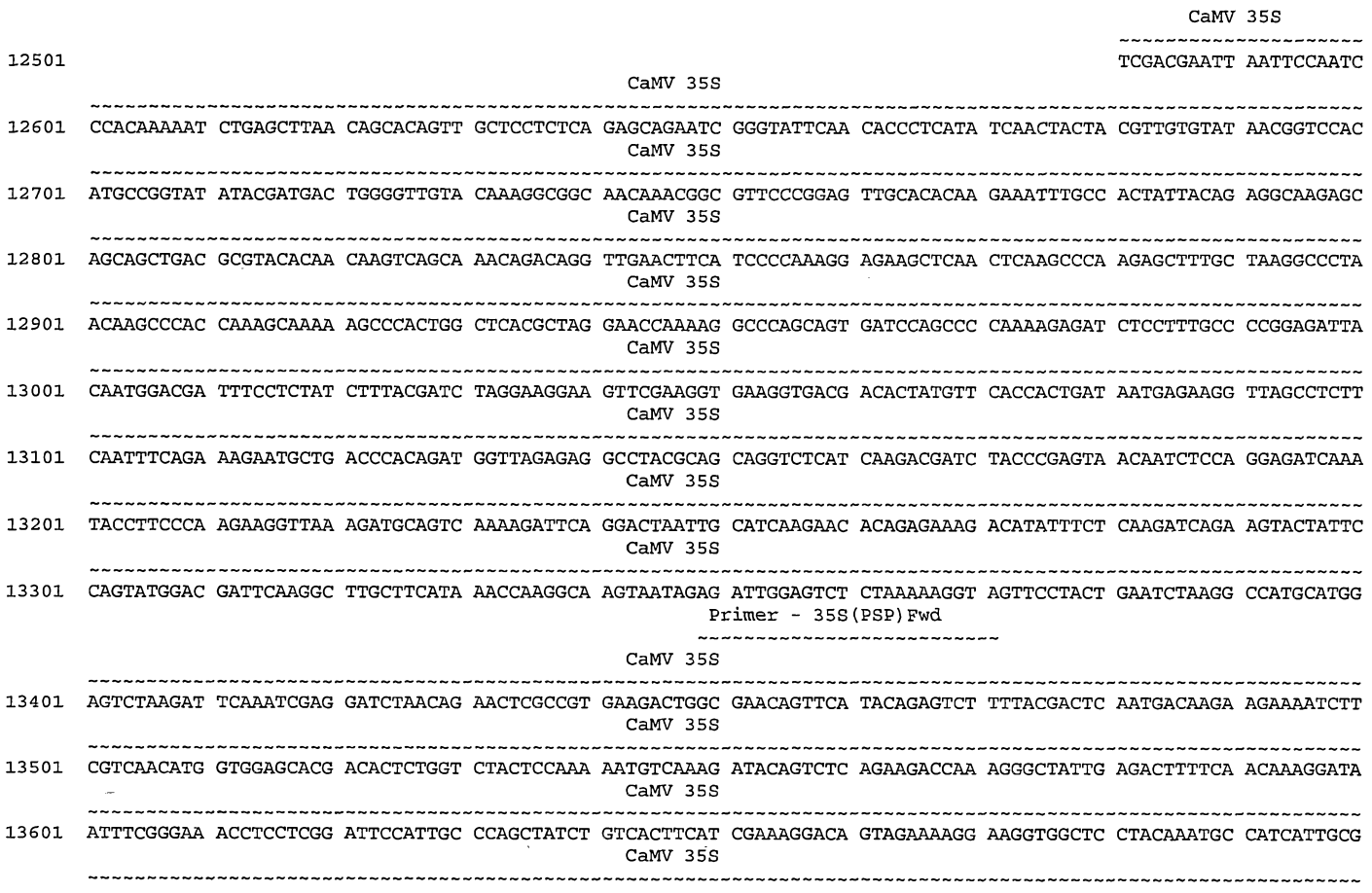

CaMV 35S
12501 TCGACGAATT AATTCCAATC
CaMV 35S
12601 CCACAAAAAT CTGAGCTTAA CAGCACAGTT GCTCCTCTCA GAGCAGAATC GGGTATTCAA CACCCTCATA TCAACTACTA CGTTGTGTAT AACGGTCCAC
CaMV 35S
12701 ATGCCGGTAT ATACGATGAC TGGGGTTGTA CAAAGGCGGC AACAAACGGC GTTCCCGGAG TTGCACACAA GAAATTTGCC ACTATTACAG AGGCAAGAGC
CaMV 35S
12801 AGCAGCTGAC GCGTACACAA CAAGTCAGCA AACAGACAGG TTGAACTTCA TCCCCAAAGG AGAAGCTCAA CTCAAGCCCA AGAGCTTTGC TAAGGCCCTA
CaMV 35S
12901 ACAAGCCCAC CAAAGCAAAA AGCCCACTGG CTCACGCTAG GAACCAAAAG GCCCAGCAGT GATCCAGCCC CAAAAGAGAT CTCCTTTGCC CCGGAGATTA
CaMV 35S
13001 CAATGGACGA TTTCCTCTAT CTTTACGATC TAGGAAGGAA GTTCGAAGGT GAAGGTGACG ACACTATGTT CACCACTGAT AATGAGAAGG TTAGCCTCTT
CaMV 35S
13101 CAATTTCAGA AAGAATGCTG ACCCACAGAT GGTTAGAGAG GCCTACGCAG CAGGTCTCAT CAAGACGATC TACCCGAGTA ACAATCTCCA GGAGATCAAA
CaMV 35S
13201 TACCTTCCCA AGAAGGTTAA AGATGCAGTC AAAAGATTCA GGACTAATTG CATCAAGAAC ACAGAGAAAG ACATATTTCT CAAGATCAGA AGTACTATTC
CaMV 35S
13301 CAGTATGGAC GATTCAAGGC TTGCTTCATA AACCAAGGCA AGTAATAGAG ATTGGAGTCT CTAAAAAGGT AGTTCCTACT GAATCTAAGG CCATGCATGG
Primer - 35S(PSP)Fwd
CaMV 35S
13401 AGTCTAAGAT TCAAATCGAG GATCTAACAG AACTCGCCGT GAAGACTGGC GAACAGTTCA TACAGAGTCT TTTACGACTC AATGACAAGA AGAAAATCTT
CaMV 35S
13501 CGTCAACATG GTGGAGCACG ACACTCTGGT CTACTCCAAA AATGTCAAAG ATACAGTCTC AGAAGACCAA AGGGCTATTG AGACTTTTCA ACAAAGGATA
CaMV 35S
13601 ATTTCGGGAA ACCTCCTCGG ATTCCATTGC CCAGCTATCT GTCACTTCAT CGAAAGGACA GTAGAAAAGG AAGGTGGCTC CTACAAATGC CATCATTGCG
CaMV 35S

Figure 105

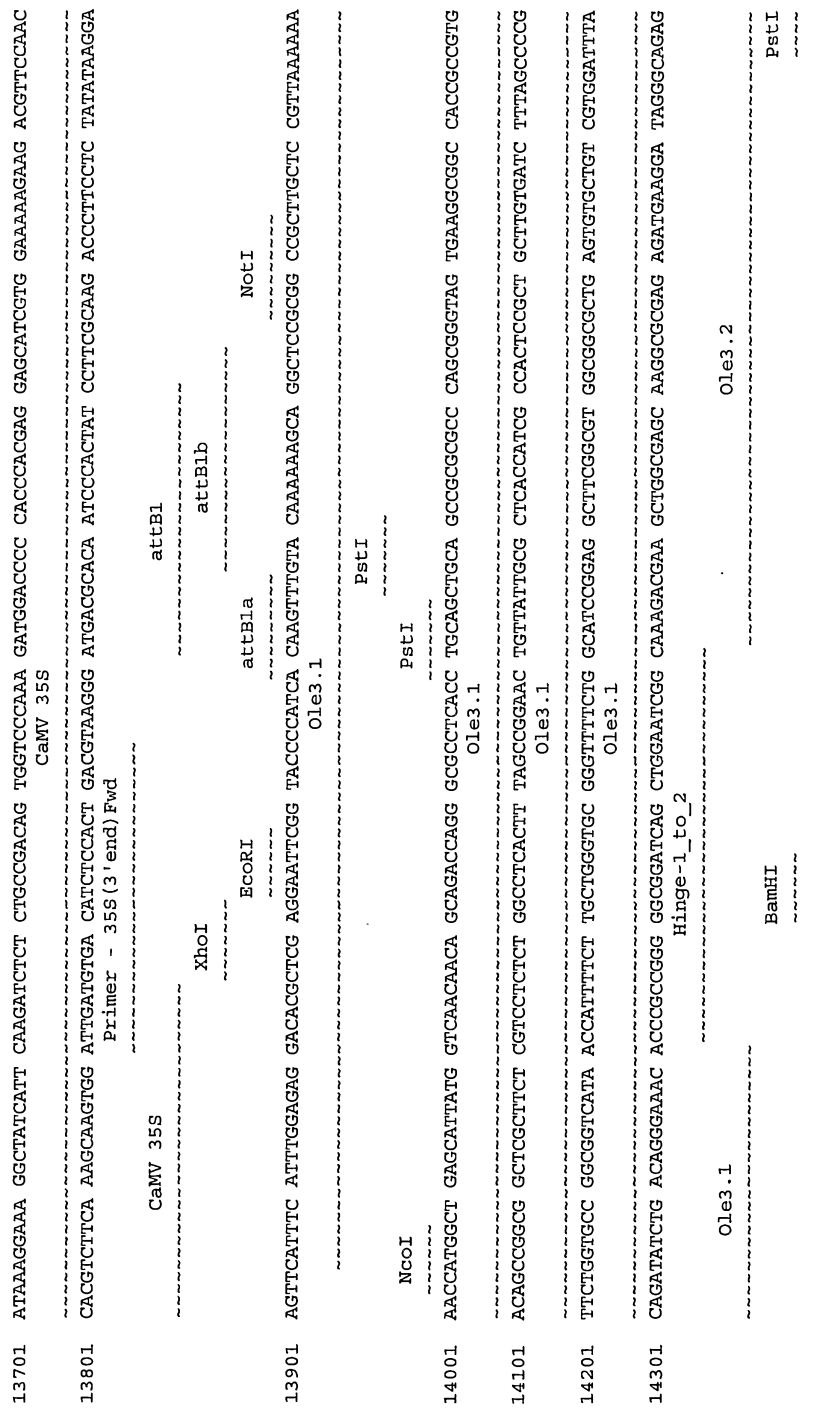


Figure 105 continued

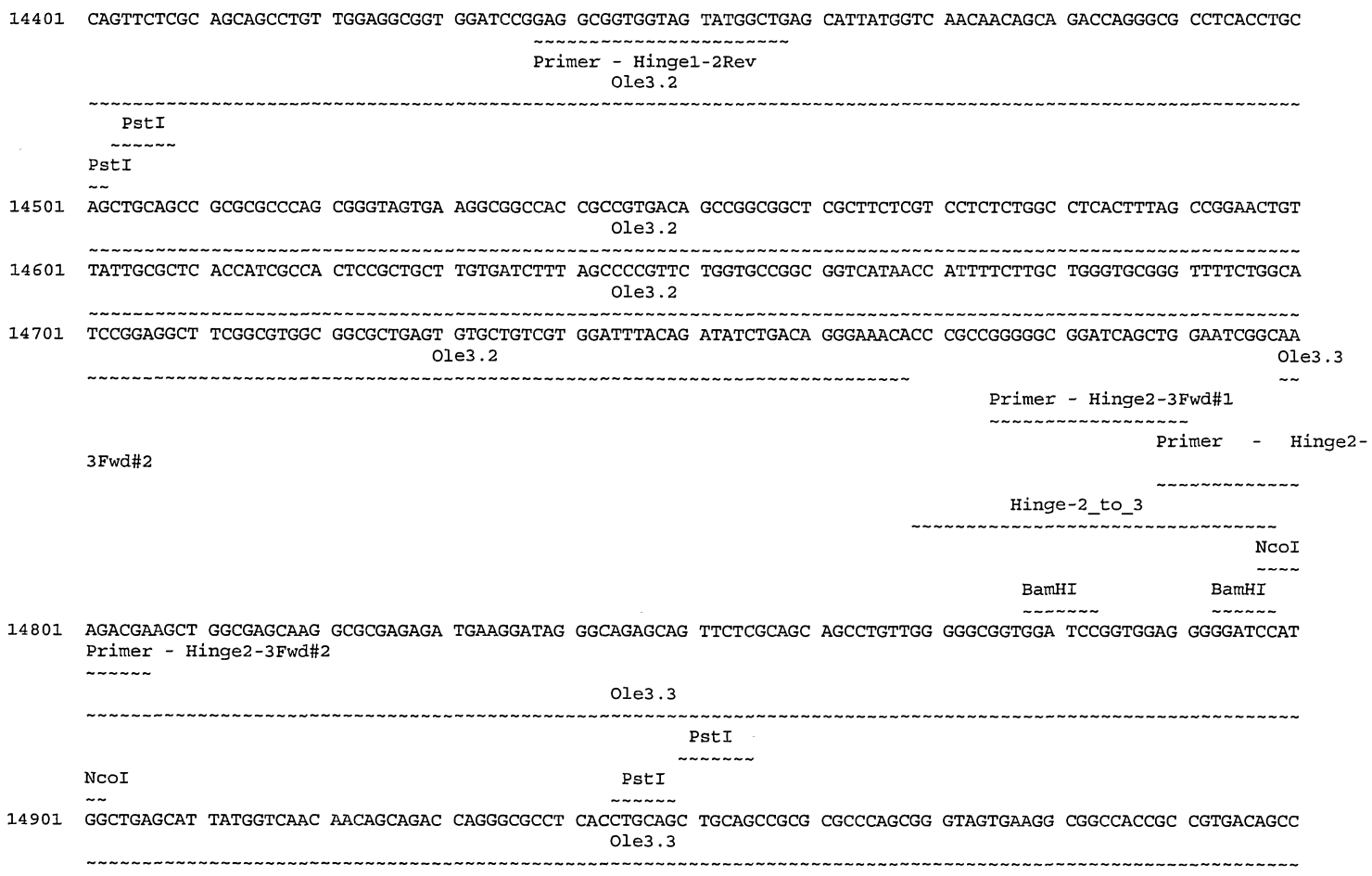


Figure 105 continued

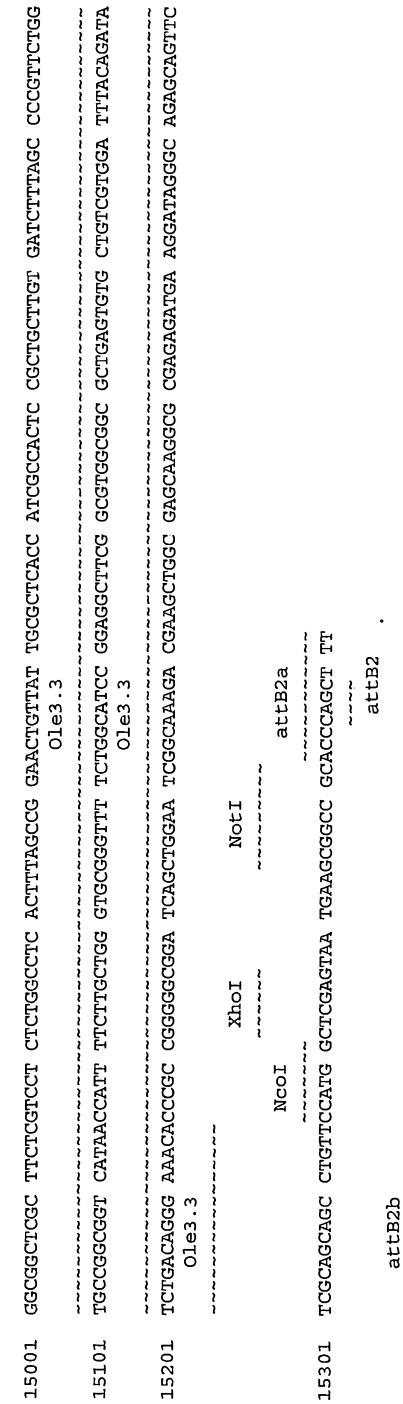


Figure 105 continued

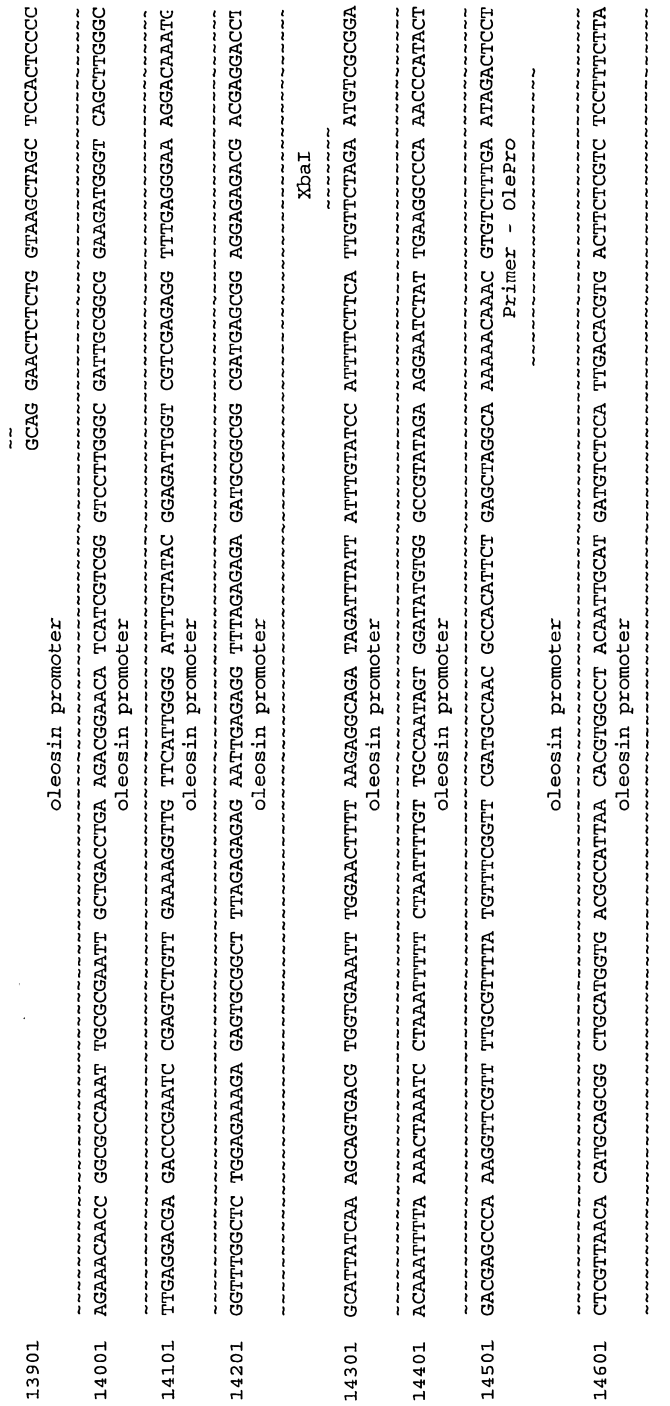

oleosin promoter
NotI
GCAG GAACTCTCTG GTAAGCTAGC TCCACTCCCC
13901
oleosin promoter
AGAAACAACC GGCGCCAAAT TGCGCGAATT GCTGACCTGA AGACGGAACA TCATCGTCGG GTCCTTGGGC GATTGCGGCG GAAGATGGGT CAGCTTGGGC
14001
oleosin promoter
TTGAGGACGA GACCCGAATC CGAGTCTGTT GAAAAGGTTG TTCATTGGGG ATTTGTATAC GGAGATTGGT CGTCGAGAGG TTTGAGGGAA AGGACAAATC
14101
oleosin promoter
GGTTTGGCTC TGGAGAAAGA GAGTGCGGCT TTAGAGAGAG AATTGAGAGG TTTAGAGAGA GATGCGGCGG CGATGAGCGG AGGAGAGACG ACGAGGACCT
14201
oleosin promoter
XbaI
GCATTATCAA AGCAGTGACG TGGTGAAATT TGGAACTTTT AAGAGGCAGA TAGATTTATT ATTTGTATCC ATTTTCTTCA TTGTTCTAGA ATGTCGCGGA
14301
oleosin promoter
ACAAATTTTA AAACTAAATC CTAAATTTTT CTAATTTTGT TGCCAATAGT GGATATGTGG GCCGTATAGA AGGAATCTAT TGAAGGCCCA AACCCATACT
14401
oleosin promoter
GACGAGCCCA AAGGTTCGTT TTGCGTTTTA TGTTTCGGTT CGATGCCAAC GCCACATTCT GAGCTAGGCA AAAAACAAAC GTGTCTTTGA ATAGACTCCT
14501
Primer - OlePro
oleosin promoter
CTCGTTAACA CATGCAGCGG CTGCATGGTG ACGCCATTAA CACGTGGCCT ACAATTGCAT GATGTCTCCA TTGACACGTG ACTTCTCGTC TCCTTTCTTA
14601
oleosin promoter

Figure 106

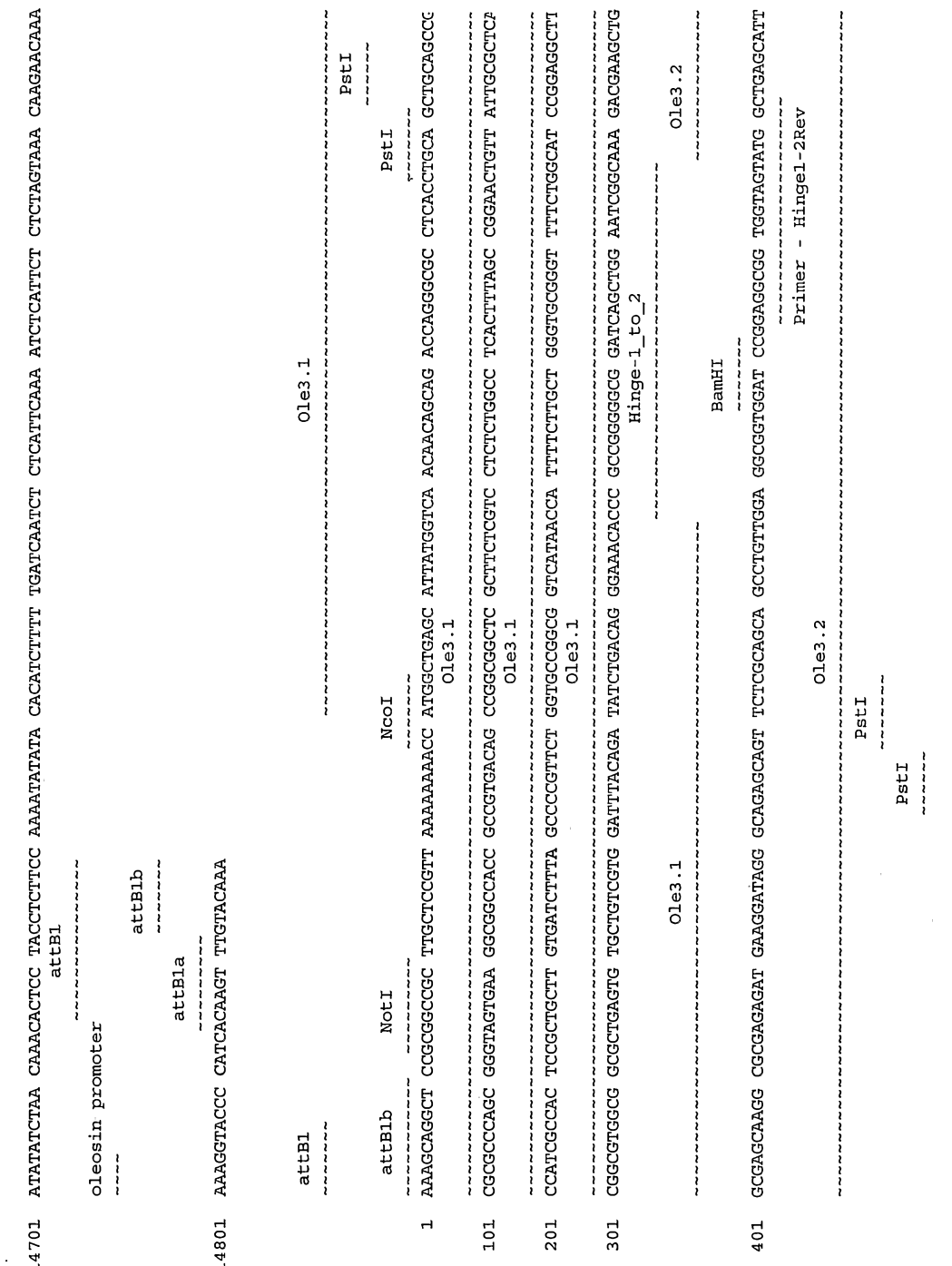


Figure 106 continued

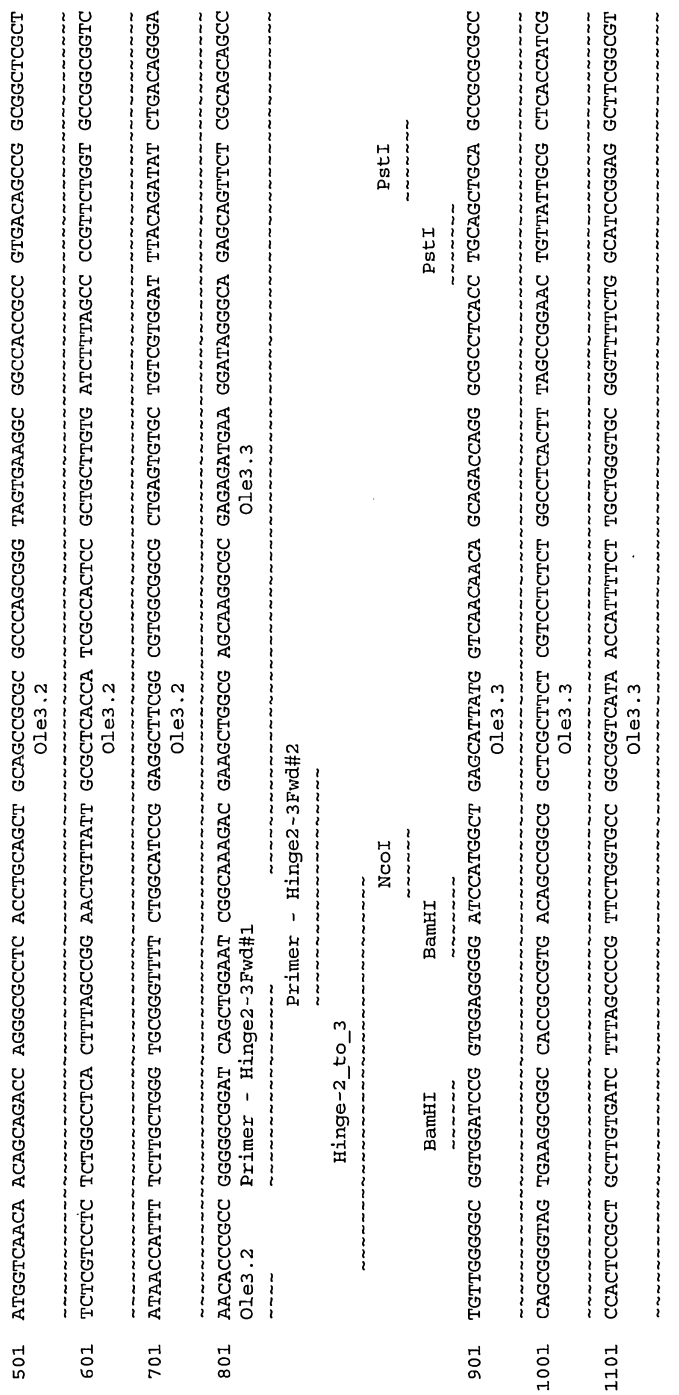


Figure 106 continued

```
1201  GGCGGCGCTG AGTGTGCTGT CGTGGATTTA CAGATATCTG ACAGGGAAAC ACCCGCCGGG GGCGGATCAG CTGGAATCGG CAAAGACGAA GCTGGCGAGC
                    Ole3.3
      ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
                                                 XhoI                NotI                      attB2b
                                                 ~~~~~~~            ~~~~~~~~                  ~~~~~~~~~~~~~~~
                                              NcoI                          attB2a
                                              ~~~~~~                        ~~~~~~~~~~~
1301  AAGGCGCGAG AGATGAAGGA TAGGGCAGAG CAGTTCTCGC AGCAGCCTGT TCCATGGCTC GAGTAATGAA GCGGCCGCAC CCAGCTTTCT TGTACAAAGT
                                                                                  ~~~~~~~~~~~~
                                                                                     attB2
      attB2b
      ~~~
1401  GGT
      ~~
      ..
```

Figure 106 continued

```
  1  MAEHYGQQQQ    TRAPHLQLQP    RAQRVVKAAT    AVTAGGSLLV
LSGLTLAGTV
 51  IALTIATPLL    VIFSPVLVPA    VITIFLLGAG    FLASGGFGVA
ALSVLSWIYR
101  YLTGKHPPGA    DQLESAKTKL    ASKAREMKDR    AEQFSQQPVA
GSQTSIDPSS
151  WLE**
```

Figure 107

```
  1 MAEHYGQQQQ TRAPHLQLQP RAQRVVKAAT AVTAGGSLLV
LSGLTLAGTV
 51 IALTIATPLL VIFSPVLVPA VITIFLLGAG FLASGGFGVA
ALSVLSWIYR
101 YLTGKHPPGA DQLESAKTKL ASKAREMKDR AEQFSQQPVA
GSQTSIDPSS
151 WHMMAEHYGQ QQQTRAPHLQ LQPRAQRVVK AATAVTAGGS
LLVLSGLTLA
201 GTVIALTIAT PLLVIFSPVL VPAVITIFLL GAGFLASGGF
GVAALSVLSW
251 IYRYLTGKHP PGADQLESAK TKLASKAREM KDRAEQFSQQ
PVAGSQTSID
301 QQVNVHMAEH YGQQQQTRAP HLQLQPRAQR VVKAATAVTA
GGSLLVLSGL
351 TLAGTVIALT IATPLLVIFS PVLVPAVITI FLLGAGFLAS
GGFGVAALSV
401 LSWIYRYLTG KHPPGADQLE SAKTKLASKA REMKDRAEQF
SQQPVAGSQT
451 SPWLE**
```

Figure 108

```
  1 MAEHYGQQQQ TRAPHLQLQP RAQRVVKAAT AVTAGGSLLV
LSGLTLAGTV
 51 IALTIATPLL VIFSPVLVPA VITIFLLGAG FLASGGFGVA
ALSVLSWIYR
101 YLTGKHPPGA DQLESAKTKL ASKAREMKDR AEQFSQQPVA
GSQTSIDPSS
151 WLEMAEHYGQ QQQTRAPHLQ LQPRAQRVVK AATAVTAGGS
LLVLSGLTLA
201 GTVIALTIAT PLLVIFSPVL VPAVITIFLL GAGFLASGGF
GVAALSVLSW
251 IYRYLTGKHP PGADQLESAK TKLASKAREM KDRAEQFSQQ
PVAGSQTSHM
301 MAEHYGQQQQ TRAPHLQLQP RAQRVVKAAT AVTAGGSLLV
LSGLTLAGTV
351 IALTIATPLL VIFSPVLVPA VITIFLLGAG FLASGGFGVA
ALSVLSWIYR
401 YLTGKHPPGA DQLESAKTKL ASKAREMKDR AEQFSQQPVA
GSQTSIDQQV
451 NVHMAEHYGQ QQQTRAPHLQ LQPRAQRVVK AATAVTAGGS
LLVLSGLTLA
501 GTVIALTIAT PLLVIFSPVL VPAVITIFLL GAGFLASGGF
GVAALSVLSW
551 IYRYLTGKHP PGADQLESAK TKLASKAREM KDRAEQFSQQ
PVAGSQTSPW
601 LE**
```

Figure 109

1 MAEHYGQQQQ TRAPHLQLQP RAQRVVKAAT AVTAGGSLLV LSGLTLAGTV

51 IALTIATPLL VIFSPVLVPA VITIFLLGAG FLASGGFGVA ALSVLSWIYR

101 YLTGKHPPGA DQLESAKTKL ASKAREMKDR AEQFSQQPVA GSQTSIDPSS

151 WLEMAEHYGQ QQQTRAPHLQ LQPRAQRVVK AATAVTAGGS LLVLSGLTLA

201 GTVIALTIAT PLLVIFSPVL VPAVITIFLL GAGFLASGGF GVAALSVLSW

251 IYRYLTGKHP PGADQLESAK TKLASKAREM KDRAEQFSQQ PVAGSQTSHM

301 FKWPSAMAEH YGQQQQTRAP HLQLQPRAQR VVKAATAVTA GGSLLVLSGL

351 TLAGTVIALT IATPLLVIFS PVLVPAVITI FLLGAGFLAS GGFGVAALSV

401 LSWIYRYLTG KHPPGADQLE SAKTKLASKA REMKDRAEQF SQQPVAGSQT

451 SSSELPWVDM AEHYGQQQQT RAPHLQLQPR AQRVVKAATA VTAGGSLLVL

501 SGLTLAGTVI ALTIATPLLV IFSPVLVPAV ITIFLLGAGF LASGGFGVAA

551 LSVLSWIYRY LTGKHPPGAD QLESAKTKLA SKAREMKDRA EQFSQQPVAG

601 SQTSSTPSSW HMMAEHYGQQ QQTRAPHLQL QPRAQRVVKA ATAVTAGGSL

651 LVLSGLTLAG TVIALTIATP LLVIFSPVLV PAVITIFLLG AGFLASGGFG

Figure 110

701 VAALSVLSWI YRYLTGKHPP GADQLESAKT KLASKAREMK
DRAEQFSQQP
751 VAGSQTSIDQ QVNVHMAEHY GQQQQTRAPH LQLQPRAQRV
VKAATAVTAG
801 GSLLVLSGLT LAGTVIALTI ATPLLVIFSP VLVPAVITIF
LLGAGFLASG
851 GFGVAALSVL SWIYRYLTGK HPPGADQLES AKTKLASKAR
EMKDRAEQFS
901 QQPVAGSQTS PWLE**

Figure 110 continued

```
  1  MAEHYGQQQQ    TRAPHLQLQP    RAQRVVKAAT    AVTAGGSLLV
LSGLTLAGTV
 51  IALTIATPLL    VIFSPVLVPA    VITIFLLGAG    FLASGGFGVA
ALSVLSWIYR
101  YLTGKHPPGA    DQLESAKTKL    ASKAREMKDR    AEQFSQQPVA
GSQTSIDPSS
151  WLE**
```

Figure 111

```
  1 MAEHYGQQQQ TRAPHLQLQP RAQRVVKAAT AVTAGGSLLV
LSGLTLAGTV
 51 IALTIATPLL VIFSPVLVPA VITIFLLGAG FLASGGFGVA
ALSVLSWIYR
101 YLTGKHPPGA DQLESAKTKL ASKAREMKDR AEQFSQQPVA
GSQTSIDPSS
151 WHMMAEHYGQ QQQTRAPHLQ LQPRAQRVVK AATAVTAGGS
LLVLSGLTLA
201 GTVIALTIAT PLLVIFSPVL VPAVITIFLL GAGFLASGGF
GVAALSVLSW
251 IYRYLTGKHP PGADQLESAK TKLASKAREM KDRAEQFSQQ
PVAGSQTSID
301 QQVNVHMAEH YGQQQQTRAP HLQLQPRAQR VVKAATAVTA
GGSLLVLSGL
351 TLAGTVIALT IATPLLVIFS PVLVPAVITI FLLGAGFLAS
GGFGVAALSV
401 LSWIYRYLTG KHPPGADQLE SAKTKLASKA REMKDRAEQF
SQQPVAGSQT
451 SPWLE**
```

Figure 112

```
  1 MAEHYGQQQQ TRAPHLQLQP RAQRVVKAAT AVTAGGSLLV
LSGLTLAGTV
 51 IALTIATPLL VIFSPVLVPA VITIFLLGAG FLASGGFGVA
ALSVLSWIYR
101 YLTGKHPPGA DQLESAKTKL ASKAREMKDR AEQFSQQPVA
GSQTSIDPSS
151 WLEMAEHYGQ QQQTRAPHLQ LQPRAQRVVK AATAVTAGGS
LLVLSGLTLA
201 GTVIALTIAT PLLVIFSPVL VPAVITIFLL GAGFLASGGF
GVAALSVLSW
251 IYRYLTGKHP PGADQLESAK TKLASKAREM KDRAEQFSQQ
PVAGSQTSHM
301 MAEHYGQQQQ TRAPHLQLQP RAQRVVKAAT AVTAGGSLLV
LSGLTLAGTV
351 IALTIATPLL VIFSPVLVPA VITIFLLGAG FLASGGFGVA
ALSVLSWIYR
401 YLTGKHPPGA DQLESAKTKL ASKAREMKDR AEQFSQQPVA
GSQTSIDQQV
451 NVHMAEHYGQ QQQTRAPHLQ LQPRAQRVVK AATAVTAGGS
LLVLSGLTLA
501 GTVIALTIAT PLLVIFSPVL VPAVITIFLL GAGFLASGGF
GVAALSVLSW
551 IYRYLTGKHP PGADQLESAK TKLASKAREM KDRAEQFSQQ
PVAGSQTSPW
601 LE**
```

Figure 113

1 MAEHYGQQQQ TRAPHLQLQP RAQRVVKAAT AVTAGGSLLV LSGLTLAGTV

51 IALTIATPLL VIFSPVLVPA VITIFLLGAG FLASGGFGVA ALSVLSWIYR

101 YLTGKHPPGA DQLESAKTKL ASKAREMKDR AEQFSQQPVA GSQTSIDPSS

151 WLEMAEHYGQ QQQTRAPHLQ LQPRAQRVVK AATAVTAGGS LLVLSGLTLA

201 GTVIALTIAT PLLVIFSPVL VPAVITIFLL GAGFLASGGF GVAALSVLSW

251 IYRYLTGKHP PGADQLESAK TKLASKAREM KDRAEQFSQQ PVAGSQTSHM

301 FKWPSAMAEH YGQQQQTRAP HLQLQPRAQR VVKAATAVTA GGSLLVLSGL

351 TLAGTVIALT IATPLLVIFS PVLVPAVITI FLLGAGFLAS GGFGVAALSV

401 LSWIYRYLTG KHPPGADQLE SAKTKLASKA REMKDRAEQF SQQPVAGSQT

451 SSSELPWVDM AEHYGQQQQT RAPHLQLQPR AQRVVKAATA VTAGGSLLVL

501 SGLTLAGTVI ALTIATPLLV IFSPVLVPAV ITIFLLGAGF LASGGFGVAA

551 LSVLSWIYRY LTGKHPPGAD QLESAKTKLA SKAREMKDRA EQFSQQPVAG

601 SQTSSTPSSW HMMAEHYGQQ QQTRAPHLQL QPRAQRVVKA ATAVTAGGSL

651 LVLSGLTLAG TVIALTIATP LLVIFSPVLV PAVITIFLLG AGFLASGGFG

701 VAALSVLSWI YRYLTGKHPP GADQLESAKT KLASKAREMK DRAEQFSQQP

Figure 114

751 VAGSQTSIDQ QVNVHMAEHY GQQQQTRAPH LQLQPRAQRV
VKAATAVTAG
801 GSLLVLSGLT LAGTVIALTI ATPLLVIFSP VLVPAVITIF
LLGAGFLASG
851 GFGVAALSVL SWIYRYLTGK HPPGADQLES AKTKLASKAR
EMKDRAEQFS
901 QQPVAGSQTS PWLE**

Figure 114 continued

```
  1 MAEHYGQQQQ TRAPHLQLQP RAQRVVKAAT AVTAGGSLLV
LSGLTLAGTV
 51 IALTIATPLL VIFSPVLVPA VITIFLLGAG FLASGGFGVA
ALSVLSWIYR
101 YLTGKHPPGA DQLESAKTKL ASKAREMKDR AEQFSQQPVG
GGGSGGGGSM
151 AEHYGQQQQT RAPHLQLQPR AQRVVKAATA VTAGGSLLVL
SGLTLAGTVI
201 ALTIATPLLV IFSPVLVPAV ITIFLLGAGF LASGGFGVAA
LSVLSWIYRY
251 LTGKHPPGAD QLESAKTKLA SKAREMKDRA EQFSQQPVGG
GGSGGGGSMA
301 EHYGQQQQTR APHLQLQPRA QRVVKAATAV TAGGSLLVLS
GLTLAGTVIA
351 LTIATPLLVI FSPVLVPAVI TIFLLGAGFL ASGGFGVAAL
SVLSWIYRYL
401 TGKHPPGADQ LESAKTKLAS KAREMKDRAE QFSQQPVPWL E**
```

Figure 115

```
  1 MAEHYGQQQQ TRAPHLQLQP RAQRVVKAAT AVTAGGSLLV
LSGLTLAGTV
 51 IALTIATPLL VIFSPVLVPA VITIFLLGAG FLASGGFGVA
ALSVLSWIYR
101 YLTGKHPPGA DQLESAKTKL ASKAREMKDR AEQFSQQPVG
GGGSGGGGSM
151 AEHYGQQQQT RAPHLQLQPR AQRVVKAATA VTAGGSLLVL
SGLTLAGTVI
201 ALTIATPLLV IFSPVLVPAV ITIFLLGAGF LASGGFGVAA
LSVLSWIYRY
251 LTGKHPPGAD QLESAKTKLA SKAREMKDRA EQFSQQPVGG
GGSGGGGSMA
301 EHYGQQQQTR APHLQLQPRA QRVVKAATAV TAGGSLLVLS
GLTLAGTVIA
351 LTIATPLLVI FSPVLVPAVI TIFLLGAGFL ASGGFGVAAL
SVLSWIYRYL
401 TGKHPPGADQ LESAKTKLAS KAREMKDRAE QFSQQPVPWL E**
```

Figure 116

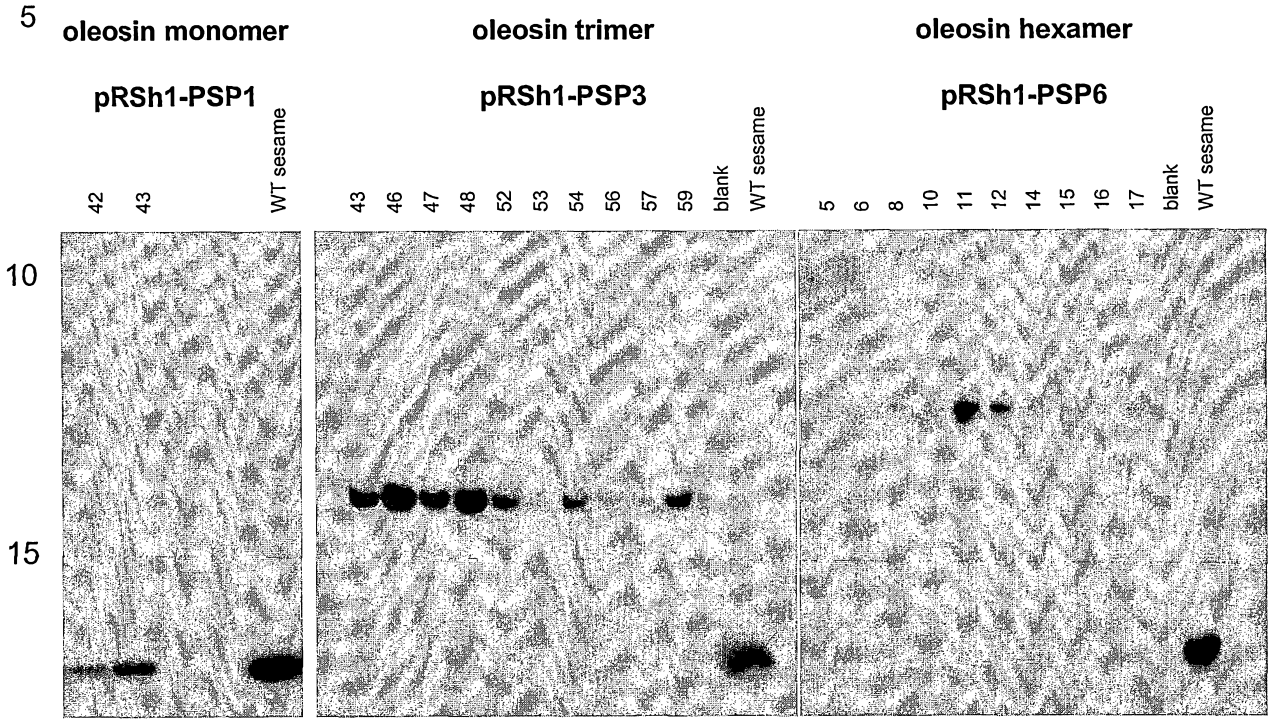

oleosin monomer
pRSh1-PSP1
42
43
WT sesame
oleosin trimer
pRSh1-PSP3
43
46
47
48
52
53
54
56
57
59
blank
WT sesame
oleosin hexamer
pRSh1-PSP6
5
6
8
10
11
12
14
15
16
17
blank
WT sesame

Figure 117

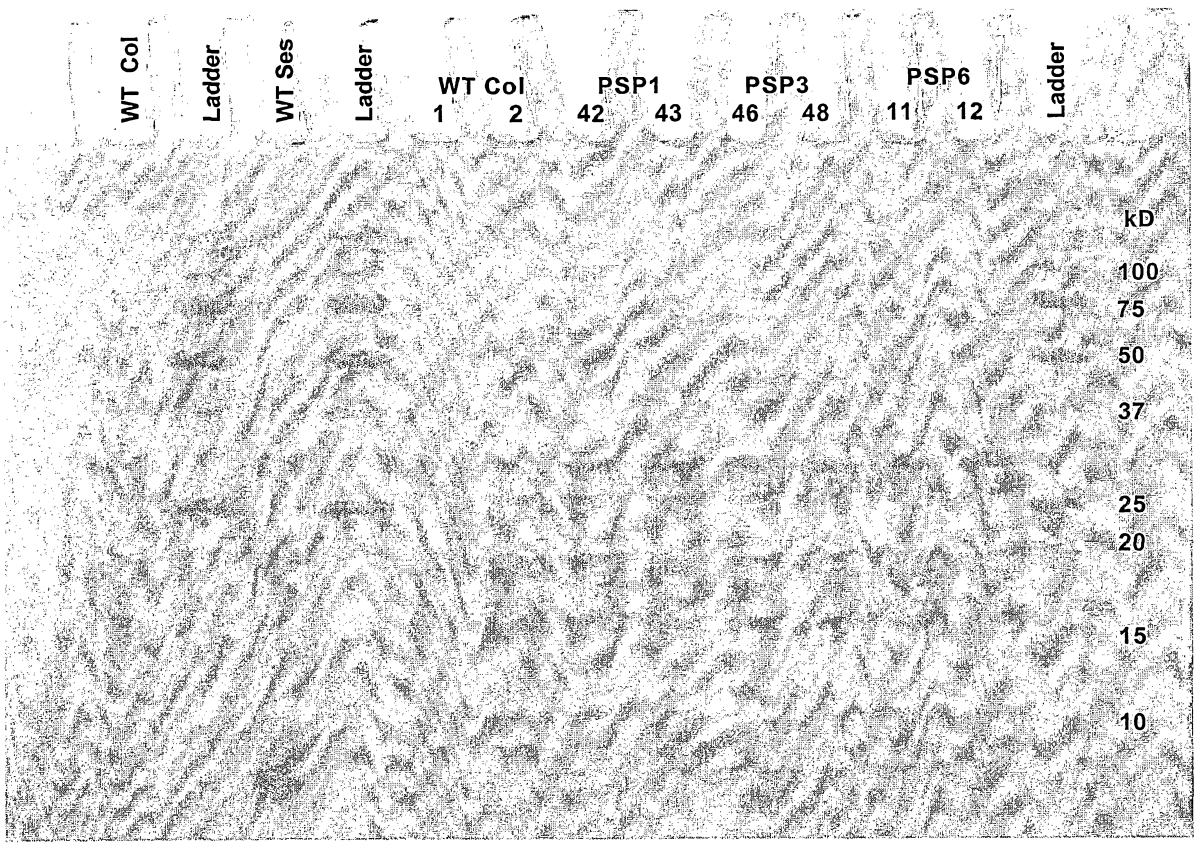

WT Col
Ladder
WT Ses
Ladder
WT Col
1
2
PSP1
42
43
PSP3
46
48
PSP6
11
12
Ladder
kD
100
75
50
37
25
20
15
10

Figure 118

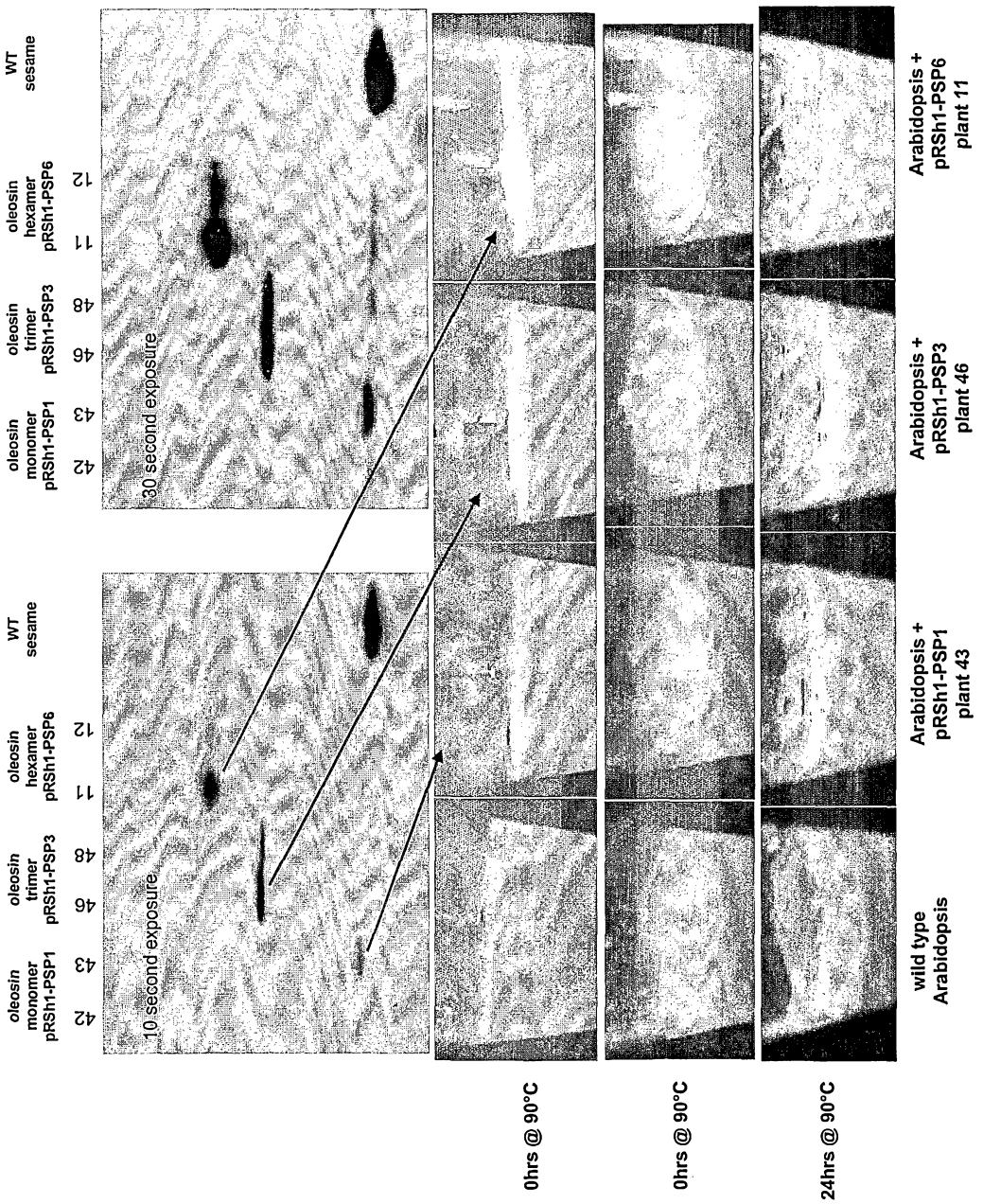

10 second exposure
30 second exposure
oleosin monomer pRSh1-PSP1
42
43
oleosin trimer pRSh1-PSP3
46
48
oleosin hexamer pRSh1-PSP6
11
12
WT sesame
0hrs @ 90°C
24hrs @ 90°C
wild type Arabidopsis
Arabidopsis + pRSh1-PSP1 plant 43
Arabidopsis + pRSh1-PSP3 plant 46
Arabidopsis + pRSh1-PSP6 plant 11

Figure 119

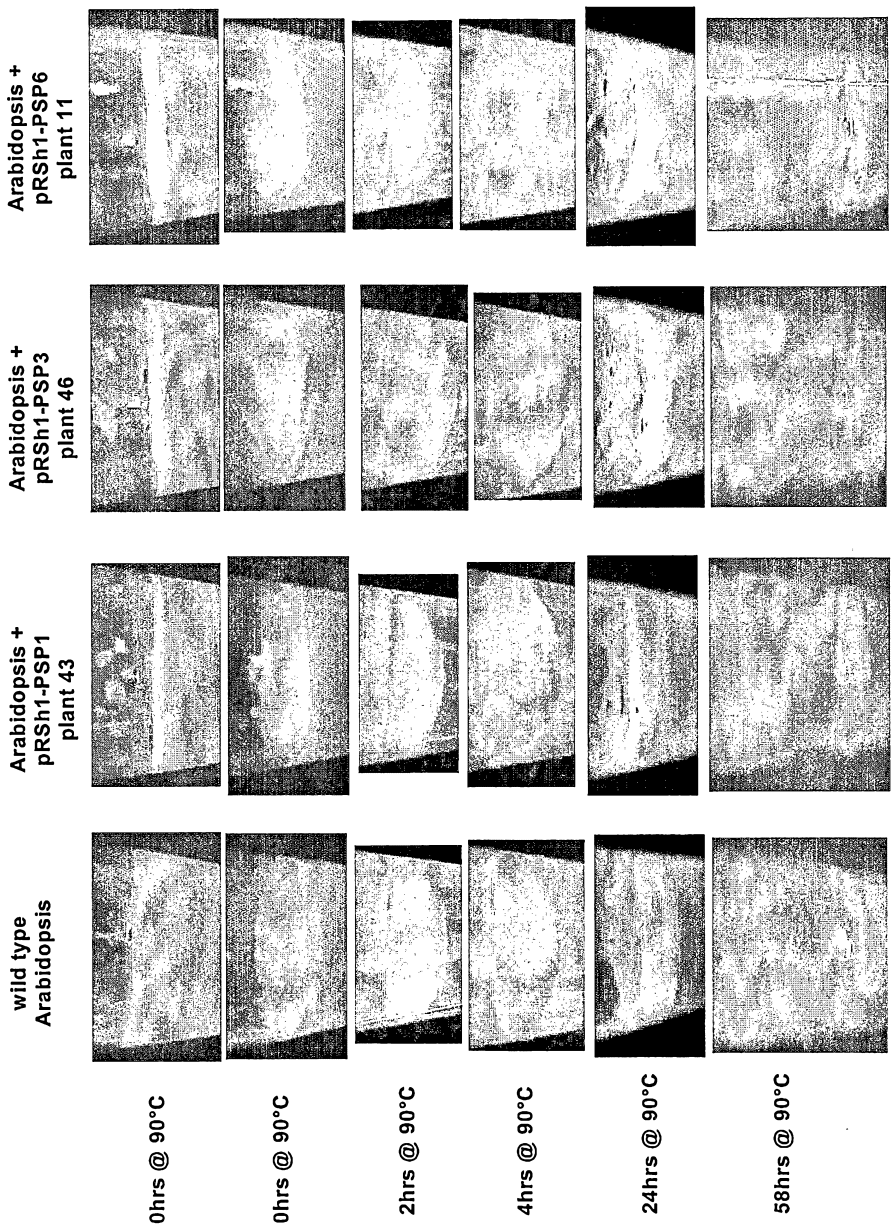

wild type Arabidopsis
Arabidopsis + pRSh1-PSP1 plant 43
Arabidopsis + pRSh1-PSP3 plant 46
Arabidopsis + pRSh1-PSP6 plant 11
0hrs @ 90°C
0hrs @ 90°C
2hrs @ 90°C
4hrs @ 90°C
24hrs @ 90°C
58hrs @ 90°C

Figure 120

SEQUENCE LISTING

<110> Agriculture Victoria Services Pty Ltd
AgResearch Limited

<120> Polyoleosins

<130> 81115824:DLT

<150> 2005905787
<151> 2005-10-19

<150> 2005906364
<151> 2005-11-16

<160> 83

<170> PatentIn version 3.3

<210> 1
<211> 4
<212> DNA
<213> Synthetic

<400> 1
cacc 4

<210> 2
<211> 15
<212> DNA
<213> Synthetic

<400> 2
tcactcgagg agctc 15

<210> 3
<211> 12
<212> DNA
<213> Synthetic

<400> 3
tctagaggta ct 12

<210> 4
<211> 12
<212> DNA
<213> Synthetic

<400> 4
actagtagta cc 12

<210> 5
<211> 12
<212> DNA
<213> Synthetic

<400> 5
cccgggggta ct 12

<210> 6
<211> 12
<212> DNA
<213> Synthetic

<400> 6
ctgcagagta cc 12

<210> 7
<211> 12
<212> DNA
<213> Synthetic

<400> 7
aagcttggta ct 12

<210> 8
<211> 12
<212> DNA
<213> Synthetic

<400> 8
atcgatagta cc 12

<210> 9
<211> 15
<212> DNA
<213> Synthetic

<400> 9
gtcgacggta cttct 15

<210> 10
<211> 6
<212> DNA
<213> Synthetic

<400> 10
ctcgag 6

<210> 11
<211> 22
<212> DNA
<213> Synthetic

<400> 11
tcagatctcg gtgacgggca gg 22

<210> 12
<211> 22
<212> DNA
<213> Synthetic

<400> 12
atgagccaag aacgacgccc gg 22

<210> 13
<211> 22
<212> DNA
<213> Synthetic

<400> 13
caccatggca caacctcaag tt 22

<210> 14
<211> 11
<212> DNA
<213> Synthetic

<400> 14
gctccaccgc g 11

<210> 15
<211> 14
<212> DNA
<213> Synthetic

<400> 15
tccactagtt ctag 14

<210> 16
<211> 15
<212> DNA
<213> Synthetic

<400> 16
ctagaactag tggat 15

<210> 17
<211> 13
<212> DNA
<213> Synthetic

<400> 17
tcctgcagag tac 13

<210> 18
<211> 13
<212> DNA
<213> Synthetic

<400> 18
gtactctgca gga 13

<210> 19
<211> 12
<212> DNA
<213> Synthetic

<400> 19
atacgatagt ac 12

<210> 20
<211> 15
<212> DNA
<213> Synthetic

<400> 20
ctatcgatac cgtcg 15

<210> 21
<211> 24
<212> DNA
<213> Synthetic

<400> 21
ctcgagtgtt gatctcttag cttc 24

<210> 22
<211> 10
<212> PRT
<213> Synthetic

<400> 22

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

<210> 23
<211> 24
<212> DNA
<213> Synthetic

<400> 23
gacacgctcg aggaattcgg tacc 24

<210> 24
<211> 24
<212> DNA
<213> Synthetic

<400> 24
gatggtgatg actgcagatc agaa 24

<210> 25
<211> 30
<212> DNA
<213> Synthetic

<400> 25
cgattaatct gagtttttct gatctgcagt 30

<210> 26
<211> 23
<212> DNA
<213> Synthetic

<400> 26
cgatcaccgt tccggccaat gtc 23

<210> 27
<211> 23
<212> DNA
<213> Synthetic

<400> 27
gacacgtgac ttctcgtctc ctt 23

<210> 28
<211> 24
<212> DNA
<213> Synthetic

<400> 28
gatggtgatg actgcagatc agaa 24

<210> 29
<211> 628
<212> DNA
<213> Trifolium repens

<400> 29
caccatggca caacctcaag ttcaagtcca ctcaacaaca acacaccgtc aagaaactgc 60

tacctaccca tcaacccaaa acattcgtaa agatgtttac gaaaatgtta actatcccgg 120

ccaacgcggt cgttataacg accgctataa tgatagtggt cgttatgatg gtggtattgc 180

ctccttttg tcagagagaa gtcctccagc ctctcaaatc ctcgctaccg ttggaggatt 240

tttcataggt ggtactctat ttttattagc tagcatttca tttatcgcca gtcttattgg 300

attggcgata atgacaccac tttttatcct ttttagcccg gttttagtcc ctgctgccct 360

cactataggg ctagcagtgg ctggaatatt gacagcagat gcttgcgggt tgacggggct 420

tatgtcgttg tcgtggaccg tgaaatatgt tagggattta caagcagtag tgcccgaaca 480

aatggattcg atgaagggac gtgtcgcgga tgtcgcgagt tatgttggac aaaagactaa 540

ggatgttgga caaaaaacta aagaggttgg acaagacata caaacaaaag cacatgaagc 600

taagagatca acagtcgacc tcgagtga 628

<210> 30
<211> 633
<212> DNA
<213> Trifolium repens

<400> 30
tctagaggta ctatggcaca acctcaagtt caagtccact caacaacaac acaccgtcaa 60

gaaactgcta cctacccatc aacccaaaac attcgtaaag atgtttacga aaatgttaac 120

tatcccggcc aacgcggtcg ttataacgac cgctataatg atagtggtcg ttatgatggt 180

ggtattgcct ccttttgtc agagagaagt cctccagcct ctcaaatcct cgctaccgtt 240

ggaggatttt tcataggtgg tactctattt ttattagcta gcatttcatt tatcgccagt 300

cttattggat tggcgataat gacaccactt tttatccttt ttagcccggt tttagtccct 360

gctgccctca ctatagggct agcagtggct ggaatattga cagcagatgc ttgcgggttg 420

acggggctta tgtcgttgtc gtggaccgtg aaatatgtta gggatttaca agcagtagtg 480

cccgaacaaa tggattcgat gaagggacgt gtcgcggatg tcgcgagtta tgttggacaa 540

aagactaagg atgttggaca aaaaactaaa gaggttggac aagacataca aacaaaagca 600

catgaagcta agagatcaac aggtactact agt 633

<210> 31
<211> 613
<212> DNA
<213> Trifolium repens

<400> 31
cccggggtac tatggcacaa cctcaagttc aagtccactc aacaacaaca caccgtcaag 60

aaactgctac ctacccatca acccaaaaca ttcgtaaaga tgtttacgaa aatgttaact 120

atcccggcca acgcggtcgt tataacgacc gctataatga tagtggtcgt tatgatggtg 180

gtattgcctc ctttttgtca gagagaagtc ctccagcctc tcaaatcctc gctaccgttg 240

gaggattttt cataggtggt actctatttt tattagctag catttcattt atcgccagtc 300

ttattggatt ggcgataatg acaccacttt ttatcctttt tagcccggtt ttagtccctg 360

ctgccctcac tatagggcta gcagtggctg gaatattgac agcagatgct tgcgggttga 420

cggggcttat gtcgttgtcg tggaccgtga aatatgttag ggatttacaa gcagtagtgc 480

ccgaacaaat ggattcgatg aagggacgtg tcgcggatgt cgcgagttat gttggacaaa 540

agactaagga tgttggacaa aaaactaaag aggttggaca agacatacaa acaaaagcac 600

aggtactctg cag 613

<210> 32
<211> 633
<212> DNA
<213> Trifolium repens

<400> 32
aagcttggta ctatggcaca acctcaagtt caagtccact caacaacaac acaccgtcaa 60

gaaactgcta cctacccatc aacccaaaac attcgtaaag atgtttacga aaatgttaac 120

tatcccggcc aacgcggtcg ttataacgac cgctataatg atagtggtcg ttatgatggt 180

ggtattgcct cctttttgtc agagagaagt cctccagcct ctcaaatcct cgctaccgtt 240

ggaggatttt tcataggtgg tactctattt ttattagcta gcatttcatt tatcgccagt 300

cttattggat tggcgataat gacaccactt tttatccttt ttagcccggt tttagtccct 360

gctgccctca ctatagggct agcagtggct ggaatattga cagcagatgc ttgcgggttg 420

acggggctta tgtcgttgtc gtggaccgtg aaatatgtta gggatttaca agcagtagtg 480

cccgaacaaa tggattcgat gaagggacgt gtcgcggatg tcgcgagtta tgttggacaa 540

aagactaagg atgttggaca aaaaactaaa gaggttggac aagacataca aacaaaagca 600

catgaagcta agagatcaac aggtactatc gat 633

<210> 33
<211> 630
<212> DNA
<213> Trifolium repens

<400> 33
gtcgacggta cttctatggc acaacctcaa gttcaagtcc actcaacaac aacacaccgt 60

caagaaactg ctacctaccc atcaacccaa aacattcgta aagatgttta cgaaaatgtt 120

aactatcccg gccaacgcgg tcgttataac gaccgctata atgatagtgg tcgttatgat 180

ggtggtattg cctccttttt gtcagagaga agtcctccag cctctcaaat cctcgctacc 240

gttggaggat ttttcatagg tggtactcta tttttattag ctagcatttc atttatcgcc 300

agtcttattg gattggcgat aatgacacca ctttttatcc tttttagccc ggttttagtc 360

cctgctgccc tcactatagg gctagcagtg gctggaatat tgacagcaga tgcttgcggg 420

ttgacggggc ttatgtcgtt gtcgtggacc gtgaaatatg ttagggattt acaagcagta 480

gtgcccgaac aaatggattc gatgaaggga cgtgtcgcgg atgtcgcgag ttatgttgga 540

caaaagacta aggatgttgg acaaaaaact aaagaggttg gacaagacat acaaacaaaa 600

gcacatgaag ctaagagatc aacactcgag 630

<210> 34
<211> 709
<212> DNA
<213> Trifolium repens

<400> 34
caccatggca caacctcaag ttcaagtcca ctcaacaaca acacaccgtc aagaaactgc 60

tacctaccca tcaacccaaa acattcgtaa agatgtttac gaaaatgtta actatcccgg 120

ccaacgcggt cgttataacg accgctataa tgatagtggt cgttatgatg gtggtattgc 180

ctccttttg tcagagagaa gtcctccagc ctctcaaatc ctcgctaccg ttggaggatt 240

tttcataggt ggtactctat ttttattagc tagcatttca tttatcgcca gtcttattgg 300

attggcgata atgacaccac tttttatcct ttttagcccg gttttagtcc ctgctgccct 360

cactataggg ctagcagtgg ctggaatatt gacagcagat gcttgcgggt tgacggggct 420

tatgtcgttg tcgtggaccg tgaaatatgt tagggattta caagcagtag tgcccgaaca 480

aatggattcg atgaagggac gtgtcgcgga tgtcgcgagt tatgttggac aaaagactaa 540

ggatgttgga caaaaaacta aagaggttgg acaagacata caaacaaaag cacatgaagc 600

taagagatca acagagctcc accgcggtgg cggccgctct agaactagtg gatcccccgg 660

gctgcaggaa ttcgatatca agcttatcga taccgtcgac ctcgagtga 709

<210> 35
<211> 1336
<212> DNA
<213> Trifolium repens

<400> 35
caccatggca caacctcaag ttcaagtcca ctcaacaaca acacaccgtc aagaaactgc 60

tacctaccca tcaacccaaa acattcgtaa agatgtttac gaaaatgtta actatcccgg 120

ccaacgcggt cgttataacg accgctataa tgatagtggt cgttatgatg gtggtattgc 180

ctccttttg tcagagagaa gtcctccagc ctctcaaatc ctcgctaccg ttggaggatt 240

tttcataggt ggtactctat ttttattagc tagcatttca tttatcgcca gtcttattgg 300

attggcgata atgacaccac tttttatcct ttttagcccg gttttagtcc ctgctgccct 360

cactataggg ctagcagtgg ctggaatatt gacagcagat gcttgcgggt tgacggggct 420

tatgtcgttg tcgtggaccg tgaaatatgt tagggattta caagcagtag tgcccgaaca 480

aatggattcg atgaagggac gtgtcgcgga tgtcgcgagt tatgttggac aaaagactaa 540

ggatgttgga caaaaaacta aagaggttgg acaagacata caaacaaaag cacatgaagc 600

taagagatca acagagctcc accgcggtgg cggccgctct agaggtacta tggcacaacc 660

tcaagttcaa gtccactcaa caacaacaca ccgtcaagaa actgctacct acccatcaac 720

ccaaaacatt cgtaaagatg tttacgaaaa tgttaactat cccggccaac gcggtcgtta 780

taacgaccgc tataatgata gtggtcgtta tgatggtggt attgcctcct ttttgtcaga 840

gagaagtcct ccagcctctc aaatcctcgc taccgttgga ggatttttca taggtggtac 900

tctattttta ttagctagca tttcatttat cgccagtctt attggattgg cgataatgac 960

accacttttt atccttttta gcccggtttt agtccctgct gccctcacta tagggctagc 1020

agtggctgga atattgacag cagatgcttg cgggttgacg gggcttatgt cgttgtcgtg 1080

gaccgtgaaa tatgttaggg atttacaagc agtagtgccc gaacaaatgg attcgatgaa 1140

gggacgtgtc gcggatgtcg cgagttatgt tggacaaaag actaaggatg ttggacaaaa 1200

aactaaagag gttggacaag acatacaaac aaaagcacat gaagctaaga gatcaacagg 1260

tactactaga actagtggat cccccgggct gcaggaattc gatatcaagc ttatcgatac 1320

cgtcgacctc gagtga 1336

<210> 36
<211> 1958
<212> DNA
<213> Trifolium repens

<400> 36
caccatggca caacctcaag ttcaagtcca ctcaacaaca acacaccgtc aagaaactgc 60

tacctaccca tcaacccaaa acattcgtaa agatgtttac ggaaaatgtt aactatcccg 120

gccaacgcgg tcgttataac gaccgctata atgatagtgg tcgttatgat ggtggtattg 180

cctccttttt gtcagagaga agtcctccag cctctcaaat cctcgctacc gttggaggat 240

ttttcatagg tggtactcta tttttattag ctagcatttc atttatcgcc agtcttattg 300

gattggcgat aatgacacca ctttttatcc tttttagccc ggttttagtc cctgctgccc 360

tcactatagg gctagcagtg gctggaatat tgacagcaga tgcttgcggg ttgacggggc 420

ttatgtcgtt gtcgtggacc gtgaaatatg ttagggattt acaagcagta gtgcccgaac 480

aaatggattc gatgaaggga cgtgtcgcgg atgtcgcgag ttatgttgga caaaagacta 540

aggatgttgg acaaaaaact aaagaggttg gacaagacat acaaacaaaa gcacatgaag 600

ctaagagatc aacagagctc caccgcggtg gcggccgctc tagaggtact atggcacaac 660

ctcaagttca agtccactca acaacaacac accgtcaaga aactgctacc tacccatcaa 720

cccaaaacat tcgtaaagat gtttacgaaa atgttaacta tcccggccaa cgcggtcgtt 780

ataacgaccg ctataatgat agtggtcgtt atgatggtgg tattgcctcc tttttgtcag 840

agagaagtcc tccagcctct caaatcctcg ctaccgttgg aggatttttc ataggtggta 900

ctctattttt attagctagc atttcattta tcgccagtct tattggattg gcgataatga 960

caccactttt tatccttttt agcccggttt tagtccctgc tgccctcact atagggctag 1020

cagtggctgg aatattgaca gcagatgctt gcgggttgac ggggcttatg tcgttgtcgt 1080

ggaccgtgaa atatgttagg gatttacaag cagtagtgcc cgaacaaatg gattcgatga 1140

agggacgtgt cgcggatgtc gcgagttatg ttggacaaaa gactaaggat gttggacaaa 1200

aaactaaaga ggttggacaa gacatacaaa caaaagcaca tgaagctaag agatcaacag 1260

gtactactag aactagtgga tcccccgggg gtactatggc acaacctcaa gttcaagtcc 1320

actcaacaac aacacaccgt caagaaactg ctacctaccc atcaacccaa aacattcgta 1380

aagatgttta cgaaaatgtt aactatcccg gccaacgcgg tcgttataac gaccgctata 1440

atgatagtgg tcgttatgat ggtggtattg cctccttttt gtcagagaga agtcctccag 1500

cctctcaaat cctcgctacc gttggaggat ttttcatagg tggtactcta tttttattag 1560

ctagcatttc atttatcgcc agtcttattg gattggcgat aatgacacca ctttttatcc 1620

tttttagccc ggttttagtc cctgctgccc tcactatagg gctagcagtg gctggaatat 1680

tgacagcaga tgcttgcggg ttgacggggc ttatgtcgtt gtcgtggacc gtgaaatatg 1740

ttagggattt acaagcagta gtgcccgaac aaatggattc gatgaaggga cgtgtcgcgg 1800

atgtcgcgag ttatgttgga caaaagacta aggatgttgg acaaaaaact aaagaggttg 1860

gacaagacat acaaacaaaa gcacatgaag ctaagagatc aacaggtact ctgcaggaat 1920

tcgatatcaa gcttatcgat accgtcgacc tcgagtga 1958

<210> 37
<211> 500
<212> DNA
<213> Trifolium repens

<400> 37
caacaaayaa cwcaccgtca agaaactgct acctacccat caacccawwa cattcgtaaa 60

gatgtttacg aaaatgttaa ctatcccggc caacgcggtc gttataacga ccgctataat 120

gatagtggtc gttatgatgg tggtattgcc tccttttgt cagagagaag tcctccagcc 180

accgttggag gatttttcat aggtggtact ctatttttat tagctagcat ttcatttatc 240

gccagtctta ttggattggc gataatgaca ccactttta tctcaaatcc tcgcttcctt 300

tttagcccgg ttttagtccc tgctgccctc actatagggc tagcagtggc tggaatattg 360

acagcagatg cttgcgggtt gacggggctt atgtcgttgt cgtggaccgt gaaatatgtt 420

agggatttac aagcagtagt gcccgaacaa atggattcga tgaagggacg tgtcgcggat 480

gtcgcgagtt atgttggaca 500

<210> 38
<211> 662
<212> DNA
<213> Trifolium repens

<400> 38
gatatcaagc ttggtactaw tggcacaacc tcaagttcaa gtccactcaa caacaacaca 60

ccgtcaagaa aactgctacc tacccatcaa cccaaaacat tcgtaaagat gtttacgaaa 120

aatgtttaac tatcccggcc aacgcggtcg ttataacgac cgctataatg atagtggtcg 180

tttatgatgg tggtattgcc tccttttgt cagagagaag tcctccagcc tctcaaatcc 240

tcgctaccgt tggaggattt ttcataggtg gtactctatt tttattagct agcatttcat 300

ttatcgccag tcttattgga ttggcgataa tgacaccact ttttatcctt tttagcccgg 360

twttagtccc tgctgccctc actatagggc tagcagtggc tggaatattg acagcagatg 420

cttgcgggtt gacggggctt atgtcgttgt cgtggaccgt gaaatatgtt agggatttac 480

aagcagtagt gcccgaacaa atggattcga tgaagggacg tgtcgcggat gtcgcgagtt 540

atgttggaca aaagactaag gatgttggac aaaaaactaa agaggttgga caagacatac 600

aaacaaaagc acatgaagct aagagatcaa caggtactat cgataccgtc gacctcgagt 660

ga 662

<210> 39
<211> 794
<212> DNA
<213> Trifolium repens

<400> 39
atacaaacaa aagcacatga agctaagaga tcaacaggta ctctgcagga attcgatatc 60

aagcttggta ctatggcaca acctcaagtt caagtccact caacaacaac acaccgtcaa 120

gaaactgcta cctacccatc aacccaaaac attcgtaaag atgtttacga aaatgttaac 180

tatcccggcc aacgcggtcg ttataacgac cgctataatg atagtggtcg ttatgatggt 240

ggtattgcct ccttttgtc agagagaagt cctccagcct ctcaaatcct cgctaccgtt 300

ggaggatttt tcataggtgg tactctattt ttattagcta gcatttcatt tatcgccagt 360

cttattggat tggcgataat gacaccactt tttatccttt ttagcccggt tttagtccct 420

gctgccctca ctatagggct agcagtggct ggaatattga cagcagatgc ttgcgggttg 480

acggggctta tgtcgttgtc gtggaccgtg aaatatgtta gggatttaca agcagtagtg 540

cccgaacaaa tggattcgat gaagggacgt gtcgcggatg tcgcgagtta tgttggacaa 600

aagactaagg atgttggaca aaaaactaaa gaggttggac aagacataca aacaaaagca 660

catgaagcta agagatcaac aggtactatc gataccgtcg acctcgagtg aaagggtggg 720

cgcgccgacc cagctttctt gtacaaagtt ggcattataa gaaagcattg cttatcaatt 780

tgttgcaacg aaca 794

<210> 40
<211> 609
<212> DNA
<213> Trifolium repens

<400> 40
atggcacaac ctcaagttca agtccactca acaacaacac accgtcaaga aactgctacc 60

tacccatcaa cccaaaacat tcgtaaagat gtttacgaaa atgttaacta tcccggccaa 120

cgcggtcgtt ataacgaccg ctataatgat agtggtcgtt atgatggtgg tattgcctcc 180

tttttgtcag agagaagtcc tccagcctct caaatcctcg ctaccgttgg aggatttttc 240

ataggtggta ctctattttt attagctagc atttcattta tcgccagtct tattggattg 300

gcgataatga caccactttt tatccttttt agcccggttt tagtccctgc tgccctcact 360

atagggctag cagtggctgg aatattgaca gcagatgctt gcgggttgac ggggcttatg 420

tcgttgtcgt ggaccgtgaa atatgttagg gatttacaag cagtagtgcc cgaacaaatg 480

gattcgatga agggacgtgt cgcggatgtc gcgagttatg ttggacaaaa gactaaggat 540

gttggacaaa aaactaaaga ggttggacaa gacatacaaa caaaagcaca tgaagctaag 600

agatcaaca 609

<210> 41
<211> 609
<212> DNA
<213> Trifolium repens

<400> 41
atggcacaac ctcaagttca agtccactca acaacaacac accgtcaaga aactgctacc 60

tacccatcaa cccaaaacat tcgtaaagat gtttacgaaa atgttaacta tcccggccaa 120

cgcggtcgtt ataacgaccg ctataatgat agtggtcgtt atgatggtgg tattgcctcc 180

tttttgtcag agagaagtcc tccagcctct caaatcctcg ctaccgttgg aggatttttc 240

ataggtggta ctctattttt attagctagc atttcattta tcgccagtct tattggattg 300

gcgataatga caccactttt tatccttttt agcccggttt tagtccctgc tgccctcact 360

atagggctag cagtggctgg aatattgaca gcagatgctt gcgggttgac ggggcttatg 420

tcgttgtcgt ggaccgtgaa atatgttagg gatttacaag cagtagtgcc cgaacaaatg 480

gattcgatga agggacgtgt cgcggatgtc gcgagttatg ttggacaaaa gactaaggat 540

gttggacaaa aaactaaaga ggttggacaa gacatacaaa caaaagcaca tgaagctaag 600

agatcaaca 609

<210> 42
<211> 869
<212> DNA
<213> Trifolium repens

<400> 42
agatcaacag gtactatcga taccgtcgac ctcgacggta cttctatggc acaacctcaa 60

gttcaagtcc actcaacaac aacacaccgt caagaaactg ctacctaccc atcaacccaa 120

aacattcgta aagatgttta cgaaaatgtt aactatcccg gccaacgcgg tcgttataac 180

gaccgctata atgatagtgg tcgttatgat ggtggtattg cctccttttt gtcagagaga 240

agtcctccag cctctcaaat cctcgctacc gttggaggat ttttcatagg tggtactcta 300

tttttattag ctagcatttc atttatcgcc agtcttattg gattggcgat aatgacacca 360

ctttttatcc tttttagccc ggttttagtc cctgctgccc tcactatagg gctagcagtg 420

gctggaatat tgacagcaga tgcttgcggg ttgacggggc ttatgtcgtt gtcgtggacc 480

gtgaaatatg ttagggattt acaagcagta gtgcccgaac aaatggattc gatgaaggga 540

cgtgtcgcgg atgtcgcgag ttatgttgga caaaagacta aggatgttgg acaaaaaact 600

aaagaggttg gacaagacat acaaacaaaa gcacatgaag ctaagagatc aacactcgag 660

tgaaagggtg ggcgcgccga cccagctttc ttgtacaaag ttggcattat aagaaagcat 720

tgcttatcaa tttgttgcaa cgaacaggtc actatcagtc aaaataaaat cattatttgc 780

catccagctg atatccccta tagtgagtcg tattacatgg tcatagctgt ttcctggcag 840

ctctggcccg tgtctcaaaa tctctgatg 869

<210> 43
<211> 818
<212> DNA
<213> Trifolium repens

<400> 43
gacatacaaa caaaagcaca tgaagctaag agatcaacag gtactatcga taccgtcgac 60

ctcgacggta cttctatggc acaacctcaa gttcaagtcc actcaacaac aacacaccgt 120

caagaaactg ctacctaccc atcaacccaa aacattcgta aagatgttta cgaaaatgtt 180

aactatcccg gccaacgcgg tcgttataac gaccgctata atgatagtgg tcgttatgat 240

ggtggtattg cctccttttt gtcagagaga agtcctccag cctctcaaat cctcgctacc 300

gttggaggat ttttcatagg tggtactcta tttttattag ctagcatttc atttatcgcc 360

agtcttattg gattggcgat aatgacacca ctttttatcc tttttagccc ggttttagtc 420

cctgctgccc tcactatagg gctagcagtg gctggaatat tgacagcaga tgcttgcggg 480

ttgacggggc ttatgtcgtt gtcgtggacc gtgaaatatg ttagggattt acaagcagta 540

gtgcccgaac aaatggattc gatgaaggga cgtgtcgcgg atgtcgcgag ttatgttgga 600

caaaagacta aggatgttgg acaaaaaact aaagaggttg gacaagacat acaaacaaaa 660

gcacatgaag ctaagagatc aacactcgag tgaaagggtg ggcgcgccga cccagctttc 720

ttgtacaaag ttgscattat aagaaagcat tgcttatcaa tttgtkgcaa cgaacaggtc 780

actatcagtc aaaataaaat ckattatttg ctcgattg 818

<210> 44
<211> 304
<212> DNA
<213> Arabidopsis thaliana

<400> 44
gtaaatttct gtgttcctta ttctctcaaa atcttcgatt ttgttttcgt tcgatcccaa 60

tttcgtatat gttctttggt ttagattctg ttaatcttag atcgaagacg attttctggg 120

tttgatcgtt agatatcatc ttaattctcg attagggttt catagatatc atccgatttg 180

ttcaaataat ttgagttttg tcgaataatt actcttcgat ttgtgatttc tatctagatc 240

tggtgttagt ttctagtttg tgcgatcgaa tttgtcgatt aatctgagtt tttctgatta 300

acag 304

<210> 45

<211> 304
<212> DNA
<213> Arabidopsis thaliana

<400> 45
gtaaatttct gtgttcctta ttctctcaaa atcttcgatt ttgttttcgt tcgatcccaa 60

tttcgtatat gttctttggt ttagattctg ttaatcttag atcgaagacg attttctggg 120

tttgatcgtt agatatcatc ttaattctcg attagggttt catagatatc atccgatttg 180

ttcaaataat ttgagttttg tcgaataatt actcttcgat ttgtgatttc tatctagatc 240

tggtgttagt ttctagtttg tgcgatcgaa tttgtcgatt aatctgagtt tttctgatct 300

gcag 304

<210> 46
<211> 301
<212> DNA
<213> Arabidopsis thaliana

<400> 46
gtaaatttct gtgttcctta ttctctcaaa atcttcgatt ttgttttcgt tcgatcccaa 60

tttcgtatat gttctttggt ttagattctg ttaatcttag atcgaagacg attttctggg 120

tttgatcgtt agatatcatc ttaattctcg attagggttt catagatatc atccgatttg 180

ttcaaataat ttgagttttg tcgaataatt actcttcgat ttgtgatttc tatctagatc 240

tggtgttagt ttctagtttg tgcgatcgaa tttgtcgatt aatctgagtt tttctgatca 300

g 301

<210> 47
<211> 870
<212> PRT
<213> Sesamum indicum

<400> 47

Met Ala Glu His Tyr Gly Gln Gln Gln Gln Thr Arg Ala Pro His Leu
1 5 10 15

Gln Leu Gln Pro Arg Ala Gln Arg Val Val Lys Ala Ala Thr Ala Val
20 25 30

Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Gly Leu Thr Leu Ala Gly
35 40 45

Thr Val Ile Ala Leu Thr Ile Ala Thr Pro Leu Leu Val Ile Phe Ser
50 55 60

Pro Val Leu Val Pro Ala Val Ile Thr Ile Phe Leu Leu Gly Ala Gly
65 70 75 80

Phe Leu Ala Ser Gly Gly Phe Gly Val Ala Ala Leu Ser Val Leu Ser
85 90 95

```
Trp Ile Tyr Arg Tyr Leu Thr Gly Lys His Pro Pro Gly Ala Asp Gln
            100                 105                 110

Leu Glu Ser Ala Lys Thr Lys Leu Ala Ser Lys Ala Arg Glu Met Lys
        115                 120                 125

Asp Arg Ala Glu Gln Phe Ser Gln Gln Pro Val Ala Gly Ser Gln Thr
    130                 135                 140

Ser Met Ala Glu His Tyr Gly Gln Gln Gln Gln Thr Arg Ala Pro His
145                 150                 155                 160

Leu Gln Leu Gln Pro Arg Ala Gln Arg Val Val Lys Ala Ala Thr Ala
                165                 170                 175

Val Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Gly Leu Thr Leu Ala
            180                 185                 190

Gly Thr Val Ile Ala Leu Thr Ile Ala Thr Pro Leu Leu Val Ile Phe
        195                 200                 205

Ser Pro Val Leu Val Pro Ala Val Ile Thr Ile Phe Leu Leu Gly Ala
    210                 215                 220

Gly Phe Leu Ala Ser Gly Gly Phe Gly Val Ala Ala Leu Ser Val Leu
225                 230                 235                 240

Ser Trp Ile Tyr Arg Tyr Leu Thr Gly Lys His Pro Pro Gly Ala Asp
                245                 250                 255

Gln Leu Glu Ser Ala Lys Thr Lys Leu Ala Ser Lys Ala Arg Glu Met
            260                 265                 270

Lys Asp Arg Ala Glu Gln Phe Ser Gln Gln Pro Val Ala Gly Ser Gln
        275                 280                 285

Thr Ser Met Ala Glu His Tyr Gly Gln Gln Gln Gln Thr Arg Ala Pro
    290                 295                 300

His Leu Gln Leu Gln Pro Arg Ala Gln Arg Val Val Lys Ala Ala Thr
305                 310                 315                 320

Ala Val Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Gly Leu Thr Leu
                325                 330                 335

Ala Gly Thr Val Ile Ala Leu Thr Ile Ala Thr Pro Leu Leu Val Ile
            340                 345                 350

Phe Ser Pro Val Leu Val Pro Ala Val Ile Thr Ile Phe Leu Leu Gly
        355                 360                 365
```

```
Ala Gly Phe Leu Ala Ser Gly Gly Phe Gly Val Ala Ala Leu Ser Val
    370                 375                 380

Leu Ser Trp Ile Tyr Arg Tyr Leu Thr Gly Lys His Pro Pro Gly Ala
385                 390                 395                 400

Asp Gln Leu Glu Ser Ala Lys Thr Lys Leu Ala Ser Lys Ala Arg Glu
                405                 410                 415

Met Lys Asp Arg Ala Glu Gln Phe Ser Gln Gln Pro Val Ala Gly Ser
            420                 425                 430

Gln Thr Ser Met Ala Glu His Tyr Gly Gln Gln Gln Gln Thr Arg Ala
        435                 440                 445

Pro His Leu Gln Leu Gln Pro Arg Ala Gln Arg Val Val Lys Ala Ala
    450                 455                 460

Thr Ala Val Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Gly Leu Thr
465                 470                 475                 480

Leu Ala Gly Thr Val Ile Ala Leu Thr Ile Ala Thr Pro Leu Leu Val
                485                 490                 495

Ile Phe Ser Pro Val Leu Val Pro Ala Val Ile Thr Ile Phe Leu Leu
            500                 505                 510

Gly Ala Gly Phe Leu Ala Ser Gly Gly Phe Gly Val Ala Ala Leu Ser
        515                 520                 525

Val Leu Ser Trp Ile Tyr Arg Tyr Leu Thr Gly Lys His Pro Pro Gly
    530                 535                 540

Ala Asp Gln Leu Glu Ser Ala Lys Thr Lys Leu Ala Ser Lys Ala Arg
545                 550                 555                 560

Glu Met Lys Asp Arg Ala Glu Gln Phe Ser Gln Gln Pro Val Ala Gly
                565                 570                 575

Ser Gln Thr Ser Met Ala Glu His Tyr Gly Gln Gln Gln Gln Thr Arg
            580                 585                 590

Ala Pro His Leu Gln Leu Gln Pro Arg Ala Gln Arg Val Val Lys Ala
        595                 600                 605

Ala Thr Ala Val Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Gly Leu
    610                 615                 620

Thr Leu Ala Gly Thr Val Ile Ala Leu Thr Ile Ala Thr Pro Leu Leu
625                 630                 635                 640
```

```
Val Ile Phe Ser Pro Val Leu Val Pro Ala Val Ile Thr Ile Phe Leu
                645                 650                 655

Leu Gly Ala Gly Phe Leu Ala Ser Gly Gly Phe Gly Val Ala Ala Leu
            660                 665                 670

Ser Val Leu Ser Trp Ile Tyr Arg Tyr Leu Thr Gly Lys His Pro Pro
        675                 680                 685

Gly Ala Asp Gln Leu Glu Ser Ala Lys Thr Lys Leu Ala Ser Lys Ala
    690                 695                 700

Arg Glu Met Lys Asp Arg Ala Glu Gln Phe Ser Gln Gln Pro Val Ala
705                 710                 715                 720

Gly Ser Gln Thr Ser Met Ala Glu His Tyr Gly Gln Gln Gln Gln Thr
                725                 730                 735

Arg Ala Pro His Leu Gln Leu Gln Pro Arg Ala Gln Arg Val Val Lys
            740                 745                 750

Ala Ala Thr Ala Val Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Gly
        755                 760                 765

Leu Thr Leu Ala Gly Thr Val Ile Ala Leu Thr Ile Ala Thr Pro Leu
    770                 775                 780

Leu Val Ile Phe Ser Pro Val Leu Val Pro Ala Val Ile Thr Ile Phe
785                 790                 795                 800

Leu Leu Gly Ala Gly Phe Leu Ala Ser Gly Gly Phe Gly Val Ala Ala
                805                 810                 815

Leu Ser Val Leu Ser Trp Ile Tyr Arg Tyr Leu Thr Gly Lys His Pro
            820                 825                 830

Pro Gly Ala Asp Gln Leu Glu Ser Ala Lys Thr Lys Leu Ala Ser Lys
        835                 840                 845

Ala Arg Glu Met Lys Asp Arg Ala Glu Gln Phe Ser Gln Gln Pro Val
    850                 855                 860

Ala Gly Ser Gln Thr Ser
865                 870
```

<210> 48
<211> 839
<212> PRT
<213> Sesamum indicum

<400> 48

```
Met Ala Glu His Tyr Gly Gln Gln Gln Gln Thr Arg Ala Pro His Leu
```

1 5 10 15

Gln Leu Gln Pro Arg Ala Gln Arg Val Val Lys Ala Ala Thr Ala Val
20 25 30

Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Gly Leu Thr Leu Ala Gly
35 40 45

Thr Val Ile Ala Leu Thr Ile Ala Thr Pro Leu Leu Val Ile Phe Ser
50 55 60

Pro Val Leu Val Pro Ala Val Ile Thr Ile Phe Leu Leu Gly Ala Gly
65 70 75 80

Phe Leu Ala Ser Gly Gly Phe Gly Val Ala Ala Leu Ser Val Leu Ser
85 90 95

Trp Ile Tyr Arg Tyr Leu Thr Gly Lys His Pro Pro Gly Ala Asp Gln
100 105 110

Leu Glu Ser Ala Lys Thr Lys Leu Ala Ser Lys Ala Arg Glu Met Lys
115 120 125

Asp Arg Ala Glu Gln Phe Ser Gln Gln Pro Val Ala Gly Ser Gln Thr
130 135 140

Ser Ile Asp Pro Ser Ser Trp Leu Glu Met Ala Glu His Tyr Gly Gln
145 150 155 160

Gln Gln Gln Thr Arg Ala Pro His Leu Gln Leu Gln Pro Arg Ala Gln
165 170 175

Arg Val Val Lys Ala Ala Thr Ala Val Thr Ala Gly Gly Ser Leu Leu
180 185 190

Val Leu Ser Gly Leu Thr Leu Ala Gly Thr Val Ile Ala Leu Thr Ile
195 200 205

Ala Thr Pro Leu Leu Val Ile Phe Ser Pro Val Leu Val Pro Ala Val
210 215 220

Ile Thr Ile Phe Leu Leu Gly Ala Gly Phe Leu Ala Ser Gly Gly Phe
225 230 235 240

Gly Val Ala Ala Leu Ser Val Leu Ser Trp Ile Tyr Arg Tyr Leu Thr
245 250 255

Gly Lys His Pro Pro Gly Ala Asp Gln Leu Glu Ser Ala Lys Thr Lys
260 265 270

Leu Ala Ser Lys Ala Arg Glu Met Lys Asp Arg Ala Glu Gln Phe Ser

275 280 285

Gln Gln Pro Val Ala Gly Ser Gln Thr Ser His Met Phe Lys Trp Pro
290 295 300

Ser Ala Met Ala Glu His Tyr Gly Gln Gln Gln Gln Thr Arg Ala Pro
305 310 315 320

His Leu Gln Leu Gln Pro Arg Ala Gln Arg Val Val Lys Ala Ala Thr
325 330 335

Ala Val Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Gly Leu Thr Leu
340 345 350

Ala Gly Thr Val Ile Ala Leu Thr Ile Ala Thr Pro Leu Leu Val Ile
355 360 365

Phe Ser Pro Val Leu Val Pro Ser Ser Ser Glu Leu Pro Trp Val Asp
370 375 380

Met Ala Glu His Tyr Gly Gln Gln Gln Gln Thr Arg Ala Pro His Leu
385 390 395 400

Gln Leu Gln Pro Arg Ala Gln Arg Val Val Lys Ala Ala Thr Ala Val
405 410 415

Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Gly Leu Thr Leu Ala Gly
420 425 430

Thr Val Ile Ala Leu Thr Ile Ala Thr Pro Leu Leu Val Ile Phe Ser
435 440 445

Pro Val Leu Val Pro Ala Val Ile Thr Ile Phe Leu Leu Gly Ala Gly
450 455 460

Phe Leu Ala Ser Gly Gly Phe Gly Val Ala Ala Leu Ser Val Leu Ser
465 470 475 480

Trp Ile Tyr Arg Tyr Leu Thr Gly Lys His Pro Pro Gly Ala Asp Gln
485 490 495

Leu Glu Ser Ala Lys Thr Lys Leu Ala Ser Lys Ala Arg Glu Met Lys
500 505 510

Asp Arg Ala Glu Gln Phe Ser Gln Gln Pro Val Ala Gly Ser Gln Thr
515 520 525

Ser His Met Glu Phe Lys Leu Ser Thr Met Ala Glu His Tyr Gly Gln
530 535 540

Gln Gln Gln Thr Arg Ala Pro His Leu Gln Leu Gln Pro Arg Ala Gln

545 550 555 560

Arg Val Val Lys Ala Ala Thr Ala Val Thr Ala Gly Gly Ser Leu Leu
565 570 575

Val Leu Ser Gly Leu Thr Leu Ala Gly Thr Val Ile Ala Leu Thr Ile
580 585 590

Ala Thr Pro Leu Leu Val Ile Phe Ser Pro Val Leu Val Pro Ala Val
595 600 605

Ile Thr Ile Phe Leu Leu Gly Ala Gly Phe Leu Ala Ser Gly Gly Phe
610 615 620

Gly Val Ala Ala Leu Ser Val Leu Ser Trp Ile Tyr Arg Tyr Leu Thr
625 630 635 640

Gly Lys His Pro Pro Gly Ala Asp Gln Leu Glu Ser Ala Lys Thr Lys
645 650 655

Leu Ala Ser Lys Ala Arg Glu Met Lys Asp Arg Ala Glu Gln Phe Ser
660 665 670

Gln Gln Pro Val Ala Gly Ser Gln Thr Ser Ile Asp Gln Gln Val Asn
675 680 685

Val His Met Ala Glu His Tyr Gly Gln Gln Gln Gln Thr Arg Ala Pro
690 695 700

His Leu Gln Leu Gln Pro Arg Ala Gln Arg Val Val Lys Ala Ala Thr
705 710 715 720

Ala Val Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Gly Leu Thr Leu
725 730 735

Ala Gly Thr Val Ile Ala Leu Thr Ile Ala Thr Pro Leu Leu Val Ile
740 745 750

Phe Ser Pro Val Leu Val Pro Ala Val Ile Thr Ile Phe Leu Leu Gly
755 760 765

Ala Gly Phe Leu Ala Ser Gly Gly Phe Gly Val Ala Ala Leu Ser Val
770 775 780

Leu Ser Trp Ile Tyr Arg Tyr Leu Thr Gly Lys His Pro Pro Gly Ala
785 790 795 800

Asp Gln Leu Glu Ser Ala Lys Thr Lys Leu Ala Ser Lys Ala Arg Glu
805 810 815

Met Lys Asp Arg Ala Glu Gln Phe Ser Gln Gln Pro Val Ala Gly Ser

820 825 830

Gln Thr Ser Pro Trp Leu Glu
835

<210> 49
<211> 869
<212> PRT
<213> Sesamum indicum

<400> 49

Met Ala Glu His Tyr Gly Gln Gln Gln Gln Thr Arg Ala Pro His Leu
1 5 10 15

Gln Leu Gln Pro Arg Ala Gln Arg Val Val Lys Ala Ala Thr Ala Val
20 25 30

Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Gly Leu Thr Leu Ala Gly
35 40 45

Thr Val Ile Ala Leu Thr Ile Ala Thr Pro Leu Leu Val Ile Phe Ser
50 55 60

Pro Val Leu Val Pro Ala Val Ile Thr Ile Phe Leu Leu Gly Ala Gly
65 70 75 80

Phe Leu Ala Ser Gly Gly Phe Gly Val Ala Ala Leu Ser Val Leu Ser
85 90 95

Trp Ile Tyr Arg Tyr Leu Thr Gly Lys His Pro Pro Gly Ala Asp Gln
100 105 110

Leu Glu Ser Ala Lys Thr Lys Leu Ala Ser Lys Ala Arg Glu Met Lys
115 120 125

Asp Arg Ala Glu Gln Phe Ser Gln Gln Pro Val Ala Gly Ser Gln Thr
130 135 140

Ser Met Ala Glu His Tyr Gly Gln Gln Gln Gln Thr Arg Ala Pro His
145 150 155 160

Leu Gln Leu Gln Pro Arg Ala Gln Arg Val Val Lys Ala Ala Thr Ala
165 170 175

Val Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Gly Leu Thr Leu Ala
180 185 190

Gly Thr Val Ile Ala Leu Thr Ile Ala Thr Pro Leu Leu Val Ile Phe
195 200 205

Ser Pro Val Leu Val Pro Ala Val Ile Thr Ile Phe Leu Leu Gly Ala
210 215 220

Gly Phe Leu Ala Ser Gly Gly Phe Gly Val Ala Ala Leu Ser Val Leu
225 230 235 240

Ser Trp Ile Tyr Arg Tyr Leu Thr Gly Lys His Pro Pro Gly Ala Asp
245 250 255

Gln Leu Glu Ser Ala Lys Thr Lys Leu Ala Ser Lys Ala Arg Glu Met
260 265 270

Lys Asp Arg Ala Glu Gln Phe Ser Gln Gln Pro Val Ala Gly Ser Gln
275 280 285

Thr Ser Met Ala Glu His Tyr Gly Gln Gln Gln Thr Arg Ala Pro
290 295 300

His Leu Gln Leu Gln Pro Arg Ala Gln Arg Val Val Lys Ala Ala Thr
305 310 315 320

Ala Val Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Gly Leu Thr Leu
325 330 335

Ala Gly Thr Val Ile Ala Leu Thr Ile Ala Thr Pro Leu Leu Val Ile
340 345 350

Phe Ser Pro Val Leu Val Pro Ala Val Ile Thr Ile Phe Leu Leu Gly
355 360 365

Ala Gly Phe Leu Ala Ser Gly Gly Phe Gly Val Ala Ala Leu Ser Val
370 375 380

Leu Ser Trp Ile Tyr Arg Tyr Leu Thr Gly Lys His Pro Pro Gly Ala
385 390 395 400

Asp Gln Leu Glu Ser Ala Lys Thr Lys Leu Ala Ser Lys Ala Arg Glu
405 410 415

Met Lys Asp Arg Ala Glu Gln Phe Ser Gln Gln Pro Val Ala Gly Ser
420 425 430

Gln Thr Ser Met Ala Glu His Tyr Gly Gln Gln Gln Thr Arg Ala
435 440 445

Pro His Leu Gln Leu Gln Pro Arg Ala Gln Arg Val Val Lys Ala Ala
450 455 460

Thr Ala Val Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Gly Leu Thr
465 470 475 480

Leu Ala Gly Thr Val Ile Ala Leu Thr Ile Ala Thr Pro Leu Leu Val
485 490 495

Ile Phe Ser Pro Val Leu Val Pro Ala Val Ile Thr Ile Phe Leu Leu
500 505 510

Gly Ala Gly Phe Leu Ala Ser Gly Gly Phe Gly Val Ala Ala Leu Ser
515 520 525

Val Leu Ser Trp Ile Tyr Arg Tyr Leu Thr Gly Lys His Pro Pro Gly
530 535 540

Ala Asp Gln Leu Glu Ser Ala Lys Thr Lys Leu Ala Ser Lys Ala Arg
545 550 555 560

Glu Met Lys Asp Arg Ala Glu Gln Phe Ser Gln Gln Pro Val Ala Gly
565 570 575

Ser Gln Thr Ser Met Ala Glu His Tyr Gly Gln Gln Gln Gln Thr Arg
580 585 590

Ala Pro His Leu Gln Leu Gln Pro Arg Ala Gln Arg Val Val Lys Ala
595 600 605

Ala Thr Ala Val Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Gly Leu
610 615 620

Thr Leu Ala Gly Thr Val Ile Ala Leu Thr Ile Ala Thr Pro Leu Leu
625 630 635 640

Val Ile Phe Ser Pro Val Leu Val Pro Ala Val Ile Thr Ile Phe Leu
645 650 655

Leu Gly Ala Gly Phe Leu Ala Ser Gly Gly Phe Gly Val Ala Ala Leu
660 665 670

Ser Val Leu Ser Trp Ile Tyr Arg Tyr Leu Thr Gly Lys His Pro Pro
675 680 685

Gly Ala Asp Gln Leu Glu Ser Ala Lys Thr Lys Leu Ala Ser Lys Ala
690 695 700

Arg Glu Met Lys Asp Arg Ala Glu Gln Phe Ser Gln Gln Pro Val Ala
705 710 715 720

Gly Ser Gln Thr Met Ala Glu His Tyr Gly Gln Gln Gln Gln Thr Arg
725 730 735

Ala Pro His Leu Gln Leu Gln Pro Arg Ala Gln Arg Val Val Lys Ala
740 745 750

Ala Thr Ala Val Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Gly Leu
755 760 765

Thr Leu Ala Gly Thr Val Ile Ala Leu Thr Ile Ala Thr Pro Leu Leu
770 775 780

Val Ile Phe Ser Pro Val Leu Val Pro Ala Val Ile Thr Ile Phe Leu
785 790 795 800

Leu Gly Ala Gly Phe Leu Ala Ser Gly Gly Phe Gly Val Ala Ala Leu
805 810 815

Ser Val Leu Ser Trp Ile Tyr Arg Tyr Leu Thr Gly Lys His Pro Pro
820 825 830

Gly Ala Asp Gln Leu Glu Ser Ala Lys Thr Lys Leu Ala Ser Lys Ala
835 840 845

Arg Glu Met Lys Asp Arg Ala Glu Gln Phe Ser Gln Gln Pro Val Ala
850 855 860

Gly Ser Gln Thr Ser
865

<210> 50
<211> 3289
<212> DNA
<213> Sesamum indicum/Synthetic

<400> 50
aataatgatt ttattttgac tgatagtgac ctgttcgttg caacaaattg atgagcaatg 60

ctttttttata atgccaactt tgtacaaaaa agcaggctcc gcggccgctt gctccgttaa 120

aaaaaaccat ggcagagcac tacgggcagc agcaacaaac acgtgccccc cacttgcaac 180

tgcaaccgcg tgctcagcgt gttgtaaagg cagcgaccgc ggttactgcc ggaggtagcc 240

tgttggtgtt atccgggttg accctggctg gaacggtcat tgcgctgaca atcgccacac 300

ctctcctggt gatcttctcg cctgtactgg tccctgcagg taaatttctg tgttccttat 360

tctctcaaaa tcttcgattt tgttttcgtt cgatcccaat ttcgtatatg ttctttggtt 420

tagattctgt taatcttaga tcgaagacga ttttctgggt ttgatcgtta gatatcatct 480

taattctcga ttagggtttc atagatatca tccgatttgt tcaaataatt tgagttttgt 540

cgaataatta ctcttcgatt tgtgatttct atctagatct ggtgttagtt tctagtttgt 600

gcgatcgaat ttgtcgatta atctgagttt ttctgatctg cagtcatcac catcttcctg 660

ttgggagcgg gttttctggc aagcggggga tttggtgttg ccgctctgtc tgtgctgtcc 720

tggatctacc gttacctgac ggggaaacat ccacccggag cggatcagtt ggagtcggcc 780

aagacaaagt tggcgagcaa ggcccgtgaa atgaaggacc gtgccgagca gttcagtcag 840

caaccggtag cagggtctca gaccagcatc gatccatcct cctggctcga gatggcggaa 900

cactacgggc aacaacagca gactcgtgct ccccacctgc aattacaacc ccgtgcccaa 960

cgtgttgtga aagcggcaac agcagtaacg gcaggggaa gtttgctggt cttatcgggg 1020

ttgaccttag cgggaaccgt gattgccctg acaattgcga ctccgctgct ggttatcttc 1080

agccccgtat tggttccggc cgtgatcacg attttttgc tgggggcagg atttttagcc 1140

agcggaggat ttggggtcgc agcgttgtct gtgctgagtt ggatctatcg ttatttgacc 1200

gggaagcacc cacctggagc agaccagctg gagagcgcga aaacgaagct ggcatcgaag 1260

gcgcgtgaaa tgaaggatcg tgctgaacaa ttctcccagc agcctgttgc cggttctcag 1320

accagccata tgtttaaatg gccaagcgct atggccgagc attatgggca gcaacagcaa 1380

acccgtgccc cgcatctgca attgcaacct cgtgcccagc gtgtcgttaa ggcggctact 1440

gcggtaacag cgggagggag cttactggta ttaagcgggc tgacattggc cggaacggtg 1500

atcgccttaa caatcgcgac acccttgctg gtcatcttca gtccggttct ggtgcccgcg 1560

gtgattacga ttttcctgct gggagccggt ttcttagcat cgggggttt tggggtagca 1620

gccttgagtg tcctgtcgtg gatctatcgt tacttaactg gaaaacaccc gccaggagct 1680

gaccagttgg agtctgcaaa aactaagctg gcgtccaaag cccgtgaaat gaaggatcgt 1740

gctgagcagt ttagccagca gccagttgcg ggaagtcaga cctcttcatc tgagctccca 1800

tgggtcgaca tggcggagca ttacggtcaa cagcaacaga cccgtgctcc gcacttacaa 1860

ttgcaaccac gtgctcaacg tgtcgtaaaa gccgccacgg cagttactgc ggggggatca 1920

ttgctggtgt taagtgggtt gacactggcg gggacagtta ttgcactgac gatcgcgacc 1980

cccttgttag tgatcttctc ccccgttctg gttccggcgg tcattacaat ctttctgttg 2040

ggtgccggat ttttagcctc tgggggattt ggagtagctg ccctgtcagt gttgagctgg 2100

atctaccgtt acttaacagg gaagcaccct cccggggcag atcagttgga aagcgccaag 2160

accaagctgg caagtaaagc gcgtgaaatg aaggaccgtg ccgaacaatt ttcgcagcaa 2220

ccggttgcgg gatcacagac ctctagtact ccatcctcct ggcatatgat ggccgagcac 2280

tatggacaac agcagcagac gcgtgccccc catctgcaac tgcaaccccg tgctcaacgt 2340

gtcgttaagg ctgcgacagc agtaaccgct gggggttctc tgttagtgtt gtcagggctg 2400

actttggcgg ggacggtaat tgcgttgacc attgccaccc cgctgttagt gattttcagc 2460

ccggtactgg tgccagcagt tatcacgatc ttcttgctgg gtgccggatt cttggcaagt 2520

ggaggttttg gagttgcggc gctgtcagtt ttatcctgga tctatcgtta tctgacagga 2580

aaacatcccc caggtgccga tcagctggag agtgccaaga caaaactggc gtctaaggca 2640

cgtgaaatga aggatcgtgc cgaacagttt tctcaacagc ccgtagcggg gtcacagacc 2700

tcgatcgatc agcaggttaa cgtgcacatg gccgaacatt acggacagca acaacagacg 2760

cgtgctccac acctgcaatt gcaaccgcgt gctcaacgtg ttgtcaaagc ggcgaccgcc 2820

gtaacagcag gaggatcact gttagtgctg tcgggtttaa ccttggccgg gaccgtcatt 2880

gcattgacta ttgcgacgcc cttactggtg atcttttctc cggtgctggt tcccgccgtt 2940

attaccatct tcttgttagg ggcaggattc ctggcatcag ggggattcgg agttgcggcg 3000

ttgagtgtct taagttggat ctaccgttat ctgactggaa agcacccgcc tggggccgat 3060

caactggagt cagccaaaac gaaattggcg tcaaaagcgc gtgaaatgaa ggaccgtgct 3120

gagcagtttt ctcagcagcc tgtggcagga tcccagacat caccatggct cgagtaatga 3180

agcggccgca cccagctttc ttgtacaaag ttggcattat aagaaagcat tgcttatcaa 3240

tttgttgcaa cgaacaggtc actatcagtc aaaataaaat cattatttg 3289

<210> 51
<211> 914
<212> PRT
<213> Sesamum indicum/Synthetic

<400> 51

Met Ala Glu His Tyr Gly Gln Gln Gln Gln Thr Arg Ala Pro His Leu
1 5 10 15

Gln Leu Gln Pro Arg Ala Gln Arg Val Val Lys Ala Ala Thr Ala Val
20 25 30

Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Gly Leu Thr Leu Ala Gly
35 40 45

Thr Val Ile Ala Leu Thr Ile Ala Thr Pro Leu Leu Val Ile Phe Ser
50 55 60

Pro Val Leu Val Pro Ala Val Ile Thr Ile Phe Leu Leu Gly Ala Gly
65 70 75 80

Phe Leu Ala Ser Gly Gly Phe Gly Val Ala Ala Leu Ser Val Leu Ser
85 90 95

Trp Ile Tyr Arg Tyr Leu Thr Gly Lys His Pro Pro Gly Ala Asp Gln
100 105 110

Leu Glu Ser Ala Lys Thr Lys Leu Ala Ser Lys Ala Arg Glu Met Lys
115 120 125

Asp Arg Ala Glu Gln Phe Ser Gln Gln Pro Val Ala Gly Ser Gln Thr
130 135 140

Ser Ile Asp Pro Ser Ser Trp Leu Glu Met Ala Glu His Tyr Gly Gln
145 150 155 160

Gln Gln Gln Thr Arg Ala Pro His Leu Gln Leu Gln Pro Arg Ala Gln
165 170 175

Arg Val Val Lys Ala Ala Thr Ala Val Thr Ala Gly Gly Ser Leu Leu
180 185 190

Val Leu Ser Gly Leu Thr Leu Ala Gly Thr Val Ile Ala Leu Thr Ile
195 200 205

Ala Thr Pro Leu Leu Val Ile Phe Ser Pro Val Leu Val Pro Ala Val
210 215 220

Ile Thr Ile Phe Leu Leu Gly Ala Gly Phe Leu Ala Ser Gly Gly Phe
225 230 235 240

Gly Val Ala Ala Leu Ser Val Leu Ser Trp Ile Tyr Arg Tyr Leu Thr
245 250 255

Gly Lys His Pro Pro Gly Ala Asp Gln Leu Glu Ser Ala Lys Thr Lys
260 265 270

Leu Ala Ser Lys Ala Arg Glu Met Lys Asp Arg Ala Glu Gln Phe Ser
275 280 285

Gln Gln Pro Val Ala Gly Ser Gln Thr Ser His Met Phe Lys Trp Pro
290 295 300

Ser Ala Met Ala Glu His Tyr Gly Gln Gln Gln Gln Thr Arg Ala Pro
305 310 315 320

His Leu Gln Leu Gln Pro Arg Ala Gln Arg Val Val Lys Ala Ala Thr
325 330 335

Ala Val Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Gly Leu Thr Leu
340 345 350

Ala Gly Thr Val Ile Ala Leu Thr Ile Ala Thr Pro Leu Leu Val Ile
355 360 365

Phe Ser Pro Val Leu Val Pro Ala Val Ile Thr Ile Phe Leu Leu Gly
370 375 380

Ala Gly Phe Leu Ala Ser Gly Gly Phe Gly Val Ala Ala Leu Ser Val
385 390 395 400

Leu Ser Trp Ile Tyr Arg Tyr Leu Thr Gly Lys His Pro Pro Gly Ala
405 410 415

Asp Gln Leu Glu Ser Ala Lys Thr Lys Leu Ala Ser Lys Ala Arg Glu
420 425 430

Met Lys Asp Arg Ala Glu Gln Phe Ser Gln Gln Pro Val Ala Gly Ser
435 440 445

Gln Thr Ser Ser Ser Glu Leu Pro Trp Val Asp Met Ala Glu His Tyr
450 455 460

Gly Gln Gln Gln Gln Thr Arg Ala Pro His Leu Gln Leu Gln Pro Arg
465 470 475 480

Ala Gln Arg Val Val Lys Ala Ala Thr Ala Val Thr Ala Gly Gly Ser
485 490 495

Leu Leu Val Leu Ser Gly Leu Thr Leu Ala Gly Thr Val Ile Ala Leu
500 505 510

Thr Ile Ala Thr Pro Leu Leu Val Ile Phe Ser Pro Val Leu Val Pro
515 520 525

Ala Val Ile Thr Ile Phe Leu Leu Gly Ala Gly Phe Leu Ala Ser Gly
530 535 540

Gly Phe Gly Val Ala Ala Leu Ser Val Leu Ser Trp Ile Tyr Arg Tyr
545 550 555 560

Leu Thr Gly Lys His Pro Pro Gly Ala Asp Gln Leu Glu Ser Ala Lys
565 570 575

Thr Lys Leu Ala Ser Lys Ala Arg Glu Met Lys Asp Arg Ala Glu Gln
580 585 590

Phe Ser Gln Gln Pro Val Ala Gly Ser Gln Thr Ser Ser Thr Pro Ser
595 600 605

Ser Trp His Met Met Ala Glu His Tyr Gly Gln Gln Gln Thr Arg
610 615 620

Ala Pro His Leu Gln Leu Gln Pro Arg Ala Gln Arg Val Val Lys Ala
625 630 635 640

Ala Thr Ala Val Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Gly Leu
645 650 655

Thr Leu Ala Gly Thr Val Ile Ala Leu Thr Ile Ala Thr Pro Leu Leu
660 665 670

Val Ile Phe Ser Pro Val Leu Val Pro Ala Val Ile Thr Ile Phe Leu
675 680 685

Leu Gly Ala Gly Phe Leu Ala Ser Gly Gly Phe Gly Val Ala Ala Leu
690 695 700

Ser Val Leu Ser Trp Ile Tyr Arg Tyr Leu Thr Gly Lys His Pro Pro
705 710 715 720

Gly Ala Asp Gln Leu Glu Ser Ala Lys Thr Lys Leu Ala Ser Lys Ala
725 730 735

Arg Glu Met Lys Asp Arg Ala Glu Gln Phe Ser Gln Gln Pro Val Ala
740 745 750

Gly Ser Gln Thr Ser Ile Asp Gln Gln Val Asn Val His Met Ala Glu
755 760 765

His Tyr Gly Gln Gln Gln Gln Thr Arg Ala Pro His Leu Gln Leu Gln
770 775 780

Pro Arg Ala Gln Arg Val Val Lys Ala Ala Thr Ala Val Thr Ala Gly
785 790 795 800

Gly Ser Leu Leu Val Leu Ser Gly Leu Thr Leu Ala Gly Thr Val Ile
805 810 815

Ala Leu Thr Ile Ala Thr Pro Leu Leu Val Ile Phe Ser Pro Val Leu
820 825 830

Val Pro Ala Val Ile Thr Ile Phe Leu Leu Gly Ala Gly Phe Leu Ala
835 840 845

Ser Gly Gly Phe Gly Val Ala Ala Leu Ser Val Leu Ser Trp Ile Tyr
850 855 860

Arg Tyr Leu Thr Gly Lys His Pro Pro Gly Ala Asp Gln Leu Glu Ser
865 870 875 880

Ala Lys Thr Lys Leu Ala Ser Lys Ala Arg Glu Met Lys Asp Arg Ala
885 890 895

Glu Gln Phe Ser Gln Gln Pro Val Ala Gly Ser Gln Thr Ser Pro Trp
900 905 910

Leu Glu

<210> 52
<211> 3291
<212> DNA
<213> Sesamum indicum/Synthetic

<400> 52
caaataatga ttttattttg actgatagtg acctgttcgt tgcaacaaat tgatgagcaa 60

tgctttttta taatgccaac tttgtacaaa aaagcaggct ccgcggccgc ttgctccgtt 120

aaaaaaaacc atggcagagc actacgggca gcagcaacaa acacgtgccc cccacttgca 180

actgcaaccg cgtgctcagc gtgttgtaaa ggcagcgacc gcggttactg ccggaggtag 240

cctgttggtg ttatccgggt tgaccctggc tggaacggtc attgcgctga caatcgccac 300

acctctcctg gtgatcttct cgcctgtact ggtccctgca ggtaaatttc tgtgttcctt 360

attctctcaa aatcttcgat tttgttttcg ttcgatccca atttcgtata tgttctttgg 420

tttagattct gttaatctta gatcgaagac gattttctgg gtttgatcgt tagatatcat 480

cttaattctc gattagggtt tcatagatat catccgattt gttcaaataa tttgagtttt 540

gtcgaataat tactcttcga tttgtgattt ctatctagat ctggtgttag tttctagttt 600

gtgcgatcga atttgtcgat taatctgagt ttttctgatc tgcagtcatc accatcttcc 660

tgttgggagc gggttttctg gcaagcgggg gatttggtgt tgccgctctg tctgtgctgt 720

cctggatcta ccgttacctg acggggaaac atccacccgg agcggatcag ttggagtcgg 780

ccaagacaaa gttggcgagc aaggcccgtg aaatgaagga ccgtgccgag cagttcagtc 840

agcaaccggt agcagggtct cagaccagca tcgatccatc ctcctggctc gagatggcgg 900

aacactacgg gcaacaacag cagactcgtg ctccccacct gcaattacaa ccccgtgccc 960

aacgtgttgt gaaagcggca acagcagtaa cggcaggggg aagtttgctg gtcttatcgg 1020

ggttgacctt agcgggaacc gtgattgccc tgacaattgc gactccgctg ctggttatct 1080

tcagccccgt attggttccg gccgtgatca cgatttttt gctgggggca ggatttttag 1140

ccagcggagg atttggggtc gcagcgttgt ctgtgctgag ttggatctat cgttatttga 1200

ccgggaagca cccacctgga gcagaccagc tggagagcgc gaaaacgaag ctggcatcga 1260

aggcgcgtga aatgaaggat cgtgctgaac aattctccca gcagcctgtt gccggttctc 1320

agaccagcca tatgtttaaa tggccaagcg ctatggccga gcattatggg cagcaacagc 1380

aaacccgtgc cccgcatctg caattgcaac ctcgtgccca gcgtgtcgtt aaggcggcta 1440

ctgcggtaac agcgggaggg agcttactgg tattaagcgg gctgacattg gccggaacgg 1500

tgatcgcctt aacaatcgcg acacccttgc tggtcatctt cagtccggtt ctggtgcccg 1560

cggtgattac gattttcctg ctgggagccg gtttcttagc atcgggggggt tttggggtag 1620

cagccttgag tgtcctgtcg tggatctatc gttacttaac tggaaaacac ccgccaggag 1680

ctgaccagtt ggagtctgca aaaactaagc tggcgtccaa agcccgtgaa atgaaggatc 1740

gtgctgagca gtttagccag cagccagttg cgggaagtca gacctcttca tctgagctcc 1800

catgggtcga catggcggag cattacggtc aacagcaaca gacccgtgct ccgcacttac 1860

aattgcaacc acgtgctcaa cgtgtcgtaa aagccgccac ggcagttact gcgggggggat 1920

cattgctggt gttaagtggg ttgacactgg cggggacagt tattgcactg acgatcgcga 1980

ccccccttgtt agtgatcttc tcccccgttc tggttccggc ggtcattaca atctttctgt 2040

tgggtgccgg atttttagcc tctgggggat ttggagtagc tgccctgtca gtgttgagct 2100

ggatctaccg ttacttaaca gggaagcacc ctcccggggc agatcagttg gaaagcgcca 2160

agaccaagct ggcaagtaaa gcgcgtgaaa tgaaggaccg tgccgaacaa ttttcgcagc 2220

aaccggttgc gggatcacag acctctagta ctccatcctc ctggcatatg atggccgagc 2280

actatggaca acagcagcag acgcgtgccc ctcatctgca actgcaaccc cgtgctcaac 2340

gtgtcgttaa ggctgcgaca gcagtaaccg ctgggggttc tctgttagtg ttgtcagggc 2400

tgactttggc ggggacggta attgcgttga ccattgccac cccgctgtta gtgattttca 2460

gcccggtact ggtgccagca gttatcacga tcttcttgct gggtgccgga ttcttggcaa 2520

gtggaggttt tggagttgcg gcgctgtcag ttttatcctg gatctatcgt tatctgacag 2580

gaaaacatcc cccaggtgcc gatcagctgg agagtgccaa gacaaaactg gcgtctaagg 2640

cacgtgaaat gaaggatcgt gccgaacagt tttctcaaca gcccgtagcg gggtcacaga 2700

cctcgatcga tcagcaggtt aacgtgcaca tggccgaaca ttacggacag caacaacaga 2760

cgcgtgctcc acacctgcaa ttgcaaccgc gtgctcaacg tgttgtcaaa gcggcgaccg 2820

ccgtaacagc aggaggatca ctgttagtgc tgtcgggttt aaccttggcc gggaccgtca 2880

ttgcattgac tattgcgacg cccttactgg tgatcttttc tccggtgctg gttcccgccg 2940

ttattaccat cttcttgtta ggggcaggat tcctggcatc agggggattc ggagttgcgg 3000

cgttgagtgt cttaagttgg atctaccgtt atctgactgg aaagcacccg cctggggccg 3060

atcaactgga gtcagccaaa acgaaattgg cgtcaaaagc gcgtgaaatg aaggaccgtg 3120

ctgagcagtt ttctcagcag cctgtggcag gatcccagac atcaccatgg ctcgagtaat 3180

gaagcggccg cacccagctt tcttgtacaa agttggcatt ataagaaagc attgcttatc 3240

aatttgttgc aacgaacagg tcactatcag tcaaaataaa atcattattt g 3291

<210> 53
<211> 1350
<212> DNA
<213> Sesamum indicum/Synthetic

<400> 53
ttgctccgtt aaaaaaaacc atggctgagc attatggtca acaacagcag accagggcgc 60

ctcacctgca gctgcagccg cgcgcccagc gggtagtgaa ggcggccacc gccgtgacag 120

ccggcggctc gcttctcgtc ctctctggcc tcactttagc cggaactgtt attgcgctca 180

ccatcgccac tccgctgctt gtgatcttta gccccgttct ggtgccggcg gtcataacca 240

ttttcttgct gggtgcgggt tttctggcat ccggaggctt cggcgtggcg gcgctgagtg 300

tgctgtcgtg gatttacaga tatctgacag ggaaacaccc gccgggggcg gatcagctgg 360

aatcggcaaa gacgaagctg gcgagcaagg cgcgagagat gaaggatagg gcagagcagt 420

tctcgcagca gcctgttgga ggcggtggat ccggaggcgg tggtagtatg gctgagcatt 480

atggtcaaca acagcagacc agggcgcctc acctgcagct gcagccgcgc gcccagcggg 540

tagtgaaggc ggccaccgcc gtgacagccg gcggctcgct tctcgtcctc tctggcctca 600

ctttagccgg aactgttatt gcgctcacca tcgccactcc gctgcttgtg atctttagcc 660

ccgttctggt gccggcggtc ataaccattt tcttgctggg tgcgggtttt ctggcatccg 720

gaggcttcgg cgtggcggcg ctgagtgtgc tgtcgtggat ttacagatat ctgacaggga 780

aacacccgcc gggggcggat cagctggaat cggcaaagac gaagctggcg agcaaggcgc 840

gagagatgaa ggatagggca gagcagttct cgcagcagcc tgttgggggc ggtggatccg 900

gtggaggggg atccatggct gagcattatg gtcaacaaca gcagaccagg gcgcctcacc 960

tgcagctgca gccgcgcgcc cagcgggtag tgaaggcggc caccgccgtg acagccggcg 1020

gctcgcttct cgtcctctct ggcctcactt tagccggaac tgttattgcg ctcaccatcg 1080

ccactccgct gcttgtgatc tttagccccg ttctggtgcc ggcggtcata accattttct 1140

tgctgggtgc gggttttctg gcatccggag gcttcggcgt ggcggcgctg agtgtgctgt 1200

cgtggattta cagatatctg acagggaaac acccgccggg ggcggatcag ctggaatcgg 1260

caaagacgaa gctggcgagc aaggcgcgag agatgaagga tagggcagag cagttctcgc 1320

agcagcctgt tccatggctc gagtaatgaa 1350

<210> 54
<211> 2878
<212> DNA
<213> Sesamum indicum/Synthetic

<400> 54
gaaggagata tacatatgaa agaaaccgct gctgctaaat tcgaacgcca gcacatggac 60

agcccagatc tgggtaccct ggtgccacgc ggttccatgg ctgagcatta tggtcaacaa 120

cagcagacca gggcgcctca cctgcagctg cagccgcgcg cccagcgggt agtgaaggcg 180

gccaccgccg tgacagccgg cggctcgctt ctcgtcctct ctggcctcac tttagccgga 240

actgttattg cgctcaccat cgccactccg ctgcttgtga tctttagccc cgttctggtg 300

ccggcggtca taaccatttt cttgctgggt gcgggttttc tggcatccgg aggcttcggc 360

gtggcggcgc tgagtgtgct gtcgtggatt tacagatatc tgacagggaa acacccgccg 420

ggggcggatc agctggaatc ggcaaagacg aagctggcga gcaaggcgcg agagatgaag 480

gatagggcag agcagttctc gcagcagcct gttccatggc tgatatcgga tccgaattcg 540

agctccgtcg acaagcttgc ggccgcactc gagatggcgg aacactacgg gcaacaacag 600

cagactcgtg ctccccacct gcaattacaa ccccgtgccc aacgtgttgt gaaagcggca 660

acagcagtaa cggcaggggg aagtttgctg gtcttatcgg ggttgacctt agcgggaacc 720

gtgattgccc tgacaattgc gactccgctg ctggttatct tcagccccgt attggttccg 780

gccgtgatca cgatttttt gctgggggca ggatttttag ccagcggagg atttggggtc 840

gcagcgttgt ctgtgctgag ttggatctat cgttatttga ccgggaagca cccacctgga 900

gcagaccagc tggagagcgc gaaaacgaag ctggcatcga aggcgcgtga aatgaaggat 960

cgtgctgaac aattctccca gcagcctgtt gccggttctc agaccagcca tatgtttaaa 1020

tggccaagcg ctatggccga gcattatggg cagcaacagc aaacccgtgc cccgcatctg 1080

caattgcaac ctcgtgccca gcgtgtcgtt aaggcggcta ctgcggtaac agcgggaggg 1140

agcttactgg tattaagcgg gctgacattg gccggaacgg tgatcgcctt aacaatcgcg 1200

acacccttgc tggtcatctt cagtccggtt ctggtgcccg cggtgattac gattttcctg 1260

ctgggagccg gtttcttagc atcgggggt tttggggtag cagccttgag tgtcctgtcg 1320

tggatctatc gttacttaac tggaaaacac ccgccaggag ctgaccagtt ggagtctgca 1380

aaaactaagc tggcgtccaa agcccgtgaa atgaaggatc gtgctgagca gtttagccag 1440

cagccagttg cgggaagtca gacctcttca tctgagctcc catgggtcga catggcggag 1500

cattacggtc aacagcaaca gacccgtgct ccgcacttac aattgcaacc acgtgctcaa 1560

cgtgtcgtaa aagccgccac ggcagttact gcgggggggat cattgctggt gttaagtggg 1620

ttgacactgg cggggacagt tattgcactg acgatcgcga ccccccttgtt agtgatcttc 1680

tcccccgttc tggttccggc ggtcattaca atctttctgt tgggtgccgg atttttagcc 1740

tctgggggat ttggagtagc tgccctgtca gtgttgagct ggatctaccg ttacttaaca 1800

gggaagcacc ctcccggggc agatcagttg gaaagcgcca agaccaagct ggcaagtaaa 1860

gcgcgtgaaa tgaaggaccg tgccgaacaa ttttcgcagc aaccggttgc gggatcacag 1920

acctctagta ctccatcctc ctggcatatg atggccgagc actatggaca acagcagcag 1980

acgcgtgccc ctcatctgca actgcaaccc cgtgctcaac gtgtcgttaa ggctgcgaca 2040

gcagtaaccg ctgggggttc tctgttagtg ttgtcagggc tgactttggc ggggacggta 2100

attgcgttga ccattgccac cccgctgtta gtgattttca gcccggtact ggtgccagca 2160

gttatcacga tcttcttgct gggtgccgga ttcttggcaa gtggaggttt tggagttgcg 2220

gcgctgtcag ttttatcctg gatctatcgt tatctgacag gaaaacatcc cccaggtgcc 2280

gatcagctgg agagtgccaa gacaaaactg gcgtctaagg cacgtgaaat gaaggatcgt 2340

gccgaacagt tttctcaaca gcccgtagcg gggtcacaga cctcgatcga tcagcaggtt 2400

aacgtgcaca tggccgaaca ttacggacag caacaacaga cgcgtgctcc acacctgcaa 2460

ttgcaaccgc gtgctcaacg tgttgtcaaa gcggcgaccg ccgtaacagc aggaggatca 2520

ctgttagtgc tgtcgggttt aaccttggcc gggaccgtca ttgcattgac tattgcgacg 2580

cccttactgg tgatcttttc tccggtgctg gttcccgccg ttattaccat cttcttgtta 2640

ggggcaggat tcctggcatc agggggattc ggagttgcgg cgttgagtgt cttaagttgg 2700

atctaccgtt atctgactgg aaagcacccg cctggggccg atcaactgga gtcagccaaa 2760

acgaaattgg cgtcaaaagc gcgtgaaatg aaggaccgtg ctgagcagtt ttctcagcag 2820

cctgtggcag gatcccagac atcaccatgg ctcgagcacc accaccacca ccactgag 2878

<210> 55
<211> 1068
<212> DNA
<213> Sesamum indicum/Synthetic

<400> 55
gaaggagata tacatatgaa agaaaccgct gctgctaaat tcgaacgcca gcacatggac 60

agcccagatc tgggtaccct ggtgccacgc ggttccatgg ctgagcatta tggtcaacaa 120

cagcagacca gggcgcctca cctgcagctg cagccgcgcg cccagcgggt agtgaaggcg 180

gccaccgccg tgacagccgg cggctcgctt ctcgtcctct ctggcctcac tttagccgga 240

actgttattg cgctcaccat cgccactccg ctgcttgtga tctttagccc cgttctggtg 300

ccggcggtca taaccatttt cttgctgggt gcgggttttc tggcatccgg aggcttcggc 360

gtggcggcgc tgagtgtgct gtcgtggatt tacagatatc tgacagggaa acacccgccg 420

ggggcggatc agctggaatc ggcaaagacg aagctggcga gcaaggcgcg agagatgaag 480

gatagggcag agcagttctc gcagcagcct gttggaggcg gtggatccgg aggcggtggt 540
agtatggctg agcattatgg tcaacaacag cagaccaggg cgcctcacct gcagctgcag 600
ccgcgcgccc agcgggtagt gaaggcggcc accgccgtga cagccggcgg ctcgcttctc 660
gtcctctctg gcctcacttt agccggaact gttattgcgc tcaccatcgc cactccgctg 720
cttgtgatct ttagccccgt tctggtgccg gcggtcataa ccattttctt gctgggtgcg 780
ggttttctgg catccggagg cttcggcgtg gcggcgctga gtgtgctgtc gtggatttac 840
agatatctga cagggaaaca cccgccgggg gcggatcagc tggaatcggc aaagacgaag 900
ctggcgagca aggcgcgaga gatgaaggat agggcagagc agttctcgca gcagcctgtt 960
gggggcggtg gatccggtgg agggggatcc atggcgatat cggatccgaa ttcgagctcc 1020
gtcgacaagc ttgcggccgc actcgagcac caccaccacc accactga 1068

<210> 56
<211> 1929
<212> DNA
<213> Sesamum indicum/Synthetic

<400> 56
gaaggagata tacatatgaa agaaaccgct gctgctaaat tcgaacgcca gcacatggac 60
agcccagatc tgggtaccct ggtgccacgc ggttccatgg ctgagcatta tggtcaacaa 120
cagcagacca gggcgcctca cctgcagctg cagccgcgcg cccagcgggt agtgaaggcg 180
gccaccgccg tgacagccgg cggctcgctt ctcgtcctct ctggcctcac tttagccgga 240
actgttattg cgctcaccat cgccactccg ctgcttgtga tctttagccc cgttctggtg 300
ccggcggtca taaccatttt cttgctgggt gcgggttttc tggcatccgg aggcttcggc 360
gtggcggcgc tgagtgtgct gtcgtggatt tacagatatc tgacagggaa acacccgccg 420
ggggcggatc agctggaatc ggcaaagacg aagctggcga gcaaggcgcg agagatgaag 480
gatagggcag agcagttctc gcagcagcct gttccatggc gatatcggat ccgaattcga 540
gctccgtcga caagcttgcg gccgcttgct ccgttaaaaa aaaccatggc tgagcattat 600
ggtcaacaac agcagaccag ggcgcctcac ctgcagctgc agccgcgcgc ccagcgggta 660
gtgaaggcgg ccaccgccgt gacagccggc ggctcgcttc tcgtcctctc tggcctcact 720
ttagccggaa ctgttattgc gctcaccatc gccactccgc tgcttgtgat ctttagcccc 780
gttctggtgc cggcggtcat aaccattttc ttgctgggtg cgggttttct ggcatccgga 840
ggcttcggcg tggcggcgct gagtgtgctg tcgtggattt acagatatct gacagggaaa 900
cacccgccgg gggcggatca gctggaatcg gcaaagacga agctggcgag caaggcgcga 960
gagatgaagg atagggcaga gcagttctcg cagcagcctg ttggaggcgg tggatccgga 1020
ggcggtggta gtatggctga gcattatggt caacaacagc agaccagggc gcctcacctg 1080
cagctgcagc cgcgcgccca gcgggtagtg aaggcggcca ccgccgtgac agccggcggc 1140
tcgcttctcg tcctctctgg cctcacttta gccggaactg ttattgcgct caccatcgcc 1200

actccgctgc ttgtgatctt tagccccgtt ctggtgccgg cggtcataac cattttcttg 1260

ctgggtgcgg gttttctggc atccggaggc ttcggcgtgg cggcgctgag tgtgctgtcg 1320

tggatttaca gatatctgac agggaaacac ccgccggggg cggatcagct ggaatcggca 1380

aagacgaagc tggcgagcaa ggcgcgagag atgaaggata gggcagagca gttctcgcag 1440

cagcctgttg ggggcggtgg atccggtgga gggggatcca tggctgagca ttatggtcaa 1500

caacagcaga ccagggcgcc tcacctgcag ctgcagccgc gcgcccagcg ggtagtgaag 1560

gcggccaccg ccgtgacagc cggcggctcg cttctcgtcc tctctggcct cactttagcc 1620

ggaactgtta ttgcgctcac catcgccact ccgctgcttg tgatctttag ccccgttctg 1680

gtgccggcgg tcataaccat tttcttgctg ggtgcgggtt ttctggcatc cggaggcttc 1740

ggcgtggcgg cgctgagtgt gctgtcgtgg atttacagat atctgacagg gaaacacccg 1800

ccgggggcgg atcagctgga atcggcaaag acgaagctgg cgagcaaggc gcgagagatg 1860

aaggataggg cagagcagtt ctcgcagcag cctgttccat ggctcgagca ccaccaccac 1920

caccactga 1929

<210> 57
<211> 2403
<212> DNA
<213> Sesamum indicum/Synthetic

<400> 57
gaaggagata tacatatgaa agaaaccgct gctgctaaat tcgaacgcca gcacatggac 60

agcccagatc tgggtaccct ggtgccacgc ggttccatgg ctgagcatta tggtcaacaa 120

cagcagacca gggcgcctca cctgcagctg cagccgcgcg cccagcgggt agtgaaggcg 180

gccaccgccg tgacagccgg cggctcgctt ctcgtcctct ctggcctcac tttagccgga 240

actgttattg cgctcaccat cgccactccg ctgcttgtga tctttagccc cgttctggtg 300

ccggcggtca taaccatttt cttgctgggt gcgggttttc tggcatccgg aggcttcggc 360

gtggcggcgc tgagtgtgct gtcgtggatt tacagatatc tgacagggaa acacccgccg 420

ggggcggatc agctggaatc ggcaaagacg aagctggcga gcaaggcgcg agagatgaag 480

gatagggcag agcagttctc gcagcagcct gttggaggcg gtggatccgg aggcggtggt 540

agtatggctg agcattatgg tcaacaacag cagaccaggg cgcctcacct gcagctgcag 600

ccgcgcgccc agcgggtagt gaaggcggcc accgccgtga cagccggcgg ctcgcttctc 660

gtcctctctg gcctcacttt agccggaact gttattgcgc tcaccatcgc cactccgctg 720

cttgtgatct ttagccccgt tctggtgccg gcggtcataa ccattttctt gctgggtgcg 780

ggttttctgg catccggagg cttcggcgtg gcggcgctga gtgtgctgtc gtggatttac 840

agatatctga cagggaaaca cccgccgggg gcggatcagc tggaatcggc aaagacgaag 900

ctggcgagca aggcgcgaga gatgaaggat agggcagagc agttctcgca gcagcctgtt 960

gggggcggtg gatccggtgg agggggatcc atgggatatc ggatccgaat tcgagctccg 1020

tcgacaagct tgcggccgct tgctccgtta aaaaaaacca tggctgagca ttatggtcaa 1080

caacagcaga ccagggcgcc tcacctgcag ctgcagccgc gcgcccagcg ggtagtgaag 1140

gcggccaccg ccgtgacagc cggcggctcg cttctcgtcc tctctggcct cactttagcc 1200

ggaactgtta ttgcgctcac catcgccact ccgctgcttg tgatctttag ccccgttctg 1260

gtgccggcgg tcataaccat tttcttgctg ggtgcgggtt ttctggcatc cggaggcttc 1320

ggcgtggcgg cgctgagtgt gctgtcgtgg atttacagat atctgacagg gaaacacccg 1380

ccgggggcgg atcagctgga atcggcaaag acgaagctgg cgagcaaggc gcgagagatg 1440

aaggataggg cagagcagtt ctcgcagcag cctgttggag gcggtggatc cggaggcggt 1500

ggtagtatgg ctgagcatta tggtcaacaa cagcagacca gggcgcctca cctgcagctg 1560

cagccgcgcg cccagcgggt agtgaaggcg gccaccgccg tgacagccgg cggctcgctt 1620

ctcgtcctct ctggcctcac tttagccgga actgttattg cgctcaccat cgccactccg 1680

ctgcttgtga tctttagccc cgttctggtg ccggcggtca taaccatttt cttgctgggt 1740

gcgggttttc tggcatccgg aggcttcggc gtggcggcgc tgagtgtgct gtcgtggatt 1800

tacagatatc tgacagggaa acacccgccg ggggcggatc agctggaatc ggcaaagacg 1860

aagctggcga gcaaggcgcg agagatgaag gatagggcag agcagttctc gcagcagcct 1920

gttgggggcg gtggatccgg tggaggggga tccatggctg agcattatgg tcaacaacag 1980

cagaccaggg cgcctcacct gcagctgcag ccgcgcgccc agcgggtagt gaaggcggcc 2040

accgccgtga cagccggcgg ctcgcttctc gtcctctctg gcctcacttt agccggaact 2100

gttattgcgc tcaccatcgc cactccgctg cttgtgatct ttagccccgt tctggtgccg 2160

gcggtcataa ccattttctt gctgggtgcg ggttttctgg catccggagg cttcggcgtg 2220

gcggcgctga gtgtgctgtc gtggatttac agatatctga cagggaaaca cccgccgggg 2280

gcggatcagc tggaatcggc aaagacgaag ctggcgagca aggcgcgaga gatgaaggat 2340

agggcagagc agttctcgca gcagcctgtt ccatggctcg agcaccacca ccaccaccac 2400

tga 2403

<210> 58
<211> 953
<212> PRT
<213> Sesamum indicum/Synthetic

<400> 58

Met Lys Glu Thr Ala Ala Ala Lys Phe Glu Arg Gln His Met Asp Ser
1 5 10 15

Pro Asp Leu Gly Thr Leu Val Pro Arg Gly Ser Met Ala Glu His Tyr
20 25 30

Gly Gln Gln Gln Gln Thr Arg Ala Pro His Leu Gln Leu Gln Pro Arg
35 40 45

Ala Gln Arg Val Val Lys Ala Ala Thr Ala Val Thr Ala Gly Gly Ser

50 55 60

Leu Leu Val Leu Ser Gly Leu Thr Leu Ala Gly Thr Val Ile Ala Leu
65 70 75 80

Thr Ile Ala Thr Pro Leu Leu Val Ile Phe Ser Pro Val Leu Val Pro
85 90 95

Ala Val Ile Thr Ile Phe Leu Leu Gly Ala Gly Phe Leu Ala Ser Gly
100 105 110

Gly Phe Gly Val Ala Ala Leu Ser Val Leu Ser Trp Ile Tyr Arg Tyr
115 120 125

Leu Thr Gly Lys His Pro Pro Gly Ala Asp Gln Leu Glu Ser Ala Lys
130 135 140

Thr Lys Leu Ala Ser Lys Ala Arg Glu Met Lys Asp Arg Ala Glu Gln
145 150 155 160

Phe Ser Gln Gln Pro Val Pro Trp Leu Ile Ser Asp Pro Asn Ser Ser
165 170 175

Ser Val Asp Lys Leu Ala Ala Ala Leu Glu Met Ala Glu His Tyr Gly
180 185 190

Gln Gln Gln Gln Thr Arg Ala Pro His Leu Gln Leu Gln Pro Arg Ala
195 200 205

Gln Arg Val Val Lys Ala Ala Thr Ala Val Thr Ala Gly Gly Ser Leu
210 215 220

Leu Val Leu Ser Gly Leu Thr Leu Ala Gly Thr Val Ile Ala Leu Thr
225 230 235 240

Ile Ala Thr Pro Leu Leu Val Ile Phe Ser Pro Val Leu Val Pro Ala
245 250 255

Val Ile Thr Ile Phe Leu Leu Gly Ala Gly Phe Leu Ala Ser Gly Gly
260 265 270

Phe Gly Val Ala Ala Leu Ser Val Leu Ser Trp Ile Tyr Arg Tyr Leu
275 280 285

Thr Gly Lys His Pro Pro Gly Ala Asp Gln Leu Glu Ser Ala Lys Thr
290 295 300

Lys Leu Ala Ser Lys Ala Arg Glu Met Lys Asp Arg Ala Glu Gln Phe
305 310 315 320

Ser Gln Gln Pro Val Ala Gly Ser Gln Thr Ser His Met Phe Lys Trp

325 330 335

Pro Ser Ala Met Ala Glu His Tyr Gly Gln Gln Gln Gln Thr Arg Ala
340 345 350

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

Thr Ala Val Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Gly Leu Thr
370 375 380

Leu Ala Gly Thr Val Ile Ala Leu Thr Ile Ala Thr Pro Leu Leu Val
385 390 395 400

Ile Phe Ser Pro Val Leu Val Pro Ala Val Ile Thr Ile Phe Leu Leu
405 410 415

Gly Ala Gly Phe Leu Ala Ser Gly Gly Phe Gly Val Ala Ala Leu Ser
420 425 430

Val Leu Ser Trp Ile Tyr Arg Tyr Leu Thr Gly Lys His Pro Pro Gly
435 440 445

Ala Asp Gln Leu Glu Ser Ala Lys Thr Lys Leu Ala Ser Lys Ala Arg
450 455 460

Glu Met Lys Asp Arg Ala Glu Gln Phe Ser Gln Gln Pro Val Ala Gly
465 470 475 480

Ser Gln Thr Ser Ser Ser Glu Leu Pro Trp Val Asp Met Ala Glu His
485 490 495

Tyr Gly Gln Gln Gln Gln Thr Arg Ala Pro His Leu Gln Leu Gln Pro
500 505 510

Arg Ala Gln Arg Val Val Lys Ala Ala Thr Ala Val Thr Ala Gly Gly
515 520 525

Ser Leu Leu Val Leu Ser Gly Leu Thr Leu Ala Gly Thr Val Ile Ala
530 535 540

Leu Thr Ile Ala Thr Pro Leu Leu Val Ile Phe Ser Pro Val Leu Val
545 550 555 560

Pro Ala Val Ile Thr Ile Phe Leu Leu Gly Ala Gly Phe Leu Ala Ser
565 570 575

Gly Gly Phe Gly Val Ala Ala Leu Ser Val Leu Ser Trp Ile Tyr Arg
580 585 590

Tyr Leu Thr Gly Lys His Pro Pro Gly Ala Asp Gln Leu Glu Ser Ala

595 600 605

Lys Thr Lys Leu Ala Ser Lys Ala Arg Glu Met Lys Asp Arg Ala Glu
610 615 620

Gln Phe Ser Gln Gln Pro Val Ala Gly Ser Gln Thr Ser Ser Thr Pro
625 630 635 640

Ser Ser Trp His Met Met Ala Glu His Tyr Gly Gln Gln Gln Gln Thr
645 650 655

Arg Ala Pro His Leu Gln Leu Gln Pro Arg Ala Gln Arg Val Val Lys
660 665 670

Ala Ala Thr Ala Val Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Gly
675 680 685

Leu Thr Leu Ala Gly Thr Val Ile Ala Leu Thr Ile Ala Thr Pro Leu
690 695 700

Leu Val Ile Phe Ser Pro Val Leu Val Pro Ala Val Ile Thr Ile Phe
705 710 715 720

Leu Leu Gly Ala Gly Phe Leu Ala Ser Gly Gly Phe Gly Val Ala Ala
725 730 735

Leu Ser Val Leu Ser Trp Ile Tyr Arg Tyr Leu Thr Gly Lys His Pro
740 745 750

Pro Gly Ala Asp Gln Leu Glu Ser Ala Lys Thr Lys Leu Ala Ser Lys
755 760 765

Ala Arg Glu Met Lys Asp Arg Ala Glu Gln Phe Ser Gln Gln Pro Val
770 775 780

Ala Gly Ser Gln Thr Ser Ile Asp Gln Gln Val Asn Val His Met Ala
785 790 795 800

Glu His Tyr Gly Gln Gln Gln Gln Thr Arg Ala Pro His Leu Gln Leu
805 810 815

Gln Pro Arg Ala Gln Arg Val Val Lys Ala Ala Thr Ala Val Thr Ala
820 825 830

Gly Gly Ser Leu Leu Val Leu Ser Gly Leu Thr Leu Ala Gly Thr Val
835 840 845

Ile Ala Leu Thr Ile Ala Thr Pro Leu Leu Val Ile Phe Ser Pro Val
850 855 860

Leu Val Pro Ala Val Ile Thr Ile Phe Leu Leu Gly Ala Gly Phe Leu

865 870 875 880

Ala Ser Gly Gly Phe Gly Val Ala Ala Leu Ser Val Leu Ser Trp Ile
885 890 895

Tyr Arg Tyr Leu Thr Gly Lys His Pro Pro Gly Ala Asp Gln Leu Glu
900 905 910

Ser Ala Lys Thr Lys Leu Ala Ser Lys Ala Arg Glu Met Lys Asp Arg
915 920 925

Ala Glu Gln Phe Ser Gln Gln Pro Val Ala Gly Ser Gln Thr Ser Pro
930 935 940

Trp Leu Glu His His His His His His
945 950

<210> 59
<211> 350
<212> PRT
<213> Sesamum indicum/Synthetic

<400> 59

Met Lys Glu Thr Ala Ala Ala Lys Phe Glu Arg Gln His Met Asp Ser
1 5 10 15

Pro Asp Leu Gly Thr Leu Val Pro Arg Gly Ser Met Ala Glu His Tyr
20 25 30

Gly Gln Gln Gln Gln Thr Arg Ala Pro His Leu Gln Leu Gln Pro Arg
35 40 45

Ala Gln Arg Val Val Lys Ala Ala Thr Ala Val Thr Ala Gly Gly Ser
50 55 60

Leu Leu Val Leu Ser Gly Leu Thr Leu Ala Gly Thr Val Ile Ala Leu
65 70 75 80

Thr Ile Ala Thr Pro Leu Leu Val Ile Phe Ser Pro Val Leu Val Pro
85 90 95

Ala Val Ile Thr Ile Phe Leu Leu Gly Ala Gly Phe Leu Ala Ser Gly
100 105 110

Gly Phe Gly Val Ala Ala Leu Ser Val Leu Ser Trp Ile Tyr Arg Tyr
115 120 125

Leu Thr Gly Lys His Pro Pro Gly Ala Asp Gln Leu Glu Ser Ala Lys
130 135 140

Thr Lys Leu Ala Ser Lys Ala Arg Glu Met Lys Asp Arg Ala Glu Gln
145 150 155 160

Phe Ser Gln Gln Pro Val Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
165 170 175

Met Ala Glu His Tyr Gly Gln Gln Gln Gln Thr Arg Ala Pro His Leu
180 185 190

Gln Leu Gln Pro Arg Ala Gln Arg Val Val Lys Ala Ala Thr Ala Val
195 200 205

Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Gly Leu Thr Leu Ala Gly
210 215 220

Thr Val Ile Ala Leu Thr Ile Ala Thr Pro Leu Leu Val Ile Phe Ser
225 230 235 240

Pro Val Leu Val Pro Ala Val Ile Thr Ile Phe Leu Leu Gly Ala Gly
245 250 255

Phe Leu Ala Ser Gly Gly Phe Gly Val Ala Ala Leu Ser Val Leu Ser
260 265 270

Trp Ile Tyr Arg Tyr Leu Thr Gly Lys His Pro Pro Gly Ala Asp Gln
275 280 285

Leu Glu Ser Ala Lys Thr Lys Leu Ala Ser Lys Ala Arg Glu Met Lys
290 295 300

Asp Arg Ala Glu Gln Phe Ser Gln Gln Pro Val Gly Gly Gly Gly Ser
305 310 315 320

Gly Gly Gly Gly Ser Met Ala Ile Ser Asp Pro Asn Ser Ser Ser Val
325 330 335

Asp Lys Leu Ala Ala Ala Leu Glu His His His His His His
340 345 350

<210> 60
<211> 637
<212> PRT
<213> Sesamum indicum/Synthetic

<400> 60

Met Lys Glu Thr Ala Ala Ala Lys Phe Glu Arg Gln His Met Asp Ser
1 5 10 15

Pro Asp Leu Gly Thr Leu Val Pro Arg Gly Ser Met Ala Glu His Tyr
20 25 30

Gly Gln Gln Gln Gln Thr Arg Ala Pro His Leu Gln Leu Gln Pro Arg
35 40 45

```
Ala Gln Arg Val Val Lys Ala Ala Thr Ala Val Thr Ala Gly Gly Ser
    50                  55                  60

Leu Leu Val Leu Ser Gly Leu Thr Leu Ala Gly Thr Val Ile Ala Leu
65                  70                  75                  80

Thr Ile Ala Thr Pro Leu Leu Val Ile Phe Ser Pro Val Leu Val Pro
                85                  90                  95

Ala Val Ile Thr Ile Phe Leu Leu Gly Ala Gly Phe Leu Ala Ser Gly
            100                 105                 110

Gly Phe Gly Val Ala Ala Leu Ser Val Leu Ser Trp Ile Tyr Arg Tyr
        115                 120                 125

Leu Thr Gly Lys His Pro Pro Gly Ala Asp Gln Leu Glu Ser Ala Lys
    130                 135                 140

Thr Lys Leu Ala Ser Lys Ala Arg Glu Met Lys Asp Arg Ala Glu Gln
145                 150                 155                 160

Phe Ser Gln Gln Pro Val Pro Trp Arg Tyr Arg Ile Arg Ile Arg Ala
                165                 170                 175

Pro Ser Thr Ser Leu Arg Pro Leu Ala Pro Leu Lys Lys Thr Met Ala
            180                 185                 190

Glu His Tyr Gly Gln Gln Gln Gln Thr Arg Ala Pro His Leu Gln Leu
        195                 200                 205

Gln Pro Arg Ala Gln Arg Val Val Lys Ala Ala Thr Ala Val Thr Ala
    210                 215                 220

Gly Gly Ser Leu Leu Val Leu Ser Gly Leu Thr Leu Ala Gly Thr Val
225                 230                 235                 240

Ile Ala Leu Thr Ile Ala Thr Pro Leu Leu Val Ile Phe Ser Pro Val
                245                 250                 255

Leu Val Pro Ala Val Ile Thr Ile Phe Leu Leu Gly Ala Gly Phe Leu
            260                 265                 270

Ala Ser Gly Gly Phe Gly Val Ala Ala Leu Ser Val Leu Ser Trp Ile
        275                 280                 285

Tyr Arg Tyr Leu Thr Gly Lys His Pro Pro Gly Ala Asp Gln Leu Glu
    290                 295                 300

Ser Ala Lys Thr Lys Leu Ala Ser Lys Ala Arg Glu Met Lys Asp Arg
305                 310                 315                 320
```

```
Ala Glu Gln Phe Ser Gln Gln Pro Val Gly Gly Gly Gly Ser Gly Gly
                325                 330                 335

Gly Gly Ser Met Ala Glu His Tyr Gly Gln Gln Gln Gln Thr Arg Ala
            340                 345                 350

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

Thr Ala Val Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Gly Leu Thr
    370                 375                 380

Leu Ala Gly Thr Val Ile Ala Leu Thr Ile Ala Thr Pro Leu Leu Val
385                 390                 395                 400

Ile Phe Ser Pro Val Leu Val Pro Ala Val Ile Thr Ile Phe Leu Leu
                405                 410                 415

Gly Ala Gly Phe Leu Ala Ser Gly Gly Phe Gly Val Ala Ala Leu Ser
            420                 425                 430

Val Leu Ser Trp Ile Tyr Arg Tyr Leu Thr Gly Lys His Pro Pro Gly
        435                 440                 445

Ala Asp Gln Leu Glu Ser Ala Lys Thr Lys Leu Ala Ser Lys Ala Arg
    450                 455                 460

Glu Met Lys Asp Arg Ala Glu Gln Phe Ser Gln Gln Pro Val Gly Gly
465                 470                 475                 480

Gly Gly Ser Gly Gly Gly Gly Ser Met Ala Glu His Tyr Gly Gln Gln
                485                 490                 495

Gln Gln Thr Arg Ala Pro His Leu Gln Leu Gln Pro Arg Ala Gln Arg
            500                 505                 510

Val Val Lys Ala Ala Thr Ala Val Thr Ala Gly Gly Ser Leu Leu Val
        515                 520                 525

Leu Ser Gly Leu Thr Leu Ala Gly Thr Val Ile Ala Leu Thr Ile Ala
    530                 535                 540

Thr Pro Leu Leu Val Ile Phe Ser Pro Val Leu Val Pro Ala Val Ile
545                 550                 555                 560

Thr Ile Phe Leu Leu Gly Ala Gly Phe Leu Ala Ser Gly Gly Phe Gly
                565                 570                 575

Val Ala Ala Leu Ser Val Leu Ser Trp Ile Tyr Arg Tyr Leu Thr Gly
            580                 585                 590
```

Lys His Pro Pro Gly Ala Asp Gln Leu Glu Ser Ala Lys Thr Lys Leu
595 600 605

Ala Ser Lys Ala Arg Glu Met Lys Asp Arg Ala Glu Gln Phe Ser Gln
610 615 620

Gln Pro Val Pro Trp Leu Glu His His His His His His
625 630 635

<210> 61
<211> 795
<212> PRT
<213> Sesamum indicum/Synthetic

<400> 61

Met Lys Glu Thr Ala Ala Ala Lys Phe Glu Arg Gln His Met Asp Ser
1 5 10 15

Pro Asp Leu Gly Thr Leu Val Pro Arg Gly Ser Met Ala Glu His Tyr
20 25 30

Gly Gln Gln Gln Gln Thr Arg Ala Pro His Leu Gln Leu Gln Pro Arg
35 40 45

Ala Gln Arg Val Val Lys Ala Ala Thr Ala Val Thr Ala Gly Gly Ser
50 55 60

Leu Leu Val Leu Ser Gly Leu Thr Leu Ala Gly Thr Val Ile Ala Leu
65 70 75 80

Thr Ile Ala Thr Pro Leu Leu Val Ile Phe Ser Pro Val Leu Val Pro
85 90 95

Ala Val Ile Thr Ile Phe Leu Leu Gly Ala Gly Phe Leu Ala Ser Gly
100 105 110

Gly Phe Gly Val Ala Ala Leu Ser Val Leu Ser Trp Ile Tyr Arg Tyr
115 120 125

Leu Thr Gly Lys His Pro Pro Gly Ala Asp Gln Leu Glu Ser Ala Lys
130 135 140

Thr Lys Leu Ala Ser Lys Ala Arg Glu Met Lys Asp Arg Ala Glu Gln
145 150 155 160

Phe Ser Gln Gln Pro Val Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
165 170 175

Met Ala Glu His Tyr Gly Gln Gln Gln Gln Thr Arg Ala Pro His Leu
180 185 190

```
Gln Leu Gln Pro Arg Ala Gln Arg Val Val Lys Ala Ala Thr Ala Val
        195                 200                 205

Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Gly Leu Thr Leu Ala Gly
    210                 215                 220

Thr Val Ile Ala Leu Thr Ile Ala Thr Pro Leu Leu Val Ile Phe Ser
225                 230                 235                 240

Pro Val Leu Val Pro Ala Val Ile Thr Ile Phe Leu Leu Gly Ala Gly
                245                 250                 255

Phe Leu Ala Ser Gly Gly Phe Gly Val Ala Ala Leu Ser Val Leu Ser
            260                 265                 270

Trp Ile Tyr Arg Tyr Leu Thr Gly Lys His Pro Pro Gly Ala Asp Gln
        275                 280                 285

Leu Glu Ser Ala Lys Thr Lys Leu Ala Ser Lys Ala Arg Glu Met Lys
    290                 295                 300

Asp Arg Ala Glu Gln Phe Ser Gln Gln Pro Val Gly Gly Gly Gly Ser
305                 310                 315                 320

Gly Gly Gly Gly Ser Met Gly Tyr Arg Ile Arg Ile Arg Ala Pro Ser
                325                 330                 335

Thr Ser Leu Arg Pro Leu Ala Pro Leu Lys Lys Thr Met Ala Glu His
            340                 345                 350

Tyr Gly Gln Gln Gln Gln Thr Arg Ala Pro His Leu Gln Leu Gln Pro
        355                 360                 365

Arg Ala Gln Arg Val Val Lys Ala Ala Thr Ala Val Thr Ala Gly Gly
    370                 375                 380

Ser Leu Leu Val Leu Ser Gly Leu Thr Leu Ala Gly Thr Val Ile Ala
385                 390                 395                 400

Leu Thr Ile Ala Thr Pro Leu Leu Val Ile Phe Ser Pro Val Leu Val
                405                 410                 415

Pro Ala Val Ile Thr Ile Phe Leu Leu Gly Ala Gly Phe Leu Ala Ser
            420                 425                 430

Gly Gly Phe Gly Val Ala Ala Leu Ser Val Leu Ser Trp Ile Tyr Arg
        435                 440                 445

Tyr Leu Thr Gly Lys His Pro Pro Gly Ala Asp Gln Leu Glu Ser Ala
    450                 455                 460
```

```
Lys Thr Lys Leu Ala Ser Lys Ala Arg Glu Met Lys Asp Arg Ala Glu
465                 470                 475                 480

Gln Phe Ser Gln Gln Pro Val Gly Gly Gly Gly Ser Gly Gly Gly Gly
                485                 490                 495

Ser Met Ala Glu His Tyr Gly Gln Gln Gln Gln Thr Arg Ala Pro His
            500                 505                 510

Leu Gln Leu Gln Pro Arg Ala Gln Arg Val Val Lys Ala Ala Thr Ala
        515                 520                 525

Val Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Gly Leu Thr Leu Ala
    530                 535                 540

Gly Thr Val Ile Ala Leu Thr Ile Ala Thr Pro Leu Leu Val Ile Phe
545                 550                 555                 560

Ser Pro Val Leu Val Pro Ala Val Ile Thr Ile Phe Leu Leu Gly Ala
                565                 570                 575

Gly Phe Leu Ala Ser Gly Gly Phe Gly Val Ala Ala Leu Ser Val Leu
            580                 585                 590

Ser Trp Ile Tyr Arg Tyr Leu Thr Gly Lys His Pro Pro Gly Ala Asp
        595                 600                 605

Gln Leu Glu Ser Ala Lys Thr Lys Leu Ala Ser Lys Ala Arg Glu Met
    610                 615                 620

Lys Asp Arg Ala Glu Gln Phe Ser Gln Gln Pro Val Gly Gly Gly Gly
625                 630                 635                 640

Ser Gly Gly Gly Gly Ser Met Ala Glu His Tyr Gly Gln Gln Gln Gln
                645                 650                 655

Thr Arg Ala Pro His Leu Gln Leu Gln Pro Arg Ala Gln Arg Val Val
            660                 665                 670

Lys Ala Ala Thr Ala Val Thr Ala Gly Gly Ser Leu Leu Val Leu Ser
        675                 680                 685

Gly Leu Thr Leu Ala Gly Thr Val Ile Ala Leu Thr Ile Ala Thr Pro
    690                 695                 700

Leu Leu Val Ile Phe Ser Pro Val Leu Val Pro Ala Val Ile Thr Ile
705                 710                 715                 720

Phe Leu Leu Gly Ala Gly Phe Leu Ala Ser Gly Gly Phe Gly Val Ala
                725                 730                 735
```

```
Ala Leu Ser Val Leu Ser Trp Ile Tyr Arg Tyr Leu Thr Gly Lys His
            740                 745                 750

Pro Pro Gly Ala Asp Gln Leu Glu Ser Ala Lys Thr Lys Leu Ala Ser
        755                 760                 765

Lys Ala Arg Glu Met Lys Asp Arg Ala Glu Gln Phe Ser Gln Gln Pro
    770                 775                 780

Val Pro Trp Leu Glu His His His His His His
785                 790                 795


<210>  62
<211>  441
<212>  PRT
<213>  Sesamum indicum/Synthetic

<400>  62

Met Ala Glu His Tyr Gly Gln Gln Gln Gln Thr Arg Ala Pro His Leu
1               5                   10                  15

Gln Leu Gln Pro Arg Ala Gln Arg Val Val Lys Ala Ala Thr Ala Val
            20                  25                  30

Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Gly Leu Thr Leu Ala Gly
        35                  40                  45

Thr Val Ile Ala Leu Thr Ile Ala Thr Pro Leu Leu Val Ile Phe Ser
    50                  55                  60

Pro Val Leu Val Pro Ala Val Ile Thr Ile Phe Leu Leu Gly Ala Gly
65                  70                  75                  80

Phe Leu Ala Ser Gly Gly Phe Gly Val Ala Ala Leu Ser Val Leu Ser
                85                  90                  95

Trp Ile Tyr Arg Tyr Leu Thr Gly Lys His Pro Pro Gly Ala Asp Gln
            100                 105                 110

Leu Glu Ser Ala Lys Thr Lys Leu Ala Ser Lys Ala Arg Glu Met Lys
        115                 120                 125

Asp Arg Ala Glu Gln Phe Ser Gln Gln Pro Val Gly Gly Gly Gly Ser
    130                 135                 140

Gly Gly Gly Gly Ser Met Ala Glu His Tyr Gly Gln Gln Gln Gln Thr
145                 150                 155                 160

Arg Ala Pro His Leu Gln Leu Gln Pro Arg Ala Gln Arg Val Val Lys
                165                 170                 175

Ala Ala Thr Ala Val Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Gly
```

180 185 190

Leu Thr Leu Ala Gly Thr Val Ile Ala Leu Thr Ile Ala Thr Pro Leu
195 200 205

Leu Val Ile Phe Ser Pro Val Leu Val Pro Ala Val Ile Thr Ile Phe
210 215 220

Leu Leu Gly Ala Gly Phe Leu Ala Ser Gly Gly Phe Gly Val Ala Ala
225 230 235 240

Leu Ser Val Leu Ser Trp Ile Tyr Arg Tyr Leu Thr Gly Lys His Pro
245 250 255

Pro Gly Ala Asp Gln Leu Glu Ser Ala Lys Thr Lys Leu Ala Ser Lys
260 265 270

Ala Arg Glu Met Lys Asp Arg Ala Glu Gln Phe Ser Gln Gln Pro Val
275 280 285

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Met Ala Glu His Tyr Gly
290 295 300

Gln Gln Gln Gln Thr Arg Ala Pro His Leu Gln Leu Gln Pro Arg Ala
305 310 315 320

Gln Arg Val Val Lys Ala Ala Thr Ala Val Thr Ala Gly Gly Ser Leu
325 330 335

Leu Val Leu Ser Gly Leu Thr Leu Ala Gly Thr Val Ile Ala Leu Thr
340 345 350

Ile Ala Thr Pro Leu Leu Val Ile Phe Ser Pro Val Leu Val Pro Ala
355 360 365

Val Ile Thr Ile Phe Leu Leu Gly Ala Gly Phe Leu Ala Ser Gly Gly
370 375 380

Phe Gly Val Ala Ala Leu Ser Val Leu Ser Trp Ile Tyr Arg Tyr Leu
385 390 395 400

Thr Gly Lys His Pro Pro Gly Ala Asp Gln Leu Glu Ser Ala Lys Thr
405 410 415

Lys Leu Ala Ser Lys Ala Arg Glu Met Lys Asp Arg Ala Glu Gln Phe
420 425 430

Ser Gln Gln Pro Val Pro Trp Leu Glu
435 440

<210> 63

<211> 149
<212> PRT
<213> Sesamum indicum/Synthetic

<400> 63

Met Ala Glu His Tyr Gly Gln Gln Gln Thr Arg Ala Pro His Leu
1 5 10 15

Gln Leu Gln Pro Arg Ala Gln Arg Val Val Lys Ala Ala Thr Ala Val
20 25 30

Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Gly Leu Thr Leu Ala Gly
35 40 45

Thr Val Ile Ala Leu Thr Ile Ala Thr Pro Leu Leu Val Ile Phe Ser
50 55 60

Pro Val Leu Val Pro Ala Val Ile Thr Ile Phe Leu Leu Gly Ala Gly
65 70 75 80

Phe Leu Ala Ser Gly Gly Phe Gly Val Ala Ala Leu Ser Val Leu Ser
85 90 95

Trp Ile Tyr Arg Tyr Leu Thr Gly Lys His Pro Pro Gly Ala Asp Gln
100 105 110

Leu Glu Ser Ala Lys Thr Lys Leu Ala Ser Lys Ala Arg Glu Met Lys
115 120 125

Asp Arg Ala Glu Gln Phe Ser Gln Gln Pro Val Pro Trp Leu Glu His
130 135 140

His His His His His
145

<210> 64
<211> 2227
<212> DNA
<213> Sesamum indicum/Synthetic

<400> 64
tcgacgaatt aattccaatc ccacaaaaat ctgagcttaa cagcacagtt gctcctctca 60

gagcagaatc gggtattcaa caccctcata tcaactacta cgttgtgtat aacggtccac 120

atgccggtat atacgatgac tggggttgta caaaggcggc aacaaacggc gttcccggag 180

ttgcacacaa gaaatttgcc actattacag aggcaagagc agcagctgac gcgtacacaa 240

caagtcagca aacagacagg ttgaacttca tccccaaagg agaagctcaa ctcaagccca 300

agagctttgc taaggcccta acaagcccac caaagcaaaa agcccactgg ctcacgctag 360

gaaccaaaag gcccagcagt gatccagccc caaaagagat ctcctttgcc ccggagatta 420

caatggacga tttcctctat ctttacgatc taggaaggaa gttcgaaggt gaaggtgacg 480

acactatgtt caccactgat aatgagaagg ttagcctctt caatttcaga aagaatgctg 540

acccacagat ggttagagag gcctacgcag caggtctcat caagacgatc tacccgagta 600

acaatctcca ggagatcaaa taccttccca agaaggttaa agatgcagtc aaaagattca 660

ggactaattg catcaagaac acagagaaag acatatttct caagatcaga agtactattc 720

cagtatggac gattcaaggc ttgcttcata aaccaaggca agtaatagag attggagtct 780

ctaaaaaggt agttcctact gaatctaagg ccatgcatgg agtctaagat tcaaatcgag 840

gatctaacag aactcgccgt gaagactggc gaacagttca tacagagtct tttacgactc 900

aatgacaaga agaaaatctt cgtcaacatg gtggagcacg acactctggt ctactccaaa 960

aatgtcaaag atacagtctc agaagaccaa agggctattg agacttttca acaaaggata 1020

atttcgggaa acctcctcgg attccattgc ccagctatct gtcacttcat cgaaaggaca 1080

gtagaaaagg aaggtggctc ctacaaatgc catcattgcg ataaaggaaa ggctatcatt 1140

caagatctct ctgccgacag tggtcccaaa gatggacccc cacccacgag gagcatcgtg 1200

gaaaaagaag acgttccaac cacgtcttca aagcaagtgg attgatgtga catctccact 1260

gacgtaaggg atgacgcaca atcccactat ccttcgcaag acccttcctc tatataagga 1320

agttcatttc atttggagag gacacgctcg aggaattcgg taccccatca caagtttgta 1380

caaaaaagca ggctccgcgg ccgcttgctc cgttaaaaaa aaccatggca gagcactacg 1440

ggcagcagca acaaacacgt gccccccact tgcaactgca accgcgtgct cagcgtgttg 1500

taaaggcagc gaccgcggtt actgccggag gtagcctgtt ggtgttatcc gggttgaccc 1560

tggctggaac ggtcattgcg ctgacaatcg ccacacctct cctggtgatc ttctcgcctg 1620

tactggtccc tgcaggtaaa tttctgtgtt ccttattctc tcaaaatctt cgattttgtt 1680

ttcgttcgat cccaatttcg tatatgttct ttggtttaga ttctgttaat cttagatcga 1740

agacgatttt ctgggtttga tcgttagata tcatcttaat tctcgattag ggtttcatag 1800

atatcatccg atttgttcaa ataatttgag ttttgtcgaa taattactct tcgatttgtg 1860

atttctatct agatctggtg ttagtttcta gtttgtgcga tcgaatttgt cgattaatct 1920

gagtttttct gatctgcagt catcaccatc ttcctgttgg gagcgggttt tctggcaagc 1980

gggggatttg gtgttgccgc tctgtctgtg ctgtcctgga tctaccgtta cctgacgggg 2040

aaacatccac ccggagcgga tcagttggag tcggccaaga caaagttggc gagcaaggcc 2100

cgtgaaatga aggaccgtgc cgagcagttc agtcagcaac cggtagcagg gtctcagacc 2160

agcatcgatc catcctcctg gctcgagtaa tgaagcggcc gcacccagct ttcttgtaca 2220

aagtggt 2227

<210> 65
<211> 3133
<212> DNA
<213> Sesamum indicum/Synthetic

<400> 65
tcgacgaatt aattccaatc ccacaaaaat ctgagcttaa cagcacagtt gctcctctca 60

gagcagaatc gggtattcaa caccctcata tcaactacta cgttgtgtat aacggtccac 120
atgccggtat atacgatgac tggggttgta caaaggcggc aacaaacggc gttcccggag 180
ttgcacacaa gaaatttgcc actattacag aggcaagagc agcagctgac gcgtacacaa 240
caagtcagca aacagacagg ttgaacttca tccccaaagg agaagctcaa ctcaagccca 300
agagctttgc taaggcccta acaagcccac caaagcaaaa agcccactgg ctcacgctag 360
gaaccaaaag gcccagcagt gatccagccc caaaagagat ctcctttgcc ccggagatta 420
caatggacga tttcctctat ctttacgatc taggaaggaa gttcgaaggt gaaggtgacg 480
acactatgtt caccactgat aatgagaagg ttagcctctt caatttcaga aagaatgctg 540
acccacagat ggttagagag gcctacgcag caggtctcat caagacgatc tacccgagta 600
acaatctcca ggagatcaaa taccttccca agaaggttaa agatgcagtc aaaagattca 660
ggactaattg catcaagaac acagagaaag acatatttct caagatcaga agtactattc 720
cagtatggac gattcaaggc ttgcttcata aaccaaggca agtaatagag attggagtct 780
ctaaaaaggt agttcctact gaatctaagg ccatgcatgg agtctaagat tcaaatcgag 840
gatctaacag aactcgccgt gaagactggc gaacagttca tacagagtct tttacgactc 900
aatgacaaga agaaaatctt cgtcaacatg gtggagcacg acactctggt ctactccaaa 960
aatgtcaaag atacagtctc agaagaccaa agggctattg agacttttca acaaaggata 1020
atttcgggaa acctcctcgg attccattgc ccagctatct gtcacttcat cgaaaggaca 1080
gtagaaaagg aaggtggctc ctacaaatgc catcattgcg ataaaggaaa ggctatcatt 1140
caagatctct ctgccgacag tggtcccaaa gatggacccc cacccacgag gagcatcgtg 1200
gaaaaagaag acgttccaac cacgtcttca aagcaagtgg attgatgtga catctccact 1260
gacgtaaggg atgacgcaca atcccactat ccttcgcaag acccttcctc tatataagga 1320
agttcatttc atttggagag gacacgctcg aggaattcgg taccccatca caagtttgta 1380
caaaaaagca ggctccgcgg ccgcttgctc cgttaaaaaa aaccatggca gagcactacg 1440
ggcagcagca acaaacacgt gccccccact tgcaactgca accgcgtgct cagcgtgttg 1500
taaaggcagc gaccgcggtt actgccggag gtagcctgtt ggtgttatcc gggttgaccc 1560
tggctggaac ggtcattgcg ctgacaatcg ccacacctct cctggtgatc ttctcgcctg 1620
tactggtccc tgcaggtaaa tttctgtgtt ccttattctc tcaaaatctt cgattttgtt 1680
ttcgttcgat cccaatttcg tatatgttct ttggtttaga ttctgttaat cttagatcga 1740
agacgatttt ctgggtttga tcgttagata tcatcttaat tctcgattag ggtttcatag 1800
atatcatccg atttgttcaa ataatttgag ttttgtcgaa taattactct tcgatttgtg 1860
atttctatct agatctggtg ttagtttcta gtttgtgcga tcgaatttgt cgattaatct 1920
gagtttttct gatctgcagt catcaccatc ttcctgttgg gagcgggttt tctggcaagc 1980
gggggatttg gtgttgccgc tctgtctgtg ctgtcctgga tctaccgtta cctgacgggg 2040
aaacatccac ccggagcgga tcagttggag tcggccaaga caaagttggc gagcaaggcc 2100

cgtgaaatga aggaccgtgc cgagcagttc agtcagcaac cggtagcagg gtctcagacc 2160
agcatcgatc catcctcctg gcatatgatg gccgagcact atggacaaca gcagcagacg 2220
cgtgcccctc atctgcaact gcaaccccgt gctcaacgtg tcgttaaggc tgcgacagca 2280
gtaaccgctg ggggttctct gttagtgttg tcagggctga ctttggcggg gacggtaatt 2340
gcgttgacca ttgccacccc gctgttagtg attttcagcc cggtactggt gccagcagtt 2400
atcacgatct tcttgctggg tgccggattc ttggcaagtg gaggttttgg agttgcggcg 2460
ctgtcagttt tatcctggat ctatcgttat ctgacaggaa aacatccccc aggtgccgat 2520
cagctggaga gtgccaagac aaaactggcg tctaaggcac gtgaaatgaa ggatcgtgcc 2580
gaacagtttt ctcaacagcc cgtagcgggg tcacagacct cgatcgatca gcaggttaac 2640
gtgcacatgg ccgaacatta cggacagcaa caacagacgc gtgctccaca cctgcaattg 2700
caaccgcgtg ctcaacgtgt tgtcaaagcg gcgaccgccg taacagcagg aggatcactg 2760
ttagtgctgt cgggtttaac cttggccggg accgtcattg cattgactat tgcgacgccc 2820
ttactggtga tcttttctcc ggtgctggtt cccgccgtta ttaccatctt cttgttaggg 2880
gcaggattcc tggcatcagg gggattcgga gttgcggcgt tgagtgtctt aagttggatc 2940
taccgttatc tgactggaaa gcacccgcct ggggccgatc aactggagtc agccaaaacg 3000
aaattggcgt caaaagcgcg tgaaatgaag gaccgtgctg agcagttttc tcagcagcct 3060
gtggcaggat cccagacatc accatggctc gagtaatgaa gcggccgcac ccagctttct 3120
tgtacaaagt ggt 3133

<210> 66
<211> 3610
<212> DNA
<213> Sesamum indicum/Synthetic

<400> 66
tcgacgaatt aattccaatc ccacaaaaat ctgagcttaa cagcacagtt gctcctctca 60
gagcagaatc gggtattcaa caccctcata tcaactacta cgttgtgtat aacggtccac 120
atgccggtat atacgatgac tggggttgta caaaggcggc aacaaacggc gttcccggag 180
ttgcacacaa gaaatttgcc actattacag aggcaagagc agcagctgac gcgtacacaa 240
caagtcagca aacagacagg ttgaacttca tccccaaagg agaagctcaa ctcaagccca 300
agagctttgc taaggcccta acaagcccac caaagcaaaa agcccactgg ctcacgctag 360
gaaccaaaag gcccagcagt gatccagccc caaaagagat ctcctttgcc ccggagatta 420
caatggacga tttcctctat ctttacgatc taggaaggaa gttcgaaggt gaaggtgacg 480
acactatgtt caccactgat aatgagaagg ttagcctctt caatttcaga aagaatgctg 540
acccacagat ggttagagag gcctacgcag caggtctcat caagacgatc tacccgagta 600
acaatctcca ggagatcaaa taccttccca agaaggttaa agatgcagtc aaaagattca 660
ggactaattg catcaagaac acagagaaag acatatttct caagatcaga agtactattc 720

cagtatggac gattcaaggc ttgcttcata aaccaaggca agtaatagag attggagtct 780

ctaaaaaggt agttcctact gaatctaagg ccatgcatgg agtctaagat tcaaatcgag 840

gatctaacag aactcgccgt gaagactggc gaacagttca tacagagtct tttacgactc 900

aatgacaaga agaaaatctt cgtcaacatg gtggagcacg acactctggt ctactccaaa 960

aatgtcaaag atacagtctc agaagaccaa agggctattg agacttttca acaaaggata 1020

atttcgggaa acctcctcgg attccattgc ccagctatct gtcacttcat cgaaaggaca 1080

gtagaaaagg aaggtggctc ctacaaatgc catcattgcg ataaaggaaa ggctatcatt 1140

caagatctct ctgccgacag tggtcccaaa gatggacccc cacccacgag gagcatcgtg 1200

gaaaaagaag acgttccaac cacgtcttca aagcaagtgg attgatgtga catctccact 1260

gacgtaaggg atgacgcaca atcccactat ccttcgcaag acccttcctc tatataagga 1320

agttcatttc atttggagag gacacgctcg aggaattcgg taccccatca caagtttgta 1380

caaaaaagca ggctccgcgg ccgcttgctc cgttaaaaaa aaccatggca gagcactacg 1440

ggcagcagca acaaacacgt gccccccact tgcaactgca accgcgtgct cagcgtgttg 1500

taaaggcagc gaccgcggtt actgccggag gtagcctgtt ggtgttatcc gggttgaccc 1560

tggctggaac ggtcattgcg ctgacaatcg ccacacctct cctggtgatc ttctcgcctg 1620

tactggtccc tgcaggtaaa tttctgtgtt ccttattctc tcaaaatctt cgattttgtt 1680

ttcgttcgat cccaatttcg tatatgttct ttggtttaga ttctgttaat cttagatcga 1740

agacgatttt ctgggtttga tcgttagata tcatcttaat tctcgattag ggtttcatag 1800

atatcatccg atttgttcaa ataatttgag ttttgtcgaa taattactct tcgatttgtg 1860

atttctatct agatctggtg ttagtttcta gtttgtgcga tcgaatttgt cgattaatct 1920

gagtttttct gatctgcagt catcaccatc ttcctgttgg gagcgggttt tctggcaagc 1980

gggggatttg gtgttgccgc tctgtctgtg ctgtcctgga tctaccgtta cctgacgggg 2040

aaacatccac ccggagcgga tcagttggag tcggccaaga caaagttggc gagcaaggcc 2100

cgtgaaatga aggaccgtgc cgagcagttc agtcagcaac cggtagcagg gtctcagacc 2160

agcatcgatc catcctcctg gctcgagatg gcggaacact acgggcaaca acagcagact 2220

cgtgctcccc acctgcaatt acaaccccgt gcccaacgtg ttgtgaaagc ggcaacagca 2280

gtaacggcag ggggaagttt gctggtctta tcggggttga ccttagcggg aaccgtgatt 2340

gccctgacaa ttgcgactcc gctgctggtt atcttcagcc ccgtattggt tccggccgtg 2400

atcacgattt ttttgctggg ggcaggattt ttagccagcg gaggatttgg ggtcgcagcg 2460

ttgtctgtgc tgagttggat ctatcgttat ttgaccggga agcacccacc tggagcagac 2520

cagctggaga gcgcgaaaac gaagctggca tcgaaggcgc gtgaaatgaa ggatcgtgct 2580

gaacaattct cccagcagcc tgttgccggt tctcagacca gccatatgat ggccgagcac 2640

tatggacaac agcagcagac gcgtgcccct catctgcaac tgcaaccccg tgctcaacgt 2700

gtcgttaagg ctgcgacagc agtaaccgct gggggttctc tgttagtgtt gtcagggctg 2760

actttggcgg ggacggtaat tgcgttgacc attgccaccc cgctgttagt gattttcagc 2820

ccggtactgg tgccagcagt tatcacgatc ttcttgctgg gtgccggatt cttggcaagt 2880

ggaggttttg gagttgcggc gctgtcagtt ttatcctgga tctatcgtta tctgacagga 2940

aaacatcccc caggtgccga tcagctggag agtgccaaga caaaactggc gtctaaggca 3000

cgtgaaatga aggatcgtgc cgaacagttt tctcaacagc ccgtagcggg gtcacagacc 3060

tcgatcgatc agcaggttaa cgtgcacatg gccgaacatt acggacagca acaacagacg 3120

cgtgctccac acctgcaatt gcaaccgcgt gctcaacgtg ttgtcaaagc ggcgaccgcc 3180

gtaacagcag gaggatcact gttagtgctg tcgggtttaa ccttggccgg gaccgtcatt 3240

gcattgacta ttgcgacgcc cttactggtg atcttttctc cggtgctggt tcccgccgtt 3300

attaccatct tcttgttagg ggcaggattc ctggcatcag ggggattcgg agttgcggcg 3360

ttgagtgtct taagttggat ctaccgttat ctgactggaa agcacccgcc tggggccgat 3420

caactggagt cagccaaaac gaaattggcg tcaaaagcgc gtgaaatgaa ggaccgtgct 3480

gagcagtttt ctcagcagcc tgtggcagga tcccagacat caccatggct cgagtaatga 3540

agcggccgca cccagctttc ttgtacaaag tggtgatggg ttcgaaatcg ataagcttgg 3600

atcctctaga 3610

<210> 67
<211> 4510
<212> DNA
<213> Sesamum indicum/Synthetic

<400> 67
tcgacgaatt aattccaatc ccacaaaaat ctgagcttaa cagcacagtt gctcctctca 60

gagcagaatc gggtattcaa caccctcata tcaactacta cgttgtgtat aacggtccac 120

atgccggtat atacgatgac tggggttgta caaaggcggc aacaaacggc gttcccggag 180

ttgcacacaa gaaatttgcc actattacag aggcaagagc agcagctgac gcgtacacaa 240

caagtcagca aacagacagg ttgaacttca tccccaaagg agaagctcaa ctcaagccca 300

agagctttgc taaggcccta acaagcccac caaagcaaaa agcccactgg ctcacgctag 360

gaaccaaaag gcccagcagt gatccagccc caaaagagat ctcctttgcc ccggagatta 420

caatggacga tttcctctat ctttacgatc taggaaggaa gttcgaaggt gaaggtgacg 480

acactatgtt caccactgat aatgagaagg ttagcctctt caatttcaga aagaatgctg 540

acccacagat ggttagagag gcctacgcag caggtctcat caagacgatc tacccgagta 600

acaatctcca ggagatcaaa taccttccca agaaggttaa agatgcagtc aaaagattca 660

ggactaattg catcaagaac acagagaaag acatatttct caagatcaga agtactattc 720

cagtatggac gattcaaggc ttgcttcata aaccaaggca agtaatagag attggagtct 780

ctaaaaaggt agttcctact gaatctaagg ccatgcatgg agtctaagat tcaaatcgag 840

gatctaacag aactcgccgt gaagactggc gaacagttca tacagagtct tttacgactc 900

aatgacaaga agaaaatctt cgtcaacatg gtggagcacg acactctggt ctactccaaa 960

aatgtcaaag atacagtctc agaagaccaa agggctattg agacttttca acaaaggata 1020
atttcgggaa acctcctcgg attccattgc ccagctatct gtcacttcat cgaaaggaca 1080
gtagaaaagg aaggtggctc ctacaaatgc catcattgcg ataaaggaaa ggctatcatt 1140
caagatctct ctgccgacag tggtcccaaa gatggacccc cacccacgag gagcatcgtg 1200
gaaaaagaag acgttccaac cacgtcttca aagcaagtgg attgatgtga catctccact 1260
gacgtaaggg atgacgcaca atcccactat ccttcgcaag acccttcctc tatataagga 1320
agttcatttc atttggagag gacacgctcg aggaattcgg taccccatca caagtttgta 1380
caaaaaagca ggctccgcgg ccgcttgctc cgttaaaaaa aaccatggca gagcactacg 1440
ggcagcagca acaaacacgt gccccccact tgcaactgca accgcgtgct cagcgtgttg 1500
taaaggcagc gaccgcggtt actgccggag gtagcctgtt ggtgttatcc gggttgaccc 1560
tggctggaac ggtcattgcg ctgacaatcg ccacacctct cctggtgatc ttctcgcctg 1620
tactggtccc tgcaggtaaa tttctgtgtt ccttattctc tcaaaatctt cgattttgtt 1680
ttcgttcgat cccaatttcg tatatgttct ttggtttaga ttctgttaat cttagatcga 1740
agacgatttt ctgggtttga tcgttagata tcatcttaat tctcgattag ggtttcatag 1800
atatcatccg atttgttcaa ataatttgag ttttgtcgaa taattactct tcgatttgtg 1860
atttctatct agatctggtg ttagtttcta gtttgtgcga tcgaatttgt cgattaatct 1920
gagtttttct gatctgcagt catcaccatc ttcctgttgg gagcgggttt tctggcaagc 1980
gggggatttg gtgttgccgc tctgtctgtg ctgtcctgga tctaccgtta cctgacgggg 2040
aaacatccac ccggagcgga tcagttggag tcggccaaga caaagttggc gagcaaggcc 2100
cgtgaaatga aggaccgtgc cgagcagttc agtcagcaac cggtagcagg gtctcagacc 2160
agcatcgatc catcctcctg gctcgagatg gcggaacact acgggcaaca acagcagact 2220
cgtgctcccc acctgcaatt acaaccccgt gcccaacgtg ttgtgaaagc ggcaacagca 2280
gtaacggcag ggggaagttt gctggtctta tcggggttga ccttagcggg aaccgtgatt 2340
gccctgacaa ttgcgactcc gctgctggtt atcttcagcc ccgtattggt tccggccgtg 2400
atcacgattt ttttgctggg ggcaggattt ttagccagcg gaggatttgg ggtcgcagcg 2460
ttgtctgtgc tgagttggat ctatcgttat ttgaccggga agcacccacc tggagcagac 2520
cagctggaga gcgcgaaaac gaagctggca tcgaaggcgc gtgaaatgaa ggatcgtgct 2580
gaacaattct cccagcagcc tgttgccggt tctcagacca gccatatgtt taaatggcca 2640
agcgctatgg ccgagcatta tgggcagcaa cagcaaaccc gtgccccgca tctgcaattg 2700
caacctcgtg cccagcgtgt cgttaaggcg gctactgcgg taacagcggg agggagctta 2760
ctggtattaa gcgggctgac attggccgga acggtgatcg ccttaacaat cgcgacaccc 2820
ttgctggtca tcttcagtcc ggttctggtg cccgcggtga ttacgatttt cctgctggga 2880
gccggtttct tagcatcggg gggttttggg gtagcagcct tgagtgtcct gtcgtggatc 2940
tatcgttact taactggaaa acacccgcca ggagctgacc agttggagtc tgcaaaaact 3000

aagctggcgt ccaaagcccg tgaaatgaag gatcgtgctg agcagtttag ccagcagcca 3060

gttgcgggaa gtcagacctc ttcatctgag ctcccatggg tcgacatggc ggagcattac 3120

ggtcaacagc aacagacccg tgctccgcac ttacaattgc aaccacgtgc tcaacgtgtc 3180

gtaaaagccg ccacggcagt tactgcgggg ggatcattgc tggtgttaag tgggttgaca 3240

ctggcgggga cagttattgc actgacgatc gcgacccccct tgttagtgat cttctccccc 3300

gttctggttc cggcggtcat tacaatcttt ctgttgggtg ccggattttt agcctctggg 3360

ggatttggag tagctgccct gtcagtgttg agctggatct accgttactt aacagggaag 3420

caccctcccg gggcagatca gttggaaagc gccaagacca agctggcaag taaagcgcgt 3480

gaaatgaagg accgtgccga acaattttcg cagcaaccgg ttgcgggatc acagacctct 3540

agtactccat cctcctggca tatgatggcc gagcactatg gacaacagca gcagacgcgt 3600

gcccctcatc tgcaactgca accccgtgct caacgtgtcg ttaaggctgc gacagcagta 3660

accgctgggg gttctctgtt agtgttgtca gggctgactt tggcggggac ggtaattgcg 3720

ttgaccattg ccaccccgct gttagtgatt ttcagcccgg tactggtgcc agcagttatc 3780

acgatcttct tgctgggtgc cggattcttg gcaagtggag gttttggagt tgcggcgctg 3840

tcagttttat cctggatcta tcgttatctg acaggaaaac atcccccagg tgccgatcag 3900

ctggagagtg ccaagacaaa actggcgtct aaggcacgtg aaatgaagga tcgtgccgaa 3960

cagttttctc aacagcccgt agcggggtca cagacctcga tcgatcagca ggttaacgtg 4020

cacatggccg aacattacgg acagcaacaa cagacgcgtg ctccacacct gcaattgcaa 4080

ccgcgtgctc aacgtgttgt caaagcggcg accgccgtaa cagcaggagg atcactgtta 4140

gtgctgtcgg gtttaacctt ggccgggacc gtcattgcat tgactattgc gacgccctta 4200

ctggtgatct tttctccggt gctggttccc gccgttatta ccatcttctt gttaggggca 4260

ggattcctgg catcaggggg attcggagtt gcggcgttga gtgtcttaag ttggatctac 4320

cgttatctga ctggaaagca cccgcctggg gccgatcaac tggagtcagc caaaacgaaa 4380

ttggcgtcaa aagcgcgtga aatgaaggac cgtgctgagc agttttctca gcagcctgtg 4440

gcaggatccc agacatcacc atggctcgag taatgaagcg gccgcaccca gctttcttgt 4500

acaaagtggt 4510

<210> 68
<211> 1706
<212> DNA
<213> Sesamum indicum/Synthetic

<400> 68
gcaggaactc tctggtaagc tagctccact ccccagaaac aaccggcgcc aaattgcgcg 60

aattgctgac ctgaagacgg aacatcatcg tcgggtcctt gggcgattgc ggcggaagat 120

gggtcagctt gggcttgagg acgagacccg aatccgagtc tgttgaaaag gttgttcatt 180

ggggatttgt atacggagat tggtcgtcga gaggtttgag ggaaaggaca aatgggtttg 240

gctctggaga aagagagtgc ggctttagag agagaattga gaggtttaga gagagatgcg 300

gcggcgatga gcggaggaga gacgacgagg acctgcatta tcaaagcagt gacgtggtga 360

aatttggaac ttttaagagg cagatagatt tattatttgt atccattttc ttcattgttc 420

tagaatgtcg cggaacaaat tttaaaacta aatcctaaat ttttctaatt ttgttgccaa 480

tagtggatat gtgggccgta tagaaggaat ctattgaagg cccaaaccca tactgacgag 540

cccaaaggtt cgttttgcgt tttatgtttc ggttcgatgc caacgccaca ttctgagcta 600

ggcaaaaaac aaacgtgtct ttgaatagac tcctctcgtt aacacatgca gcggctgcat 660

ggtgacgcca ttaacacgtg gcctacaatt gcatgatgtc tccattgaca cgtgacttct 720

cgtctccttt cttaatatat ctaacaaaca ctcctacctc ttccaaaata tatacacatc 780

tttttgatca atctctcatt caaaatctca ttctctctag taaacaagaa caaaaaaggt 840

accccatcac aagtttgtac aaaaaagcag gctccgcggc cgcttgctcc gttaaaaaaa 900

accatggcag agcactacgg gcagcagcaa caaacacgtg ccccccactt gcaactgcaa 960

ccgcgtgctc agcgtgttgt aaaggcagcg accgcggtta ctgccggagg tagcctgttg 1020

gtgttatccg ggttgaccct ggctggaacg gtcattgcgc tgacaatcgc cacacctctc 1080

ctggtgatct tctcgcctgt actggtccct gcaggtaaat ttctgtgttc cttattctct 1140

caaaatcttc gattttgttt tcgttcgatc ccaatttcgt atatgttctt tggtttagat 1200

tctgttaatc ttagatcgaa gacgattttc tgggtttgat cgttagatat catcttaatt 1260

ctcgattagg gtttcataga tatcatccga tttgttcaaa taatttgagt tttgtcgaat 1320

aattactctt cgatttgtga tttctatcta gatctggtgt tagtttctag tttgtgcgat 1380

cgaatttgtc gattaatctg agtttttctg atctgcagtc atcaccatct tcctgttggg 1440

agcgggtttt ctggcaagcg ggggatttgg tgttgccgct ctgtctgtgc tgtcctggat 1500

ctaccgttac ctgacgggga aacatccacc cggagcggat cagttggagt cggccaagac 1560

aaagttggcg agcaaggccc gtgaaatgaa ggaccgtgcc gagcagttca gtcagcaacc 1620

ggtagcaggg tctcagacca gcatcgatcc atcctcctgg ctcgagtaat gaagcggccg 1680

cacccagctt tcttgtacaa agtggt 1706

<210> 69
<211> 2612
<212> DNA
<213> Sesamum indicum/Synthetic

<400> 69
gcaggaactc tctggtaagc tagctccact ccccagaaac aaccggcgcc aaattgcgcg 60

aattgctgac ctgaagacgg aacatcatcg tcgggtcctt gggcgattgc ggcggaagat 120

gggtcagctt gggcttgagg acgagacccg aatccgagtc tgttgaaaag gttgttcatt 180

ggggatttgt atacggagat tggtcgtcga gaggtttgag ggaaaggaca aatgggtttg 240

gctctggaga aagagagtgc ggctttagag agagaattga gaggtttaga gagagatgcg 300

gcggcgatga gcggaggaga gacgacgagg acctgcatta tcaaagcagt gacgtggtga 360

aatttggaac ttttaagagg cagatagatt tattatttgt atccattttc ttcattgttc 420
tagaatgtcg cggaacaaat tttaaaacta aatcctaaat ttttctaatt ttgttgccaa 480
tagtggatat gtgggccgta tagaaggaat ctattgaagg cccaaaccca tactgacgag 540
cccaaaggtt cgttttgcgt tttatgtttc ggttcgatgc caacgccaca ttctgagcta 600
ggcaaaaaac aaacgtgtct ttgaatagac tcctctcgtt aacacatgca gcggctgcat 660
ggtgacgcca ttaacacgtg gcctacaatt gcatgatgtc tccattgaca cgtgacttct 720
cgtctccttt cttaatatat ctaacaaaca ctcctacctc ttccaaaata tatacacatc 780
tttttgatca atctctcatt caaaatctca ttctctctag taaacaagaa caaaaaaggt 840
accccatcac aagtttgtac aaaaaagcag gctccgcggc cgcttgctcc gttaaaaaaa 900
accatggcag agcactacgg gcagcagcaa caaacacgtg ccccccactt gcaactgcaa 960
ccgcgtgctc agcgtgttgt aaaggcagcg accgcggtta ctgccggagg tagcctgttg 1020
gtgttatccg ggttgaccct ggctggaacg gtcattgcgc tgacaatcgc cacacctctc 1080
ctggtgatct tctcgcctgt actggtccct gcaggtaaat ttctgtgttc cttattctct 1140
caaaatcttc gattttgttt tcgttcgatc ccaatttcgt atatgttctt tggtttagat 1200
tctgttaatc ttagatcgaa gacgattttc tgggtttgat cgttagatat catcttaatt 1260
ctcgattagg gtttcataga tatcatccga tttgttcaaa taatttgagt tttgtcgaat 1320
aattactctt cgatttgtga tttctatcta gatctggtgt tagtttctag tttgtgcgat 1380
cgaatttgtc gattaatctg agtttttctg atctgcagtc atcaccatct tcctgttggg 1440
agcgggtttt ctggcaagcg ggggatttgg tgttgccgct ctgtctgtgc tgtcctggat 1500
ctaccgttac ctgacgggga aacatccacc cggagcggat cagttggagt cggccaagac 1560
aaagttggcg agcaaggccc gtgaaatgaa ggaccgtgcc gagcagttca gtcagcaacc 1620
ggtagcaggg tctcagacca gcatcgatcc atcctcctgg catatgatgg ccgagcacta 1680
tggacaacag cagcagacgc gtgcccctca tctgcaactg caaccccgtg ctcaacgtgt 1740
cgttaaggct gcgacagcag taaccgctgg gggttctctg ttagtgttgt cagggctgac 1800
tttggcgggg acggtaattg cgttgaccat tgccaccccg ctgttagtga ttttcagccc 1860
ggtactggtg ccagcagtta tcacgatctt cttgctgggt gccggattct tggcaagtgg 1920
aggttttgga gttgcggcgc tgtcagtttt atcctggatc tatcgttatc tgacaggaaa 1980
acatccccca ggtgccgatc agctggagag tgccaagaca aaactggcgt ctaaggcacg 2040
tgaaatgaag gatcgtgccg aacagttttc tcaacagccc gtagcggggt cacagacctc 2100
gatcgatcag caggttaacg tgcacatggc cgaacattac ggacagcaac aacagacgcg 2160
tgctccacac ctgcaattgc aaccgcgtgc tcaacgtgtt gtcaaagcgg cgaccgccgt 2220
aacagcagga ggatcactgt tagtgctgtc gggtttaacc ttggccggga ccgtcattgc 2280
attgactatt gcgacgccct tactggtgat cttttctccg gtgctggttc ccgccgttat 2340
taccatcttc ttgttagggg caggattcct ggcatcaggg ggattcggag ttgcggcgtt 2400

gagtgtctta agttggatct accgttatct gactggaaag cacccgcctg gggccgatca 2460

actggagtca gccaaaacga aattggcgtc aaaagcgcgt gaaatgaagg accgtgctga 2520

gcagttttct cagcagcctg tggcaggatc ccagacatca ccatggctcg agtaatgaag 2580

cggccgcacc cagctttctt gtacaaagtg gt 2612

<210> 70
<211> 3089
<212> DNA
<213> Sesamum indicum/Synthetic

<400> 70
gcaggaactc tctggtaagc tagctccact ccccagaaac aaccggcgcc aaattgcgcg 60

aattgctgac ctgaagacgg aacatcatcg tcgggtcctt gggcgattgc ggcggaagat 120

gggtcagctt gggcttgagg acgagacccg aatccgagtc tgttgaaaag gttgttcatt 180

ggggatttgt atacggagat tggtcgtcga gaggtttgag ggaaaggaca aatgggtttg 240

gctctggaga aagagagtgc ggctttagag agagaattga gaggtttaga gagagatgcg 300

gcggcgatga gcggaggaga gacgacgagg acctgcatta tcaaagcagt gacgtggtga 360

aatttggaac ttttaagagg cagatagatt tattatttgt atccattttc ttcattgttc 420

tagaatgtcg cggaacaaat tttaaaacta aatcctaaat ttttctaatt ttgttgccaa 480

tagtggatat gtgggccgta tagaaggaat ctattgaagg cccaaaccca tactgacgag 540

cccaaaggtt cgttttgcgt tttatgtttc ggttcgatgc caacgccaca ttctgagcta 600

ggcaaaaaac aaacgtgtct ttgaatagac tcctctcgtt aacacatgca gcggctgcat 660

ggtgacgcca ttaacacgtg gcctacaatt gcatgatgtc tccattgaca cgtgacttct 720

cgtctccttt cttaatatat ctaacaaaca ctcctacctc ttccaaaata tatacacatc 780

tttttgatca atctctcatt caaaatctca ttctctctag taaacaagaa caaaaaaggt 840

accccatcac aagtttgtac aaaaaagcag gctccgcggc cgcttgctcc gttaaaaaaa 900

accatggcag agcactacgg gcagcagcaa caaacacgtg ccccccactt gcaactgcaa 960

ccgcgtgctc agcgtgttgt aaaggcagcg accgcggtta ctgccggagg tagcctgttg 1020

gtgttatccg ggttgaccct ggctggaacg gtcattgcgc tgacaatcgc cacacctctc 1080

ctggtgatct tctcgcctgt actggtccct gcaggtaaat ttctgtgttc cttattctct 1140

caaaatcttc gattttgttt tcgttcgatc ccaatttcgt atatgttctt tggtttagat 1200

tctgttaatc ttagatcgaa gacgattttc tgggtttgat cgttagatat catcttaatt 1260

ctcgattagg gtttcataga tatcatccga tttgttcaaa taatttgagt tttgtcgaat 1320

aattactctt cgatttgtga tttctatcta gatctggtgt tagtttctag tttgtgcgat 1380

cgaatttgtc gattaatctg agtttttctg atctgcagtc atcaccatct tcctgttggg 1440

agcgggtttt ctggcaagcg ggggatttgg tgttgccgct ctgtctgtgc tgtcctggat 1500

ctaccgttac ctgacgggga aacatccacc cggagcggat cagttggagt cggccaagac 1560

aaagttggcg agcaaggccc gtgaaatgaa ggaccgtgcc gagcagttca gtcagcaacc 1620

ggtagcaggg tctcagacca gcatcgatcc atcctcctgg ctcgagatgg cggaacacta 1680

cgggcaacaa cagcagactc gtgctcccca cctgcaatta caaccccgtg cccaacgtgt 1740

tgtgaaagcg gcaacagcag taacggcagg gggaagtttg ctggtcttat cggggttgac 1800

cttagcggga accgtgattg ccctgacaat tgcgactccg ctgctggtta tcttcagccc 1860

cgtattggtt ccggccgtga tcacgatttt tttgctgggg gcaggatttt tagccagcgg 1920

aggatttggg gtcgcagcgt tgtctgtgct gagttggatc tatcgttatt tgaccgggaa 1980

gcacccacct ggagcagacc agctggagag cgcgaaaacg aagctggcat cgaaggcgcg 2040

tgaaatgaag gatcgtgctg aacaattctc ccagcagcct gttgccggtt ctcagaccag 2100

ccatatgatg gccgagcact atggacaaca gcagcagacg cgtgcccctc atctgcaact 2160

gcaaccccgt gctcaacgtg tcgttaaggc tgcgacagca gtaaccgctg ggggttctct 2220

gttagtgttg tcagggctga ctttggcggg gacggtaatt gcgttgacca ttgccacccc 2280

gctgttagtg attttcagcc cggtactggt gccagcagtt atcacgatct tcttgctggg 2340

tgccggattc ttggcaagtg gaggttttgg agttgcggcg ctgtcagttt tatcctggat 2400

ctatcgttat ctgacaggaa aacatccccc aggtgccgat cagctggaga gtgccaagac 2460

aaaactggcg tctaaggcac gtgaaatgaa ggatcgtgcc gaacagtttt ctcaacagcc 2520

cgtagcgggg tcacagacct cgatcgatca gcaggttaac gtgcacatgg ccgaacatta 2580

cggacagcaa caacagacgc gtgctccaca cctgcaattg caaccgcgtg ctcaacgtgt 2640

tgtcaaagcg gcgaccgccg taacagcagg aggatcactg ttagtgctgt cgggtttaac 2700

cttggccggg accgtcattg cattgactat tgcgacgccc ttactggtga tcttttctcc 2760

ggtgctggtt cccgccgtta ttaccatctt cttgttaggg gcaggattcc tggcatcagg 2820

gggattcgga gttgcggcgt tgagtgtctt aagttggatc taccgttatc tgactggaaa 2880

gcacccgcct ggggccgatc aactggagtc agccaaaacg aaattggcgt caaaagcgcg 2940

tgaaatgaag gaccgtgctg agcagttttc tcagcagcct gtggcaggat cccagacatc 3000

accatggctc gagtaatgaa gcggccgcac ccagctttct tgtacaaagt ggtgatgggt 3060

tcgaaatcga taagcttgga tcctctaga 3089

<210> 71
<211> 3989
<212> DNA
<213> Sesamum indicum/Synthetic

<400> 71
gcaggaactc tctggtaagc tagctccact ccccagaaac aaccggcgcc aaattgcgcg 60

aattgctgac ctgaagacgg aacatcatcg tcgggtcctt gggcgattgc ggcggaagat 120

gggtcagctt gggcttgagg acgagacccg aatccgagtc tgttgaaaag gttgttcatt 180

ggggatttgt atacggagat tggtcgtcga gaggtttgag ggaaaggaca aatgggtttg 240

gctctggaga aagagagtgc ggctttagag agagaattga gaggtttaga gagagatgcg 300

gcggcgatga gcggaggaga gacgacgagg acctgcatta tcaaagcagt gacgtggtga 360
aatttggaac ttttaagagg cagatagatt tattatttgt atccattttc ttcattgttc 420
tagaatgtcg cggaacaaat tttaaaacta aatcctaaat ttttctaatt ttgttgccaa 480
tagtggatat gtgggccgta tagaaggaat ctattgaagg cccaaaccca tactgacgag 540
cccaaaggtt cgttttgcgt tttatgtttc ggttcgatgc caacgccaca ttctgagcta 600
ggcaaaaaac aaacgtgtct ttgaatagac tcctctcgtt aacacatgca gcggctgcat 660
ggtgacgcca ttaacacgtg gcctacaatt gcatgatgtc tccattgaca cgtgacttct 720
cgtctccttt cttaatatat ctaacaaaca ctcctacctc ttccaaaata tatacacatc 780
tttttgatca atctctcatt caaaatctca ttctctctag taaacaagaa caaaaaaggt 840
accccatcac aagtttgtac aaaaaagcag gctccgcggc cgcttgctcc gttaaaaaaa 900
accatggcag agcactacgg gcagcagcaa caaacacgtg ccccccactt gcaactgcaa 960
ccgcgtgctc agcgtgttgt aaaggcagcg accgcggtta ctgccggagg tagcctgttg 1020
gtgttatccg ggttgaccct ggctggaacg gtcattgcgc tgacaatcgc cacacctctc 1080
ctggtgatct tctcgcctgt actggtccct gcaggtaaat ttctgtgttc cttattctct 1140
caaaatcttc gattttgttt tcgttcgatc ccaatttcgt atatgttctt tggtttagat 1200
tctgttaatc ttagatcgaa gacgattttc tgggtttgat cgttagatat catcttaatt 1260
ctcgattagg gtttcataga tatcatccga tttgttcaaa taatttgagt tttgtcgaat 1320
aattactctt cgatttgtga tttctatcta gatctggtgt tagtttctag tttgtgcgat 1380
cgaatttgtc gattaatctg agtttttctg atctgcagtc atcaccatct tcctgttggg 1440
agcgggtttt ctggcaagcg ggggatttgg tgttgccgct ctgtctgtgc tgtcctggat 1500
ctaccgttac ctgacgggga aacatccacc cggagcggat cagttggagt cggccaagac 1560
aaagttggcg agcaaggccc gtgaaatgaa ggaccgtgcc gagcagttca gtcagcaacc 1620
ggtagcaggg tctcagacca gcatcgatcc atcctcctgg ctcgagatgg cggaacacta 1680
cgggcaacaa cagcagactc gtgctcccca cctgcaatta caaccccgtg cccaacgtgt 1740
tgtgaaagcg gcaacagcag taacggcagg gggaagtttg ctggtcttat cggggttgac 1800
cttagcggga accgtgattg ccctgacaat tgcgactccg ctgctggtta tcttcagccc 1860
cgtattggtt ccggccgtga tcacgatttt tttgctgggg gcaggatttt tagccagcgg 1920
aggatttggg gtcgcagcgt tgtctgtgct gagttggatc tatcgttatt tgaccgggaa 1980
gcacccacct ggagcagacc agctggagag cgcgaaaacg aagctggcat cgaaggcgcg 2040
tgaaatgaag gatcgtgctg aacaattctc ccagcagcct gttgccggtt ctcagaccag 2100
ccatatgttt aaatggccaa gcgctatggc cgagcattat gggcagcaac agcaaacccg 2160
tgccccgcat ctgcaattgc aacctcgtgc ccagcgtgtc gttaaggcgg ctactgcggt 2220
aacagcggga gggagcttac tggtattaag cgggctgaca ttggccggaa cggtgatcgc 2280
cttaacaatc gcgacaccct tgctggtcat cttcagtccg gttctggtgc ccgcggtgat 2340

tacgattttc ctgctgggag ccggtttctt agcatcgggg ggttttgggg tagcagcctt 2400

gagtgtcctg tcgtggatct atcgttactt aactggaaaa cacccgccag gagctgacca 2460

gttggagtct gcaaaaacta agctggcgtc caaagcccgt gaaatgaagg atcgtgctga 2520

gcagtttagc cagcagccag ttgcgggaag tcagacctct tcatctgagc tcccatgggt 2580

cgacatggcg gagcattacg gtcaacagca acagacccgt gctccgcact tacaattgca 2640

accacgtgct caacgtgtcg taaaagccgc cacggcagtt actgcggggg gatcattgct 2700

ggtgttaagt gggttgacac tggcggggac agttattgca ctgacgatcg cgacccccctt 2760

gttagtgatc ttctcccccg ttctggttcc ggcggtcatt acaatctttc tgttgggtgc 2820

cggatttta gcctctgggg gatttggagt agctgccctg tcagtgttga gctggatcta 2880

ccgttactta acagggaagc accctcccgg ggcagatcag ttggaaagcg ccaagaccaa 2940

gctggcaagt aaagcgcgtg aaatgaagga ccgtgccgaa caattttcgc agcaaccggt 3000

tgcgggatca cagacctcta gtactccatc ctcctggcat atgatggccg agcactatgg 3060

acaacagcag cagacgcgtg cccctcatct gcaactgcaa ccccgtgctc aacgtgtcgt 3120

taaggctgcg acagcagtaa ccgctggggg ttctctgtta gtgttgtcag ggctgacttt 3180

ggcggggacg gtaattgcgt tgaccattgc caccccgctg ttagtgattt tcagcccggt 3240

actggtgcca gcagttatca cgatcttctt gctgggtgcc ggattcttgg caagtggagg 3300

ttttggagtt gcggcgctgt cagttttatc ctggatctat cgttatctga caggaaaaca 3360

tcccccaggt gccgatcagc tggagagtgc caagacaaaa ctggcgtcta aggcacgtga 3420

aatgaaggat cgtgccgaac agttttctca acagcccgta gcggggtcac agacctcgat 3480

cgatcagcag gttaacgtgc acatggccga acattacgga cagcaacaac agacgcgtgc 3540

tccacacctg caattgcaac cgcgtgctca acgtgttgtc aaagcggcga ccgccgtaac 3600

agcaggagga tcactgttag tgctgtcggg tttaaccttg gccgggaccg tcattgcatt 3660

gactattgcg acgcccttac tggtgatctt ttctccggtg ctggttcccg ccgttattac 3720

catcttcttg ttaggggcag gattcctggc atcaggggga ttcggagttg cggcgttgag 3780

tgtcttaagt tggatctacc gttatctgac tggaaagcac ccgcctgggg ccgatcaact 3840

ggagtcagcc aaaacgaaat tggcgtcaaa agcgcgtgaa atgaaggacc gtgctgagca 3900

gttttctcag cagcctgtgg caggatccca gacatcacca tggctcgagt aatgaagcgg 3960

ccgcacccag ctttcttgta caaagtggt 3989

<210> 72
<211> 2787
<212> DNA
<213> Sesamum indicum/Synthetic

<400> 72
tcgacgaatt aattccaatc ccacaaaaat ctgagcttaa cagcacagtt gctcctctca 60

gagcagaatc gggtattcaa caccctcata tcaactacta cgttgtgtat aacggtccac 120

atgccggtat atacgatgac tggggttgta caaaggcggc aacaaacggc gttcccggag 180
ttgcacacaa gaaatttgcc actattacag aggcaagagc agcagctgac gcgtacacaa 240
caagtcagca aacagacagg ttgaacttca tccccaaagg agaagctcaa ctcaagccca 300
agagctttgc taaggcccta acaagcccac caaagcaaaa agcccactgg ctcacgctag 360
gaaccaaaag gcccagcagt gatccagccc caaaagagat ctcctttgcc ccggagatta 420
caatggacga tttcctctat ctttacgatc taggaaggaa gttcgaaggt gaaggtgacg 480
acactatgtt caccactgat aatgagaagg ttagcctctt caatttcaga aagaatgctg 540
acccacagat ggttagagag gcctacgcag caggtctcat caagacgatc tacccgagta 600
acaatctcca ggagatcaaa taccttccca agaaggttaa agatgcagtc aaaagattca 660
ggactaattg catcaagaac acagagaaag acatatttct caagatcaga agtactattc 720
cagtatggac gattcaaggc ttgcttcata aaccaaggca agtaatagag attggagtct 780
ctaaaaaggt agttcctact gaatctaagg ccatgcatgg agtctaagat tcaaatcgag 840
gatctaacag aactcgccgt gaagactggc gaacagttca tacagagtct tttacgactc 900
aatgacaaga agaaaatctt cgtcaacatg gtggagcacg acactctggt ctactccaaa 960
aatgtcaaag atacagtctc agaagaccaa agggctattg agacttttca acaaaggata 1020
atttcgggaa acctcctcgg attccattgc ccagctatct gtcacttcat cgaaaggaca 1080
gtagaaaagg aaggtggctc ctacaaatgc catcattgcg ataaaggaaa ggctatcatt 1140
caagatctct ctgccgacag tggtcccaaa gatggacccc cacccacgag gagcatcgtg 1200
gaaaaagaag acgttccaac cacgtcttca aagcaagtgg attgatgtga catctccact 1260
gacgtaaggg atgacgcaca atcccactat ccttcgcaag acccttcctc tatataagga 1320
agttcatttc atttggagag gacacgctcg aggaattcgg taccccatca caagtttgta 1380
caaaaaagca ggctccgcgg ccgcttgctc cgttaaaaaa aaccatggct gagcattatg 1440
gtcaacaaca gcagaccagg gcgcctcacc tgcagctgca gccgcgcgcc cagcgggtag 1500
tgaaggcggc caccgccgtg acagccggcg gctcgcttct cgtcctctct ggcctcactt 1560
tagccggaac tgttattgcg ctcaccatcg ccactccgct gcttgtgatc tttagccccg 1620
ttctggtgcc ggcggtcata accattttct tgctgggtgc gggttttctg gcatccggag 1680
gcttcggcgt ggcggcgctg agtgtgctgt cgtggattta cagatatctg acagggaaac 1740
acccgccggg ggcggatcag ctggaatcgg caaagacgaa gctggcgagc aaggcgcgag 1800
agatgaagga tagggcagag cagttctcgc agcagcctgt tggaggcggt ggatccggag 1860
gcggtggtag tatggctgag cattatggtc aacaacagca gaccagggcg cctcacctgc 1920
agctgcagcc gcgcgcccag cgggtagtga aggcggccac cgccgtgaca gccggcggct 1980
cgcttctcgt cctctctggc ctcactttag ccggaactgt tattgcgctc accatcgcca 2040
ctccgctgct tgtgatcttt agccccgttc tggtgccggc ggtcataacc attttcttgc 2100
tgggtgcggg ttttctggca tccggaggct tcggcgtggc ggcgctgagt gtgctgtcgt 2160

ggatttacag atatctgaca gggaaacacc cgccgggggc ggatcagctg gaatcggcaa 2220

agacgaagct ggcgagcaag gcgcgagaga tgaaggatag ggcagagcag ttctcgcagc 2280

agcctgttgg gggcggtgga tccggtggag ggggatccat ggctgagcat tatggtcaac 2340

aacagcagac cagggcgcct cacctgcagc tgcagccgcg cgcccagcgg gtagtgaagg 2400

cggccaccgc cgtgacagcc ggcggctcgc ttctcgtcct ctctggcctc actttagccg 2460

gaactgttat tgcgctcacc atcgccactc cgctgcttgt gatctttagc cccgttctgg 2520

tgccggcggt cataaccatt ttcttgctgg gtgcgggttt tctggcatcc ggaggcttcg 2580

gcgtggcggc gctgagtgtg ctgtcgtgga tttacagata tctgacaggg aaacacccgc 2640

cgggggcgga tcagctggaa tcggcaaaga cgaagctggc gagcaaggcg cgagagatga 2700

aggatagggc agagcagttc tcgcagcagc ctgttccatg gctcgagtaa tgaagcggcc 2760

gcacccagct ttcttgtaca aagtggt 2787

<210> 73
<211> 2263
<212> DNA
<213> Sesamum indicum/Synthetic

<400> 73
gcaggaactc tctggtaagc tagctccact ccccagaaac aaccggcgcc aaattgcgcg 60

aattgctgac ctgaagacgg aacatcatcg tcgggtcctt gggcgattgc ggcggaagat 120

gggtcagctt gggcttgagg acgagacccg aatccgagtc tgttgaaaag gttgttcatt 180

ggggatttgt atacggagat tggtcgtcga gaggtttgag ggaaaggaca aatgggtttg 240

gctctggaga aagagagtgc ggctttagag agagaattga gaggtttaga gagagatgcg 300

gcggcgatga gcggaggaga gacgacgagg acctgcatta tcaaagcagt gacgtggtga 360

aatttggaac ttttaagagg cagatagatt tattatttgt atccattttc ttcattgttc 420

tagaatgtcg cggaacaaat tttaaaacta aatcctaaat ttttctaatt ttgttgccaa 480

tagtggatat gtgggccgta tagaaggaat ctattgaagg cccaaaccca tactgacgag 540

cccaaaggtt cgttttgcgt tttatgtttc ggttcgatgc caacgccaca ttctgagcta 600

ggcaaaaaac aaacgtgtct ttgaatagac tcctctcgtt aacacatgca gcggctgcat 660

ggtgacgcca ttaacacgtg gcctacaatt gcatgatgtc tccattgaca cgtgacttct 720

cgtctccttt cttaatatat ctaacaaaca ctcctacctc ttccaaaata tatacacatc 780

tttttgatca atctctcatt caaaatctca ttctctctag taaacaagaa caaaaaaggt 840

accccatcac aagtttgtac aaaaaagcag gctccgcggc cgcttgctcc gttaaaaaaa 900

accatggctg agcattatgg tcaacaacag cagaccaggg cgcctcacct gcagctgcag 960

ccgcgcgccc agcgggtagt gaaggcggcc accgccgtga cagccggcgg ctcgcttctc 1020

gtcctctctg gcctcacttt agccggaact gttattgcgc tcaccatcgc cactccgctg 1080

cttgtgatct ttagccccgt tctggtgccg gcggtcataa ccattttctt gctgggtgcg 1140

ggttttctgg catccggagg cttcggcgtg gcggcgctga gtgtgctgtc gtggatttac 1200

agatatctga cagggaaaca cccgccgggg gcggatcagc tggaatcggc aaagacgaag 1260

ctggcgagca aggcgcgaga gatgaaggat agggcagagc agttctcgca gcagcctgtt 1320

ggaggcggtg gatccggagg cggtggtagt atggctgagc attatggtca acaacagcag 1380

accagggcgc ctcacctgca gctgcagccg cgcgcccagc gggtagtgaa ggcggccacc 1440

gccgtgacag ccggcggctc gcttctcgtc ctctctggcc tcactttagc cggaactgtt 1500

attgcgctca ccatcgccac tccgctgctt gtgatcttta gccccgttct ggtgccggcg 1560

gtcataacca ttttcttgct gggtgcgggt tttctggcat ccggaggctt cggcgtggcg 1620

gcgctgagtg tgctgtcgtg gatttacaga tatctgacag ggaaacaccc gccgggggcg 1680

gatcagctgg aatcggcaaa gacgaagctg gcgagcaagg cgcgagagat gaaggatagg 1740

gcagagcagt tctcgcagca gcctgttggg ggcggtggat ccggtggagg gggatccatg 1800

gctgagcatt atggtcaaca acagcagacc agggcgcctc acctgcagct gcagccgcgc 1860

gcccagcggg tagtgaaggc ggccaccgcc gtgacagccg gcggctcgct tctcgtcctc 1920

tctggcctca ctttagccgg aactgttatt gcgctcacca tcgccactcc gctgcttgtg 1980

atctttagcc ccgttctggt gccggcggtc ataaccattt tcttgctggg tgcgggtttt 2040

ctggcatccg gaggcttcgg cgtggcggcg ctgagtgtgc tgtcgtggat ttacagatat 2100

ctgacaggga aacacccgcc gggggcggat cagctggaat cggcaaagac gaagctggcg 2160

agcaaggcgc gagagatgaa ggatagggca gagcagttct cgcagcagcc tgttccatgg 2220

ctcgagtaat gaagcggccg cacccagctt tcttgtacaa agt 2263

<210> 74
<211> 153
<212> PRT
<213> Sesamum indicum/Synthetic

<400> 74

Met Ala Glu His Tyr Gly Gln Gln Gln Gln Thr Arg Ala Pro His Leu
1 5 10 15

Gln Leu Gln Pro Arg Ala Gln Arg Val Val Lys Ala Ala Thr Ala Val
20 25 30

Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Gly Leu Thr Leu Ala Gly
35 40 45

Thr Val Ile Ala Leu Thr Ile Ala Thr Pro Leu Leu Val Ile Phe Ser
50 55 60

Pro Val Leu Val Pro Ala Val Ile Thr Ile Phe Leu Leu Gly Ala Gly
65 70 75 80

Phe Leu Ala Ser Gly Gly Phe Gly Val Ala Ala Leu Ser Val Leu Ser
85 90 95

Trp Ile Tyr Arg Tyr Leu Thr Gly Lys His Pro Pro Gly Ala Asp Gln
100 105 110

Leu Glu Ser Ala Lys Thr Lys Leu Ala Ser Lys Ala Arg Glu Met Lys
115 120 125

Asp Arg Ala Glu Gln Phe Ser Gln Gln Pro Val Ala Gly Ser Gln Thr
130 135 140

Ser Ile Asp Pro Ser Ser Trp Leu Glu
145 150

<210> 75
<211> 455
<212> PRT
<213> Sesamum indicum/Synthetic

<400> 75

Met Ala Glu His Tyr Gly Gln Gln Gln Thr Arg Ala Pro His Leu
1 5 10 15

Gln Leu Gln Pro Arg Ala Gln Arg Val Val Lys Ala Ala Thr Ala Val
20 25 30

Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Gly Leu Thr Leu Ala Gly
35 40 45

Thr Val Ile Ala Leu Thr Ile Ala Thr Pro Leu Leu Val Ile Phe Ser
50 55 60

Pro Val Leu Val Pro Ala Val Ile Thr Ile Phe Leu Leu Gly Ala Gly
65 70 75 80

Phe Leu Ala Ser Gly Gly Phe Gly Val Ala Ala Leu Ser Val Leu Ser
85 90 95

Trp Ile Tyr Arg Tyr Leu Thr Gly Lys His Pro Pro Gly Ala Asp Gln
100 105 110

Leu Glu Ser Ala Lys Thr Lys Leu Ala Ser Lys Ala Arg Glu Met Lys
115 120 125

Asp Arg Ala Glu Gln Phe Ser Gln Gln Pro Val Ala Gly Ser Gln Thr
130 135 140

Ser Ile Asp Pro Ser Ser Trp His Met Met Ala Glu His Tyr Gly Gln
145 150 155 160

Gln Gln Gln Thr Arg Ala Pro His Leu Gln Leu Gln Pro Arg Ala Gln
165 170 175

```
Arg Val Val Lys Ala Ala Thr Ala Val Thr Ala Gly Gly Ser Leu Leu
            180                 185                 190

Val Leu Ser Gly Leu Thr Leu Ala Gly Thr Val Ile Ala Leu Thr Ile
        195                 200                 205

Ala Thr Pro Leu Leu Val Ile Phe Ser Pro Val Leu Val Pro Ala Val
    210                 215                 220

Ile Thr Ile Phe Leu Leu Gly Ala Gly Phe Leu Ala Ser Gly Gly Phe
225                 230                 235                 240

Gly Val Ala Ala Leu Ser Val Leu Ser Trp Ile Tyr Arg Tyr Leu Thr
                245                 250                 255

Gly Lys His Pro Pro Gly Ala Asp Gln Leu Glu Ser Ala Lys Thr Lys
            260                 265                 270

Leu Ala Ser Lys Ala Arg Glu Met Lys Asp Arg Ala Glu Gln Phe Ser
        275                 280                 285

Gln Gln Pro Val Ala Gly Ser Gln Thr Ser Ile Asp Gln Gln Val Asn
    290                 295                 300

Val His Met Ala Glu His Tyr Gly Gln Gln Gln Gln Thr Arg Ala Pro
305                 310                 315                 320

His Leu Gln Leu Gln Pro Arg Ala Gln Arg Val Val Lys Ala Ala Thr
                325                 330                 335

Ala Val Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Gly Leu Thr Leu
            340                 345                 350

Ala Gly Thr Val Ile Ala Leu Thr Ile Ala Thr Pro Leu Leu Val Ile
        355                 360                 365

Phe Ser Pro Val Leu Val Pro Ala Val Ile Thr Ile Phe Leu Leu Gly
    370                 375                 380

Ala Gly Phe Leu Ala Ser Gly Gly Phe Gly Val Ala Ala Leu Ser Val
385                 390                 395                 400

Leu Ser Trp Ile Tyr Arg Tyr Leu Thr Gly Lys His Pro Pro Gly Ala
                405                 410                 415

Asp Gln Leu Glu Ser Ala Lys Thr Lys Leu Ala Ser Lys Ala Arg Glu
            420                 425                 430

Met Lys Asp Arg Ala Glu Gln Phe Ser Gln Gln Pro Val Ala Gly Ser
        435                 440                 445
```

Gln Thr Ser Pro Trp Leu Glu
450 455

<210> 76
<211> 602
<212> PRT
<213> Sesamum indicum/Synthetic

<400> 76

Met Ala Glu His Tyr Gly Gln Gln Gln Gln Thr Arg Ala Pro His Leu
1 5 10 15

Gln Leu Gln Pro Arg Ala Gln Arg Val Val Lys Ala Ala Thr Ala Val
20 25 30

Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Gly Leu Thr Leu Ala Gly
35 40 45

Thr Val Ile Ala Leu Thr Ile Ala Thr Pro Leu Leu Val Ile Phe Ser
50 55 60

Pro Val Leu Val Pro Ala Val Ile Thr Ile Phe Leu Leu Gly Ala Gly
65 70 75 80

Phe Leu Ala Ser Gly Gly Phe Gly Val Ala Ala Leu Ser Val Leu Ser
85 90 95

Trp Ile Tyr Arg Tyr Leu Thr Gly Lys His Pro Pro Gly Ala Asp Gln
100 105 110

Leu Glu Ser Ala Lys Thr Lys Leu Ala Ser Lys Ala Arg Glu Met Lys
115 120 125

Asp Arg Ala Glu Gln Phe Ser Gln Gln Pro Val Ala Gly Ser Gln Thr
130 135 140

Ser Ile Asp Pro Ser Ser Trp Leu Glu Met Ala Glu His Tyr Gly Gln
145 150 155 160

Gln Gln Gln Thr Arg Ala Pro His Leu Gln Leu Gln Pro Arg Ala Gln
165 170 175

Arg Val Val Lys Ala Ala Thr Ala Val Thr Ala Gly Gly Ser Leu Leu
180 185 190

Val Leu Ser Gly Leu Thr Leu Ala Gly Thr Val Ile Ala Leu Thr Ile
195 200 205

Ala Thr Pro Leu Leu Val Ile Phe Ser Pro Val Leu Val Pro Ala Val
210 215 220

Ile Thr Ile Phe Leu Leu Gly Ala Gly Phe Leu Ala Ser Gly Gly Phe

225 230 235 240

Gly Val Ala Ala Leu Ser Val Leu Ser Trp Ile Tyr Arg Tyr Leu Thr
245 250 255

Gly Lys His Pro Pro Gly Ala Asp Gln Leu Glu Ser Ala Lys Thr Lys
260 265 270

Leu Ala Ser Lys Ala Arg Glu Met Lys Asp Arg Ala Glu Gln Phe Ser
275 280 285

Gln Gln Pro Val Ala Gly Ser Gln Thr Ser His Met Met Ala Glu His
290 295 300

Tyr Gly Gln Gln Gln Gln Thr Arg Ala Pro His Leu Gln Leu Gln Pro
305 310 315 320

Arg Ala Gln Arg Val Val Lys Ala Ala Thr Ala Val Thr Ala Gly Gly
325 330 335

Ser Leu Leu Val Leu Ser Gly Leu Thr Leu Ala Gly Thr Val Ile Ala
340 345 350

Leu Thr Ile Ala Thr Pro Leu Leu Val Ile Phe Ser Pro Val Leu Val
355 360 365

Pro Ala Val Ile Thr Ile Phe Leu Leu Gly Ala Gly Phe Leu Ala Ser
370 375 380

Gly Gly Phe Gly Val Ala Ala Leu Ser Val Leu Ser Trp Ile Tyr Arg
385 390 395 400

Tyr Leu Thr Gly Lys His Pro Pro Gly Ala Asp Gln Leu Glu Ser Ala
405 410 415

Lys Thr Lys Leu Ala Ser Lys Ala Arg Glu Met Lys Asp Arg Ala Glu
420 425 430

Gln Phe Ser Gln Gln Pro Val Ala Gly Ser Gln Thr Ser Ile Asp Gln
435 440 445

Gln Val Asn Val His Met Ala Glu His Tyr Gly Gln Gln Gln Gln Thr
450 455 460

Arg Ala Pro His Leu Gln Leu Gln Pro Arg Ala Gln Arg Val Val Lys
465 470 475 480

Ala Ala Thr Ala Val Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Gly
485 490 495

Leu Thr Leu Ala Gly Thr Val Ile Ala Leu Thr Ile Ala Thr Pro Leu

500 505 510

Leu Val Ile Phe Ser Pro Val Leu Val Pro Ala Val Ile Thr Ile Phe
515 520 525

Leu Leu Gly Ala Gly Phe Leu Ala Ser Gly Gly Phe Gly Val Ala Ala
530 535 540

Leu Ser Val Leu Ser Trp Ile Tyr Arg Tyr Leu Thr Gly Lys His Pro
545 550 555 560

Pro Gly Ala Asp Gln Leu Glu Ser Ala Lys Thr Lys Leu Ala Ser Lys
565 570 575

Ala Arg Glu Met Lys Asp Arg Ala Glu Gln Phe Ser Gln Gln Pro Val
580 585 590

Ala Gly Ser Gln Thr Ser Pro Trp Leu Glu
595 600

<210> 77
<211> 914
<212> PRT
<213> Sesamum indicum/Synthetic

<400> 77

Met Ala Glu His Tyr Gly Gln Gln Gln Gln Thr Arg Ala Pro His Leu
1 5 10 15

Gln Leu Gln Pro Arg Ala Gln Arg Val Val Lys Ala Ala Thr Ala Val
20 25 30

Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Gly Leu Thr Leu Ala Gly
35 40 45

Thr Val Ile Ala Leu Thr Ile Ala Thr Pro Leu Leu Val Ile Phe Ser
50 55 60

Pro Val Leu Val Pro Ala Val Ile Thr Ile Phe Leu Leu Gly Ala Gly
65 70 75 80

Phe Leu Ala Ser Gly Gly Phe Gly Val Ala Ala Leu Ser Val Leu Ser
85 90 95

Trp Ile Tyr Arg Tyr Leu Thr Gly Lys His Pro Pro Gly Ala Asp Gln
100 105 110

Leu Glu Ser Ala Lys Thr Lys Leu Ala Ser Lys Ala Arg Glu Met Lys
115 120 125

Asp Arg Ala Glu Gln Phe Ser Gln Gln Pro Val Ala Gly Ser Gln Thr
130 135 140

```
Ser Ile Asp Pro Ser Ser Trp Leu Glu Met Ala Glu His Tyr Gly Gln
145                 150                 155                 160

Gln Gln Gln Thr Arg Ala Pro His Leu Gln Leu Gln Pro Arg Ala Gln
                165                 170                 175

Arg Val Val Lys Ala Ala Thr Ala Val Thr Ala Gly Gly Ser Leu Leu
            180                 185                 190

Val Leu Ser Gly Leu Thr Leu Ala Gly Thr Val Ile Ala Leu Thr Ile
        195                 200                 205

Ala Thr Pro Leu Leu Val Ile Phe Ser Pro Val Leu Val Pro Ala Val
    210                 215                 220

Ile Thr Ile Phe Leu Leu Gly Ala Gly Phe Leu Ala Ser Gly Gly Phe
225                 230                 235                 240

Gly Val Ala Ala Leu Ser Val Leu Ser Trp Ile Tyr Arg Tyr Leu Thr
                245                 250                 255

Gly Lys His Pro Pro Gly Ala Asp Gln Leu Glu Ser Ala Lys Thr Lys
            260                 265                 270

Leu Ala Ser Lys Ala Arg Glu Met Lys Asp Arg Ala Glu Gln Phe Ser
        275                 280                 285

Gln Gln Pro Val Ala Gly Ser Gln Thr Ser His Met Phe Lys Trp Pro
    290                 295                 300

Ser Ala Met Ala Glu His Tyr Gly Gln Gln Gln Gln Thr Arg Ala Pro
305                 310                 315                 320

His Leu Gln Leu Gln Pro Arg Ala Gln Arg Val Val Lys Ala Ala Thr
                325                 330                 335

Ala Val Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Gly Leu Thr Leu
            340                 345                 350

Ala Gly Thr Val Ile Ala Leu Thr Ile Ala Thr Pro Leu Leu Val Ile
        355                 360                 365

Phe Ser Pro Val Leu Val Pro Ala Val Ile Thr Ile Phe Leu Leu Gly
    370                 375                 380

Ala Gly Phe Leu Ala Ser Gly Gly Phe Gly Val Ala Ala Leu Ser Val
385                 390                 395                 400

Leu Ser Trp Ile Tyr Arg Tyr Leu Thr Gly Lys His Pro Pro Gly Ala
                405                 410                 415
```

Asp Gln Leu Glu Ser Ala Lys Thr Lys Leu Ala Ser Lys Ala Arg Glu
420 425 430

Met Lys Asp Arg Ala Glu Gln Phe Ser Gln Gln Pro Val Ala Gly Ser
435 440 445

Gln Thr Ser Ser Ser Glu Leu Pro Trp Val Asp Met Ala Glu His Tyr
450 455 460

Gly Gln Gln Gln Gln Thr Arg Ala Pro His Leu Gln Leu Gln Pro Arg
465 470 475 480

Ala Gln Arg Val Val Lys Ala Ala Thr Ala Val Thr Ala Gly Gly Ser
485 490 495

Leu Leu Val Leu Ser Gly Leu Thr Leu Ala Gly Thr Val Ile Ala Leu
500 505 510

Thr Ile Ala Thr Pro Leu Leu Val Ile Phe Ser Pro Val Leu Val Pro
515 520 525

Ala Val Ile Thr Ile Phe Leu Leu Gly Ala Gly Phe Leu Ala Ser Gly
530 535 540

Gly Phe Gly Val Ala Ala Leu Ser Val Leu Ser Trp Ile Tyr Arg Tyr
545 550 555 560

Leu Thr Gly Lys His Pro Pro Gly Ala Asp Gln Leu Glu Ser Ala Lys
565 570 575

Thr Lys Leu Ala Ser Lys Ala Arg Glu Met Lys Asp Arg Ala Glu Gln
580 585 590

Phe Ser Gln Gln Pro Val Ala Gly Ser Gln Thr Ser Ser Thr Pro Ser
595 600 605

Ser Trp His Met Met Ala Glu His Tyr Gly Gln Gln Gln Gln Thr Arg
610 615 620

Ala Pro His Leu Gln Leu Gln Pro Arg Ala Gln Arg Val Val Lys Ala
625 630 635 640

Ala Thr Ala Val Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Gly Leu
645 650 655

Thr Leu Ala Gly Thr Val Ile Ala Leu Thr Ile Ala Thr Pro Leu Leu
660 665 670

Val Ile Phe Ser Pro Val Leu Val Pro Ala Val Ile Thr Ile Phe Leu
675 680 685

Leu Gly Ala Gly Phe Leu Ala Ser Gly Gly Phe Gly Val Ala Ala Leu
690 695 700

Ser Val Leu Ser Trp Ile Tyr Arg Tyr Leu Thr Gly Lys His Pro Pro
705 710 715 720

Gly Ala Asp Gln Leu Glu Ser Ala Lys Thr Lys Leu Ala Ser Lys Ala
725 730 735

Arg Glu Met Lys Asp Arg Ala Glu Gln Phe Ser Gln Gln Pro Val Ala
740 745 750

Gly Ser Gln Thr Ser Ile Asp Gln Gln Val Asn Val His Met Ala Glu
755 760 765

His Tyr Gly Gln Gln Gln Gln Thr Arg Ala Pro His Leu Gln Leu Gln
770 775 780

Pro Arg Ala Gln Arg Val Val Lys Ala Ala Thr Ala Val Thr Ala Gly
785 790 795 800

Gly Ser Leu Leu Val Leu Ser Gly Leu Thr Leu Ala Gly Thr Val Ile
805 810 815

Ala Leu Thr Ile Ala Thr Pro Leu Leu Val Ile Phe Ser Pro Val Leu
820 825 830

Val Pro Ala Val Ile Thr Ile Phe Leu Leu Gly Ala Gly Phe Leu Ala
835 840 845

Ser Gly Gly Phe Gly Val Ala Ala Leu Ser Val Leu Ser Trp Ile Tyr
850 855 860

Arg Tyr Leu Thr Gly Lys His Pro Pro Gly Ala Asp Gln Leu Glu Ser
865 870 875 880

Ala Lys Thr Lys Leu Ala Ser Lys Ala Arg Glu Met Lys Asp Arg Ala
885 890 895

Glu Gln Phe Ser Gln Gln Pro Val Ala Gly Ser Gln Thr Ser Pro Trp
900 905 910

Leu Glu

<210> 78
<211> 153
<212> PRT
<213> Sesamum indicum/Synthetic

<400> 78

```
Met Ala Glu His Tyr Gly Gln Gln Gln Thr Arg Ala Pro His Leu
1               5                  10                  15

Gln Leu Gln Pro Arg Ala Gln Arg Val Val Lys Ala Ala Thr Ala Val
            20                  25                  30

Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Gly Leu Thr Leu Ala Gly
        35                  40                  45

Thr Val Ile Ala Leu Thr Ile Ala Thr Pro Leu Leu Val Ile Phe Ser
    50                  55                  60

Pro Val Leu Val Pro Ala Val Ile Thr Ile Phe Leu Leu Gly Ala Gly
65                  70                  75                  80

Phe Leu Ala Ser Gly Gly Phe Gly Val Ala Ala Leu Ser Val Leu Ser
                85                  90                  95

Trp Ile Tyr Arg Tyr Leu Thr Gly Lys His Pro Pro Gly Ala Asp Gln
            100                 105                 110

Leu Glu Ser Ala Lys Thr Lys Leu Ala Ser Lys Ala Arg Glu Met Lys
        115                 120                 125

Asp Arg Ala Glu Gln Phe Ser Gln Gln Pro Val Ala Gly Ser Gln Thr
    130                 135                 140

Ser Ile Asp Pro Ser Ser Trp Leu Glu
145                 150

<210>  79
<211>  455
<212>  PRT
<213>  Sesamum indicum/Synthetic

<400>  79

Met Ala Glu His Tyr Gly Gln Gln Gln Gln Thr Arg Ala Pro His Leu
1               5                  10                  15

Gln Leu Gln Pro Arg Ala Gln Arg Val Val Lys Ala Ala Thr Ala Val
            20                  25                  30

Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Gly Leu Thr Leu Ala Gly
        35                  40                  45

Thr Val Ile Ala Leu Thr Ile Ala Thr Pro Leu Leu Val Ile Phe Ser
    50                  55                  60

Pro Val Leu Val Pro Ala Val Ile Thr Ile Phe Leu Leu Gly Ala Gly
65                  70                  75                  80
```

```
Phe Leu Ala Ser Gly Gly Phe Gly Val Ala Ala Leu Ser Val Leu Ser
                85                  90                  95

Trp Ile Tyr Arg Tyr Leu Thr Gly Lys His Pro Pro Gly Ala Asp Gln
            100                 105                 110

Leu Glu Ser Ala Lys Thr Lys Leu Ala Ser Lys Ala Arg Glu Met Lys
        115                 120                 125

Asp Arg Ala Glu Gln Phe Ser Gln Gln Pro Val Ala Gly Ser Gln Thr
    130                 135                 140

Ser Ile Asp Pro Ser Ser Trp His Met Met Ala Glu His Tyr Gly Gln
145                 150                 155                 160

Gln Gln Gln Thr Arg Ala Pro His Leu Gln Leu Gln Pro Arg Ala Gln
                165                 170                 175

Arg Val Val Lys Ala Ala Thr Ala Val Thr Ala Gly Gly Ser Leu Leu
            180                 185                 190

Val Leu Ser Gly Leu Thr Leu Ala Gly Thr Val Ile Ala Leu Thr Ile
        195                 200                 205

Ala Thr Pro Leu Leu Val Ile Phe Ser Pro Val Leu Val Pro Ala Val
    210                 215                 220

Ile Thr Ile Phe Leu Leu Gly Ala Gly Phe Leu Ala Ser Gly Gly Phe
225                 230                 235                 240

Gly Val Ala Ala Leu Ser Val Leu Ser Trp Ile Tyr Arg Tyr Leu Thr
                245                 250                 255

Gly Lys His Pro Pro Gly Ala Asp Gln Leu Glu Ser Ala Lys Thr Lys
            260                 265                 270

Leu Ala Ser Lys Ala Arg Glu Met Lys Asp Arg Ala Glu Gln Phe Ser
        275                 280                 285

Gln Gln Pro Val Ala Gly Ser Gln Thr Ser Ile Asp Gln Gln Val Asn
    290                 295                 300

Val His Met Ala Glu His Tyr Gly Gln Gln Gln Gln Thr Arg Ala Pro
305                 310                 315                 320

His Leu Gln Leu Gln Pro Arg Ala Gln Arg Val Val Lys Ala Ala Thr
                325                 330                 335

Ala Val Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Gly Leu Thr Leu
            340                 345                 350
```

Ala Gly Thr Val Ile Ala Leu Thr Ile Ala Thr Pro Leu Leu Val Ile
355 360 365

Phe Ser Pro Val Leu Val Pro Ala Val Ile Thr Ile Phe Leu Leu Gly
370 375 380

Ala Gly Phe Leu Ala Ser Gly Gly Phe Gly Val Ala Ala Leu Ser Val
385 390 395 400

Leu Ser Trp Ile Tyr Arg Tyr Leu Thr Gly Lys His Pro Pro Gly Ala
405 410 415

Asp Gln Leu Glu Ser Ala Lys Thr Lys Leu Ala Ser Lys Ala Arg Glu
420 425 430

Met Lys Asp Arg Ala Glu Gln Phe Ser Gln Gln Pro Val Ala Gly Ser
435 440 445

Gln Thr Ser Pro Trp Leu Glu
450 455

<210> 80
<211> 602
<212> PRT
<213> Sesamum indicum/Synthetic

<400> 80

Met Ala Glu His Tyr Gly Gln Gln Gln Gln Thr Arg Ala Pro His Leu
1 5 10 15

Gln Leu Gln Pro Arg Ala Gln Arg Val Val Lys Ala Ala Thr Ala Val
20 25 30

Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Gly Leu Thr Leu Ala Gly
35 40 45

Thr Val Ile Ala Leu Thr Ile Ala Thr Pro Leu Leu Val Ile Phe Ser
50 55 60

Pro Val Leu Val Pro Ala Val Ile Thr Ile Phe Leu Leu Gly Ala Gly
65 70 75 80

Phe Leu Ala Ser Gly Gly Phe Gly Val Ala Ala Leu Ser Val Leu Ser
85 90 95

Trp Ile Tyr Arg Tyr Leu Thr Gly Lys His Pro Pro Gly Ala Asp Gln
100 105 110

Leu Glu Ser Ala Lys Thr Lys Leu Ala Ser Lys Ala Arg Glu Met Lys
115 120 125

Asp Arg Ala Glu Gln Phe Ser Gln Gln Pro Val Ala Gly Ser Gln Thr

130 135 140

Ser Ile Asp Pro Ser Ser Trp Leu Glu Met Ala Glu His Tyr Gly Gln
145 150 155 160

Gln Gln Gln Thr Arg Ala Pro His Leu Gln Leu Gln Pro Arg Ala Gln
165 170 175

Arg Val Val Lys Ala Ala Thr Ala Val Thr Ala Gly Gly Ser Leu Leu
180 185 190

Val Leu Ser Gly Leu Thr Leu Ala Gly Thr Val Ile Ala Leu Thr Ile
195 200 205

Ala Thr Pro Leu Leu Val Ile Phe Ser Pro Val Leu Val Pro Ala Val
210 215 220

Ile Thr Ile Phe Leu Leu Gly Ala Gly Phe Leu Ala Ser Gly Gly Phe
225 230 235 240

Gly Val Ala Ala Leu Ser Val Leu Ser Trp Ile Tyr Arg Tyr Leu Thr
245 250 255

Gly Lys His Pro Pro Gly Ala Asp Gln Leu Glu Ser Ala Lys Thr Lys
260 265 270

Leu Ala Ser Lys Ala Arg Glu Met Lys Asp Arg Ala Glu Gln Phe Ser
275 280 285

Gln Gln Pro Val Ala Gly Ser Gln Thr Ser His Met Met Ala Glu His
290 295 300

Tyr Gly Gln Gln Gln Gln Thr Arg Ala Pro His Leu Gln Leu Gln Pro
305 310 315 320

Arg Ala Gln Arg Val Val Lys Ala Ala Thr Ala Val Thr Ala Gly Gly
325 330 335

Ser Leu Leu Val Leu Ser Gly Leu Thr Leu Ala Gly Thr Val Ile Ala
340 345 350

Leu Thr Ile Ala Thr Pro Leu Leu Val Ile Phe Ser Pro Val Leu Val
355 360 365

Pro Ala Val Ile Thr Ile Phe Leu Leu Gly Ala Gly Phe Leu Ala Ser
370 375 380

Gly Gly Phe Gly Val Ala Ala Leu Ser Val Leu Ser Trp Ile Tyr Arg
385 390 395 400

Tyr Leu Thr Gly Lys His Pro Pro Gly Ala Asp Gln Leu Glu Ser Ala

405 410 415

Lys Thr Lys Leu Ala Ser Lys Ala Arg Glu Met Lys Asp Arg Ala Glu
420 425 430

Gln Phe Ser Gln Gln Pro Val Ala Gly Ser Gln Thr Ser Ile Asp Gln
435 440 445

Gln Val Asn Val His Met Ala Glu His Tyr Gly Gln Gln Gln Gln Thr
450 455 460

Arg Ala Pro His Leu Gln Leu Gln Pro Arg Ala Gln Arg Val Val Lys
465 470 475 480

Ala Ala Thr Ala Val Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Gly
485 490 495

Leu Thr Leu Ala Gly Thr Val Ile Ala Leu Thr Ile Ala Thr Pro Leu
500 505 510

Leu Val Ile Phe Ser Pro Val Leu Val Pro Ala Val Ile Thr Ile Phe
515 520 525

Leu Leu Gly Ala Gly Phe Leu Ala Ser Gly Gly Phe Gly Val Ala Ala
530 535 540

Leu Ser Val Leu Ser Trp Ile Tyr Arg Tyr Leu Thr Gly Lys His Pro
545 550 555 560

Pro Gly Ala Asp Gln Leu Glu Ser Ala Lys Thr Lys Leu Ala Ser Lys
565 570 575

Ala Arg Glu Met Lys Asp Arg Ala Glu Gln Phe Ser Gln Gln Pro Val
580 585 590

Ala Gly Ser Gln Thr Ser Pro Trp Leu Glu
595 600

<210> 81
<211> 914
<212> PRT
<213> Sesamum indicum/Synthetic

<400> 81

Met Ala Glu His Tyr Gly Gln Gln Gln Gln Thr Arg Ala Pro His Leu
1 5 10 15

Gln Leu Gln Pro Arg Ala Gln Arg Val Val Lys Ala Ala Thr Ala Val
20 25 30

Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Gly Leu Thr Leu Ala Gly
35 40 45

```
Thr Val Ile Ala Leu Thr Ile Ala Thr Pro Leu Leu Val Ile Phe Ser
    50                  55                  60

Pro Val Leu Val Pro Ala Val Ile Thr Ile Phe Leu Leu Gly Ala Gly
65                  70                  75                  80

Phe Leu Ala Ser Gly Gly Phe Gly Val Ala Ala Leu Ser Val Leu Ser
                85                  90                  95

Trp Ile Tyr Arg Tyr Leu Thr Gly Lys His Pro Pro Gly Ala Asp Gln
            100                 105                 110

Leu Glu Ser Ala Lys Thr Lys Leu Ala Ser Lys Ala Arg Glu Met Lys
        115                 120                 125

Asp Arg Ala Glu Gln Phe Ser Gln Gln Pro Val Ala Gly Ser Gln Thr
    130                 135                 140

Ser Ile Asp Pro Ser Ser Trp Leu Glu Met Ala Glu His Tyr Gly Gln
145                 150                 155                 160

Gln Gln Gln Thr Arg Ala Pro His Leu Gln Leu Gln Pro Arg Ala Gln
                165                 170                 175

Arg Val Val Lys Ala Ala Thr Ala Val Thr Ala Gly Gly Ser Leu Leu
            180                 185                 190

Val Leu Ser Gly Leu Thr Leu Ala Gly Thr Val Ile Ala Leu Thr Ile
        195                 200                 205

Ala Thr Pro Leu Leu Val Ile Phe Ser Pro Val Leu Val Pro Ala Val
    210                 215                 220

Ile Thr Ile Phe Leu Leu Gly Ala Gly Phe Leu Ala Ser Gly Gly Phe
225                 230                 235                 240

Gly Val Ala Ala Leu Ser Val Leu Ser Trp Ile Tyr Arg Tyr Leu Thr
                245                 250                 255

Gly Lys His Pro Pro Gly Ala Asp Gln Leu Glu Ser Ala Lys Thr Lys
            260                 265                 270

Leu Ala Ser Lys Ala Arg Glu Met Lys Asp Arg Ala Glu Gln Phe Ser
        275                 280                 285

Gln Gln Pro Val Ala Gly Ser Gln Thr Ser His Met Phe Lys Trp Pro
    290                 295                 300

Ser Ala Met Ala Glu His Tyr Gly Gln Gln Gln Gln Thr Arg Ala Pro
305                 310                 315                 320
```

His Leu Gln Leu Gln Pro Arg Ala Gln Arg Val Val Lys Ala Ala Thr
325 330 335

Ala Val Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Gly Leu Thr Leu
340 345 350

Ala Gly Thr Val Ile Ala Leu Thr Ile Ala Thr Pro Leu Leu Val Ile
355 360 365

Phe Ser Pro Val Leu Val Pro Ala Val Ile Thr Ile Phe Leu Leu Gly
370 375 380

Ala Gly Phe Leu Ala Ser Gly Gly Phe Gly Val Ala Ala Leu Ser Val
385 390 395 400

Leu Ser Trp Ile Tyr Arg Tyr Leu Thr Gly Lys His Pro Pro Gly Ala
405 410 415

Asp Gln Leu Glu Ser Ala Lys Thr Lys Leu Ala Ser Lys Ala Arg Glu
420 425 430

Met Lys Asp Arg Ala Glu Gln Phe Ser Gln Gln Pro Val Ala Gly Ser
435 440 445

Gln Thr Ser Ser Ser Glu Leu Pro Trp Val Asp Met Ala Glu His Tyr
450 455 460

Gly Gln Gln Gln Gln Thr Arg Ala Pro His Leu Gln Leu Gln Pro Arg
465 470 475 480

Ala Gln Arg Val Val Lys Ala Ala Thr Ala Val Thr Ala Gly Gly Ser
485 490 495

Leu Leu Val Leu Ser Gly Leu Thr Leu Ala Gly Thr Val Ile Ala Leu
500 505 510

Thr Ile Ala Thr Pro Leu Leu Val Ile Phe Ser Pro Val Leu Val Pro
515 520 525

Ala Val Ile Thr Ile Phe Leu Leu Gly Ala Gly Phe Leu Ala Ser Gly
530 535 540

Gly Phe Gly Val Ala Ala Leu Ser Val Leu Ser Trp Ile Tyr Arg Tyr
545 550 555 560

Leu Thr Gly Lys His Pro Pro Gly Ala Asp Gln Leu Glu Ser Ala Lys
565 570 575

Thr Lys Leu Ala Ser Lys Ala Arg Glu Met Lys Asp Arg Ala Glu Gln
580 585 590

Phe Ser Gln Gln Pro Val Ala Gly Ser Gln Thr Ser Ser Thr Pro Ser
595 600 605

Ser Trp His Met Met Ala Glu His Tyr Gly Gln Gln Gln Thr Arg
610 615 620

Ala Pro His Leu Gln Leu Gln Pro Arg Ala Gln Arg Val Val Lys Ala
625 630 635 640

Ala Thr Ala Val Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Gly Leu
645 650 655

Thr Leu Ala Gly Thr Val Ile Ala Leu Thr Ile Ala Thr Pro Leu Leu
660 665 670

Val Ile Phe Ser Pro Val Leu Val Pro Ala Val Ile Thr Ile Phe Leu
675 680 685

Leu Gly Ala Gly Phe Leu Ala Ser Gly Gly Phe Gly Val Ala Ala Leu
690 695 700

Ser Val Leu Ser Trp Ile Tyr Arg Tyr Leu Thr Gly Lys His Pro Pro
705 710 715 720

Gly Ala Asp Gln Leu Glu Ser Ala Lys Thr Lys Leu Ala Ser Lys Ala
725 730 735

Arg Glu Met Lys Asp Arg Ala Glu Gln Phe Ser Gln Gln Pro Val Ala
740 745 750

Gly Ser Gln Thr Ser Ile Asp Gln Gln Val Asn Val His Met Ala Glu
755 760 765

His Tyr Gly Gln Gln Gln Gln Thr Arg Ala Pro His Leu Gln Leu Gln
770 775 780

Pro Arg Ala Gln Arg Val Val Lys Ala Ala Thr Ala Val Thr Ala Gly
785 790 795 800

Gly Ser Leu Leu Val Leu Ser Gly Leu Thr Leu Ala Gly Thr Val Ile
805 810 815

Ala Leu Thr Ile Ala Thr Pro Leu Leu Val Ile Phe Ser Pro Val Leu
820 825 830

Val Pro Ala Val Ile Thr Ile Phe Leu Leu Gly Ala Gly Phe Leu Ala
835 840 845

Ser Gly Gly Phe Gly Val Ala Ala Leu Ser Val Leu Ser Trp Ile Tyr
850 855 860

```
Arg Tyr Leu Thr Gly Lys His Pro Pro Gly Ala Asp Gln Leu Glu Ser
865                 870                 875                 880

Ala Lys Thr Lys Leu Ala Ser Lys Ala Arg Glu Met Lys Asp Arg Ala
                885                 890                 895

Glu Gln Phe Ser Gln Gln Pro Val Ala Gly Ser Gln Thr Ser Pro Trp
            900                 905                 910

Leu Glu
```

<210> 82
<211> 441
<212> PRT
<213> Sesamum indicum/Synthetic

<400> 82

```
Met Ala Glu His Tyr Gly Gln Gln Gln Thr Arg Ala Pro His Leu
1               5                   10                  15

Gln Leu Gln Pro Arg Ala Gln Arg Val Val Lys Ala Ala Thr Ala Val
            20                  25                  30

Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Gly Leu Thr Leu Ala Gly
        35                  40                  45

Thr Val Ile Ala Leu Thr Ile Ala Thr Pro Leu Leu Val Ile Phe Ser
    50                  55                  60

Pro Val Leu Val Pro Ala Val Ile Thr Ile Phe Leu Leu Gly Ala Gly
65                  70                  75                  80

Phe Leu Ala Ser Gly Gly Phe Gly Val Ala Ala Leu Ser Val Leu Ser
                85                  90                  95

Trp Ile Tyr Arg Tyr Leu Thr Gly Lys His Pro Pro Gly Ala Asp Gln
            100                 105                 110

Leu Glu Ser Ala Lys Thr Lys Leu Ala Ser Lys Ala Arg Glu Met Lys
        115                 120                 125

Asp Arg Ala Glu Gln Phe Ser Gln Gln Pro Val Gly Gly Gly Gly Ser
    130                 135                 140

Gly Gly Gly Gly Ser Met Ala Glu His Tyr Gly Gln Gln Gln Gln Thr
145                 150                 155                 160

Arg Ala Pro His Leu Gln Leu Gln Pro Arg Ala Gln Arg Val Val Lys
                165                 170                 175
```

```
Ala Ala Thr Ala Val Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Gly
            180                 185                 190

Leu Thr Leu Ala Gly Thr Val Ile Ala Leu Thr Ile Ala Thr Pro Leu
        195                 200                 205

Leu Val Ile Phe Ser Pro Val Leu Val Pro Ala Val Ile Thr Ile Phe
    210                 215                 220

Leu Leu Gly Ala Gly Phe Leu Ala Ser Gly Gly Phe Gly Val Ala Ala
225                 230                 235                 240

Leu Ser Val Leu Ser Trp Ile Tyr Arg Tyr Leu Thr Gly Lys His Pro
                245                 250                 255

Pro Gly Ala Asp Gln Leu Glu Ser Ala Lys Thr Lys Leu Ala Ser Lys
            260                 265                 270

Ala Arg Glu Met Lys Asp Arg Ala Glu Gln Phe Ser Gln Gln Pro Val
        275                 280                 285

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Met Ala Glu His Tyr Gly
    290                 295                 300

Gln Gln Gln Gln Thr Arg Ala Pro His Leu Gln Leu Gln Pro Arg Ala
305                 310                 315                 320

Gln Arg Val Val Lys Ala Ala Thr Ala Val Thr Ala Gly Gly Ser Leu
                325                 330                 335

Leu Val Leu Ser Gly Leu Thr Leu Ala Gly Thr Val Ile Ala Leu Thr
            340                 345                 350

Ile Ala Thr Pro Leu Leu Val Ile Phe Ser Pro Val Leu Val Pro Ala
        355                 360                 365

Val Ile Thr Ile Phe Leu Leu Gly Ala Gly Phe Leu Ala Ser Gly Gly
    370                 375                 380

Phe Gly Val Ala Ala Leu Ser Val Leu Ser Trp Ile Tyr Arg Tyr Leu
385                 390                 395                 400

Thr Gly Lys His Pro Pro Gly Ala Asp Gln Leu Glu Ser Ala Lys Thr
                405                 410                 415

Lys Leu Ala Ser Lys Ala Arg Glu Met Lys Asp Arg Ala Glu Gln Phe
            420                 425                 430

Ser Gln Gln Pro Val Pro Trp Leu Glu
        435                 440
```

<210> 83
<211> 441
<212> PRT
<213> Sesamum indicum/Synthetic

<400> 83

Met Ala Glu His Tyr Gly Gln Gln Gln Gln Thr Arg Ala Pro His Leu
1 5 10 15

Gln Leu Gln Pro Arg Ala Gln Arg Val Val Lys Ala Ala Thr Ala Val
20 25 30

Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Gly Leu Thr Leu Ala Gly
35 40 45

Thr Val Ile Ala Leu Thr Ile Ala Thr Pro Leu Leu Val Ile Phe Ser
50 55 60

Pro Val Leu Val Pro Ala Val Ile Thr Ile Phe Leu Leu Gly Ala Gly
65 70 75 80

Phe Leu Ala Ser Gly Gly Phe Gly Val Ala Ala Leu Ser Val Leu Ser
85 90 95

Trp Ile Tyr Arg Tyr Leu Thr Gly Lys His Pro Pro Gly Ala Asp Gln
100 105 110

Leu Glu Ser Ala Lys Thr Lys Leu Ala Ser Lys Ala Arg Glu Met Lys
115 120 125

Asp Arg Ala Glu Gln Phe Ser Gln Gln Pro Val Gly Gly Gly Gly Ser
130 135 140

Gly Gly Gly Gly Ser Met Ala Glu His Tyr Gly Gln Gln Gln Gln Thr
145 150 155 160

Arg Ala Pro His Leu Gln Leu Gln Pro Arg Ala Gln Arg Val Val Lys
165 170 175

Ala Ala Thr Ala Val Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Gly
180 185 190

Leu Thr Leu Ala Gly Thr Val Ile Ala Leu Thr Ile Ala Thr Pro Leu
195 200 205

Leu Val Ile Phe Ser Pro Val Leu Val Pro Ala Val Ile Thr Ile Phe
210 215 220

Leu Leu Gly Ala Gly Phe Leu Ala Ser Gly Gly Phe Gly Val Ala Ala
225 230 235 240

```
Leu Ser Val Leu Ser Trp Ile Tyr Arg Tyr Leu Thr Gly Lys His Pro
                245                 250                 255

Pro Gly Ala Asp Gln Leu Glu Ser Ala Lys Thr Lys Leu Ala Ser Lys
            260                 265                 270

Ala Arg Glu Met Lys Asp Arg Ala Glu Gln Phe Ser Gln Gln Pro Val
        275                 280                 285

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Met Ala Glu His Tyr Gly
    290                 295                 300

Gln Gln Gln Gln Thr Arg Ala Pro His Leu Gln Leu Gln Pro Arg Ala
305                 310                 315                 320

Gln Arg Val Val Lys Ala Ala Thr Ala Val Thr Ala Gly Gly Ser Leu
                325                 330                 335

Leu Val Leu Ser Gly Leu Thr Leu Ala Gly Thr Val Ile Ala Leu Thr
            340                 345                 350

Ile Ala Thr Pro Leu Leu Val Ile Phe Ser Pro Val Leu Val Pro Ala
        355                 360                 365

Val Ile Thr Ile Phe Leu Leu Gly Ala Gly Phe Leu Ala Ser Gly Gly
    370                 375                 380

Phe Gly Val Ala Ala Leu Ser Val Leu Ser Trp Ile Tyr Arg Tyr Leu
385                 390                 395                 400

Thr Gly Lys His Pro Pro Gly Ala Asp Gln Leu Glu Ser Ala Lys Thr
                405                 410                 415

Lys Leu Ala Ser Lys Ala Arg Glu Met Lys Asp Arg Ala Glu Gln Phe
            420                 425                 430

Ser Gln Gln Pro Val Pro Trp Leu Glu
        435                 440
```