Isolated Spinach Ribulose-1,5-Bisphosphate Carboxylase/Oxygenase Large Subunit ε-N-Methyltransferase and Method of Inactivating Ribulose-1,5-Bisphosphatase Carboxylase/Oxygenase Large Subunit ε-N-Methyltransferase Activity

Robert L. Houtz
University of Kentucky, rhoutz@uky.edu

Follow this and additional works at: https://uknowledge.uky.edu/horticulture_patents
Part of the Horticulture Commons
Right click to open a feedback form in a new tab to let us know how this document benefits you.

Recommended Citation
Houtz, Robert L., "Isolated Spinach Ribulose-1,5-Bisphosphate Carboxylase/Oxygenase Large Subunit ε-N-Methyltransferase and Method of Inactivating Ribulose-1,5-Bisphosphatase Carboxylase/Oxygenase Large Subunit ε-N-Methyltransferase Activity" (1999). Horticulture Faculty Patents. 4. https://uknowledge.uky.edu/horticulture_patents/4

This Patent is brought to you for free and open access by the Horticulture at UKnowledge. It has been accepted for inclusion in Horticulture Faculty Patents by an authorized administrator of UKnowledge. For more information, please contact UKnowledge@lsv.uky.edu.
ISOLATED SPINACH RIBULOSE-1,5-BISPHOSPHATE CARBOXYLASE/OXYGENASE LARGE SUBUNIT ‘N’-METHYLTRANSFERASE AND METHOD OF INACTIVATING RIBULOSE-1,5-BISPHOSPHATE CARBOXYLASE/OXYGENASE LARGE SUBUNIT ‘N’-METHYLTRANSFERASE ACTIVITY

Inventor: Robert L. Houtz, Lexington, Ky.

Assignee: University of Kentucky Research Foundation, Lexington, Ky.

Filed: Jul. 29, 1996

Related U.S. Application Data

Continuation-in-part of application No. 08/391,000, Feb. 21, 1995, Pat. No. 5,723,752.

References Cited

PUBLICATIONS


Eckes et al., Isolation and characterization of a light-inducible, organ-specific gene from potato and analysis of its expression after tagging and transfer into tobacco and potato shoot, Mol Gen Genet vol. 205, pp. 14–22 (1986).


Primary Examiner—Robert A. Wax
Assistant Examiner—Ickchand Saidha
Attorney, Agent, or Firm—Burns, Doane, Swecker & Mathis, L.L.P.

ABSTRACT

The gene sequence for ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) large subunit (LS) ‘N’-methyltransferase (protein methylase III or Rubisco LSMT) from a plant which has a des(methyl) lysyl residue in the LS is disclosed. In addition, the full-length cDNA clones for Rubisco LSMT are disclosed. Transgenic plants and methods of producing same which have the Rubisco LSMT gene inserted into the DNA are also provided. Further, methods of inactivating the enzymatic activity of Rubisco LSMT are also disclosed.

21 Claims, 15 Drawing Sheets
FIG. 2B
FIG. 2D
FIG. 2E
SF-11
AGAAGGTGTTGCAACAGATTGATGAGGAATTCAAAGAAGAGAGATGGGATTTTG
KVLLQQIDDEEFKEREEMELGGY 2911
KVLLQQIDDEEFKEREEMELGGY 468

SR-2*
ACGAATACCTACCAAGAACCGGAGGCCTAAGGATCTTGGATTTGGCAGGGGCACAGGGAGAGA 2971
EYQERRRLKDGLAGAQGEK 488
EYQERRRLKDGLAGAQGEK 484

AACTACCCCTGGAATAGGAGAGGTCTAATTATTTATAGAACACTTTTTCTACTTCTTGTTTCTT
LPWIGEV 3031
LPWIGEV 495

TACTTCACTTCCTCCTCAGAAATTACTATTTATCTGAAATTGTAGAACAATAGTGAT
3091

SR-1
GATTTTGCCTGTAATGGCTTCACTTGAAAGTGGAACACTACAAATCGAAACTpoly(A)20 3144

FIG. 2F
<table>
<thead>
<tr>
<th>Protein</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPINACH40</td>
<td>MATTLFTLIPS - SNSTFLNPFFKTQHSHKLHFATPSPTFKNPLSIRCRRPETDTPPEIQKFWGW</td>
</tr>
<tr>
<td>SPINACH38</td>
<td>MATTLFTLIPS - SNSTFLNPFFKTQHSHKLHFATPSPTFKNPLSIRCRRPETDTPPEIQKFWGW</td>
</tr>
<tr>
<td>PEA</td>
<td>MATIFSGG - - SVSPFLFHTNKGSFTPKAPILHLKRFSASKVSASVGEPSLSPAVQTFWKW</td>
</tr>
<tr>
<td>TOBACCO</td>
<td>MASVFSVHPLPSSFLCPLKTKSKRTKHHQTYTQKTLINSQLTEDFDPKIPQPQVQTFWQW</td>
</tr>
<tr>
<td>SPINACH40</td>
<td>LSDKGIISSPKCPVKGIPVEGLGLVAQKDISRNEVVELTVQKFWINPDVAASEIGSVCNGLK</td>
</tr>
<tr>
<td>SPINACH38</td>
<td>LSDKGIISSPKCPVKGIPVEGLGLVAQKDISRNEVVELTVQKFWINPDVAASEIGSVCNGLK</td>
</tr>
<tr>
<td>PEA</td>
<td>LQEEGVIKTAKTPVKSSTVTEGLVALKISRNDVIQLPKRLWPVDAAASEIGVRVCSELK</td>
</tr>
<tr>
<td>TOBACCO</td>
<td>LQKEGVVTTTKPQVKGIPVEGLVAQRDIAKGETVLPQKRFWINPDAAASEIGNSCSGLK</td>
</tr>
<tr>
<td>SPINACH40</td>
<td>PWVSVALFLREKKLGNSSWKPYIDILPDSTNISTYWSEEELSEQGLNTTLGVKELVA</td>
</tr>
<tr>
<td>SPINACH38</td>
<td>PWVSVALFLREKKLGNSSWKPYIDILPDSTNISTYWSEEELSEQGLNTTLGVKELVA</td>
</tr>
<tr>
<td>PEA</td>
<td>PWLSVILFLIRER - SREDSVWKHYFGILPQETDSTYWSEEELSEQGLKTTVSKEYVK</td>
</tr>
<tr>
<td>TOBACCO</td>
<td>PWIVSLAFLREK - WRDDSKWKKYMVLPKSTDSTYWSEEELSEQGTQLLSTTSVKSVDYQ</td>
</tr>
</tbody>
</table>

**FIG. 3A**
FIG. 3C
Fig. 6

1 2 3 4 5

- 775 bases
- 457 bases
- 306 bases
FIG. 7B

Activity (pmoles CH$_3$-min$^{-1}$-mg protein$^{-1}$)

S-40  P-55  P-55-174  S-38

Rubisco LSMT Construct
ISOLATED SPINACH RIBULOSE-1,5-BISPHOSPHATE CARBOXYLASE/OXYGENASE LARGE SUBUNIT ‘N’ METHYLTRANSFERASE AND METHOD OF INACTIVATING RIBULOSE-1,5-BISPHOSPHATE CARBOXYLASE/OXYGENASE LARGE SUBUNIT ‘N’ METHYLTRANSFERASE ACTIVITY

RELATED APPLICATIONS

This application is continuation-in-part of U.S. patent application Ser. No. 08/391,000, filed on Feb. 21, 1995, now U.S. Pat. No. 5,723,752, which is hereby incorporated by reference in its entirety.

ACKNOWLEDGEMENT OF GOVERNMENT SUPPORT

This invention was made with Government support under Grant No. DE-FG05-92ER20075, awarded by the Department of Energy. The Government may have certain rights in this invention.

BACKGROUND OF THE INVENTION

1. Field of the Invention

This invention relates to ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) large subunit (LS) ‘N’-methyltransferase (protein methylase III or Rubisco LSMT). This enzyme catalyzes methylation of the ε-amino of lysine-14 in the large subunit of Rubisco. Many plant species contain methylated Lys-14 in the LS of Rubisco but some do not (i.e., a des(methyl) lysyl residue in the LS). In addition, the present invention relates to a gene and full-length cDNA clones for Rubisco LSMT. The present invention further relates to transgenic plants and methods of producing same which have the Rubisco LSMT gene inserted into the DNA. This invention also relates to a four amino acid insert (WVQQ) which inactivates the enzymatic activity of Rubisco LSMT and thereby accounts for the subsequent absence of trimethyllysine-14 in the LS of Rubisco.

2. Description of the Related Art

transgenic tobacco plants,” Proc. Natl. Acad. Sci. USA 89:8394—8398 (1992). These studies have been inconclusive as to the exact biological role of methylation of the ε-amino of protein bound lysyl residues.


Rubisco LSMT has been affinity purified ~8000-fold from pea chloroplasts and identified as a monomeric protein with a molecular mass of ~57 kDa (Wang et al., “Affinity Purification of Ribulose-1,5-bisphosphate Carboxylase/Oxygenase Large Subunit ‘N-Methyltransferase,’ accepted by Protein Expression and Purification (1995)). Recently, Rubisco LSMT cDNAs have been cloned and sequenced from pea and tobacco (Klein et al., “Cloning and developmental expression of pea ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit N-methyltransferase,” Plant Molecular Biol. 27:249—261 (1995); Ying et al., “Organization and characterization of the ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit ‘N-methyltransferase gene in tobacco,” Plant Molecular Biology (In press)). The deduced amino acid sequences of tobacco Rubisco LSMT has 64.5% identity and 75.3% similarity with the sequence of pea Rubisco LSMT, and both proteins contain several copies of a conserved imperfect leucine-rich repeat motifs (Ying et al., “Organization and characterization of the ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit ‘N-methyltransferase gene in tobacco,” Plant Molecular Biology (In press)).

Rubisco LSMT has high specific specificity, methyllating only Rubisco and only lysyl residue 14 in the LS. Of many plant species examined several contain methylated Lys-14 in the LS of Rubisco, such as pea and tobacco, but some do not, such as spinach and alfalfa (Houtz et al., “Post-translational modifications in the large subunit of ribulose bisphosphate carboxylase/oxygenase,” Proc. Natl. Acad. Sci. USA 86:1855—1859 (1989); Houtz et al., “Post-translational modifications in the amino-terminal region of the large subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase from several plant species,” “Plant Physiol. 98:1170—1174 (1992); and unpublished data). There has been no explanation for the existence of Lys-14 in the LS of Rubisco in a non-methylated state (i.e., a des(methyl) lysyl residue in the LS). Further, since some plant species, such as spinach, wheat, corn (maize) and lettuce do not contain methylated Lys-14 in the LS of Rubisco (Houtz et al., “Post-translational modifications in the amino-terminal region of the large subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase from several plant species,” “Plant Physiol. 98:1170—1174 (1992); and unpublished data), it was once assumed that these same plant species did not possess the Rubisco LSMT gene.

SUMMARY OF THE INVENTION

In view of the state of the art as previously described, there thus exists a need in the art for a better understanding of post-translational protein methylation in plants. More specifically, a better understanding for the molecular basis for the absence of trimethylation-14 in the LS of Rubisco from certain plant species.

It is thus an object of the present invention to provide a Rubisco LSMT gene.

It is a further object of the present invention to provide the DNA and amino acid sequence for a Rubisco LSMT enzyme.

It is a still further object of the present invention to provide full-length cDNA clones for Rubisco LSMT.

In a first aspect, the present invention relates to a Rubisco LSMT gene which exists in a higher plant with a des(methyl) lysyl residue in the LS of Rubisco. A particularly preferred higher plant includes the spinach plant.

In a second aspect, the present invention relates to the DNA and amino acid sequence for a Rubisco LSMT enzyme.

In a third aspect, the present invention relates to a recombinant vector including the Rubisco LSMT gene described above. The vector is suitable for transforming higher plants.

In a fourth aspect, the present invention relates to an isolated or recombinant Rubisco LSMT enzyme encoded by the Rubisco LSMT gene described above.

In a fifth aspect, the present invention relates to a recombinant or transgenic plant transformed with the Rubisco LSMT gene described above.

In a sixth aspect, the present invention relates to a method of inactivating Rubisco LSMT activity which comprises inserting a 4 amino acid sequence (SEQ ID NO:1) insert (WVQQ) into Rubisco LSMT.

In a further aspect, the present invention relates to a method for preventing or reducing Rubisco LSMT activity in a photosynthesizing plant comprising transforming a
photosynthesizing plant with a recombinant vector wherein the vector comprises a Rubisco LSMT gene with the 12 nucleotide insert.

With the foregoing and other objects, advantages and features of the invention that will become hereinafter apparent, the nature of the invention may be more clearly understood by reference to the following detailed description of the preferred embodiments of the invention and to the appended claims.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1A illustrates the genomic organization and restriction map of rbcMT-S. Exons are shown as heavy black bars, introns as horizontal lines, and the axon is indicated by an arrow.

FIG. 1B is a diagrammatic representation of the S38 and S40 cDNAs with coding regions as heavy black bars, untranslated regions as open bars and the axon as a shaded bar.

FIG. 1C shows Probe I, which is a 1056-bp SfiI fragment with the 12-bp axon, and Probe II, which is a riboprobe for the RNAase protection assay which results in only one 775-bp fragment protected by S40 mRNA, and two 500-bp and 457-bp fragments protected by S38 mRNA.

FIG. 1D depicts the strategy for PCR cloning and joining different cDNA fragments. The restriction enzymes labeled with stars were used to ligate corresponding fragments. Abbreviations for restriction sites: B, BglII; E, EcoRI; S, SacI; Sc, Scal; Sf, SfiI; Sp, SphI and X, XbaI.

FIGS. 2A, 2B, 2C, 2D, 2E and 2F show the nucleotide sequence of the rbcMT-S and the corresponding deduced amino acid sequences. Introns are printed in lower case letters and exons in upper case letters. The putative start and stop codons are underlined. The 12 nucleotides and corresponding 4 amino acids representing the axon sequence are indicated by bold italic letters. The deduced polypeptide for the S38 cDNA is underneath the one for the S40 cDNA that contains the axon. The oligonucleotide primers for sequencing, PCR and RACE, are indicated by arrows above the nucleotide sequence. The primers labeled with a star are derived from the conserved regions of pea and tobacco Rubisco LSMTs.

FIGS. 3A, 3B and 3C are a comparison of the deduced amino acid sequences of S38, S40, with tobacco and pea Rubisco LSMTs. Identical residues are indicated by vertical lines and similar residues by colons. Gaps introduced to maximize alignment are indicated by dashes. Potential N-glycosylation sites are shown in bold. Leucine-rich repeat-like motifs are underlined. The four amino acid repeat sequence, WVRW, deduced from the 12-bp axon is shown in bold italic letters. The conserved peptide sequences, from which the primers are derived to clone the rbcMT-S, are indicated by arrows.

FIG. 4 (SEQ ID NOS.: 26-30) illustrates alternative splicing of intron M of rbcMT-S mRNA. The top portion shows the sequence of intron III and flanking regions. Shown below are the two types of mRNAs (S40 and S38) produced by alternative splicing. When the second 3'splice site is utilized, the 12-bp axon is retained to produce S40 mRNA (center), which encodes a 55.5 kD polypeptide. If the first 3'splice site is utilized, the axon is absent and S38 mRNA is produced (bottom), which encodes a 55.0 kD polypeptide.

FIG. 5 is an analysis of the spinach genomic DNA. An aliquot of 20 μg of spinach genomic DNA was digested with Scal and EcoRI respectively, electrophoresed on a 0.7% agarose gel and processed for DNA gel-blot analysis by hybridization to the rbcMT-S cDNA probe labeled with digoxigenin-UTP. A rbcMT-S cDNA clone in BlueScript II KS(+) digested with EcoRI corresponding to one copy was used for copy number reconstitution.

FIG. 6 shows expression of both S38 and S40 mRNA in spinach leaves. RNase protection assays using a 785-bp antisense riboprobe designed to protect a 775-bp of the S40 mRNA from nt-455 to nt-1229, and a 306-bp and 457-bp of the S38 mRNA from nt-455 to nt-760 and from nt-761 to nt-1217 respectively, were carried out. Lanes 1, 2, 3, 4 and 5 are 2.5, 5, 10, 20 and 20 μg of spinach leaf total RNA. After hybridization all but lane 5 were digested with 1:100 dilution of RNases. Lane 5 was digested with a 1:5 dilution of RNases (Ambion).

FIG. 7A is a Western blot analysis of S-40, S-38, P-55 and P-55-174 mRNAs expressed in C. coli. Lane 1, standard markers; lanes 2 and 3, S-40; lanes 4 and 5, P-55; lanes 6 and 7, P-55-174; lanes 8 and 9, S-38; lanes 2, 4, 6 and 8, soluble protein; lanes 3, 5, 7 and 9, insoluble protein.

FIG. 7B is a bar graph representing Rubisco LSMT activity from the different constructs corresponding to the lanes in FIG. 7A.

DETAILED DESCRIPTION OF THE INVENTION

The present invention relates to a Rubisco LSMT gene, its DNA and amino acid sequence encoding therefor, cDNA clones therefor, and a four amino acid sequence insert which inactivates the enzymatic activity of Rubisco LSMT.

In the present application, naturally occurring amino acid residues in peptides are abbreviated as recommended by the IUPAC IUB Biochemical Nomenclature Commission as follows: Phenyalanine is Phe or F; Leucine is Leu or L; Isoleucine is Ile or I; Methionine is Met or M; Norleucine is Nle; Valine is Val or V; Serine is Ser or S; Proline is Pro or P; Threonine is Thr or T; Alanine is Ala or A; Tyrosine is Tyr of Y; Histidine is His or H; Glutamine is Gln or Q; Asparagine is Asn or N; Lysine is Lys or K; Aspartic Acid is Asp or D; Glutamic Acid is Gln or E; Cysteine is Cys or C; Tryptophan is Trp or W; Arginine is Arg or R; Glycine is Gly or G, and X is any amino acid.

Synthetic or non-naturally occurring amino acids refer to amino acids which do not naturally occur in vivo but which, nevertheless, can be incorporated into the peptide structures described herein. Preferred synthetic amino acids are the D-amino acids of naturally occurring L-amino acids as well as non-naturally occurring D and L amino acids represented by the formula HNC(NH)COOH, wherein R is: (1) a lower alkyl group; (2) a cycloalkyl group of from 3 to 7 carbon atoms; (3) a heterocycle of from 3 to 7 carbon atoms and 1 to 2 heteroatoms selected from the group consisting of oxygen, sulfur, and nitrogen; (4) an aromatic or alkylalkyl residue of from 6 to 15 carbon atoms optionally having from 1 to 3 substituents on the aromatic nucleus selected from the group consisting of hydroxy, lower alkoxy, amino, and carboxyl; (5) alkylene-Y where alkylene is an alkylene group of from 1 to 7 carbon atoms and Y is selected from the group consisting of hydroxy, amino, cycloalkyl of from 3 to 7 carbon atoms, heterocyclic of from 3 to 7 carbon atoms and 1 to 2 heteroatoms selected from the group consisting of oxygen, sulfur and nitrogen, and --COR where R is selected from the group consisting of hydrogen, lower alkyl, lower alkoxy, and --NR2R4 where R2 and R4 are independently selected from the group consisting of hydrogen and...
lower alkyl; (6) alkylene-S(O)nR[^5] where n is 1 or 2, and R[^5] is a lower alkyl or lower alkylene.

Particularly preferred synthetic amino acids include, by way of example, the D-amino acids of naturally occurring L-amino acids, L-1-naphthylalanine, L-2-naphthylalanine, L-cyclohexylalanine, L-2-amino isobutyric acid, the sulfoxide and sulfone derivatives of methionine, and the lower alkyl derivatives of methionine. “Peptide mimetics” are also encompassed by the present invention and include peptides having one or more of the following modifications:

- peptides wherein one or more of the peptidyl [—(O)—NH—] linkages (bonds) have been replaced by a non-peptidyl linkage such as carbamate linkage [—(O)CON—], phosphate linkage, amide linkage, sulfonamide linkage, and secondary amine linkage or with an alkylated peptidyl linkage [C(ONR—)] where R[^5] is a lower alkyl.
- peptides wherein the N-terminus is derivatized to a —NR[^6]R[^6] group, to a —N(CO)R[^6] group where R[^6] and R[^6] are independently selected from hydrogen and lower alcohols with the proviso that R[^6] and R[^6] are both not hydrogen, to a succinimide group, to a benzoxycarbonyl-NH—(CBZ-NH—) group, to a benzoylcarbonyl-NH— group having from 1 to 3 substituents on the phenyl ring selected from the group consisting of lower alkyl, lower alchohol, chloro, and bromo,

Other abbreviations are as follows: aa, amino acid(s); auxon, auxiliary exon; bp, base pair(s); nt, nucleotide(s); Rubisco LSMT, Ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit N-methyltransferase; RACE, rapid amplification of cDNA ends; RT-PCR, reverse transcription-polymerase chain reaction.

Although the present invention is described with respect to spinach, it will be appreciated that the techniques employed herein are applicable to other plants species which contain a des(methyl) form of Rubisco with regards to trimethylation of lysyl residue 14 in the large subunit (LS). Examples of such plant species include alfalfa, wheat, corn (maize) and lettuce.

Ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) large subunit (LS) N-methyltransferase (referred to herein as “Rubisco LSMT”) catalyzes methylation of the ε-amine of lysine-14 in the LS of Rubisco. Rubisco is the world’s most abundant protein, and serves as the only significant link between the inorganic and organic carbon pools in the Earth’s biosphere by catalyzing the reduction of atmospheric carbon dioxide to carbohydrates during photosynthesis. Perturbations of Rubisco activity translate directly into similar changes in plant growth and yield. Thus, there is significant interest in the art in the potential manipulation and control of Rubisco activity through genetic engineering.

However, the complexity and multimeric nature of Rubisco have proven to be substantial obstacles to achieving this goal, which have not yet been overcome. Rubisco LSMT provides an opportunity for the selective manipulation of Rubisco activity through changes in the structure and stability of the N-terminal region in the LS, an area known to be essential for catalytic activity. Rubisco LSMT is a highly specific enzyme which is found to interact only with Rubisco and does not interact with any other protein in the plant cell. Since Rubisco catalyzes the reduction of atmospheric CO₂ during photosynthesis, Rubisco and Rubisco LSMT are critical to the plant cell for viability. Furthermore, the exceptionally tight and specific nature of the interaction between Rubisco LSMT and des(methyl) forms of Rubisco creates the possibility for the development of novel synthetic polypeptide herbicides, whose target is the in vivo interaction between Rubisco LSMT and Rubisco, whose specificity crosses a group of plant species related only by the presence of Rubisco LSMT, and whose target protein has no homologue in the entire animal kingdom. Finally, this same affinity of Rubisco LSMT for des(methyl) forms of Rubisco also creates the possibility for the site and protein specific delivery of compounds into the chloroplast and to Rubisco, for the potential manipulation of Rubisco activity and/or stability.

Ribulose bisphosphate carboxylase/oxygenase (Rubisco) from spinach (Spinacia oleracea) is a des(methyl) form of Rubisco with regards to trimethylation of lysyl residue 14 in the large subunit (LS). In investigating the molecular basis for the absence of trimethylation-14 in the LS of spinach Rubisco, the inventor has isolated and sequenced two full-length cDNAs (S40 and S38) and the gene for spinach Rubisco LSMT (rbcMT-S). This discovery was quite unexpected since it was once thought that spinach did not possess the Rubisco LSMT gene because it contained a des(methyl) lysyl residue in the LS of Rubisco. The gene for spinach Rubisco LSMT, covering all 6 exons and 5 introns, has an organization similar to the tobacco Rubisco LSMT gene (rbcMT-T). Southern blot analysis of spinach genomic DNA shows that the rbcMT-S is present as a single copy. The reduced amino acid sequence from the rbcMT-S cDNAs shows 60% and 62% identity with the amino acid sequences of pea and tobacco Rubisco LSMT, respectively.

Moreover, the particular sequence disclosed herein for the spinach Rubisco LSMT gene may be used to determine the particular sequence in other photosynthesizing plants. The sequence of the gene may be used as a probe to screen cDNA or genomic DNA libraries from other plants and, due to the expected homology between the gene sequences in the various plant species, the particular sequence for the Rubisco LSMT gene in other species may then be found.

In a further aspect, the present invention relates to a recombinant or transgenic plant transformed with the Rubisco LSMT gene described above. The methods employed for transforming the plants are generally known in the art. For example, the transformation method described in Bechtold et al, *Plant Agrobacterium Mediated Gene Transfer By Infiltration of Adult Arabidopsis Thaliana Plants*, C.R. Acad. Sci., Paris 316:1194–1199 (1993) and Valvekens et al, “Agrobacterium tumefaciens-mediated transformation of Arabidopsis thaliana root explants by using kanamycin selection,” *Proc. Natl. Acad. Sci.* USA 85:5536–5540 (1988), may be used in the method of the present invention.

To achieve the present invention, a full-length cDNA clone was isolated by the present inventor utilizing polymerase chain reaction (PCR)-based technology and conventional bacteriophage library screening. PCR techniques are disclosed, for example, in Klein et al, “Cloning and Developmental Expression of the Sucrose-Phosphate-Synthase Gene From Spinach,” *Planta* 190:498-510 (1993); in Ampli-Taq PCR kit by Perkin Elmer-Cetus, Emeryville, Calif.); and in the manufacturer’s instruction manual. Bacteriophage library screening is described, for example, in Gantt et al, “Transfer of rpl22 to the Nucleus Greatly Proceeded its loss from the Chloroplast and Involved the Gain of an Intron,” *EMBO J.* 10:3073–3078 (1991), and in the information provided by the manufacturer of the screening membrane (Stratagene, La Jolla, Calif.).
A cDNA of the Rubisco LSMT gene from spinach was isolated and studies of Rubisco LSMT gene expression initiated. Utilizing amino acid sequence information derived from purified peptic polypeptide fragments from proteolyzed Rubisco LSMT, a full-length cDNA of Rubisco LSMT was obtained. The cDNA of Rubisco LSMT, rbcMT, was used to examine organ-specific and developmental parameters affecting rbcMT gene expression.

The present specification details the purification of peptic fragments from spinach Rubisco LSMT and a PCR-based cloning strategy for isolating a full-length cDNA. A similar strategy was previously utilized to obtain a full-length cDNA of sucrose-phosphate synthase from spinach (Klein et al., “Cloning and developmental expression of the sucrose-phosphate-synthase gene from spinach,” *Planta* 190:498–510 (1993)) and to obtain the cDNA of the Rubisco LSMT gene from pea and from tobacco. The sequence information obtained from peptic fragments permitted the confirmation of clones encoding for Rubisco LSMT. Hence, a molecular probe of the spinach Rubisco LSMT gene was rapidly obtained thereby permitting identification of protein and nucleotide sequence and characterization of its gene expression.

This amino acid sequence deduced from the S40 cDNA, as described in the Examples and in FIGS. 2, 3 and 4, contains a 4-amino acid (SEQ ID NO.: 1) insert (WVQQ) located near the center of the protein, which is a consequence of alternative 3′mRNA splicing and inclusion of 12 nucleotides from the 3′end of intron III. For example, the 4-amino acid sequence was determined to be a 12 nucleotide (SEQ ID NO.: 2) WGAFGRRRSA, 5′CGATGGGCACTTTGGGAATCTCTAGATCAAGGGC. The SR-2 sequence (SEQ ID NO.: 5), including a BglIII site and encoding the peptide (SEQ ID NO.: 5) ERRKLKDGLGL, is 5′GGCCAAAGCCACAGACCTTTAAGCTTCTTCTTCC. Conditions for PCR were 35 cycles of: 94°C 1 min, 50°C 1 min, 72°C 1.5 min and final extension 72°C 10 min. The PCR product was digested with EcoRI and BglIII, and gel-purified. The purified fragment was cloned into Bluescript II KS+(+) vector for sequencing. For sequencing, two different clones were identified, one with a 12bp auxiliary exon (auxon) designated as S40 and another without the auxon designated as S38.

For 3RACE, reverse-transcription was the same as described above except for using an rbcMT-S-specific primer (SR-3, FIG. 2) anchored in the mid-coding region and followed by poly d(C)-tailing as described in Ying et al., “Isolation and characterization of auxon, a *Knesus laevis* ortholog of the chicken nov gene,” *Gene* 171:243–248 (1996). The resulting dC-tailed products were amplified using a nested primer (SR-5) which included a XbaI restriction site, and a poly(dG/dC)-containing oligonucleotide (SEQ ID NO.: 7) (AP-2, 5′GGCCTAGCTCCTAGCTGGCGGIGGGIGGGGG, SacI). The PCR products were digested with ScaI and XbaI, gel-purified and cloned into Bluescript II KS+(+) vector for sequencing. After sequencing, two different clones were identified, one with a 12bp auxiliary exon (auxon) designated as S40' and another without the auxon designated as S38'.

For 3RACE, 5 μg of total RNA from spinach leaves was reverse-transcribed with an adapter-primer (AP-1, 5′GGCCACCGTGCTGGAGTACTG) Amplification by PCR was as described above except for using the AP-1 and spinach specific primer (SEQ ID NO.: 8) (SF-9). The PCR product was cloned into pCR-Script Direct SK(+) vector (Stratagene) for sequencing, designated as S2′ (FIG. 1D).

Two to five independent clones were chosen for sequencing from each of the above constructs. Both strands of each clone were sequenced by the dideoxy chain termination method (Sanger et al., “DNA sequencing with chain-terminating inhibitors,” *Proc. Natl. Acad. Sci. USA* 74:5463–5467 (1977)) using Sequenase (US Biochemical) and 3′S-dATP (NEN) with M13 reverse and 40mers. In
addition, 18 to 27-mer oligonucleotides synthesized according to sequence information obtained were used directly as primers for further sequencing.

Both full-length S38 and S40 cDNAs were obtained by ligation of clones S2' and S25' to S38' and S40', accordingly, based on restriction sites within the overlapped regions (FIG. 1D).

Example 3

Isolation and Southern Analysis of the rbMT-S

The rbMT-S gene was cloned by PCR. Spinach nuclear DNA was isolated using Floraclean (Bio101, Inc.). Approximately 100 ng of the nuclear DNA was amplified by PCR with Taq polymerase (Gibco/BRL) using a forward primer (SF-1) and a reverse primer (SR-1). The PCR product was cloned into PCR-Script SK(++) for sequencing and restriction mapping.

For Southern analysis, spinach nuclear DNA was digested with EcoRI or SacI, electrophoresed on a 0.7% agarose gel and transferred onto nylon membranes (MSI) (Sambrook et al., Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Lab., Cold Spring Harbor, N.Y.), 2nd Ed. (1989)). The DNA blot was hybridized with the cDNA probe I (SfiI fragment, 1056-bp long, FIG. 1C) labeled with digoxigenin-UTP according to the procedure provided by the manufacturer (Boehringer Mannheim).

Example 4

Genetic Engineering of the (12-bp) Auxon into the Pea LSMT

A 5'end-truncated pea LSMT cDNA cloned in pET23d (P-55) (Cheng and Houtz, unpublished data) was digested with KpnI which generated a 802-bp fragment I and a 4300-bp fragment H which were gel-purified. The purified 802-bp fragment was self-ligated and then amplified by Taq polymerase with a forward primer (SEQ ID NO.: 9) (P-R, 5'AGT CCC GGG TGC AAC AGA TTA ACC ACA GTC CAG GAG TTA C, Smal). Note: 12 nucleotides, including one in the reverse primer, are in bold italic letters and consist of the auxon) and a reverse primer (SEQ ID NO.: 10) (P-R, 5'AGT TTT AAA GGT CTT CCA TGG AAA CAC C, Dral) at 35 cycles of: 94°C, 1 min, 56°C, 1 min, 72°C, 4 sec and 72°C, 10 min. The PCR product was digested with Smal and Dral, and self-religated. The circular DNA was digested with KpnI, ligated into KpnI-fragment I, and transformed into DH5α cells (BRL/GIBCO). After screening 180 colonies, two of them (designated as P-55-84, and P-55-174) were selected for sequencing to confirm that the 12-bp auxon was engineered into the P-55 and no other mutation was caused by PCR. The full-length encoding regions of S40 and S38 cDNA were also cloned into the pET-23d E. coli expression vector (designated as S-40 and S-38 respectively).

Example 5

RNase Protection Assay

The antisense riboprobe (probe II) was made by transcribing a rbMT-S cDNA clone 210-1 (which contained a 775-bp EcoRI-Sacl fragment with the 12-bp auxon and was linearized by EcoRI, FIG. 1C) with T7 RNA polymerase, (α-32)PUTP (800 Ci/mmol, 10 mCi/ml) and cold NTP. Probe III generated a 775-nucleotide was fully protected by the S40 mRNA but only partially protected by the S38 mRNA. The 2.5, 5, 10, 20 and 20 μg of total RNA isolated from spinach leaves were hybridized with 1x10⁵ cpm of the probe II according to the manufacturer’s instructions (Ambion).

Example 6

Rubisco LSMT Activity Assay and Western Blot Analysis

Individual clones (S-40, S-38, P-55 and P-55-174) in pLysS host cells were cultured at 37°C. For 3.5 hrs in 5 mL LB broth with 50 μg/ml carbenicillin and 35 μg/ml chloramphenicol and induced by the addition of IPTG to the growing cells at a final concentration of 0.5 mM. After induction cell cultures were continued for 2.5 hrs at 35°C. After induction the cells were harvested by centrifugation at 5000 x g for 5 min at 4°C, washed twice with deionized water, and resuspended in 100 μl of buffer (50 mM TRIS-KCl, pH 8.2, 5 mM MgCl₂, 1 mM EDTA) with protease inhibitors (1 mM PMSF, 10 μg/ml leupeptin) and frozen at -80°C. The activity of Rubisco LSMT was determined as described previously (Wang et al., “Affinity purification of ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit N-methyltransferase,” Prot. Expr. Pur. 6:528–536 (1995)).

For Western analysis protein extracts prepared as described above were separated by SDS-PAGE (15% acrylamide) and transferred to PVDF-membranes (Millipore Corp.). The membranes were probed with antibody raised against the precursor form of pea Rubisco LSMT expressed in E. coli.

Example 7

Isolation of rbMT-S cDNA

The high homology between pea and tobacco Rubisco LSMT enabled the inventor to design Rubisco LSMT-specific primers for amplifying a 786-bp fragment (S25, FIG. 1D) from a spinach first-strand cDNA pool reverse-transcribed from total RNA isolated from spinach leaves. Cloning and sequencing of the 786-bp fragment showed that it was a truncated rbMT-S cDNA which lacked 5' and 3'ends. The remaining 5' and 3' sequences were obtained by 5' and 3' RACE, respectively (for review see Forhman, “RACE: rapid amplification of cDNA ends,” In PCR protocols: a guide to methods and applications, pp 28–38, Innis et al., eds. Academic Press, San Diego (1990)).

For 5' RACE using an rbMT-S-specific primer (SR-3, FIG. 2) for first-strand cDNA synthesis and a second nested gene-specific antisense primer (SR-5, FIG. 2) for PCR amplification, resulted in the identification of two 5' RACE products (836-bp and 848-bp fragments) after sequencing, one with a 12-bp insertion designated as S40, and the other without the insertion designated as S38. In the region where the 5' RACE products and the PCR product (S25) have sequence in common, complete sequence identity was observed and 118-bp overlapped in the cDNA sequences excepting the 12-bp insertion in S40 (FIG. 2).

For 3' RACE using an adapter-primer (AP-1) for first-strand cDNA synthesis and also as a reverse primer, and SR-9 as the rbMT-S-specific primer for PCR amplification, a single 761-bp PCR product was obtained. Sequence analysis confirmed the identity of the 3' RACE product as encoding the predicted 3' portion of the rbMT-S protein including the 3' untranslated region (FIG. 1D, FIG. 2). Given these overlapping clones, the inventor was able to assemble the two cDNA sequences (S40 and S38) of the rbMT-S as shown in the FIG. 1B and FIG. 2.

Both rbMT-S cDNAs contain a 5' leader of 31-nt and encode for proteins of 495-aa (S40) and 491-aa (S38) with predicted molecular mass of 55.5 kD for S40 and 55.0 kD for S38, which are similar to that of pea (55.0 kD) and tobacco (56.0 kD) (FIG. 3). The deduced rbMT-S proteins contain four potential N-linked glycosylation site which fit.
the consensus sequence Asn-Xaa-Ser/Thr (NXS/T), one of which is conserved in the pea and tobacco Rubisco LSMTs (FIG. 3), and like that of pea and tobacco, they also contain five imperfect copies of a motif similar to leucine-rich repeats (LLR) (FIG. 3) (Kobe et al., “The leucine-rich repeat: a versatile binding motif,” Trends Biochem. Sci., 19:415–21 (1994)).

Example 8

Characterization of rbcMT-S.

The rbcMT-S covering the entire coding region was cloned and sequenced in the overall length of 3144-bp (FIG. 2). Comparison of the genomic DNA and cDNA sequences allowed the precise location of the six exons and five introns to be mapped (FIG. 1A). It has the similar genomic organization of the tobacco Rubisco LSMT gene (rbcMT-T). The size of the exons is fairly constant while that of the introns is quite variable. Intron III of rbcMT-S occurs at a position corresponding to the 12-bp insertion in the rbcMT-S S40 cDNA (FIG. 2). An identical 12-bp sequence was found to be present at the 3′end of the intron. Examination of the DNA sequence of this intron and flanking regions suggested that either of two 3′splice sites (separated by the 12-bp sequence) is utilized during splicing of the rbcMT-S transcripts. Thus, as illustrated in FIG. 4, when the intron III sequence is completely removed, S38 mRNA encoding a 55.0 kD polypeptide is produced. However, if splicing occurs at the alternative site, S40 mRNA that retains a 12-bp portion of the 3′end of the intron III is generated, and subsequently a 4-amino acid longer polypeptide of 55.5 kD is produced.

A sequence comparison between the rbcMT-S gene and a Drosophila trn gene (O’Neil et al., “Interspecific comparison of the transformer gene of Drosophila reveals an unusually high degree of evolutionary divergence,” Genetics 131:113–128 (1992)) which has been studied for alternative 3′splicing events (Mckean, “Alternative mRNA splicing,” Annu. Rev. Cell Biol. 8:133–155 (1992)) shows two striking TC-rich regions of primary sequence homology between these genes (SEQ ID NO: 11–14) (CTTTTTCTC and TCTTTTTCCCTGTTTCT for rbcMT-S, and TCTTTTTGTT and TTTTTTTCCCT for trn) in the region preceding the regulated splice site of both genes, and what is likely to be the regulated splice site of rbcMT-S.

Southern blot analysis suggests that the rbcMT-S is a single copy gene. FIG. 5 shows hybridization of probe I of a 32P-labeled rbcMT-S cDNA fragment (FIG. 1C) to spinach genomic DNA digested with EcoRI and Scal. Probe I detected a predicted major 2424-bp EcoRI fragment. Additionally, a predicted 876-bp and two other Scal fragments were also detected (FIG. 5). The intensity of the signals in each lane is equivalent to a single copy standard (Croy et al., “Plant Nucleic Acids,” In: Croy, R. R. D. (eds.) Plant Molecular Biology, pp. 21–48. BIOS Scientific Publishers Limited, Oxford (1999)) on the left side of the blot. Therefore, we conclude that rbcMT-S is a single copy gene in the spinach genome as rbcMT-T is in the tobacco genome.

Example 9

The rbcMT-S mRNA Present In Vivo and E. coli Expression in Vitro

To determine whether both S38 and S40 mRNA are present in the spinach leaves, total RNA from spinach leaves was subjected to an RNase protection analysis using probe II directed toward the middle region of both S38 and S40 mRNAs (FIG. 1C), where the auxon is present in S40 mRNA. Probe II was designed to protect a single fragment (775-nt) of S40 mRNA and two fragments (306-nt and 457-nt) of S38 mRNA. FIG. 6 shows that S38 mRNA is 20 fold more than S40 mRNA in spinach leaves based on quantitative analysis with a PhosphorImager 445SI (Molecular Dynamic). S40 mRNA is very low in abundance but detectable when high concentrations of total RNA are used. However, S38 and S40 mRNAs are undetectable in spinach roots, stems, and flowers by RNase protection assay (data not shown).

In vitro bacterial expression of the S40 cDNA (S-40) using a pET expression vector did yield a protein (FIG. 7A) at detectable levels but with undetectable activity (FIG. 7B). Furthermore, engineering of the 4 amino acid insert encoded by the 12-bp auxon into the corresponding position in pea Rubisco LSMT (P-55), and bacterial expression of the engineered pea Rubisco LSMT (P-55-174, FIG. 7A) demonstrated that the 4 amino acid insert resulted in complete inactivation of pea Rubisco LSMT activity (FIG. 7B). Therefore, alternative 3′mRNA splicing may result in the inactivation of S40 LSMT. Investigation of the mechanism for inactivation of S38 LSMT is still under way. For some unknown reason, bacterial expression of S38 cDNA (S-38) has been unsuccessful (FIG. 7A).

While the invention has been described and illustrated herein by references to various specific material, procedures and examples, it is understood that the invention is not restricted to the particular material, combinations of material, and procedures selected for that purpose. Numerous variations of such details can be implied and will be appreciated by those skilled in the art. Furthermore, all of the publications, patents and patent applications cited herein are incorporated by reference in their entirety.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(iii) NUMBER OF SEQUENCES: 30

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

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

Trp Val Gln Gln

1

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

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

TG66T5CACA AG

12

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

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

Trp Ala Phe Gly Ile Leu Arg Ser Arg Ala

1 5 10

(2) INFORMATION FOR SEQ ID NO:4:

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

(ii) MOLECULE TYPE: DNA (genomic)

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

CGATGCCGAT TGGGATCT CGATCAAGG GC

32

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

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

Glu Arg Arg Leu Lys Asp Leu Gly Leu Ala

1 5 10

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 31 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: DNA (genomic)

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

G GCCAAGGGC A AGATCTTT A AGCTCTTT C

31

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

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

GCTAGCTTC T AGACTGCG GGGGGGG

28

(2) INFORMATION FOR SEQ ID NO:8:

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

(ii) MOLECULE TYPE: DNA (genomic)

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

GCCACGC GTAGCTAC T

21

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 40 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

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

AGTCCC GGGTT GCAACAGATT AACCAGGTG CAGGAGTTAC

40

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

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

AGTTTTAAG GTCTCCATT GGAACCAC

28

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

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

CTTTTCTC
(2) INFORMATION FOR SEQ ID NO:12:

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

(ii) MOLECULE TYPE: DNA (genomic)

(x) SEQUENCE DESCRIPTION: SEQ ID NO:12:

TCTTTTCCCT TGTCTCT

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(x) SEQUENCE DESCRIPTION: SEQ ID NO:13:

TCTTTTTGTT

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 11 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(x) SEQUENCE DESCRIPTION: SEQ ID NO:14:

TCTTTTTCTC

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 3144 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

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

AAATCTTAAT CTCAGAATGA GTGAGCTAAA AATGCGAACT TTAATTCATC TCATCCTCCTC 60
ATCAAACTCT ACCTTTCCTA ACCCTTTCAA AACCACCCA CACTTCAAAC TTCTATCCGG 120
AACCCATCTT CCCCCTTCTA AAAACCCCTT CTCATTCAGA TTTTTCGTGG CACCCGAAAA 180
GGTACTACCA CGGAAAATCC AGAATATCTG GGGTGCGCTT TCCGACAAAG GAATATCTC 240
ACCAAAATGC CCTGTA AACG AGUTATTTG CCTGAAAGGA TTAGGTATAG TACGCCAAAA 300
AGATATACCC AGAAGACGAG TGCTTGGAGA GGTCCCCAG AGGTTTGGAG TAAACCGAA 360
TCAGATTTCA GGTCTCAGAG TTGGGTCAAG TTTAAATTGG CTTGAGCTCT GGGTTTCCTG 420
GCTTCTTGT CTGATGAGAG AGAAAAAATT GGGAATTTCT CTATCTTGGGA AACCCTATCAT 480
TGATATTTGG CCGTTAATCTA TCAATTCACG AAATATTTG ATATTTTTTT TGTTAAATT 540
GACTGGTTT AGTTCTTGGG TACGTATTAT GTTGGCAATC TTAATTTTTT AATTGTTGA 600
TTTAAGCTAA ATGAAAGTTG GGTGTTGCTT TCACAGAGGAG TCAAGAAGA GAACCTTCTG 660
AGCTTCAGG TTGTTCCTCA TTTTCGACTT AGATGTGCTT GTGATTATAC TCMTCCAAAG 720
TGCTGTGGTT TATGGCTGTT TTTTTTCGTT TCTTCTGAAG CCATTTATTT AGCTTCTTNG 780
TTGAACTTTA TTAACGCAGT ATTGAGTGCAT AGTTGCACCT CATTACTAAC TGAGATTGCGC 840
TACACGCTG TAAATGATGC CAGGAGAGTG CAGTGGTATG TACTAACTTC GCCAGATTCG 900
AGTTCGACA TTATATTTTAC ATTAACGTGC ATTGAGAAGCA TTGTCTGCTT TACAGAGGAGG 960
TACACGCTG TTTTATTTTG AGTTGAGAT TATTTTCTTG TGAGAGGAAA CTCGAAACTGA 1020
ATGCTGCTGT ACTAGTGCTG CGAGAGGTGA CACCTGCTTAT AACATCAA2A GAGAGGAGGT 1080
GCGTAGCAGA AATACCTGCTG CTTGATGTTCG AGCTTTTCTAC TCTATCTTCT CTCACATGG 1140
TTTACCGCTT GTATTATGAAG GAGAGGAAA AATACATGCA TACCTCCACG TATTTGCTAAAC 1200
TTTTGACCG TTGAACCTAA CAGGTCGCTT ATATAGAAAC TAATACTCCAA TTTCACATCG 1260
AAAAACCGG TACGTCTAGTA TCTGCATCGCA TTGGATGCTA CAGTITTTATG ATCTITTTGAA 1320
CTTTCGAGGC GTCTGCTTAA GGGTCACGGA TCTCTGTCTT TCAACAGCCT CTACACTAAA 1380
CTGCGGCTTT CATTTCTGAG ATTTCTGATT TCCAGACGG GTTGGACCTTG GAGGCAAGGA 1440
CGAATCTCAA CTTACATCAG CCAACCTCAAT GTCGATTATT ATGTCCTGTTG AGCTCTCAG 1500
AGTGTGCCGA TATGTTCTCG TTTGAAGATG AGAATGAGTG AGAATGATAG ATCCCTTTTT 1560
AATTTGCTTG TAGATGATTGT TTAATCTTGG AAATGCTTTG CCAATCTAAA CAATCTCTAT 1620
TTGCACGTGT CAGAAGAATA ACACCTAATG TACAGACGTA AGGGTTCTTT TTCTGCTTGAT 1680
TTTCAATTTT CCTGATCTCA CAGTTGACCTA TGAGACTATA AATCATCTGCAG GATATGCATT 1740
GCTGAGACCA ACATTTGGTG TGAAGGGATT AGTCAACAAAT GATTTTGGCA AGCTGAAAGA 1800
AGAATCTCAA GTTCCACGCA ACCAATCTTT CTTTTGTATG GTAACTCAAAG ATGACCTCTT 1860
TTGGACATTG GGAATCTGCG GACTCAAGAC ATTCACCTGT CTGGAAGGCC AAAGCTCTTG 1920
TCTAACACCC TTGGGCGGATT TGTAATATCAT CTTTTTCTCT GCCCTCTATT TCAATTATATAA 1980
AAAAAAAAC ATGACTTTTT TCTCATGTTA TCTGATTATA ATGATGAATA TTTAAATTAC 2040
ATGGAAAAAG GTGGCAACAG CTCTACCAAG CTCTGATATG AACAGCACCG AGATATCCCT 2100
GGAAATCAG AGAAGCTGCTGT CATATCTTCGA GAAACTTGGT ATTTCTACGT AGGAATCCTA 2160
CCCCATTGGA GCCTCTGTCTG CATGGTATTG TTTTCCCTCT GCTTCAGACA ATGAAATATA 2220
TATAACGTAC TTAAATTTTA TTAGAAGCTG TAGAAGCTGCA ATGAGAGATA TTGAAGATGC 2280
ATTACCTCAG ACATTTGTCT CCAATACATG TTGACACGAG GCAAATCAGGA AATACCCTTG 2340
GATTAGGTT GAAGCGAAGG CGATGATAGG AGAATAAGCAT ACACCGTAAAC ACTGGAATGA 2400
CCGGAATTG AGTTTTTATT TAGGGAGCG TAAATGACAT GCTGACAGGAA AGGAGGTTG 2460
GATAAGCTGT ACCTTTGATAT GCTTTTGAGGA CCACCATGTC CAAACGAAAT GCTGATCACAT 2520
TTGAAGCTGG TGGACTCTGGT GAGAGCTAGCT GTTTTTCTGG TGAAGCTCTAT ATCCAGGAAC 2580
TCTTACATATG GTATGCACCGT TCCTCTTATT GACCTTCGCA ATGAGGACGT CATATGTCCAA 2640
GTTTGCTCTG ATGCTTGTAC ATCTGCTCTT TTCTGTTTAC GACATCAATA ATCAGAGGATA 2700
ACACTATAG ATTTATTTACT CTTCTGCTTG TTATACAAAG AGTATGCTTG 2760
TTTTTTTTTAT TTATGTAGGA TGAGACGGTG TTAGAGAGAA AGAAGGCTGT TGCAACAGAT 2820
GGAAATCTTGA TTAATGACAG GTTGAAGAGAA GAAATGCTTG TGAAGATGAGA 2880
TTCAAAGAAA GAAATGACAG ATGCTTGTGT TGCAATACT ACCAAAGACG GAGGGTTTAG 2940
GATCTTGGAT TGCGGCGGCC AGAAGGAGAG ACAAATCCTTG GAATAGAGGA GGCTCAATTTA 3000
TTATTAGAAC ATTTTTTCTAC TTTTCTTCTT TACCTTCACT TCACTTCACT TCAAGAAATAC 3060
(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 163 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:
Met Ala Thr Leu Phe Thr Leu Ile Pro Ser Ser Asn Ser Thr Phe Leu
1   5   10   15
Asn Pro Phe Lys Thr Thr Gln His Ser Lys Leu His Phe Ala Thr Pro
20  25   30
Ser Pro Thr Phe Lys Asn Pro Leu Ser Ile Arg Cys Phe Arg Pro Pro
35  40   45
Glu Thr Asp Thr Pro Pro Glu Ile Gin Lys Phe Trp Gly Trp Leu Ser
50  55   60
Asp Lys Gly Ile Ile Ser Pro Lys Cys Pro Val Lys Pro Gly Ile Val
65  70   75   80
Pro Glu Gly Leu Gly Leu Val Ala Gin Lys Asp Ile Ser Arg Asn Glu
85  90   95
Val Val Leu Glu Val Pro Gin Lys Phe Trp Ile Asn Pro Asp Thr Val
100 105  110
Ala Ala Ser Glu Ile Gly Ser Val Cys Asn Gly Leu Lys Pro Trp Val
115 120  125
Ser Val Ala Leu Phe Leu Met Arg Glu Lys Leu Gly Asn Ser Ser
130 135  140
Ser Thr Lys Pro Tyr Ile Asp Ile Leu Pro Asp Ser Thr Asn Ser Thr
145 150  155  160
Ile Tyr Trp

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:
Ser Glu Glu Glu Leu Ser Glu Leu Gin Gly
1   5   10

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 78 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:
Ser Gin Leu Leu Asn Thr Thr Leu Gly Val Lys Glu Leu Val Ala Asn
(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 45 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

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

Trp Val Gln Gln Ala Asn His Ser Pro Asp Ile Thr Ala Pro Lys Tyr
1  2  3  4  5  6  7  8  9 10 11 12
Ala Trp Glu Ile Arg Gly Ala Gly Leu Phe Ser Arg Glu Leu Val Phe
13 14 15 16 17 18 19 20 21 22 23 24
Ser Leu Arg Asn Pro Thr Pro Val Lys Ala Gly Asp Gln
25 26 27 28 29 30 31 32 33 34 35 36

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 135 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

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

Val Leu Ile Gln Tyr Asp Leu Asn Lys Ser Asn Ala Glu Leu Ala Leu
1  2  3  4  5  6  7  8  9 10 11 12
Asp Tyr Gly Leu Thr Glu Ser Arg Ser Glu Arg Asn Ala Tyr Thr Leu
13 14 15 16 17 18 19 20 21 22 23 24
Thr Leu Glu Ile Pro Glu Ser Asp Ser Phe Tyr Gly Asp Lys Leu Asp
25 26 27 28 29 30 31 32 33 34 35 36
Ile Ala Glu Ser Asn Gly Met Gly Glu Ser Ala Tyr Phe Asp Ile Val
37 38 39 40 41 42 43 44 45 46 47 48
Leu Glu Gln Pro Leu Pro Ala Asn Met Leu Pro Tyr Leu Arg Leu Val
49 50 51 52 53 54 55 56 57 58 59 60
Ala Leu Gly Gly Glu Asp Ala Phe Leu Leu Glu Ser Ile Phe Arg Asn
61 62 63 64 65 66 67 68 69 70 71 72
Ser Ile Trp Gly His Leu Asp Leu Pro Ile Ser Pro Ala Asn Glu Glu
73 74 75 76 77 78 79 80 81 82 83 84
Leu Ile Cys Gln Val Ile Arg Asp Ala Cys Thr Ser Ala Leu Ser Gly
85 86 87 88 89 90 91 92 93 94 95 96
Tyr Ser Thr Thr Ile Ala Glu
97 98 99 100 101 102 103 104 105 106 107 108

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 73 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

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

```
Asp Glu Lys Lys Leu Leu Leu Ala Glu Gly Asp Ile Asp Pro Arg Leu Glu 1 5 10 15
Ile Ala Ile Thr Ile Arg Leu Gly Glu Lys Val Leu Gln Gln Ile 20 25 30
Asp Glu Glu Phe Lys Glu Arg Glu Met Glu Leu Gly Tyr Glu Tyr 35 40 45
Tyr Gln Glu Arg Arg Leu Lys Asp Leu Gly Ala Gly Ala Gln Gly 50 55 60
Glu Lys Leu Pro Trp Ile Gly Glu Val 65 70
```

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 495 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

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

```
Met Ala Thr Leu Phe Thr Leu Ile Pro Ser Ser Asn Ser Thr Phe Leu 1 5 10 15
Asn Pro Phe Lys Thr Thr His Ser Lys Leu His Phe Ala Thr Pro 20 25 30
Ser Pro Thr Phe Lys Asn Pro Leu Ser Ile Arg Cys Phe Arg Pro Pro 35 40 45
Glu Thr Asp Thr Pro Pro Glu Ile Gln Lys Phe Trp Gly Trp Leu Ser 50 55 60
Asp Lys Gly Ile Ile Ser Pro Pro Lys Pro Val Lys Pro Gly Ile Val 65 70 75 80
Pro Glu Gly Leu Gly Leu Val Ala Gln Lys Asp Ile Ser Arg Asn Glu 85 90 95
Val Val Leu Glu Val Pro Gln Lys Phe Trp Ile Asn Pro Asp Thr Val 100 105 110
Ala Ala Ser Glu Ile Gly Ser Val Cys Asn Gly Leu Lys Pro Trp Val 115 120 125
Ser Val Ala Leu Phe Leu Met Arg Glu Lys Leu Gly Asn Ser Ser 130 135 140
Ser Thr Lys Pro Tyr Ile Asp Ile Leu Pro Asp Ser Thr Asn Ser Thr 145 150 155 160
Ile Tyr Trp Ser Glu Glu Glu Leu Ser Glu Leu Gln Gly Ser Gln Leu 165 170 175
Leu Asn Thr Thr Leu Gly Val Lys Glu Leu Val Ala Asn Glu Phe Ala 180 185 190
Lys Leu Glu Glu Glu Val Leu Val Pro His Lys Gln Leu Phe Pro Phe 195 200 205
Asp Val Thr Gln Asp Phe Phe Trp Ala Phe Gly Met Leu Arg Ser 210 215 220
Arg Ala Phe Thr Cys Leu Glu Gly Gln Ser Leu Val Leu Ile Pro Leu
```
(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 491 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

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

Met Ala Thr Leu Phe Thr Leu Ile Pro Ser Ser Ser Ser Thr Phe Leu 1 5 10 15
Asn Pro Phe Lys Thr Thr Gln His Ser Lys Leu His Phe Ala Thr Pro 20 25 30
Ser Pro Thr Phe Lys Asn Pro Leu Ser Ile Arg Cys Phe Arg Pro Pro 35 40 45
Glu Thr Asp Thr Pro Glu Ile Gln Lys Phe Trp Gly Trp Leu Ser 50 55 60
Asp Lys Gly Ile Ile Ser Pro Lys Cys Pro Val Lys Pro Gly Ile Val 65 70 75 80
Pro Glu Gly Leu Gly Leu Val Ala Gln Lys Asp Ile Ser Arg Asn 85 90 95
Val Val Leu Glu Val Pro Gln Lys Phe Trp Ile Asn Pro Asp Thr Val
100 105 110
Ala Ala Ser Glu Ile Gly Ser Val Cys Asn Gly Leu Lys Pro Trp Val
115 120 125
Ser Val Ala Leu Phe Leu Met Arg Glu Lys Leu Gly Asn Ser Ser
130 135 140
Ser Trp Lys Pro Tyr Ile Asp Ile Leu Pro Asp Ser Thr Asn Ser Thr
145 150 155 160
Ile Tyr Trp Ser Glu Glu Leu Ser Glu Glu Gln Gly Ser Gln Leu
165 170 175
Leu Asn Thr Thr Leu Gly Val Lys Glu Leu Val Ala Asn Glu Phe Ala
180 185 190
Lys Leu Glu Glu Glu Val Leu Val Pro His Lys Glu Leu Phe Pro Phe
195 200 205
Asp Val Thr Gln Asp Asp Phe Phe Trp Ala Asp Gly Met Leu Arg Ser
210 215 220
Arg Ala Phe Thr Cys Leu Gly Glu Gln Ser Leu Val Leu Ile Pro Leu
225 230 235 240
Ala Asp Leu Ala Asn His Ser Pro Asp Ile Thr Ala Pro Lys Tyr Ala
245 250 255
Trp Glu Ile Arg Gly Ala Gly Leu Phe Ser Arg Glu Leu Val Phe Ser
260 265 270
Leu Arg Asn Pro Thr Pro Val Lys Ala Gly Asp Glu Val Leu Ile Gln
275 280 285
Tyr Asp Leu Asn Lys Ser Asn Ala Glu Leu Ala Leu Asp Tyr Gly Leu
290 295 300
Thr Glu Ser Arg Ser Glu Arg Asn Ala Tyr Thr Leu Thr Leu Glu Ile
305 310 315 320
Pro Glu Ser Asp Ser Phe Tyr Gly Asp Lys Leu Asp Ile Ala Ala Ser
325 330 335
Asn Gly Met Gly Glu Ser Ala Tyr Phe Asp Ile Val Leu Glu Glu Pro
340 345 350
Leu Pro Ala Asn Met Leu Pro Tyr Leu Arg Leu Val Ala Leu Gly Gly
355 360 365
Glu Asp Val Phe Leu Leu Glu Ser Ile Phe Arg Asn Ser Ile Trp Gly
370 375 380
His Leu Asp Leu Pro Ile Ser Pro Ala Asn Glu Leu Ile Cys Gln
385 390 395 400
Val Ile Arg Asp Ala Cys Thr Ser Ala Leu Ser Gly Tyr Ser Thr Thr
405 410 415
Ile Ala Glu Asp Glu Lys Leu Ala Glu Gly Asp Ile Asp Pro Arg
420 425 430
Leu Glu Ile Ala Ile Thr Arg Leu Gly Glu Lys Val Leu Gln
435 440 445
Gln Ile Asp Glu Phe Lys Glu Glu Arg Glu Met Glu Leu Gly Gly Gly
450 455 460
Glu Tyr Tyr Glu Arg Arg Leu Lys Asp Leu Gly Leu Ala Gly Glu
465 470 475 480
Gln Gly Glu Lys Leu Pro Trp Ile Gly Glu Val
485 490

(2) INFORMATION FOR SEQ ID NO:24:
(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 490 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

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

<table>
<thead>
<tr>
<th>1</th>
<th>5</th>
<th>10</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>25</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>40</td>
<td>45</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>55</td>
<td>60</td>
<td></td>
</tr>
<tr>
<td>65</td>
<td>70</td>
<td>75</td>
<td>80</td>
</tr>
<tr>
<td>95</td>
<td>90</td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>105</td>
<td>110</td>
<td></td>
</tr>
<tr>
<td>115</td>
<td>120</td>
<td>125</td>
<td></td>
</tr>
<tr>
<td>130</td>
<td>135</td>
<td>140</td>
<td></td>
</tr>
<tr>
<td>145</td>
<td>150</td>
<td>155</td>
<td>160</td>
</tr>
<tr>
<td>165</td>
<td>170</td>
<td>175</td>
<td></td>
</tr>
<tr>
<td>180</td>
<td>185</td>
<td>190</td>
<td></td>
</tr>
<tr>
<td>195</td>
<td>200</td>
<td>205</td>
<td></td>
</tr>
<tr>
<td>210</td>
<td>215</td>
<td>220</td>
<td></td>
</tr>
<tr>
<td>225</td>
<td>230</td>
<td>235</td>
<td>240</td>
</tr>
<tr>
<td>245</td>
<td>250</td>
<td>255</td>
<td></td>
</tr>
<tr>
<td>260</td>
<td>265</td>
<td>270</td>
<td></td>
</tr>
<tr>
<td>275</td>
<td>280</td>
<td>285</td>
<td></td>
</tr>
<tr>
<td>290</td>
<td>295</td>
<td>300</td>
<td></td>
</tr>
<tr>
<td>305</td>
<td>310</td>
<td>315</td>
<td>320</td>
</tr>
<tr>
<td>325</td>
<td>330</td>
<td>335</td>
<td></td>
</tr>
<tr>
<td>340</td>
<td>345</td>
<td>350</td>
<td></td>
</tr>
<tr>
<td>355</td>
<td>360</td>
<td>365</td>
<td></td>
</tr>
<tr>
<td>370</td>
<td>375</td>
<td>380</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Met</th>
<th>Ala</th>
<th>Thr</th>
<th>Ile</th>
<th>Phe</th>
<th>Ser</th>
<th>Gly</th>
<th>Ser</th>
<th>Val</th>
<th>Ser</th>
<th>Pro</th>
<th>Phe</th>
<th>Leu</th>
<th>Phe</th>
<th>His</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thr</td>
<td>Asn</td>
<td>Lys</td>
<td>Gly</td>
<td>Thr</td>
<td>Ser</td>
<td>Phe</td>
<td>Thr</td>
<td>Pro</td>
<td>Lys</td>
<td>Ala</td>
<td>Pro</td>
<td>Ile</td>
<td>Leu</td>
<td>His</td>
</tr>
<tr>
<td>Lys</td>
<td>Arg</td>
<td>Ser</td>
<td>Phe</td>
<td>Ser</td>
<td>Ala</td>
<td>Lys</td>
<td>Ser</td>
<td>Val</td>
<td>Ala</td>
<td>Ser</td>
<td>Val</td>
<td>Gly</td>
<td>Thr</td>
<td>Glu</td>
</tr>
<tr>
<td>Ser</td>
<td>Leu</td>
<td>Ser</td>
<td>Pro</td>
<td>Ala</td>
<td>Val</td>
<td>Gln</td>
<td>Thr</td>
<td>Phe</td>
<td>Trp</td>
<td>Lys</td>
<td>Trp</td>
<td>Leu</td>
<td>Gln</td>
<td>Glu</td>
</tr>
<tr>
<td>Gly</td>
<td>Val</td>
<td>Ile</td>
<td>Thr</td>
<td>Ala</td>
<td>Lys</td>
<td>Thr</td>
<td>Pro</td>
<td>Val</td>
<td>Lys</td>
<td>Ala</td>
<td>Ser</td>
<td>Val</td>
<td>Val</td>
<td>Thr</td>
</tr>
<tr>
<td>Gly</td>
<td>Leu</td>
<td>Gly</td>
<td>Leu</td>
<td>Val</td>
<td>Ala</td>
<td>Lys</td>
<td>Asp</td>
<td>Ile</td>
<td>Ser</td>
<td>Arg</td>
<td>Asn</td>
<td>Asp</td>
<td>Val</td>
<td>Ile</td>
</tr>
<tr>
<td>Leu</td>
<td>Gln</td>
<td>Val</td>
<td>Pro</td>
<td>Lys</td>
<td>Arg</td>
<td>Leu</td>
<td>Trp</td>
<td>Ile</td>
<td>Asn</td>
<td>Pro</td>
<td>Asp</td>
<td>Ala</td>
<td>Val</td>
<td>Ala</td>
</tr>
<tr>
<td>Ser</td>
<td>Glu</td>
<td>Ile</td>
<td>Gly</td>
<td>Arg</td>
<td>Val</td>
<td>Cys</td>
<td>Ser</td>
<td>Glu</td>
<td>Leu</td>
<td>Lys</td>
<td>Pro</td>
<td>Trp</td>
<td>Leu</td>
<td>Ser</td>
</tr>
<tr>
<td>Ile</td>
<td>Leu</td>
<td>Phe</td>
<td>Leu</td>
<td>Ile</td>
<td>Arg</td>
<td>Glu</td>
<td>Arg</td>
<td>Ser</td>
<td>Arg</td>
<td>Glu</td>
<td>Asp</td>
<td>Ser</td>
<td>Val</td>
<td>Trp</td>
</tr>
<tr>
<td>His</td>
<td>Tyr</td>
<td>Phe</td>
<td>Gly</td>
<td>Ile</td>
<td>Leu</td>
<td>Pro</td>
<td>Gln</td>
<td>Glu</td>
<td>Thr</td>
<td>Asp</td>
<td>Ser</td>
<td>Thr</td>
<td>Ile</td>
<td>Tyr</td>
</tr>
<tr>
<td>Ser</td>
<td>Glu</td>
<td>Glu</td>
<td>Leu</td>
<td>Gln</td>
<td>Glu</td>
<td>Leu</td>
<td>Gln</td>
<td>Gly</td>
<td>Ser</td>
<td>Gln</td>
<td>Leu</td>
<td>Lys</td>
<td>Thr</td>
<td></td>
</tr>
<tr>
<td>Thr</td>
<td>Val</td>
<td>Ser</td>
<td>Val</td>
<td>Lys</td>
<td>Glu</td>
<td>Tyr</td>
<td>Val</td>
<td>Lys</td>
<td>Asn</td>
<td>Glu</td>
<td>Cys</td>
<td>Leu</td>
<td>Lys</td>
<td>Glu</td>
</tr>
<tr>
<td>Gln</td>
<td>Glu</td>
<td>Ile</td>
<td>Leu</td>
<td>Pro</td>
<td>Asn</td>
<td>Lys</td>
<td>Arg</td>
<td>Leu</td>
<td>Phe</td>
<td>Pro</td>
<td>Asp</td>
<td>Pro</td>
<td>Val</td>
<td>Thr</td>
</tr>
<tr>
<td>Leu</td>
<td>Asp</td>
<td>Asp</td>
<td>Phe</td>
<td>Phe</td>
<td>Trp</td>
<td>Ala</td>
<td>Phe</td>
<td>Gly</td>
<td>Ile</td>
<td>Leu</td>
<td>Arg</td>
<td>Ser</td>
<td>Arg</td>
<td>Ala</td>
</tr>
<tr>
<td>Ser</td>
<td>Arg</td>
<td>Leu</td>
<td>Arg</td>
<td>Asn</td>
<td>Glu</td>
<td>Asn</td>
<td>Leu</td>
<td>Val</td>
<td>Val</td>
<td>Pro</td>
<td>Met</td>
<td>Ala</td>
<td>Asp</td>
<td>Leu</td>
</tr>
<tr>
<td>Ile</td>
<td>Asn</td>
<td>His</td>
<td>Ser</td>
<td>Ala</td>
<td>Gly</td>
<td>Val</td>
<td>Thr</td>
<td>Glu</td>
<td>Asp</td>
<td>His</td>
<td>Ala</td>
<td>Tyr</td>
<td>Glu</td>
<td>Val</td>
</tr>
<tr>
<td>Lys</td>
<td>Gly</td>
<td>Ala</td>
<td>Gly</td>
<td>Leu</td>
<td>Phe</td>
<td>Ser</td>
<td>Trp</td>
<td>Asp</td>
<td>Tyr</td>
<td>Leu</td>
<td>Phe</td>
<td>Ser</td>
<td>Leu</td>
<td>Lys</td>
</tr>
<tr>
<td>Ser</td>
<td>Pro</td>
<td>Leu</td>
<td>Ser</td>
<td>Val</td>
<td>Lys</td>
<td>Ala</td>
<td>Gly</td>
<td>Glu</td>
<td>Val</td>
<td>Tyr</td>
<td>Ile</td>
<td>Gln</td>
<td>Tyr</td>
<td>Asp</td>
</tr>
<tr>
<td>Leu</td>
<td>Asn</td>
<td>Lys</td>
<td>Ser</td>
<td>Asn</td>
<td>Ala</td>
<td>Glu</td>
<td>Ala</td>
<td>Leu</td>
<td>Asp</td>
<td>Tyr</td>
<td>Gly</td>
<td>Phe</td>
<td>Ile</td>
<td>Glu</td>
</tr>
<tr>
<td>Pro</td>
<td>Asn</td>
<td>Glu</td>
<td>Asn</td>
<td>Arg</td>
<td>His</td>
<td>Ala</td>
<td>Tyr</td>
<td>Thr</td>
<td>Leu</td>
<td>Thr</td>
<td>Leu</td>
<td>Glu</td>
<td>Ile</td>
<td>Ser</td>
</tr>
<tr>
<td>Ser</td>
<td>Asp</td>
<td>Pro</td>
<td>Phe</td>
<td>Phe</td>
<td>Asp</td>
<td>Asp</td>
<td>Lys</td>
<td>Leu</td>
<td>Asp</td>
<td>Val</td>
<td>Ala</td>
<td>Glu</td>
<td>Ser</td>
<td>Asn</td>
</tr>
<tr>
<td>Phe</td>
<td>Ala</td>
<td>Gin</td>
<td>Thr</td>
<td>Ala</td>
<td>Tyr</td>
<td>Phe</td>
<td>Asp</td>
<td>Ile</td>
<td>Phe</td>
<td>Tyr</td>
<td>Asn</td>
<td>Arg</td>
<td>Thr</td>
<td>Leu</td>
</tr>
<tr>
<td>Pro</td>
<td>Gly</td>
<td>Leu</td>
<td>Leu</td>
<td>Pro</td>
<td>Tyr</td>
<td>Leu</td>
<td>Arg</td>
<td>Leu</td>
<td>Val</td>
<td>Ala</td>
<td>Leu</td>
<td>Gly</td>
<td>Glu</td>
<td>Thr</td>
</tr>
<tr>
<td>Ala</td>
<td>Phe</td>
<td>Leu</td>
<td>Leu</td>
<td>Glu</td>
<td>Ser</td>
<td>Ile</td>
<td>Phe</td>
<td>Arg</td>
<td>Asn</td>
<td>Ser</td>
<td>Val</td>
<td>Trp</td>
<td>Gly</td>
<td>His</td>
</tr>
<tr>
<td>Gly</td>
<td>Leu</td>
<td>Pro</td>
<td>Val</td>
<td>Ser</td>
<td>Arg</td>
<td>Ala</td>
<td>Asn</td>
<td>Glu</td>
<td>Glu</td>
<td>Leu</td>
<td>Ile</td>
<td>Cys</td>
<td>Lys</td>
<td>Val</td>
</tr>
<tr>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>385</td>
<td>390</td>
<td>395</td>
<td>400</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Arg</th>
<th>Asp</th>
<th>Ala</th>
<th>Cys</th>
<th>Lys</th>
<th>Ser</th>
<th>Ala</th>
<th>Ser</th>
<th>Gly</th>
<th>Tyr</th>
<th>His</th>
<th>Thr</th>
<th>Thr</th>
<th>Ile</th>
<th>Glu</th>
</tr>
</thead>
<tbody>
<tr>
<td>405</td>
<td>410</td>
<td>415</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Glu</th>
<th>Asp</th>
<th>Gly</th>
<th>Lys</th>
<th>Leu</th>
<th>Met</th>
<th>Glu</th>
<th>Gly</th>
<th>Asn</th>
<th>Leu</th>
<th>Ser</th>
<th>Thr</th>
<th>Arg</th>
<th>Leu</th>
<th>Gln</th>
</tr>
</thead>
<tbody>
<tr>
<td>420</td>
<td>425</td>
<td>430</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ile</th>
<th>Ala</th>
<th>Val</th>
<th>Gly</th>
<th>Ile</th>
<th>Arg</th>
<th>Gly</th>
<th>Glu</th>
<th>Lys</th>
<th>Met</th>
<th>Val</th>
<th>Leu</th>
<th>Gln</th>
<th>Gln</th>
<th>Ile</th>
</tr>
</thead>
<tbody>
<tr>
<td>435</td>
<td>440</td>
<td>445</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Asp</th>
<th>Gly</th>
<th>Ile</th>
<th>Phe</th>
<th>Glu</th>
<th>Gln</th>
<th>Lys</th>
<th>Glu</th>
<th>Leu</th>
<th>Asp</th>
<th>Gln</th>
<th>Leu</th>
<th>Glu</th>
<th>Tyr</th>
</tr>
</thead>
<tbody>
<tr>
<td>450</td>
<td>455</td>
<td>460</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Tyr</th>
<th>Gln</th>
<th>Glu</th>
<th>Arg</th>
<th>Arg</th>
<th>Leu</th>
<th>Asp</th>
<th>Leu</th>
<th>Gly</th>
<th>Leu</th>
<th>Asp</th>
<th>Gly</th>
<th>Glu</th>
<th>Asn</th>
<th>Gln</th>
</tr>
</thead>
<tbody>
<tr>
<td>465</td>
<td>470</td>
<td>475</td>
<td>480</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Asp</th>
<th>Ile</th>
<th>Leu</th>
<th>Gly</th>
<th>Asp</th>
<th>Leu</th>
<th>Gly</th>
<th>Lys</th>
<th>Phe</th>
<th>Phe</th>
</tr>
</thead>
<tbody>
<tr>
<td>485</td>
<td>490</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 490 amino acids
   (B) TYPE: amino acid
   (C) STRANDNESS: single
   (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

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

<table>
<thead>
<tr>
<th>Met</th>
<th>Ala</th>
<th>Ser</th>
<th>Val</th>
<th>Phe</th>
<th>Ser</th>
<th>Val</th>
<th>His</th>
<th>Pro</th>
<th>Leu</th>
<th>Pro</th>
<th>Ser</th>
<th>Ser</th>
<th>Ser</th>
<th>Ser</th>
<th>Phe</th>
<th>Leu</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5</td>
<td>10</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cys</th>
<th>Pro</th>
<th>Leu</th>
<th>Lys</th>
<th>Thr</th>
<th>Thr</th>
<th>Lys</th>
<th>Ser</th>
<th>Arg</th>
<th>Thr</th>
<th>Lys</th>
<th>His</th>
<th>His</th>
<th>Gln</th>
<th>Thr</th>
<th>Phe</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>25</td>
<td>30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Tyr</th>
<th>Thr</th>
<th>Tyr</th>
<th>Gln</th>
<th>Lys</th>
<th>Thr</th>
<th>Ile</th>
<th>Leu</th>
<th>Ile</th>
<th>Asn</th>
<th>Ser</th>
<th>Leu</th>
<th>Gln</th>
<th>Leu</th>
<th>Thr</th>
<th>Gln</th>
</tr>
</thead>
<tbody>
<tr>
<td>35</td>
<td>40</td>
<td>45</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Leu</th>
<th>Asp</th>
<th>Pro</th>
<th>Lys</th>
<th>Ile</th>
<th>Pro</th>
<th>Gln</th>
<th>Pro</th>
<th>Val</th>
<th>Gln</th>
<th>Thr</th>
<th>Phe</th>
<th>Trp</th>
<th>Gln</th>
<th>Trp</th>
<th>Leu</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>55</td>
<td>60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cys</th>
<th>Lys</th>
<th>Glu</th>
<th>Gly</th>
<th>Val</th>
<th>Val</th>
<th>Thr</th>
<th>Thr</th>
<th>Lys</th>
<th>Pro</th>
<th>Val</th>
<th>Lys</th>
<th>Gln</th>
<th>Gln</th>
<th>Thr</th>
<th>Phe</th>
</tr>
</thead>
<tbody>
<tr>
<td>65</td>
<td>70</td>
<td>75</td>
<td>80</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Val</th>
<th>Pro</th>
<th>Glu</th>
<th>Gly</th>
<th>Leu</th>
<th>Val</th>
<th>Ala</th>
<th>Lys</th>
<th>Arg</th>
<th>Asp</th>
<th>Ile</th>
<th>Ala</th>
<th>Lys</th>
<th>Gly</th>
</tr>
</thead>
<tbody>
<tr>
<td>85</td>
<td>90</td>
<td>95</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Glu</th>
<th>Thr</th>
<th>Val</th>
<th>Gln</th>
<th>Val</th>
<th>Pro</th>
<th>Lys</th>
<th>Arg</th>
<th>Phe</th>
<th>Trp</th>
<th>Ile</th>
<th>Asn</th>
<th>Pro</th>
<th>Asp</th>
<th>Ala</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>105</td>
<td>110</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Val</th>
<th>Ala</th>
<th>Glu</th>
<th>Ser</th>
<th>Glu</th>
<th>Ile</th>
<th>Gly</th>
<th>Asn</th>
<th>Val</th>
<th>Cys</th>
<th>Ser</th>
<th>Gly</th>
<th>Leu</th>
<th>Lys</th>
<th>Pro</th>
<th>Trp</th>
</tr>
</thead>
<tbody>
<tr>
<td>115</td>
<td>120</td>
<td>125</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ile</th>
<th>Ser</th>
<th>Val</th>
<th>Ala</th>
<th>Leu</th>
<th>Phe</th>
<th>Leu</th>
<th>Arg</th>
<th>Glu</th>
<th>Lys</th>
<th>Trp</th>
<th>Arg</th>
<th>Asp</th>
<th>Asp</th>
<th>Ser</th>
</tr>
</thead>
<tbody>
<tr>
<td>130</td>
<td>135</td>
<td>140</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Lys</th>
<th>Trp</th>
<th>Lys</th>
<th>Tyr</th>
<th>Tyr</th>
<th>Met</th>
<th>Asp</th>
<th>Val</th>
<th>Leu</th>
<th>Pro</th>
<th>Lys</th>
<th>Ser</th>
<th>Thr</th>
<th>Asp</th>
<th>Ser</th>
</tr>
</thead>
<tbody>
<tr>
<td>145</td>
<td>150</td>
<td>155</td>
<td>160</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ile</th>
<th>Tyr</th>
<th>Trp</th>
<th>Ser</th>
<th>Glu</th>
<th>Glu</th>
<th>Leu</th>
<th>Ser</th>
<th>Glu</th>
<th>Ile</th>
<th>Gln</th>
<th>Gly</th>
<th>Thr</th>
<th>Gln</th>
<th>Leu</th>
</tr>
</thead>
<tbody>
<tr>
<td>165</td>
<td>170</td>
<td>175</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Leu</th>
<th>Ser</th>
<th>Thr</th>
<th>Thr</th>
<th>Met</th>
<th>Ser</th>
<th>Val</th>
<th>Lys</th>
<th>Asp</th>
<th>Tyr</th>
<th>Gln</th>
<th>Asn</th>
<th>Glu</th>
<th>Phe</th>
<th>Gln</th>
</tr>
</thead>
<tbody>
<tr>
<td>180</td>
<td>185</td>
<td>190</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Lys</th>
<th>Val</th>
<th>Glu</th>
<th>Glu</th>
<th>Val</th>
<th>Ile</th>
<th>Leu</th>
<th>Arg</th>
<th>Asn</th>
<th>Gln</th>
<th>Leu</th>
<th>Phe</th>
<th>Pro</th>
<th>Phe</th>
</tr>
</thead>
<tbody>
<tr>
<td>195</td>
<td>200</td>
<td>205</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Pro</th>
<th>Ile</th>
<th>Thr</th>
<th>Leu</th>
<th>Asp</th>
<th>Phe</th>
<th>Phe</th>
<th>Trp</th>
<th>Ala</th>
<th>Phe</th>
<th>Gly</th>
<th>Ile</th>
<th>Leu</th>
<th>Arg</th>
<th>Ser</th>
</tr>
</thead>
<tbody>
<tr>
<td>210</td>
<td>215</td>
<td>220</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Arg</th>
<th>Ala</th>
<th>Phe</th>
<th>Ser</th>
<th>Arg</th>
<th>Leu</th>
<th>Arg</th>
<th>Asn</th>
<th>Gln</th>
<th>Asn</th>
<th>Leu</th>
<th>Ile</th>
<th>Leu</th>
<th>Val</th>
<th>Pro</th>
<th>Phe</th>
</tr>
</thead>
<tbody>
<tr>
<td>225</td>
<td>230</td>
<td>235</td>
<td>240</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Ala  Asp  Leu  Thr  Asn  His  Asn  Ala  Arg  Val  Thr  Thr  Glu  Asp  His  Ala  
    245  250  255  
His  Glu  Val  Arg  Gly  Pro  Ala  Gly  Leu  Phe  Ser  Trp  Asp  Leu  Leu  Phe  
    260  265  270  
Ser  Leu  Arg  Ser  Pro  Leu  Lys  Leu  Ala  Gly  Asp  Gln  Leu  Phe  Ile  
    275  280  285  
Gln  Tyr  Asp  Leu  Asn  Lys  Ser  Asn  Ala  Asp  Met  Ala  Leu  Asp  Tyr  Gly  
    290  295  300  
Phe  Ile  Glu  Pro  Ser  Ser  Ala  Arg  Asp  Ala  Phe  Thr  Leu  Thr  Leu  Glu  
    305  310  315  320  
Ile  Ser  Glu  Ser  Asp  Glu  Phe  Tyr  Gly  Asp  Lys  Leu  Asp  Ile  Ala  Glu  
    325  330  335  
Thr  Asn  Gly  Ile  Gly  Glu  Thr  Ala  Tyr  Phe  Asp  Ile  Lys  Ile  Gly  Gln  
    340  345  350  
Ser  Leu  Pro  Pro  Thr  Met  Ile  Pro  Tyr  Leu  Arg  Leu  Val  Ala  Leu  Gly  
    355  360  365  
Gly  Thr  Asp  Ala  Phe  Leu  Leu  Glu  Ser  Leu  Phe  Arg  Asp  Thr  Ile  Trp  
    370  375  380  
Gly  His  Leu  Glu  Leu  Ser  Val  Ser  Arg  Asp  Ala  Glu  Leu  Leu  Cys  
    385  390  395  400  
Lys  Ala  Val  Arg  Glu  Ala  Cys  Lys  Ser  Ala  Leu  Ala  Gly  Tyr  His  Thr  
    405  410  415  
Thr  Ile  Glu  Gln  Asp  Arg  Glu  Leu  Lys  Glu  Gly  Asn  Leu  Asp  Ser  Arg  
    420  425  430  
Leu  Ala  Ile  Ala  Val  Gly  Ile  Arg  Leu  Gly  Lys  Arg  Val  Leu  Lys  
    435  440  445  
Gln  Ile  Asp  Asp  Ile  Phe  Arg  Glu  Glu  Leu  Asp  Glu  Leu  Asp  Leu  
    450  455  460  
Glu  Tyr  Tyr  Gly  Glu  Arg  Arg  Leu  Lys  Asp  Leu  Gly  Leu  Val  Gly  Glu  
    465  470  475  480  
Gln  Gly  Asp  Ile  Ile  Phe  Trp  Glu  Pro  Lys  
    485  490  

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 60 base pairs
    (B) TYPE: nucleic acid
    (C) STRANDEDNESS: single
    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

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

TTGCGGAACT TGCTTAAACT CTTTTAAC ATGAAATGGG TGCAACAGG CAACACAGG

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 36 base pairs
    (B) TYPE: nucleic acid
    (C) STRANDEDNESS: single
    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

UUGGCGGAU UUGGGGUC ACGGCUAAC CACAGU

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 12 amino acids
   (B) TYPE: amino acid
   (C) STRANDEDNESS: single
   (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:
Leu Ala Asp Leu Trp Val Gln Gln Ala Asn His Ser
   1    5

(2) INFORMATION FOR SEQ ID NO:29:

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

   (ii) MOLECULE TYPE: DNA (genomic)

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:
UUGGCGOAUUGGCAUACCAAG

(2) INFORMATION FOR SEQ ID NO:30:

   (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 8 amino acids
       (B) TYPE: amino acid
       (C) STRANDEDNESS: single
       (D) TOPOLOGY: linear

   (ii) MOLECULE TYPE: protein

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:
Leu Ala Asp Leu Ala Asn His Ser
   1    5

What is claimed is:
1. An isolated ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) large subunit "N-methyltransferase (LSMT) gene, wherein said gene is obtained from spinach with a des(methyl) lysyl residue in the large subunit of Rubisco.
2. The isolated gene of claim 1, wherein said spinach is Spinacia oleracea.
3. The isolated gene of claim 1, wherein said gene encodes an amino acid sequence of SEQ ID NO.: 22 or 23.
4. The isolated gene of claim 1, wherein said gene has a nucleotide sequence of SEQ ID NO.: 15.
5. A recombinant vector comprising the isolated Rubisco LSMT gene of claim 1.
6. The recombinant vector of claim 5, wherein said gene encodes an amino acid sequence of SEQ ID NO.: 22 or 23.
7. The recombinant vector of claim 5, wherein said gene has a nucleotide sequence of SEQ ID NO.: 15.
8. The recombinant vector of claim 5, wherein said vector is capable of transforming a plant.
10. The method of claim 9, wherein said gene encodes an amino acid sequence of SEQ ID NO.: 22 or 23.
11. The method of claim 9, wherein said gene has a nucleotide sequence of SEQ ID NO.: 15.
12. The method of claim 9, wherein said plant is a photosynthesizing plant.
14. The recombinant plant of claim 13, wherein said gene encodes an amino acid sequence of SEQ ID NO.: 22 or 23.
15. The recombinant plant of claim 13, wherein said gene has a nucleotide sequence of SEQ ID NO.: 15.
16. The recombinant plant of claim 13, wherein said plant is a photosynthesizing plant.
17. The isolated gene of claim 1, comprising the nucleotide sequence of SEQ ID NO.: 2 which encodes the amino acid sequence: WVOO (SEQ ID NO.: 1) of S40.
19. The method of claim 18, wherein said fragment has the nucleotide sequence (SEQ ID NO.: 2): TGGGTGCAACAG.
20. A method for preventing or reducing Rubisco LSMT activity in a photosynthesizing plant, comprising transforming a photosynthesizing plant with a recombinant vector wherein said vector comprises a Rubisco LSMT gene with the fragment of claim 17.
21. The method of claim 20, wherein said fragment has the nucleotide sequence (SEQ ID NO. 2): TGGGTGCAACAG.