Office de la Propriété Intellectuelle du Canada
Un organisme d'Industrie Canada

Canadian Intellectual Property Office
An agency of Industry Canada

CA 2272128 C 2009/03/17

(11)(21) 2 272 128

(12) BREVET CANADIEN
CANADIAN PATENT

(13) C

(86) Date de dépôt PCT/PCT Filing Date: 1997/11/18
(87) Date publication PCT/PCT Publication Date: 1998/05/28
(45) Date de délivrance/Issue Date: 2009/03/17
(85) Entrée phase nationale/National Entry: 1999/05/17
(86) N° demande PCT/PCT Application No.: US 1997/021145
(87) N° publication PCT/PCT Publication No.: 1998/022601
(30) Priorité/Priority: 1996/11/18 (US60/031,045)

(51) Cl.Int./Int.Cl. *C12N 15/54* (2006.01), *A01H 3/00* (2006.01), *A01H 5/00* (2006.01), *C12N 15/82* (2006.01), *C12N 5/10* (2006.01), *C12N 9/12* (2006.01), *C12Q 1/68* (2006.01)

(72) Inventeurs/Inventors:
HANNAH, L. CURTIS, US;
GREENE, TOM W., US

(73) Propriétaire/Owner:
UNIVERSITY OF FLORIDA, US

(74) Agent: MACRAE & CO.

(54) Titre : MUTANTS THERMOSTABLES D'ENZYMES DE BIOSYNTHESE D'AMIDON
(54) Title: HEAT STABLE MUTANTS OF STARCH BIOSYNTHESIS ENZYMES

(57) Abrégé/Abstract:
The subject invention pertains to novel mutant polynucleotide molecules that encode enzymes that have increased heat stability. These polynucleotides, when expressed in plants, result in increased yield in plants grown under conditions of heat stress. The polynucleotide molecules of the subject invention encode maize endosperm ADP glucose pyrophosphorylase (AGP) and soluble starch synthase (SSS) enzyme activities. Plants and plant tissue bred to contain, or transformed with, the mutant polynucleotides, and expressing the polypeptides encoded by the polynucleotides, are also contemplated by the present invention. The subject invention also concerns methods for isolating polynucleotides and polypeptides contemplated within the scope of the invention. Methods for increasing yield in plants grown under conditions of best stress are also provided.

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/82, 15/54, 9/12, C12Q 1/68, A01H 5/00

A1

(11) International Publication Number: WO 98/22601

(43) International Publication Date: 28 May 1998 (28.05.98)

(21) International Application Number: PCT/US97/21145

(22) International Filing Date: 18 November 1997 (18.11.97)

(30) Priority Data:
60/031,045 18 November 1996 (18.11.96) US

(71) Applicant: UNIVERSITY OF FLORIDA [US/US]; 223 Grinter Hall, Gainesville, FL 32611 (US).

(72) Inventors: HANNAH, L., Curtis; 4400 N.W. 39th Avenue #434, Gainesville, FL 32606 (US). GREENE, Tom, W.; 925 S.W. 82nd Terrace, Gainesville, FL 32607 (US).

(74) Agents: PACE, Doran, R. et al.; Saliwanchik, Lloyd & Saliwanchik, Suite A–1, 2421 N.W. 41st Street, Gainesville, FL 32606–6669 (US).

(81) Designated States: AL, AU, BA, BB, BG, BR, CA, CN, CU, CZ, EE, GE, HU, ID, IL, IS, JP, KP, KR, LC, LK, LR, LT, LV, MG, MK, MN, MX, NO, NZ, PL, RO, SG, SI, SK, SL, TR, TT, UA, UZ, VN, YU, ZW, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published

With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: HEAT STABLE MUTANTS OF STARCH BIOSYNTHESIS ENZYMES

(57) Abstract

The subject invention pertains to novel mutant polynucleotide molecules that encode enzymes that have increased heat stability. These polynucleotides, when expressed in plants, result in increased yield in plants grown under conditions of heat stress. The polynucleotide molecules of the subject invention encode maize endosperm ADP glucose pyrophosphorylase (AGP) and soluble starch synthase (SSS) enzyme activities. Plants and plant tissue bred to contain, or transformed with, the mutant polynucleotides, and expressing the polypeptides encoded by the polynucleotides, are also contemplated by the present invention. The subject invention also concerns methods for isolating polynucleotides and polypeptides contemplated within the scope of the invention. Methods for increasing yield in plants grown under conditions of best stress are also provided.

DESCRIPTION

HEAT STABLE MUTANTS OF STARCH BIOSYNTHESIS ENZYMES

This invention was made with government support under National Science Foundation grant number 9316887. The government has certain rights in the invention.

Background of the Invention

The sessile nature of plant life generates a constant exposure to environmental factors that exert positive and negative effects on its growth and development. One of the major impedients facing modern agriculture is adverse environmental conditions. One important factor which causes significant crop loss is heat stress. Temperature stress greatly reduces grain yield in many cereal crops such as maize, wheat, and barley. Yield decreases due to heat stress range from 7 to 35% in the cereals of world-wide importance.

A number of studies have identified likely physiological consequences of heat stress. Early work by Hunter *et al.* (Hunter, R.B., Tollenaar, M., and Breuer, C.M. [1977] *Can. J. Plant Sci.* 57:1127-1133) using growth chamber conditions showed that temperature decreased the duration of grain filling in maize. Similar results in which the duration of grain filling was adversely altered by increased temperatures were identified by Tollenaar and Bruulsema (Tollenaar, M. and Bruulsema, T.W. [1988] *Can. J. Plant Sci.* 68:935-940). Badu-Apraku *et al.* (Badu-Apraku, B., Hunter, R.B., and Tollenaar, M. [1983] *Can. J. Plant. Sci.* 63:357-363) measured a marked reduction in the yield of maize plants grown under the day/night temperature regime of 35/15°C compared to growth in a 25/15°C temperature regime. Reduced yields due to increased temperatures is also supported by historical as well as climatological studies (Thompson, L.M. [1986]

Agron. J. 78:649-653; Thompson, L. M. [1975] *Science* 188:535-541; Chang, J. [1981] *Agricul. Metero.* 24:253-262; and Conroy, J. P., Seneweera, S., Basra, A. S., Rogers, G., and Nissen-Wooller, B. [1994] *Aust. J. Plant Physiol.* 21:741-758).

That the physiological processes of the developing seed are adversely affected by heat stress is evident from studies using an *in vitro* kernel culture system (Jones, R.J., Gengenbach, B.G., and Cardwell, V.B. [1981] *Crop Science* 21:761-766; Jones, R.J., Ouattar, S., and Crookston, R.K. [1984] *Crop Science* 24:133-137; and Cheikh, N., and Jones, R.J. [995] *Physiol. Plant.* 95:59-66). Maize kernels cultured at the above-optimum temperature of 35 °C exhibited a dramatic reduction in weight.

Work with wheat identified the loss of soluble starch synthase (SSS) activity as a hallmark of the wheat endosperm's response to heat stress (Hawker, J. S. and Jenner, C. F. [1993] *Aust. J. Plant Physiol.* 20:197-209; Denyer, K., Hylton, C. M., and Smith, A. M. [1994] *Aust. J. Plant Physiol.* 21:783-789; Jenner, C. F. [1994] *Aust. J. Plant Physiol.* 21:791-806). Additional studies with SSS of wheat endosperm show that it is heat labile (Rijven, A.H.G.C. [1986] *Plant Physiol.* 81:448-453; Keeling, P.L., Bacon, P.J., Holt, D.C. [1993] *Planta.* 191:342-348; Jenner, C. F., Denyer, K., and Guerin, J. [1995] *Aust. J. Plant Physiol.* 22:703-709).

The roles of SSS and ADP glucose pyrophosphorylase (AGP) under heat stress conditions in maize is less clear. (AGP) catalyzes the conversion of ATP and α-glucose-1-phosphate to ADP-glucose and pyrophosphate. ADP-glucose is used as a glycosyl donor in starch biosynthesis by plants and in glycogen biosynthesis by bacteria. The importance of ADP-glucose pyrophosphorylase as a key enzyme in the regulation of starch biosynthesis was noted in the study of starch deficient mutants of maize (*Zea mays*) endosperm (Tsai, C.Y., and Nelson, Jr., O.E. [1966] *Science* 151:341-343; Dickinson, D.B., J. Preiss [1969] *Plant Physiol.* 44:1058-1062).

Ou-Lee and Setter (Ou-Lee, T. and Setter, T.L. [1985] *Plant Physiol.* 79:852-855) examined the effects of temperature on the apical or tip regions of maize ears. With elevated temperatures, AGP activity was lower in apical kernels when compared to basal kernels during the time of intense starch deposition. In contrast, in kernels developed at normal temperatures, AGP activity was similar in apical and basal kernels during this period. However, starch synthase activity during this period was not differentially

affected in apical and basal kernels. Further, heat-treated apical kernels exhibited an increase in starch synthase activity over control. This was not observed with AGP activity. Singletary *et al.* (Singletary, G.W., Banisadra, R., and Keeling, P.L. [1993] *Plant Physiol.* 102: 6 (suppl); Singletary, G.W., Banisadra, R., Keeling, P.L. [1994] *Aust. J. Plant Physiol.* 21:829-841) using an *in vitro* culture system quantified the effect of various temperatures during the grain fill period. Seed weight decreased steadily as temperature increased from 22-36°C. A role for AGP in yield loss is also supported by work from Duke and Doehlert (Duke, E.R. and Doehlert, D.C. [1996] *Envrion. Exp. Botany.* 36:199-208).

Work by Keeling *et al.* (1994, *supra*) quantified SSS activity in maize and wheat using Q_{10} analysis, and showed that SSS is an important control point in the flux of carbon into starch.

In vitro biochemcial studies with AGP and SSS clearly show that both enzymes are heat labile. Maize endosperm AGP loses 96% of its activity when heated at 57°C for five minutes (Hannah, L.C., Tuschall, D.M., and Mans, R.J. [1980] *Genetics* 95: 961-970). This is in contrast to potato AGP which is fully stable at 70°C (Sowokinos, J.R. and Preiss, J. [1982] *Plant Physiol.* 69:1459-1466; Okita, T.W., Nakata, P.A., Anderson, J.M., Sowokinos, J., Morell, J., and Preiss, J. [1990] *Plant Physiol.* 93:785-90). Heat inactivation studies with SSS showed that it is also labile at higher temperatures, and kinetic studies determined that the K*m* value for amylopectin rose exponentially when temperature increased from 25-45°C (Jenner *et al.*, 1995, *supra*).

Biochemical and genetic evidence has identified AGP as a key enzyme in starch biosynthesis in higher plants and glycogen biosynthesis in *E. coli* (Preiss, J. and Romeo, T. [1994] *Progress in Nuc. Acid Res. and Mol Biol.* 47:299-329; Preiss, J. and Sivak, M. [1996] "Starch synthesis in sinks and sources," In *Photoassimilate distribution in plants and crops: source-sink relationships*. Zamski, E., ed., Marcil Dekker Inc. pp. 139-168). AGP catalyzes what is viewed as the initial step in the starch biosynthetic pathway with the product of the reaction being the activated glucosyl donor, ADP glucose. This is utilized by starch synthase for extension of the polysaccharide polymer (reviewed in Hannah, L. Curtis (1997) "Starch Synthesis in the Maize Seed" In: Cellular and Molecular Biology of Plant Seed Development, B.A. Larkins and I.K. Vasil (eds.), Kluwer Academic Publishers,

printed in the Netherlands, pg. 375-405.

Initial studies with potato AGP showed that expression in *E. coli* yielded an enzyme with allosteric and kinetic properties very similar to the native tuber enzyme (Iglesias, A., Barry, G.F., Myer, C., Bloksberg, L., Nakata, P., Greene, T., Laughlin, M.J., Okita, T.W., Kishore, G.M., and Preiss, J. [1993] *J. Biol Chem.* 268:1081-86; Ballicora, M.A., Laughlin, M.J., Fu, Y., Okita, T.W., Barry, G.F., and Preiss, J. [1995] *Plant Physiol.* 109:245-251). Greene *et al.* (Greene, T.W., Chantler, S.E., Kahn, M.L., Barry, G.F., Preiss, J., and Okita, T.W. [1996] *Proc. Natl. Acad. Sci.* 93:1509-1513; Greene, Thomas W., Ronald L. Woodbury, Thomas W. Okita (1996) "Aspartic Acid 413 Is Important for the Normal Allosteric Functioning of ADP-Glucose Pyrophosphorylase" *Plant Physiol.* 112:1315-1320) showed the usefulness of the bacterial expression system in their structure-function studies with the potato AGP. Multiple mutations important in mapping allosteric and substrate binding sites were identified (Okita, T.W., Greene, T.W., Laughlin, M.J., Salamone, P., Woodbury, R., Choi, S., Ito, H., Kavakli, H., and Stephens, K. [1996] "Engineering Plant Starches by the Generation of Modified Plant Biosynthetic Enzymes," In *Engineering Crops for Industrial End Uses*, Shewry, P.R., Napier, J.A., and Davis, P., eds., Portland Press Ltd., London (pp.1-18).

AGP enzymes have been isolated from both bacteria and plants. Bacterial AGP consists of a homotetramer, while plant AGP from photosynthetic and non-photosynthetic tissues is a heterotetramer composed of two different subunits. The plant enzyme is encoded by two different genes, with one subunit being larger than the other. This feature has been noted in a number of plants. The AGP subunits in spinach leaf have molecular weights of 54 kDa and 51 kDa, as estimated by SDS-PAGE. Both subunits are immunoreactive with antibody raised against purified AGP from spinach leaves (Copeland, L., J. Preiss (1981) *Plant Physiol.* 68:996-1001; Morell, M., M. Bloon, V. Knowles, J. Preiss [1988] *J. Bio. Chem.* 263:633). Immunological analysis using antiserum prepared against the small and large subunits of spinach leaf showed that potato tuber AGP is also encoded by two genes (Okita *et al.*, 1990, *supra*). The cDNA clones of the two subunits of potato tuber (50 and 51 kDa) have also been isolated and sequenced (Muller-Rober, B.T., J. Kossmann, L.C. Hannah, L. Willmitzer, U. Sounewald [1990] *Mol. Gen. Genet.* 224:136-146; Nakata, P.A., T.W. Greene, J. M. Anderson, B.J.

Smith-White, T.W. Okita, J. Preiss [1991] *Plant Mol. Biol.* 17:1089-1093). The large subunit of potato tuber AGP is heat stable (Nakata *et al.* [1991], *supra*).

As Hannah and Nelson (Hannah, L.C., O.E. Nelson (1975) *Plant Physiol.* 55:297-302.; Hannah, L.C., and Nelson, Jr., O.E. [1976] *Biochem. Genet.* 14:547-560) postulated, both *Shrunken-2* (*Sh2*) (Bhave, M.R., S. Lawrence, C. Barton, L.C. Hannah [1990] *Plant Cell* 2:581-588) and *Brittle-2* (*Bt2*) (Bae, J.M., M. Giroux, L.C. Hannah [1990] *Maydica* 35:317-322) are structural genes of maize endosperm ADP-glucose pyrophosphorylase. *Sh2* and *Bt2* encode the large subunit and small subunit of the enzyme, respectively. From cDNA sequencing, *Sh2* and *Bt2* proteins have predicted molecular weight of 57,179 Da (Shaw, J.R., L.C. Hannah [1992] *Plant Physiol.* 98:1214-1216) and 52,224 Da, respectively. The endosperm is the site of most starch deposition during kernel development in maize. *Sh2* and *bt2* maize endosperm mutants have greatly reduced starch levels corresponding to deficient levels of AGP activity. Mutations of either gene have been shown to reduce AGP activity by about 95% (Tsai and Nelson, 1966, *supra*; Dickinson and Preiss, 1969, *supra*). Furthermore, it has been observed that enzymatic activities increase with the dosage of functional wild type *Sh2* and *Bt2* alleles, whereas mutant enzymes have altered kinetic properties. AGP is the rate limiting step in starch biosynthesis in plants. Stark *et al.* placed a mutant form of *E. coli* AGP in potato tuber and obtained a 35% increase in starch content (Stark *et al.* [1992] *Science* 258:287).

The cloning and characterization of the genes encoding the AGP enzyme subunits have been reported for various plants. These include *Sh2* cDNA (Bhave *et al.*, 1990, *supra*), *Sh2* genomic DNA (Shaw and Hannah, 1992, *supra*), and *Bt2* cDNA (Bae *et al.*, 1990, *supra*) from maize; small subunit cDNA (Anderson, J.M., J. Hnilo, R. Larson, T.W. Okita, M. Morell, J. Preiss [1989] *J. Biol. Chem.* 264:12238-12242) and genomic DNA (Anderson, J.M., R. Larson, D. Landencia, W.T. Kim, D. Morrow, T.W. Okita, J. Preiss [1991] *Gene* 97:199-205) from rice; and small and large subunit cDNAs from spinach leaf (Morell *et al.*, 1988, *supra*) and potato tuber (Muller-Rober *et al.*, 1990, *supra*; Nakata, P.A., Greene, T.W., Anderson, J.W., Smith-White, B.J., Okita, T.W., and Preiss, J. [1991] *Plant Mol. Biol.* 17:1089-1093). In addition, cDNA clones have been isolated from wheat endosperm and leaf tissue (Olive, M.R., R.J. Ellis, W.W. Schuch

[1989] *Plant Physiol. Mol. Biol.* 12:525-538) and *Arabidopsis thaliana* leaf (Lin, T., Caspar, T., Sommerville, C.R., and Preiss, J. [1988] *Plant Physiol.* 88:1175-1181).

AGP functions as an allosteric enzyme in all tissues and organisms investigated to date. The allosteric properties of AGP were first shown to be important in *E. coli.* A glycogen-overproducing *E. coli* mutant was isolated and the mutation mapped to the structural gene for AGP, designated as *gly*C. The mutant *E. coli*, known as *gly*C-16, was shown to be more sensitive to the activator, fructose 1,6 bisphosphate, and less sensitive to the inhibitor, cAMP (Preiss, J. [1984] *Ann. Rev. Microbiol.* 419-458). Although plant AGP's are also allosteric, they respond to different effector molecules than bacterial AGP's. In plants, 3-phosphoglyceric acid (3-PGA) functions as an activator while phosphate (PO_4) serves as an inhibitor (Dickinson and Preiss, 1969, *supra*).

Using an *in vivo* mutagenesis system created by the *Ac*-mediated excision of a *Ds* transposable element fortuitously located close to a known activator binding site, Giroux *et al.* (Giroux, M.J., Shaw, J., Barry, G., Cobb, G.B., Greene, T., Okita, T.W., and Hannah, L. C. [1996] *Proc. Natl. Acad. Sci.* 93:5824-5829) were able to generate site-specific mutants in a functionally important region of maize endosperm AGP. One mutant, *Rev 6,* contained a tyrosine-serine insert and conditioned a 11-18% increase in seed weight.

Brief Summary of the Invention

The subject invention pertains to materials and methods useful for improving crop yields in plants, such as those plants that produce cereal crops. In one embodiment, the subject invention provides heat stable AGP enzymes and nucleotide sequences which encode these enzymes. In a preferred embodiment, the heat stable enzymes can be used to provide plants having greater tolerance to higher temperatures, thus enhancing the crop yields from these plants. In a particularly preferred embodiment, the improved plant is a cereal. Cereals to which this invention applies include, for example, maize, wheat, rice, and barley.

Brief Description of the Drawings

Figure 1. Heat stable maize endosperm AGP large subunit mutants. Percentage of AGP activity remaining after five minutes of heat treatment at 60°C is shown.

Figure 2. Primary sequence alignment of region surrounding HS 33 mutation with large subunits of maize, wheat, barley, and potato. Conserved regions are boxed.

Figure 3. Primary sequence alignment of region surrounding HS 40 mutation with large subunits of maize, wheat, barley, and potato. Conserved regions are boxed. Bolded aspartic acid residue corresponds to D413A allosteric mutant of potato LS (Greene, Thomas W., Ronald L. Woodbury, Thomas W. Okita (1996) "Aspartic Acid 413 Is Important for the Normal Allosteric Functioning of ADP-Glucose Pyrophosphorylase" *Plant Physiol.* 112:1315-1320). Spinach leaf AGP sequence is the activator site 2 peptide identified in 3-PGA analogue studies (Ball, K. and Preiss, J. [1994] *J. Biol. Chem.* 269:24706-24711). The labeled lysine residue is bolded.

Detailed Disclosure of the Invention

The subject invention concerns novel mutant polynucleotide molecules, and the polypeptides encoded thereby, that confer increased yield in plants grown under conditions of heat stress relative to plants having wild type genotype. In specific embodiments, the polynucleotide molecules of the subject invention encode maize endosperm ADP glucose pyrophosphorylase (AGP) and soluble starch synthase (SSS) enzyme activities. The mutant enzymes confer increased stability of seeds to heat stress conditions during seed development as compared with wild type enzyme activities.

In one embodiment, a mutant polynucleotide of the present invention encodes a large subunit of AGP containing a histidine-to-tyrosine amino acid substitution in the sequence of the polypeptide. This substitution occurs at amino acid residue number 333, according to the accepted number of the amino acids in this protein (Shaw and Hannah, 1992, *supra*). The position of this substitution can be readily identified by a person skilled in the art. A second mutation exemplified in the subject invention is a threonine-to-isoleucine substitution at position number 460 of the AGP protein. Additional mutants confering increased heat stability are shown below in Table 1.

Table 1.

Mutant	Amino Acid Change
HS 13	Ala to Pro at position 177
HS 14	Asp to His at position 400, and Val to Ile at position 454
HS 16	Arg to Thr at position 104
HS 33	His to Tyr at position 333
HS 39	His to Tyr at position 333
HS 40	His to Tyr at position 333, and Thr to Ile at position 460
HS 47	Arg to Pro at position 216, and His to Tyr at position 333

cDNA clones for the subunits of the maize endosperm AGP (SH2 and BT2) and an *E. coli* strain deficient in the endogenous bacterial AGP (*glg* C$^-$) (AC70R1-504) have facilitated the establishment of a bacterial expression system to study the maize endosperm AGP. Expression of a single subunit is unable to complement the *glg* C$^-$ mutant, and no glycogen is produced (Iglesias, A., Barry, G. F., Meyer, C., Bloksberg, L., Nakata, P., Greene, T., Laughlin, M. J., Okita, T. W., Kishore, G. M., and Preiss, J. [1993] *J. Biol Chem.* 268: 1081-86). However, expression of both the large and small subunits on compatible expression vectors fully complements the *glg* C$^-$ mutation and restores glycogen production as evidenced by a dark, reddish-brown staining of colonies exposed to iodine. Thus, complementation is easily identified by simply exposing the colonies to iodine.

In one embodiment, *E. coli glg* C$^-$ cells expressing the structural genes for either potato or maize endosperm AGP were used. Cells containing potato AGP genes can synthesize copious levels of glycogen when grown at 37° or at 42°C. However, cells expressing maize endosperm AGP only synthesize glycogen at 37°C. This result demonstrates the heat sensitivity of wild-type maize endosperm AGP. That there is a difference between potato and maize AGP's in this regard provides an efficient system for screening for mutant cells that have heat stable variants of the maize endosperm AGP.

One aspect of the subject invention pertains to the efficient identification of AGP which is heat stable. Accordingly, a plasmid comprising a polynucleotide encoding for the SH2 subunit of maize AGP was chemically mutagenized, as described below, placed into mutant *E. coli* cells expressing the BT2 subunit, and grown at 42°C. Other mutagens known in the art can also be used. Eleven heritable, iodine staining mutants, termed heat stable (HS) mutants, were isolated. Crude extracts of these mutants were prepared and the heat stability of the resulting AGP was monitored. The mutants retain between 8-59% of their activity after incubation at 60°C for five minutes (Figure 1). This compares to the 1-4% routinely observed for wild-type AGP at this temperature.

The results show that heat stable forms of enzymes can be created according to the subject invention by mutation. Unexpectedly, total activity of the maize endosperm AGP before heat treatment was elevated about 10-fold in the majority of these mutants. This surprising result makes these mutants particularly advantageous for use in agriculture. Mutagenesis techniques as described herein can be used according to the subject invention to identify other genes encoding heat stable starch biosynthesis enzymes.

The genes encoding several of the heat stable mutants, including the two most heat stable HS mutants, HS 33 and HS 40, were completely sequenced. HS 33, which retains 59% of its activity after heat treatment, contains a single base pair mutation that changes a histidine residue at position 333 of the amino acid sequence of the polypeptide to a tyrosine (Figure 2). Primary sequence alignments with the large subunits from wheat and barley AGPs show that a histidine is also present at the analogous residue (Figure 3) (Ainsworth, C., Hosein, F., Tarvis, M., Weir, F., Burrell, M., Devos, K.M., Gale, M.D. [1995] *Planta* 197:1-10). Sequence analysis of HS 40, which retains 41% of its activity post heat treatment, also contained a histidine to tyrosine mutation at position 333. An additional point mutation was identified that generated a threonine to isoleucine substitution. The threonine residue is highly conserved in AGP large subunits, while in AGP small subunits the analogous residue is either a cysteine or serine (Ainsworth *et al.*, 1995, *supra*). The threonine to isoleucine substitution is located close to the carboxyl terminus of the large subunit, and close to a known binding site for the activator 3-PGA (Figure 3).

The subject invention also concerns heat stable mutants of AGP that have mutations in the small subunit of the enzyme. Also encompassed within the scope of the invention are polynucleotides that encode the mutant small subunits of AGP. Mutations in the small subunit of AGP that confer heat stability to the enzyme can also be readily prepared and identified using the methods of the subject invention.

Plants and plant tissue bred to contain or transformed with the mutant polynucleotides, and expressing the polypeptides encoded by the polynucleotides, are also contemplated by the present invention. Plants and plant tissue expressing the mutant polynucleotides produce tissues that have, for example, lower heat-induced loss in weight or yield when subjected to heat stress during development.

The subject invention also concerns methods for producing and identifying polynucleotides and polypeptides contemplated within the scope of the invention. In one embodiment, gene mutation, followed by selection using a bacterial expression system, can be used to isolate polynucleotide molecules that encode enzymes that can alleviate heat-induced loss in starch synthesis in plants.

The subject invention further concerns plants and plant tissue that have an AGP mutant gene incorporated into its genome. Other alleles disclosed herein can also be incorporated into a plant genome. In a preferred embodiment, the plant is a cereal plant. More preferably, the plant is *Zea mays*. Plants having an AGP mutant gene can be grown from seeds that comprise a mutant gene in their genome. In addition, techniques for transforming plants with a gene are known in the art.

Because of the degeneracy of the genetic code, a variety of different polynucleotide sequences can encode each of the variant AGP polypeptides disclosed herein. In addition, it is well within the skill of a person trained in the art to create alternative polynucleotide sequences encoding the same, or essentially the same, polypeptide of the subject invention. These variant or alternative polynucleotide sequences are within the scope of the subject invention. As used herein, references to "essentially the same" sequence refers to sequences which encode amino acid substitutions, deletions, additions, or insertions which do not materially alter the functional activity of the polypeptide encoded by the AGP mutant polypeptide described herein.

Substitution of amino acids other than those specifically exemplified in the mutants disclosed herein are also contemplated within the scope of the present invention. Amino acids can be placed in the following classes: non-polar, uncharged polar, basic, and acidic. Conservative substitutions whereby a mutant AGP polypeptide having an amino acid of one class is replaced with another amino acid of the same class fall within the scope of the subject invention so long as the mutant AGP polypeptide having the substitution still retains increased heat stability relative to a wild type polypeptide. Table 2 below provides a listing of examples of amino acids belonging to each class.

Table 2.

Class of Amino Acid	Examples of Amino Acids
Nonpolar	Ala, Val, Leu, Ile, Pro, Met, Phe, Trp
Uncharged Polar	Gly, Ser, Thr, Cys, Tyr, Asn, Gln
Acidic	Asp, Glu
Basic	Lys, Arg, His

For example, substitution of the tyrosine at position 333 in the HS 33, HS 39, HS 40 and HS 47 mutant maize endosperm AGP with other amino acids, such as Glycine, Serine, Threonine, Cysteine, Asparagine, and Glutamine, are encompassed within the scope of the invention.

The subject invention also concerns polynucleotides which encode fragments of the full length mutant polypeptide, so long as those fragments retain substantially the same functional activity as full length polypeptide. The fragments of mutant AGP polypeptide encoded by these polynucleotides are also within the scope of the present invention.

The subject invention also contemplates those polynucleotide molecules having sequences which are sufficiently homologous with the wild type *Sh2* DNA sequence so as to permit hybridization with that sequence under standard high-stringency conditions. Such hybridization conditions are conventional in the art (see, *e.g.*, Maniatis, T., E.F.

Fritsch, J. Sambrook [1989] *Molecular Cloning: A Laboratory Manual*, 2nd Edition, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York).

The polynucleotide molecules of the subject invention can be used to transform plants to express the mutant heat stable AGP enzyme in those plants. In addition, the polynucleotides of the subject invention can be used to express the recombinant variant AGP enzyme. They can also be used as a probe to detect related enzymes. The polynucleotides can also be used as DNA sizing standards.

The polynucleotide molecules of the subject invention also include those polynucleotides that encode AGP enzymes that confer increased seed weight, in addition to enhanced heat stability, on a plant. Combination of a heat stabilizing mutation, HS, with a mutation conferring increased seed weight, *e.g.*, Rev 6, is specifically contemplated for the present invention. See, for example, U.S. Patent Nos. 5,589,618 and 5,650,557.

Mutations in the AGP subunits that confer heat stability can be combined according to the subject invention with phosphate insensitive mutants of maize, such as the *Rev6* mutation, to enhance the stability of the *Rev6* encoded large subunit.

It is expected that enzymic activity of SSS will be impaired at the higher temperature as found with AGP. Thus, mutagenized forms of SSS can be expressed under increased thermal conditions (42°C), to isolate heat stable variants. These heat stable mutagenized forms of SSS are further aspects of the subject invention.

Following are examples which illustrate procedures for practicing the invention. These examples should not be construed as limiting. All percentages are by weight and all solvent mixture proportions are by volume unless otherwise noted.

Example 1 - Use of Mutagenesis to Obtain Maize Endosperm AGP Heat Stable Variants

The chemical mutagen hydroxylamine-HCl was initially used for the random mutagenesis of the large subunit expression plasmid. Hydroxylamine preferentially hydroxylates the amino nitrogen at the C-4 position of cytosine, and leads to a GC to AT transition (Suzuki D.T., Griffith, A.J.F., Miller, J.H., and Lewontin, R.C. [1989]. In *Introduction to genetic analysis*, Freeman, NY, 4th ed., pp. 475-499). The chemical

mutagen was chosen for its high mutation frequency. Limitations of the chemical mutagen are recognized, and if a large variety of genetic variants are not isolated, PCR based random mutagenesis can be performed. PCR mutagenesis generates a broader spectrum of mutations that include similar frequencies of transitions and transversion, and provides an excellent alternative to the chemical method. The method outlined by Cadwell and Joyce (Cadwell, R.C. and Joyce, G.F. [1992] *PCR Methods and Applications* 2:28-33) can be followed for the PCR based method.

Since the complete expression plasmid is used in the random mutagenesis, it is possible that mutations will occur outside of the coding region. Although it is expected that such mutations will not have any effect on the heat stability of the maize endosperm AGP, each variant can be subcloned into an unmutated expression plasmid before any additional characterization at the enzyme level is conducted. Both the large and small subunit expression plasmids can be constructed so that a *Nco*I/*Sac*I digestion will liberate the complete coding region. This can easily be cloned back into a unmutated *Nco*I/*Sac*I digested expression plasmid.

Example 2 – Molecular Characterization and Analysis of Heat Stable AGP Variants

Initially, 11 heat stable variants of the maize endosperm large subunit were obtained. Two have been completely sequenced. Sequencing is done using DuPont and ABI instrumentation.

Sequence data can be routinely compared to the progenitor wild-type allele. This analysis reveals the extent of diversity of changes conditioning heat stability.

Both sequenced HS mutants contain the identical histidine to tyrosine change in the large subunit. PCR-derived HS mutants can be quickly screened for the histidine to tyrosine alteration by use of site-specific mutagenesis using primers that change the tyrosine back to histidine.

Example 3 – Expression, Purification, and Kinetic Analysis of Genetic Variants

Conditions for the expression of the wild-type maize endosperm AGP in *E. coli* have been fully characterized. Optimum growth and induction conditions vary somewhat from those previously published for potato AGP expressed in *E. coli* (Iglesias *et al.*,

1993, *supra*; Ballicora *et al.*, 1995, *supra*). Induction at room temperature for 12-14 hrs in the presence of 0.3 mM IPTG and 25 μg/ml nalidixic acid consistently gives high levels of expression and activity. Addition of 30% ammonium sulfate and 10 mM $KH_2PO_4^-/K_2HPO_4^-$ to the extraction buffer stabilizes the maize AGP in the crude extract.

Ammonium sulfate concentrated AGP is further purified by Hydrophobic Interaction Chromatography using Tentacle C3™ aminopropyl media (EM Separations) packed into a Pharmacia HR™ 10/10 column. Protein binds to the column in a buffer containing 1 M ammonium sulfate. AGP is eluted from the column by successive step gradient washes of buffer that contains 0.75 M, 0.5 M, 0.25 M, and 0 M ammonium sulfate. Wild-type maize endosperm AGP typically elutes in the 0.25 M wash. C3 purified maize endosperm AGP is further purified by anion exchange chromatography using Macro-Prep™ DEAE (BioRad) anion exchange media packed into a Pharmacia HR 10/10 column. AGP is eluted by a linear gradient of 100-500 mM KCl, and typically elutes at a salt concentration around 300 mM. A Pharmacia FPLC™ system is used for all chromatography steps. The conditions for the individual purification steps are fully characterized. AGP activity during the purification is monitored by the pyrophosphorylysis assay, and purification steps are monitored by SDS-PAGE, Coomassie staining, and Western analysis using polyclonal antibodies specific to the maize endosperm AGP large and small subunits.

Example 4 - Enhanced Subunit Interaction

A totally unexpected pleiotropic effect of the HS maize endosperm AGP mutants is a 2-fold elevation of activity before heat treatment. One possible explanation for this result is that we have, by mutational change, shifted the ratio of SH2 and BT2 monomers and polymers existing within the *E. coli* cell. Perhaps, in wild-type, only 10% or less of the total proteins exist in the active heterotetrameric form whereas in the mutants, this percentage is much higher. If the polymer is more heat resistant than are the monomers, then the phenotype of the mutants would be identical to what has been observed Kinetic analysis can be used to determine changes in affinities for substrates and/or allosteric effectors.

To test the idea that the monomer/polymer ratio may be altered in these mutants, the amounts of monomers and polymers in wild-type and in selected mutants both before and after heat treatment can be monitored. The availability of antibodies (Giroux, M.J., and Hannah, L.C. [1994] *Mol. Gen. Genetics* 243:400-408) for both subunits makes this approach feasible. This can be examined both through sucrose gradient ultracentrifugation and through gel chromatography and will readily determine which method is most efficient and definitive.

Since the higher plant AGP consists of two similar but distinct subunits that oligomerize to form the native heterotetrameric structure, mutations that enhance this interaction can provide added stability to the enzyme. A yeast two-hybrid system (CLONTECH Laboratories, Palo Alto, CA) can be used to evaluate subunit interactions. Specific primers for the amplification of the coding regions can be constructed. These primers add unique restriction sites to the 5′- and 3′- ends so that cloning facilitates the translational fusion of the individual subunit to the GAL4 DNA binding domain (pGBT9) or GAL4 activation domain (pGAD424). If the proteins cloned into the vectors interact, the DNA binding domain and the activation domain will form a functional transcription activator. This in turn activates expression of the reporter gene, lac Z, cloned behind a GAL4 promoter.

Initially, conditions can be characterized with the wild-type subunits. The coding regions of the wild-type large and small subunits can be cloned into the pGBT9 and pGAD424 yeast expression vectors. All possible combinations can be generated and tested. pGBT9 and pGAD424 vectors containing *Sh2* and *Bt2* can be cotransformed into the same yeast strain, and selected for growth on media lacking tryptophan (pGBT9) and leucine (pGAD424). Subunit interaction as a function of lacZ expression can be detected two ways. Positive colonies are visually identified by a B-galactosidase filter assay. With this assay colonies are bound to the filter, lysed, and incubated with an X-gal solution. Colonies that exhibit a blue color can be analyzed. Subunit interaction can be further analyzed by an enzyme assay specific for B-galactosidase. This allows the quantification of the interaction. Mutations that enhance subunit interactions will give higher levels of B-galactosidase activity when assayed.

Example 5 – Further Enhancement of Stability

The large subunit mutants isolated vary in their heat stability characteristics, suggesting the possibility of multiple mutations. While sequence analysis of mutants HS 33 and HS 40 reveal that the mutant sequences are not identical, both mutants contained the identical histidine to tyrosine change. Given the identification of different HS alterations within the SH2 protein, it is possible to efficiently pyramid these changes into one protein. Furthermore, any HS mutations within the small subunit can be co-expressed with HS SH2 mutants to further enhance the stability of the maize endosperm enzyme.

Multiple HS mutants within one subunit can easily be combined. For example, different unique restriction sites that divide the coding regions of *Sh2* into three distinct fragments can be used. Where appropriate, mutation combinations can be generated by subcloning the corresponding fragment containing the added mutation. If two mutations are in close proximity, then site-directed mutagenesis can be used to engineer such combinations. One method for site specific mutations involves PCR, mutagenic primer, and the use of *Dpn*I restriction endonuclease. Primers can be constructed to contain the mutation in the 5′ end, and used to PCR amplify using the proofreading polymerase Vent. Amplified DNA can then be digested with *Dpn*I. Parental DNA isolated from *E. coli* is methylated and hence susceptible to *Dpn*I. Digested DNA is size fractionated by gel electrophoresis, ligated, and cloned into the expression vectors. Mutations are confirmed by sequence analysis and transformed into the AC70R1-504 strain carrying the wild-type small subunit. Combinatorial mutants can then be analyzed.

Example 6 – Combination of Heat Stability Mutations with *Rev6*

According to the subject invention, the heat stable mutations can be combined with a mutation associated with increased seed weight, such as, for example, the *Rev6* mutation. The goal is to maintain the desired phosphate insensitivity characteristic of *Rev6* while enhancing its stability. Rev 6/HS double mutants can be constructed and confirmed as described herein. Double mutants can be transformed into AC70R1-504 carrying the wild-type small subunit. Increased heat stability can be easily identified by a positive glycogen staining on a low glucose media. *Rev6* does not stain when grown

on this media. Initially all mutant combinations can be screened enzymatically for maintenance of phosphate insensitivity, and only combinations that maintain phosphate insensitivity are further analyzed.

Example 7 – Cloning of SSS I Mutants

A *glg* A⁻ *E. coli* strain deficient in the endogenous bacterial glycogen synthase can be obtained from the *E. coli* Stock Center. Bacterial expression vectors currently used for the expression of AGP can be used for expression of SSS.

One cloning strategy, as used, for example, with *Sh2* and *Bt2* (Giroux *et al.*, 1996, *supra*), is the following: One primer contains a unique restriction plus the 5′ terminus of the transcript while the other primer contains another unique restriction site and sequences 3′ to the translational termination codon of the gene under investigation. Subsequent cloning of these gives rise to a translational fusion within the plasmid. These gene specific primers are initially used in RT-PCR reactions using poly A+RNA from developing endosperms.

Expression of the maize endosperm SSS I will complement the lack of glycogen synthase activity in the *glg* A⁻ strain. Complementation should be easily visualized with iodine staining as it is with the expression of AGP in the *glg* C⁻ strain. Crude extracts can be incubated at various temperatures and lengths of time to determine the heat stability of SSS I. The *glg* A⁻ strain expressing the maize endosperm SSS I can be grown at various temperatures to determine if function is temperature sensitive as it is with the AGP bacterial expression system. Once a restrictive temperature is established, a random mutagenesis can be conducted with the SSS I clone. Mutant forms of SSS I can be transformed into the *glg* A⁻ strain, grown at the restrictive temperature, and heat stable variants identified by their ability to produce iodine-staining glycogen at the restriction temperature.

It should be understood that the examples and embodiments described herein are for illustrative purposes only and that various modifications or changes in light thereof will be suggested to persons skilled in the art and are to be included within the spirit and purview of this application and the scope of the appended claims.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: University of Florida

(ii) TITLE OF INVENTION: Heat Stable Mutants of Starch Biosynthesis Enzymes

(iii) NUMBER OF SEQUENCES: 12

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
(B) STREET: 2421 N.W. 41st Street, Suite A-1
(C) CITY: Gainesville
(D) STATE: FL
(E) COUNTRY: US
(F) ZIP: 32606

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: PCT/US97/21145
(B) FILING DATE: 18-NOV-1997
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 60/031,045
(B) FILING DATE: 18-NOV-1996
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Pace, Doran R.
(B) REGISTRATION NUMBER: 38,261
(C) REFERENCE/DOCKET NUMBER: UF-178 Canada

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 352-375-8100
(B) TELEFAX: 352-372-5800

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 516 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Amino Acid

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

Met Gln Phe Ala Leu Ala Leu Asp Thr Asn Ser Gly Pro His Gln Ile
1 5 10 15

Arg Ser Cys Glu Gly Asp Gly Ile Asp Arg Leu Glu Lys Leu Ser Ile
20 25 30

Gly Gly Arg Lys Gln Glu Lys Ala Leu Arg Asn Arg Cys Phe Gly Gly
35 40 45

Arg Val Ala Ala Thr Thr Gln Cys Ile Leu Thr Ser Asp Ala Cys Pro
50 55 60

Glu Thr Leu His Ser Gln Thr Gln Ser Ser Arg Lys Asn Tyr Ala Asp
65 70 75 80

Ala Asn Arg Val Ser Ala Ile Ile Leu Gly Gly Gly Thr Gly Ser Gln
85 90 95

Leu Phe Pro Leu Thr Ser Thr Arg Ala Thr Pro Ala Val Pro Val Gly
100 105 110

Gly Cys Tyr Arg Leu Ile Asp Ile Pro Met Ser Asn Cys Phe Asn Ser
115 120 125

Gly Ile Asn Lys Ile Phe Val Met Ser Gln Phe Asn Ser Thr Ser Leu
130 135 140

Asn Arg His Ile His Arg Thr Tyr Leu Glu Gly Gly Ile Asn Phe Ala
145 150 155 160

Asp Gly Ser Val Gln Val Leu Ala Ala Thr Gln Met Pro Glu Glu Pro
165 170 175

Ala Gly Trp Phe Gln Gly Thr Ala Asp Ser Ile Arg Lys Phe Ile Trp
180 185 190

Val Leu Glu Asp Tyr Tyr Ser His Lys Ser Ile Asp Asn Ile Val Ile
195 200 205

Leu Ser Gly Asp Gln Leu Tyr Arg Met Asn Tyr Met Glu Leu Val Gln
210 215 220

Lys His Val Glu Asp Asp Ala Asp Ile Thr Ile Ser Cys Ala Pro Val
225 230 235 240

Asp Glu Ser Arg Ala Ser Lys Asn Gly Leu Val Lys Ile Asp His Thr
245 250 255

Gly Arg Val Leu Gln Phe Phe Glu Lys Pro Lys Gly Ala Asp Leu Asn
260 265 270

Ser Met Arg Val Glu Thr Asn Phe Leu Ser Tyr Ala Ile Asp Asp Ala
275 280 285

Gln Lys Tyr Pro Tyr Leu Ala Ser Met Gly Ile Tyr Val Phe Lys Lys
290 295 300

Asp Ala Leu Leu Asp Leu Leu Lys Ser Lys Tyr Thr Gln Leu His Asp
305 310 315 320

Phe Gly Ser Glu Ile Leu Pro Arg Ala Val Leu Asp Tyr Ser Val Gln
325 330 335

Ala Cys Ile Phe Thr Gly Tyr Trp Glu Asp Val Gly Thr Ile Lys Ser
340 345 350

Phe Phe Asp Ala Asn Leu Ala Leu Thr Glu Gln Pro Ser Lys Phe Asp
355 360 365

Phe Tyr Asp Pro Lys Thr Pro Phe Phe Thr Ala Pro Arg Cys Leu Pro
370 375 380

Pro Thr Gln Leu Asp Lys Cys Lys Met Lys Tyr Ala Phe Ile Ser Asp
385 390 395 400

Gly Cys Leu Leu Arg Glu Cys Asn Ile Glu His Ser Val Ile Gly Val
405 410 415

Cys Ser Arg Val Ser Ser Gly Cys Glu Leu Lys Asp Ser Val Met Met
420 425 430

Gly Ala Asp Ile Tyr Glu Thr Glu Glu Glu Ala Ser Lys Leu Leu Leu
435 440 445

Ala Gly Lys Val Pro Ile Gly Ile Gly Arg Asn Thr Lys Ile Arg Asn
450 455 460

Cys Ile Ile Asp Met Asn Ala Arg Ile Gly Lys Asn Val Val Ile Thr
465 470 475 480

Asn Ser Lys Gly Ile Gln Glu Ala Asp His Pro Glu Glu Gly Tyr Tyr
485 490 495

Ile Arg Ser Gly Ile Val Val Ile Leu Lys Asn Ala Thr Ile Asn Asp
500 505 510

Gly Ser Val Ile
515

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 516 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Amino Acid

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

Met Gln Phe Ala Leu Ala Leu Asp Thr Asn Ser Gly Pro His Gln Ile
1 5 10 15

Arg Ser Cys Glu Gly Asp Gly Ile Asp Arg Leu Glu Lys Leu Ser Ile
20 25 30

Gly Gly Arg Lys Gln Glu Lys Ala Leu Arg Asn Arg Cys Phe Gly Gly
35 40 45

Arg Val Ala Ala Thr Thr Gln Cys Ile Leu Thr Ser Asp Ala Cys Pro
50 55 60

Glu Thr Leu His Ser Gln Thr Gln Ser Ser Arg Lys Asn Tyr Ala Asp
65 70 75 80

Ala Asn Arg Val Ser Ala Ile Ile Leu Gly Gly Gly Thr Gly Ser Gln
85 90 95

Leu Phe Pro Leu Thr Ser Thr Arg Ala Thr Pro Ala Val Pro Val Gly
100 105 110

Gly Cys Tyr Arg Leu Ile Asp Ile Pro Met Ser Asn Cys Phe Asn Ser
115 120 125

Gly Ile Asn Lys Ile Phe Val Met Ser Gln Phe Asn Ser Thr Ser Leu
130 135 140

Asn Arg His Ile His Arg Thr Tyr Leu Glu Gly Gly Ile Asn Phe Ala
145 150 155 160

```
Asp Gly Ser Val Gln Val Leu Ala Ala Thr Gln Met Pro Glu Glu Pro
                 165                 170                 175

Ala Gly Trp Phe Gln Gly Thr Ala Asp Ser Ile Arg Lys Phe Ile Trp
            180                 185                 190

Val Leu Glu Asp Tyr Tyr Ser His Lys Ser Ile Asp Asn Ile Val Ile
        195                 200                 205

Leu Ser Gly Asp Gln Leu Tyr Arg Met Asn Tyr Met Glu Leu Val Gln
    210                 215                 220

Lys His Val Glu Asp Asp Ala Asp Ile Thr Ile Ser Cys Ala Pro Val
225                 230                 235                 240

Asp Glu Ser Arg Ala Ser Lys Asn Gly Leu Val Lys Ile Asp His Thr
                245                 250                 255

Gly Arg Val Leu Gln Phe Phe Glu Lys Pro Lys Gly Ala Asp Leu Asn
            260                 265                 270

Ser Met Arg Val Glu Thr Asn Phe Leu Ser Tyr Ala Ile Asp Asp Ala
        275                 280                 285

Gln Lys Tyr Pro Tyr Leu Ala Ser Met Gly Ile Tyr Val Phe Lys Lys
    290                 295                 300

Asp Ala Leu Leu Asp Leu Leu Lys Ser Lys Tyr Thr Gln Leu His Asp
305                 310                 315                 320

Phe Gly Ser Glu Ile Leu Pro Arg Ala Val Leu Asp Tyr Ser Val Gln
                325                 330                 335

Ala Cys Ile Phe Thr Gly Tyr Trp Glu Asp Val Gly Thr Ile Lys Ser
            340                 345                 350

Phe Phe Asp Ala Asn Leu Ala Leu Thr Glu Gln Pro Ser Lys Phe Asp
        355                 360                 365

Phe Tyr Asp Pro Lys Thr Pro Phe Phe Thr Ala Pro Arg Cys Leu Pro
    370                 375                 380

Pro Thr Gln Leu Asp Lys Cys Lys Met Lys Tyr Ala Phe Ile Ser Asp
385                 390                 395                 400

Gly Cys Leu Leu Arg Glu Cys Asn Ile Glu His Ser Val Ile Gly Val
                405                 410                 415
```

```
Cys Ser Arg Val Ser Ser Gly Cys Glu Leu Lys Asp Ser Val Met Met
            420                 425                 430

Gly Ala Asp Ile Tyr Glu Thr Glu Glu Glu Ala Ser Lys Leu Leu Leu
        435                 440                 445

Ala Gly Lys Val Pro Ile Gly Ile Gly Arg Asn Ile Lys Ile Arg Asn
    450                 455                 460

Cys Ile Ile Asp Met Asn Ala Arg Ile Gly Lys Asn Val Val Ile Thr
465                 470                 475                 480

Asn Ser Lys Gly Ile Gln Glu Ala Asp His Pro Glu Glu Gly Tyr Tyr
                485                 490                 495

Ile Arg Ser Gly Ile Val Val Ile Leu Lys Asn Ala Thr Ile Asn Asp
            500                 505                 510

Gly Ser Val Ile
        515
```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 516 Amino Acid
(B) TYPE: Amino Acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Amino Acid

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

```
Met Gln Phe Ala Leu Ala Leu Asp Thr Asn Ser Gly Pro His Gln Ile
  1               5                  10                  15

Arg Ser Cys Glu Gly Asp Gly Ile Asp Arg Leu Glu Lys Leu Ser Ile
            20                  25                  30

Gly Gly Arg Lys Gln Glu Lys Ala Leu Arg Asn Arg Cys Phe Gly Gly
        35                  40                  45

Arg Val Ala Ala Thr Thr Gln Cys Ile Leu Thr Ser Asp Ala Cys Pro
    50                  55                  60

Glu Thr Leu His Ser Gln Thr Gln Ser Ser Arg Lys Asn Tyr Ala Asp
65                  70                  75                  80
```

Ala Asn Arg Val Ser Ala Ile Ile Leu Gly Gly Gly Thr Gly Ser Gln
85 90 95

Leu Phe Pro Leu Thr Ser Thr Arg Ala Thr Pro Ala Val Pro Val Gly
100 105 110

Gly Cys Tyr Arg Leu Ile Asp Ile Pro Met Ser Asn Cys Phe Asn Ser
115 120 125

Gly Ile Asn Lys Ile Phe Val Met Ser Gln Phe Asn Ser Thr Ser Leu
130 135 140

Asn Arg His Ile His Arg Thr Tyr Leu Glu Gly Gly Ile Asn Phe Ala
145 150 155 160

Asp Gly Ser Val Gln Val Leu Ala Ala Thr Gln Met Pro Glu Glu Pro
165 170 175

Ala Gly Trp Phe Gln Gly Thr Ala Asp Ser Ile Arg Lys Phe Ile Trp
180 185 190

Val Leu Glu Asp Tyr Tyr Ser His Lys Ser Ile Asp Asn Ile Val Ile
195 200 205

Leu Ser Gly Asp Gln Leu Tyr Arg Met Asn Tyr Met Glu Leu Val Gln
210 215 220

Lys His Val Glu Asp Asp Ala Asp Ile Thr Ile Ser Cys Ala Pro Val
225 230 235 240

Asp Glu Ser Arg Ala Ser Lys Asn Gly Leu Val Lys Ile Asp His Thr
245 250 255

Gly Arg Val Leu Gln Phe Phe Glu Lys Pro Lys Gly Ala Asp Leu Asn
260 265 270

Ser Met Arg Val Glu Thr Asn Phe Leu Ser Tyr Ala Ile Asp Asp Ala
275 280 285

Gln Lys Tyr Pro Tyr Leu Ala Ser Met Gly Ile Tyr Val Phe Lys Lys
290 295 300

Asp Ala Leu Leu Asp Leu Leu Lys Ser Lys Tyr Thr Gln Leu His Asp
305 310 315 320

Phe Gly Ser Glu Ile Leu Pro Arg Ala Val Leu Asp His Ser Val Gln
325 330 335

Ala Cys Ile Phe Thr Gly Tyr Trp Glu Asp Val Gly Thr Ile Lys Ser
340 345 350

Phe Phe Asp Ala Asn Leu Ala Leu Thr Glu Gln Pro Ser Lys Phe Asp
355 360 365

Phe Tyr Asp Pro Lys Thr Pro Phe Phe Thr Ala Pro Arg Cys Leu Pro
370 375 380

Pro Thr Gln Leu Asp Lys Cys Lys Met Lys Tyr Ala Phe Ile Ser Asp
385 390 395 400

Gly Cys Leu Leu Arg Glu Cys Asn Ile Glu His Ser Val Ile Gly Val
405 410 415

Cys Ser Arg Val Ser Ser Gly Cys Glu Leu Lys Asp Ser Val Met Met
420 425 430

Gly Ala Asp Thr Tyr Glu Thr Glu Glu Glu Ala Ser Lys Leu Leu Leu
435 440 445

Ala Gly Lys Val Pro Val Gly Ile Gly Arg Asn Thr Lys Ile Arg Asn
450 455 460

Cys Ile Ile Asp Met Asn Ala Arg Ile Gly Lys Asn Val Val Ile Thr
465 470 475 480

Asn Ser Lys Gly Ile Gln Glu Ala Asp His Pro Glu Glu Gly Tyr Tyr
485 490 495

Ile Arg Ser Gly Ile Val Val Ile Leu Lys Asn Ala Thr Ile Asn Asp
500 505 510

Gly Ser Val Ile
515

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7320 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Nucleic acid

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

tgatgctttt tcctgggcag ggagagctat gagacgtatg tcctcaaagc cactttgcat 60
tgtgtgaaac caatatcgat ctttgttact tcatcatgca tgaacatttg tggaaactac 120
tagcttacaa gcattagtga cagctcagaa aaaagttatc tctgaaaggt ttcatgtgta 180
ccgtgggaaa tgagaaatgt tgccaactca aacaccttca atatgttgtt tgcaggcaaa 240
ctcttctgga agaaaggtgt ctaaaactat gaacgggtta cagaaaggta taaaccacgg 300
ctgtgcattt tggaagtatc atctatagat gtctgttgag gggaaagccg tacgccaacg 360
ttatttactc agaaacagct tcaacacaca gttgtctgct ttatgatggc atctccaccc 420
aggcacccac catcacctat tcacctatct ctcgtgcctg tttatttttct tgccctttct 480
gatcataaaa aatcattaag agtttgcaaa catgcatagg catatcaata tgctcattta 540
ttaatttgct agcagatcat cttcctactc tttactttat ttattgtttg aaaaatatgt 600
cctgcaccta gggagctcgt atacagtacc aatgcatctt cattaaatgt gaatttcaga 660
aaggaagtag gaacctatga gagtattttt caaaattaat tagcggcttc tattatgttt 720
atagcaaagg ccaagggcaa aatcggaaca ctaatgatgg ttggttgcat gagtctgtcg 780
attacttgca agaaatgtga acctttgttt ctgtgcgtgg gcataaaaca aacagcttct 840
agcctctttt acggtacttg cacttgcaag aaatgtgaac tccttttcat ttctgtatgt 900
ggacataatg ccaaagcatc caggcttttt catggttgtt gatgtcttta cacagttcat 960
ctccaccagt atgccctcct catactctat ataaacacat caacagcatc gcaattagcc 1020
acaagatcac ttcgggaggc aagtgtgatt tcgaccttgc agccaccttt ttttgttctg 1080
ttgtaagtat actttccctt accatcttta tctgttagtt taatttgtaa ttgggaagta 1140
ttagtggaaa gaggatgaga tgctatcatc tatgtactct gcaaatgcat ctgacgttat 1200
atgggctgct tcatataatt tgaattgctc cattcttgcc gacaatatat tgcaaggtat 1260
atgcctagtt ccatcaaaag ttctgttttt tcattctaaa agcattttag tggcacgcaa 1320
ttttgtccat gagggaaagg aaatctgttt tggttacttt gcttgaggtg cattcttcat 1380
atgtccagtt ttatggaagt aataaacttc agtttggtca taagatgtca tattaaaggg 1440
caaacatata ttcaatgttc aattcatcgt aaatgttccc tttttgtaaa agattgcata 1500
ctcatttatt tgagttgcag gtgtatctag tagttggagg agatatgcag tttgcacttg 1560
cattggacac gaactcaggt cctcaccaga taagatcttg tgagggtgat gggattgaca 1620
ggttggaaaa attaagtatt gggggcagaa agcaggagaa agctttgaga aataggtgct 1680
ttggtggtag agttgctgca actacacaat gtattcttac ctcagatgct tgtcctgaaa 1740
ctcttgtaag tatccacctc aattattact cttacatgtt ggtttacttt acgtttgtct 1800
tttcaaggga aatttactgt atttttgtg ttttgtggga gttctatact tctgttggac 1860
tggttattgt aaagatttgt tcaaataggg tcatctaata attgtttgaa atctgggaac 1920
tgtggtttca ctgcgttcag gaaaaagtga attattggtt actgcatgaa taacttatgg 1980
aaatagacct tagagttgct gcatgattat cacaaatcat tgctacgata tcttataata 2040
gttctttcga cctcgcatta catatataac tgcaactcct agttgcgttc aaaaaaaaaa 2100
atgcaactct tagaacgctc accagtgtaa tctttcctga attgttattt aatggcatgt 2160
atgcactact tgtatactta tctaggatta agtaatctaa ctctaggccc catatttgca 2220
gcattctcaa acacagtcct ctaggaaaaa ttatgctgat gcaaaccgtg tatctgctat 2280
cattttgggc ggaggcactg gatctcagct ctttcctctg acaagcacaa gagctacgcc 2340
tgctgtaagg gataacactg aacatccaac gttgattact ctattatagt attatacaga 2400
ctgtactttt cgaatttatc ttagttttct acaatattta gtggattctt ctcattttca 2460
agatacacaa ttgatccata atcgaagtgg tatgtaagac agtgagttaa aagattatat 2520
tttttgggag acttccagtc aaattttctt agaagttttt ttggtccaga tgttcataaa 2580
gtcgccgctt tcatactttt tttaattttt taattggtgc actattaggt acctgttgga 2640
ggatgttaca ggcttattga tatccctatg agtaactgct tcaacagtgg tataaataag 2700
atatttgtga tgagtcagtt caattctact tcgcttaacc gccatattca tcgtacatac 2760

cttgaaggcg ggatcaactt tgctgatgga tctgtacagg tgatttacct catcttgttg 2820
atgtgtaata ctgtaattag gagtagattt gtgtggagag aataataaac agatgccgag 2880
attcttttct aaaagtctag atccaaaggc attgtggttc aaaacactat ggacttctac 2940
catttatgtc attactttgc cttaatgttc cattgaatgg ggcaaattat tgattctaca 3000
agtgtttaat taaaaactaa ttgttcatcc tgcaggtatt agcggctaca caaatgcctg 3060
aagagccagc tggatggttc cagggtacag cagactctat cagaaaattt atctgggtac 3120
tcgaggtagt tgatattttc tcgtttatga atgtccattc actcattcct gtagcattgt 3180
ttctttgtaa ttttgagttc tcctgtattt ctttaggatt attacagtca caaatccatt 3240
gacaacattg taatcttgag tggcgatcag ctttatcgga tgaattacat ggaacttgtg 3300
caggtatggt gttctcttgt tcctcatgtt tcacgtaatg tcctgatttt ggattaacca 3360
actacttttg gcatgcatta tttccagaaa catgtcgagg acgatgctga tatcactata 3420
tcatgtgctc ctgttgatga gaggtaatca gttgtttata tcatcctaat atgaatatgt 3480
catcttgtta tccaacacag gatgcatatg gtctaatctg ctttcctttt ttttcccttc 3540
ggaagccgag cttctaaaaa tgggctagtg aagattgatc atactggacg tgtacttcaa 3600
ttctttgaaa aaccaaaggg tgctgatttg aattctatgg ttagaaattc cttgtgtaat 3660
ccaattcttt tgttttcctt tctttcttga gatgaacccc tcttttagtt atttccatgg 3720
ataacctgta cttgacttat tcagaaatga ttttctattt tgctgtagaa tctgacacta 3780
aagctaatag cactgatgtt gcagagagtt gagaccaact tcctgagcta tgctatagat 3840
gatgcacaga aatatccata ccttgcatca atgggcattt atgtcttcaa gaaagatgca 3900
cttttagacc ttctcaagta atcactttcc tgtgacttat ttctatccaa ctcctagttt 3960
accttctaac agtgtcaatt cttaggtcaa aatatactca attacatgac tttggatctg 4020
aaatcctccc aagagctgta ctagatcata gtgtgcaggt aagtctgatc tgtctggagt 4080
atgtgttctg taaactgtaa attcttcatg tcaaaaagtt gtttttgttt ccagtttcca 4140
ctaccaatgc acgatttatg tattttcgct tccatgcatc atacatacta acaatacatt 4200
ttacgtattg tgttaggcat gcatttttac gggctattgg gaggatgttg gaacaatcaa 4260
atcattcttt gatgcaaact tggccctcac tgagcaggta ctctgtcatg tattctgtac 4320
tgcatatata ttacctggaa ttcaatgcat agaatgtgtt agaccatctt agttccatcc 4380
tgttttcttc aattagctta tcatttaata gttgttggct agaatttaaa cacaaattta 4440
cctaatatgt ttctctcttc agccttccaa gtttgatttt tacgatccaa aaacaccttt 4500
cttcactgca ccccgatgct tgcctccgac gcaattggac aagtgcaagg tatatgtctt 4560
actgagcaca attgttacct gagcaagatt ttgtgtactt gacttgttct cctccacaga 4620
tgaaatatgc atttatctca gatggttgct tactgagaga atgcaacatc gagcattctg 4680
tgattggagt ctgctcacgt gtcagctctg gatgtgaact caaggtacat actctgccaa 4740
tgtatctact cttgagtata ccatttcaac accaagcatc accaaatcac acagaacaat 4800
agcaacaaag ccttttagtt ccaagcaatt tagggtagcc tagagttgaa atctaacaaa 4860
acaaaagtca aagctctatc acgtggatag ttgttttcca tgcactctta tttaagctaa 4920
ttttttgggt atactacatc catttaatta ttgttttatt gcttcttccc tttgcctttc 4980
ccccattact atcgcgtctt aagatcatac tacgcactag tgtctttaga ggtctctggt 5040
ggacatgttc aaaccatctc aatcggtgtt ggacaagttt ttcttgaatt tgtgctacac 5100
ctaacctatc acgtatgtca tcgtttcaaa ctcgatcctt cctgtatcat cataaatcca 5160
atgcaacata cgcatttatg caacatttat ctgttgaaca tgtcatcttt ttgtaggtta 5220
acattatgca ccatacaatg tagcatgtct aatcatcatc ctataaaatt tacattttag 5280
cttatgtggt atcctcttgc cacttagaac accatatgct tgatgccatt tcatccaccc 5340
tgctttgatt ctatggctaa catcttcatt aatatcctcg cctctctgta tcattggtcc 5400
taaatatgga aatacattct ttctgggcac tacttgacct tccaaactaa cgtctccttt 5460
gctccttttct tgtgtgtagt agtaccgaag tcacatctca tatattcggt tttagttcta 5520
ctaagtcccg ggttcgatcc ccctcagggg tgaatttcgg gcttggtaaa aaaaatcccc 5580
tcgctgtgtc ccgcccgctc tcggggatcg atatcctgcg cgccaccctc cggctgggca 5640
ttgcagagtg agcagttgat cggctcgtta gtgatgggga gcggggttca agggttttct 5700
cggccgggac catgtttcgg tctcttaata taatgccggg agggcagtct ttccctcccc 5760
ggtcgagttt tagttctacc gagtctaaaa cctttggact ctagagtccc ctgtcacaac 5820

tcacaactct agttttctat ttacttctac ctagcgttta ttaatgatca ctatatcgtc 5880
tgtaaaaagc atacaccaat gtaatcccct tgtatgtccc ttgtaatatt atccatcaca 5940
agaaaaaaag gtaaggctca aagttgactt ttgatatagt cctattctaa tcgagaagtc 6000
atctgtatct tcgtctcttg ttcgaacact agtcacaaaa ttttttgtac atgttcttaa 6060
tgagtccaac gtaatattcc ttgatatttt gtcataagcc ctcatcaagt caatgaaaat 6120
cacgtgtagg tccttcattt gttccttata ctgctccatc acttgtctca ttaagaaaat 6180
ctctctcata gttaaccttt tggcatgaaa caaaatcaca cagaagttgt ttccttttt 6240
taagatccca cacaaaagag gtttgatcta aggaatctgg atccctgaca ggtttatcaa 6300
aatcctttgt gtttttctta aaactgaata ttcctccagc ttctagtatt gatgtaatat 6360
tcaatctgtt tagcaagtga acaccttggt tcttgttgtt actgtacccc ccccccccc 6420
ccccccccga ggcccagatt accacgacat gaatacaaga atattgaacc cagatctaga 6480
gtttgtttgt actgttgaaa atcggtgaca attcattttg ttattgcgct ttctgataac 6540
gacaggactc cgtgatgatg ggagcggaca cctatgaaac tgaagaagaa gcttcaaagc 6600
tactgttagc tgggaaggtc ccagttggaa taggaaggaa cacaaagata aggtgagtat 6660
ggatgtggaa ccaccggtta gttcccaaaa atatcactca ctgatacctg atggtatcct 6720
ctgattattt tcaggaactg tatcattgac atgaatgcta ggattgggaa gaacgtggtg 6780
atcacaaaca gtaaggtgag cgagcgcacc tacatgggtg cagaatcttg tgtgctcatc 6840
tatcctaatt cggtaattcc tatccagcgc tagtcttgtg accatggggc atgggttcga 6900
ctctgtgaca gggcatccaa gaggctgatc acccggaaga agggtactac ataaggtctg 6960
gaatcgtggt gatcttgaag aatgcaacca tcaacgatgg gtctgtcata tagatcggct 7020
gcgtgtgcgt ctacaaaaca agaacctaca atggtattgc atcgatggat cgtgtaacct 7080
tggtatggta agagccgctt gacagaaagt cgagcgttcg ggcaagatgc gtagtctggc 7140
atgctgttcc ttgaccattt gtgctgctag tatgtactgt tataagctgc cctagaagtt 7200
gcagcaaacc tttttatgaa cctttgtatt tccattacct gctttggatc aactatatct 7260
gtcatcctat atattactaa atttttacgt gtttttctaa ttcggtgctg cttttgggat 7320

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 522 Amino Acid
(B) TYPE: Amino Acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Amino Acid

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

Met Ser Ser Met Gln Phe Ser Ser Val Leu Pro Leu Glu Gly Lys Ala
1 5 10 15

Cys Ile Ser Pro Val Arg Arg Glu Gly Ser Ala Ser Glu Arg Leu Lys
20 25 30

Val Gly Asp Ser Ser Ser Ile Arg His Glu Arg Ala Ser Arg Arg Met
35 40 45

```
Cys Asn Gly Gly Arg Gly Pro Ala Ala Thr Gly Ala Gln Cys Val Leu
     50                  55                  60

Thr Ser Asp Ala Ser Pro Ala Asp Thr Leu Val Leu Arg Thr Ser Phe
 65                  70                  75                  80

Arg Arg Asn Tyr Ala Asp Pro Asn Glu Val Ala Ala Val Ile Leu Gly
                 85                  90                  95

Gly Gly Thr Gly Thr Gln Leu Phe Pro Leu Thr Ser Thr Arg Ala Thr
            100                 105                 110

Pro Ala Val Pro Ile Gly Gly Cys Tyr Arg Leu Ile Asp Ile Pro Met
        115                 120                 125

Ser Asn Cys Phe Asn Ser Gly Ile Asn Lys Ile Phe Val Met Thr Gln
    130                 135                 140

Phe Asn Ser Ala Ser Leu Asn Arg His Ile His Arg Thr Tyr Leu Gly
145                 150                 155                 160

Gly Gly Ile Asn Phe Thr Asp Gly Ser Val Glu Val Leu Ala Ala Thr
                165                 170                 175

Gln Met Pro Gly Glu Ala Ala Gly Trp Phe Arg Gly Thr Ala Asp Ala
            180                 185                 190

Val Arg Lys Phe Ile Trp Val Leu Glu Asp Tyr Tyr Lys Asn Lys Ser
        195                 200                 205


Ile Glu His Ile Leu Ile Leu Ser Gly Asp Gln Leu Tyr Arg Met Asp
    210                 215                 220

Tyr Met Glu Leu Val Gln Lys His Val Asp Asp Asn Ala Asp Ile Thr
225                 230                 235                 240

Leu Ser Cys Ala Pro Val Gly Glu Ser Arg Ala Ser Glu Tyr Gly Leu
                245                 250                 255

Val Lys Phe Asp Ser Ser Gly Arg Val Val Gln Phe Ser Glu Lys Pro
            260                 265                 270

Lys Gly Asp Asp Leu Glu Ala Met Lys Val Asp Thr Ser Phe Leu Asn
        275                 280                 285

Phe Ala Ile Asp Asp Pro Ala Lys Tyr Pro Tyr Ile Ala Ser Met Gly
    290                 295                 300
```

Val Tyr Val Phe Lys Arg Asp Val Leu Leu Asn Leu Leu Lys Ser Arg
305 310 315 320

Tyr Ala Glu Leu His Asp Phe Gly Ser Glu Ile Leu Pro Arg Ala Leu
325 330 335

His Asp His Asn Val Gln Ala Tyr Val Phe Thr Asp Tyr Trp Glu Asp
340 345 350

Ile Gly Thr Ile Arg Ser Phe Phe Asp Ala Asn Met Ala Leu Cys Glu
355 360 365

Gln Pro Pro Lys Phe Glu Phe Tyr Asp Pro Lys Thr Pro Phe Phe Thr
370 375 380

Ser Pro Arg Tyr Leu Pro Pro Thr Lys Ser Asp Lys Cys Arg Ile Lys
385 390 395 400

Glu Ala Ile Ile Ser His Gly Cys Phe Leu Arg Glu Cys Lys Ile Glu
405 410 415

His Ser Ile Ile Gly Val Arg Ser Arg Leu Asn Ser Gly Ser Glu Leu
420 425 430

Lys Asn Ala Met Met Met Gly Ala Asp Ser Tyr Glu Thr Glu Asp Glu
435 440 445

Ile Ser Arg Leu Met Ser Glu Gly Lys Val Pro Ile Gly Val Gly Glu
450 455 460

Asn Thr Lys Ile Ser Asn Cys Ile Ile Asp Met Asn Ala Arg Ile Gly
465 470 475 480

Arg Asp Val Val Ile Ser Asn Lys Glu Gly Val Gln Glu Ala Asp Arg
485 490 495

Pro Glu Glu Gly Tyr Tyr Ile Arg Ser Gly Ile Val Val Ile Gln Lys
500 505 510

Asn Ala Thr Ile Lys Asp Gly Thr Val Val
515 520

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2084 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Nucleic Acid

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

ggctcgaccg tctctcctct ctctctctct ctccccctca ccaccacatc agcggggcca 60
cccccggcct cttcttccct gcatttgatt gatccgtcgc ttgcccggtc gcccgtcgat 120
ccgtttaccc gccgtccgcc cgtccccggc agccgcagcc gcaccgcagt tgcaggtgga 180
ctgcgaatgt catcgatgca gttcagcagc gtgctgcccc tggagggcaa agcgtgcatc 240
tcccccgtca ggagagaggg ctccgcctcc gagcgcctca aggtcgggga cagcagcagc 300
atcaggcacg agagggcctc caggaggatg tgcaacggtg gcaggggccc cgccgccacc 360
ggcgcgcagt gcgtgctcac ctccgacgcc agcccagcag acacccttgt tctccggacg 420
tccttccgca ggaattacgc cgatccgaac gaggtcgcgg ccgtcatact cggcggtggc 480
accgggactc agctcttccc tctcacaagc acaagggcca cacctgctgt tcctattgga 540
ggatgttaca ggctcatcga tattcccatg agcaactgct tcaatagtgg catcaacaag 600
atattcgtca tgacgcagtt caactcggcc tctcttaatc gtcacattca ccgcacctac 660
ctcggcgggg gaatcaattt cactgatgga tctgttgagg tattggccgc gacgcaaatg 720
cccggggagg ctgctggatg gttccgcgga acagcggatg ccgtcagaaa atttatctgg 780
gtgcttgagg actattataa gaataaatcc atagagcaca ttttgatctt gtcgggcgat 840
cagctttatc gcatggatta catggagctt gtgcagaaac atgtggatga caatgctgac 900
attactttat catgtgcccc tgttggagag agccgggcat ctgagtacgg gctagtgaag 960
ttcgacagtt caggccgtgt ggtccaattt tctgagaagc caaagggtga cgatctggaa 1020
gcgatgaaag tggacaccag ttttctcaat ttcgccatcg acgaccctgc taaatatcca 1080
tacattgctt ctatgggagt ctatgtcttc aaaagagatg ttctgctcaa ccttctaaag 1140
tcaagatacg cagaactaca tgactttggg tctgaaatcc tcccgagagc tctgcatgac 1200
cacaatgtac aggcttatgt cttcactgac tactgggagg acattggaac aatcagatcc 1260
ttcttcgatg cgaacatggc cctctgcgag cagcccccaa agttcgagtt ttatgatccc 1320
aaaactccct tcttcacctc gcctcgatac ttgccaccaa caaagtcaga caagtgcagg 1380
atcaaagaag cgatcatttc gcacggctgc ttcttgcgtg aatgcaaaat cgagcactcc 1440
atcatcggcg ttcgttcacg cctaaactcc ggaagcgaac tcaagaacgc gatgatgatg 1500
ggtgcggatt cgtacgagac cgaggatgag atctcgaggc tgatgtccga gggcaaggtc 1560
cccatcggcg tcggggagaa cacaaagatc agcaactgca tcatcgacat gaacgcgagg 1620
ataggaaggg acgtggtcat ctcaaacaag gagggagtgc aagaagccga caggccggag 1680
gaggggtact acatcaggtc cgggatcgtg gtgatccaga agaacgcgac catcaaggac 1740
ggcaccgtcg tgtagagcac cgccgggtcg gcgtgacggg ttctgcgaca acctctgcgc 1800
tgcgttgatc gtcgtcgtcg tctcgaggcc cgggagggac tgaagaagtg acccggggac 1860
gggaggcgtt tgaagcttga atgactgaga aggcgcgcgc gggcagcatt agtagtaagt 1920
agtagtaagg agcagtggaa caaagtaata gtcgttcgtt tttcccctgt aataaataag 1980
aggctgtgtg ttgaggtaaa gaagtggcag cgagcaaaca aactcccggg ggatgttcgt 2040
gtaaataaaa ctctatctag acctgtgaaa ttttcaaaaa aaaa 2084

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 503 Amino Acid
(B) TYPE: Amino Acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Amino Acid

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

```
Met Asp Leu Arg Val Ala Ala Pro Ala Ser Val Ala Ala Ala Ala Arg
 1               5                  10                  15

Arg Gly Ala Leu Gly Cys Ala Arg Val Arg Pro Leu Gln Gly Arg Arg
            20                  25                  30

Gln Cys Arg Pro Ser Val Arg Val Ser Val Ala Thr Thr Glu Ser Ala
        35                  40                  45

Ala Ala Ala Ala Ala Val Ser Ala Ser Ala Asp Glu Asp Ala Glu Thr
    50                  55                  60

Thr Asn Pro Arg Thr Val Val Ala Val Ile Leu Gly Gly Gly Ala Gly
65                  70                  75                  80

Thr Arg Leu Phe Pro Leu Thr Lys Arg Arg Ala Lys Pro Ala Val Pro
                85                  90                  95

Ile Gly Gly Ala Tyr Arg Leu Ile Asp Val Pro Met Ser Asn Cys Ile
            100                 105                 110

Asn Ser Gly Ile Asn Lys Val Tyr Val Leu Thr Gln Phe Asn Ser Ala
        115                 120                 125

Ser Leu Asn Arg His Leu Phe Arg Ala Tyr Asn Phe Ser Asn Gly Val
    130                 135                 140

Gly Phe Gly Asp Gly Phe Val Glu Val Leu Ala Ala Thr Gln Arg Pro
145                 150                 155                 160

Gly Ser Glu Gly Lys Arg Trp Phe Gln Gly Thr Ala Asp Ala Val Arg
                165                 170                 175

Gln Phe Ala Trp Leu Phe Asp Asp Ala Lys Ser Lys Asp Ile Glu Asp
            180                 185                 190
```

```
Val Leu Ile Leu Ser Gly Asp His Leu Tyr Arg Met Asp Tyr Met Asp
        195                 200                 205

Phe Val Gln Ser His Arg Gln Arg Asp Ala Gly Ile Ser Ile Cys Cys
    210                 215                 220

Leu Pro Ile Asp Asp Ser Arg Ala Ser Asp Phe Gly Leu Met Lys Ile
225                 230                 235                 240

Asp Asp Thr Gly Arg Val Ile Ser Phe Ser Glu Lys Pro Lys Gly Ala
                245                 250                 255

Asp Leu Lys Ala Met Gln Val Asp Thr Thr Leu Leu Gly Leu Pro Lys
            260                 265                 270

Glu Glu Ala Glu Lys Lys Pro Tyr Ile Ala Ser Met Gly Val Tyr Ile
        275                 280                 285

Phe Lys Lys Glu Ile Leu Leu Asn Leu Leu Arg Trp Arg Phe Pro Thr
    290                 295                 300

Ala Asn Asp Phe Gly Ser Glu Ile Ile Pro Ala Ala Ala Arg Glu Ile
305                 310                 315                 320

Asn Val Lys Ala Tyr Leu Phe Asn Asp Tyr Trp Glu Asp Ile Gly Thr
                325                 330                 335

Ile Lys Ser Phe Phe Glu Ala Asn Leu Ala Leu Ala Glu Gln Pro Ser
            340                 345                 350

Lys Phe Ser Phe Tyr Asp Ala Ser Lys Pro Met Tyr Thr Ser Arg Arg
        355                 360                 365

Asn Leu Pro Pro Ser Met Ile Ser Gly Ser Lys Ile Thr Asp Ser Ile
    370                 375                 380

Ile Ser His Gly Cys Phe Leu Asp Lys Cys Arg Val Glu His Ser Val
385                 390                 395                 400

Val Gly Ile Arg Ser Arg Ile Gly Ser Asn Val His Leu Lys Asp Thr
                405                 410                 415

Val Met Leu Gly Ala Asp Phe Tyr Glu Thr Asp Ala Glu Arg Gly Asp
            420                 425                 430

Gln Leu Ala Glu Gly Lys Val Pro Ile Gly Ile Gly Glu Asn Thr Ser
        435                 440                 445
```

Ile Gln Asn Cys Ile Ile Asp Lys Asn Ala Arg Ile Gly Lys Asn Val
450 455 460

Thr Ile Ala Asn Thr Glu Gly Val Gln Glu Ser Asp Arg Thr Ser Glu
465 470 475 480

Gly Phe His Ile Arg Ser Gly Ile Thr Val Val Leu Lys Asn Ser Val
485 490 495

Ile Ala Asp Gly Leu Val Ile
500

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1599 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Nucleic Acid

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

caaaccaagc tcccggccgg gcggattcat tggttgcgcg atggacctgc gcgtcgctgc 60
ccccgcctcc gtggccgccg ccgcgcgccg gggcgcgctc ggatgcgcgc gcgtccggcc 120
gctccaaggc cgccgccagt gccgcccctc ggtgcgcgtc tccgtcgcga ccacagagtc 180
cgccgccgcc gccgccgccg tatccgcctc ggcggacgag gacgcggaga cgacgaaccc 240
gaggacggtg gtggcggtca tcctcggcgg cggcgccggg accaggctct tccctctcac 300
caagcgcagg gccaagccgg cggtgccgat aggcggcgcg tacaggctca tcgacgtgcc 360
catgagcaac tgcatcaaca gcgggatcaa caaggtgtac gtcctcaccc agttcaactc 420
cgcctccctc aaccgccacc tcttcagggc ctacaacttc agcaatggcg tcggcttcgg 480
agacggtttc gtcgaggttc tggcagcaac tcagaggcct ggatcggagg gaaagaggtg 540
gtttcagggt accgccgacg cggttcggca attcgcttgg cttttcgatg atgccaaatc 600
taaagacata gaggatgtat tgattctttc gggcgatcac ctctaccgta tggactacat 660
ggactttgtt cagagtcatc ggcagagaga tgcaggcatc agcatctgtt gcttgcctat 720
tgacgacagc cgggcgtctg attttggtct aatgaagata gacgacacag gaagagttat 780
ttcatttagt gaaaaaccga aaggagctga cttgaaagca atgcaagttg acaccactct 840
cctcggctta ccgaaggagg aagcagaaaa gaaaccatac atagcttcaa tgggggtata 900
catattcaag aaagagatac ttctaaatct tttgagatgg cgttttccca ctgcaaatga 960
ttttggctct gaaataattc ctgctgcagc aagagagatt aatgtaaagg catatctttt 1020
caatgattac tgggaagaca ttggaactat taaatccttc ttcgaagcaa atcttgccct 1080
tgctgaacag ccttcaaagt tcagcttcta tgatgctagc aaaccgatgt acacatcacg 1140
gagaaaccta ccaccatcta tgatcagcgg tagtaagatc actgattcga tcatttccca 1200
cggatgtttc ttggataaat gcagggtaga gcacagtgtc gttggaatcc gttctcggat 1260
aggctccaac gtacacctca aggatacggt aatgctcggt gctgatttct atgaaactga 1320
tgcggaaaga ggtgaccagc tggctgaagg aaaggttcca attgggatcg gggagaacac 1380
ttcgattcaa aactgcatca ttgacaagaa tgcgaggata gggaagaacg tgaccattgc 1440

taacaccgag ggtgtacagg aatcagacag gacatcagaa ggcttccaca tccggtccgg 1500
catcacggtt gtgctgaaga actcagtgat cgcggatgga ttagtcatat gatctgaaag 1560
aagtgactct ccagtctcca gttaagcaag agaaataaa 1599

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 470 Amino Acid
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Amino Acid

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

Asn Lys Ile Lys Pro Gly Val Ala Tyr Ser Val Ile Thr Thr Glu Asn
1 5 10 15

Asp Thr Gln Thr Val Phe Val Asp Met Pro Arg Leu Glu Arg Arg Arg
20 25 30

Ala Asn Pro Lys Asp Val Ala Ala Val Ile Leu Gly Gly Gly Glu Gly
35 40 45

Thr Lys Leu Phe Pro Leu Thr Ser Arg Thr Ala Thr Pro Ala Val Pro
50 55 60

Val Gly Gly Cys Tyr Arg Leu Ile Asp Ile Pro Met Ser Asn Cys Ile
65 70 75 80

Asn Ser Ala Ile Asn Lys Ile Phe Val Leu Thr Gln Tyr Asn Ser Ala
85 90 95

Pro Leu Asn Arg His Ile Ala Arg Thr Tyr Phe Gly Asn Gly Val Ser
100 105 110

Phe Gly Asp Gly Phe Val Glu Val Leu Ala Ala Thr Gln Thr Pro Gly
115 120 125

Glu Ala Gly Lys Lys Trp Phe Gln Gly Thr Ala Asp Ala Val Arg Lys
130 135 140

Phe Ile Trp Val Phe Glu Asp Ala Lys Asn Lys Asn Ile Glu Asn Ile
145 150 155 160

Val Val Leu Ser Gly Asp His Leu Tyr Arg Met Asp Tyr Met Glu Leu
165 170 175

Val Gln Asn His Ile Asp Arg Asn Ala Asp Ile Thr Leu Ser Cys Ala
180 185 190

Pro Ala Glu Asp Ser Arg Ala Ser Asp Phe Gly Leu Val Lys Ile Asp
195 200 205

Ser Arg Gly Arg Val Val Gln Phe Ala Glu Lys Pro Lys Gly Phe Asp
210 215 220

Leu Lys Ala Met Gln Val Asp Thr Thr Leu Val Gly Leu Ser Pro Gln
225 230 235 240

Asp Ala Lys Lys Ser Pro Tyr Ile Ala Ser Met Gly Val Tyr Val Phe
245 250 255

Lys Thr Asp Val Leu Leu Lys Leu Leu Lys Trp Ser Tyr Pro Thr Ser
260 265 270

Asn Asp Phe Gly Ser Glu Ile Ile Pro Ala Ala Ile Asp Asp Tyr Asn
275 280 285

Val Gln Ala Tyr Ile Phe Lys Asp Tyr Trp Glu Asp Ile Gly Thr Ile
290 295 300

Lys Ser Phe Tyr Asn Ala Ser Leu Ala Leu Thr Gln Glu Phe Pro Glu
305 310 315 320

Phe Gln Phe Tyr Asp Pro Lys Thr Pro Phe Tyr Thr Ser Pro Arg Phe
325 330 335

Leu Pro Pro Thr Lys Ile Asp Asn Cys Lys Ile Lys Asp Ala Ile Ile
340 345 350

Ser His Gly Cys Phe Leu Arg Asp Cys Ser Val Glu His Ser Ile Val
355 360 365

Gly Glu Arg Ser Arg Leu Asp Cys Gly Val Glu Leu Lys Asp Thr Phe
370 375 380

Met Met Gly Ala Asp Tyr Tyr Gln Thr Glu Ser Glu Ile Ala Ser Leu
385 390 395 400

Leu Ala Glu Gly Lys Val Pro Ile Gly Ile Gly Glu Asn Thr Lys Ile
405 410 415

Arg Lys Cys Ile Ile Asp Lys Asn Ala Lys Ile Gly Lys Asn Val Ser
420 425 430

Ile Ile Asn Lys Asp Gly Val Gln Glu Ala Asp Arg Pro Glu Glu Gly
435 440 445

Phe Tyr Ile Arg Ser Gly Ile Ile Ile Ile Leu Glu Lys Ala Thr Ile
450 455 460

Arg Asp Gly Thr Val Ile
465 470

(2) INFORMATION FOR SEQ ID NO:10

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1413 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Nucleic Acid

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

aacaagatca aacctggggt tgcttactct gtgatcacta ctgaaaatga cacacagact 60
gtgttcgtag atatgccacg tcttgagaga cgccgggcaa atccaaagga tgtggctgca 120
gtcatactgg gaggaggaga agggaccaag ttattcccac ttacaagtag aactgcaacc 180
cctgctgttc cggttggagg atgctacagg ctaatagaca tcccaatgag caactgtatc 240
aacagtgcta ttaacaagat ttttgtgctg acacagtaca attctgctcc cctgaatcgt 300
cacattgctc gaacatattt tggcaatggt gtgagctttg gagatggatt tgtcgaggta 360
ctagctgcaa ctcagacacc cggggaagca ggaaaaaaat ggtttcaagg aacagcagat 420
gctgttagaa aatttatatg ggtttttgag gacgctaaga acaagaatat tgaaaatatc 480
gttgtactat ctggggatca tctttatagg atggattata tggagttggt gcagaaccat 540
attgacagga atgctgatat tactctttca tgtgcaccag ctgaggacag ccgagcatca 600
gattttgggc tggtcaagat tgacagcaga ggcagagtag tccagtttgc tgaaaaacca 660
aaaggttttg atcttaaagc aatgcaagta gatactactc ttgttggatt atctccacaa 720
gatgcgaaga aatccccctа tattgcttca atgggagttt atgtattcaa gacagatgta 780
ttgttgaagc tcttgaaatg gagctatccc acttctaatg attttggctc tgaaattata 840
ccagcagcta ttgacgatta caatgtccaa gcatacattt tcaaagacta ttgggaagac 900
attggaacaa ttaaatcgtt ttataatgct agcttggcac tcacacaaga gtttccagag 960
ttccaatttt acgatccaaa aacacctttt tacacatctc ctaggttcct tccaccaacc 1020
aagatagaca attgcaagat taaggatgcc ataatctctc atggatgttt cttgcgagat 1080
tgttctgtgg aacactccat agtgggtgaa agatcgcgct tagattgtgg tgttgaactg 1140
aaggatactt tcatgatggg agcagactac taccaaacag aatctgagat tgcctccctg 1200
ttagcagagg ggaaagtacc gattggaatt ggggaaaata caaaaataag gaaatgtatc 1260
attgacaaga acgcaaagat aggaaagaat gtttcaatca taaataaaga cggtgttcaa 1320

gaggcagacc gaccagagga aggattctac atacgatcag ggataatcat tatattagag 1380
aaagccacaa ttagagatgg aacagtcata tga 1413

(2) INFORMATION FOR SEQ ID NO:11

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 444 Amino Acid
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Amino Acid

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

Asn Ser Gln Thr Cys Leu Asp Pro Glu Ala Ser Arg Ser Val Leu Gly
1 5 10 15

Ile Ile Leu Gly Gly Gly Ala Gly Thr Arg Leu Tyr Pro Leu Thr Lys
20 25 30

Lys Arg Ala Lys Pro Ala Val Pro Leu Gly Ala Asn Tyr Arg Leu Ile
35 40 45

Asp Ile Pro Val Ser Asn Cys Leu Asn Ser Asn Ile Ser Lys Ile Tyr
50 55 60

Val Leu Thr Gln Phe Asn Ser Ala Ser Leu Asn Arg His Leu Ser Arg
65 70 75 80

Ala Tyr Ala Ser Asn Leu Gly Gly Tyr Lys Asn Glu Gly Phe Val Glu
85 90 95

Val Leu Ala Ala Gln Gln Ser Pro Glu Asn Pro Asp Trp Phe Gln Gly
100 105 110

Thr Ala Asp Ala Val Arg Gln Tyr Leu Trp Leu Phe Glu Glu His Asn
115 120 125

Val Met Glu Phe Leu Ile Leu Ala Gly Asp His Leu Tyr Arg Met Asp
130 135 140

Tyr Glu Arg Phe Ile Gln Ala His Arg Glu Thr Asp Ala Asp Ile Thr
145 150 155 160

Val Ala Ala Leu Pro Met Asp Glu Lys Arg Ala Thr Ala Phe Gly Leu
165 170 175

Met Lys Ile Asp Glu Glu Gly Arg Ile Ile Glu Phe Ala Glu Lys Pro
180 185 190

Lys Gly Glu Gln Leu Gln Ala Met Lys Val Asp Thr Thr Ile Leu Gly
195 200 205

Leu Asp Asp Glu Arg Ala Lys Glu Met Pro Tyr Ile Ala Ser Met Gly
210 215 220

Ile Tyr Val Ile Ser Lys Asp Val Met Leu Asn Leu Leu Arg Asp Lys
225 230 235 240

Phe Pro Gly Ala Asn Asp Phe Gly Ser Glu Val Ile Pro Gly Ala Thr
245 250 255

Ser Ile Gly Leu Thr Val Gln Ala Tyr Leu Tyr Asp Gly Tyr Trp Glu
260 265 270

Asp Ile Gly Thr Ile Glu Ala Phe Tyr Asn Ala Asn Leu Gly Ile Thr
275 280 285

Lys Lys Pro Val Pro Asp Phe Ser Phe Tyr Asp Arg Ser Ser Pro Ile
290 295 300

Tyr Thr Gln Pro Arg Tyr Leu Pro Pro Ser Lys Met Leu Asp Ala Asp
305 310 315 320

Ile Thr Asp Ser Val Ile Gly Glu Gly Cys Val Ile Lys Asn Cys Lys
325 330 335

Ile His His Ser Val Ile Gly Leu Arg Ser Cys Ile Ser Glu Gly Ala
340 345 350

Ile Ile Glu Asp Thr Leu Leu Met Gly Ala Asp Tyr Tyr Glu Thr Asp
355 360 365

Ala Asp Arg Lys Leu Leu Ala Ala Lys Gly Ser Val Val Leu Gly Ile
370 375 380

Gly Gln Asn Ser His Ile Lys Arg Ala Ile Ile Asp Lys Asn Ala Arg
385 390 395 400

Ile Gly Asp Asn Val Lys Ile Ile Asn Ser Asp Asn Val Gln Glu Ala
405 410 415

Ala Arg Glu Thr Asp Gly Tyr Phe Ile Lys Ser Gly Ile Val Thr Val
420 425 430

Ile Lys Asp Ala Leu Ile Pro Ser Gly Thr Val Ile
435 440

(2) INFORMATION FOR SEQ ID NO:12

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1601 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Nucleic Acid

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

gaattcgcag acttgtcttg atcccgaagc tagccgcagt gttctaggta ttattcttgg 60
aggtggtgct ggtacacgtt tgtacccact tactaagaaa agagcaaagc ctgcggtgcc 120
gcttggtgcc aattataggc ttattgacat cccagtgagc aactgtttaa acagtaacat 180
ttccaaaata tatgttctta ctcaattcaa ttctgcatcc ctcaaccgtc atctttcacg 240
ggcgtatgct agtaatctgg gaggttacaa gaacgaaggg tttgtagaag ttcttgccgc 300
tcagcaaagt cctgagaatc ctgactggtt ccagggtaca gctgatgctg tcaggcaata 360
cttgtggcta ttcgaagagc acaatgtcat ggaattcttg attcttgctg gtgatcattt 420
atatcgtatg gattatgaaa gatttatcca agctcataga gaaactgatg cagacattac 480
tgttgctgca ctgccgatgg atgaaaagcg tgctactgca tttggtttaa tgaaaattga 540
cgaagaagga aggattattg agtttgctga gaaacccaaa ggagaacaat tgcaagctat 600
gaaggttgat accactatct tgggtctgga tgatgagaga gcaaaagaaa tgccatacat 660
agctagtatg ggaatttatg ttatcagcaa agatgtaatg ctaaatttgc ttcgggataa 720
atttcctggt gctaacgatt ttggcagtga agttattcca ggagcgactt ccatagggct 780
aacagttcaa gcttacttgt atgatggtta ttgggaagat atcggtacaa ttgaggcttt 840
ttacaatgca aacttgggta tcaccaaaaa gccggtgcca gatttcagct tctatgatcg 900
ttcgtctcca atatatacac aacctcggta tttgcctcct tcaaagatgc ttgacgctga 960
tataactgac agtgtcatcg gtgaaggctg tgtgattaag aactgcaaga ttcatcattc 1020
tgttattggt ctccgatctt gcatttcgga gggtgcaatt attgaggata cgctgttaat 1080
gggagctgat tattatgaga ctgatgctga tcggaaactc ctagctgcta agggtagcgt 1140
agtacttggc attgggcaga atagccatat aaaaagagca ataattgata agaatgctcg 1200
gattggggac aatgtcaaga ttatcaacag tgacaatgta caagaagcag caagagaaac 1260
agacggttac ttcatcaaga gtggaattgt aactgttatc aaagacgcct tgattccgag 1320
tggaaccgta atctaaaatg gtgcaaatac cattttcctc tttctttttt tttttctttt 1380
ttttttttcca agtacttcaa gtgtcccact acaaggcctc tctacctctt gtttcagtct 1440
ttgagtatgt aaatattttg atcttttgcc gtattcggtg cgcggtagtg ttataagttg 1500
gacaggggggt atgaggaacg gcattgtaaa catattttga ccactttagg cactataaaa 1560
tttgtttgat gttataccaa aataaataat taccggaatt c 1601

CLAIMS:

1. A polynucleotide encoding a mutant large subunit of a plant ADP-glucose pyrophosphorylase polypeptide, or a biologically-active fragment or a biologically active variant of said mutant polypeptide, wherein said mutant polypeptide comprises an amino acid mutation in the amino acid sequence of said polypeptide, and wherein when said mutant polypeptide is expressed to form a mutant ADP-glucose pyrophosphorylase enzyme, said mutant enzyme exhibits increased heat stability relative to a wild type ADP-glucose pyrophosphorylase enzyme, and wherein said mutant polypeptide encoded by said polynucleotide comprises an amino acid mutation wherein the amino acid corresponding to position 333 in the amino acid sequence of the wild type large subunit of ADP-glucose pyrophosphorylase polypeptide of maize endosperm is replaced by an amino acid that confers said increased heat stability on said mutant enzyme.

2. The polynucleotide according to claim 1, wherein said amino acid corresponding to position 333 is replaced with a tyrosine.

3. The polynucleotide according to claim 1, wherein said polynucleotide encodes a large subunit of ADP-glucose pyrophosphorylase of maize endosperm wherein the histidine amino acid at position 333 in the amino acid sequence of the wild type large subunit of ADP-glucose pyrophosphorylase polypeptide of maize endosperm is replaced with a tyrosine.

4. A polynucleotide encoding a mutant large subunit of a plant ADP-glucose pyrophosphorylase polypeptide, or a biologically-active fragment or a biologically active variant of said mutant polypeptide, wherein said mutant polypeptide comprises an amino acid mutation in the amino acid sequence of said polypeptide, and wherein when said mutant polypeptide is expressed to form a mutant ADP-glucose pyrophosphorylase enzyme, said mutant enzyme exhibits increased heat stability relative to a wild type ADP-glucose pyrophosphorylase enzyme, and wherein said mutant

polypeptide encoded by said polynucleotide comprises an amino acid mutation wherein the amino acid corresponding to position 177 in the amino acid sequence of the wild type large subunit of ADP-glucose pyrophosphorylase polypeptide of maize endosperm is replaced by an amino acid that confers said increased heat stability on said mutant enzyme.

5. The polynucleotide according to claim 4, wherein said amino acid corresponding to position 177 is replaced with a proline.

6. The polynucleotide according to claim 4, wherein said polynucleotide encodes a large subunit of ADP-glucose pyrophosphorylase of maize endosperm wherein the alanine amino acid at position 177 in the amino acid sequence of the wild type large subunit of ADP-glucose pyrophosphorylase polypeptide of maize endosperm is replaced with a proline.

7. A polynucleotide encoding a mutant large subunit of a plant ADP-glucose pyrophosphorylase polypeptide, or a biologically-active fragment or a biologically active variant of said mutant polypeptide, wherein said mutant polypeptide comprises an amino acid mutation in the amino acid sequence of said polypeptide, and wherein when said mutant polypeptide is expressed to form a mutant ADP-glucose pyrophosphorylase enzyme, said mutant enzyme exhibits increased heat stability relative to a wild type ADP-glucose pyrophosphorylase enzyme, and wherein said mutant polypeptide encoded by said polynucleotide comprises an amino acid mutation wherein the amino acid corresponding to position 400 in the amino acid sequence of the wild type large subunit of ADP-glucose pyrophosphorylase polypeptide of maize endosperm is replaced by an amino acid that confers said increased heat stability on said mutant enzyme.

8. The polynucleotide according to claim 7, wherein said amino acid corresponding to position 400 is replaced with a histidine.

9. The polynucleotide according to claim 7, wherein said polynucleotide encodes a large subunit of ADP-glucose pyrophosphorylase of maize endosperm wherein the aspartic acid amino acid at position 400 in the amino acid sequence of the wild type large subunit of ADP-glucose pyrophosphorylase polypeptide of maize endosperm is replaced with a histidine.

10. A polynucleotide encoding a mutant large subunit of a plant ADP-glucose pyrophosphorylase polypeptide, or a biologically-active fragment or a biologically active variant of said mutant polypeptide, wherein said mutant polypeptide comprises an amino acid mutation in the amino acid sequence of said polypeptide, and wherein when said mutant polypeptide is expressed to form a mutant ADP-glucose pyrophosphorylase enzyme, said mutant enzyme exhibits increased heat stability relative to a wild type ADP-glucose pyrophosphorylase enzyme, and wherein said mutant polypeptide encoded by said polynucleotide comprises an amino acid mutation wherein the amino acid corresponding to position 454 in the amino acid sequence of the wild type large subunit of ADP-glucose pyrophosphorylase polypeptide of maize endosperm is replaced by an amino acid that confers said increased heat stability on said mutant enzyme.

11. The polynucleotide according to claim 10, wherein said amino acid corresponding to position 454 is replaced with an isoleucine.

12. The polynucleotide according to claim 10, wherein said polynucleotide encodes a large subunit of ADP-glucose pyrophosphorylase of maize endosperm wherein the valine amino acid at position 454 in the amino acid sequence of the wild type large subunit of ADP-glucose pyrophosphorylase polypeptide of maize endosperm is replaced with an isoleucine.

13. A polynucleotide encoding a mutant large subunit of a plant ADP-glucose pyrophosphorylase polypeptide, or a biologically-active fragment or a biologically active variant of said mutant polypeptide, wherein said mutant polypeptide

comprises an amino acid mutation in the amino acid sequence of said polypeptide, and wherein when said mutant polypeptide is expressed to form a mutant ADP-glucose pyrophosphorylase enzyme, said mutant enzyme exhibits increased heat stability relative to a wild type ADP-glucose pyrophosphorylase enzyme, and wherein said mutant polypeptide encoded by said polynucleotide comprises an amino acid mutation wherein the amino acid corresponding to position 104 in the amino acid sequence of the wild type large subunit of ADP-glucose pyrophosphorylase polypeptide of maize endosperm is replaced by an amino acid that confers said increased heat stability on said mutant enzyme.

14. The polynucleotide according to claim 13, wherein said amino acid corresponding to position 104 is replaced with a threonine.

15. The polynucleotide according to claim 13, wherein said polynucleotide encodes a large subunit of ADP-glucose pyrophosphorylase of maize endosperm wherein the arginine amino acid at position 104 in the amino acid sequence of the wild type large subunit of ADP-glucose pyrophosphorylase polypeptide of maize endosperm is replaced with an threonine.

16. A polynucleotide encoding a mutant large subunit of a plant ADP-glucose pyrophosphorylase polypeptide, or a biologically-active fragment or a biologically active variant of said mutant polypeptide, wherein said mutant polypeptide comprises an amino acid mutation in the amino acid sequence of said polypeptide, and wherein when said mutant polypeptide is expressed to form a mutant ADP-glucose pyrophosphorylase enzyme, said mutant enzyme exhibits increased heat stability relative to a wild type ADP-glucose pyrophosphorylase enzyme, and wherein said mutant polypeptide encoded by said polynucleotide comprises an amino acid mutation wherein the amino acid corresponding to position 460 in the amino acid sequence of the wild type large subunit of ADP-glucose pyrophosphorylase polypeptide of maize endosperm is replaced by an amino acid that confers said increased heat stability on said mutant enzyme.

17. The polynucleotide according to claim 16, wherein said amino acid corresponding to position 460 is replaced with a isoleucine.

18. The polynucleotide according to claim 16, wherein said polynucleotide encodes a large subunit of ADP-glucose pyrophosphorylase of maize endosperm wherein the threonine amino acid at position 460 in the amino acid sequence of the wild type large subunit of ADP-glucose pyrophosphorylase polypeptide of maize endosperm is replaced with an isoleucine.

19. A polynucleotide encoding a mutant large subunit of a plant ADP-glucose pyrophosphorylase polypeptide, or a biologically-active fragment or a biologically active variant of said mutant polypeptide, wherein said mutant polypeptide comprises an amino acid mutation in the amino acid sequence of said polypeptide, and wherein when said mutant polypeptide is expressed to form a mutant ADP-glucose pyrophosphorylase enzyme, said mutant enzyme exhibits increased heat stability relative to a wild type ADP-glucose pyrophosphorylase enzyme, and wherein said mutant polypeptide encoded by said polynucleotide comprises an amino acid mutation wherein the amino acid corresponding to position 216 in the amino acid sequence of the wild type large subunit of ADP-glucose pyrophosphorylase polypeptide of maize endosperm is replaced by an amino acid that confers said increased heat stability on said mutant enzyme.

20. The polynucleotide according to claim 19, wherein said amino acid corresponding to position 216 is replaced with a proline.

21. The polynucleotide according to claim 19, wherein said polynucleotide encodes a large subunit of ADP-glucose pyrophosphorylase of maize endosperm wherein the arginine amino acid at position 216 in the amino acid sequence of the wild type large subunit of ADP-glucose pyrophosphorylase polypeptide of maize endosperm is replaced with a proline.

22. The polynucleotide according to any of claims 1 to 3, wherein said mutant polypeptide encoded by said polynucleotide further comprises a second amino acid mutation wherein the amino acid corresponding to position 460 in the amino acid sequence of the wild type large subunit of ADP-glucose pyrophosphorylase polypeptide of maize endosperm is replaced by an amino acid that confers said increased heat stability on said mutant enzyme.

23. The polynucleotide according to claim 22, wherein said amino acid corresponding to position 460 is replaced with an isoleucine.

24. The polynucleotide according to claim 22, wherein said polynucleotide encodes a large subunit of ADP-glucose pyrophosphorylase of maize endosperm wherein the threonine amino acid at position 460 in the amino acid sequence of the wild type large subunit of ADP-glucose pyrophosphorylase polypeptide of maize endosperm is replaced with an isoleucine.

25. The polynucleotide according to any of claims 1 to 3, wherein said mutant polypeptide encoded by said polynucleotide further comprises a second amino acid mutation wherein the amino acid corresponding to position 216 in the amino acid sequence of the wild type large subunit of ADP-glucose pyrophosphorylase polypeptide of maize endosperm is replaced by an amino acid that confers said increased heat stability on said mutant enzyme.

26. The polynucleotide according to claim 25, wherein said amino acid corresponding to position 216 is replaced with a proline.

27. The polynucleotide according to claim 25, wherein said polynucleotide encodes a large subunit of ADP-glucose pyrophosphorylase of maize endosperm wherein the arginine amino acid at position 216 in the amino acid sequence of the wild type large subunit of ADP-glucose pyrophosphorylase polypeptide of maize endosperm is replaced with a proline.

28. A polynucleotide encoding a mutant large subunit of a plant ADP-glucose pyrophosphorylase polypeptide, or a biologically-active fragment or a biologically active variant of said mutant polypeptide, wherein said mutant polypeptide comprises an amino acid mutation in the amino acid sequence of said polypeptide, and wherein when said mutant polypeptide is expressed to form a mutant ADP-glucose pyrophosphorylase enzyme, said mutant enzyme exhibits increased heat stability relative to a wild type ADP-glucose pyrophosphorylase enzyme, and wherein said mutant polypeptide encoded by said polynucleotide comprises an amino acid mutation wherein the amino acid corresponding to position 400 in the amino acid sequence of the wild type large subunit of ADP-glucose pyrophosphorylase polypeptide of maize endosperm is replaced by a histidine amino acid and the amino acid corresponding to position 454 in the amino acid sequence of the wild type large subunit of ADP-glucose pyrophosphorylase polypeptide of maize endosperm is replaced by an isoleucine amino.

29. The polynucleotide according to any of claims 1 to 28, wherein said mutant polypeptide encoded by said polynucleotide further comprises the Rev6 mutation that confers increased seed weight to a plant expressing said polynucleotide.

30. The polynucleotide according to claim 29, wherein said mutation comprises the insertion of at least one serine residue between amino acids corresponding to positions 494 and 495 in the amino acid sequence of the wild type large subunit of ADP-glucose pyrophosphorylase polypeptide of maize endosperm.

31. The polynucleotide according to claim 29, wherein said mutation comprises the insertion of the amino acid pair tyrosine:serine between amino acids corresponding to positions 494 and 495 in the amino acid sequence of the wild type large subunit of ADP-glucose pyrophosphorylase polypeptide of maize endosperm.

32. The polynucleotide according to claim 29, wherein said mutation comprises the insertion of the amino acid pair serine:tyrosine between amino acids corresponding to positions 495 and 496 in the amino acid sequence of the wild type

large subunit of ADP-glucose pyrophosphorylase polypeptide of maize endosperm.

33. The polynucleotide according to any of claims 1 to 32, wherein said plant is a cereal.

34. The polynucleotide according to claim 33, wherein said cereal is selected from the group consisting of maize, wheat, rice and barley.

35. The polynucleotide according to claim 34, wherein said cereal is maize.

36. A method for increasing heat resistance of a plant, said method comprising the step of introducing a polynucleotide according to any one of claims 1 to 35 into said plant in a manner which enables expression of the protein encoded by said polynucleotide molecule.

37. The method according to claim 36, wherein said plant is a cereal.

38. The method according to claim 37, wherein said cereal is selected from the group consisting of maize, wheat, rice, and barley.

39. The method according to claim 36, wherein said plant is maize.

40. A plant cell comprising the polynucleotide molecule of any of claims 1 to 35.

41. The plant cell according to claim 40, wherein said plant cell is from a cereal plant.

42. The plant cell according to claim 41, wherein said cereal plant is selected from the group consisting of maize, wheat, rice, and barley.

43. The plant cell according to claim 40, wherein said plant cell is from a maize plant.

44. The plant cell according to claim 40, wherein said mutant polypeptide encoded by said polynucleotide comprises an amino acid mutation wherein the amino acid corresponding to position 333 in the amino acid sequence of the wild type large subunit of ADP-glucose pyrophosphorylase polypeptide of maize endosperm is replaced by an amino acid that confers said increased heat stability on said mutant enzyme.

45. The plant cell according to claim 44, wherein said amino acid corresponding to position 333 is replaced with a tyrosine.

46. The plant cell according to claim 44, wherein said polynucleotide encodes a large subunit of ADP-glucose pyrophosphorylase of maize endosperm wherein the histidine amino acid at position 333 in the amino acid sequence of the wild type large subunit of ADP-glucose pyrophosphorylase polypeptide of maize endosperm is replaced with a tyrosine.

47. The plant cell according to claim 40, wherein said mutant polypeptide encoded by said polynucleotide comprises an amino acid mutation wherein the amino acid corresponding to position 177 in the amino acid sequence of the wild type large subunit of ADP-glucose pyrophosphorylase polypeptide of maize endosperm is replaced by an amino acid that confers said increased heat stability on said mutant enzyme.

48. The plant cell according to claim 47, wherein said amino acid corresponding to position 177 is replaced with a proline.

49. The plant cell according to claim 47, wherein said polynucleotide encodes a large subunit of ADP-glucose pyrophosphorylase of maize endosperm wherein the alanine amino acid at position 177 in the amino acid sequence of the wild type large subunit of ADP-glucose pyrophosphorylase polypeptide of maize endosperm

is replaced with a proline.

50. The plant cell according to claim 40, wherein said mutant polypeptide encoded by said polynucleotide comprises an amino acid mutation wherein the amino acid corresponding to position 400 in the amino acid sequence of the wild type large subunit of ADP-glucose pyrophosphorylase polypeptide of maize endosperm is replaced by an amino acid that confers said increased heat stability on said mutant enzyme.

51. The plant cell according to claim 50, wherein said amino acid corresponding to position 400 is replaced with a histidine.

52. The plant cell according to claim 50, wherein said polynucleotide encodes a large subunit of ADP-glucose pyrophosphorylase of maize endosperm wherein the aspartic acid amino acid at position 400 in the amino acid sequence of the wild type large subunit of ADP-glucose pyrophosphorylase polypeptide of maize endosperm is replaced with a histidine.

53. The plant cell according to claim 40, wherein said mutant polypeptide encoded by said polynucleotide comprises an amino acid mutation wherein the amino acid corresponding to position 454 in the amino acid sequence of the wild type large subunit of ADP-glucose pyrophosphorylase polypeptide of maize endosperm is replaced by an amino acid that confers said increased heat stability on said mutant enzyme.

54. The plant cell according to claim 53, wherein said amino acid corresponding to position 454 is replaced with an isoleucine.

55. The plant cell according to claim 53, wherein said polynucleotide encodes a large subunit of ADP-glucose pyrophosphorylase of maize endosperm wherein the valine amino acid at position 454 in the amino acid sequence of the wild type large subunit of ADP-glucose pyrophosphorylase polypeptide of maize endosperm is replaced with an isoleucine.

56. The plant cell according to claim 40, wherein said mutant polypeptide encoded by said polynucleotide comprises an amino acid mutation wherein the amino acid corresponding to position 104 in the amino acid sequence of the wild type large subunit of ADP-glucose pyrophosphorylase polypeptide of maize endosperm is replaced by an amino acid that confers said increased heat stability on said mutant enzyme.

57. The plant cell according to claim 56, wherein said amino acid corresponding to position 104 is replaced with a threonine.

58. The plant cell according to claim 56, wherein said polynucleotide encodes a large subunit of ADP-glucose pyrophosphorylase of maize endosperm wherein the arginine amino acid at position 104 in the amino acid sequence of the wild type large subunit of ADP-glucose pyrophosphorylase polypeptide of maize endosperm is replaced with an threonine.

59. The plant cell according to claim 40, wherein said mutant polypeptide encoded by said polynucleotide comprises an amino acid mutation wherein the amino acid corresponding to position 460 in the amino acid sequence of the wild type large subunit of ADP-glucose pyrophosphorylase polypeptide of maize endosperm is replaced by an amino acid that confers said increased heat stability on said mutant enzyme.

60. The plant cell according to claim 59, wherein said amino acid corresponding to position 460 is replaced with a isoleucine.

61. The plant cell according to claim 59, wherein said polynucleotide encodes a large subunit of ADP-glucose pyrophosphorylase of maize endosperm wherein the threonine amino acid at position 460 in the amino acid sequence of the wild type large subunit of ADP-glucose pyrophosphorylase polypeptide of maize endosperm is replaced with an isoleucine.

62. The plant cell according to claim 40, wherein said mutant polypeptide encoded by said polynucleotide comprises an amino acid mutation wherein the amino acid corresponding to position 216 in the amino acid sequence of the wild type large subunit of ADP-glucose pyrophosphorylase polypeptide of maize endosperm is replaced by an amino acid that confers said increased heat stability on said mutant enzyme.

63. The plant cell according to claim 62, wherein said amino acid corresponding to position 216 is replaced with a proline.

64. The plant cell according to claim 62, wherein said polynucleotide encodes a large subunit of ADP-glucose pyrophosphorylase of maize endosperm wherein the arginine amino acid at position 216 in the amino acid sequence of the wild type large subunit of ADP-glucose pyrophosphorylase polypeptide of maize endosperm is replaced with a proline.

65. The plant cell according to any of claims 44 to 46, wherein said mutant polypeptide encoded by said polynucleotide further comprises a second amino acid mutation wherein the amino acid corresponding to position 460 in the amino acid sequence of the wild type large subunit of ADP-glucose pyrophosphorylase polypeptide of maize endosperm is replaced by an amino acid that confers said increased heat stability on said mutant enzyme.

66. The plant cell according to claim 65, wherein said amino acid corresponding to position 460 is replaced with an isoleucine.

67. The plant cell according to claim 65, wherein said polynucleotide encodes a large subunit of ADP-glucose pyrophosphorylase of maize endosperm wherein the threonine amino acid at position 460 in the amino acid sequence of the wild type large subunit of ADP-glucose pyrophosphorylase polypeptide of maize endosperm is replaced with an isoleucine.

68. The plant cell according to any of claims 44 to 46, wherein said mutant polypeptide encoded by said polynucleotide further comprises a second amino acid mutation wherein the amino acid corresponding to position 216 in the amino acid sequence of the wild type large subunit of ADP-glucose pyrophosphorylase polypeptide of maize endosperm is replaced by an amino acid that confers said increased heat stability on said mutant enzyme.

69. The plant cell according to claim 68, wherein said amino acid corresponding to position 216 is replaced with a proline.

70. The plant cell according to claim 68, wherein said polynucleotide encodes a large subunit of ADP-glucose pyrophosphorylase of maize endosperm wherein the arginine amino acid at position 216 in the amino acid sequence of the wild type large subunit of ADP-glucose pyrophosphorylase polypeptide of maize endosperm is replaced with a proline.

71. The plant cell according to claim 40, wherein said mutant polypeptide encoded by said polynucleotide comprises an amino acid mutation wherein the amino acid corresponding to position 400 in the amino acid sequence of the wild type large subunit of ADP-glucose pyrophosphorylase polypeptide of maize endosperm is replaced by a histidine amino acid and the amino acid corresponding to position 454 in the amino acid sequence of the wild type large subunit of ADP-glucose pyrophosphorylase polypeptide of maize endosperm is replaced by an isoleucine amino.

72. The plant cell according to claim 40, wherein said mutant polypeptide encoded by said polynucleotide further comprises the Rev6 mutation that confers increased seed weight to a plant expressing said polynucleotide.

73. The plant cell according to claim 72, wherein said mutation comprises the insertion of at least one serine residue between amino acids corresponding to positions 494 and 495 in the amino acid sequence of the wild type large subunit of

ADP-glucose pyrophosphorylase polypeptide of maize endosperm.

74. The plant cell according to claim 72, wherein said mutation comprises the insertion of the amino acid pair tyrosine:serine between amino acids corresponding to positions 494 and 495 in the amino acid sequence of the wild type large subunit of ADP-glucose pyrophosphorylase polypeptide of maize endosperm.

75. The plant cell according to claim 72, wherein said mutation comprises the insertion of the amino acid pair serine:tyrosine between amino acids corresponding to positions 495 and 496 in the amino acid sequence of the wild type large subunit of ADP-glucose pyrophosphorylase polypeptide of maize endosperm.

76. A mutant plant ADP-glucose pyrophosphorylase polypeptide encoded by a polynucleotide as defined in any of claims 1 to 35.

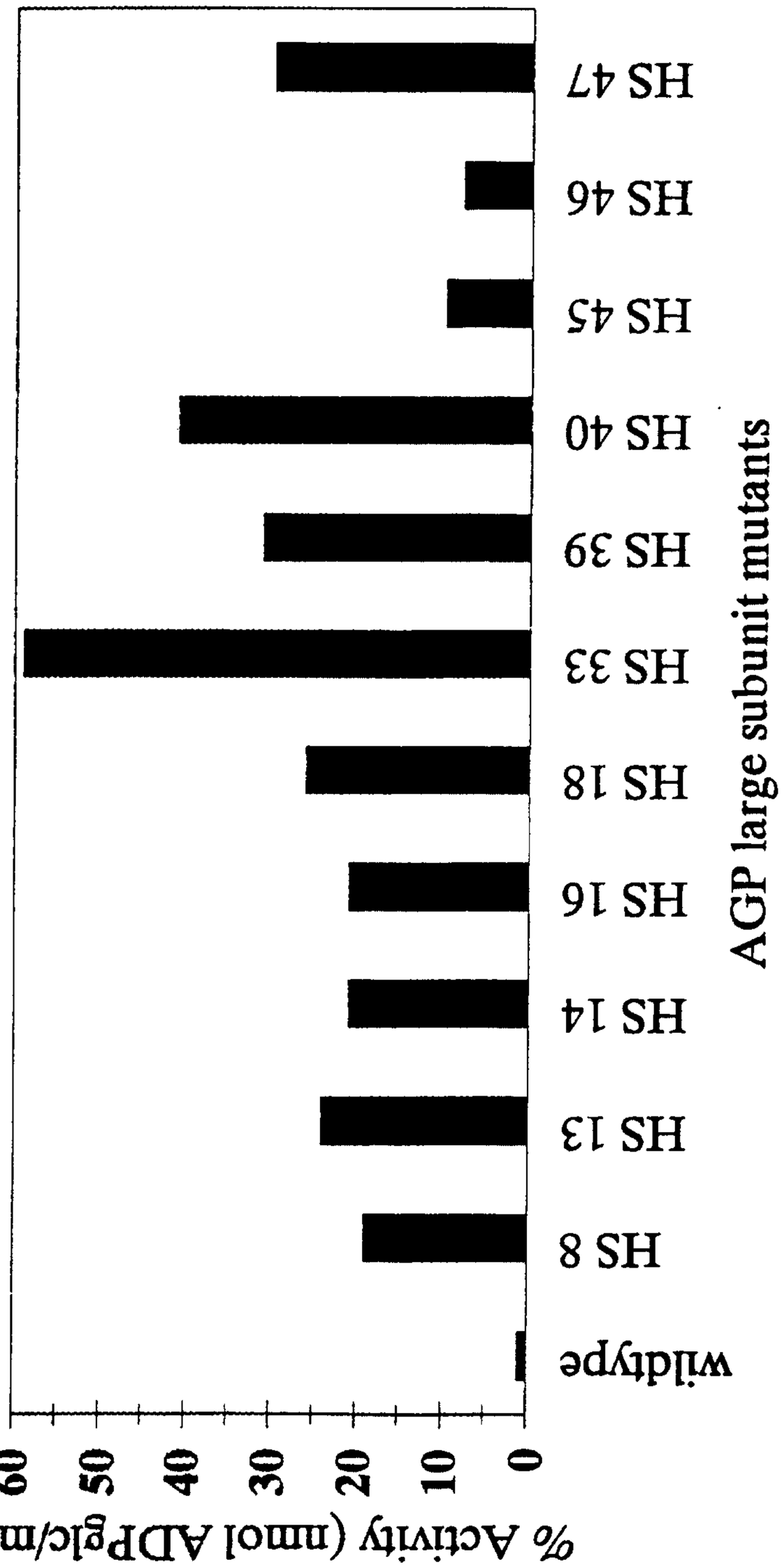
wildtype
8 SH
13 SH
14 SH
16 SH
18 SH
33 SH
39 SH
40 SH
45 SH
46 SH
47 SH
AGP large subunit mutants
% Activity (nmol ADPglc/min)
0
10
20
30
40
50
60

Figure 1

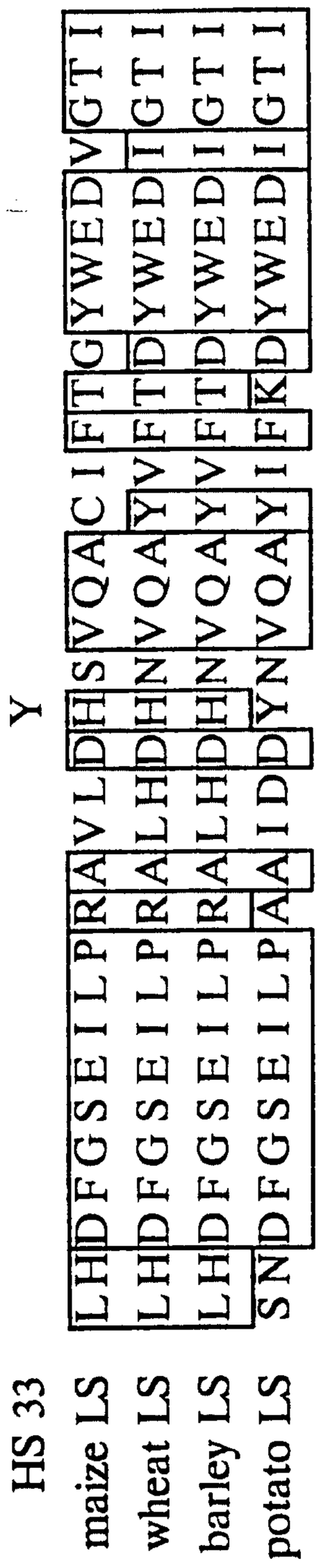
HS 33
maize LS
wheat LS
barley LS
potato LS
Y
VQACIFTGYWEDVGTI
VQAYVFTDIGTI
VQAYVFTDIGTI
VQAYIFKDIGTI
LHDFGSEILPRAVLDHSY
LHDFGSEILPRALHDHNY
LHDFGSEILPRALHDHNY
SNDFGSEILPAAIDDYNY

Figure 2

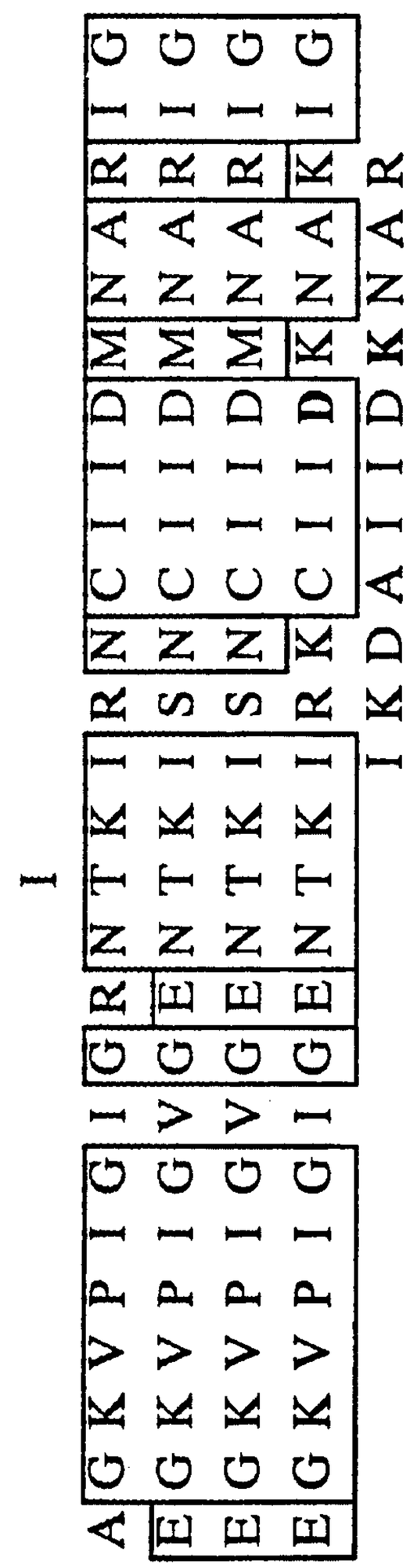

Figure 3