

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
18 September 2003 (18.09.2003)

PCT

(10) International Publication Number
WO 03/076580 A2

- (51) International Patent Classification⁷: C12N
- (21) International Application Number: PCT/US03/06908
- (22) International Filing Date: 5 March 2003 (05.03.2003)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:

10/091,912	5 March 2002 (05.03.2002)	US
10/092,227	5 March 2002 (05.03.2002)	US
60/362,372	5 March 2002 (05.03.2002)	US
60/403,921	15 August 2002 (15.08.2002)	US
- (71) Applicant (for all designated States except US): GENENCOR INTERNATIONAL, INC. [US/US]; 925 Page Mill Road, Palo Alto, CA 94304 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): BOTT, Richard, R. [US/US]; 3032 Hillside Drive, Burlingame, CA 94010

(US). KELLIS, James, T., Jr. [US/US]; 111 Tan Oak Drive, Portola Valley, CA 94028 (US). MORRISON, Thomas, B. [US/US]; 18 Blossom Road, Winchester, MA 01890 (US).

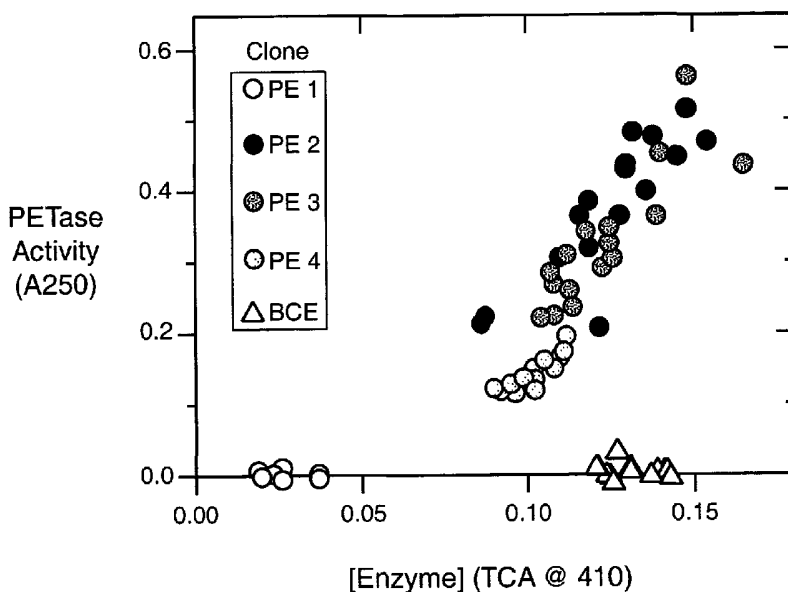
(74) Agent: MACKNIGHT, Kamrin, T.; Genencor International, Inc., 925 Page Mill Road, Palo Alto, CA 94304 (US).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO,

[Continued on next page]

(54) Title: HIGH THROUGHPUT MUTAGENESIS SCREENING METHOD



(57) Abstract: The present invention provides mutagenesis methods for protein engineering. In particular, the present invention provides mutagenesis methods that include feedback adjustment from systematic result evaluation. More specifically, the invention provides site-saturation mutagenesis methods that screen for variants with one or more desirable protein properties and evaluates screening results to provide feedback for repeat screening and construction of new libraries. The invention further provides cutinase variants that provide improved stability and activity compared with the wild type cutinase.

WO 03/076580 A2



SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

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

Published:

— *without international search report and to be republished upon receipt of that report*

HIGH THROUGHPUT MUTAGENESIS SCREENING METHOD

5

FIELD OF THE INVENTION

The present invention provides mutagenesis methods for protein engineering. In particular, the present invention provides mutagenesis methods that include feedback adjustment from systematic result evaluation. More specifically, the invention provides site-saturation mutagenesis methods that screen for variants with one or more desirable protein properties and evaluates screening results to provide feedback for repeat screening and construction of new libraries. The invention further provides cutinase variants that provide improved stability and activity compared with the wild type cutinase.

BACKGROUND OF THE INVENTION

Various mutagenesis procedures are used in the modification of proteins to obtain proteins with desired properties and/or capabilities. In most of these methods, the nucleotide sequence of a cloned gene encoding a protein is mutated and the modified gene is expressed to produce mutants, which are then typically screened for activities of interest. Then, the activities of the mutant proteins are usually compared with those of wild-type proteins.

A number of mutagenesis methods are known in the art. One general method involves site-saturation mutagenesis, which results in a mutated gene population that consists of otherwise identical genes, but includes randomly introduced codons. The random introduced codon can code for any amino acid, and the position is said to be "saturated" (Airaksinen and Hovi, *Nucleic Acids Res.*, 26:576-581[1998]). There are several different site-saturation methods known to those in the art.

Because most mutagenesis methods provide a great number of amino acid mutation options, screening of a large number of variants generally is required to produce a desired protein property. Generally, screening is repeated multiple times in order to produce a beneficial variant of interest. Thus, there is a continuing need in the art for mutagenesis and screening methods that provide relative ease of use as well as being efficient in terms of protocols and data analysis, in order to obtain desired protein engineering results.

- 2 -

SUMMARY OF THE INVENTION

The present invention provides mutagenesis methods for protein engineering. In particular, the present invention provides to mutagenesis methods that include feedback adjustment from systematic result evaluation. More specifically, the invention provides site-saturation mutagenesis methods that screen for variants with one or more desirable protein properties and evaluates screening results to provide feedback for repeat screening and construction of new libraries. The invention further provides cutinase variants that provide improved stability and activity compared with the wild type cutinase.

The high throughput mutagenesis screening methods of the present invention select sites for saturation scanning using protein structural considerations. In preferred embodiments, site-saturation libraries are created and screened using assays selected to identify clones having protein properties of interest. The mutation sites are categorized based on their properties of interest, and trends, if any, are identified. A variant or multiple variants having at least one desired property is/are selected for additional library creation procedures. These additional libraries are created using feedback from the determined categories. One set of additional libraries repeats construction of the previous libraries using the feedback. A second set of additional libraries represents new libraries created at sites that are expected to be beneficial based on the feedback. In some embodiments of the present invention, these categories are coded (*i.e.*, using color or other indicia), to allow easy identification of trends.

The present invention also provides numerous cutinase variants. In some embodiments, these variants comprise a substitution of one or more amino acids at residue positions corresponding to sites 40, 57-68, 85-86, 87, 125-127, 130, 148-152, 154-155, 176-183, 186, 190, 204-211, or 249 of *Pseudomonas mendocina* cutinase SEQ ID NO: 2, and homologous cutinases thereof, wherein the variant has improved activity and/or stability compared with the wild type cutinase. In some particularly preferred embodiments, the variants provide improved activity under various conditions and/or improved stability compared with the wild type cutinase. In some embodiments, the present invention provides cutinase variants of *Pseudomonas mendocina* cutinase SEQ ID NO: 2 and homologous cutinases thereof, wherein the variants comprise substitution of Phe 180 with Pro, substitution of Ser 205 with Gly, and one or more additional substitutions in at least one amino acid position selected from the amino acid positions 40, 64, 67, 177, 182, 186, 190, and 249.

In some embodiments, the variants have amino acid position 40 substituted with Ser, Glu, or Thr. In other embodiments, the amino acid position 64 is substituted with Val. In further embodiments, the amino acid position 67 is substituted with Ser. In still further

- 3 -

embodiments, amino acid position 177 is substituted with Asn or His. In other
embodiments, amino acid position 182 is substituted with Leu or Ala. In additional
embodiments, amino acid position 186 is substituted with Leu. In yet additional
embodiments, amino acid position 190 is substituted with Leu. In further embodiments
5 amino acid position 249 is substituted with Pro.

The present invention also provides cutinase variants comprising a substitution of
one or more amino acids at residue positions corresponding to sites 40, 57-68, 85-86, 88,
125-127, 130, 148-152, 154-155, 176-183, 186, 190, 204-211, or 249 of *Pseudomonas*
mendocina cutinase SEQ ID NO: 2, and homologous cutinase variants thereof, wherein
10 the cutinase variants have polyesterase activity. In some embodiments, the cutinase
variants comprise the substitution of GLY 59 with an amino acid selected from the group
consisting of Phe, Lys, Leu, and Val, and wherein the cutinase variant has increased
stability as compared to wild-type cutinase. In other embodiments, the cutinase variant
comprises the substitution of Thr 177 with an amino acid selected from the group
15 consisting of His, Leu, and Tyr, and wherein the cutinase variant has increased stability as
compared to wild-type cutinase. In alternative embodiments, the cutinase variant
comprises the substitution of Thr 64 with an amino acid selected from the group consisting
of Ala, Leu and Val, and wherein the cutinase variant has increased stability as compared
to wild-type cutinase. In some embodiments, the cutinase variant comprises the
20 substitution of Tyr 150 with Phe, and wherein the cutinase variant has increased stability
as compared to wild-type cutinase. In still other embodiments, the cutinase variant
comprises the substitution of Tyr 182 with an amino acid selected from the group
consisting of Ala, Leu, and Pro, and wherein the cutinase variant has increased stability as
compared to wild-type cutinase. In yet additional embodiments, the cutinase variant
25 comprises the substitution of Phe 180 and substitution of Ser 205, and wherein the
cutinase variant has increased stability as compared to wild-type cutinase. In further
embodiments, the Phe at position 180 is substituted with an amino acid selected from the
group consisting of Ala, His, Lys, Leu, Asn, Pro, and Gly, and wherein Ser 205 is
substituted with Gly. In some particularly preferred embodiments, the cutinase variant has
30 increased polyesterase activity as compared to wild-type cutinase. In some embodiments,
the cutinase variant comprises the substitution of Ala 66 with Glu, and wherein the
cutinase variant has increased stability as compared to wild-type cutinase. In additional
preferred embodiments, the cutinase variant comprises the substitution of Phe 180 with
Ile, and the cutinase variant has increased polyesterase activity as compared to wild-type
35 cutinase. In still further embodiments, the cutinase variant comprises the substitution of
Phe 180 with an amino acid selected from the group consisting of Lys and Leu, and the

- 4 -

substitution of Ser 205 with Gly, and the cutinase variant has increased polyesterase activity as compared to wild-type cutinase. In yet other embodiments, the cutinase variant comprises the substitution of Phe 180 with Asn, and wherein the cutinase variant has increased polyesterase activity as compared to wild-type cutinase. In further

5 embodiments, the cutinase variant comprises the substitution of Phe 180 with an amino acid selected from the group consisting of Asn, Pro, and Ser, and the substitution of Ser 205 with Gly, and wherein the cutinase variant has increased polyesterase activity as compared to wild-type cutinase. In yet additional embodiments, the cutinase variant comprises the substitution of Gly 59 with an amino acid selected from the group consisting

10 of Phe and Leu, and wherein the cutinase variant has increased polyesterase activity as compared to wild-type cutinase. In further embodiments, the cutinase variant comprises the substitution of Gly 61 with an amino acid selected from the group consisting of Asp and Glu, and wherein the cutinase variant has increased polyesterase activity as compared to wild-type cutinase. In some embodiments, the variant comprises the

15 following substitutions Ile 178 with Met; Phe 180 with Val; and Ser 205 with Gly. In particularly preferred embodiments, these cutinase variants have increased polyesterase activity as compared to wild-type cutinase. In additional embodiments, the cutinase variant comprises the substitution of Arg 20 with Gln, and the substitution of Tyr 112, and has increased polyesterase activity as compared to wild-type cutinase. In still further

20 embodiments, the cutinase variant comprises the substitution of Ser 205 with Asn, and substitution of Phe 207 with Leu, and has increased polyesterase activity as compared to wild-type cutinase. In yet other embodiments, the cutinase variant comprises substitution of Ser 263 with Arg, and has increased polyesterase activity as compared to wild-type cutinase. In additional embodiments, the cutinase variant comprises the substitution of

25 Ser 85 with Lys, and has increased polyesterase activity as compared to wild-type cutinase. In further embodiments, the cutinase variant comprises the substitution of Thr 177 with an amino acid selected from the group consisting of His, Leu, and Tyr, and has increased polyesterase activity as compared to wild-type cutinase. In alternative embodiments, the cutinase variant comprises the substitution of Thr 64 with Leu, and has

30 increased polyesterase activity as compared to wild-type cutinase. In further embodiments, the cutinase variant comprises the substitution of Tyr 182 with an amino acid selected from the group consisting of Ala and Pro, and has increased polyesterase activity as compared to wild-type cutinase. In some embodiments, the cutinase variant comprises the substitution of Phe 180 with an amino acid selected from the group

35 consisting of Ile, Leu, Asn, and Pro, and has increased stability as compared to wild-type cutinase. In other embodiments, the cutinase variant comprises the substitution of Arg

- 5 -

190 with Leu, and has increased stability as compared to wild-type cutinase. In additional
embodiments, the cutinase variant comprises the substitution of Arg 190 with Leu, and
has increased polyesterase activity as compared to wild-type cutinase. In yet further
embodiments, the cutinase variant comprises the substitution of Leu 249 with Pro, and
5 has increased stability as compared to wild-type cutinase. In other embodiments, the
cutinase variant comprises the substitution of Leu 249 with Pro, and has increased
polyesterase activity as compared to wild-type cutinase. In some embodiments, the
mutations described above are present in combination, while in other embodiments, they
are present as single mutations.

10 The present invention further provides cutinase variants comprising a substitution
of one or more amino acids at residue positions corresponding to sites 40, 57-68, 85-86,
88, 125-127, 130, 148-152, 154-155, 176-183, 190, 204-211, or 249 of *Pseudomonas*
mendocina cutinase SEQ ID NO: 2, and homologous cutinase variants thereof wherein the
cutinase variant is thermostable and has hydrolytic activity on polyester.

15 The present invention also provides methods for obtaining a variant having one or
more desired protein properties comprising the steps of: selecting amino acid sites in a
protein for mutation; performing mutagenesis at the selected mutation sites to create a
library; screening the library for variants having one or more desired protein properties;
grading the mutation sites of the variants for the one or more desired protein properties;
20 selecting one or more variants having a desirable grade as a template for, with feedback
from the grading, creating and screening additional libraries, whereby the method utilizes
cooperative mutations to obtain a variant having at least two mutations. In some
embodiments, the mutagenesis step is performed by site-saturation mutagenesis and
wherein selecting amino acid sites is performed by utilizing protein structural
25 considerations. In additional embodiments, the steps of creating and screening additional
libraries are performed by repeating site-saturation mutagenesis at mutation sites having
desirable grades and performing site-saturation mutagenesis at new sites on new libraries.
In further embodiments, the step of performing repeating site-saturation is performed by
selecting sites located near mutation sites having desirable grades. In yet additional
30 embodiments, protein structural considerations comprise such factors as the binding site
location, three-dimensional structure, amino acid sequence, nature of chemical reaction,
or nature of chemical binding. In some preferred embodiments, the protein property
comprises an enzyme property. In some particularly preferred embodiments, the enzyme
property comprises at least one property selected from the group consisting of catalysis,
35 binding, and stability. In alternative embodiments, the step of screening for one or more
variants is performed by selecting and conducting assays for the one or more protein

properties of interest. In yet other embodiments, grading is performed by identifying trends. In some preferred embodiments, the step of identifying trends is performed by plotting a spatial distribution of graded sites on a three-dimensional rendition of the protein. In some further preferred embodiments, the step of identifying trends is performed by plotting amino acid mutation identities. In yet other embodiments, the step of identifying trends is performed by plotting a distribution of graded mutation sites. In additional embodiments, the steps of creating and screening additional libraries are performed by screening the additional libraries for the desired protein properties and repeating site-saturation mutagenesis until a desired protein property goal is attained.

10 The present invention also provides methods or obtaining a variant enzyme having one or more desired properties comprising: utilizing a three-dimensional rendition of the enzyme to select mutation sites based on the amino acids; performing site-saturation mutagenesis at the selected mutation sites to create a library; screening the library for variants having one or more desired properties; grading the mutation sites of the variants for one or more desired properties; selecting one or more variants having a desirable grade as a template; using the template and feedback to repeat site-saturation mutagenesis at mutation sites having desirable grades and to perform site-saturation mutagenesis on new libraries at new sites. In yet additional embodiments, one or more desired properties are selected from the group consisting of substrate activity, thermostability, stability relative to reaction environment, ionic strength range of stability, pressure stability, and pH range of stability. In some preferred embodiments, one or more desired properties is selected from the group consisting of substrate activity and thermostability. In some particularly preferred embodiments, the enzyme is cutinase.

25 The present invention also provides processes (*i.e.*, methods) for the production of a cutinase variant with hydrolytic activity on polyester, wherein the cutinase is a variant of *Pseudomonas* cutinase, the process comprising: utilizing a three-dimensional model to select for mutation amino acid sites likely to demonstrate hydrolytic activity; performing site-saturation mutagenesis at the selected mutation amino acid sites to create a library having variants with mutated sites; screening the library for variants having mutated sites, using assays to detect polyesterase activity and thermostability; grading the mutated sites as beneficial, neutral or detrimental for both polyesterase activity and thermostability; selecting a variant having at least one beneficial grade; creating new and repeat libraries using the selected variant and feedback from the grading.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 provides a diagrammatic view illustrating the general principle of the invention.

Figure 2 provides a graph showing the relationship between polyesterase activity and protein concentration, as determined by TCA assay for several different *Bacillus subtilis* transformants.

Figure 3 provides a graph showing the polyesterase activity of "Library A" cutinase variants in two site-saturation libraries relative to TCA assay results. Wild-type enzyme controls are shown as well.

Figure 4 provides a graph showing the polyesterase activity of "Library B" cutinase variants in two site-saturation libraries relative to TCA assay results. Wild-type enzyme controls are shown as well.

Figure 5 provides a graph showing the thermal stability in laundry detergent of "Library A" cutinase variants in two site-saturation libraries relative to pNB activity results. Wild-type enzyme controls are also shown.

Figure 6 provides a graph showing the thermal stability in laundry detergent of "Library B" cutinase variants in two site-saturation libraries relative to pNB activity results. Wild-type enzyme controls are also shown.

Figure 7 provides a diagram showing the 3-dimensional structure of a protein showing the catalytic triad (dark color) active site face of *Pseudomonas mendocina* cutinase.

Figure 8 provides an alternate rendition of Figure 7. In this Figure, the initial target amino acids used for the site-saturation mutagenesis procedure of the present invention are shown in a darker color.

Figure 9 provides an alternate rendition of Figure 7. In this Figure, sites demonstrating at least 2% of variants with improved enzyme activity are shown in a darker color.

Figure 10 provides another alternate rendition of Figure 7. In this Figure, sites in which the majority of mutations had no effect on the enzyme activity are shown in dark color.

Figure 11 provides another rendition of Figure 7. In this Figure, sites in which the majority of mutations had a detrimental effect on the enzyme-activity are shown in dark color.

Figure 12 provides another rendition of Figure 7. In this Figure, sites demonstrating at least 2% of variants with improved protein stability are shown in dark color.

Figure 13 shows yet another rendition of Figure 7. In this Figure, sites in which the majority of mutations had no effect on protein stability are shown in dark color.

Figure 14 provides another rendition of Figure 7. In this Figure, sites in which the majority of mutations had a detrimental effect on protein stability are shown in dark color.

5 Figure 15 provides a graph showing the activity of a second generation variant in comparison with the wild type enzyme.

Figure 16 provides a graph showing the increased thermostability of a second generation variant versus the wild type.

10 Figure 17 provides the nucleotide sequence (SEQ ID NO:1) of the parent *P. mendocina* cutinase expressed in *Bacillus subtilis* #3934.

Figure 18 provides the amino acid (SEQ ID NO:2) and DNA sequence (SEQ ID NO:1) for the parent *P. mendocina* cutinase.

DESCRIPTION OF THE INVENTION

15 The present invention provides mutagenesis methods for protein engineering. In particular, the present invention provides to mutagenesis methods that include feedback adjustment from systematic result evaluation. More specifically, the invention provides site-saturation mutagenesis methods that screen for variants with one or more desirable protein properties and evaluates screening results to provide feedback for repeat
20 screening and construction of new libraries. The invention further provides cutinase variants that provide improved stability and activity compared with the wild type cutinase.

The present invention provides scanning and site-saturation mutagenesis methods for producing improved protein variants using as few as two sequential amino acid substitutions. In preferred embodiments, the methods include feedback adjustment from
25 systematic screening result evaluation for multiple protein properties following amino acid substitution. The method both identifies and grades amino acid substitutions and the mutation sites of the substitutions based upon the results of assays selected to detect at least one desired, improved protein property. The method further correlates the mutation sites and the specific mutations with the assay results, facilitating the identification of
30 trends and enabling construction of new molecular libraries and graded identification of mutation sites for further improvement. The method also involves repeated mutagenesis of the initial libraries, resulting in cooperative mutations that improve protein properties with as few as two mutations. New libraries also are created at new sites selected with grading feedback.

35 The general methods of the present invention are set out generally in Figure 1. In preferred embodiments, sites are selected for saturation based upon protein structural

- 9 -

information and/or any hypotheses related to structural-functional relationships at step 2 (See, Figure 1), and site-saturation libraries are created at step 4 (See, Figure 1). The created libraries are then screened at step 6 (See, Figure 1) for their property(ies) of interest, and the screening results are used to categorize sites at step 8 (See, Figure 1).
5 For example, in some preferred embodiments, the sites are ranked as "beneficial," "neutral," or "detrimental," while in other embodiments, other categories are designated as desired. At step 10 (See, Figure 1), an improved variant is selected based on the results obtained at step 8 (See, Figure 1), and the variant is used as a template for additional library construction. Library construction at step 12 (See, Figure 1) utilizes the template
10 variant from step 10 (See, Figure 1) and creates new sites expected to be beneficial based on feedback from the categorized results from step 8 (See, Figure 1). In addition to new libraries, the improved variant from step 10 (See, Figure 1) is used to repeat the previous libraries at the beneficial sites at step 14 using feedback from the categorized results from step 8 (See, Figure 1).

15 The site-saturation scanning mutagenesis methods of the present invention are conducted in an iterative fashion, with feedback adjustment, and in some embodiments, are broken down into the following four sections: site-saturation mutagenesis; activity screening; systematic result evaluation; new library construction/repeat site-saturation mutagenesis of previous libraries.

20

Definitions

Unless defined otherwise herein, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention pertains. For example, Singleton and Sainsbury, *Dictionary of*
25 *Microbiology and Molecular Biology*, 2d Ed., John Wiley and Sons, NY (1994); and Hale and Marham, *The Harper Collins Dictionary of Biology*, Harper Perennial, NY (1991) provide those of skill in the art with a general dictionaries of many of the terms used in the invention. Although any methods and materials similar or equivalent to those described herein find use in the practice of the present invention, the preferred methods and
30 materials are described herein. Accordingly, the terms defined immediately below are more fully described by reference to the Specification as a whole. Also, as used herein, the singular "a", "an" and "the" includes the plural reference unless the context clearly indicates otherwise.

As used herein, "*Pseudomonas*" refers to all of the members of the bacterial genus
35 *Pseudomonas*, including but not limited to *P. fluorescens*, *P. synxantha*, *P. mendocina*, *P. aeruginosa*, *P. aureofaciens*, and any other species within the genus *Pseudomonas*. It is

intended that the term encompass wild-type strains, as well as mutant and recombinant strains. It is also intended that the genus include species that have been reclassified to another genus.

As used herein, "the genus *Bacillus*" includes all members known to those of skill
5 in the art, including but not limited to *B. subtilis*, *B. licheniformis*, *B. lentus*, *B. brevis*, *B. stearothermophilus*, *B. alkalophilus*, *B. amyloliquefaciens*, *B. clausii*, *B. halodurans*, *B. megaterium*, *B. coagulans*, *B. circulans*, *B. lautus*, and *B. thuringiensis*. It is recognized that the genus *Bacillus* continues to undergo taxonomical reorganization. Thus, it is
10 intended that the genus include species that have been reclassified, including but not limited to such organisms as *B. stearothermophilus*, which is now named "*Geobacillus stearothermophilus*." The production of resistant endospores in the presence of oxygen is considered the defining feature of the genus *Bacillus*, although this characteristic also applies to the recently named *Alicyclobacillus*, *Amphibacillus*, *Aneurinibacillus*,
15 *Anoxybacillus*, *Brevibacillus*, *Filobacillus*, *Gracilibacillus*, *Halobacillus*, *Paenibacillus*, *Salibacillus*, *Thermobacillus*, *Ureibacillus*, and *Virgibacillus*. It is intended that the term encompass wild-type strains, as well as mutant and recombinant strains.

As used herein, "wild-type" and "native" proteins are those found in nature. The terms "wild-type sequence," and "wild-type gene" are used interchangeably herein, to refer to a sequence that is native or naturally occurring in a host cell. In some embodiments,
20 the wild-type sequence refers to a sequence of interest that is the starting point of a protein engineering project. The genes encoding the naturally-occurring (*i.e.*, precursor) protein may be obtained in accord with the general methods known to those skilled in the art. The methods generally comprise synthesizing labeled probes having putative sequences encoding regions of the protein of interest, preparing genomic libraries from
25 organisms expressing the protein, and screening the libraries for the gene of interest by hybridization to the probes. Positively hybridizing clones are then mapped and sequenced.

As used herein, "recombinant protein," refers to a protein in which the DNA sequence encoding the protein has been modified to produce a variant (or mutant) DNA
30 sequence which encodes the substitution, deletion or insertion of one or more amino acids in the naturally-occurring amino acid sequence. Suitable methods to produce such modification, and which may be combined with those disclosed herein, include, but are not limited to those disclosed in US Patent 4,760,025 (US RE 34,606), US Patent 5,204,015 and US Patent 5,185,258, all of which are incorporated herein by reference.

35 As used herein, "protein" refers to any composition comprised of amino acids and recognized as a protein by those of skill in the art. The terms "protein," "peptide" and

- 11 -

polypeptide are used interchangeably herein. Amino acids may be referred to by their complete names (*e.g.*, alanine) or by the accepted one letter (*e.g.*, A), or three letter (*e.g.*, ala) abbreviations. Wherein a peptide is a portion of a protein, those skill in the art understand the use of the term in context. The term "protein" encompasses mature forms of proteins, as well as the pro- and prepro-forms of related proteins. Prepro forms of proteins comprise the mature form of the protein having a prosequence operably linked to the amino terminus of the protein, and a "pre-" or "signal" sequence operably linked to the amino terminus of the prosequence.

In some preferred embodiments, the term protein refers to natural, synthetic, and engineered enzymes such as oxidoreductases, transferases, isomerases, ligases, hydrolases; antibodies; polypeptides; peptides; hormones; cytokines; growth factors; other biological modulators; protein-based diagnostic reagents; and biosensors. Examples of enzymes that may be enhanced by the method of the invention include, but are not limited to, proteases, carbohydrases, lipases, peroxidases, cellulases, and dioxygenases.

As used herein, the term "derivative" refers to a protein (*e.g.*, a cutinase) which is derived from a precursor protein (*e.g.*, the native cutinase) by addition of one or more amino acids to either or both the C- and N-terminal end(s), substitution of one or more amino acids at one or a number of different sites in the amino acid sequence, and/or deletion of one or more amino acids at either or both ends of the protein or at one or more sites in the amino acid sequence, and/or insertion of one or more amino acids at one or more sites in the amino acid sequence. The preparation of a cutinase derivative is preferably achieved by modifying a DNA sequence which encodes for the native protein, transformation of that DNA sequence into a suitable host, and expression of the modified DNA sequence to form the derivative cutinase.

Related (and derivative) proteins include "variant proteins." In preferred embodiments, variant proteins differ from a parent protein and one another by a small number of amino acid residues. The number of differing amino acid residues may be one or more, preferably 1, 2, 3, 4, 5, 10, 15, 20, 30, 40, 50, or more amino acid residues. In one preferred embodiment, the number of different amino acids between variants is between 1 and 10. In particularly preferred embodiments, related proteins and particularly variant proteins comprise at least 50%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, or 99% amino acid sequence identity.

As used herein, "protein properties" refers to protein structural integrity and/or function, including binding and/or catalysis relative to one or more conditions or substrates, protein activity including but not limited to binding and catalysis; protein stability relative to exposure to selected conditions, protein stability including but not

- 12 -

limited to thermal stability, pH range stability, ionic strength range stability, pressure stability; and/or reaction substance environment stability, such as stability in organic solvents.

As used herein, "protein structural considerations" refers to available information
5 about the physical structure of the protein of interest, including but not limited to substrate binding site location, three-dimensional structure, amino acid sequences; and, the chemical nature of the reaction or binding event to be addressed by mutagenesis of the protein.

The use of protein structural considerations is illustrated by the following two
10 examples which are not meant to be limiting. When both the site of substrate binding and the three-dimensional structure are known, site-saturation mutagenesis may be limited to a substrate binding site if the target property is catalytic activity. When only the amino acid sequence of the target protein is known, mutagenesis sites can be selected by comparing the known sequence to homologous sequences to allow selection of suspected
15 functionally important sites for mutagenesis. Both examples further include consideration of the nature of the problem to be solved (*i.e.* alteration of catalytic activity, binding, stability, etc.).

As used herein, "site-saturation mutagenesis" refers to a method of mutagenesis wherein oligonucleotide primers are randomized at the site of interest using the triplet
20 sequence NNG/C, where N is a mixture of all 4 bases. For each site selected, the amino acid is randomly replaced with all 19 possible alternatives. Site-saturation mutagenesis is not limited to mutations produced by a single codon substitution. In some preferred embodiments, mutant DNA sequences are generated with site saturation mutagenesis in at least one codon. In other preferred embodiments, site saturation mutagenesis is
25 performed for two or more codons. In a further embodiment, mutant DNA sequences have more than 40%, more than 45%, more than 50%, more than 55%, more than 60%, more than 65%, more than 70%, more than 75%, more than 80%, more than 85%, more than 90%, more than 95%, or more than 98% homology with the wild-type sequence. Alternatively, mutant DNA may be generated *in vivo* using any known mutagenic
30 procedure (*e.g.*, radiation, nitrosoguanidine, etc.). The DNA construct sequences may be wild-type, mutant or modified. In addition, the sequences may be homologous or heterologous.

As used herein, the term "targeted randomization" refers to a process that produces a plurality of sequences where one or several positions have been randomized.
35 In some embodiments, randomization is complete (*i.e.*, all four nucleotides, A, T, G, and C can occur at a randomized position. In alternative embodiments, randomization of a

nucleotide is limited to a subset of the four nucleotides. Targeted randomization can be applied to one or several codons of a sequence, coding for one or several proteins of interest. When expressed, the resulting libraries produce protein populations in which one or more amino acid positions can contain a mixture of all 20 amino acids or a subset of amino acids, as determined by the randomization scheme of the randomized codon. In some embodiments, the individual members of a population resulting from targeted randomization differ in the number of amino acids, due to targeted or random insertion or deletion of codons. In further embodiments, synthetic amino acids are included in the protein populations produced.

10 As used herein, "corresponding to," refers to a residue at the enumerated position in a protein or peptide, or a residue that is analogous, homologous, or equivalent to an enumerated residue in another protein or peptide.

As used herein, "corresponding region" generally refers to an analogous position within related proteins or a parent protein.

15 As used herein, the term "analogous sequence" refers to a sequence within a protein that provides similar function, tertiary structure, and/or conserved residues as the protein of interest. In particularly preferred embodiments, the analogous sequence involves sequence(s) at or near an epitope. For example, in epitope regions that contain an alpha helix or a beta sheet structure, the replacement amino acids in the analogous sequence preferably maintain the same specific structure. The term also refers to
20 nucleotide sequences, as well as amino acid sequences.

As used herein, "homologous protein" refers to a protein (*e.g.*, cutinase) that has similar catalytic action, structure, antigenic, and/or immunogenic response as the protein (*i.e.*, cutinase) of interest. It is not intended that a homolog and a protein (*e.g.*, cutinase) of interest be necessarily related evolutionarily. Thus, it is intended that the term encompass the same functional protein obtained from different species. In some preferred embodiments, it is desirable to identify a homolog that has a tertiary and/or primary structure similar to the protein of interest, as replacement for the epitope in the protein of interest with an analogous segment from the homolog will reduce the
25 disruptiveness of the change. Thus, in most cases, closely homologous proteins provide the most desirable sources of epitope substitutions. Alternatively, it is advantageous to look to human analogs for a given protein.

30 As used herein, "homologous genes" refers to at least a pair of genes from different, but usually related species, which correspond to each other and which are identical or very similar to each other. The term encompasses genes that are separated
35

by speciation (*i.e.*, the development of new species) (*e.g.*, orthologous genes), as well as genes that have been separated by genetic duplication (*e.g.*, paralogous genes).

As used herein, "ortholog" and "orthologous genes" refer to genes in different species that have evolved from a common ancestral gene (*i.e.*, a homologous gene) by speciation. Typically, orthologs retain the same function in during the course of evolution. Identification of orthologs finds use in the reliable prediction of gene function in newly sequenced genomes.

As used herein, "paralog" and "paralogous genes" refer to genes that are related by duplication within a genome. While orthologs retain the same function through the course of evolution, paralogs evolve new functions, even though some functions are often related to the original one. Examples of paralogous genes include, but are not limited to genes encoding trypsin, chymotrypsin, elastase, and thrombin, which are all serine proteinases and occur together within the same species.

The degree of homology between sequences may be determined using any suitable method known in the art (See *e.g.*, Smith and Waterman, *Adv. Appl. Math.*, 2:482 [1981]; Needleman and Wunsch, *J. Mol. Biol.*, 48:443 [1970]; Pearson and Lipman, *Proc. Natl. Acad. Sci. USA* 85:2444 [1988]; programs such as GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package (Genetics Computer Group, Madison, WI); and Devereux *et al.*, *Nucl. Acid Res.*, 12:387-395 [1984]). For example, PILEUP is a useful program to determine sequence homology levels. PILEUP creates a multiple sequence alignment from a group of related sequences using progressive, pairwise alignments. It can also plot a tree showing the clustering relationships used to create the alignment. PILEUP uses a simplification of the progressive alignment method of Feng and Doolittle, (Feng and Doolittle, *J. Mol. Evol.*, 35:351-360 [1987]). The method is similar to that described by Higgins and Sharp (Higgins and Sharp, *CABIOS* 5:151-153 [1989]). Useful PILEUP parameters including a default gap weight of 3.00, a default gap length weight of 0.10, and weighted end gaps. Another example of a useful algorithm is the BLAST algorithm, described by Altschul *et al.*, (Altschul *et al.*, *J. Mol. Biol.*, 215:403-410, [1990]; and Karlin *et al.*, *Proc. Natl. Acad. Sci. USA* 90:5873-5787 [1993]). One particularly useful BLAST program is the WU-BLAST-2 program (See, Altschul *et al.*, *Meth. Enzymol.*, 266:460-480 [1996]). parameters "W," "T," and "X" determine the sensitivity and speed of the alignment. The BLAST program uses as defaults a wordlength (W) of 11, the BLOSUM62 scoring matrix (See, Henikoff and Henikoff, *Proc. Natl. Acad. Sci. USA* 89:10915 [1989]) alignments (B) of 50, expectation (E) of 10, M⁵, N⁴, and a comparison of both strands.

- 15 -

As used herein, "percent (%) nucleic acid sequence identity" is defined as the percentage of nucleotide residues in a candidate sequence that are identical with the nucleotide residues of the sequence.

As used herein, the term "hybridization" refers to the process by which a strand of nucleic acid joins with a complementary strand through base pairing, as known in the art.

As used herein, "maximum stringency" refers to the level of hybridization that typically occurs at about $T_m - 5^\circ\text{C}$ (5°C below the T_m of the probe); "high stringency" at about 5°C to 10°C below T_m ; "intermediate stringency" at about 10°C to 20°C below T_m ; and "low stringency" at about 20°C to 25°C below T_m . As will be understood by those of skill in the art, a maximum stringency hybridization can be used to identify or detect identical polynucleotide sequences while an intermediate or low stringency hybridization can be used to identify or detect polynucleotide sequence homologs.

The phrases "substantially similar and "substantially identical" in the context of two nucleic acids or polypeptides typically means that a polynucleotide or polypeptide comprises a sequence that has at least 75% sequence identity, preferably at least 80%, more preferably at least 90%, still more preferably 95%, most preferably 97%, sometimes as much as 98% and 99% sequence identity, compared to the reference (*i.e.*, wild-type) sequence. Sequence identity may be determined using known programs such as BLAST, ALIGN, and CLUSTAL using standard parameters. (See *e.g.*, Altschul, *et al.*, J. Mol. Biol. 215:403-410 [1990]; Henikoff *et al.*, Proc. Natl. Acad. Sci. USA 89:10915 [1989]; Karin *et al.*, Proc. Natl. Acad. Sci. USA 90:5873 [1993]; and Higgins *et al.*, Gene 73:237 - 244 [1988]). Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information. Also, databases may be searched using FASTA (Pearson *et al.*, Proc. Natl. Acad. Sci. USA 85:2444-2448 [1988]).

In some embodiments, "equivalent residues" are defined by determining homology at the level of tertiary structure for a precursor protein (*i.e.*, protein of interest) whose tertiary structure has been determined by x-ray crystallography. Equivalent residues are defined as those for which the atomic coordinates of two or more of the main chain atoms of a particular amino acid residue of the precursor protein and another protein are within 0.13nm and preferably 0.1nm after alignment. Alignment is achieved after the best model has been oriented and positioned to give the maximum overlap of atomic coordinates of non-hydrogen protein atoms of the protein. In most embodiments, the best model is the crystallographic model giving the lowest R factor for experimental diffraction data at the highest resolution available.

In some embodiments, modification is preferably made to the "precursor DNA sequence" which encodes the amino acid sequence of the precursor enzyme, but can be

- 16 -

by the manipulation of the precursor protein. In the case of residues which are not conserved, the replacement of one or more amino acids is limited to substitutions which produce a variant which has an amino acid sequence that does not correspond to one found in nature. In the case of conserved residues, such replacements should not result
5 in a naturally-occurring sequence. Derivatives provided by the present invention further include chemical modification(s) that change the characteristics of the cutinase.

In some preferred embodiments, the protein gene is ligated into an appropriate expression plasmid. The cloned protein gene is then used to transform or transfect a host cell in order to express the protein gene. This plasmid may replicate in hosts in the sense
10 that it contains the well-known elements necessary for plasmid replication or the plasmid may be designed to integrate into the host chromosome. The necessary elements are provided for efficient gene expression (e.g., a promoter operably linked to the gene of interest). In some embodiments, these necessary elements are supplied as the gene's own homologous promoter if it is recognized, (i.e., transcribed, by the host), a transcription
15 terminator (a polyadenylation region for eukaryotic host cells) which is exogenous or is supplied by the endogenous terminator region of the protein gene. In some embodiments, a selection gene such as an antibiotic resistance gene that enables continuous cultural maintenance of plasmid-infected host cells by growth in antimicrobial-containing media is also included.

20 As used herein, "host cell" refers to a cell that has the capacity to act as a host and expression vehicle for an incoming sequence. In one embodiment, the host cell is a microorganism.

As used herein, the terms "DNA construct" and "transforming DNA" are used interchangeably to refer to DNA used to introduce sequences into a host cell or organism.
25 The DNA may be generated *in vitro* by PCR or any other suitable technique(s) known to those in the art. In particularly preferred embodiments, the DNA construct comprises a sequence of interest (e.g., as an incoming sequence). In some embodiments, the sequence is operably linked to additional elements such as control elements (e.g., promoters, etc.). The DNA construct may further comprise a selectable marker. It may
30 further comprise an incoming sequence flanked by homology boxes. In a further embodiment, the transforming DNA comprises other non-homologous sequences, added to the ends (e.g., stuffer sequences or flanks). In some embodiments, the ends of the incoming sequence are closed such that the transforming DNA forms a closed circle. The transforming sequences may be wild-type, mutant or modified. In some embodiments, the
35 DNA construct comprises sequences homologous to the host cell chromosome. In other embodiments, the DNA construct comprises non-homologous sequences. Once the DNA

construct is assembled *in vitro* it may be used to: 1) insert heterologous sequences into a desired target sequence of a host cell, and/or 2) mutagenize a region of the host cell chromosome (*i.e.*, replace an endogenous sequence with a heterologous sequence), 3) delete target genes; and/or introduce a replicating plasmid into the host.

5 In one embodiment, the present invention involves assembling a DNA construct *in vitro*, followed by cloning of such construct into a host cell. For example, PCR fusion and/or ligation can be employed to assemble a DNA construct *in vitro*. In some embodiments, the DNA construct comprises a DNA into which a mutation has been introduced.

10 An "incoming sequence" as used herein means a DNA sequence that is newly introduced into the host cell. In some embodiments, the incoming sequence becomes integrated into the host chromosome or genome. The sequence may encode one or more proteins of interest. Thus, as used herein, the term "sequence of interest" refers to an incoming sequence or a sequence to be generated by the host cell. The terms "gene of
15 interest" and "sequence of interest" are used interchangeably herein. The incoming sequence may comprise a promoter operably linked to a sequence of interest. An incoming sequence comprises a sequence that may or may not already present in the genome of the cell to be transformed (*i.e.*, homologous and heterologous sequences find use with the present invention).

20 As used herein, the terms "hybrid cutinases" and "fusion cutinases" refer to proteins that are engineered from at least two different or "parental" proteins. In preferred embodiments, these parental proteins are homologs of one another. For example, in some embodiments, a preferred hybrid cutinases or fusion cutinases protein contains the N-terminus of a protein and the C-terminus of a homolog of the protein. In some preferred
25 embodiment, the two terminal ends are combined to correspond to the full-length active protein.

 The term "sample" as used herein is used in its broadest sense. However, in preferred embodiments, the term is used in reference to a sample (*e.g.*, an aliquot) that comprises a variant (*e.g.*, a variant of a protein of interest) that is being analyzed,
30 identified, modified, and/or compared with other variants and/or the wild-type protein.

 As used herein, "mutations" refer to insertions, deletions, duplications, or alterations to peptide sequences or nucleic acid sequences (*e.g.*, wild-type sequences), including, but not limited to, point mutations.

 As used herein, "cooperative mutations" refers to two or more amino acid
35 mutations. In some preferred embodiments, these mutations function together in a complementary means.

As used herein, "cutinase" refers to a lipolytic enzyme capable of hydrolyzing cutin substrates. It is intended that cutinases from any source be encompassed within this definition. For example, cutinases are known to be produced by various fungi and bacterial species. Many cutinases have been described (See *e.g.*, Kolattukudy, *in*
5 Borgstrom and Brockman (eds.), Lipases, Elsevier, at pages 471-504 [1984]; Longhi *et al.*, *J. Mol. Biol.*, 268:779-799 [1997]; U.S Pat. No. 5,827,719; WO 94/14963; WO 94/14964; WO 00/05389; *Appl. Environ. Microbiol.*, 64: 2794-2799 [1998]; *Proteins: Structure, Function and Genetics* 26:442-458 [1996]; *J. Comput. Chem.*, 17:1783-1803 [1996]; *Protein Engineer.*, 6:157-165 [1993]). In some embodiments, the "parent" cutinase
10 (*i.e.*, the cutinases from which variant cutinases are designed and provided) has a nucleotide and/or amino acid sequence that is at least 50% homologous to nucleic acid and/or amino acid sequence of the cutinases of *Pseudomonas mendocina* cutinase SEQ ID NOS: 1 and SEQ ID NO:2. In some embodiments, the parent cutinases is 60%, 70%, 80%, 90%, or greater than 90% homologous to the *Pseudomonas mendocina* cutinase
15 nucleotide and/or amino acid sequences set forth in SEQ ID NO:1 and SEQ ID NO:2). In some embodiments, the cutinase is naturally occurring, while in other embodiments, it is a genetically modified cutinases. In some embodiments, the genetically modified cutinases is obtained by UV irradiation, N-methyl-N'-nitrosoguanidine (NTG) treatment, ethyl methanesulfonate (EMS) treatment, nitrous acid treatment, acridine treatment or the like.
20 In further embodiments, the modified cutinase is a recombinant cutinases produced by genetic engineering procedures such as cell fusion, gene recombination, and/or other methods known in the art.

As used herein, "grading" refers to the ranking, rating, and categorization of protein properties, including but not limited to identifying the properties as beneficial, neutral, and
25 detrimental, and including grouping such ranked properties and marking the groups by color-coding the data and three-dimensional renditions, and including identifying structural-functional trend analysis.

As used herein, "molecular library" refers to a vector-based assembly of at least one nucleic acid molecule that is vector based, non-vector based, or combined vector and
30 non-vector based.

As used herein, "polyester" refers to aliphilic and aromatic linear polymeric molecules containing in-chain ester groups and which are derived from the condensation of a diacid with a diol or from the polymerization of hydroxyl acids. The principal polyesters in industrial usage include polyethylene terephthalate (PET), tetramethylene
35 terephthalate (PTMT), polybutylene terphthalate (PBT), polytrimethylene terephthalate (PTT), polylactic acid (PLA), and polyethylene naphthalate (PEN),

- 19 -

polycyclohexanedimethylene terephthalate (CHDMT), poly(ethylene-4-oxybenzoate)A-Tell, polyglycolide, PHBA and 2GN.

As used herein, "polyesterase" refers to an enzyme that has significant capability to catalyze the hydrolysis and/or surface modification of polyester. Suitable polyesterases are isolated from animal, plant, fungal and bacterial sources. In some embodiments, the enzyme is obtained from any mutant strain (e.g., obtained by UV irradiation, N-methyl-N'-nitrosoguanidine (NTG) treatment, ethyl methanesulfonate (EMS) treatment, nitrous acid treatment, acridine treatment or the like), as well as recombinant strains induced by the genetic engineering procedures such as cell fusion and gene recombination, etc.

DETAILED DESCRIPTION OF THE INVENTION

The present invention provides mutagenesis methods for protein engineering. In particular, the present invention provides to mutagenesis methods that include feedback adjustment from systematic result evaluation. More specifically, the invention provides site-saturation mutagenesis methods that screen for variants with one or more desirable protein properties and evaluates screening results to provide feedback for repeat screening and construction of new libraries. The invention further provides cutinase variants that provide improved stability and activity compared with the wild type cutinase.

In preferred embodiments, the present invention provides high throughput mutagenesis screening methods to select sites for saturation scanning using protein structural considerations. Site-saturation libraries are created and screened using assays for sites having protein properties of interest. The sites are categorized for the properties of interest and trends, if any, are identified. Then, at least one variant having a desired property is selected for additional library creation procedures. The additional libraries are created using feedback from the determined categories. One set of additional libraries repeats construction of the previous libraries using the feedback. A second set of additional libraries are new libraries created at sites that are expected to be beneficial based on the feedback. In some embodiments, the categories are coded, using color or other indicia, to allow easy identification of trends.

The methods utilized in the present invention are described below. This description is divided into the four sections for clarity of presentation and is not meant to be limiting.

Site-Saturation Mutagenesis

As illustrated in Figure 1, in some preferred embodiments, the site-saturation mutagenesis method of the present invention is limited to a selected subset of amino acid

- 20 -

positions in the protein defined after a thorough analysis of the protein's structure and function.

Following selection of the subset of amino acid positions in the protein for mutagenesis (*i.e.*, step 2 in Figure 1), site-saturation libraries are made at the selected sites by random substitutions of all 19 amino acids using conventional oligonucleotide-directed mutagenesis (*i.e.*, step 4 in Figure 1). During the development of the present invention, the Stratagene Quik-Change™ kit was used. However, it is not intended that the present invention be limited to any particular methods as conventional library production is well-known to those skilled in the art (for example, see also, Airaksinen and Hovi, Nucl. Acids Res., 26:576-581 [1998]; and the Sculptor Mutagenesis Kit (Amersham, UK)).

The mutagenic oligonucleotide employed is complementary to the areas contiguous to the selected sites. The codon for the selected sites is replaced by the sequence N,N, (G/C mixture). Because the mutagenesis method is not limited to mutations produced by a single codon substitution, the method determines site importance relative to desired protein properties, and identifies the amino-acid substitutions that produce the most efficacious results. The method of the present invention is applicable to a variety of mutagenesis methods.

20 Activity Screening

Following site-saturation mutagenesis, members of the library undergo screening for the protein properties of interest using appropriate assays for the protein of interest (*i.e.*, step 6 in Figure 1). The screening preferably is conducted for more than one property to provide additional information regarding the quality and mutational effects of the screening data. The probability, "P," that a collection of variants contains at least one clone with a given amino acid is determined from the binomial distribution formula.

$$P=1 - (1-p)^n$$

p= probability of finding the particular amino acid in the clone mixture

n= the number of clones screened.

30 For example, if 90 clones are screened, there is a 94% probability of identifying an amino acid coded for by one of the 32 possible codons.

Systematic Result Evaluation

Next, the activity screening results for each variant are evaluated and categorized relative to the position of the mutation in the protein to identify any structure-based trends (*i.e.*, step 8 in Figure 1). Evaluations for trends include, but are not limited to, one or more

- 21 -

of the following examples using the assay results: (1) plotting the spatial distribution of variant sites within a three-dimensional rendition of the structure of the protein; (2) plotting the identity of the amino acid mutation; (3) plotting the distribution of beneficial mutation sites compared to amino acid sequences of similar proteins; and/or (4) grading and categorizing results, for example, as beneficial, detrimental, and neutral. The evaluated results, particularly categories, may be color-coded. The color-coding may be added to plotted data and superimposed on known, three-dimensional structural renditions of the protein. Such an analysis produces data that clearly reveal non-uniform site distribution for the various categories, and quickly points out trends for beneficial, detrimental and neutral amino acid substitutions. The analysis and result evaluation feedback enables informed selection of further mutagenesis sites and construction of new libraries.

New Library Construction/Repeat Mutagenesis

The data from the systematic result evaluation are used as feedback to construct new libraries and to select promising mutation sites for the previously tested libraries (*i.e.*, steps 12, 14 in Figure 1). These new libraries are constructed using the optimum variants identified as having beneficial mutations following primary mutagenesis. The optimum variants are used as templates for an additional mutagenesis procedure (*i.e.*, step 10 in Figure 1). In this construction method, libraries are made at each of the sites that produced improved or beneficial variants in the previous mutagenesis and screening steps. Following a second site-saturation mutagenesis procedure using as the parent DNA a clone with an amino acid mutation identified as beneficial, the optimum variants each have two beneficial mutations. The procedure is repeated until any protein improvement goal is reached, or until the number of beneficial sites is exhausted. This new library construction essentially repeats previous libraries using an improved starting protein chosen from one of the previous libraries and using identified beneficial mutations.

A second method of constructing a new library utilizes the trend analysis described above to mutate additional sites located near any observed and categorized beneficial mutation sites (*i.e.*, step 12 in Figure 1). This step also extends to producing mutations based on trends other than spatial proximity observed in beneficial sites. This library construction extends to new, potentially beneficial amino acid sites, and allows exclusion of categorized detrimental sites. In preferred embodiments, this construction also leads to additional improvements.

Cutinase Variants

In order to exemplify the utility of the present invention, numerous variant cutinases were produced using the methods of the present invention. However, it is not intended that the present invention be limited to cutinases as it provides advantages in the improvement of any protein.

The present invention further provides cutinase variants with improved activity and/or stability compared with the wild type cutinases. In some embodiments, the cutinase variant comprises one or more amino acid substitutions at residue positions corresponding to sites 40, 57-68, 85, 86, 88, 125-127, 130, 148-152, 154, 155, 176-183, 186 or 204-211 of *Pseudomonas mendocina* cutinase SEQ ID NO: 2, and homologous cutinases thereof.

A preferred cutinase variant is a variant of *Pseudomonas mendocina* cutinase SEQ ID NO: 2 and homologous cutinases thereof, wherein the variant comprises the substitutions of Phe 180 with Pro, Ser 205 with Gly, and one or more additional substitutions at amino acid positions including but not limited to 40, 64, 67, 177, 182, 186, 190, 249. For example, in some preferred embodiments, the amino acid position 40 is substituted with Ser, Glu, or Thr; the amino acid position 64 is substituted with Val; the amino acid position 67 is substituted with Ser; amino acid position 177 is substituted with Asn or His; the amino acid position 182 is substituted with Leu or Ala; the amino acid position 186 is substituted with Leu; the amino acid position 190 is substituted with Leu; the amino acid position 249 is substituted with Pro. Thus, in some embodiments, more than one amino acid is substituted in the cutinase variants as compared to the wild-type cutinase. For example, in some preferred embodiments, there are mutations at two, three, four, five, six, or more amino acid residues. In some embodiments, the variants were derived from a parent variant (e.g., F180P-S205G, referred to in the Tables below). Thus, in some embodiments, multiple amino acids are substituted so as to provide additional benefits to the variant cutinases.

EXPERIMENTAL

The following examples are provided in order to demonstrate and further illustrate certain preferred embodiments and aspects of the present invention and are not to be construed as limiting the scope thereof. These Examples describe the application of the methods of the present invention to produce an enzyme having increased stability and increased hydrolytic activity. However, it is not intended that the present invention be limited to these specific enzymes.

- 23 -

In the experimental disclosure which follows, the following abbreviations apply: °C (degrees Centigrade); rpm (revolutions per minute); H₂O (water); dH₂O (deionized water); (HCl (hydrochloric acid); aa (amino acid); bp (base pair); kb (kilobase pair); kD (kilodaltons); gm (grams); µg (micrograms); mg (milligrams); ng (nanograms); 5 µl (microliters); ml (milliliters); mm (millimeters); nm (nanometers); µm (micrometer); M (molar); mM (millimolar); µM (micromolar); U (units); V (volts); MW (molecular weight); sec (seconds); min(s) (minute/minutes); hr(s) (hour/hours); MgCl₂ (magnesium chloride); NaCl (sodium chloride); OD₂₈₀ (optical density at 280 nm); OD₆₀₀ (optical density at 600 nm); PAGE (polyacrylamide gel electrophoresis); PBS (phosphate buffered saline [150 mM NaCl, 10 mM sodium phosphate buffer, pH 7.2]); MOPS (3-(N-10 morpholino)propanesulfonic acid); TCA (trichloro-acetic acid); PEG (polyethylene glycol); PCR (polymerase chain reaction); RT-PCR (reverse transcription PCR); SDS (sodium dodecyl sulfate); Tris (tris(hydroxymethyl)aminomethane); w/v (weight to volume); v/v (volume to volume); LA medium (per liter: Difco Tryptone Peptone 20g, Difco Yeast 15 Extract 10g, EM Science NaCl 1g, EM Science Agar 17.5g, dH₂O to 1L); ATCC (American Type Culture Collection, Rockville, MD); Clontech (CLONTECH Laboratories, Palo Alto, CA); Difco (Difco Laboratories, Detroit, MI); GIBCO BRL or Gibco BRL (Life Technologies, Inc., Gaithersburg, MD); Invitrogen (Invitrogen Corp., San Diego, CA); NEB (New England Biolabs, Beverly, MA); Sigma (Sigma Chemical Co., St. Louis, MO); Takara (Takara Bio 20 Inc. Otsu, Japan); Roche (Hoffmann-La Roche, Basel, Switzerland); Testfabrics (Testfabrics, Inc., West Pittston, PA); Scientific Polymer (Scientific Polymer Products, Ontario, NY); V&P Scientific (V&P Scientific, San Diego, CA); Marsh (Marsh BioProducts, Rochester, NY); EM Science (EM Science, Gibbstown, NJ); Qiagen (Qiagen, Inc., Valencia, CA); and Stratagene (Stratagene Cloning Systems, La Jolla, CA).

25

EXAMPLE 1

Site-Saturation Mutagenesis of *Pseudomonas mendocina* Cutinase

Cutinase from *P. mendocina* (See, Figures 17 and 18, SEQ ID NO: 1 and SEQ ID NO:2) was subjected to the methods of the present invention to identify mutants that 30 improve hydrolytic activity on polyester in poly(ethylene terephthalate) fibers and resin and/or enzyme stability.

The culture media used was buffered by MOPS and also contained 5 µg/mL chloramphenicol, to select for transformants harboring the cutinase gene. All chemicals used in these Examples were obtained from commercial vendors. The culture medium 35 was prepared as follows:

- 24 -

MOPS 1A Starter Medium – to prepare 1000 mL (filter sterilize)

100 mL 10x MOPS (see below)
 12.6 g Glucose
 0.23 g K_2HPO_4
 0.53 g NH_4Cl
 Adjust pH to 7.3 with KOH

MOPS-Urea Medium – to prepare 1000 mL (filter sterilize)

100 mL 10x MOPS (see below)
 0.52 g K_2HPO_4
 3.6 g Urea
 21 g Glucose
 35 g Maltrin-150
 Adjust pH to 7.4 with KOH

10x MOPS – to prepare 1000 mL (filter sterilize)

83.7 g MOPS
 7.2 g Tricine
 12.0 g KOH
 0.48 g K_2SO_4
 1.07 g $MgCl_2 \cdot 6H_2O$
 29.2 g NaCl
 100 mL 100x Micronutrients (see below)

100x Micronutrients – to prepare 1000 mL (filter sterilize)

400 mg $FeSO_4 \cdot 7H_2O$
 100 mg $MnSO_4 \cdot H_2O$
 100 mg $ZnSO_4 \cdot 7H_2O$
 50 mg $CuCl_2 \cdot 2H_2O$
 100 mg $CoCl_2 \cdot 6H_2O$
 100 mg $NaMoO_4 \cdot 2H_2O$
 100 mg $Na_2B_4O_7 \cdot 10H_2O$
 1.1 g $CaCl_2$ (or 10 mL of 1.0 M stock)
 1.47 g Sodium citrate

I. Primary Screening Site-Saturation Mutagenesis

Site-saturation variant libraries were created at amino acid positions corresponding to residue positions 57-66, 68, 85, 86, 88, 125-127, 130, 148-152, 154, 155, 176-183, and 204-211 of SEQ ID NO:2 in *P. mendocina* cutinase, expressed in *Bacillus subtilis* strain 3934. The nucleic acid sequence of the parent cutinase is attached hereto as Figure 17 (SEQ ID NO:1). Libraries were created using the Stratagene Quik-Change™ kit using oligonucleotide primers randomized with NN(G/C) at the target position. Each selected amino acid was randomly replaced with all 19 possible alternatives. Colonies obtained from the variant libraries were selected and placed into 96-well growth plates that contained 100 μ L/well of MOPS 1A starter medium. The plates were incubated at 37°C overnight with continuous humidified shaking at 260 rpm. A sample from each well was

- 25 -

replicated into a Millipore 96-well filter plate that contained 200 μ L/well of MOPS-Urea medium. Then, 50 μ L of 45% glycerol was added to each well of the initial growth plates, which then were stored at -70°C.

The filter plates were incubated for 4 days at 37° C with continuous humidified shaking at 260 rpm, to allow growth and enzyme production. Next, the plates were filtered and assayed for hydrolysis activity using a polyester hydrolysis assay. In addition, the extracellular cutinase concentration was determined using a TCA turbidity assay, and the thermal stability in laundry detergent was also assessed.

II. Assays

The assays used to characterize the variant enzymes included the following assay systems.

A. Assay for Hydrolysis Activity on Polyester:

This polyester hydrolysis assay monitors the release of soluble terephthalate-containing fragments resulting from the enzymatic hydrolysis of insoluble polyester (polyethylene terephthalate or PET). The fragments have a significant absorbance at approximately 240-250 nm.

Insoluble PET polyester substrates in the form of fabric samples (Dacron 54 from Testfabrics) or PET resin powder (Scientific Polymer Products catalog # 138, cryogenically ground to < 400 μ m particle size) were used. The powdered PET samples were placed into a dispensing device (V&P Scientific), to distribute approximately 18 mg uniform samples into the wells of 96-well conventional assay plates. When PET fabric samples were used, 5/8-inch circular pieces of the fabric were cut and placed into 12-well assay plates.

For the PET powder hydrolysis assay, two plates were prepared, one containing PET powder and one with no substrate. Each of the 96 wells of the two assay plates received 200 μ L of a reaction buffer containing 50 mM Tris pH 8.6 and 0.01% Brij-35 (Sigma #430AG-6). The plate without PET substrate was used as an assay blank.

Twenty μ l of each enzyme variant samples were added to appropriate wells in each of the two assay plates. Both assay plates were sealed with plate sealers (Marsh). The assay plate containing the PET sample was placed in a covered 40°C incubator, and the plate was incubated for 18 hours with shaking at 250 rpm. The covered blank assay plate was maintained at room temperature.

- 26 -

Following incubation, 180 μ L samples from each well of the PET substrate plate were transferred to a filter plate with a 12-channel pipettor device. The ends of the pipettes were removed by cutting to avoid clogging with the PET powder. Each sample of the PET filter plate was filtered into a well of a new 96-well assay plate.

5 Then, 100 μ l from each well of the substrate plate and the blank plate were transferred to a UV-transparent plate, and the absorbance of each sample was determined in a plate reader at 250 nm. The absorbance of the wells from the blank plate without substrate was subtracted from the corresponding wells in the assay plate with PET substrate.

10

B. TCA Assay to Determine Cutinase Concentration

Twenty μ L of the filtered culture supernatant were added to each well of a 96-well flat bottom plate which contained 100 μ L/well of purified water. Light scattering/absorbance was determined at 410 nm using 5 second mixing mode in a plate reader to determine base-line blank values. Then, 100 μ L of 15% w/v TCA were added to each well prior to incubating the plate for between 5 and 30 minutes at room temperature. Light scattering/absorbance caused by protein precipitation was determined at 410 nm using a plate reader set for 5 second mixing mode.

20 The results were calculated by subtracting the blank reading, which did not contain TCA, from the sample readings with TCA. The light scattering/absorbance results were determined to be linear with respect to the enzyme concentrations present in the samples, thereby allowing direct plotting of concentration against enzyme performance as shown in Figure 2. Figure 2 shows the results obtained from the untransformed host strain (PE-1), three wild-type cutinase transformants (PE-2 through PE-4) and one negative control, transformed with a cellulase gene (BCE). Good correlation is observed between polyester hydrolysis activity and the TCA signal. These graphs help facilitate identification and selection of variants with elevated hydrolytic activity.

30 Figures 3 and 4 provide data from two site-saturation libraries involving amino acid positions Leu 154 (Library A) and Phe 180 (Library B). As indicated, the wild type controls fall along the same activity-concentration line; the line shown is the best fit to these data points calculated using linear regression. The graphs also show that the variants fall above, below, and along this line, signifying better, worse, or similar specific activity relative to the wild-type controls. In addition, as indicated by the graphs, the library at position 180 exhibits more diversity and more improved variants than the library at position 154.

35

C. Enzyme Stability Assays

Stability assays were conducted to determine thermostability of the variants in a detergent medium. The stability of the variants obtained from the site-saturation mutagenesis procedure was determined by adding 10 μL of the filter plate samples into each well of two 96-well plates containing 200 μL /well of liquid detergent (1.6 g/L, 6.3 gpg hardness). The 2 plates were mixed, and one plate was maintained on ice or in an ice-water bath (unstressed plate) while the second plate was incubated in a 40° C water bath for 10 minutes. (thermostressed plate). The stressed plate was transferred to the ice-water bath, and esterase activity assays were conducted on both plates as described below. Wild type cutinase samples also were run as above for comparison with the variants.

The enzyme activity assay used in these experiments is based upon the cleavage of colorless p-nitrophenyl butyrate (pNB) with esterase in the presence of water to produce butyrate and yellow colored p-nitrophenolate. The assay was conducted using an assay buffer containing 100 mM Tris pH 8.0 and 0.1% (v/v) Triton X-100. The substrate stock was 100 mM p-nitrophenyl butyrate in DMSO, prepared by adding 174.3 μL pNB to 10 mL DMSO. The substrate stock was divided into 1-mL aliquots and stored at -20° C until use.

The pNB assay procedure was as follows: substrate stock (400 μL) was diluted into 40 mL assay buffer for each pair of assay plates. Then, 200 μL of this diluted substrate solution were added to each well of conventional 96-well plates. The plates were equilibrated at 25° C, and the reaction was initiated by adding 10 μL of the enzyme/detergent solution from a stressed or unstressed plate. The plate was mixed and inserted into the plate reader and data acquisition was begun. The activity rate ($\Delta\text{mAU}_{410}/\text{minute}$) was determined. The pNB reaction rate is proportional to the concentration of active enzyme. For each well, the fraction of active enzyme after stress was calculated using the following formula:

$$\text{Fraction active} = (\text{stressed pNB activity} - \text{blank}) / (\text{unstressed pNB activity} - \text{blank}).$$

The fraction of active enzyme after stress was plotted against the pNB activity of the unstressed enzyme for each variant, as shown in Figures 5 and 6, which depict the results for site-saturation libraries at amino acid positions Leu 154 (Library A) and Phe 180 (Library B). The results shown in Figures 5 and 6 show thermostable variants having unstressed pNB activity, and also identify some variants which have apparently elevated thermostability only because the pNB activity is too low to produce an accurate ratio. Figures 5 and 6 further demonstrate that mutations in different sites produce different

- 28 -

enzymatic properties. For instance, the library shown in Figure 5 produced only a few promising variants with activity greater than wild-type, while the library shown in Figure 6 produced a number of variants with considerably greater activity than the wild-type enzyme.

5

D. Systematic Result Evaluations

The results from the TCA assay, the PET hydrolysis assay, and the stability assay were analyzed using conventional, known formulas as set out above, using Excel™ software. Variants were graded and ranked for stability and hydrolysis activity. Table 1 shows the results obtained from the site-saturation mutagenesis procedure described in this Example. The results were graded and color coded as follows: green=more than 2% of the variants at the listed sites had significantly greater activity than the control (beneficial mutation); blue=the majority of the variants had unchanged activity (neutral mutation); red=the majority of the variants had significantly decreased activity relative to the control (detrimental mutation) ; and brown=approximately half of the variants had unchanged activity and half had decreased activity (neutral mutation). The color-coding was repeated using the same colors and same percentage criteria for stability results. As indicated in Table 1, color coding easily identifies those variants having both desired enzyme properties.

20

//

//

Table 1

Site	wt Res.	# Clones	Specific PETase Activity			Stability		
			% Up	% Neutral	% Down	% Up	% Neutral	% Down
57	Gly	135	0.0	8.1	91.9	0.0	28.9	71.1
58	Thr	168	1.2	67.3	31.5	0.0	53.6	46.4
59	Gly	131	0.8	61.8	37.4	9.9	74.0	16.0
60	Ala	179	1.7	77.1	21.2	0.0	76.0	24.0
61	Gly	132	0.0	75.0	25.0	0.0	81.1	18.9
62	Pro	100	0.0	48.0	52.0	0.0	68.0	32.0
63	Ser	159	11.9	77.4	10.7	1.3	97.5	1.3
64	Thr	178	2.2	77.5	20.2	5.1	75.3	19.7
65	Tyr	179	4.5	43.0	52.5	0.0	67.0	33.0
66	Ala	138	0.7	82.6	16.7	0.7	92.8	6.5
68	Leu	134	0.0	27.6	72.4	0.0	50.7	49.3
85	Ser	133	1.5	78.2	20.3	0.0	87.2	12.8
86	Asn	174	1.1	47.7	51.1	0.6	97.7	1.7
88	Gly	145	0.0	4.8	95.2	0.0	33.8	66.2
125	His	110	0.9	10.9	88.2	0.0	45.5	54.5
126	Ser	158	0.0	9.5	90.5	0.0	13.3	86.7
127	Gln	179	0.6	38.0	61.5	0.0	39.7	60.3
130	Gly	166	1.2	25.9	72.9	0.0	48.2	51.8
148	Gln	180	1.1	38.3	60.6	0.0	27.8	72.2
149	Pro	179	0.0	2.8	97.2	0.0	34.1	65.9
150	Tyr	179	1.7	30.2	68.2	4.5	37.4	58.1
151	Thr	167	1.8	28.7	69.5	0.0	69.5	30.5
152	Leu	180	1.1	67.2	31.7	0.0	78.3	21.7
154	Leu	180	0.6	67.8	31.7	0.6	70.6	28.9
155	Gly	148	0.7	57.4	41.9	0.0	70.9	29.1
176	Asp	180	0.0	12.8	87.2	0.0	45.6	54.4
177	Thr	180	0.6	47.2	52.2	6.7	73.3	20.0
178	Ile	179	3.9	60.9	35.2	0.0	54.7	45.3
179	Ala	173	2.3	16.2	81.5	0.0	39.3	60.7
180	Phe	173	8.1	61.8	30.1	2.3	81.5	16.2
181	Pro	178	2.3	34.3	62.9	0.0	54.5	45.5
182	Tyr	166	9.6	83.7	6.6	15.7	83.1	1.2
183	Leu	179	1.7	41.3	57.0	0.0	45.3	54.7
204	Val	147	0.0	21.1	78.9	0.0	28.6	71.4
205	Ser	179	2.8	69.8	27.4	2.2	74.9	22.9
206	His	160	0.6	3.8	95.6	0.0	26.9	73.1
207	Phe	177	1.1	61.0	37.9	0.0	48.0	52.0
208	Glu	179	0.0	1.1	98.9	0.0	5.0	95.0
209	Pro	136	4.4	41.9	53.7	0.0	47.1	52.9
210	Val	180	1.7	56.7	41.7	0.6	75.6	23.9
211	Gly	165	2.4	33.3	64.2	3.6	53.3	43.0

Table 1 shows the results obtained from a typical enzyme system. The results are color-coded as follows:

- 5 Green (■) -- more than 2% of the clones had significantly greater activity than control.
- Blue (▨) -- the majority were unchanged. Red (▩) -- the majority had significantly decreased activity relative to the control. Brown (■) -- approximately half were unchanged and half were decreased.

- 30 -

Secondary screening to repeat the previous libraries and create new libraries was conducted using the variants having beneficial properties, as shown in Table 1. The frozen glycerol stock samples for each improved variant were streaked onto appropriate media and incubated to obtain single colonies. The colonies from each variant were inoculated into tubes containing 3 mL LB medium and shaken at 37° C for 6-18 hours.

Then, 800 μ L of broth from each incubated tube was combined with 400 μ L of 45% glycerol and stored at -70° C. One mL of broth was centrifuged, the supernatant was removed, and the pellet was stored at -20° C for DNA sequencing.

Then, 100 μ L of each broth was pipetted in multiple replicates into a 96-well plate. A sample from each well was replicated into a filter plate containing 200 μ L/well of MOPS-Urea medium. The samples present on the filter plates were allowed to grow for 4 days at 37° C with humidified shaking at 260 rpm. The plates were filtered and the TCA, PET hydrolysis, and stability assays described above were conducted. The results were calculated as described above, and the samples of the highest performing variants were pooled. The enzyme concentration in the pooled sample was determined by SDS-PAGE using densitometry or LC/MS using an ¹⁵N-labeled wild-type cutinase standard. The variants were tested for activity on polyester fabric in buffer and in detergent. The results are shown in Table 2, which lists the site and substitution made for variants with elevated stability. Table 3 indicates the variants with improved hydrolysis of polyester (PET). The F180-S205 last seven variants shown in Table 2 represent a second library as created in step 14 of Figure 1 (using the S205G variant as a template), showing two beneficial mutations which greatly increased relative stability of these seven variants. A number of variants with two mutations are also indicated in Table 3.

25

Table 2. Polyesterase Variants with Increased Stability (Secondary Screen)

Mutations in Variants	Relative Stability of Variants Compared to Wild-Type
F180I	1.65
F180L	1.89
F180N	1.66
F180P	1.63
G59F	2.02
G59K	2.62
G59L	2.78
G59V	3.19
S205G	2.47
S20R	1.38
S63R	2.23
S85H	2.33

- 31 -

S85K	2.56
T177H	2.46
T177L	2.68
T177Y	2.36
T64A	2.26
T64L	2.63
Y150F	1.87
Y182A	2.28
Y182L	2.14
Y182P	3.21
F180A-S205G	8.35
F180H-S205G	13.17
F180K-S205G	10.14
F180L-S205G	8.84
F180N-S205G	11.58
F180P-S205G	10.56
F180S-S205G	7.27

Table 3. Polyesterase Variants and Performance Summary (Tertiary Screen)

Mutations in Variants	Relative Buffer Activity of Variants Compared to Wild-Type	Relative Detergent Activity of Variants Compared to Wild-Type
A66E	1.41	1.22
F180A-S205G	1.31	2.31
F180H-S205G	1.18	1.75
F180I	2.21	2.24
F180K-S205G	1.01	2.87
F180L-S205G	1.18	1.95
F180N	1.77	1.83
F180N-S205G	1.24	2.02
F180P-S205G	1.64	3.43
F180S-S205G	1.54	2.62
G59F	1.13	1.09
G59L	1.18	1.30
G61D	1.85	2.07
G61E	1.73	2.16
I178M-F180V-S205G	2.07	2.99
R40Q-Y112Y	1.33	1.10
S205N-F207L	0.82	1.15
S63R	0.77	1.12
S85K	1.18	1.58
T177H	1.62	2.55
T177L	1.50	2.24
T177Y	1.58	2.73
T64L	1.30	1.27
Y182A	1.28	1.57
Y182P	1.21	1.38

- 32 -

The data in Table 3 were obtained by comparing samples of the variants having the indicated mutations with wild-type cutinase. These comparisons were made using equal enzyme concentrations in the "Tertiary Screen" (*i.e.*, overnight hydrolysis of 5/8" swatches of Dacron 54 polyester fabric in 12-well plates held at 40°C). Liquid laundry detergent was used in these experiments at a concentration standard in the industry.

The improvements in the properties of the second generation variants when tested in a detergent environment are shown graphically in Figures 15 and 16. Figure 15 indicates that the second generation variant has a 3-fold increase in activity as compared to the wild type when plotted against enzyme concentration. Figure 16 shows that the thermostability of the second generation variant is greatly increased as compared to the wild type.

A summary of the results of the method of the present invention includes the initial screening procedure results shown in Table 1. Table 1 provides a list of sites where mutations produced variants having at least some polyesterase activity and some stability. Additionally, Table 1 shows that mutations at the following sites produced variants having more than two percent of the clones with greater activity than controls: sites 63 -65, 178-182, 205, 209, and 211. Mutations at the following sites resulted in variants having more than two percent of the clones with greater stability than controls include sites 59, 64, 150, 177, 180, and 182. The following sites represent variants with both improved activity and stability include sites 64, 180, 182, and 211. Clones were selected based upon the promising sites shown in Table 1, and secondary screening was conducted by repeating the previous libraries and by creating new libraries at new sites expected to be beneficial.

Tables 2 and 3 represent, respectively, secondary stability screening steps and tertiary activity screening steps using the selected clones identified during primary screening as templates. Tables 2 and 3 identify the mutations made at the selected sites and include combinations of mutations. The substitutions of serine at locations 20 and 85 represent the creation of new libraries at new sites which are expected to be beneficial.

Table 2 shows that substitution of phenylalanine at site 180 with isoleucine, leucine, asparagines, or proline resulted in variants having improved stability as compared to the wild-type; substitution of glycine at site 59 with phenylalanine, lysine, leucine, or valine produced variants with improved stability as compared to the wild type. Substitutions of phenylalanine at site 180 in combination with substitutions of serine at site 205 resulted in variants with significantly increased relative stability.

As indicated in Table 3, specific substitutions at one to three sites produced variants with improved polyesterase activity.

- 33 -

D. Structural Context Evaluation

The graded screening results (amino acid changes that resulted in improvement, detriment, or neutral) were examined in the context of position within the protein to elucidate any underlying structurally based trends. The results are shown in bold superimposed on maps of the structure of the protein, as shown in Figures 7-14. The results also were color-coded (not shown) using the same colors used in Table 1. Figure 7 shows the catalytic triad active site face (in bold) of the *Pseudomonas mendocina* cutinase. (Ser 126, Asp 176, His 206). Figure 8 shows the addition of the initial target amino acids (in bold) that were used for the site-saturation scanning mutagenesis procedure. Figure 9 shows the sites (in bold) that resulted in at least 2% of the variants having significantly improved polyester hydrolysis activity. Figure 10 shows the variant sites (in bold) in which the majority of mutations had no effect on activity. Figure 11 shows the variant sites (in bold) in which the majority of mutations generally had a detrimental effect on activity. Figure 12 shows the variant sites (in bold) which resulted in a least 2% of the variants having significantly improved stability. Figure 13 shows the variant sites (in bold) in which the majority of mutations generally had no effect on stability. Figure 14 shows the variant sites (in bold) in which the majority of mutations generally had a detrimental effect on stability.

Figures 7-14 relate the variants to structure, thereby suggesting further modification sites. For instance, in the examples shown, detrimental sites tend to be buried in both the activity and stability maps and confined to generally the same locations. Improvement sites for activity and stability tend to occupy similar locations within the structure of the protein. Sites with no effect also occupy similar positions within the protein structure for both activity and stability results.

25

EXAMPLE 2

Additional Cutinase Variants

Clone number PE051-2-A2, which has mutations F180P-S205G, was subjected to additional site-saturation mutagenesis to create additional mutations according to protocols described in Example 1. Table 4 lists the site and substitution made for variants with elevated stability compared with the parent variant. Table 5 lists the site and substitution made for variants with elevated activity compared with the parent variant.

35

- 34 -

Table 4. Polyesterase Variants with Increased Stability (Secondary Screen)

Parent Variant	Additional Mutations	Relative Stability of Additional Mutations Compared to Parent Variant
F180P-S205G	T177N	1.59
F180P-S205G	T177H	1.73
F180P-S205G	Y182L-Q186L	1.50
F180P-S205G	Y182L	1.32
F180P-S205G	Y182A	1.35
F180P-S205G	T64V	1.28
F180P-S205G	R40T-T177N	1.23
F180P-S205G	R40S-T177H	1.34
F180P-S205G	R40T-T177N	1.87
F180P-S205G	R40T-T64V-T177N-L249P	3.15
F180P-S205G	R40T-T64V-T177N-Y182L-R190L-L248P	3.45

Table 5. Polyesterase Variants with Increased Activity (Tertiary Screen)

Parent Variant	Mutations	Relative Activity in Buffer of Variant Mutations Compared to Parent Variant	Relative Activity in Detergent of Variant Mutations Compared to Parent Variant
F180P-S205G	T177N	1.20	0.71
F180P-S205G	Y182L-Q186L	1.37	0.64
F180P-S205G	Y182L	1.27	0.81
F180P-S205G	Y182A	1.25	0.92
F180P-S205G	R40S	2.06	1.06
F180P-S205G	R40T	2.00	1.13
F180P-S205G	T64V	1.69	0.92
F180P-S205G	R40E	1.41	1.31
F180P-S205G	R40T-T177N	1.26	1.05
F180P-S205G	R40E-Y182L	1.81	1.48
F180P-S205G	R40E-T177N	1.52	1.27
F180P-S205G	R40S-Y182A-G67S	1.42	1.46
F180P-S205G	R40T	1.45	1.11
F180P-S205G	R40T-T177N	1.50	1.09
F180P-S205G	R40T-T64V-T177N-L249P	1.83	1.29
F180P-S205G	R40T-T64V-T177N-Y182L-R190L-L249P	2.39	1.61

5

The methods of the present invention find use in the engineering of proteins with enhanced properties for use in fields other than textile modifications as described above. Such fields include, but are not limited to, healthcare, foods and feed, brewing, cosmetics,

- 35 -

agriculture, textiles, detergents, environmental waste conversion, biopulp processing, biomass conversion to fuel, and other chemical procedures. Indeed, the methods of the present invention find use in the improvement of proteins for any application.

5 The examples above are not meant to be limiting, and the invention is intended to apply to mutagenesis of any protein. For those proteins where three dimensional maps are not available and only the amino acid sequence is known, the method may be carried out by comparing the known sequence to other sequences which are homologous. Using this method, dissimilar sites or sites with suspected functional importance are mutated because non-conserved sites generally are more readily available to accommodate amino
10 acid substitutions.

All publications and patents mentioned in the above specification are herein incorporated by reference. Various modifications and variations of the described method and system of the invention will be apparent to those skilled in the art without departing
15 from the scope and spirit of the invention. Although the invention has been described in connection with specific preferred embodiments, it should be understood that the invention as should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention that are obvious to those skilled in the art and/or related fields are intended to be within the scope of the present invention.

20

CLAIMS

1. A cutinase variant comprising a substitution of one or more amino acids at residue positions corresponding to sites 40, 57-68, 85-86, 88, 125-127, 130, 148-152, 154-155, 176-183, 190, 204-211, or 249 of *Pseudomonas mendocina* cutinase SEQ ID NO: 2, and homologous cutinase variants thereof.
2. The cutinase variant of Claim 1, wherein said cutinase variant provides improved activity and/or improved stability compared with the wild type cutinase.
3. A cutinase variant of *Pseudomonas mendocina* cutinase SEQ ID NO: 2 and homologous cutinases thereof, wherein said variant comprises the substitutions of Phe 180 with Pro, Ser 205 with Gly, and one or more additional substitutions selected from the amino acid positions 40, 64, 67, 177, 182, 186, 190, and 249.
4. The cutinase variant of Claim 3, wherein the amino acid position 40 is substituted with Ser, Glu, or Thr.
5. The cutinase variant of Claim 3, wherein the amino acid position 64 is substituted with Val.
6. The cutinase variant of Claim 3, wherein the amino acid position 67 is substituted with Ser.
7. The cutinase variant of Claim 3, wherein the amino acid position 177 is substituted with Asn or His.
8. The cutinase variant of Claim 3, wherein the amino acid position 182 is substituted with Leu or Ala.
9. The cutinase variant of Claim 3, wherein the amino acid position 186 is substituted with Leu.
10. The cutinase variant of Claim 3, wherein the amino acid position 190 is substituted with Leu.

- 37 -

11. The cutinase variant of Claim 3, wherein the amino acid position 249 is substituted with Pro.

12. A cutinase variant comprising a substitution of one or more amino acids at residue positions corresponding to sites 40, 57-68, 85-86, 88, 125-127, 130, 148-152, 154-155, 176-183, 186, 190, 204-211, or 249 of *Pseudomonas mendocina* cutinase SEQ ID NO: 2, and homologous cutinase variants thereof, wherein said cutinase variant has polyesterase activity.

13. The cutinase variant of Claim 12, wherein said cutinase variant comprises the substitution of GLY 59 with an amino acid selected from the group consisting of Phe, Lys, Leu, and Val, and wherein said cutinase variant has increased stability as compared to wild-type cutinase.

14. The cutinase variant of Claim 12, wherein said cutinase variant comprises the substitution of Thr 177 with an amino acid selected from the group consisting of His, Leu, and Tyr, and wherein said cutinase variant has increased stability as compared to wild-type cutinase.

15. The cutinase variant of Claim 12, wherein said cutinase variant comprises the substitution of Thr 64 with an amino acid selected from the group consisting of Ala, Leu and Val, and wherein said cutinase variant has increased stability as compared to wild-type cutinase.

16. The cutinase variant of Claim 12, wherein said cutinase variant comprises the substitution of Tyr 150 with Phe, and wherein said cutinase variant has increased stability as compared to wild-type cutinase.

17. The cutinase variant of Claim 12, wherein said cutinase variant comprises the substitution of Tyr 182 with an amino acid selected from the group consisting of Ala, Leu, and Pro, and wherein said cutinase variant has increased stability as compared to wild-type cutinase.

18. The cutinase variant of Claim 12, wherein said cutinase variant comprises the substitution of Phe 180 and substitution of Ser 205, and wherein said cutinase variant has increased stability as compared to wild-type cutinase.

19. The cutinase variant of Claim 18, wherein said Phe 180 is substituted with an amino acid selected from the group consisting of Ala, His, Lys, Leu, Asn, Pro, and Gly; and wherein said Ser 205 is substituted with Gly.

20. The cutinase variant of Claim 19, wherein said cutinase variant has increased polyesterase activity as compared to wild-type cutinase.

21. The cutinase variant of Claim 12, wherein said cutinase variant comprises the substitution of Ala 66 with Glu, and wherein said cutinase variant has increased stability as compared to wild-type cutinase.

22. The cutinase variant of Claim 12, wherein said cutinase variant comprises the substitution of Phe 180 with Ile, and wherein said cutinase variant has increased polyesterase activity as compared to wild-type cutinase.

23. The cutinase variant of Claim 12, wherein said cutinase variant comprises the substitution of Phe 180 with an amino acid selected from the group consisting of Lys and Leu, and the substitution of Ser 205 with Gly, and wherein said cutinase variant has increased polyesterase activity as compared to wild-type cutinase.

24. The cutinase variant of Claim 12, wherein said cutinase variant comprises the substitution of Phe 180 with Asn, and wherein said cutinase variant has increased polyesterase activity as compared to wild-type cutinase.

25. The cutinase variant of Claim 12, wherein said cutinase variant comprises the substitution of Phe 180 with an amino acid selected from the group consisting of Asn, Pro, and Ser, and the substitution of Ser 205 with Gly, and wherein said cutinase variant has increased polyesterase activity as compared to wild-type cutinase.

26. The cutinase variant of Claim 12, wherein said cutinase variant comprises the substitution of Gly 59 with an amino acid selected from the group consisting of Phe and Leu, and wherein said cutinase variant has increased polyesterase activity as compared to wild-type cutinase.

27. The cutinase variant of Claim 12, wherein said cutinase variant comprises the substitution of Gly 61 with an amino acid selected from the group consisting of Asp and Glu, and wherein said cutinase variant has increased polyesterase activity as compared to wild-type cutinase.

28. The cutinase variant of Claim 12, wherein said cutinase variant comprises the following substitutions Ile 178 with Met; Phe 180 with Val; and Ser 205 with Gly, and wherein said cutinase variant has increased polyesterase activity as compared to wild-type cutinase.

29. The cutinase variant of Claim 12, wherein said cutinase variant comprises the substitution of Arg 20 with Gln, and the substitution of Tyr 112, and wherein said cutinase variant has increased polyesterase activity as compared to wild-type cutinase.

30. The cutinase variant of Claim 12, wherein said cutinase variant comprises the substitution of Ser 205 with Asn, and substitution of Phe 207 with Leu, and wherein said cutinase variant has increased polyesterase activity as compared to wild-type cutinase.

31. The cutinase variant of Claim 12, wherein said cutinase variant comprises substitution of Ser 263 with Arg, and wherein said cutinase variant has increased polyesterase activity as compared to wild-type cutinase.

32. The cutinase variant of Claim 12, wherein said cutinase variant comprises the substitution of Ser 85 with Lys, and wherein said cutinase variant has increased polyesterase activity as compared to wild-type cutinase.

33. The cutinase variant of Claim 12, wherein said cutinase variant comprises the substitution of Thr 177 with an amino acid selected from the group consisting of His, Leu, and Tyr, and wherein said cutinase variant has increased polyesterase activity as compared to wild-type cutinase.

34. The cutinase variant of Claim 12, wherein said cutinase variant comprises the substitution of Thr 64 with Leu, and wherein said cutinase variant has increased polyesterase activity as compared to wild-type cutinase.

35. The cutinase variant of Claim 12, wherein said cutinase variant comprises the substitution of Tyr 182 with an amino acid selected from the group consisting of Ala and Pro, and wherein said cutinase variant has increased polyesterase activity as compared to wild-type cutinase.

36. The cutinase variant of Claim 12, wherein said cutinase variant comprises the substitution of Phe 180 with an amino acid selected from the group consisting of Ile, Leu, Asn, and Pro, and wherein said cutinase variant has increased stability as compared to wild-type cutinase.

37. The cutinase variant of Claim 12, wherein said cutinase variant comprises the substitution of Arg 190 with Leu, and wherein said cutinase variant has increased stability as compared to wild-type cutinase.

38. The cutinase variant of Claim 12, wherein said cutinase variant comprises the substitution of Arg 190 with Leu, and wherein said cutinase variant has increased polyesterase activity as compared to wild-type cutinase.

39. The cutinase variant of Claim 12, wherein said cutinase variant comprises the substitution of Leu 249 with Pro, and wherein said cutinase variant has increased stability as compared to wild-type cutinase.

40. The cutinase variant of Claim 12, wherein said cutinase variant comprises the substitution of Leu 249 with Pro, and wherein said cutinase variant has increased polyesterase activity as compared to wild-type cutinase.

41. A cutinase variant comprising a substitution of one or more amino acids at residue positions corresponding to sites 40, 57-68, 85-86, 88, 125-127, 130, 148-152, 154-155, 176-183, 190, 204-211, or 249 of *Pseudomonas mendocina* cutinase SEQ ID NO: 2, and homologous cutinase variants thereof wherein said cutinase variant is thermostable and has hydrolytic activity on polyester.

42. A method for obtaining a variant having one or more desired protein properties comprising the steps of: selecting amino acid sites in a protein for mutation; performing mutagenesis at the selected mutation sites to create a library; screening the library for variants having one or more desired protein properties; grading the mutation

sites of the variants for the one or more desired protein properties; selecting one or more variants having a desirable grade as a template for, with feedback from the grading, creating and screening additional libraries, whereby the method utilizes cooperative mutations to obtain a variant having at least two mutations.

43. The method of Claim 42, wherein said mutagenesis is performed by site-saturation mutagenesis and wherein selecting amino acid sites is performed by utilizing protein structural considerations.

44. The method of Claim 43, wherein creating and screening additional libraries is performed by repeating site-saturation mutagenesis at mutation sites having desirable grades and performing site-saturation mutagenesis at new sites on new libraries.

45. The method of Claim 44, wherein performing repeating site-saturation is performed by selecting sites located near mutation sites having desirable grades.

46. The method of Claim 43, wherein said protein structural considerations comprise binding site location, three-dimensional structure, amino acid sequence, nature of chemical reaction, or nature of chemical binding.

47. The method of Claim 42, wherein said protein property comprises an enzyme property.

48. The method of Claim 47, wherein said enzyme property comprises at least one property selected from the group consisting of catalysis, binding, and stability.

49. The method of Claim 42, wherein said screening for one or more variants is performed by selecting and conducting assays for the one or more protein properties of interest.

50. The method of Claim 42, wherein said grading is performed by identifying trends.

51. The method of Claim 50, wherein said identifying trends is performed by plotting a spatial distribution of graded sites on a three-dimensional rendition of the protein.

52. The method of Claim 51, wherein said identifying trends is performed by plotting amino acid mutation identities.

53. The method of Claim 51, wherein said identifying trends is performed by plotting a distribution of graded mutation sites.

54. The method of Claim 43, wherein said creating and screening additional libraries is performed by screening the additional libraries for the desired protein properties and repeating site-saturation mutagenesis until a desired protein property goal is attained.

55. A method for obtaining a variant enzyme having one or more desired properties comprising: utilizing a three-dimensional rendition of the enzyme to select mutation sites based on the amino acids; performing site-saturation mutagenesis at said selected mutation sites to create a library; screening said library for variants having one or more desired properties; grading the mutation sites of the variants for said one or more desired properties; selecting one or more variants having a desirable grade as a template; using said template and feedback to repeat site-saturation mutagenesis at mutation sites having desirable grades and to perform site-saturation mutagenesis on new libraries at new sites.

56. The method of Claim 55, wherein said one or more desired properties are selected from the group consisting of substrate activity, thermostability, stability relative to reaction environment, ionic strength range of stability, pressure stability, and pH range of stability.

57. The method of Claim 55, wherein said one or more desired properties is selected from the group consisting of substrate activity and thermostability.

58. The method of Claim 55, wherein said enzyme is cutinase.

59. A process for the production of a cutinase variant with hydrolytic activity on polyester, wherein said cutinase is a variant of *Pseudomonas* cutinase, the process comprising: utilizing a three-dimensional model to select for mutation amino acid sites likely to demonstrate hydrolytic activity; performing site-saturation mutagenesis at the

- 43 -

selected mutation amino acid sites to create a library having variants with mutated sites; screening the library for variants having mutated sites, using assays to detect polyesterase activity and thermostability; grading the mutated sites as beneficial, neutral or detrimental for both polyesterase activity and thermostability; selecting a variant having at least one beneficial grade; creating new and repeat libraries using the selected variant and feedback from the grading.

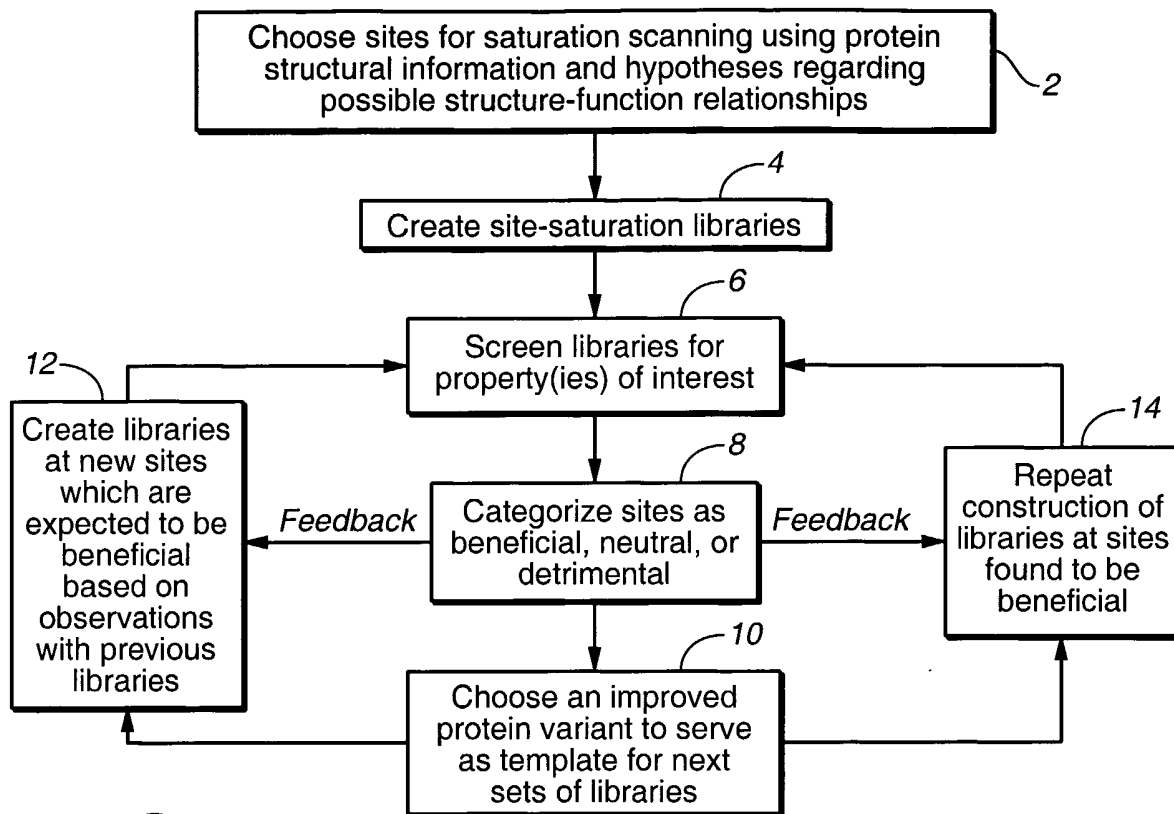


FIG. 1

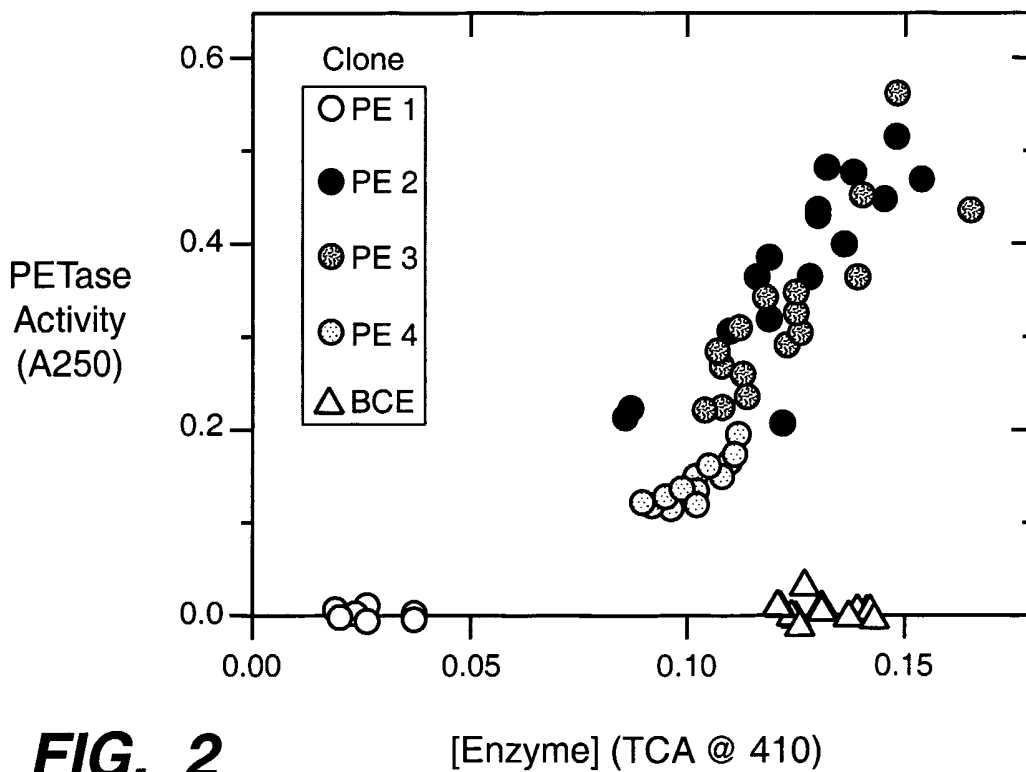
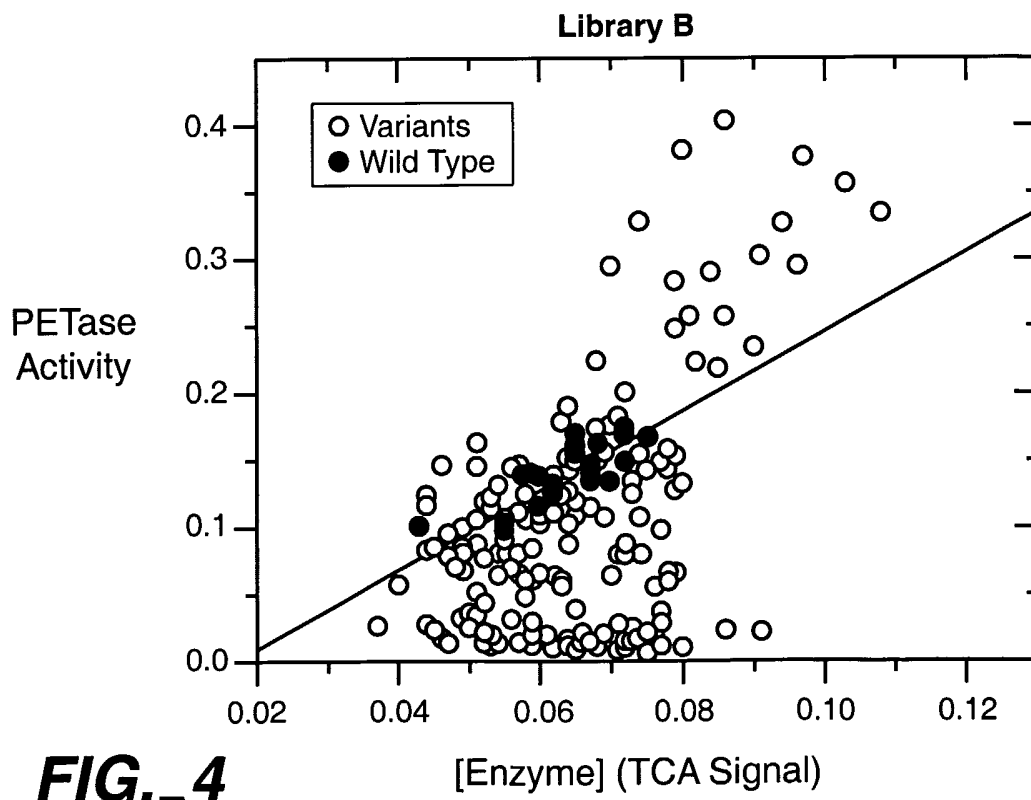
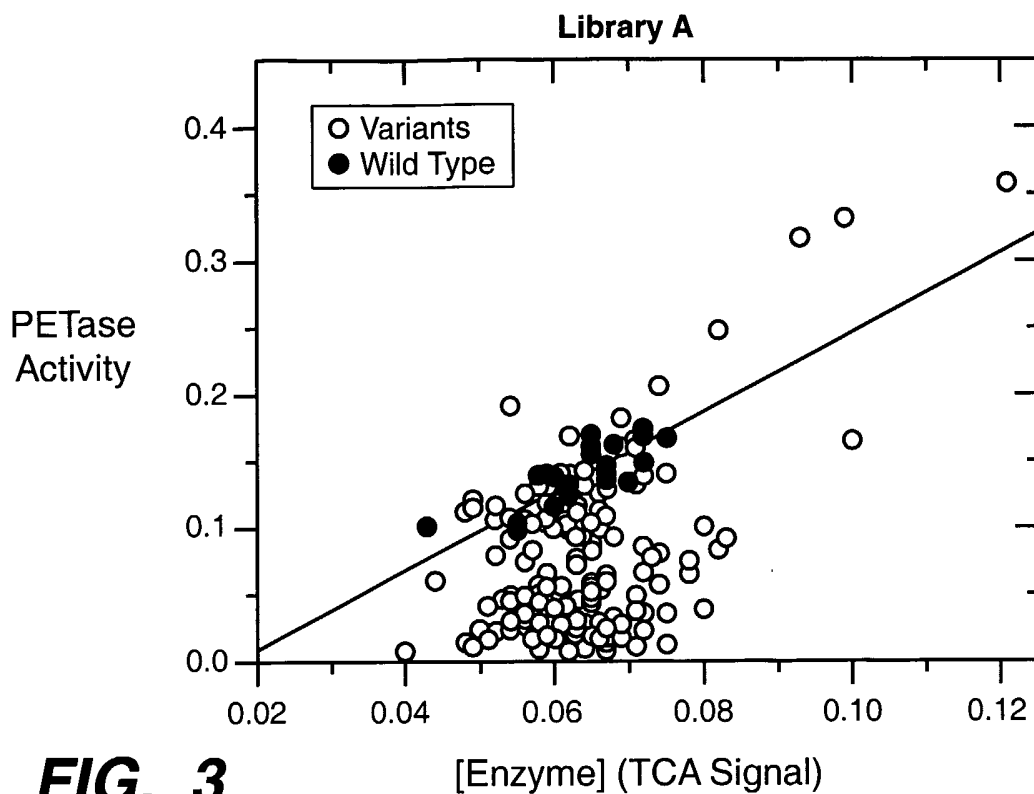


FIG. 2



3 / 10

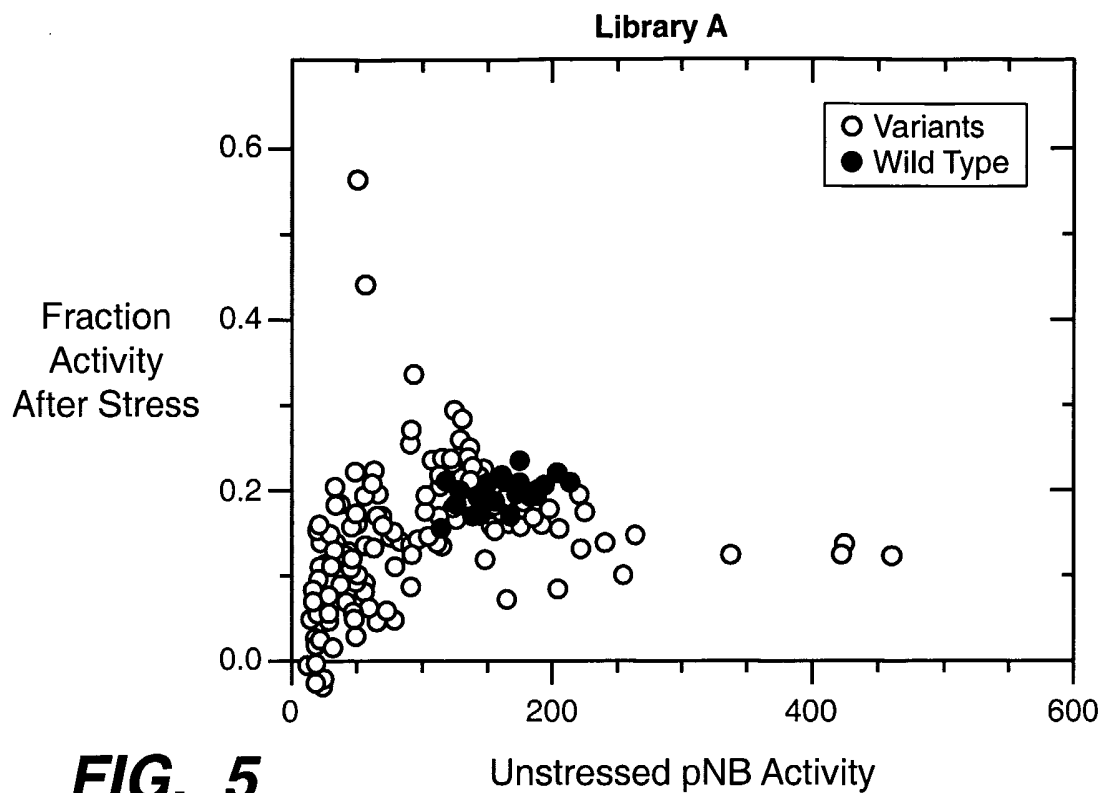


FIG._5

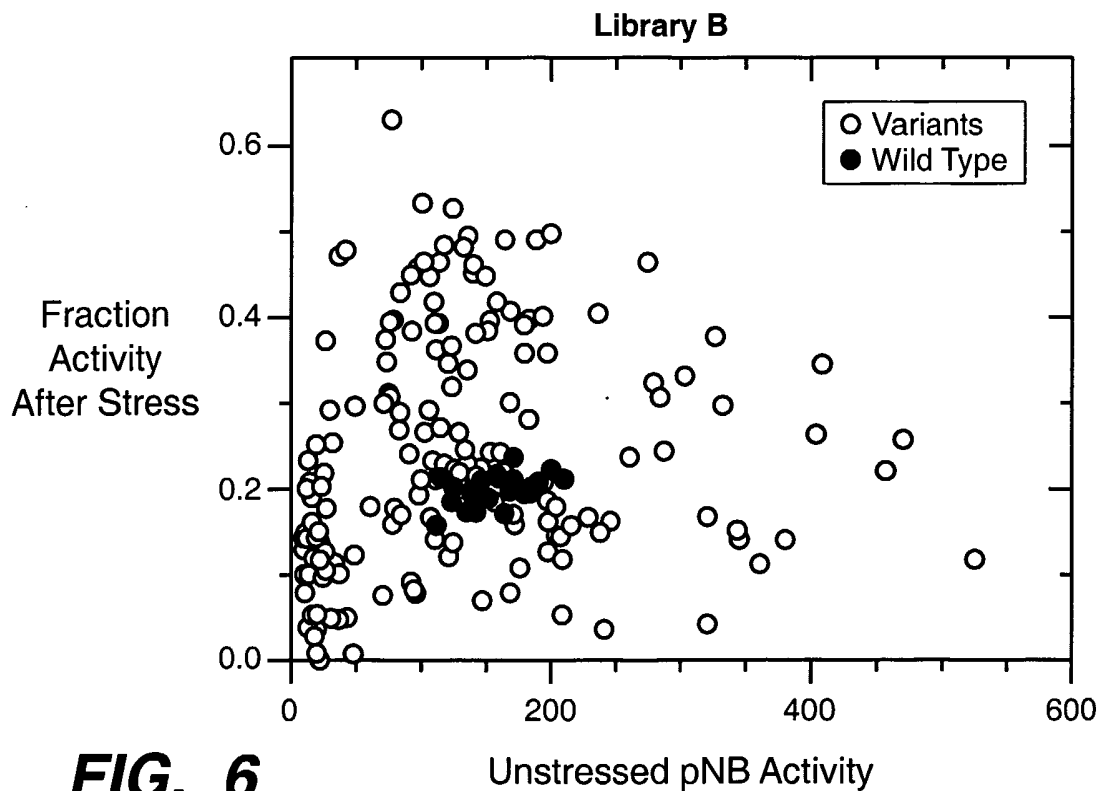


FIG._6



FIG._7



FIG._8

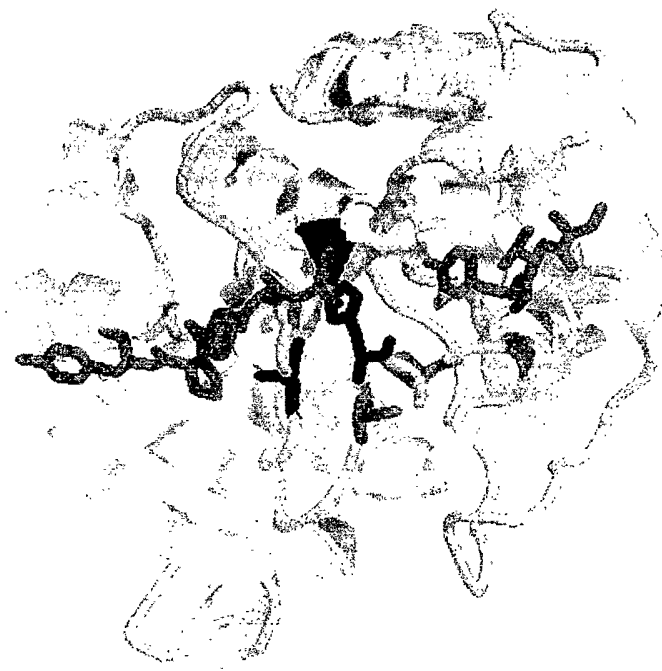


FIG._9



FIG._10



FIG. 11

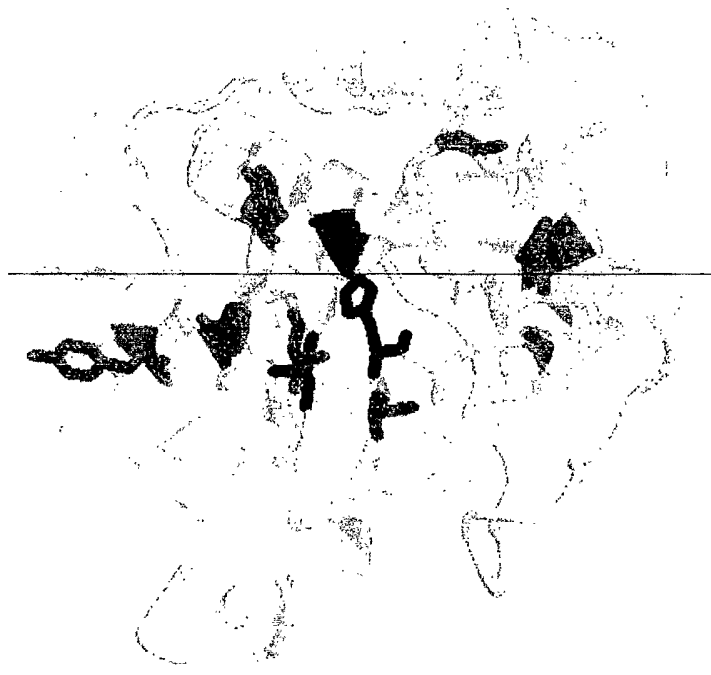


FIG. 12

7/10



FIG. 13



FIG. 14

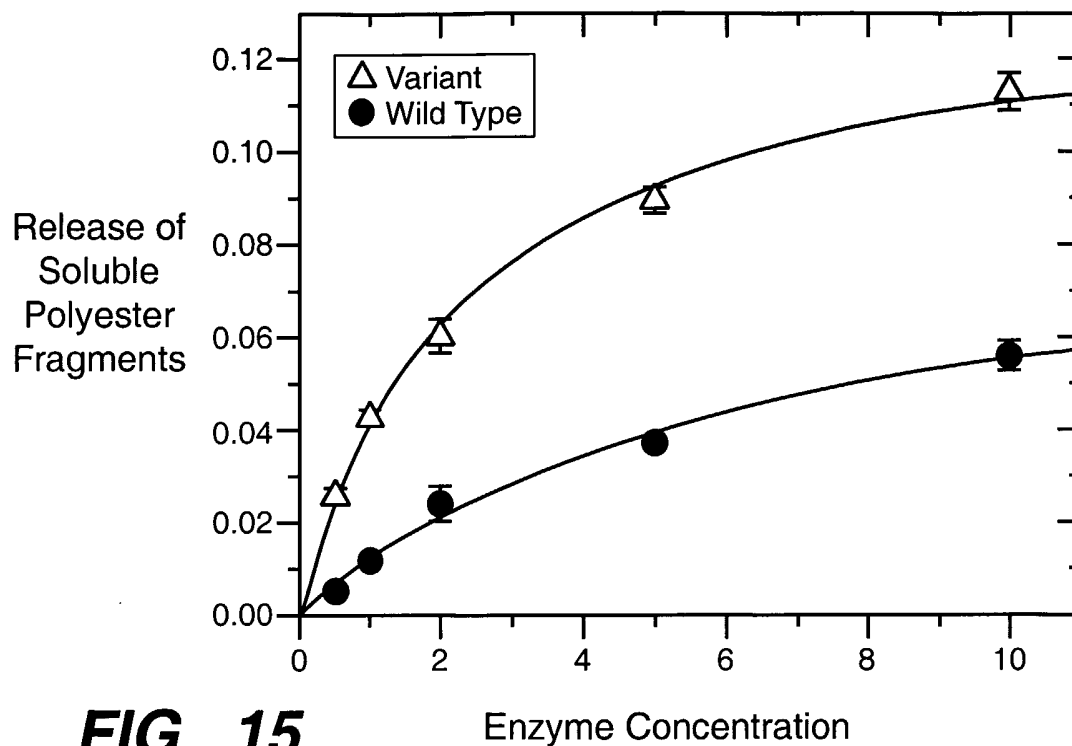


FIG. 15

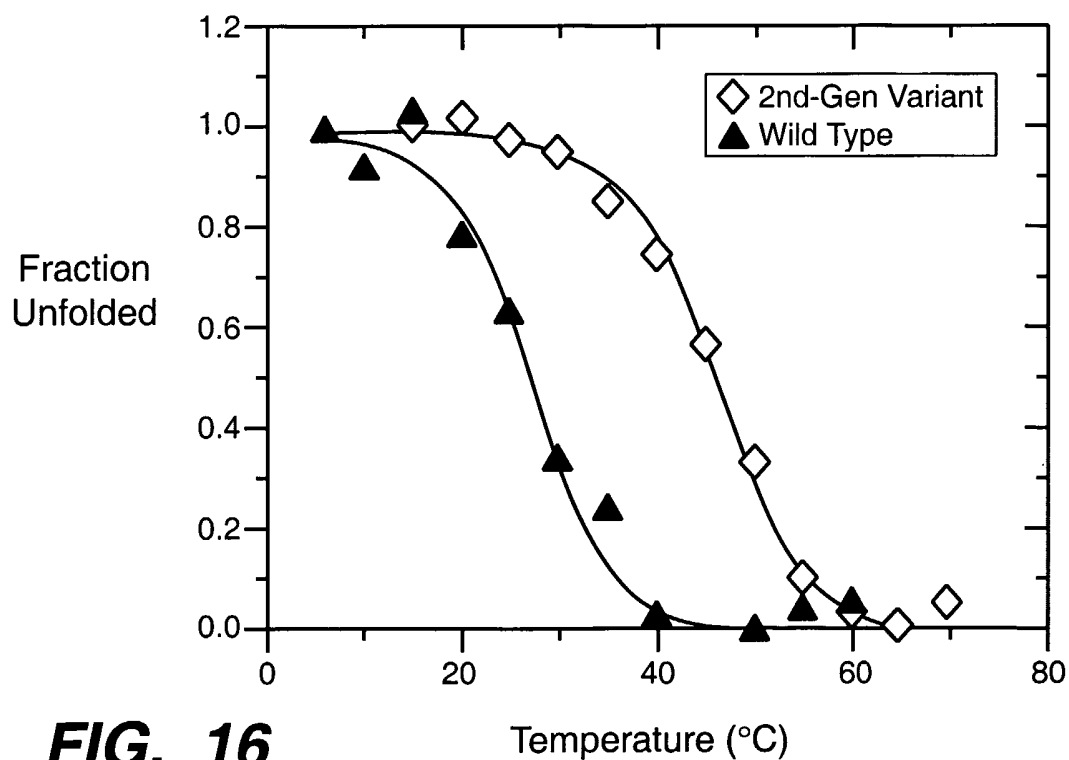


FIG. 16

9 / 10

TGGCGGCCTCTTGCCCTGTCCGTCTGTGCCACTGTCGCGGC 40
GGCTCCCCTGCCGGATACACCGGGAGCGCCATTTCCGGCT 80
GTCGCCAATTTTCGACCGCAGTGGCCCCCTACACCACCAGCA 120
GCCAGAGCGAGGGGCGGAGCTGTTCGCATCTATCGGCCCCG 160
CGACCTGGGTTCAGGGGGGGCGTGCATCCGGTGATTCTC 200
TGGGGCAATGGCACCGGTGCCGGGCCGTCCACCTATGCCG 240
GCTTGCTATCGCACTGGGCAAGCCACGGTTTCGTGGTGGC 280
GGCGGCGGAAACCTCCAATGCCGGTACCGGGCGGGAAATG 320
CTCGCCTGCCTGGACTATCTGGTACGTGAGAACGACACC 360
CCTACGGCACCTATTCCGGCAAGCTCAATACCGGGCGAGT 400
CGGCACTTCTGGGCATTTCCAGGGTGGTGGCGGCTCGATC 440
ATGGCCGGGCAGGATACGAGGGTGCCTACCACGGCGCCGA 480
TCCAGCCCTACACCCTCGGCCTGGGGCACGACAGCGCCTC 520
GCAGCGGCGGCAGCAGGGGCCGATGTTCCCTGATGTCCGGT 560
GGCGGTGACACCATCGCCTTTCCCTACCTCAACGCTCAGC 600
CGGTCTACCGGCGTGCCAATGTGCCGGTGTTCCTGGGGCGA 640
ACGGCGTTACGTCAGCCACTTCGAGCCGGTCGGTAGCGGT 680
GGGGCCTATCGCGGCCCGAGCACGGCATGGTTCGGCTTCC 720
AGCTGATGGATGACCAAGACGCCCGCGCTACCTTCTACGG 760
CGCGCAGTGCAGTCTGTGCACCAGCCTGCTGTGGTCCGGT 800
GAGCGCCGCGGGCTTTAA 818

FIG. 17

63 TGGCGGCTCTTGCCGTGTCCGTGTGCCACTGTCCGGGGTCCCTCCCTGCGGATACACCGG
 MetAlaSerCysLeuSerValCysAlaThrValAlaAlaProLeuProAspThrPro
 126 GAGGCCATTCCGGGTGCGCCAAATTCGACCGCAGTGGCCCTACACACCAGCAGCCAGA
 GlyAlaProPheProAlaValAlaAsnPheAspArgSerGlyProTyrThrThrSerSerGln
 189 GCGAGGGCCGAGCTGTCGCATCTATCGGCCCCCGACCTGGGTCAAGGGGGTGGTGCATC
 SerGluGlyProSerCysArgIleTyrArgProArgAspLeuGlyGlnGlyValArgHis
 252 CCGTGATTCTGGGGCAATGGCACCGGTGCCGGCCGTCACCTATGCCGGCTTGCTATCGC
 ProValIleLeuTrpGlyAsnGlyThrGlyAlaGlyProSerThrTyrAlaGlyLeuLeuSer
 315 ACTGGGCAAGCCACGGTTTCGTGGTGGCGGGGAAACCTCCAAATGCCGGTACCGGGCGGG
 HisTrpAlaSerHisGlyPheValAlaAlaAlaGluThrSerAsnAlaGlyThrGlyArg
 378 AAATGCTCGCCTGGACTATCTGGTACGTGAGAACGACACCCCTACGGCACCTATTCCG
 GluMetLeuAlaCysLeuAspTyrLeuValArgGluAsnAspThrProTyrGlyThrTyrSer
 441 GCAAGCTCAATACCGGGGAGTCGGCACCTTCTGGGCATTCACCAGGGTGGTGGCGGCTCGATCA
 GlyLysLeuAsnThrGlyArgValGlyThrSerGlyHisSerGlnGlyGlyGlySerIle
 504 TGGCCGGCAGGATACGAGGGTGCGTACCACGGCCCGATCCAGCCCTACACCCCTCGGCCCTGG
 MetAlaGlyGlnAspThrArgValArgThrThrAlaProIleGlnProTyrThrLeuGlyLeu
 567 GGCACGACAGCCCTCGCAGCGGGCAGCAGGGCCGATGTTCTGTGATGTCCGGTGGCGGTG
 GlyHisAspSerAlaSerGlnArgArgGlnGlyProMetPheLeuMetSerGlyGlyGly
 630 ACACCATCGCCTTCCCTACCTCAACGCTCAGCCGGTCTACGGCGTGCCAAATGTCCGGTGT
 AspThrIleAlaPheProTyrLeuAsnAlaGlnProValTyrArgArgAlaAsnValProVal
 693 TCTGGGGCAACGGCTTACGTACGCCACTTCGAGCCGGTCCGGTACGGTGGGGCCTATCGCG
 PheTrpGlyGluArgArgTyrValSerHisPheGluProValGlySerGlyGlyAlaTyrArg
 756 GCCCGACACGGCATGGTTCGGCTTCCAGCTGATGGATGACCAAGACGCCCGCGCTACCTTCT
 GlyProSerThrAlaTrpPheArgPheGlnLeuMetAspAspGlnAspAlaArgAlaThrPhe
 818 ACGCGCGCAGTGCAGTCTGTGCACCAGCCTGCTGTGGTCCGAGCGCCGCGGCTTTAA
 TyrGlyAlaGlnCysSerLeuCysThrSerLeuLeuTrpSerValGluArgArgGlyLeu*