PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 :		(11) International Publication Number: WO 92/19726
C12N 9/00, 15/57	A1	(43) International Publication Date: 12 November 1992 (12.11.92)

(21) International Application Number: PCT/DK92/00142

(22) International Filing Date: 1 May 1992 (01.05.92)

(30) Priority data:
91610035.7 1 May 1991 (01.05.91) EP
(34) Countries for which the regional or international application was filed: AT et al.

(71) Applicant *(for all designated States except US):* NOVO NORDISK A/S [DK/DK]; Novo Allé, DK-2880 Bagsvaerd (DK).

(72) Inventors; and
(75) Inventors/Applicants *(for US only) :* SVENDSEN, Allan [DK/DK]; Bakkeleddet 28, DK-3460 Birkerød (DK). VON DER OSTEN, Claus [DK/DK]; Boulevarden 14, L66, DK-2800 Lyngby (DK). CLAUSEN, Ib, Groth [DK/DK]; Ordrupjagtvej 153, st.t.v., DK-2920 Charlottenlund (DK). PATKAR, Shamkant, Anant [US/DK]; Christoffers Allé 91, DK-2800 Lyngby (DK). BORCH, Kim [DK/DK]; Klerkegade 12, II.t.v., DK-1308 Copenhagen K (DK).

(74) Common Representative: NOVO NORDISK A/S; Patent Department, Novo Allé, DK-2880 Bagsvaerd (DK).

(81) Designated States: AT (European patent), BE (European patent), CH (European patent), DE (European patent), DK (European patent), ES (European patent), FI, FR (European patent), GB (European patent), GR (European patent), IT (European patent), JP, KR, LU (European patent), MC (European patent), NL (European patent), SE (European patent), US.

Published
With international search report.

(54) Title: STABILIZED ENZYMES

(57) Abstract

This invention relates to novel stabilized enzymes. More specifically the invention relates to novel stabilized enzymes, in which a naturally occurring amino acid residue (other than proline) has been substituted with a proline residue at one or more positions; at which position(s) the dihedral angles φ (phi) constitute values within the interval [-90° < φ < -40°]; preferably the dihedral angles φ (phi) and ψ (psi) constitute values within the intervals [-90° < φ < -40°] and [-180° < ψ < -150° or -80 < $g c) < 10 or 100 < ψ < 180]; and which position(s) is/are not located in regions in which the enzyme is characterized by possessing α-helical or β-sheet structure. The invention also relates to nucleotide sequences encoding the novel stabilized enzymes, and expression vectors and host organisms containing the nucleotide sequences. The invention also relates to detergent compositions comprising the stabilized enzymes.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	ES	Spain	MG	Madagascar
AU	Australia	FI	Finland	ML	Mali
BB	Barbados	FR	France	MN	Mongolia
BE	Belgium	GA	Gabon	MR	Mauritania
BF	Burkina Faso	GB	United Kingdom	MW	Malawi
BG	Bulgaria	GN	Guinea	NL	Netherlands
BJ	Benin	GR	Greece	NO	Norway
BR	Brazil	HU	Hungary	PL	Poland
CA	Canada	IT	Italy	RO	Romania
CF	Central African Republic	JP	Japan	RU	Russian Federation
CG	Congo	KP	Democratic People's Republic of Korea	SD	Sudan
CH	Switzerland			SE	Sweden
CI	Côte d'Ivoire	KR	Republic of Korea	SN	Senegal
CM	Cameroon	LI	Liechtenstein	SU	Soviet Union
CS	Czechoslovakia	LK	Sri Lanka	TD	Chad
DE	Germany	LU	Luxembourg	TG	Togo
DK	Denmark	MC	Monaco	US	United States of America

STABILIZED ENZYMES

TECHNICAL FIELD

This invention relates to novel stabilized enzymes. More specifically the invention relates to novel stabilized enzymes, in which a naturally occurring amino acid residue (other than proline) has been substituted with a proline residue at one or more positions; at which position(s) the dihedral angles ϕ (phi) constitute values within the interval [-90°<ϕ<-40°]; preferably the dihedral angles ϕ (phi) and ψ (psi) constitute values within the intervals [-90°<ϕ<-40°] and [-180°<ψ<-150° or -80<ψ<10 or 100<ψ<180]; and which position(s) is/are not located in regions in which the enzyme is characterized by possessing α-helical or β-sheet structure.

The invention also relates to nucleotide sequences encoding the novel stabilized enzymes, and expression vectors and host organisms containing the nucleotide sequences. The invention also relates to detergent compositions comprising the stabilized enzymes.

BACKGROUND ART

The Structure of Proteins

Enzymes are globular proteins and quite compact due to the considerable amount of folding of the long polypeptide chain. The polypeptide chain essentially consists of the "backbone" and its "side-groups". As the peptide bond is planar, only rotations around the C_α-N axis and the C_α-C' axis are permitted. Rotation around the C_α-N bond of the peptide backbone is denoted by the torsion angle ϕ (phi), rotation around the C_α-C' bond by ψ (psi) [*vide* e.g. *Creighton, T.E.* (1984), Proteins; W.H. Freeman and Company, New York]. The choice of the values of these angles of rotation is made by assigning the maximum value of +180° (which is identical to -180°) to the maximally extended chain. In the fully extended polypeptide chain, the N, C_α and C' atoms are all "trans" to each other. In the "cis" configuration, the angles ϕ and ψ are assigned the value of 0°. Rotation from this position around the bonds, so that the atoms viewed behind the rotated bond move

"counterclockwise", are assigned negative values by definition, those "clockwise" are assigned positive values. Thus, the values of the torsion angles lie within the range -180° to +180°.

Since the C_α-atoms are the swivel point for the chain, the side-groups (R-groups) associated with the C_α-atoms become extremely important with respect to the conformation of the molecule.

The term "conformation" defines the participation of the secondary and tertiary structures of the polypeptide chains in moulding the overall structure of a protein. The correct conformation of a protein is of prime importance to the specific structure of a protein and contributes greatly to the unique catalytic properties (i.e. activity and specificity) of enzymes and their stability.

The amino acids of polypeptides can be divided into four general groups: nonpolar, uncharged polar, and negatively or positively charged polar amino acids. A protein molecule, when submerged in its aqueous environment in which it normally occurs, tends to expose a maximum number of its polar side-groups to the surrounding environment, while a majority of its nonpolar sidegroups are oriented internally. Orientation of the side-groups in this manner leads to a stabilization of protein conformation.

Proteins, thus, exist in a dynamic equilibrium between a folded and ordered state, and an unfolded and disordered state. This equilibrium in part reflects the short range interactions among the different segments of the polypeptide chain, which tends to stabilize the overall structure of proteins. Thermodynamic forces simultaneously tend to promote randomization of the unfolding molecule.

A way to engineer stabilized proteins is to reduce the extent of unfolding by decreasing the flexibility of the polypeptide backbone, and simultaneously decreasing the entropy of the unfolded chain. So far only few attempts have been made to implement this rationale in the development of novel stabilized enzymes.

A general principle of increasing protein thermostability has been provided [*Suzuki, Y.* (1989); Proc. Japan Acad.; **65** Ser. B]. In this article Suzuki states that the thermostability of a globular protein can be enhanced cumulatively to a great extent by increasing the frequency of proline occurrence at the second site of β-turns without significant alterations in the secondary and tertiary structures as

well as in the catalytic function of enzymes. The principle is based on various facts and findings, among these the fact that proline residues show a strong tendency to occur preferentially at the second site of β-turns [*Levitt, M* (1978); Biochemistry; **17** 4277-4285; and *Chou, P.Y. & Fasman, G.D* (1977); J. Mol. Biol.; **115** 135-175]. The principle is restricted to insertion of proline into the second site of β-turns in proteins, no other sites are mentioned.

International Patent Publication WO 89/01520 (Cetus Corporation, USA) provides a method for increasing the stability of a protein by decreasing the configurational entropy of unfolding the protein. The method is applied on a *Streptomyces rubiqinosus* xylose isomerase, and it involves substitution of an amino acid with proline, or replacement of glycine with alanine, at predicted substitution sites.

It is an object of this invention to provide novel enzymes having improved stability.

SUMMARY OF THE INVENTION

The present invention provides novel stabilized enzymes, in which a naturally occurring amino acid residue (other than proline) has been substituted with a proline residue at one or more positions, at which position(s) the dihedral angles ϕ (phi) constitute values within the interval [-90°<ϕ<-40°], preferably the dihedral angles ϕ (phi) and ψ (psi) constitute values within the intervals [-90°<ϕ<-40°] and [-180°<ψ<-150° or -80<ψ<10 or 100<ψ<180], and which position(s) is/are not located in regions in which the enzyme is characterized by possessing α-helical or β-sheet structure.

In another aspect, the invention relates to nucleotide sequences encoding novel stabilized enzymes, expression vectors comprising these nucleotide sequences, and host organisms containing these expression vectors.

Enzymes

In the context of this invention, an enzyme may be any enzyme apart from proteases, such as e.g. a lipase, a cellulase, a peroxidase, a xylanase, or an amylase.

Amino Acids

As abbreviations for amino acids the following symbols are used:

A	=	Ala	=	Alanine
C	=	Cys	=	Cysteine
D	=	Asp	=	Aspartic acid
E	=	Glu	=	Glutamic acid
F	=	Phe	=	Phenylalanine
G	=	Gly	=	Glycine
H	=	His	=	Histidine
I	=	Ile	=	Isoleucine
K	=	Lys	=	Lysine
L	=	Leu	=	Leucine
M	=	Met	=	Methionine
N	=	Asn	=	Asparagine
P	=	Pro	=	Proline
Q	=	Gln	=	Glutamine
R	=	Arg	=	Arginine
S	=	Ser	=	Serine
T	=	Thr	=	Threonine
V	=	Val	=	Valine
W	=	Trp	=	Tryptophan
Y	=	Tyr	=	Tyrosine
B	=	Asx	=	Asp (D) or Asn (N)
Z	=	Glx	=	Glu (E) or Gln (Q)
X	=	an arbitrary amino acid		
*	=	deletion or absent amino acid		

Enzyme Variants

A stabilized enzyme of this invention is an enzyme variant or mutated enzyme. By an enzyme variant or mutated enzyme is meant an enzyme obtainable by alteration of a DNA nucleotide sequence of the parent gene or its derivatives. The enzyme variant or mutated enzyme may be expressed and produced when the DNA nucleotide sequence encoding the enzyme is inserted into a suitable vector in a

suitable host organism. The host organism is not necessarily identical to the organism from which the parent gene originated.

Amino Acid Numbering

In the context of this invention, a specific numbering of amino acid residue positions in enzyme is employed.

In describing the various enzyme variants produced or contemplated according to the invention, the following nomenclatures were adapted for ease of reference:

[Original amino acid; Position; Substituted amino acid]

According to this, the substitution of glycine with proline in position 225 is designated as G225P.

If a substitution, e.g. G225P, is made by mutation in e.g. Humicola lanuginosa lipase, the product is designated "Humicola lanuginosa/G225P lipase".

Enzymatic Activity

In the context of this invention, the enzymatic activity of lipases is expressed in Lipase Units. A Lipase Unit (LU) is the amount of enzyme which under standard conditions, i.e. 30.0°C; pH 7.0; tributyrine substrate, liberates 1 μmol titrable butyric acid per minute. A folder AF 95/5 describing this analytical method is available upon request to Novo Nordisk A/S, Denmark, which folder is hereby included by reference.

In the context of this invention, the enzymatic activity of cellulases is expressed in Novo Cellulase Units (NCU). One unit is defined as the amount of enzyme which, under standard conditions (i.e. at pH 4.80; 0.1 M acetate buffer; 10 g/l Hercules CMC type 7 LFD as substrate; an incubation temp. of 40.0°C; an incubation time of 20 min; and an enzyme concentration of approximately 0.041 NCU/ml) forms an amount of reducing carbohydrates equivalent to 1 μmol glucose per minute. A folder AF 187.2/1 describing this analytical method is available upon request to Novo Nordisk A/S, Denmark, which folder is hereby included by reference.

BRIEF DESCRIPTION OF DRAWINGS

The present invention is further illustrated by reference to the accompanying drawings, in which:

Fig. 1 (sheets 1/17-9/17) shows the nucleotide sequence of an expression cassette of the Humicola lanuginosa lipase;

Fig. 2 shows the restriction map of plasmid pAHL;

Figs. 3-4 show the PCR technique for producing Humicola lipase variants;

Fig. 5 (sheets 13/17-14/17) shows the amino acid sequence (and a corresponding DNA nucleotide sequence) of the Humicola lanuginosa lipase, obtained according to European Patent Application 305,216;

Fig. 6 (sheets 15/17-16/17) shows the amino acid sequence of the Humicola insolens cellulase obtained according to International Patent Application WO 91/17243; and

Fig. 7 shows the stability at 70°C and pH 7 of Humicola lanuginosa/-G225P lipase (♦) of the invention compared to wildtype enzyme (□).

DETAILED DISCLOSURE OF THE INVENTION

The present invention provides novel stabilized enzymes in which a naturally occurring amino acid residue (other than proline) has been substituted with a proline residue at one or more positions, at which position(s) the dihedral angles ϕ (phi) constitute values in the interval [-90°<ϕ<-40°], preferably the dihedral angles ϕ (phi) and ψ (psi) constitute values in the intervals [-90°<ϕ<-40°] and [-180°-<ψ<150° or -80<ψ<10 or 100<ψ<180], and which position(s) is/are not located in regions, in which the enzyme is characterized by possessing α-helical or β-sheet structure.

In the context of this invention, a stabilized enzyme of the invention is an enzyme variant or mutated enzyme, being functionally equivalent or having structural features similar to a naturally occurring enzyme, and in which enzyme a naturally occurring amino acid residue (other than proline) has been substituted with

a proline residue at one or more positions, at which position(s) the dihedral angles ϕ (phi) constitute values within the intervals [-90°<ϕ<-40°], preferably the dihedral angles ϕ (phi) and ψ (psi) constitute values within the intervals [-90°<ϕ<-40°] and [-180°<ψ<-150° or -80<ψ<10 or 100<ψ<180], and which position(s) is/are not located in regions in which the enzyme is characterized by possessing α-helical or β-sheet structure.

Moreover, in the context of this invention, a stabilized enzyme is an enzyme having improved stability, e.g. in respect to thermal stability, storage stability, etc., when compared to the parent enzyme.

Defining Secondary Structure of Proteins

The stabilized enzymes of the invention may be obtained by subjecting the enzyme in question to analysis for secondary structure, identifying residues in the enzyme having dihedral angles ϕ (phi) confined to the interval [-90°<ϕ<-40°], preferably having dihedral angles ϕ (phi) and ψ (psi) confined to the intervals [-90°<ϕ<-40°] and [-180°<ψ<-150° or -80<ψ<10 or 100<ψ<180], excluding residues located in regions in which the enzyme is characterized by possessing α-helical or β-sheet structure, if a proline residue is not already at the identified position(s), substitution of the naturally occurring amino acid residue with a proline residue at the identified position(s), preferably by site directed mutagenesis applied on a gene encoding the enzyme in question, and gene expression by insertion of the gene encoding the stabilized enzyme in a suitable host organism, followed by cultivation of said host organism in a suitable nutrient medium, and recovery of the desired enzyme.

This method of obtaining stabilized enzymes includes subjecting the enzyme in question to analysis for secondary structure. In one way to perform such analysis the atomic structure of the enzyme has to be elucidated. The atomic structure may be determined by X-ray diffraction techniques. X-ray diffraction techniques are described by e.g. Hendrickson [*Hendrickson, W.A.* (1987); X-ray diffraction; in Protein Engineering (Ed: Oxender, D.L and Fox, C.F.), ch. 1; Alan R. Liss, Inc.] and Creighton [*Creighton*, T.E.; *supra*; ch. 6].

The crystal structure of a *Rhizomucor miehei* lipase has been deduced [*vide Brady L., Brzozowski A.M., Derewenda Z.S., Dodson E., Dodson G., Tolley S., Turkenburg J.P., Christiansen L., Huge-Jensen B., Norskov L. Thim L. & Menge U.* (1990); "A serine protease triad forms the catalytic centre of a triacylglycerol lipase"; Nature, Vol. 343 **6260** 767-770], and the coordinates have been deposited and are available from the Brookhaven Protein Data Bank [*Bernstein et al.* (1977); J. Mol. Biol. **112** 535-542].

When the atomic structure has been determined, it is possible to compute dihedral angles from the atomic coordinates. Moreover, it is possible to assign secondary structure elements. The secondary structure elements are defined on the basis of hydrogen bindings. Cooperative secondary structure is recognized as repeats of the elementary hydrogen-bonding patterns "turn" and "bridge". Repeating turns are "helixes", repeating bridges are "ladders", connected ladders are "sheets".

Analysis for secondary structure elements requires a computerized compilation of structure assignments and geometrical features extracted from atomic coordinates. The conventional method to elucidate the secondary structure of a protein, based on its atomic coordinates, is described by Kabsch and Sander [*Kabsch, W. and Sander, C.* (1983); Biopolymers, **22** 2577-2637]. In this article an algorithm for extracting structural features from the atomic coordinates by a pattern-recognition process is provided. First, H-bonds are identified based on electrostatic interactions between pairs of H-bonding groups. Next, the patterns of H-bonding are used to define secondary structure elements such as turns (T), bends (S), bridges (B), helices (G,H,I), β-ladders (E) and β-sheets (E).

A computer program DSSP (Define Secondary Structure of Proteins), enabling the computation of Kabsch & Sander files and written in standard PASCAL, is available from the Protein Data Bank, Chemistry Dept., Brookhaven National Laboratory, Upton, N.Y. 11973.

Analysis for secondary structure and calculation of the dihedral angles may also be carried out by other methods, e.g. by model building [*vide* e.g. *Sutcliffe, Haneef, Carney and Blundell* (1987); Protein Engineering, 1 (5) 377-384].

After the dihedral angles ϕ (phi) and ψ (psi) for the amino acids have been calculated, based on the atomic structure in the crystalline enzyme, it is possible to select position(s) which has/have dihedral phi and psi angles favourable for substitution with a proline residue. The aliphatic side chain of proline residues is bonded covalently to the nitrogen atom of the peptide group. The resulting cyclic five-membered ring consequently imposes a rigid constraint on the rotation about the N-C_α bond of the peptide backbone and simultaneously prevents the formation of hydrogen bonding to the backbone N-atom. For these structural reasons, prolines are generally not compatible with α-helical and β-sheet secondary conformations. Due to the same rotational constraint about the C_α-N bond, and due to the requirement that neighbouring amino acids in the chain are not disturbed, the magnitudes of the dihedral angles phi and psi (and in particular phi) are confined to limited intervals for proline residues in polypeptides. The dihedral angles for proline residues in polypeptides are almost exclusively within the intervals [-90°<ø<-40°], preferably the intervals [-90°<ϕ<-40°] and [-180°<ψ<-150° or -80<ψ<10 or 100<ψ<180]. In this context, both cis- and trans-proline residues are considered.

A proline residue may already occur at one or more positions pointed out by the procedure described above, and then a substitution is, of course, irrelevant. Otherwise, the method includes substitution of the naturally occurring amino acid residue with a proline residue.

When performing this method on a Humicola lanuginosa lipase, obtained as described in European Patent Application No. 305,216, the positions cited in Table 1 are revealed. When performing this method on a *Rhizomucor insolens* cellulase, obtained as described in International Patent Application WO 91/17243, the positions cited in Table 2 are revealed.

Certainly it is not to be expected that a substitution into proline at every of the predicted positions would bring about improved stability of the enzyme. At some of the positions revealed by the method, a substitution of a naturally occurring amino acid residue into a proline residue may cause destabilisation due to unpredictable factors, such as loss of essential flexibility, loss of H-bond possibilities, unpredictable sterical hindrance, etc. Such "critical sites" are not always to be foreseen.

However, it is to be expected that the stabilizing (or destabilizing) effects of individual substitutions are additive [*vide* e.g. *Wells, J.A.* (1990); Biochemistry; **29** (37) 8510-8517].

Preferred Enzymes

Preferably, an enzyme of the invention is a stabilized lipase, a stabilized cellulase, a stabilized peroxidase, a stabilized xylanase, or a stabilized amylase.

In one aspect an enzyme of the invention is a stabilized lipase, the lipase being obtainable from a member of the genera *Humicola*, *Rhizomucor*, *Candida*, or *Pseudomonas*.

In a specific embodiment of the invention, the enzyme is a stabilized lipase, the lipase being obtainable from a strain of *Humicola lanuginosa*, *Rhizomucor miehei*, *Candida antarctica*, or *Pseudomonas cepacia*. Lipases obtainable from *Humicola* are described in e.g. European Patent Application No. 305,216 and International Patent Application WO 89/01969. Preferably, the lipase is a stabilized Humicola lanuginosa lipase, the lipase being obtained as described in European Patent Application No. 305,216.

In another specific embodiment of the invention, the enzyme is a stabilized lipase, the lipase being obtainable from a member of the genus *Humicola*, which stabilized lipase comprises a substitution into a proline residue at one or more of the positions listed in Table 1, or positions analogous hereto. Preferably, the enzyme is a stabilized lipase, the lipase being obtainable from a strain of *Humicola lanuginosa*.

In yet another specific embodiment, an enzyme of the invention is a stabilized lipase, the lipase being obtainable from a member of the genus *Humicola*, which stabilized lipase comprises one or more of the following substitutions: A28P, G61P, N101P, S105P, D111P, D165P, R209P, S224P, G225P, T226P, R232P, N233P, I241P, T244P, (according to the amino acid sequence presented in Fig. 5) or at positions analogous hereto. Preferably, the enzyme is a stabilized lipase, the lipase being obtainable from a strain of *Humicola lanuginosa*.

In a more specific aspect, an enzyme of the invention is Humicola lanuginosa/G225P lipase or Humicola lanuginosa/T244P lipase.

Table 1

Proline Mutants Proposed in Humicola lanuginosa Lipase Based on phi and psi Angles.

Criteria -90°<phi<-40°;
Neither part of an alpha helix nor a beta sheet structure.

AA numbers	phi angle	psi angle	Amino acid	Structure	Mutant & Comments
19	-57	-22	A	T	
20	-59	-39	A	T	
21	-62	-42	Y	T	
23	-65	175	G	S	
28	-84	-103	A	S	
43	-72	-23	E	T	
44	-41	-52	V	T	
45	-44	-48	E	T	
46	-86	33	K	T	
48	-81	82	D		
61	-88	-105	G	T	
71	-68	0	N	T	
83	-66	155	S		
84	-83	8	R	S	
101	-64	127	N	T	
103	-89	-24	I	T	
105	-86	-79	S	S	
111	-68	172	D	T	
160	-80	174	R	T	
165	-74	129	D		
173	-83	142	A		
175	-76	145	R		
178	-81	-160	N		
203	-85	164	V	T	
213	-85	148	Y		
216	-74	167	S	S	
223	-58	-22	K		
224	-83	159	S	S	
225	-70	-178	G		
226	-77	138	T	T	
232	-45	-24	R	T	
234	-89	9	D	S	
241	-63	134	I	T	

244	-84	140	T	T
247	-54	-27	N	S
255	-78	-34	I	S
257	-79	-23	A	T
258	-69	-13	H	T

In another aspect an enzyme of the invention is a stabilized cellulase, the cellulase being obtainable from a member of the genera *Rhizomucor.* In a specific embodiment an enzyme of the invention is a stabilized cellulase being obtainable from a strain of *Rhizomucor miehei*. Cellulases obtainable from *Rhizomucor miehei* are described in e.g. International Patent Application WO 91/17243.

In another specific embodiment of the invention, the enzyme is a stabilized *Rhizomucor miehei* cellulase, which stabilized cellulase comprises a substitution into a proline residue at one or more of the positions listed in Table 2, or positions analogous hereto.

In yet another specific embodiment, an enzyme of the invention the enzyme is a stabilized *Rhizomucor miehei* cellulase, which stabilized cellulase comprises a substitution into a proline residue at one or more of the following positions: A33P, A78P, I131P, A162P, or at positions analogous hereto.

Table 2

Proline Mutants Proposed in Rhizomucor miehei Cellulase Based on phi and psi Angles.

Criteria -90°<phi<-40°;
Neither part of an alpha helix nor a beta sheet struc- ture.

AA numbers	phi angle	psi angle	Amino acid	Struc- ture	Mutant & Comments
19	-68		A		
25	-42		N		
28	-62		V		
29	-64		F		
32	-78		N		
33	-64		A		
34	-84		N		
37	-66		R		
39	-89		T		
40	-82		D		
41	-73		F		
42	-95		D		
43	-61		A		
44	-58		K		
45	-62		S		
46	-66		G		
51	-80		G		
55	-69		S		
57	-70		A		
58	-86		D		
66	-75		D		
78	-52		A		
27	-75		G		
129	-70		V		
131	-78		I		
146	-53		R		
150	-78		I		
161	-61		D		
162	-64		A		
177	-76		A		
181	-84		S		
188	-61		Q		
189			C		
201	-75		R		

202	-68	N
204	-55	D
205	-77	G

The Effect of Prolinstabilization

The thermostability of purified lipase variants has been tested by a differential scanning calorimetry (DSC) method, and by activity determination at elevated temperatures (*vide* Example 3 for experimental data). The result of this experiment is shown in Table 3, below, and in Fig. 7.

Table 3

Stabilization relative to Humicola lanuginosa Lipase

Variant	Relative Stabilization Δ DSC
Humicola lanuginosa/S224P	-6.0 °C
Humicola lanuginosa/G225P	2.0 °C
Humicola lanuginosa/I241P	-6.2 °C

It appears from Table 3 that 3 of the variants constructed possess significantly improved thermostability, and 2 of the variants posses significantly decreased thermostability, when compared to the wild-type enzyme. These results clearly demonstrate that although a rationale exists for stabilization by introduction of proline residues into a protein by the Phi-Psi-Concept described in this specification, no conclusion as to the stabilizing effect of the individual variants are predictable.

Recombinantly Produced Enzymes

In the past, numerous processes have been developed for the production of polypeptides or proteins by means of the recombinant DNA technology. Mostly used for this purpose are *E. coli*, *Bacillus subtilis*, *Saccharomyces cerevisiae* and different *Aspergillus* strains, e.g. *A. oryzae* and *A. niger*. Especially the *Aspergillii* are attractive candidates as host microorganisms for recombinant DNA vectors being well-characterized and widely used microorganisms for the commercial production of enzymes. In *Aspergillus oryzae*, methods have been developed for transformation of the organism, and production of several enzymes, among these the *Humicola lanuginosa* and *Rhizomucor miehei* lipases (*vide* e.g. European Patent Applications Nos. 238,023 and 305,216, and International Patent Application No. WO 89/01969), and the *Humicola insolens* and *Fusarium oxysporium* cellulases (*vide* e.g. International Patent Application WO 91/17243), has also been demonstrated, which publications are hereby included by reference.

Expression of Polypeptides Biosynthetically

Upon transformation of an organism where the intention is production of a polypeptide or a protein, a DNA sequence is introduced into the organism. The sequence contains the coding region of the gene of interest flanked by transcription/translation start signals and transcription/translation termination signals. The coding region contains units of three base pairs, called codons, which upon translation of the transcribed gene are translated into amino acids, which again are assembled to give the polypeptide of interest.

Introducing Mutations in Polypeptides

By changing one or more specific codons in the coding region and transforming the host microorganism with these new coding regions, new polypeptides can be produced, which differ from the original polypeptide by one or more amino acids. Such alterations can be introduced by means of a technique generally known as "site-directed in vitro mutagenesis". A number of methods have been published. An early method is described by *Zoller & Smith* (1984); DNA, 3 (6) 479-488, and involves use of the single-stranded M13 bacteriophage. A preferred

method using PCR (polymerase chain reaction) is described by *Nelson & Long* (1989); Analytical Biochemistry, 180, 147-151. It involves a 3-step generation of a PCR fragment containing the desired mutation by using a chemically synthesized DNA oligonucleotide as one of the primers in the PCR reactions. From the PCR-generated fragment, a DNA fragment carrying the mutation can be isolated by cleavage with restriction enzymes and re-inserted into the expression plasmid. A third mutagenesis method takes advantage of restriction sites in the DNA coding region. By digesting the DNA with restriction enzymes at sites flanking the mutagenesis target, synthesizing a new fragment synthetically containing the desired mutation and cloning this new fragment between the restriction sites, a mutant coding region can be constructed.

All methods are generally applicable to investigations in the field called protein engineering which deals with the development of polypeptides with new or altered characteristics.

Transformation and expression may be accomplished by methods known in the art, e.g. as described in European Patent Application No. 305,216, which specification is hereby included by reference.

The microorganisms able to produce a stabilized enzyme of this invention can be cultivated by conventional fermentation methods in a nutrient medium containing assimilable carbon and nitrogen together with other essential nutrients, the medium being composed in accordance with the principles of the known art. Purification and recovery of the stabilized enzyme may also be conducted in accordance with methods known *per se*.

Nucleotide Sequences, Expression Vectors And Microorganisms

This invention also relates to DNA nucleotide sequences encoding a stabilized enzyme of the invention. The stabilized enzyme may be expressed and produced when DNA nucleotide sequence encoding this enzyme is inserted into a suitable vector in a suitable host organism. The host organism is not necessarily identical to the organism from which the parent gene originated. The construction of the mutated genes, vectors and mutant and transformed microorganisms may be carried out by any appropriate recombinant DNA technique, known in the art.

The invention also relates to expression vectors and host organisms containing a DNA nucleotide encoding a stabilized enzyme of this invention.

Detergent compositions

The present invention also comprises the use of stabilized enzymes of the invention in cleaning and detergent compositions and such compositions comprising one or more stabilized enzymes of the invention.

The detergent composition of the invention may comprise one or more surfactants, which may be of an anionic, non-ionic, cat-ionic, amphoteric or zwitter-ionic type, or a mixture of these. Typical examples of anionic surfactants are linear alkyl benzene sulfonates (LAS), alkyl sulfates (AS), alpha olefin sulfonates (AOS), alcohol ethoxy sulfates (AES) and alkali metal salts of natural fatty acids. Examples of non-ionic surfactants are alkyl polyethylene glycol ethers, nonylphenol poly-ethylene glycol ethers, fatty acids esters of sucrose and glucose, and esters of polyethoxylated alkyl glucoside.

The detergent composition of the invention may also contain other detergent ingredients known in the art such as builders, bleaching agents, bleach activators, anti-corrosion agents, sequestering agents, anti soil-redeposition agents, perfumes, stabilizers for the enzymes and bleaching agents, formulations aids, optical brighteners, foam boosters, chelating agents, fillers, fabric softeners, etc. The detergent composition of the invention may be formulated substantially as described in J. Falbe [*Falbe J.*; Surfactants in Consumer Products. Theory, Technology and Application; Springer Verlag 1987, *vide* in particular the section entitled "Frame formulations for liquid/powder heavy-duty detergents"].

It is at present contemplated that the detergent composition of the invention may contain a protease preparation in an amount corresponding to 0.0005-0.5 CPU of the proteolytic enzyme per litre of washing liquor. Generally, detergent compositions are used in dosages within the range 0.3 to 15 g detergent per litre wash liquor.

The detergent compositions of the invention can be formulated in any convenient form, such as powders, liquids, etc.

The enzymes of the invention may be included in a detergent composition by adding separate additives containing one or more enzymes, or by adding a combined additive comprising all of these enzymes.

The additive of the invention, i.e. a separated additive or a combined additive, can be formulated e.g. as granulates, liquids, slurries, etc. Preferred detergent additive formulations are non-dusting granulates, liquids, in particular stabilized liquids, slurries, or protected enzymes.

Dust free granulates may be produced according to e.g. GB Patent No. 1,362,365 or US Patent No. 4,106,991, and may optionally be coated by methods known in the art. The detergent enzymes may be mixed before or after granulation.

Liquid enzyme preparations may, for instance, be stabilized by adding a polyol such as e.g. propylene glycol, a sugar or sugar alcohol, lactic acid or boric acid, according to established methods. Other enzyme stabilizers are well known in the art.

Protected enzymes may be prepared according to the method disclosed in EP Patent Application No. 238,216.

The invention is further illustrated in the following examples, which are not intended to be in any way limiting to the scope of the invention as claimed.

EXAMPLE 1

Construction Of A Plasmid Expressing the G225P Variant of Humicola lanuginosa Lipase

The expression plasmid used in this specification is identical to p960, described in European Patent Application No. 305,216, except for minor modifications just 3' to the lipase coding regions. The modifications were made the following way:

p960 was digested with NruI and BamHI restriction enzymes. Between these two sites the BamHI/NheI fragment from plasmid pBR322, in which the NheI fragment was filled in with Klenow polymerase, was cloned, thereby creating plasmid

pAOI, which contained unique BamHI and NheI sites. Between these unique sites BamHI/XbaI fragments from p960 were cloned to give pAHL.

The sequence of the SalI/EcoRI fragment comprising the expression cassette of the *Humicola lanuginosa* lipase is shown in Fig. 1. The restriction map of plasmid pAHL containing this expression cassette is shown in Fig. 2. The plasmid was used as template for construction of some of the *Humicola* lipase variants. The method used is described in the following and is further outlined in Figs. 3 and 4. It was originally published by Nelson & Long in Analytical Biochemistry, 180, 147-151 (1989).

Linearization of Plasmid pAHL

The circular plasmid pAHL was linearized with the restriction enzyme SphI in the following 50 µl reaction mixture: 50 mM NaCl, 10 mM Tris-HCl, pH 7.9, 10 mM $MgCl_2$, 1 mM dithiothreitol, 1 µg plasmid and 2 units of SphI. The digestion was carried out for 2 hours at 37°C. The reaction mixture was extracted with phenol (equilibrated with Tris-HCl, pH 7.5) and precipitated by adding 2 volumes of ice-cold 96% ethanol. After centrifugation and drying of the pellet, the linearized DNA was dissolved in 50 µl H_2O and the concentration estimated on an agarose gel.

3-Step PCR Mutagenesis

As is also shown in Fig. 4, 3-step mutagenisation involved the use of four primers:

Mutagenisation primer (=A):

5'-GGGGACAAGGGTTGGAGATTTGATCCA-3'

PCR Helper 1 (=B):

5'-GGTCATCCAGTCACTGAGACCCTCTACCTATTAAATCGGC-3'

PCR Helper 2(=C):

5'-CCATGGCTTTCACGGTGTCT-3'

PCR Handle (=D):

5'-GGTCATCCAGTCACTGAGAC-3'

All 3 steps were carried out in the following buffer containing: 10 mM Tris-HCl, pH 8.3, 50 mM KCl, 1.5 mM $MgCl_2$, 0.001% gelatin, 0.2 mM dATP, 0.2 mM dCTP, 0.2 mM dGTP, 0.2 mM TTP, 2.5 units Taq polymerase.

In step 1, 100 pmol primer A, 100 pmol primer b, and 1 fmol linearized plasmid were added to a total of 100 μl of the above buffer and 15 cycles consisting of 2 minutes at 95°C, 2 minutes at 37°C and 3 minutes at 72°C were carried out.

The concentration of the PCR product was estimated on an agarose gel. Then, step 2 was carried out. 0.6 pmol step 1 product and 1 fmol linearized plasmid were contained in a total of 100 μl of the previously mentioned buffer and 1 cycle consisting of 5 minutes at 95°C, 2 minutes at 37°C and 10 minutes at 72°C was carried out.

To the step 2 reaction mixture, 100 pmol primer C and 100 pmol primer D were added (1 μl of each) and 20 cycles consisting of 2 minutes at 95°C, 2 minutes at 37°C and 3 minutes at 72°C were carried out. This manipulation comprised step 3 in the mutagenisation procedure.

Isolation of Mutated Restriction Fragment

The product from step 3 was isolated from an agarose gel and re-dissolved in 20 μl H_2O. Then, it was digested with the restriction enzymes BstXI and BglII in a total volume of 50 μl with the following composition: 100 mM NaCl, 50 mM Tris-HCl, pH 7.9, 10 mM $MgCl_2$, 1 mM DTT, 5 units of BglII and 10 units of BstXI. Incubation was at 37°C for 2 hours. The 200 bp BstXI/BglII fragment was isolated from an agarose gel.

Ligation To Expression Vector pAHL

The expression plasmid pAHL was cleaved with BstXI and BglII under conditions indicated above and the large fragment was isolated from an agarose gel. To this vector, the mutated fragment isolated above was ligated and the ligation mix was used to transform E. coli. Sequence analysis was carried out on the double-stranded plasmid using the di-deoxy chain termination procedure developed by Sanger. The plasmid was named pAHLG225P and is identical to pAHL, except for the altered codon.

The plasmid was then ready to be transformed into the host organism. Preferred host organisms are of the genus *Aspergillus*. Transformation and expression are carried out essentially as described in European Patent Application No. 305,216, *supra*.

Construction of S224P, I241P, and T244P Variants

The plasmids pAHLS224P, pAHLI241P, and pAHLT244P, encoding Humicola lanuginosa/S224P, Humicola lanuginosa/I241P, and Humicola lanuginosa/T244P, respectively, were constructed using the method described above, with the exception that the following primers were used as mutagenisation primer ("Primer A").

S224P: 5'-GACAAGGGTTCCTGGTTTGATCCAGTA-3'

I241P: 5'-GCCGGTGGCATCTGGGCCTTCTATCTT-3'

T244P: 5'-TATTGCCGCCGGGGGCATCGATGCC-3'

EXAMPLE 2

Purification Example

About 200 ml of supernatant were centrifuged and precipitate was discarded by decanting. Slowly ice cold ethanol was added to the supernatant on ice bath. The precipitate was centrifuged and discarded and pH af the supernatant was adjusted to 7 with NaOH.

The 70% ethanol supernatant were applied on 200 ml DEAE fast flow sepharose column equilibrated with 50 mM Tris acetate buffer, pH 7, and flow rate

5 ml/min. The column was washed until OD_{280} was less than 0.05. The bound protein was eluted with linear NaCl gradient up to 0.5 M NaCl in the same buffer, using 5 times column volume. The lipase activity was eluted between 0.1 and 0.15 M NaCl salt concentration.

The fractions containing the lipase activity were pooled and ammonium acetate was added to a final concentration of 0.8 M. The pool was applied on a Toyopearl™-Butyl column at a flow rate of 5 ml/min. The column was equilibrated with 0.8 M ammonium acetate and washed until OD_{280} was below 0.05.

The bound activity was eluted with Mili-Q water. The fractions with lipase activity were pooled and the conductivity of the pool was adjusted to less than that of a 50 mM Tris acetate buffer, pH 7.

The pool was applied on a 30 ml HPQ-sepharose column equilibrated with 50 mM Tris acetate buffer, pH 7, at a flow rate of 1 ml/min. The bound activity was eluted with a linear salt gradient up to 1 M NaCl.

EXAMPLE 3

Differential Scanning Calorimetry

Purified lipase variants of the invention were subjected to thermal analysis by Differential Scanning Calorimetry (DSC). Using this technique, the thermal denaturation temperature, T_d, is determined by heating an enzyme solution at a constant programmed rate.

A Differential Scanning Calorimeter, MC-2D, from MicroCal Inc., was used for the investigations. Enzyme solutions were preparaed in 50 mM TRIS-acetate, pH 7.0. Enzyme concentration ranged between 0.6 and 0.9 mg/ml, and a total volume of approximately 1.2 ml was used for each experiment. All samples were heated from 25°C to 90°C at a scan rate of 90°C/hour.

The results from this experiment are presented in Table 3 above.

Temperature Stability

The thermostability of a purified lipase variant of the invention was also tested by activity determination at 70°C.

Samples were diluted to 1 mg/ml using 0.1 M Tris buffer, pH 7.0. Test tubes containing 100 µl of enzyme solution were placed in a waterbath at 70°C for 10, 25, and 45 minutes, respectively.

The residual activity (LU/mg) was measured using the analytical method described in this specification.

The results from this experiment are presented in Fig. 7.

CLAIMS

1. A stabilized enzyme, in which a naturally occurring amino acid residue (other than proline) has been substituted with a proline residue at one or more positions; at which position(s) the dihedral angles ϕ (phi) constitute values within the interval [-90°<ϕ<-40°]; preferably the dihedral angles ϕ (phi) and ψ (psi) constitute values within the intervals [-90°<ϕ<-40°] and [-180°<ψ<-150° or -80<ψ<10 or 100<ψ<180]; and which position(s) is/are not located in regions in which the enzyme is characterized by possessing α-helical or β-sheet structure.

2. An enzyme of claim 1, being a stabilized lipase, a stabilized cellulase, a stabilized peroxidase, a stabilized xylanase, or a stabilized amylase.

3. An enzyme of claim 1, being a stabilized lipase, the lipase being obtainable from a member of the genus *Humicola*, *Rhizomucor*, *Candida*, or *Pseudomonas*.

4. A lipase of claim 3, the lipase being obtainable from a strain of *Humicola lanuginosa*, *Rhizomucor miehei*, *Candida antarctica*, or *Pseudomonas cepacia*.

5. A lipase of claim 4, being a stabilized Humicola lanuginosa lipase.

6. A lipase of claim 5, the lipase being obtainable from a member of the genus *Humicola*, which stabilized lipase comprises a substitution into a proline residue at one or more of the positions listed in Table 1, or positions analogous hereto.

7. A lipase of claim 6, the lipase being obtainable from a strain of *Humicola lanuginosa*.

8. A lipase of claim 6, the lipase being obtainable from a member of the genus *Humicola*, which stabilized lipase comprises a substitution into a proline residue at one or more of the positions: A28P; G61P; N101P; S105P; D111P; D165P; R209P; S224P; G225P; T226P; R232P; N233P; I241P; T244P, or positions analogous hereto.

9. A lipase of claim 8, the lipase being obtainable from a strain of *Humicola lanuginosa*.

10. *Humicola lanuginosa*/G225P lipase.

11. *Humicola lanuginosa*/T244P lipase.

12. An enzyme of claim 1, being a stabilized cellulase, the cellulase being obtainable from a member of the genus *Humicola*, *Fusarium*, *Trichoderma*, *Myceliophthora*, *Phanerochaete*, *Schizophyllum*, *Penicillium*, *Aspergillus*, or *Geotricum*.

13. A cellulase of claim 12, the cellulase being obtainable from a strain of *Fusarium oxysporum*.

14. A cellulase of claim 12, the cellulase being obtainable from a strain *Humicola insolens*, which stabilized cellulase comprises a substitution into a proline residue at one or more of the positions listed in Table 2, or positions analogous hereto.

15. A cellulase of claim 14, which stabilized lipase comprises a substitution into a proline residue at one or more of the positions: A33P, A78P, I131P, or A162P, or positions analogous hereto.

16. A nucleotide sequence encoding a stabilized enzyme of any of claims 1-15.

17. An expression vector containing the nucleotide sequence according to claim 16 encoding a stabilized enzyme.

18. A host organism containing an expression vector carrying a nucleotide sequence encoding a stabilized enzyme of any of claims 1-15.

19. A detergent composition comprising an enzyme according to any of claims 1-15.

20. A detergent additive comprising an enzyme according to any of claims 1-15, provided in the form of a granulate, preferably a non-dusting granulate, a liquid, in particular a stabilized liquid, a slurry, or a protected enzyme.

1/17

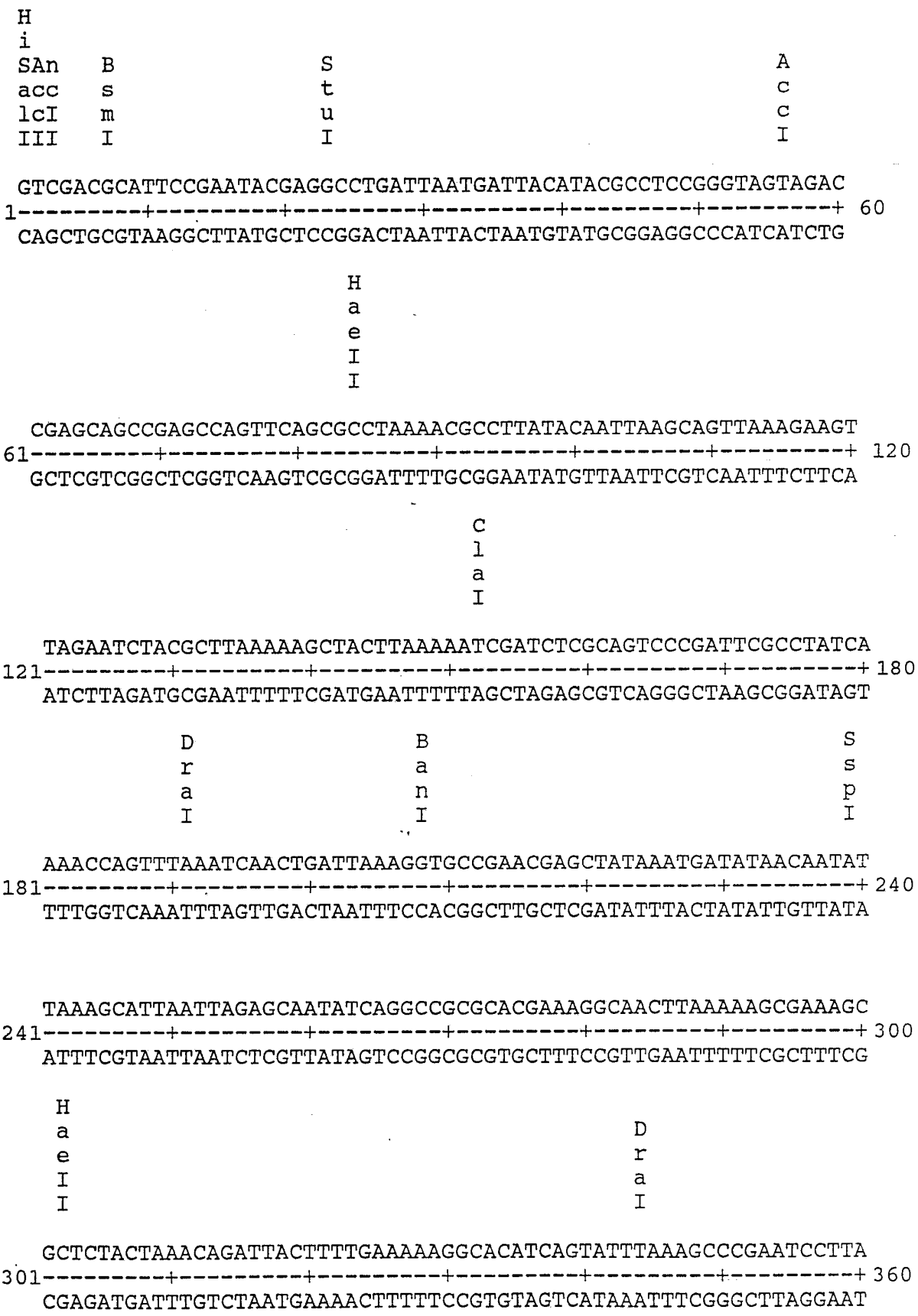

HindIII
SAnaclcIII
BsmI
StuI
AccI
GTCGACGCATTCCGAATACGAGGCCTGATTAATGATTACATACGCCTCCGGGTAGTAGAC
CAGCTGCGTAAGGCTTATGCTCCGGACTAATTACTAATGTATGCGGAGGCCCATCATCTG
HaeIII
CGAGCAGCCGAGCCAGTTCAGCGCCTAAAACGCCTTATACAATTAAGCAGTTAAAGAAGT
GCTCGTCGGCTCGGTCAAGTCGCGGATTTTGCGGAATATGTTAATTCGTCAATTTCTTCA
ClaI
TAGAATCTACGCTTAAAAAGCTACTTAAAAATCGATCTCGCAGTCCCGATTCGCCTATCA
ATCTTAGATGCGAATTTTTCGATGAATTTTTAGCTAGAGCGTCAGGGCTAAGCGGATAGT
DraI
BanI
SspI
AAACCAGTTTAAATCAACTGATTAAAGGTGCCGAACGAGCTATAAATGATATAACAATAT
TTTGGTCAAATTTAGTTGACTAATTTCCACGGCTTGCTCGATATTTACTATATTGTTATA
TAAAGCATTAATTAGAGCAATATCAGGCCGCGCACGAAAGGCAACTTAAAAAGCGAAAGC
ATTTCGTAATTAATCTCGTTATAGTCCGGCGCGTGCTTTCCGTTGAATTTTTCGCTTTCG
HaeIII
DraI
GCTCTACTAAACAGATTACTTTTGAAAAAGGCACATCAGTATTTAAAGCCCGAATCCTTA
CGAGATGATTTGTCTAATGAAAACTTTTTCCGTGTAGTCATAAATTTCGGGCTTAGGAAT
1
60
61
120
121
180
181
240
241
300
301
360

Fig. 1

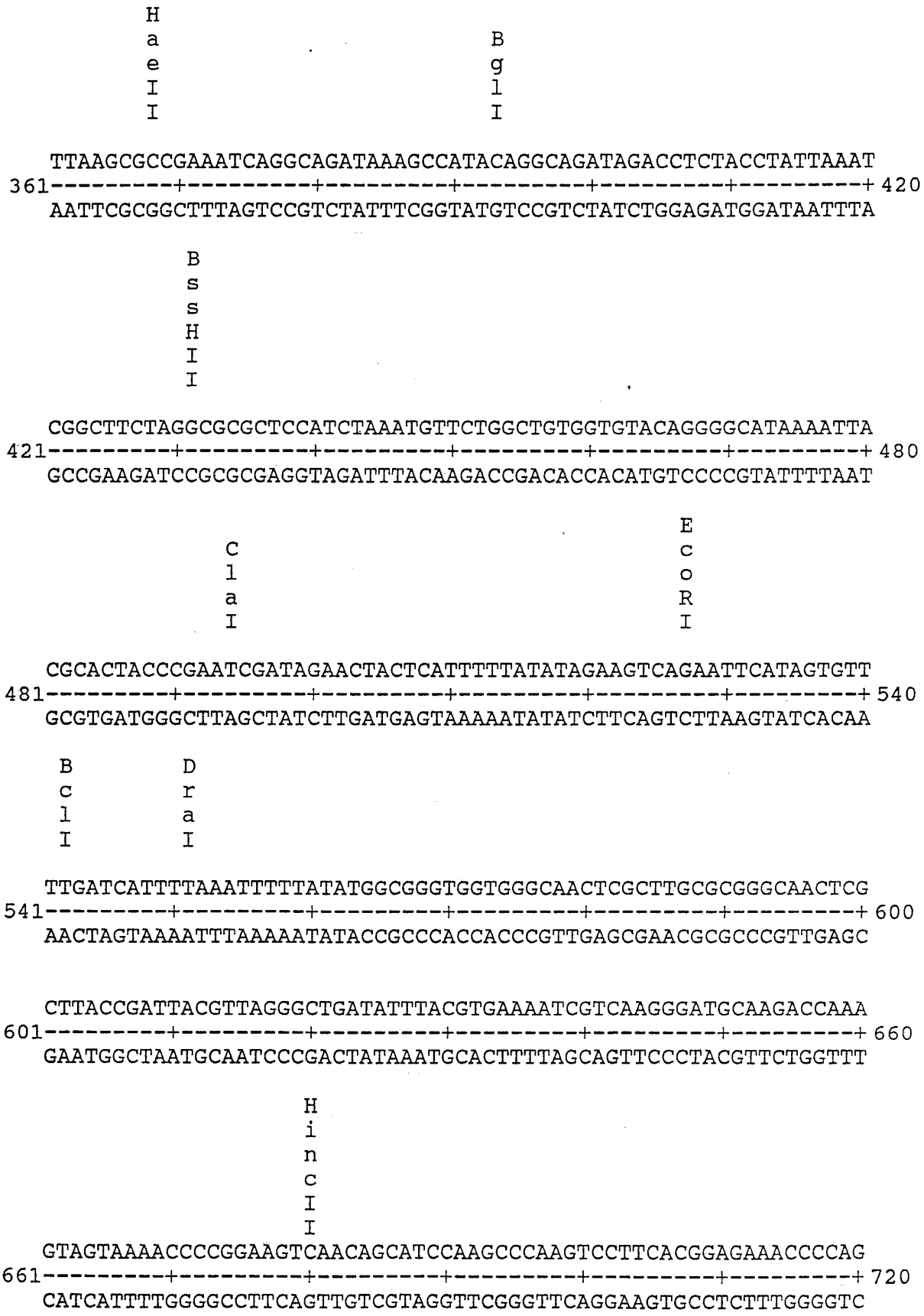

HaeIII
BglI
TTAAGCGCCGAAATCAGGCAGATAAAGCCATACAGGCAGATAGACCTCTACCTATTAAAT
361
AATTCGCGGCTTTAGTCCGTCTATTTCGGTATGTCCGTCTATCTGGAGATGGATAATTTA
420
BssHII
CGGCTTCTAGGCGCGCTCCATCTAAATGTTCTGGCTGTGGTGTACAGGGGCATAAAATTA
421
GCCGAAGATCCGCGCGAGGTAGATTTACAAGACCGACACCACATGTCCCCGTATTTTAAT
480
ClaI
EcoRI
CGCACTACCCGAATCGATAGAACTACTCATTTTTATATAGAAGTCAGAATTCATAGTGTT
481
GCGTGATGGGCTTAGCTATCTTGATGAGTAAAAATATATCTTCAGTCTTAAGTATCACAA
540
BclI
DraI
TTGATCATTTTAAATTTTTATATGGCGGGTGGTGGGCAACTCGCTTGCGCGGGCAACTCG
541
AACTAGTAAAATTTAAAAATATACCGCCCACCACCCGTTGAGCGAACGCGCCCGTTGAGC
600
CTTACCGATTACGTTAGGGCTGATATTTACGTGAAAATCGTCAAGGGATGCAAGACCAAA
601
GAATGGCTAATGCAATCCCGACTATAAATGCACTTTTAGCAGTTCCCTACGTTCTGGTTT
660
HincII
GTAGTAAAACCCCGGAAGTCAACAGCATCCAAGCCCAAGTCCTTCACGGAGAAACCCCAG
661
CATCATTTTGGGGCCTTCAGTTGTCGTAGGTTCGGGTTCAGGAAGTGCCTCTTTGGGGTC
720

Fig. 1

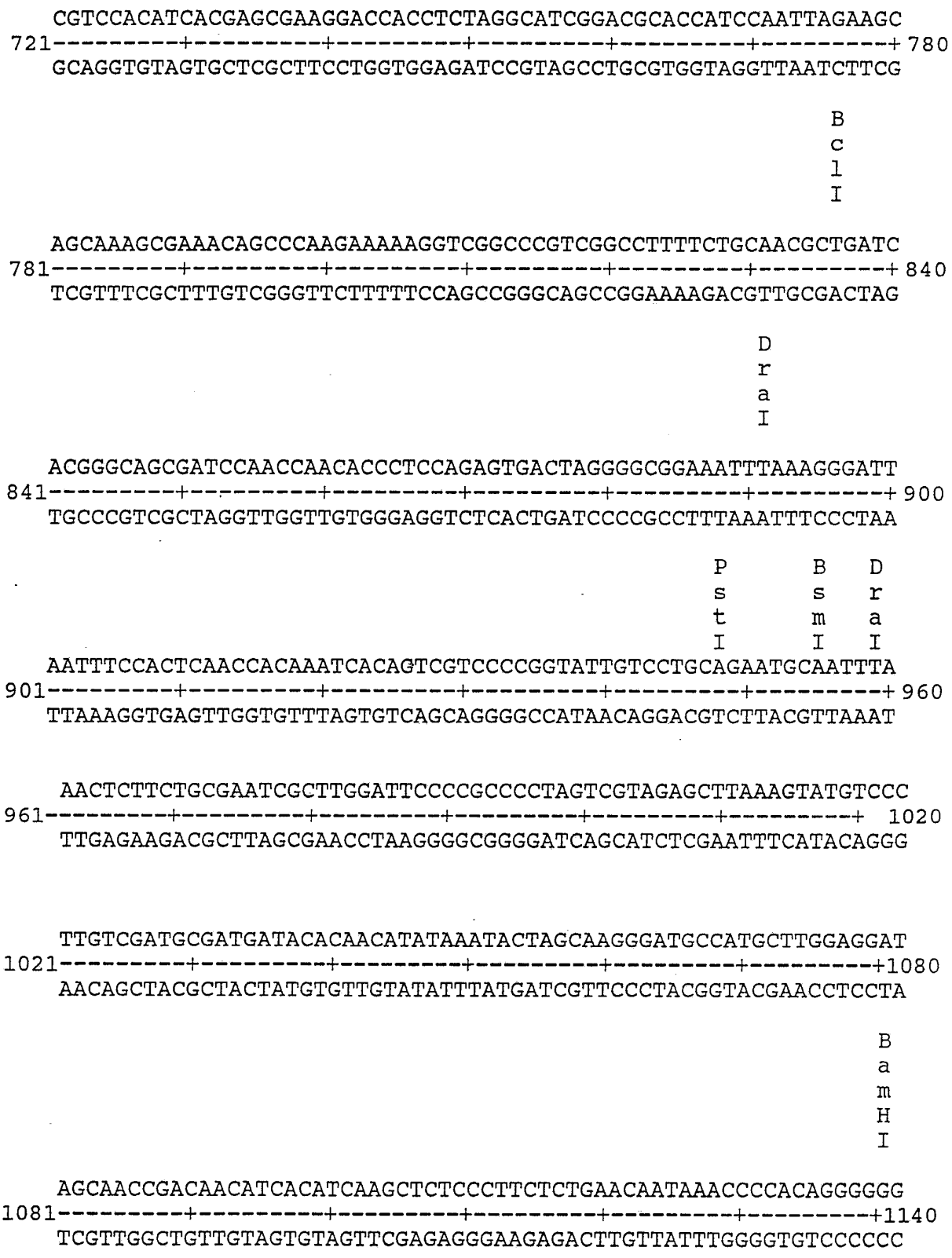

CGTCCACATCACGAGCGAAGGACCACCTCTAGGCATCGGACGCACCATCCAATTAGAAGC
721
780
GCAGGTGTAGTGCTCGCTTCCTGGTGGAGATCCGTAGCCTGCGTGGTAGGTTAATCTTCG
B
c
l
I
AGCAAAGCGAAACAGCCCAAGAAAAAGGTCGGCCCGTCGGCCTTTTCTGCAACGCTGATC
781
840
TCGTTTCGCTTTGTCGGGTTCTTTTTCCAGCCGGGCAGCCGGAAAAGACGTTGCGACTAG
D
r
a
I
ACGGGCAGCGATCCAACCAACACCCTCCAGAGTGACTAGGGGCGGAAATTTAAAGGGATT
841
900
TGCCCGTCGCTAGGTTGGTTGTGGGAGGTCTCACTGATCCCCGCCTTTAAATTTCCCTAA
P
s
t
I
B
s
m
I
D
r
a
I
AATTTCCACTCAACCACAAATCACAGTCGTCCCCGGTATTGTCCTGCAGAATGCAATTTA
901
960
TTAAAGGTGAGTTGGTGTTTAGTGTCAGCAGGGGCCATAACAGGACGTCTTACGTTAAAT
AACTCTTCTGCGAATCGCTTGGATTCCCCGCCCCTAGTCGTAGAGCTTAAAGTATGTCCC
961
1020
TTGAGAAGACGCTTAGCGAACCTAAGGGGCGGGGATCAGCATCTCGAATTTCATACAGGG
TTGTCGATGCGATGATACACAACATATAAATACTAGCAAGGGATGCCATGCTTGGAGGAT
1021
1080
AACAGCTACGCTACTATGTGTTGTATATTTATGATCGTTCCCTACGGTACGAACCTCCTA
B
a
m
H
I
AGCAACCGACAACATCACATCAAGCTCTCCCTTCTCTGAACAATAAACCCCACAGGGGGG
1081
1140
TCGTTGGCTGTTGTAGTGTAGTTCGAGAGGGAAGAGACTTGTTATTTGGGGTGTCCCCCC

Fig. 1

4/17

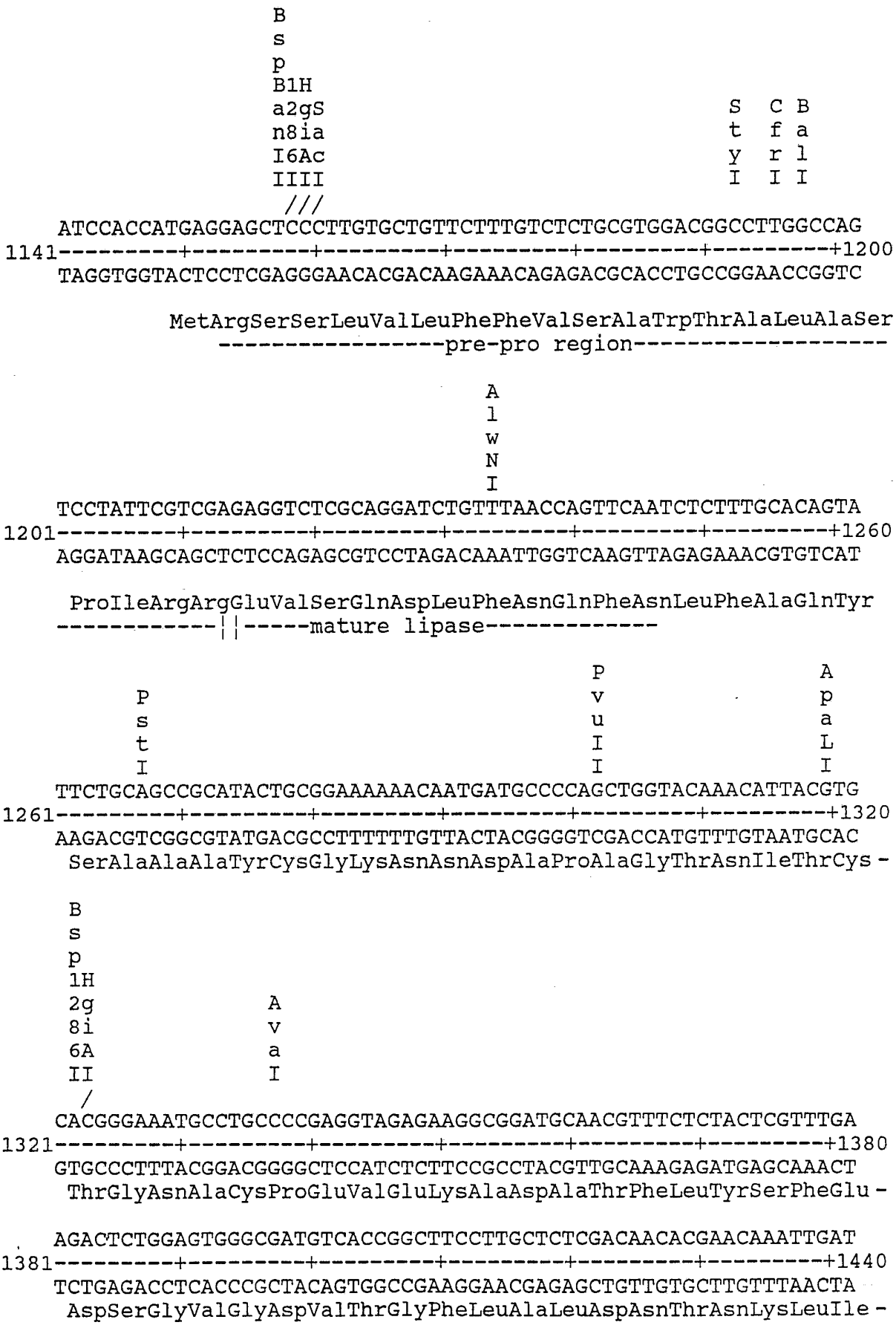

B
s
p
B1H
a2gS
n8ia
I6Ac
IIII
///
S t y I
C f r I
B a l I
ATCCACCATGAGGAGCTCCCTTGTGCTGTTCTTTGTCTCTGCGTGGACGGCCTTGGCCAG
1141---------+---------+---------+---------+---------+---------+1200
TAGGTGGTACTCCTCGAGGGAACACGACAAGAAACAGAGACGCACCTGCCGGAACCGGTC
MetArgSerSerLeuValLeuPhePheValSerAlaTrpThrAlaLeuAlaSer
----------pre-pro region----------
A
l
w
N
I
TCCTATTCGTCGAGAGGTCTCGCAGGATCTGTTTAACCAGTTCAATCTCTTTGCACAGTA
1201---------+---------+---------+---------+---------+---------+1260
AGGATAAGCAGCTCTCCAGAGCGTCCTAGACAAATTGGTCAAGTTAGAGAAACGTGTCAT
ProIleArgArgGluValSerGlnAspLeuPheAsnGlnPheAsnLeuPheAlaGlnTyr
----------||-----mature lipase----------
P s t I
P v u I I
A p a L I
TTCTGCAGCCGCATACTGCGGAAAAAACAATGATGCCCCAGCTGGTACAAACATTACGTG
1261---------+---------+---------+---------+---------+---------+1320
AAGACGTCGGCGTATGACGCCTTTTTTGTTACTACGGGGTCGACCATGTTTGTAATGCAC
SerAlaAlaAlaTyrCysGlyLysAsnAsnAspAlaProAlaGlyThrAsnIleThrCys-
B
s
p
1H
2g
8i
6A
II
/
A
v
a
I
CACGGGAAATGCCTGCCCCGAGGTAGAGAAGGCGGATGCAACGTTTCTCTACTCGTTTGA
1321---------+---------+---------+---------+---------+---------+1380
GTGCCCTTTACGGACGGGGCTCCATCTCTTCCGCCTACGTTGCAAAGAGATGAGCAAACT
ThrGlyAsnAlaCysProGluValGluLysAlaAspAlaThrPheLeuTyrSerPheGlu-
AGACTCTGGAGTGGGCGATGTCACCGGCTTCCTTGCTCTCGACAACACGAACAAATTGAT
1381---------+---------+---------+---------+---------+---------+1440
TCTGAGACCTCACCCGCTACAGTGGCCGAAGGAACGAGAGCTGTTGTGCTTGTTTAACTA
AspSerGlyValGlyAspValThrGlyPheLeuAlaLeuAspAsnThrAsnLysLeuIle-

Fig. 1

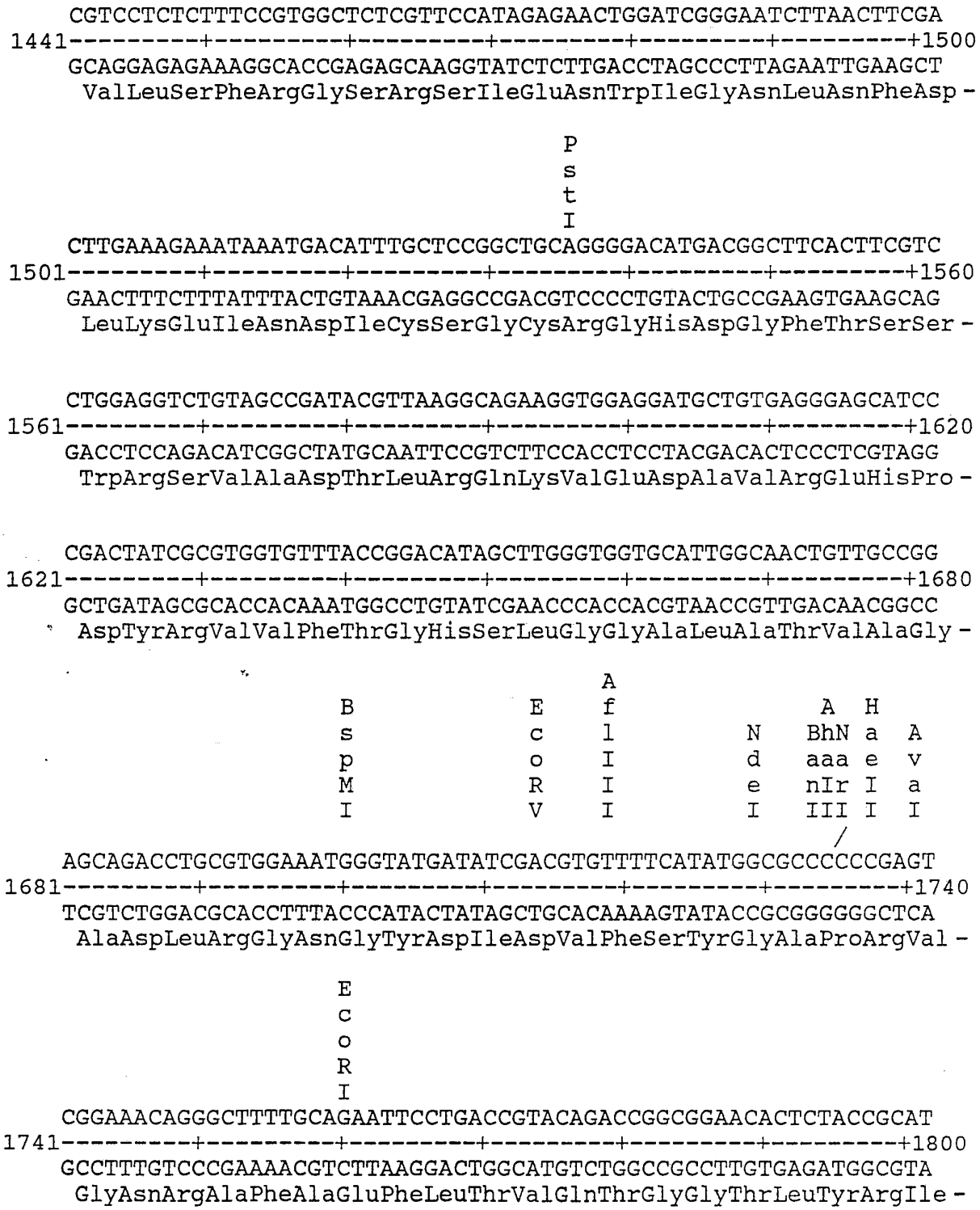

CGTCCTCTCTTTCCGTGGCTCTCGTTCCATAGAGAACTGGATCGGGAATCTTAACTTCGA
1441---------+---------+---------+---------+---------+---------+1500
GCAGGAGAGAAAGGCACCGAGAGCAAGGTATCTCTTGACCTAGCCCTTAGAATTGAAGCT
ValLeuSerPheArgGlySerArgSerIleGluAsnTrpIleGlyAsnLeuAsnPheAsp-
PstI
CTTGAAAGAAATAAATGACATTTGCTCCGGCTGCAGGGGACATGACGGCTTCACTTCGTC
1501---------+---------+---------+---------+---------+---------+1560
GAACTTTCTTTATTTACTGTAAACGAGGCCGACGTCCCCTGTACTGCCGAAGTGAAGCAG
LeuLysGluIleAsnAspIleCysSerGlyCysArgGlyHisAspGlyPheThrSerSer-
CTGGAGGTCTGTAGCCGATACGTTAAGGCAGAAGGTGGAGGATGCTGTGAGGGAGCATCC
1561---------+---------+---------+---------+---------+---------+1620
GACCTCCAGACATCGGCTATGCAATTCCGTCTTCCACCTCCTACGACACTCCCTCGTAGG
TrpArgSerValAlaAspThrLeuArgGlnLysValGluAspAlaValArgGluHisPro-
CGACTATCGCGTGGTGTTTACCGGACATAGCTTGGGTGGTGCATTGGCAACTGTTGCCGG
1621---------+---------+---------+---------+---------+---------+1680
GCTGATAGCGCACCACAAATGGCCTGTATCGAACCCACCACGTAACCGTTGACAACGGCC
AspTyrArgValValPheThrGlyHisSerLeuGlyGlyAlaLeuAlaThrValAlaGly-
BspMI EcoRI AflII NdeI BhnIrIII AhaIII AvaI
AGCAGACCTGCGTGGAAATGGGTATGATATCGACGTGTTTTCATATGGCGCCCCCCGAGT
1681---------+---------+---------+---------+---------+---------+1740
TCGTCTGGACGCACCTTTACCCATACTATAGCTGCACAAAAGTATACCGCGGGGGGCTCA
AlaAspLeuArgGlyAsnGlyTyrAspIleAspValPheSerTyrGlyAlaProArgVal-
EcoRI
CGGAAACAGGGCTTTTGCAGAATTCCTGACCGTACAGACCGGCGGAACACTCTACCGCAT
1741---------+---------+---------+---------+---------+---------+1800
GCCTTTGTCCCGAAAACGTCTTAAGGACTGGCATGTCTGGCCGCCTTGTGAGATGGCGTA
GlyAsnArgAlaPheAlaGluPheLeuThrValGlnThrGlyGlyThrLeuTyrArgIle-

Fig. 1

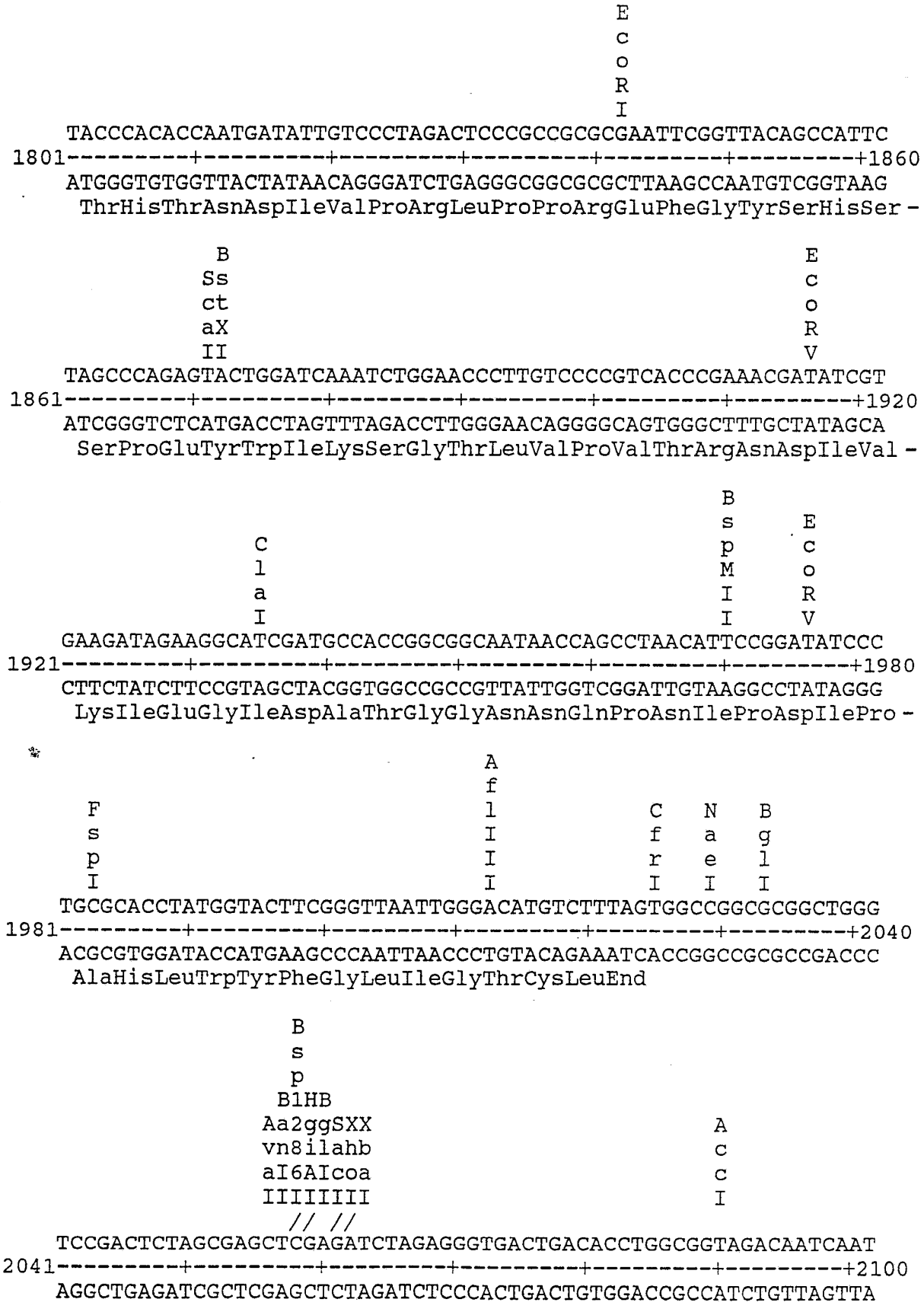

EcoRI
TACCCACACCAATGATATTGTCCCTAGACTCCCGCCGCGCGAATTCGGTTACAGCCATTC
1801---------+---------+---------+---------+---------+---------+1860
ATGGGTGTGGTTACTATAACAGGGATCTGAGGGCGGCGCGCTTAAGCCAATGTCGGTAAG
ThrHisThrAsnAspIleValProArgLeuProProArgGluPheGlyTyrSerHisSer -
BSstaXII
EcoRV
TAGCCCAGAGTACTGGATCAAATCTGGAACCCTTGTCCCCGTCACCCGAAACGATATCGT
1861---------+---------+---------+---------+---------+---------+1920
ATCGGGTCTCATGACCTAGTTTAGACCTTGGGAACAGGGGCAGTGGGCTTTGCTATAGCA
SerProGluTyrTrpIleLysSerGlyThrLeuValProValThrArgAsnAspIleVal -
ClaI
BspMI
EcoRV
GAAGATAGAAGGCATCGATGCCACCGGCGGCAATAACCAGCCTAACATTCCGGATATCCC
1921---------+---------+---------+---------+---------+---------+1980
CTTCTATCTTCCGTAGCTACGGTGGCCGCCGTTATTGGTCGGATTGTAAGGCCTATAGGG
LysIleGluGlyIleAspAlaThrGlyGlyAsnAsnGlnProAsnIleProAspIlePro -
AflIII
FspI
CfrI
NaeI
BglI
TGCGCACCTATGGTACTTCGGGTTAATTGGGACATGTCTTTAGTGGCCGGCGCGGCTGGG
1981---------+---------+---------+---------+---------+---------+2040
ACGCGTGGATACCATGAAGCCCAATTAACCCTGTACAGAAATCACCGGCCGCGCCGACCC
AlaHisLeuTrpTyrPheGlyLeuIleGlyThrCysLeuEnd
BspB1HBAa2ggSXXvn8ilahbaI6AIcoaIIIIIIII
AccI
// //
TCCGACTCTAGCGAGCTCGAGATCTAGAGGGTGACTGACACCTGGCGGTAGACAATCAAT
2041---------+---------+---------+---------+---------+---------+2100
AGGCTGAGATCGCTCGAGCTCTAGATCTCCCACTGACTGTGGACCGCCATCTGTTAGTTA

Fig. 1

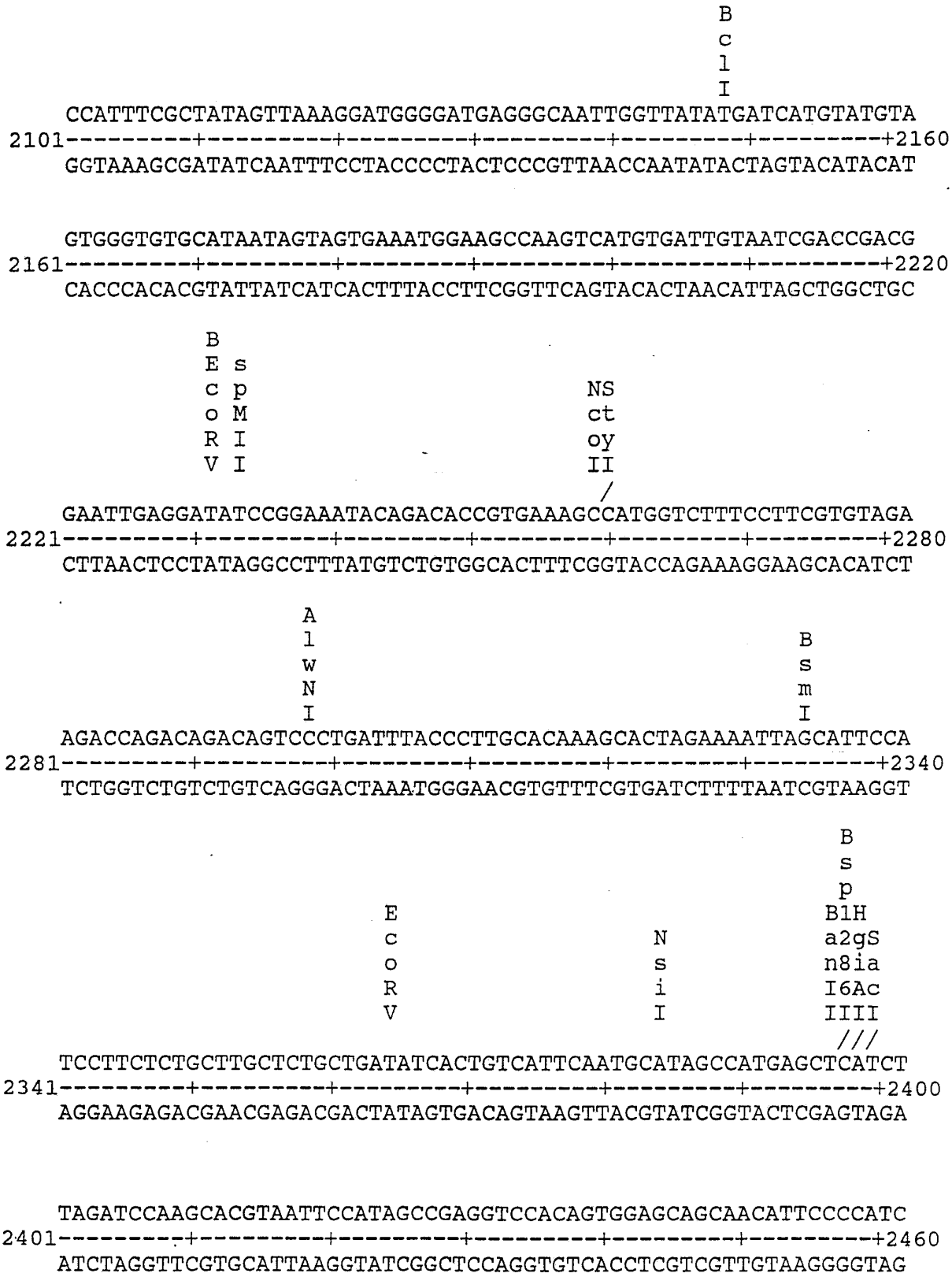

B
c
l
I
CCATTTCGCTATAGTTAAAGGATGGGGATGAGGGCAATTGGTTATATGATCATGTATGTA
2101---------+---------+---------+---------+---------+---------+2160
GGTAAAGCGATATCAATTTCCTACCCCTACTCCCGTTAACCAATATACTAGTACATACAT
GTGGGTGTGCATAATAGTAGTGAAATGGAAGCCAAGTCATGTGATTGTAATCGACCGACG
2161---------+---------+---------+---------+---------+---------+2220
CACCCACACGTATTATCATCACTTTACCTTCGGTTCAGTACACTAACATTAGCTGGCTGC
B
E s
c p
o M
R I
V I
NS
ct
oy
II
GAATTGAGGATATCCGGAAATACAGACACCGTGAAAGCCATGGTCTTTCCTTCGTGTAGA
2221---------+---------+---------+---------+---------+---------+2280
CTTAACTCCTATAGGCCTTTATGTCTGTGGCACTTTCGGTACCAGAAAGGAAGCACATCT
A
l
w
N
I
B
s
m
I
AGACCAGACAGACAGTCCCTGATTTACCCTTGCACAAAGCACTAGAAAATTAGCATTCCA
2281---------+---------+---------+---------+---------+---------+2340
TCTGGTCTGTCTGTCAGGGACTAAATGGGAACGTGTTTCGTGATCTTTTAATCGTAAGGT
B
s
p
E
c
o
R
V
N
s
i
I
B1H
a2gS
n8ia
I6Ac
IIII
TCCTTCTCTGCTTGCTCTGCTGATATCACTGTCATTCAATGCATAGCCATGAGCTCATCT
2341---------+---------+---------+---------+---------+---------+2400
AGGAAGAGACGAACGAGACGACTATAGTGACAGTAAGTTACGTATCGGTACTCGAGTAGA
TAGATCCAAGCACGTAATTCCATAGCCGAGGTCCACAGTGGAGCAGCAACATTCCCCATC
2401---------+---------+---------+---------+---------+---------+2460
ATCTAGGTTCGTGCATTAAGGTATCGGCTCCAGGTGTCACCTCGTCGTTGTAAGGGGTAG

Fig. 1

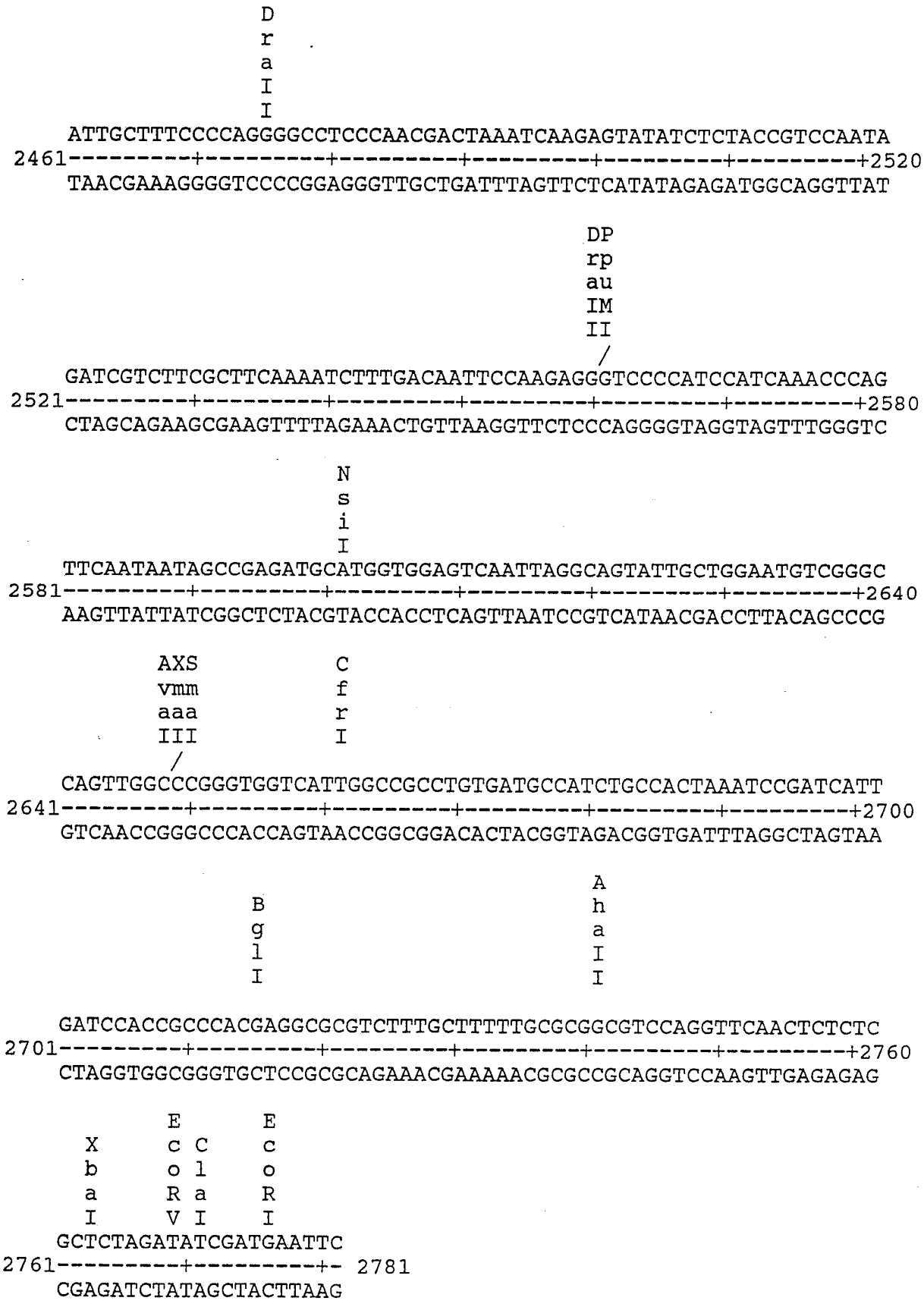

DraII
ATTGCTTTCCCCAGGGGCCTCCCAACGACTAAATCAAGAGTATATCTCTACCGTCCAATA
2461---------+---------+---------+---------+---------+---------+2520
TAACGAAAGGGGTCCCCGGAGGGTTGCTGATTTAGTTCTCATATAGAGATGGCAGGTTAT
DPrpauIMII
GATCGTCTTCGCTTCAAAATCTTTGACAATTCCAAGAGGGTCCCCATCCATCAAACCCAG
2521---------+---------+---------+---------+---------+---------+2580
CTAGCAGAAGCGAAGTTTTAGAAACTGTTAAGGTTCTCCCAGGGGTAGGTAGTTTGGGTC
NsiI
TTCAATAATAGCCGAGATGCATGGTGGAGTCAATTAGGCAGTATTGCTGGAATGTCGGGC
2581---------+---------+---------+---------+---------+---------+2640
AAGTTATTATCGGCTCTACGTACCACCTCAGTTAATCCGTCATAACGACCTTACAGCCCG
AXSvmmaaaIII
CfrI
CAGTTGGCCCGGGTGGTCATTGGCCGCCTGTGATGCCATCTGCCACTAAATCCGATCATT
2641---------+---------+---------+---------+---------+---------+2700
GTCAACCGGGCCCACCAGTAACCGGCGGACACTACGGTAGACGGTGATTTAGGCTAGTAA
BglI
AhaII
GATCCACCGCCCACGAGGCGCGTCTTTGCTTTTTGCGCGGCGTCCAGGTTCAACTCTCTC
2701---------+---------+---------+---------+---------+---------+2760
CTAGGTGGCGGGTGCTCCGCGCAGAAACGAAAAACGCGCCGCAGGTCCAAGTTGAGAGAG
XbaI
EcoRV
EcoRI
GCTCTAGATATCGATGAATTC
2761---------+---------+- 2781
CGAGATCTATAGCTACTTAAG

Fig. 1

Enzymes that do cut:

AccI	AflIII	AhaII	AlwNI	ApaLI	AvaI	BalI	BamHI	
BanI	BanII	BclI	BglI	BglII	BsmI	Bsp1286I	BspMI	
BspMII	BssHII	BstXI	CfrI	ClaI	DraI	DraII	EcoRI	
EcoRV	FspI	HaeII	HgiAI	HincII	NaeI	NarI	NcoI	
NdeI	NsiI	PpuMI	PstI	PvuII	SacI	SalI	ScaI	
SmaI	SspI	StuI	StyI	XbaI	XhoI	XmaI		

Enzymes that do not cut:

AatII	AflII	ApaI	AvrII	BspHI	BstEII	DraIII	EcoNI	
HindIII	HpaI	KpnI	MluI	NheI	NotI	NruI	PflMI	
PvuI	RsrII	SacII	SfiI	SnaBI	SpeI	SphI	Tth111I	

Fig. 1

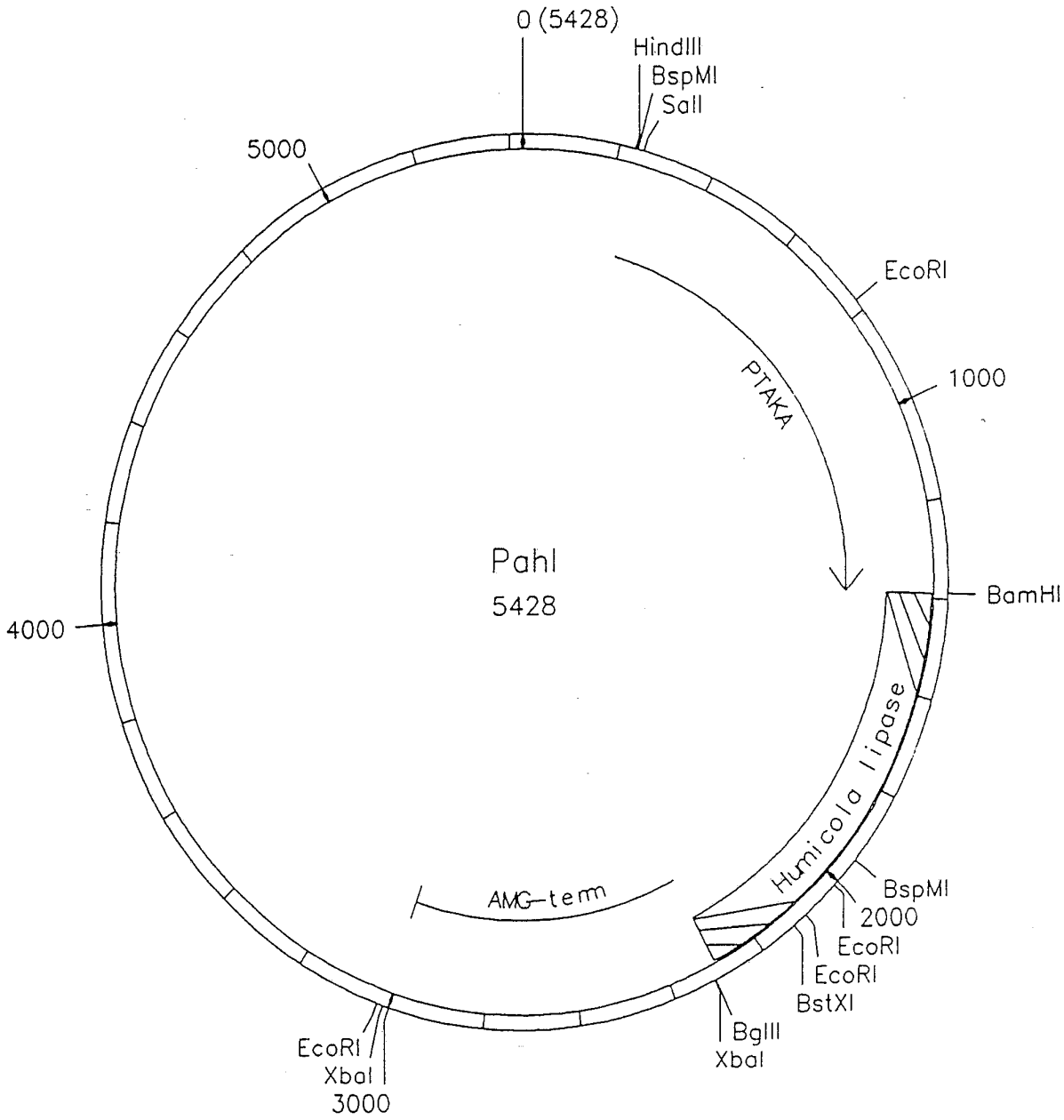

0 (5428)
HindIII
BspMI
SalI
5000
EcoRI
1000
PTAKA
PahI
5428
BamHI
4000
Humicola lipase
AMG-term
BspMI
2000
EcoRI
EcoRI
BstXI
BglII
XbaI
EcoRI
XbaI
3000

Fig. 2

Mutant construction using PCR

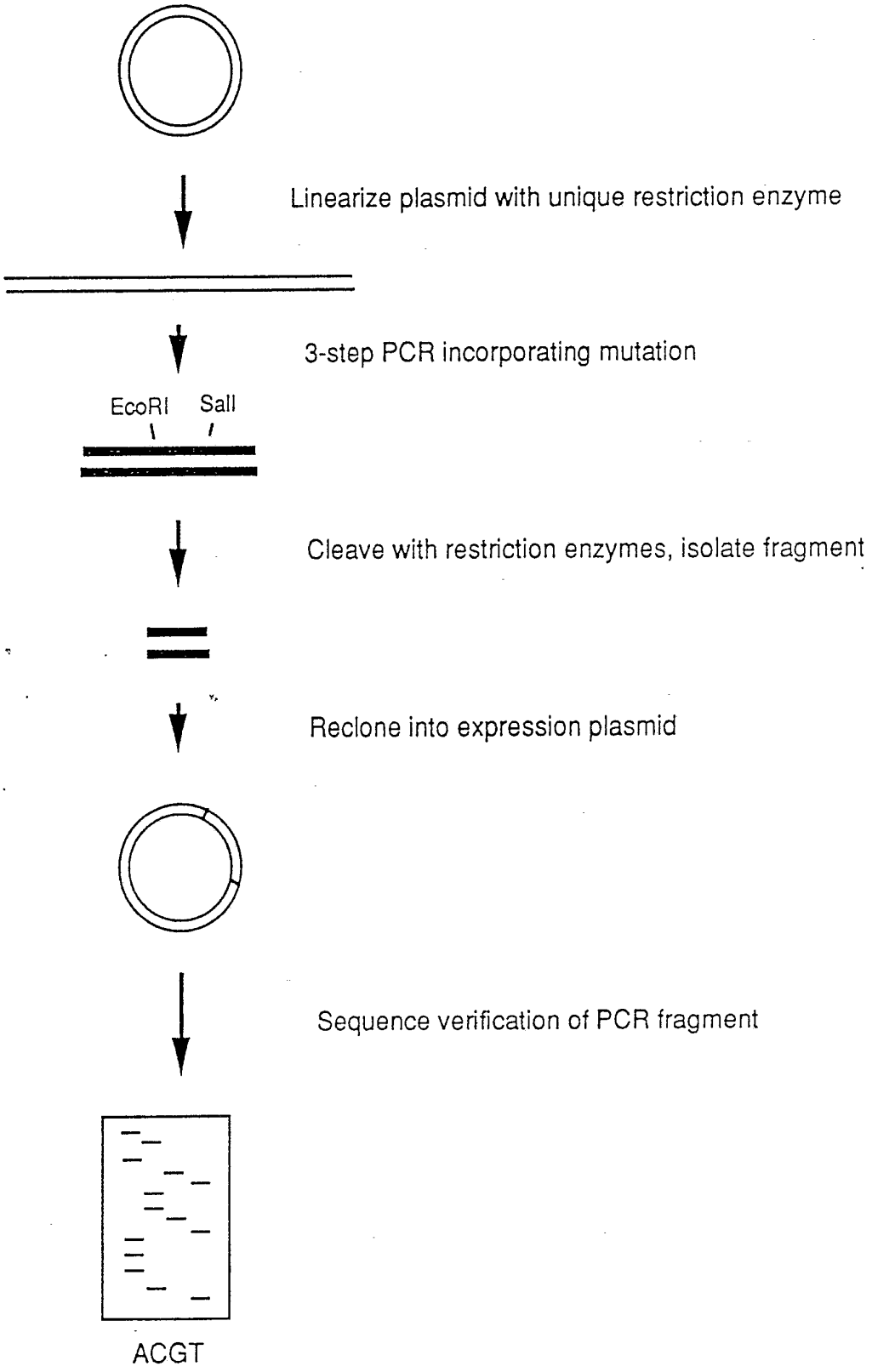


Fig. 3

PCR-mutagenesis

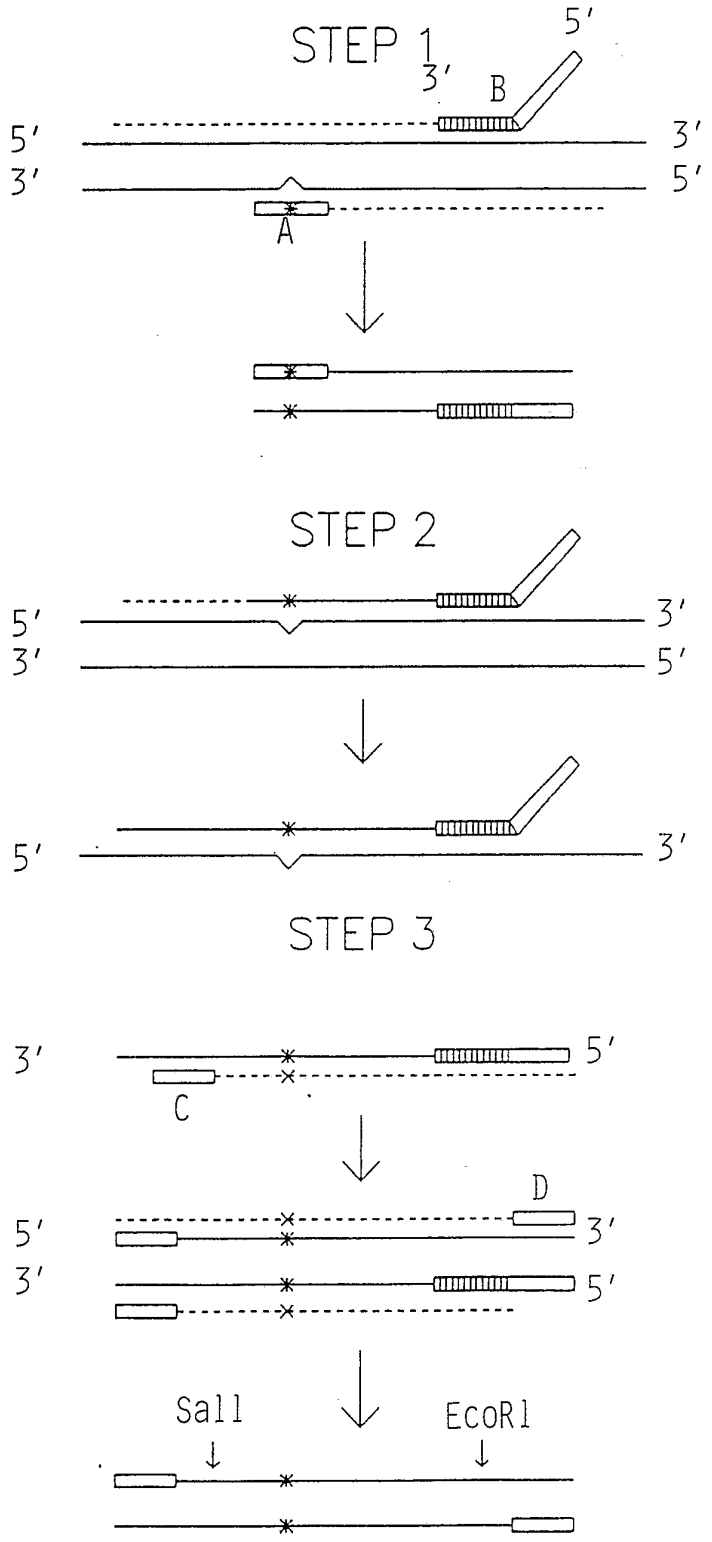


Fig. 4

```
          ATGAGGAGCTCCCTTGTGCTGTTCTTTGTCTCTGCGTGGACGGCCTTGGCCAGT
          MetArgSerSerLeuValLeuPhePheValSerAlaTrpThrAlaLeuAlaSer
             ------------------pre-pro region-------------------

CCTATTCGTCGAGAGGTCTCGCAGGATCTGTTTAACCAGTTCAATCTCTTTGCACAGTAT
ProIleArgArgGluValSerGlnAspLeuPheAsnGlnPheAsnLeuPheAlaGlnTyr
-----------¦¦-----mature lipase-------------

TCTGCAGCCGCATACTGCGGAAAAAACAATGATGCCCCAGCTGGTACAAACATTACGTGC
SerAlaAlaAlaTyrCysGlyLysAsnAsnAspAlaProAlaGlyThrAsnIleThrCys

ACGGGAAATGCCTGCCCCGAGGTAGAGAAGGCGGATGCAACGTTTCTCTACTCGTTTGAA
ThrGlyAsnAlaCysProGluValGluLysAlaAspAlaThrPheLeuTyrSerPheGlu

GACTCTGGAGTGGGCGATGTCACCGGCTTCCTTGCTCTCGACAACACGAACAAATTGATC
AspSerGlyValGlyAspValThrGlyPheLeuAlaLeuAspAsnThrAsnLysLeuIle

GTCCTCTCTTTCCGTGGCTCTCGTTCCATAGAGAACTGGATCGGGAATCTTAACTTCGAC
ValLeuSerPheArgGlySerArgSerIleGluAsnTrpIleGlyAsnLeuAsnPheAsp

TTGAAAGAAATAAATGACATTTGCTCCGGCTGCAGGGGACATGACGGCTTCACTTCGTCC
LeuLysGluIleAsnAspIleCysSerGlyCysArgGlyHisAspGlyPheThrSerSer

TGGAGGTCTGTAGCCGATACGTTAAGGCAGAAGGTGGAGGATGCTGTGAGGGAGCATCCC
TrpArgSerValAlaAspThrLeuArgGlnLysValGluAspAlaValArgGluHisPro

GACTATCGCGTGGTGTTTACCGGACATAGCTTGGGTGGTGCATTGGCAACTGTTGCCGGA
AspTyrArgValValPheThrGlyHisSerLeuGlyGlyAlaLeuAlaThrValAlaGly

GCAGACCTGCGTGGAAATGGGTATGATATCGACGTGTTTTCATATGGCGCCCCCCGAGTC
AlaAspLeuArgGlyAsnGlyTyrAspIleAspValPheSerTyrGlyAlaProArgVal
```

Fig. 5

```
GGAAACAGGGCTTTTGCAGAATTCCTGACCGTACAGACCGGCGGAACACTCTACCGCATT
GlyAsnArgAlaPheAlaGluPheLeuThrValGlnThrGlyGlyThrLeuTyrArgIle

ACCCACACCAATGATATTGTCCCTAGACTCCCGCCGCGCGAATTCGGTTACAGCCATTCT
ThrHisThrAsnAspIleValProArgLeuProProArgGluPheGlyTyrSerHisSer

AGCCCAGAGTACTGGATCAAATCTGGAACCCTTGTCCCCGTCACCCGAAACGATATCGTG
SerProGluTyrTrpIleLysSerGlyThrLeuValProValThrArgAsnAspIleVal

AAGATAGAAGGCATCGATGCCACCGGCGGCAATAACCAGCCTAACATTCCGGATATCCCT
LysIleGluGlyIleAspAlaThrGlyGlyAsnAsnGlnProAsnIleProAspIlePro

GCGCACCTATGGTACTTCGGGTTAATTGGGACATGTCTTTAGTGGCCGGCGCGGCTGGG
AlaHisLeuTrpTyrPheGlyLeuIleGlyThrCysLeuEnd
```

Fig. 5

```
GGATCCAAG ATG CGT TCC TCC CCC CTC CTC CCG TCC GCC GTT GTG GCC     48
          Met Arg Ser Ser Pro Leu Leu Pro Ser Ala Val Val Ala
          -21 -20                 -15                 -10

GCC CTG CCG GTG TTG GCC CTT GCC GCT GAT GGC AGG TCC ACC CGC TAC   96
Ala Leu Pro Val Leu Ala Leu Ala Ala Asp Gly Arg Ser Thr Arg Tyr
            -5                   1               5

TGG GAC TGC TGC AAG CCT TCG TGC GGC TGG GCC AAG AAG GCT CCC GTG  144
Trp Asp Cys Cys Lys Pro Ser Cys Gly Trp Ala Lys Lys Ala Pro Val
     10                  15                  20

AAC CAG CCT GTC TTT TCC TGC AAC GCC AAC TTC CAG CGT ATC ACG GAC  192
Asn Gln Pro Val Phe Ser Cys Asn Ala Asn Phe Gln Arg Ile Thr Asp
 25                  30                  35                  40

TTC GAC GCC AAG TCC GGC TGC GAG CCG GGC GGT GTC GCC TAC TCG TGC  240
Phe Asp Ala Lys Ser Gly Cys Glu Pro Gly Gly Val Ala Tyr Ser Cys
                 45                  50                  55

GCC GAC CAG ACC CCA TGG GCT GTG AAC GAC GAC TTC GCG CTC GGT TTT  288
Ala Asp Gln Thr Pro Trp Ala Val Asn Asp Asp Phe Ala Leu Gly Phe
            60                  65                  70

GCT GCC ACC TCT ATT GCC GGC AGC AAT GAG GCG GGC TGG TGC TGC GCC  336
Ala Ala Thr Ser Ile Ala Gly Ser Asn Glu Ala Gly Trp Cys Cys Ala
        75                  80                  85

TGC TAC GAG CTC ACC TTC ACA TCC GGT CCT GTT GCT GGC AAG AAG ATG  384
Cys Tyr Glu Leu Thr Phe Thr Ser Gly Pro Val Ala Gly Lys Lys Met
     90                  95                  100

GTC GTC CAG TCC ACC AGC ACT GGC GGT GAT CTT GGC AGC AAC CAC TTC  432
Val Val Gln Ser Thr Ser Thr Gly Gly Asp Leu Gly Ser Asn His Phe
105                 110                 115                 120

GAT CTC AAC ATC CCC GGC GGC GGC GTC GGC ATC TTC GAC GGA TGC ACT  480
Asp Leu Asn Ile Pro Gly Gly Gly Val Gly Ile Phe Asp Gly Cys Thr
                125                 130                 135
```

Fig. 6

```
CCC CAG TTC GGC GGT CTG CCC GGC CAG CGC TAC GGC GGC ATC TCG TCC  528
Pro Gln Phe Gly Gly Leu Pro Gly Gln Arg Tyr Gly Gly Ile Ser Ser
            140                 145                 150

CGC AAC GAG TGC GAT CGG TTC CCC GAC GCC CTC AAG CCC GGC TGC TAC  576
Arg Asn Glu Cys Asp Arg Phe Pro Asp Ala Leu Lys Pro Gly Cys Tyr
        155                 160                 165

TGG CGC TTC GAC TGG TTC AAG AAC GCC GAC AAT CCG AGC TTC AGC TTC  624
Trp Arg Phe Asp Trp Phe Lys Asn Ala Asp Asn Pro Ser Phe Ser Phe
    170                 175                 180

CGT CAG GTC CAG TGC CCA GCC GAG CTC GTC GCT CGC ACC GGA TGC CGC  672
Arg Gln Val Gln Cys Pro Ala Glu Leu Val Ala Arg Thr Gly Cys Arg
185                 190                 195                 200

CGC AAC GAC GAC GGC AAC TTC CCT GCC GTC CAG ATC CCC TCC AGC AGC  720
Arg Asn Asp Asp Gly Asn Phe Pro Ala Val Gln Ile Pro Ser Ser Ser
                205                 210                 215

ACC AGC TCT CCG GTC AAC CAG CCT ACC AGC ACC AGC ACC ACG TCC ACC  768
Thr Ser Ser Pro Val Asn Gln Pro Thr Ser Thr Ser Thr Thr Ser Thr
            220                 225                 230

TCC ACC ACC TCG AGC CCG CCA GTC CAG CCT ACG ACT CCC AGC GGC TGC  816
Ser Thr Thr Ser Ser Pro Pro Val Gln Pro Thr Thr Pro Ser Gly Cys
        235                 240                 245

ACT GCT GAG AGG TGG GCT CAG TGC GGC GGC AAT GGC TGG AGC GGC TGC  864
Thr Ala Glu Arg Trp Ala Gln Cys Gly Gly Asn Gly Trp Ser Gly Cys
    250                 255                 260

ACC ACC TGC GTC GCT GGC AGC ACT TGC ACG AAG ATT AAT GAC TGG TAC  912
Thr Thr Cys Val Ala Gly Ser Thr Cys Thr Lys Ile Asn Asp Trp Tyr
265                 270                 275                 280

CAT CAG TGC CTG TAGACGCAGG GCAGCTTGAG GGCCTTACTG GTGGCCGCAA      964
His Gln Cys Leu
                285

CGAAATGACA CTCCCAATCA CTGTATTAGT TCTTGTACAT AATTTCGTCA          1014

TCCCTCCAGG GATTGTCACA TAAATGCAAT GAGGAACAAT GAGTAC              1060
```

Fig. 6

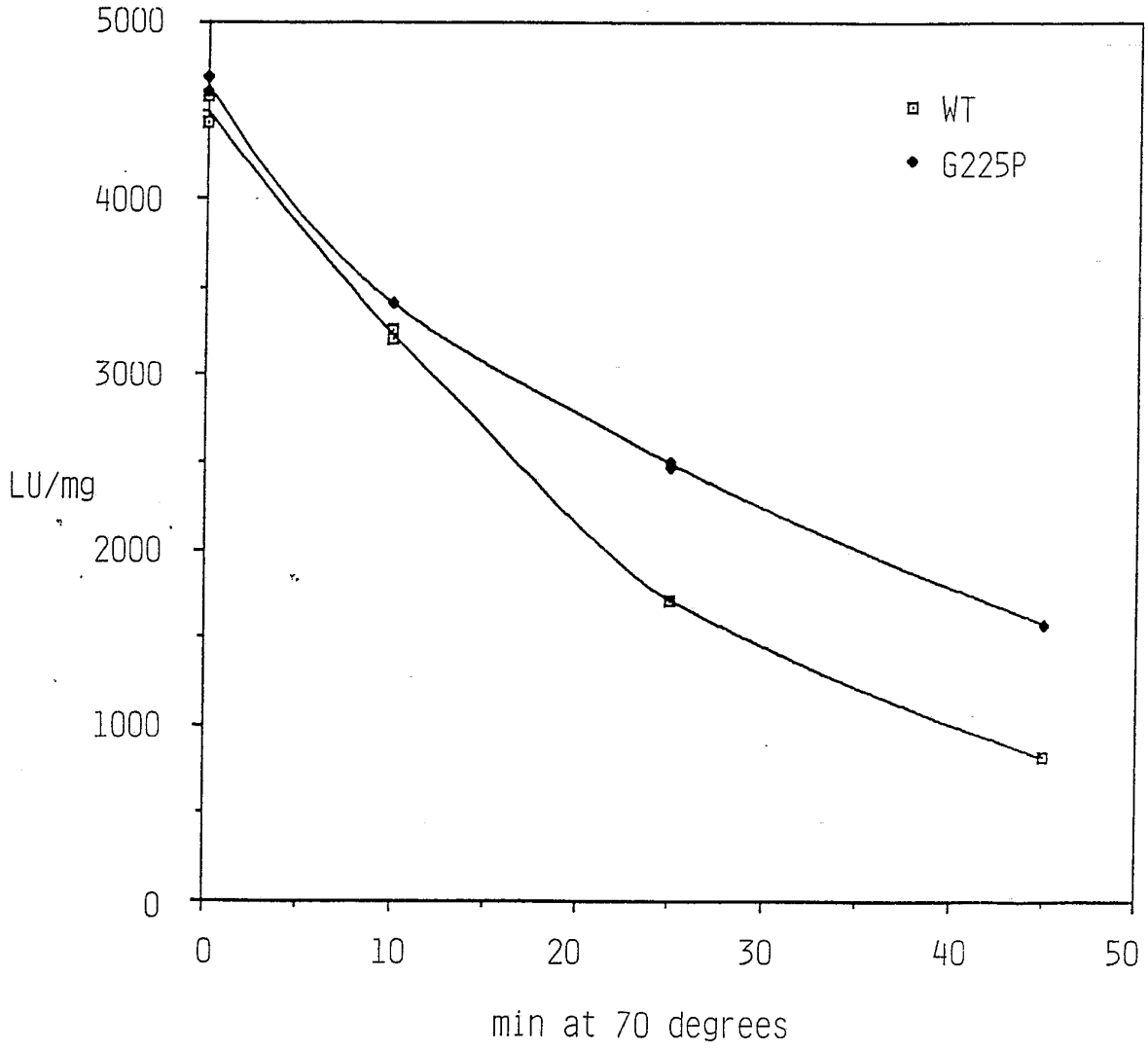

5000
4000
3000
2000
1000
0
LU/mg
WT
G225P
0
10
20
30
40
50
min at 70 degrees

Fig. 7

INTERNATIONAL SEARCH REPORT

International Application No PCT/DK 92/00142

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all)[6]

According to International Patent Classification (IPC) or to both National Classification and IPC

IPC5: C 12 N 9/00, 15/57

II. FIELDS SEARCHED

Minimum Documentation Searched[7]

Classification System	Classification Symbols
IPC5	C 12 N

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in Fields Searched[8]

SE,DK,FI,NO classes as above

III. DOCUMENTS CONSIDERED TO BE RELEVANT[9]

Category *	Citation of Document,[11] with indication, where appropriate, of the relevant passages[12]	Relevant to Claim No.[13]
X	WO, A1, 8901520 (CETUS CORPORATION) 23 February 1989, see especially claim 1 --	1-20

* **Special categories of cited documents:** [10]

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance, the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance, the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search	Date of Mailing of this International Search Report
13th August 1992	1992 -08- 14
International Searching Authority	**Signature of Authorized Officer**
SWEDISH PATENT OFFICE	Mikael G:son Bergstrand

Form PCT/ISA/210 (second sheet) (January 1985)

ANNEX TO THE INTERNATIONAL SEARCH REPORT ON INTERNATIONAL PATENT APPLICATION NO. PCT/DK 92/00142

This annex lists the patent family members relating to the patent documents cited in the above-mentioned international search report. The members are as contained in the Swedish Patent Office EDP file on 01/07/92
The Swedish Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
WO-A1- 8901520	89-02-23	AU-D-	2421588	89-03-09
		US-A-	5041378	91-08-20