



- (72) ANDERSEN, CARSTEN, DK  
(72) NIELSEN, BJARNE RONFELDT, DK  
(72) DIJKHUIZEN, LUBBERT, NL  
(72) DIJKSTRA, BAUKE, NL  
(71) NOVO NORDISK A/S, DK  
(51) Int.Cl.<sup>6</sup> C12N 9/10, C12P 19/18  
(30) 1997/09/24 (1098/97) DK  
(30) 1997/10/08 (60/062,659) US  
(54) **NOUVELLES VARIANTES DE LA CYCLOMALTODEXTRINE  
TRANSFERASE**  
(54) **NOVEL CYCLOMALTODEXTRIN GLUCANOTRANSFERASE  
VARIANTS**

(57) La présente invention porte sur des variantes de la cyclomaltodextrine transférase présentant une plus grande spécificité.

(57) The present invention relates to variants of cyclomaltodextrin glucanotransferase of increased product specificity.

**PCT**WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau

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

<b>(51) International Patent Classification <sup>6</sup> :</b> <b>C12N 9/10, C12P 19/18</b>	<b>A1</b>	<b>(11) International Publication Number:</b> <b>WO 99/15633</b> <b>(43) International Publication Date:</b> 1 April 1999 (01.04.99)
<b>(21) International Application Number:</b> PCT/DK98/00412 <b>(22) International Filing Date:</b> 23 September 1998 (23.09.98) <b>(30) Priority Data:</b> 1098/97                      24 September 1997 (24.09.97)    DK 60/062,659                    8 October 1997 (08.10.97)        US <b>(71) Applicant:</b> NOVO NORDISK A/S [DK/DK]; Novo Allé, DK-2880 Bagsværd (DK). <b>(72) Inventors:</b> ANDERSEN, Carsten; Højeloft Vænge 162, DK-3500 Værløse (DK). NIELSEN, Bjarne, Rønfeldt; Ligustervænget 37, DK-2830 Virum (DK). DIJKHUIZEN, Lubbert; University of Groningen, Kerklaan 30, NL-9751 NN Haren (NL). DIJKSTRA, Bauke; University of Groningen, Kerklaan 30, NL-9751 NN Haren (NL).	<b>(81) Designated States:</b> AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).  <b>Published</b> <i>With international search report.</i>	
<b>(54) Title:</b> NOVEL CYCLOMALTODEXTRIN GLUCANOTRANSFERASE VARIANTS		
<b>(57) Abstract</b>  The present invention relates to variants of cyclomalto-dextrin glucanotransferase of increased product specificity.		

**NOVEL CYCLOMALTODEXTRIN GLUCANOTRANSFERASE VARIANTS****TECHNICAL FIELD**

The present invention relates to variants of cyclomaltodextrin glucanotransferase of increased product specificity.

5

**BACKGROUND ART**

Cyclomaltodextrin glucanotransferase (E.C. 2.4.1.19), also designated cyclodextrin glucanotransferase or cyclodextrin glycosyltransferase, in the following termed CGTase, catalyses the conversion of starch and similar substrates into  
10 cyclomaltodextrins via an intramolecular transglycosylation reaction, thereby forming cyclomaltodextrins, in the following termed cyclodextrins (or CD), of various sizes. Commercially most important are cyclodextrins of 6, 7 and 8 glucose units, which are termed  $\alpha$ -,  $\beta$ - and  $\gamma$ -cyclodextrins, respectively. Commercially  
15 less important are cyclodextrins of 9, 10, and 11 glucose units, which are termed  $\delta$ -,  $\epsilon$ -, and  $\zeta$ -cyclodextrins, respectively.

Cyclodextrins are thus cyclic glucose oligomers with a hydrophobic internal cavity. They are able to form inclusion complexes with many small hydrophobic molecules in aqueous  
20 solutions, resulting in changes in physical properties, e.g. increased solubility and stability and decreased chemical reactivity and volatility. Cyclodextrins find applications particularly in the food, cosmetic, chemical and pharmaceutical industries.

25 Most CGTases have both starch-degrading activity and transglycosylation activity. Although some CGTases produce mainly  $\alpha$ -cyclodextrins and some CGTases produce mainly  $\beta$ -cyclodextrins, CGTases usually form a mixture of  $\alpha$ -,  $\beta$ - and  $\gamma$ -cyclodextrins. Selective precipitation steps with organic solvents may be used  
30 for the isolation of separate  $\alpha$ -,  $\beta$ - and  $\gamma$ -cyclodextrins. To

avoid expensive and environmentally harmful procedures, the availability of CGTases capable of producing an increased ratio of one particular type of cyclodextrin, in particular with respect to  $\alpha$ -,  $\beta$ - or  $\gamma$ -cyclodextrin, is desirable.

5 WO 96/33267 (Novo Nordisk) describes CGTase variants showing a modified substrate binding and/or product selectivity. Although CGTase variants produced by mutation at positions 47, 145, 146, 147, 196 or 371 have been described, the specific CGTase variants of this invention have never been described or even suggested.

10

#### SUMMARY OF THE INVENTION

The present invention provided novel CGTase variants of increased product specificity. Although CGTase variants of increased product specificity have been described in the prior art (WO 15 96/33267), the CGTase variants of the present invention have never been described or suggested.

Among the tremendous number of possible CGTase variants we have now succeeded in finding a limited number of variants showing increased product specificity when compared to the wild-type 20 enzyme.

Accordingly the invention provides a CGTase variant of increased product specificity, in which one or more of the amino acid residues corresponding to the following positions have been introduced by substitution and/or insertion (CGTase Numbering):

(i) Position 47: 47C; 47D; 47E; 47F; 47G; 47I; 47K; 47N; 47P; 47R; 47S; 47T; 47V; 47W; or 47Y;

(ii) Position 145: 145D; 145H; 145I; 145N; 145Q; or 145V;

(iii) Position 146: 146H, 146K; 146L; 146T; 146V; or 146Y;

(iv) Position 147: 147C; 147D; 147E; 147N; 147Q;

(v) Position 196: 196C; 196E; 196F; 196G; 196H; 196I; 196K; 196L; 196M; 196P; 196Q; 196R; 196T; 196V; or 196W; 196Y and/or

(vi) Position 371: 371C; 371E; 371F; 371H; 371I; 371K; 371L; 371M; 371Q; 371R; 371T; 371V; or 371W.

### Amino Acids

In the context of this invention the following symbols and abbreviations for amino acids and amino acid residues are used:

5	A	=	Ala	=	Alanine
	C	=	Cys	=	Cysteine
	D	=	Asp	=	Aspartic acid
	E	=	Glu	=	Glutamic acid
	F	=	Phe	=	Phenylalanine
10	G	=	Gly	=	Glycine
	H	=	His	=	Histidine
	I	=	Ile	=	Isoleucine
	K	=	Lys	=	Lysine
	L	=	Leu	=	Leucine
15	M	=	Met	=	Methionine
	N	=	Asn	=	Asparagine
	P	=	Pro	=	Proline
	Q	=	Gln	=	Glutamine
	R	=	Arg	=	Arginine
20	S	=	Ser	=	Serine
	T	=	Thr	=	Threonine
	V	=	Val	=	Valine
	W	=	Trp	=	Tryptophan
	Y	=	Tyr	=	Tyrosine
25					
	B	=	Asx	=	Asp or Asn
	Z	=	Glx	=	Glu or Gln
	X	=	Xaa	=	Any amino acid
	*	=		=	Deletion or absent amino acid

**CGTase Variants**

A CGTase variant of this invention is a CGTase variant or mutated CGTase, having an amino acid sequence not found in nature.

A CGTase variant or mutated CGTase of this invention may be considered a functional derivative of a precursor CGTase enzyme (i.e. the native, parental, or wild-type enzyme), and may be obtained by alteration of a DNA nucleotide sequence of a precursor gene or its derivatives, encoding the precursor enzyme. The CGTase variant or mutated CGTase may be expressed and produced when the DNA nucleotide sequence encoding the CGTase variant is inserted into a suitable vector in a suitable host organism. The host organism is not necessarily identical to the organism from which the precursor gene originated.

In the literature, enzyme variants have also been referred to as mutants or muteins.

**CGTase Numbering**

In the context of this invention a specific numbering of amino acid residue positions in CGTase enzymes is employed. By alignment of the amino acid sequences of various known CGTases it is possible to unambiguously allot a CGTase amino acid position number to any amino acid residue position in any CGTase enzyme, which amino acid sequence is known.

Using the numbering system originating from the amino acid sequence of the CGTase obtained from *Bacillus circulans* Strain 251, aligned with the amino acid sequence of a number of other known CGTases, it is possible to indicate the position of an amino acid residue in a CGTase enzyme unambiguously.

This CGTase Numbering system has been described in WO 96/33267, see Table 1, pages 9-31 (in which table *Bacillus circulans* Strain 251 is represented as a). Table 1 of WO 96/33267 also shows the

protein sequences of a number of relevant CGTase and is hereby incorporated by reference.

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

[Original amino acid; Position; Substituted amino acid]

Accordingly, the substitution of serine with alanine in position  
10 145 is designated as S145A.

Amino acid residues which represent insertions in relation to the amino acid sequence of the CGTase from *Bacillus circulans* Strain 251, are numbered by the addition of letters in alphabetical  
15 order to the preceding CGTase number, such as e.g. position 91aF for the "insert" Phe between Thr at position 91 and Gly at position 92 of the amino acid sequence of the CGTase from *Thermoanaerobacter* sp. ATCC 53627, cf. Table 1 (j).

Deletion of a proline at position 149 is indicated as P149\*, and  
20 an insertion between position 147 and 148 where no amino acid residue is present, is indicated as \*147aD for insertion of an aspartic acid in position 147a.

Multiple mutations are separated by slash marks ("/"), e.g. S145A/D147L, representing mutations in positions 145 and 147  
25 substituting serine with alanine and aspartic acid with leucine, respectively.

If a substitution is made by mutation in e.g. a CGTase derived from a strain of *Bacillus circulans*, the product is designated e.g. "*B. circulans*/S145A".

30 All positions referred to in this application by CGTase numbering refer to the CGTase numbers described above.

**BRIEF DESCRIPTION OF THE DRAWINGS**

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

Figure 1 shows a cartoon of the substrate binding cleft of *B. circulans* st. 251 CGTase. Dotted lines indicate hydrogen bonds between the enzyme (water) and the substrate. Cleavage point is defined between glucose residue -1 and 1.

Figure 2, 3 and 4 show cyclodextrins formed during incubation of (mutant) CGTase proteins from *Thermoanaerobacterium thermosulfurigenes*. A. wild-type (Figure 2); B. Variant D196H (Figure 3); C. variant D371R (Figure 4).

**DETAILED DISCLOSURE OF THE INVENTION**

The present invention provides novel CGTase variants, *i.e.* CGTase variants having an amino acid sequence not found in nature. Formally, the CGTase variant of the invention may be regarded as a functional derivative of a precursor CGTase enzyme (*i.e.* the native, parental, or wild-type enzyme), by substitution, insertion and/or deletion of one or more amino acid residue(s) of the precursor enzyme.

In the context of this invention, a CGTase variant of increased product specificity is a CGTase variant capable of producing an increased ratio of one particular type of cyclodextrin, when compared to the wild-type enzyme.

In a CGTase variant of the invention, one or more amino acid residues corresponding to the following positions (CGTase Numbering) have been introduced by substitution and/or insertion:

(i) Position 47: 47C; 47D; 47E; 47F; 47G; 47I; 47K; 47N; 47P; 47R; 47S; 47T; 47V; 47W; or 47Y;

(ii) Position 145: 145D; 145H; 145I; 145N; 145Q; or 145V;

(iii) Position 146: 146H; 146K; 146L; 146T; 146V; or 146Y;

WO 99/15633

PCT/DK98/00412

7

(iv) Position 147: 147C; 147D; 147E; 147N; 147Q;

(v) Position 196: 196C; 196E; 196F; 196G; 196H; 196I; 196K; 196L; 196M; 196P; 196Q; 196R; 196T; 196V; 196W; or 196Y and/or

(vi) Position 371: 371C; 371E; 371F; 371H; 371I; 371K; 371L; 371M; 371Q; 371R; 371T; 371V; or 371W.

In a preferred embodiment, CGTase variants showing an increased product specificity with respect to the production of  $\alpha$ -cyclodextrin are provided, in which variants one or more of the amino acid residues corresponding to the following positions have been introduced by substitution and/or insertion (CGTase Numbering):

(i) Position 47: 47F; 47K; 47R; 47W; or 47Y;

(ii) Position 145: 145D; 145H; 145N; or 145Q;

(iii) Position 146: 146H, 146K; 146L; 146T; 146V; or 146Y;

(iv) Position 147: 147C; 147D; 147E; 147N; 147Q;

(v) Position 196: 196C; 196E; 196F; 196G; 196H; 196I; 196K; 196L; 196M; 196P; 196Q; 196R; 196T; 196V; 196W; or 196Y and/or

(vi) Position 371: 371C; 371H; 371K; 371R; or 371T.

In another preferred embodiment, CGTase variants showing an increased product specificity with respect to the production of  $\beta$ -cyclodextrin are provided, in which variants one or more of the amino acid residues corresponding to the following positions have been introduced by substitution and/or insertion (CGTase Numbering):

(i) Position 47: 47C; 47D; 47E; 47F; 47G; 47I; 47N; 47P; 47S; 47T; 47V; 47W; or 47Y;

WO 99/15633

PCT/DK98/00412

8

(ii) Position 145: 145D; 145I; 145N; or 145V;

(iii) Position 147: 147E;

(iv) Position 196: 196C; 196E; 196F; 196G; 196H; 196I; 196K; 196L; 196M; 196P; 196Q; 196R; 196T; 196V; 196W; or 196Y and/or

(v) Position 371: 371C; 371E; 371F; 371H; 371I; 371K; 371L; 371M; 371Q; 371R; 371T; 371V; or 371W.

In yet another preferred embodiment, CGTase variants showing an increased product specificity with respect to the production of  $\gamma$ -cyclodextrin are provided, in which variants one or more of the amino acid residues corresponding to the following positions have been introduced by substitution and/or insertion (CGTase Numbering):

(i) Position 47: 47C; 47D; 47E; 47F; 47G; 47I; 47N; 47P; 47S; 47T; 47V; 47W; or 47Y;

(ii) Position 145: 145D; 145I; 145N; or 145V;

(iii) Position 147: 147E;

(iv) Position 196: 196C; 196E; 196F; 196G; 196H; 196I; 196K; 196L; 196M; 196P; 196Q; 196R; 196T; 196V; 196W; or 196Y and/or

(v) Position 371: 371C; 371E; 371F; 371H; 371K; 371M; 371Q; 371R; 371T; or 371W.

The CGTase variant of the invention may be derived from any CGTase enzyme found in nature. However, the CGTase variant of the invention preferably is derived from a microbial enzyme, preferably a bacterial enzyme, and preferably the CGTase variant  
5 is derived from a strain of *Bacillus*, a strain of *Brevibacterium*, a strain of *Clostridium*, a strain of *Corynebacterium*, a strain of *Klebsiella*, a strain of *Micrococcus*, a strain of *Thermoanaerobium*, a strain of *Thermoanaerobacter*, a strain of

*Thermoanaerobacterium*, a strain of *Thermoanaerobacterium*, or a strain of *Thermoactinomyces*.

In a more preferred embodiment, the CGTase variant of the invention is derived from a strain of *Bacillus autolyticus*, a  
5 strain of *Bacillus cereus*, a strain of *Bacillus circulans*, a strain of *Bacillus circulans* var. *alkalophilus*, a strain of *Bacillus coagulans*, a strain of *Bacillus firmus*, a strain of *Bacillus halophilus*, a strain of *Bacillus macerans*, a strain of *Bacillus megaterium*, a strain of *Bacillus ohbensis*, a strain of  
10 *Bacillus stearothermophilus*, a strain of *Bacillus subtilis*, a strain of *Klebsiella pneumonia*, a strain of *Thermoanaerobacter ethanolicus*, a strain of *Thermoanaerobacter finnii*, a strain of *Clostridium thermoamylolyticum*, a strain of *Clostridium thermosaccharolyticum*, or a strain of *Thermoanaerobacterium*  
15 *thermosulfurigenes*.

In a most preferred embodiment, the CGTase variant of the invention is derived from the strain *Bacillus* sp. Strain 1011, the strain *Bacillus* sp. Strain 38-2, the strain *Bacillus* sp. Strain 17-1, the strain *Bacillus* sp. 1-1, the strain *Bacillus* sp.  
20 Strain B1018, the strain *Bacillus circulans* Strain 8, the strain *Bacillus circulans* Strain 251, or the strain *Thermoanaerobacter* sp. ATCC 53627, or mutants or variants thereof.

If the CGTase variant of the invention is derived from a strain of *Bacillus circulans*, one or more of the amino acid residues  
25 corresponding to the following positions may be introduced:

(i) Position R47: R47C; R47D; R47E; R47F; R47G; R47I; R47K; R47N; R47P; R47S; R47T; R47V; R47W; or R47Y;

(ii) Position S145: S145D; S145H; S145I; S145N; S145Q; or S145V;

(iii) Position S146: S146H, S146K; S146L; S146T; S146V; or S146Y;

(iv) Position D147: D147C; D147E; D147N; D147Q;

(v) Position D196: D196C; D196E; D196F; D196G; D196H; D196I; D196K; D196L; D196M; D196P; D196Q; D196R; D196T; D196V; D196W; or D196Y and/or

(vi) Position D371; D371C; D371E; D371F; D371H; D371I; D371K; D371L; D371M; D371Q; D371R; D371T; D371V; or D371W.

Preferably the CGTase variant is derived from *Bacillus circulans* Strain 251, or a mutant or a variant thereof.

If the CGTase variant is derived from a strain of *Thermoanaerobacter* sp., one or more of the amino acid residues corresponding to the following positions may be introduced:

(i) Position K47; K47C; K47D; K47E; K47F; K47G; K47I; K47N; K47P; K47R; K47S; K47T; K47V; K47W; or K47Y;

(ii) Position S145: S145D; S145H; S145I; S145N; S145Q; or S145V;

(iii) Position E146: E146H, E146K; E146L; E146T; E146V; or E146Y;

(iv) Position T147: T147C; T147D; T147E; T147N; T147Q;

(v) Position D196: D196C; D196E; D196F; D196G; D196H; D196I; D196K; D196L; D196M; D196P; D196Q; D196R; D196T; D196V; D196W; or D196Y and/or

(vi) Position D371: D371C; D371E; D371F; D371H; D371I; D371K; D371L; D371M; D371Q; D371R; D371T; D371V; or D371W.

Preferably the CGTase variant is derived from the strain *Thermoanaerobacter* sp. ATCC 53627, or a mutant or a variant thereof.

Example 1 describes the construction of *T. thermosulfurigenes* CGTase variants Asp196His (D196H) and Asp371Arg (D371R) with modified product specificity, in which site-directed mutagenesis has lead to an altered number of hydrogen bonds in the subsite of  
 5 the active site cleft. The variants are derived from a *Thermoan-*  
*aerobacter thermosulfurigenes* EM1 CGTase (i.e. the wild-type),  
 obtained as described by Haeckel and Bahl [Haeckel, K., and

Bahl, H. (1989) *FEMS Microbiol. Lett.* 60, 333-338 or Knegtel R.M.A., Wind R.D., Rozeboom H.J., Kalk K.H., Buitelaar R.M., Dijkhuizen L., Dijkstra B.W. *J. Mol. Biol.* 256:611-622 (1996)].

In another preferred embodiment, the CGTase variant of the invention comprises one or more of the following amino acid residues (CGTase Numbering):

- (i) 47K/145E/146V/147N;
- 5 (ii) 47K/145E/146E/147N;
- (iii) 47K/145D/146R/147D;
- (iv) 47K/145D/146E/147D;
- (v) 47K/145E/146V/147N/196H;
- (vi) 47K/145E/146E/147N/196H;
- 10 (vii) 47K/145E/146V/147N/196H/371R;
- (viii) 47K/145E/146E/147N/196H/371R;
- (ix) 47K/145D/146R/147D/196H;
- (x) 47K/145D/146E/147D/196H;
- (xi) 47K/145D/146R/147D/196H/371R; and/or
- 15 (xii) 47K/145D/146R/147D/196H/371R.
- (xiii) 47K/196H;
- (xiv) 47R/196H
- (xv) 145E/146V/147N;
- (xvi) 145E/146E/147N;
- 20 (xvii) 145D/146R/147D;
- (xviii) 145D/146E/147D;
- (xix) 47K/371R;
- (xx) 47R/371R;

If the CGTase variant is derived from a strain of *Bacillus circulans* one or more of the following amino acid residues may be introduced:

- 5 (i) R47K/S145E/S146V/D147N;
- (ii) R47K/S145E/S146E/D147N;
- (iii) R47K/S145D/S146R;
- (iv) R47K/S145D/S146E;
- (v) R47K/S145E/S146V/D147N/D196H;
- 10 (vi) R47K/S145E/S146E/D147N/D196H;
- (vii) R47K/S145E/S146V/D147N/D196H/D371R;
- (viii) R47K/S145E/S146E/D147N/D196H/D371R;
- (ix) R47K/S145D/S146R/D196H;
- (x) R47K/S145D/S146E/D196H;
- 15 (xi) R47K/S145D/S146R/D196H/D371R;
- (xii) R47K/S145D/S146R/D196H/D371R.
- (xiii) R47K/D196H;
- (xiv) S145E/S146V/D147N;
- (xv) S145E/S146E/D147N;
- (xvi) S145D/S146R;
- 20 (xvii) S145D/S146E;
- (xviii) R47K/D371R;

Preferably the CGTase variant is derived from *Bacillus circulans* Strain 251, or a mutant or a variant thereof.

If the CGTase variant is derived from a strain of *Thermoanaerobacter sp.*, one or more of the following amino acid residues may be introduced:

- (i) S145E/E146V/T147N;

- (ii) S145E/T147N;
- (iii) S145D/E146R/T147D;
- (iv) S145D/T147D;
- (v) S145E/E146V/T147N/D196H;
- 5 (vi) S145E/T147N/D196H;
- (vii) S145E/E146V/T147N/D196H/D371R;
- (viii) S145E/T147N/D196H/D371R;
- (ix) S145D/E146R/T147D/D196H;
- (x) S145D/T147D/D196H;
- 10 (xi) S145D/E146R/T147D/D196H/D371R;
- (xii) S145D/E146R/T147D/D196H/D371R.
- (xiii) S145E/E146V/T147N;
- (xiv) S145E/T147N;
- (xv) S145D/E146R/T147D;
- (xvi) S145D/T147D; and/or
- 15 (xvii) K47R/D371R;
- (xviii) K47R/D196H

Preferably the CGTase variant is derived from the strain *Thermoanaerobacter* sp. ATCC 53627, or a mutant or a variant thereof.

In a further aspect the invention relates to the use of a CGTase variant of the invention for increasing the  $\alpha$  or  $\beta$  or  $\gamma$ -cyclodextrin content of the final cyclodextrin product of cyclodextrins processes.

In a final aspect the invention relates to a method of increasing the product specificity with respect to the production of  $\alpha$  or  $\beta$  or  $\gamma$ -cyclodextrins, wherein one or more amino acid residues, corresponding to the positions of the CGTase

variants of the invention mentioned above, have been introduced by substitution and/or insertion.

## **EXAMPLES**

### **EXAMPLE 1**

**Construction of *T. thermosulfurigenes* CGTase variants with  
5 modified product specificity.**

#### **Bacterial strains, plasmids and growth conditions**

*Escherichia coli* JM109 [*endA1 recA1 gyrA96 thi hsdR17(rK-,mK+) relA1 supE44 (lac-proAB) [F' traD36 proAB lacIqZ M15]* (Yanish-  
10 Perron et al. 1985 *Gene* 33, 103-119) was used for recombinant DNA manipulations. *Escherichia coli* PC1990 (Lazzaroni and Port-  
talier 1979 *FEMS Microbiol. Lett.* 5, 411-416), known to leak pe-  
riplasmic proteins into the supernatant because of a mutation in  
its *tolB* locus, was used for production of CGTase (mutant) pro-  
15 teins. Plasmid pCT2, a derivative of pUC18 containing the *amyA*  
gene of *Thermoanaerobacterium thermosulfurigenes* EM1 (Knegtel  
R.M.A., Wind R.D., Rozeboom H.J., Kalk K.H., Buitelaar R.M.,  
Dijkhuizen L., Dijkstra B.W. *J. Mol. Biol.* 256:611-622 (1996)),  
was used for site-directed mutagenesis, sequencing and expres-  
20 sion of the CGTase (mutant) proteins. Plasmid-carrying bacterial  
strains were grown on LB medium in the presence of 100 µg/ml am-  
picillin. When appropriate, IPTG (isopropyl-β-D-thiogalacto-  
pyranoside) was added at a concentration of 0.1 mM for induction  
of protein expression.

25

#### **DNA manipulations**

DNA manipulations and transformation of *E. coli* were essentially as described by Sambrook et al. [Sambrook, J., Fritsch, E.J.,

and Maniatis, T. (1989) *Molecular cloning: a laboratory manual*, Cold Spring Harbor Laboratory Press, New York]. Transformation of *E. coli* by electroporation was performed using the Bio-Rad gene pulser apparatus (Bio-Rad, Veenendaal, The Netherlands).  
5 The selected conditions were 2.5 kV, 25  $\mu$ F and 200  $\Omega$ .

#### **Site-directed mutagenesis**

Mutant CGTase genes were constructed via a double PCR method using the *Pfu* DNA polymerase from Stratagene (Westburg, Leusden, The Netherlands). A first PCR reaction was carried out with the mutagenesis primer for the coding strand plus a primer 195-715 bp downstream on the template strand. The reaction product was subsequently used as primer in a second PCR reaction together with a primer 295-815 bp upstream on the coding strand. The  
15 product of the last reaction was cut with *NcoI* and *MunI*, and exchanged with the corresponding fragment (900 bp) from the vector pCT2. The resulting (mutant) plasmid was transformed to *E. coli* JM109 for sequencing and *E. coli* PC1990 for production of the (mutant) proteins. The following oligonucleotides were used:

20

D196H 5'-CGTAACTTATTTCATTTAGCAGATCTAAATCAACAG-3'

(SEQ ID No. 1)

D371R 5'-GACAGGCAATGGACGTCCTTATAATAGAGC-3'

25 (SEQ ID NO. 2).

Successful mutations resulted in the underlined restriction sites (*BglIII* for D196H and *AatII* for D371R), which allowed quick screening of transformants. Mutations were verified by DNA sequencing (Sanger et al. 1977 *Proc. Natl. Sci. USA* 74, 5463-5467). All 900 bp on the *MunI-NcoI* fragment obtained by PCR were  
30 checked by DNA-sequencing.

**Production and purification of CGTase proteins**

For production of CGTase proteins, *E. coli* PC1990 was grown in a 2-liter fermentor at pH 7.0 and 30 C. The medium contained 2% (w/w) tryptone (Oxoid, Boom BV, Meppel, The Netherlands), 1% (w/w) yeast extract (Oxoid), 1% (w/w) sodium chloride, 1% (w/w) casein hydrolysate (Merck, Darmstadt, Germany), 100 µg/l ampicillin and 0.1 mM IPTG. Growth was monitored by measuring the optical density at 450 nm. When the optical density at 450 nm reached 2 to 3, an extra amount of 50 g tryptone was added to the fermentor. Cells were harvested after 20-24 hours of growth (8,000 g, 30 minutes, 4°C), at an optical density of 8-12, and the supernatant was used for further purification of the CGTases. Supernatant was directly applied to an α-CD-sepharose-6FF affinity column (Monma et al. 1988 *Biotechnol. Bioeng.* 32, 404-407). After washing the column with 10 mM sodium acetate (pH 5.5), the CGTase was eluted with the same buffer supplemented with 1% (w/w) α-CD. Purity and molecular weight of the CGTase (mutant) proteins were checked on SDS-PAGE (Wind et al. 1995 *Appl. Environ. Microb.* 61, 1257-1265). Protein concentrations were determined by the method of Bradford, using the Coomassie protein assay reagent of Pierce (Pierce Europe bv, Oud-Beijerland, The Netherlands).

**25 Enzyme assays**

All assays were standardly performed at pH 6.0 and 50°C. Cyclization and saccharifying assays were performed as described by Penninga et al. (Penninga, D., Strokopytov, B., Rozeboom, H.J., Lawson, C.L., Dijkstra, B.W., Bergsma, J., and Dijkhuizen, L. (1995) *Biochemistry* 34, 3368-3376). Units for the different reactions were defined as the amount of enzyme producing 1 µmol of substrate at pH 6.0 and 50°C.

**HPLC product analysis**

Formation of cyclodextrins was measured under industrial process conditions by incubation of 0.1 U/ml CGTase ( $\beta$ -CD forming activity) with 10% Paselli WA4 (pregelatinized drum-dried starch with a high degree of polymerization; AVEBE, The Netherlands) in 10 mM sodium citrate buffer (pH 6.0) at 60°C for 45 hours. Samples were taken at regular time intervals and boiled for 10 minutes. Products formed were analyzed by HPLC, using a 25-cm Econosil-10 NH2 10- $\mu$ m column (Alltech Nederland bv, Breda, The Netherlands) eluted with acetonitrile-water (65:45) at 1 ml/minute. Products were detected by a refractive index detector (Waters 410, Waters Chromatography Division, Milford, USA). The temperature of the flow cell and column was set at 50°C, to avoid possible precipitation of starch. Formation of linear products was directly analyzed. Formation of CD's was analyzed after incubation of the samples with an appropriate amount of  $\beta$ -amylase (type I-B from Sweet potato, Sigma, Boom BV, Meppel, The Netherlands).

**20 Cyclodextrin product specificity**

In order to change the product specificity of *Thermoanaerobacterium* CGTase, we replaced Asp196 by His (D196H) and Asp371 by Arg (D371R). In mutant D196H production of  $\alpha$ -CD was increased at the expense of the production of  $\beta$ -CD when compared to the wild-type 25 CGTase (Figure 2). The cyclodextrin product ratio was changed from 28:58:14 ( $\alpha$ : $\beta$ : $\gamma$ ) for the wild-type CGTase to 35:49:16 for mutant D196H (Table 1). This is according to expectations on the basis of the structural work. By the replacement of Asp197 by His a larger residue is introduced, this would block the so-called straight substrate binding mode. Furthermore, His 197 is 30 probably also able to form hydrogen bonds with the substrate bound in the 'bent' mode. These results lean more vigour to the

theory that the bent conformation is correlated with  $\alpha$ -cyclodextrin production, and show that it can be used to rationally engineer a CGTase with desired product specificity.

5 10	Cyclization activity	Product (%) (U/ml)	Ratio			Conversion of Starch into cyclodextrins(%)
			$\alpha$	$\beta$	$\gamma$	
	Wild-					
	type	163	28	58	14	35
	D196H	144	35	49	16	37
	D371R	14	6	68	26	29

**Table 1:** Starch conversion of *T. thermosulfurigenes* wild-type and mutant CGTase proteins. Proteins (0.1 U/ml  $\beta$ -CD forming activity) were incubated for 45H at pH 6.0 and 60°C with 10% Paselli WA4.

Asp 371 has a very important role in substrate binding at subsite 2, both in the *BC251* and *Tabium* CGTase. Replacement of Asp371 by Arg resulted in the introduction of a very bulky amino acid, which would probably interfere severely with binding of all kinds of substrates at subsite 2, thereby explaining the overall decrease of activity. Apart from this low activity, the product ratio changed from 28:58:14 for the wild-type enzyme to 6:68:26 for mutant D371H (Table 1, Figure 2). This suggests that the cyclisation reaction leading to a  $\alpha$ -cyclodextrin is more hampered by the Arg 371 than the other cyclisation reactions. Probably this bulky residue is sterically hindering the 'bent' conformation. These results show that the *Tabium* CGTase can be changed from a  $\alpha/\beta$ -cyclodextrin producer to a  $\beta/\gamma$ -cyclodextrin producer by just one mutation, illustrating the feasibility of CGTase protein engineering.

WO 99/15633

PCT/DK98/00412

19

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

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

- 5 (i) APPLICANT:  
 (A) NAME: Novo Nordisk A/S  
 (B) STREET: Novo Alle  
 (C) CITY: Bagsvaerd  
 (E) COUNTRY: Denmark  
 10 (F) POSTAL CODE (ZIP): DK 2880  
 (G) TELEPHONE: +45 4444 8888  
 (H) TELEFAX: +45 4449 3256  
 (ii) TITLE OF INVENTION: Novel CGTase variants  
 (iii) NUMBER OF SEQUENCES: 2  
 15 (iv) COMPUTER READABLE FORM:  
 (A) MEDIUM TYPE: Floppy disk  
 (B) COMPUTER: IBM PC compatible  
 (C) OPERATING SYSTEM: PC-DOS/MS-DOS  
 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
 20
- 2) INFORMATION FOR SEQ ID NO: 1:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 36 base pairs  
 (B) TYPE: nucleic acid  
 25 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: other nucleic acid  
 (ix) FEATURE:  
 (A) NAME/KEY: misc-feature  
 30 (B) OTHER INFORMATION: /desc: "D196H oligo"  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

CGTAACTTAT TTCATTTAGC AGATCTAAAT CAACAG

36

- 35 2) INFORMATION FOR SEQ ID NO: 2:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 30 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 40 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: other nucleic acid  
 (ix) FEATURE:  
 (A) NAME/KEY: misc-feature  
 (B) OTHER INFORMATION: /desc: "D371R oligo"  
 45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

GACAGGCAAT GGACGTCCTT ATAATAGAGC

30

**CLAIMS**

1. A CGTase variant of increased product specificity, in which one or more of the amino acid residues corresponding to the following positions have been introduced by substitution and/or insertion (CGTase Numbering):

(i) Position 47: 47C; 47D; 47E; 47F; 47G; 47I; 47K; 47N; 47P; 47R; 47S; 47T; 47V; 47W; or 47Y;

(ii) Position 145: 145D; 145H; 145I; 145N; 145Q; or 145V;

(iii) Position 146: 146H, 146K; 146L; 146T; 146V; or 146Y;

(iv) Position 147: 147C; 147D; 147E; 147N; 147Q;

(v) Position 196: 196C; 196E; 196F; 196G; 196H; 196I; 196K; 196L; 196M; 196P; 196Q; 196R; 196T; 196V; 196W; or 196Y and/or

(vi) Position 371: 371C; 371E; 371F; 371H; 371I; 371K; 371L; 371M; 371Q; 371R; 371T; 371V; or 371W.

2. The CGTase variant of claim 1, being K47R.

3. The CGTase variant of claim 1, being D196H.

4. The CGTase variant of claim 1, being D371R.

5. The CGTase variant according to claim 1, showing an increased product specificity with respect to the production of  $\alpha$ -cyclodextrin, in which one or more of the amino acid residues corresponding to the following positions have been introduced by substitution and/or insertion (CGTase Numbering):

(i) Position 47: 47F; 47K; 47R; 47W; or 47Y;

(ii) Position 145: 145D; 145H; 145N; or 145Q;

(iii) Position 146: 146H, 146K; 146L; 146T; 146V; or 146Y;

(iv) Position 147: 147C; 147D; 147E; 147N; 147Q;

(v) Position 196: 196C; 196E; 196F; 196G; 196H; 196I; 196K; 196L; 196M; 196P; 196Q; 196R; 196T; 196V; 196W; 196Y and/or

(vi) Position 371: 371C; 371H; 371K; 371R; or 371T.

6. The CGTase variant according to claim 1, showing an increased product specificity with respect to the production of  $\beta$ -cyclodextrin, in which one or more of the amino acid residues corresponding to the following positions have been introduced by substitution and/or insertion (CGTase Numbering):

(i) Position 47: 47C; 47D; 47E; 47F; 47G; 47I; 47N; 47P; 47S; 47T; 47V; 47W; or 47Y;

(ii) Position 145: 145D; 145I; 145N; or 145V;

(iii) Position 147: 147E;

(iv) Position 196: 196C; 196E; 196F; 196G; 196H; 196I; 196K; 196L; 196M; 196P; 196Q; 196R; 196T; 196V; 196W; or 196Y and/or

(v) Position 371: 371C; 371E; 371F; 371H; 371I; 371K; 371L; 371M; 371Q; 371R; 371T; 371V; or 371W.

7. The CGTase variant according to claim 1, showing an increased product specificity with respect to the production of  $\gamma$ -cyclodextrin, in which one or more of the amino acid residues corresponding to the following positions have been introduced by substitution and/or insertion (CGTase Numbering):

(i) Position 47: 47C; 47D; 47E; 47F; 47G; 47I; 47N; 47P; 47S; 47T; 47V; 47W; or 47Y;

(ii) Position 145: 145D; 145I; 145N; or 145V;

(iii) Position 147: 147E;

(iv) Position 196: 196C; 196E; 196F; 196G; 196H; 196I; 196K; 196L; 196M; 196P; 196Q; 196R; 196T; 196V; 196W; or 196Y and/or

(v) Position 371: 371C; 371E; 371F; 371H; 371K; 371M; 371Q; 371R; 371T; or 371W.

8. The CGTase variant according to any of claims 1-7, which is derived from a strain of *Bacillus*, a strain of *Brevibacterium*, a strain of *Clostridium*, a strain of *Corynebacterium*, a strain of *Klebsiella*, a strain of *Micrococcus*, a strain of *Thermoanaerobium*, a strain of *Thermoanaerobacterium*, a strain of *Thermoanaerobacter*, or a strain of *Thermoactinomyces*.

9. The CGTase variant according to claim 8, which is derived from a strain of *Bacillus autolyticus*, a strain of *Bacillus cereus*, a strain of *Bacillus circulans*, a strain of *Bacillus circulans* var. *alkalophilus*, a strain of *Bacillus coagulans*, a strain of *Bacillus firmus*, a strain of *Bacillus halophilus*, a strain of *Bacillus macerans*, a strain of *Bacillus megaterium*, a strain of *Bacillus ohbensis*, a strain of *Bacillus stearothermophilus*, or a strain of *Bacillus subtilis*.

10. The CGTase variant according to claim 9, which is derived from the strain *Bacillus* sp. Strain 1011, the strain *Bacillus* sp. Strain 38-2, the strain *Bacillus* sp. Strain 17-1, the strain *Bacillus* sp. 1-1, the strain *Bacillus* sp. Strain B1018, the strain *Bacillus circulans* Strain 8, or the strain *Bacillus circulans* Strain 251, or a mutant or a variant thereof.

11. The CGTase variant according to claim 10, which is derived from the strain *Bacillus circulans* Strain 251, or a mutant or a variant thereof.

12. The CGTase variant according to claim 8, which is derived from a strain of *Klebsiella pneumonia*, a strain of *Thermoanaerobacter ethanolicus*, a strain of *Thermoanaerobacter finnii*, a strain of *Clostridium thermoamylolyticum*, a strain of

*Clostridium thermosaccharolyticum*, or a strain of *Thermoanaerobacterium thermosulfurigenes*.

13. The CGTase variant according to claim 8, which is derived from the strain *Thermoanaerobacter sp.* ATCC 53627, or a mutant or a variant thereof.

14. The CGTase variant according to any of claims 1-7, which is derived from a strain of *Bacillus circulans*, and in which variant one or more of the amino acid residues corresponding to the following positions have been introduced by substitution and/or insertion (CGTase Numbering):

(i) Position R47: R47C; R47D; R47E; R47F; R47G; R47I; R47K; R47N; R47P; R47S; R47T; R47V; R47W; or R47Y;

(ii) Position S145: S145D; S145H; S145I; S145N; S145Q; or S145V;

(iii) Position S146: S146H, S146K; S146L; S146T; S146V; or S146Y;

(iv) Position D147: D147C; D147E; D147N; D147Q;

(v) Position D196: D196C; D196E; D196F; D196G; D196H; D196I; D196K; D196L; D196M; D196P; D196Q; D196R; D196T; D196V; D196W; or D196Y and/or

(vi) Position D371: D371C; D371E; D371F; D371H; D371I; D371K; D371L; D371M; D371Q; D371R; D371T; D371V; or D371W.

15. The CGTase variant according to claim 11 which is derived from *Bacillus circulans* Strain 251, or a mutant or a variant thereof.

16. The CGTase variant according to any of claims 1-7, which is derived from a strain of *Thermoanaerobacter sp.*, and in which variant one or more of the amino acid residues corresponding to

WO 99/15633

PCT/DK98/00412

24

the following positions have been introduced by substitution and/or insertion (CGTase Numbering):

- (i) Position K47: K47C; K47D; K47E; K47F; K47G; K47I; K47N; K47P; K47R; K47S; K47T; K47V; K47W; or K47Y;
- (ii) Position S145: S145D; S145H; S145I; S145N; S145Q; or S145V;
- (iii) Position E146: E146H, E146K; E146L; E146T; E146V; or E146Y;
- (iv) Position T147: T147C; T147D; T147E; T147N; T147Q;
- (v) Position D196: D196C; D196E; D196F; D196G; D196H; D196I; D196K; D196L; D196M; D196P; D196Q; D196R; D196T; D196V; D196W; D196Y and/or
- (vi) Position D371: D371C; D371E; D371F; D371H; D371I; D371K; D371L; D371M; D371Q; D371R; D371T; D371V; or D371W.

17. The CGTase variant according to claim 16, which is derived from the strain *Thermoanaerobacter sp.* ATCC 53627, or a mutant or a variant thereof.

18. The CGTase variant according to any of claims 1-13, which comprises one or more of the following amino acid residues (CGTase Numbering):

- (i) 47K/145E/146V/147N;
- (ii) 47K/145E/146E/147N;
- (iii) 47K/145D/146R/147D;
- (iv) 47K/145D/146E/147D;
- 5 (v) 47K/145E/146V/147N/196H;
- (vi) 47K/145E/146E/147N/196H;
- (vii) 47K/145E/146V/147N/196H/371R;
- (viii) 47K/145E/146E/147N/196H/371R;
- (ix) 47K/145D/146R/147D/196H;
- 10 (x) 47K/145D/146E/147D/196H;

WO 99/15633

PCT/DK98/00412

25

- (xi) 47K/145D/146R/147D/196H/371R;
- (xii) 47K/145D/146R/147D/196H/371R.
- (xiii) 47K/196H;
- (xiv) 47R/196H;
- (xv) 145E/146V/147N;
- 5 (xvi) 145E/146E/147N;
- (xvii) 145D/146R/147D;
- (xviii) 145D/146E/147D; and/or
- (xix) 47K/371R;
- (xx) 47R/371R;

10

19. The CGTase variant according to claim 18, which is derived from a strain of *Bacillus circulans*, and which comprises one or more of the following amino acid residues (CGTase Numbering):

- (i) R47K/S145E/S146V/D147N;
- 15 (ii) R47K/S145E/S146E/D147N;
- (iii) R47K/S145D/S146R;
- (iv) R47K/S145D/S146E;
- (v) R47K/S145E/S146V/D147N/D196H;
- (vi) R47K/S145E/S146E/D147N/D196H;
- 20 (vii) R47K/S145E/S146V/D147N/D196H/D371R;
- (viii) R47K/S145E/S146E/D147N/D196H/D371R;
- (ix) R47K/S145D/S146R/D196H;
- (x) R47K/S145D/S146E/D196H;
- (xi) R47K/S145D/S146R/D196H/D371R;
- 25 (xii) R47K/S145D/S146R/D196H/D371R.
- (xiii) R47K/D196H;
- (xiv) S145E/S146V/D147N;
- (xv) S145E/S146E/D147N;

(xvi) S145D/S146R;

(xvii) S145D/S146E;

(xviii) R47K/D371R;

20. The CGTase variant according to claim 19, which is derived from *Bacillus circulans* Strain 251, or a mutant or a variant thereof.

21. The CGTase variant according to claim 18, which is derived from a strain of *Thermoanaerobacter* sp., and which comprises one or more of the following amino acid residues (CGTase Numbering):

- 5 (i) S145E/E146V/T147N;
- (ii) S145E/T147N;
- (iii) S145D/E146R/T147D;
- (iv) S145D/T147D;
- (v) S145E/E146V/T147N/D196H;
- 10 (vi) S145E/T147N/D196H;
- (vii) S145E/E146V/T147N/D196H/D371R;
- (viii) S145E/T147N/D196H/D371R;
- (ix) S145D/E146R/T147D/D196H;
- (x) S145D/T147D/D196H;
- 15 (xi) S145D/E146R/T147D/D196H/D371R;
- (xii) S145D/E146R/T147D/D196H/D371R.
- (xiii) S145E/E146V/T147N;
- (xiv) S145E/T147N;
- (xv) S145D/E146R/T147D;
- (xvi) S145D/T147D; and/or
- 20 (xvii) K47R/D371R;
- (xviii) K47R/D196H

22. The CGTase variant according to claim 21, which is derived from the strain *Thermoanaerobacter* sp. ATCC 53627, or a mutant or a variant thereof.

23. Use of a CGTase variant according to any of claims 1-22 for increasing the  $\alpha$ -cyclodextrin content of the final cyclodextrin product of cyclodextrins processes.

24. Use of a CGTase variant according to any of claims 1-22 for increasing the  $\beta$ -cyclodextrin content of the final cyclodextrin product of cyclodextrins processes.

25. Use of a CGTase variant according to any of claims 1-22 for increasing the  $\gamma$ -cyclodextrin content of the final cyclodextrin product of cyclodextrins processes.

26. A method of increasing the product specificity with respect to the production of  $\alpha$ -cyclodextrins, wherein one or more amino acid residues corresponding to the positions of the variants of any of claims 1-22 have been introduced by substitution and/or insertion.

27. A method of increasing the product specificity with respect to the production of  $\beta$ -cyclodextrins, wherein one or more of the amino acid residues corresponding to the positions of the variant of any of claims 1-22 have been introduced by substitution and/or insertion.

28. A method of increasing the product specificity with respect to the production of  $\gamma$ -cyclodextrins, wherein one or more of the amino acid residues corresponding to the positions of the variant

WO 99/15633

PCT/DK98/00412

28

of any of claims 1-22 have been introduced by substitution and/or insertion.

1/4

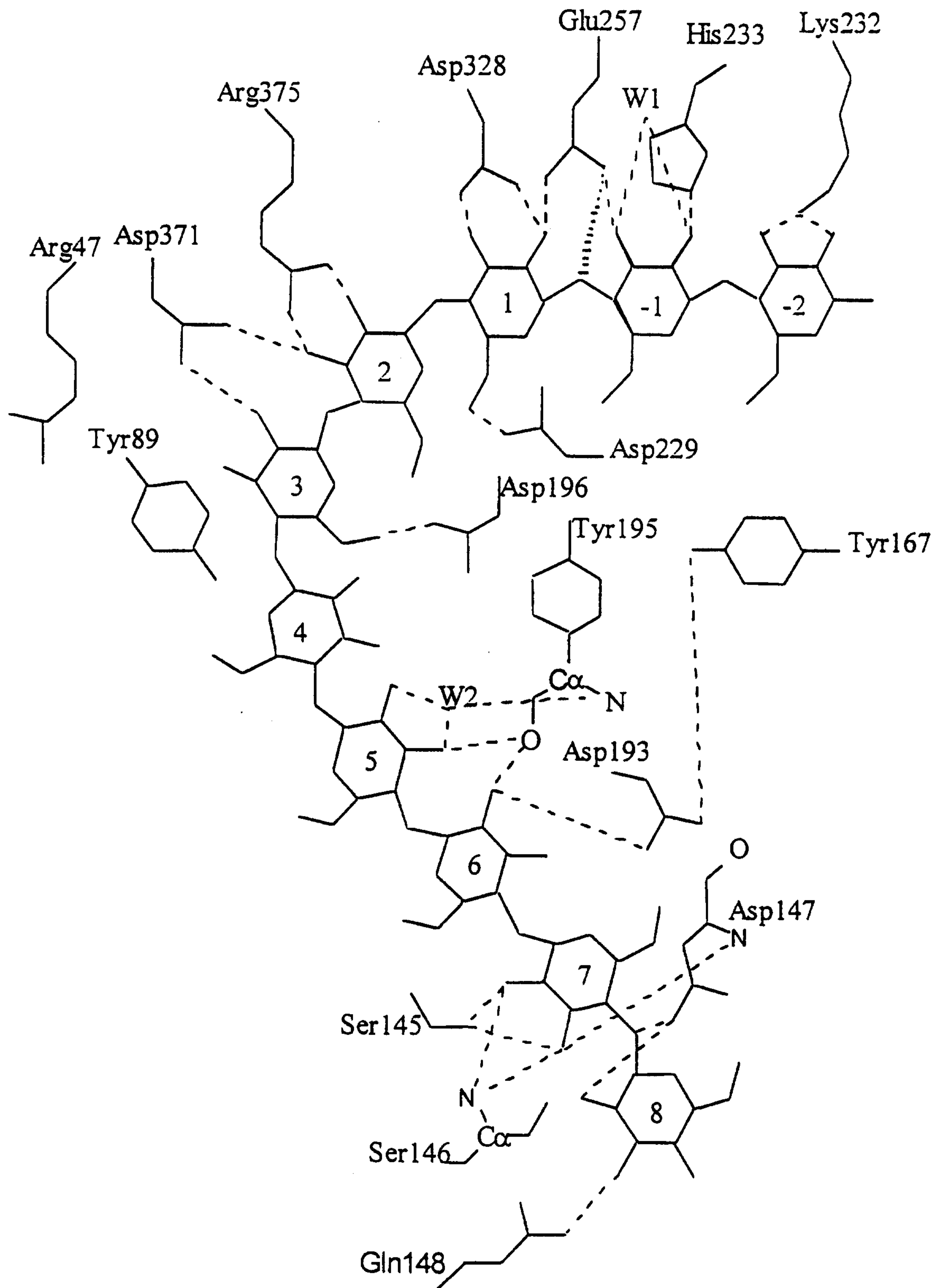


Fig. 1

2/4

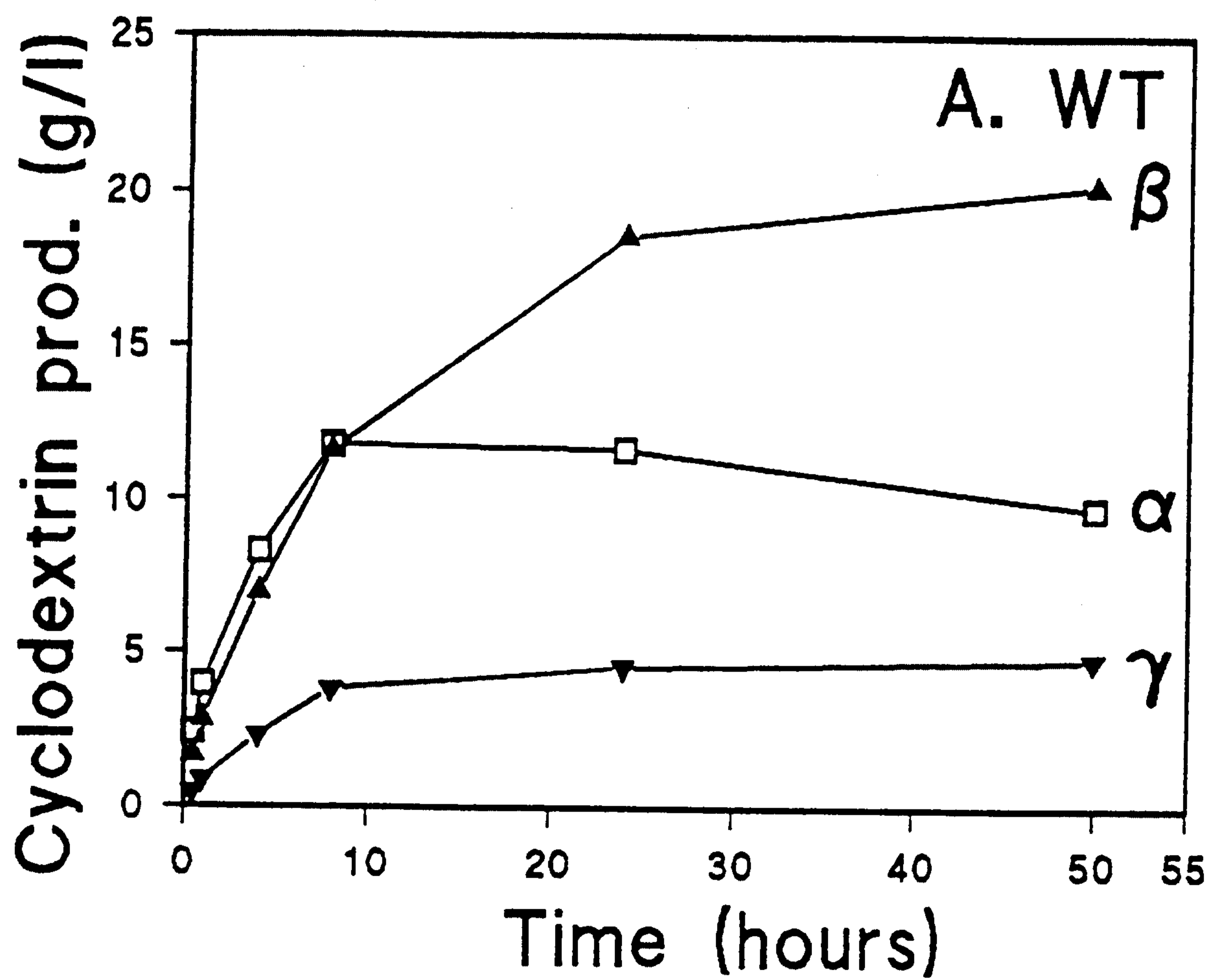


Fig. 2

3/4

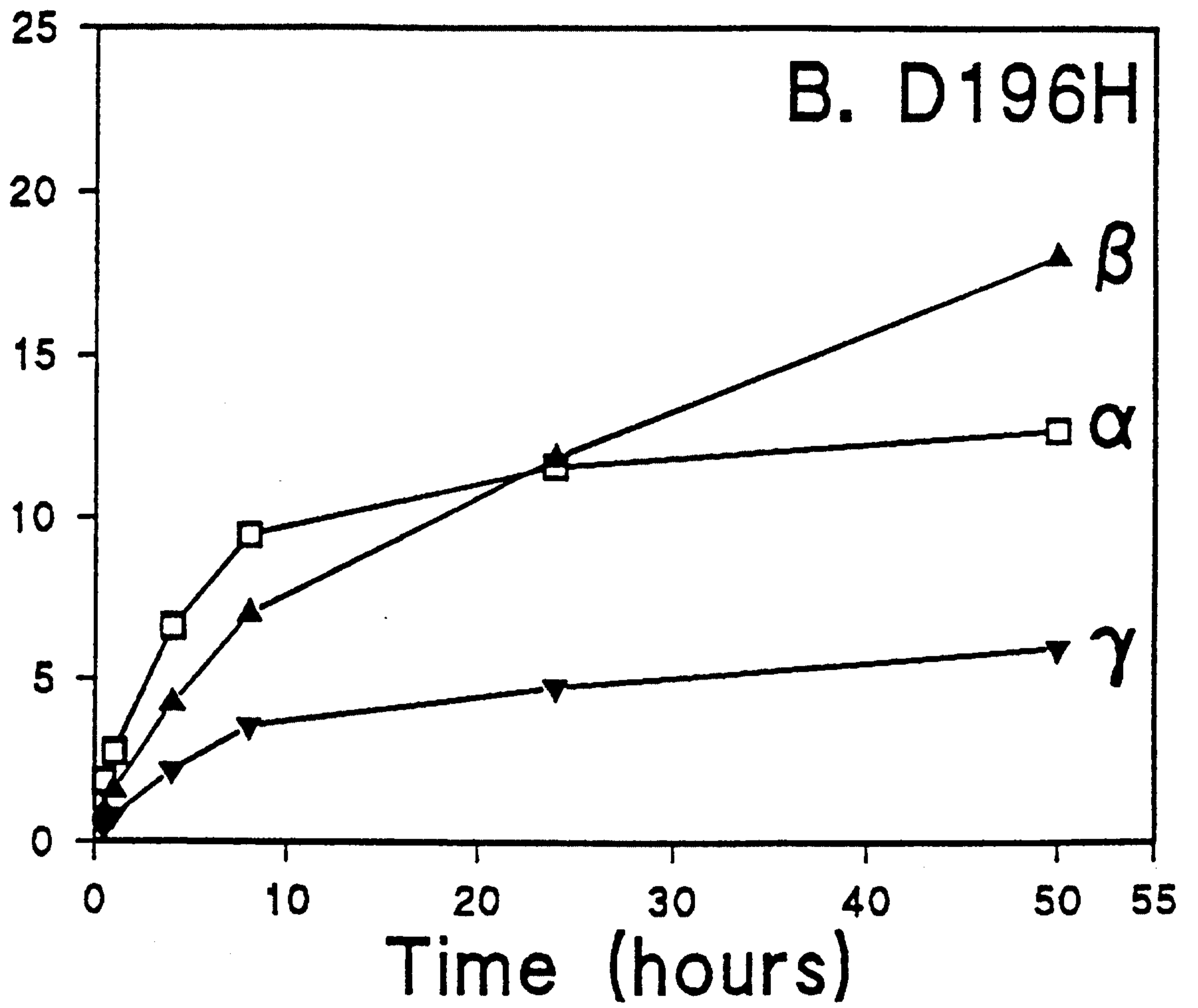


Fig. 3

4/4

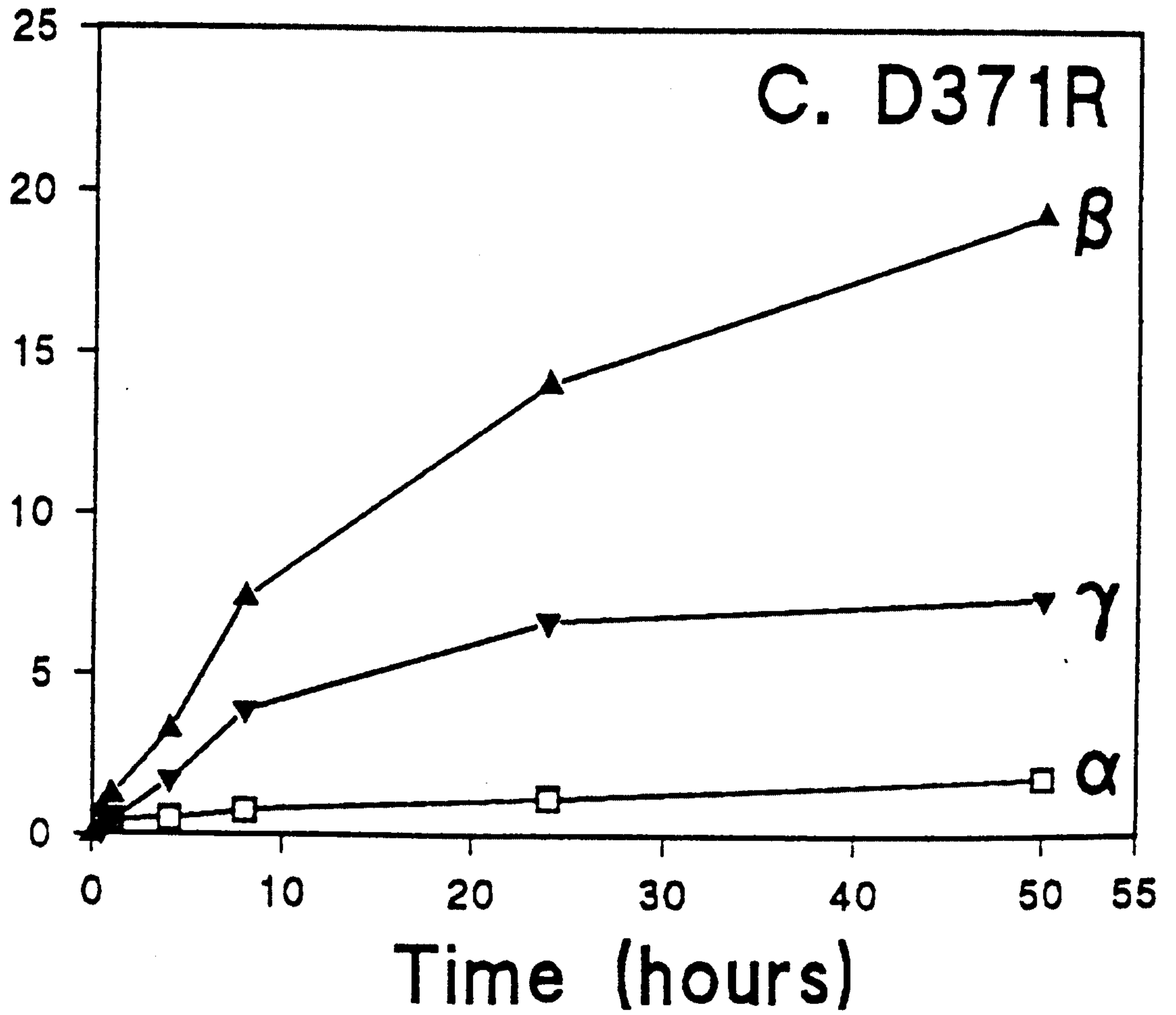


Fig. 4