



(86) Date de dépôt PCT/PCT Filing Date: 2009/01/16
(87) Date publication PCT/PCT Publication Date: 2009/07/23
(85) Entrée phase nationale/National Entry: 2010/07/14
(86) N° demande PCT/PCT Application No.: CA 2009/000053
(87) N° publication PCT/PCT Publication No.: 2009/089630
(30) Priorité/Priority: 2008/01/18 (US61/022,101)

(51) Cl.Int./Int.Cl. *C12N 9/42* (2006.01),
C12N 1/15 (2006.01), *C12N 1/19* (2006.01),
C12N 15/56 (2006.01), *C12N 15/80* (2006.01),
C12P 19/02 (2006.01)

(71) Demandeur/Applicant:
IOGEN ENERGY CORPORATION, CA

(72) Inventeurs/Inventors:
LAVIGNE, JAMES A., CA;
HILL, CHRISTOPHER M.D., CA;
ST-PIERRE, PATRICK, CA;
TOMASHEK, JOHN J., CA;
TREMBLAY, ANNIE, CA

(74) Agent: GOWLING LAFLEUR HENDERSON LLP

(54) Titre : VARIANTS DE CELLULASE PRESENTANT UNE INHIBITION REDUITE PAR LE GLUCOSE
(54) Title: CELLULASE VARIANTS WITH REDUCED INHIBITION BY GLUCOSE

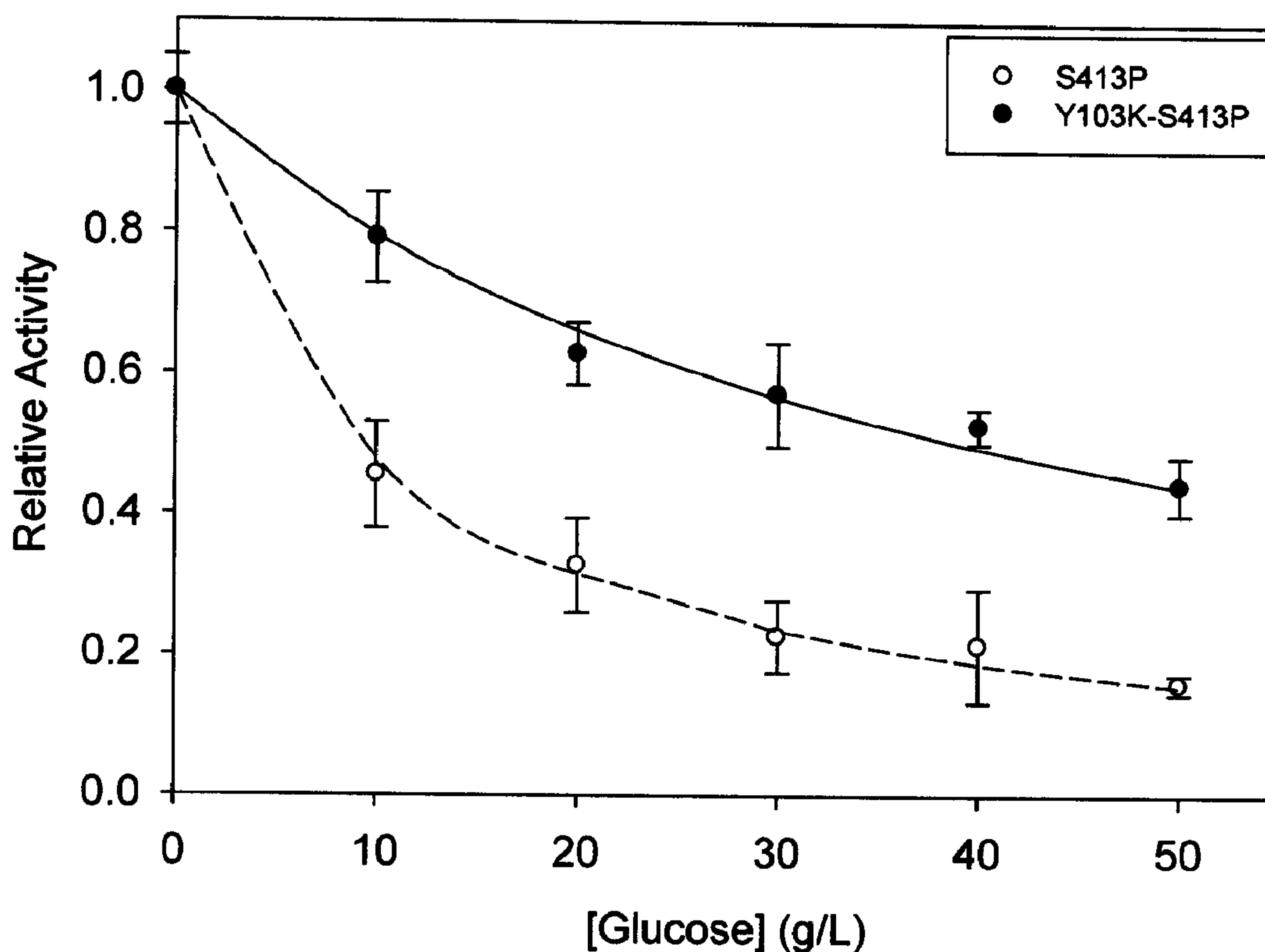


FIG. 6

(57) Abrégé/Abstract:

A Family 6 cellulase variant enzyme comprising one or more than one amino acid substitution selected from a basic, polar or non-polar amino acid at position 103, a valine or isoleucine at position 136, a tyrosine at position 186, a glutamic acid or glutamine at

(57) Abrégé(suite)/Abstract(continued):

position 365 and a glutamine at position 410 is provided (said position determined from alignment of the parental Family 6 with SEQ ID NO: 1). Genetic constructs and genetically modified microbes comprising DNA sequences encoding the Family 6 cellulase variant are also provided. Family 6 cellulases of the invention display reduced inhibition by glucose relative to the parent Family 6 cellulases. Such cellulases find use in a variety of applications in industry, e.g., in the hydrolysis of pretreated lignocellulosic feedstock, that require cellulose activity in the presence glucose concentrations that would otherwise inhibit the activity of the parental enzyme.

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

(19) World Intellectual Property Organization
International Bureau

PCT

(43) International Publication Date
23 July 2009 (23.07.2009)(10) International Publication Number
WO 2009/089630 A1

(51) International Patent Classification:

C12N 9/42 (2006.01) *C12N 15/56* (2006.01)
C12N 1/15 (2006.01) *C12N 15/80* (2006.01)
C12N 1/19 (2006.01) *C12P 19/02* (2006.01)

(21) International Application Number:

PCT/CA2009/000053

(22) International Filing Date: 16 January 2009 (16.01.2009)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

61/022,101 18 January 2008 (18.01.2008) US

(71) Applicant (for all designated States except US): **IOGEN ENERGY CORPORATION** [CA/CA]; 310 Hunt Club Road, East, Ottawa, Ontario K1V 1C1 (CA).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **LAVIGNE, James, A.** [CA/CA]; 219 Gershwin Private, Ottawa, Ontario K2H 1G6 (CA). **HILL, Christopher, M.D.** [CA/CA]; 1149 Shillington Avenue, Unit 4, Ottawa, Ontario K1Z 7Z3 (CA). **TREMBLAY, Annie** [CA/CA]; 6 rue du Printemps,Gatineau, Quebec J9A 3N9 (CA). **ST-PIERRE, Patrick** [CA/CA]; 284 Rue de Morency, Gatineau, Quebec J8V 3R2 (CA). **TOMASHEK, John, J.** [CA/CA]; 157 Daly Avenue, #1, Ottawa, Ontario K1N 6E8 (CA).(74) Agent: **SECHLEY, Konrad, A.**; Gowling Lafleur Henderson LLP, P.O. Box 30, Suite 2300, 550 Burrard Street, Vancouver, British Columbia V6C 2B5 (CA).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK,

[Continued on next page]

(54) Title: CELLULASE VARIANTS WITH REDUCED INHIBITION BY GLUCOSE

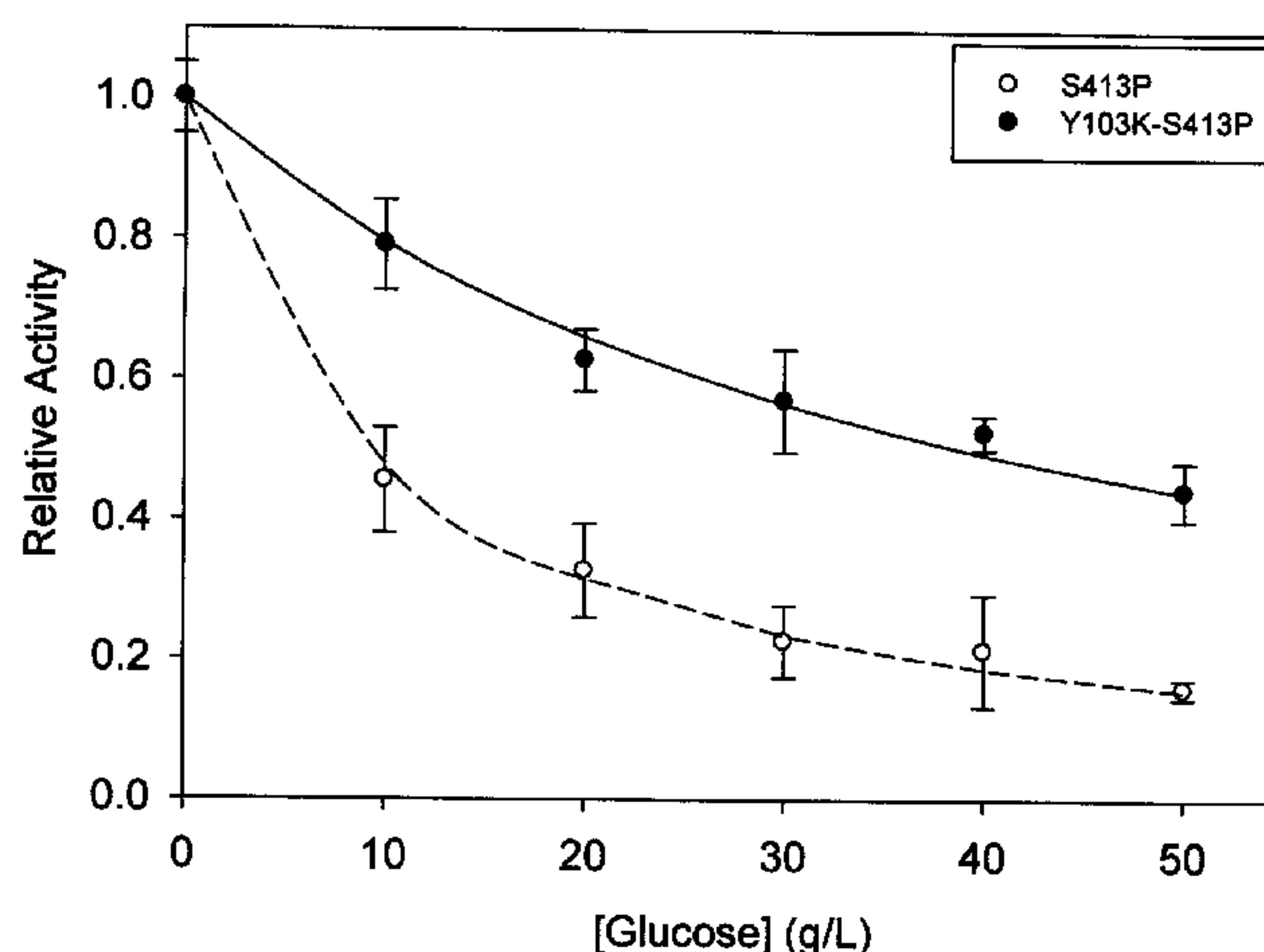


FIG. 6

(57) Abstract: A Family 6 cellulase variant enzyme comprising one or more than one amino acid substitution selected from a basic, polar or non-polar amino acid at position 103, a valine or isoleucine at position 136, a tyrosine at position 186, a glutamic acid or glutamine at position 365 and a glutamine at position 410 is provided (said position determined from alignment of the parental Family 6 with SEQ ID NO: 1). Genetic constructs and genetically modified microbes comprising DNA sequences encoding the Family 6 cellulase variant are also provided. Family 6 cellulases of the invention display reduced inhibition by glucose relative to the parent Family 6 cellulases. Such cellulases find use in a variety of applications in industry, e.g., in the hydrolysis of pretreated lignocellulosic feedstock, that require cellulase activity in the presence of glucose concentrations that would otherwise inhibit the activity of the parental enzyme.

WO 2009/089630 A1

WO 2009/089630 A1



MT, NL, NO, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

Published:

— *with international search report*

CELLULASE VARIANTS WITH REDUCED INHIBITION BY GLUCOSE**TECHNICAL FIELD**

[0001] This application claims benefit of prior provisional application No. 61/022,101 filed
5 January 18, 2008.

[0002] The present invention relates to cellulase variants. More specifically, the invention relates to Family 6 cellulase variants with reduced inhibition by glucose. The present invention also relates to genetic constructs comprising nucleotide sequences encoding for Family 6 cellulase variants, methods for the production of Family 6 cellulase variants from host strains
10 and the use of Family 6 cellulase variants in the hydrolysis of cellulose.

BACKGROUND OF THE INVENTION

[0003] Cellulose is an unbranched polymer of glucose linked by $\beta(1\rightarrow4)$ -glycosidic bonds. Cellulose chains can interact with each other via hydrogen bonding to form a crystalline solid of high mechanical strength and chemical stability. The cellulose chains must be depolymerized
15 into glucose and short oligosaccharides before organisms, such as the fermenting microbes used in ethanol production, can use them as metabolic fuel. Cellulase enzymes catalyze the hydrolysis of the cellulose (hydrolysis of β -1,4-D-glucan linkages) in the feedstock into products such as glucose, cellobiose, and other cellooligosaccharides. Cellulase is a generic term denoting a multienzyme mixture comprising exo-acting cellobiohydrolases (CBHs), endoglucanases (EGs)
20 and β -glucosidases (β G) that can be produced by a number of plants and microorganisms. Enzymes in the cellulase of *Trichoderma reesei* include CBH1 (more generally, Cel7A), CBH2 (Cel6A), EG1 (Cel7B), EG2 (Cel5), EG3 (Cel12), EG4 (Cel61A), EG5 (Cel45A), EG6 (Cel74A), Cip1, Cip2, β -glucosidases (including, e.g., Cel3A), acetyl xylan esterase, β -mannanase, and swollenin.

25 [0004] Cellulase enzymes work synergistically to hydrolyze cellulose to glucose. CBH1 and CBH2 act on opposing ends of cellulose chains (Barr *et al.*, 1996), while the endoglucanases act at internal locations in the cellulose. The primary product of these enzymes is cellobiose, which is further hydrolyzed to glucose by one or more β -glucosidases.

[0005] The kinetics of the enzymatic hydrolysis of insoluble cellulosic substrates by cellulases
30 do not follow simple Michaelis-Menten behaviour (Zhang *et al.*, 1999). Specifically, increasing the dosage of cellulase in a hydrolysis reaction does not provide a linearly dependent increase in the amount of glucose produced in a given time. There is also a significant decrease in the rate

of reaction as cellulose hydrolysis proceeds (Tolan, 2002). Several explanations have been proposed to explain the decline in the reaction rate; the major hypotheses include substrate heterogeneity (Nidetsky and Steiner, 1993; Zhang *et al.*, 1999), enzyme inactivation (Caminal *et al.*, 1985; Converse *et al.*, 1988; Gusakov and Sinitsyn, 1992; Eriksson *et al.*, 2002), and product inhibition (Lee and Fan, 1983; Caminal *et al.*, 1985; Holtzapple *et al.*, 1990; Gusakov and Sinitsyn, 1992; Eriksson *et al.*, 2002; Gruno *et al.*, 2004).

[0006] Inhibition of enzymes by the products of the reactions they catalyze has long been recognized; the phenomenon was known to Henri, Michaelis, and Menten, all pioneers in the field of enzymology (Frieden and Walter, 1963). The nature of product inhibition may be competitive, as product competes with substrate to form the same interactions with the enzyme, but other forms of inhibition are possible. Indeed, due to the insoluble nature of cellulose and the challenges it poses as a substrate in kinetic studies, there have been many conflicting reports as to the nature of inhibition in the cellulase system (Holtzapple *et al.*, 1990, and references therein). The cellobiohydrolases are subject to inhibition by their direct product, cellobiose, and to a lesser degree by the glucose produced by the further hydrolysis of the cellobiose by β -glucosidase. One technique for reducing cellulase inhibition is to increase the amount of β -glucosidase in the system (U.S. Patent No. 6,015,703), as cellobiose is more inhibitory to cellulases than glucose (Holtzapple *et al.*, 1990; Teleman *et al.*, 1995). Inhibition can be mitigated by altering the primary sequence of the protein using DNA mutagenesis guided by rational design or applied randomly. For example, rational design was used to target the Y245 residue in Cel5A, an endoglucanase, for mutagenesis, which resulted in an increase in its cellobiose inhibition constant (U.S. Publication No. 2003/0054535).

[0007] There are relatively few reports of engineering Cel6A (also known as cellobiohydrolase II or CBH2), a major cellobiohydrolase of the *T. reesei* (also known as *Hypocrea jecorina*) cellulase system, for desirable properties. St-Pierre *et al.* (U. S. Publication No. 2008/0076152) have shown that substitution of the naturally occurring amino acids at the equivalent of positions 231, 305, 410 and 413 in the *T. reesei* Cel6A sequence to serine or threonine (positions 231 and 305), glutamine or asparagines (position 410) or proline (position 413) increase thermostability, thermophilicity and/or alkalophilicity of a Family 6 cellulase. Wohlfahrt *et al.* have enhanced the stability of the protein by forming amide-carboxylate pairs through mutagenesis at residues E107, D170 and D366 (U.S. Publication No. 2004/0152872). Rational design was also applied to a related cellobiohydrolase, Cel6B from *Thermobifida fusca*, to relieve inhibition by cellobiose (Zhang *et al.*, 2000). Mutations at Cel6B residues equivalent to W269, H266, and E399 in the

Cel6A sequence were shown to reduce cellobiose inhibition, but at a significant cost to activity on crystalline cellulose. Another approach, based on the consensus sequence derived from an alignment of Cel6A sequences from several species (U.S. Publication No. 2006/0205042), identified 38 amino acids associated with improved thermostability (specifically: V94, P98, G118, M120, M134, T142, M145, T148, T154, L179, Q204, V206, I212, L215, G231, T232, V250, Q276, N285, S291, G308, T312, S316, V323, N325, I333, G334, S343, T349, G360, S380, A381, S390, F411, S413, A416, Q426 and A429). The authors speculate that these mutations may also affect product inhibition and/or enzyme processivity, but offer no data or specific hypotheses based on modeling to associate changes in these properties with the claimed residues. The consensus approach is designed to generate protein variants with improved thermodynamic stability (Steipe, 2004) and it does not provide predictive power for the improvement of any other biochemical property.

[0008] Although cellulase compositions have been described previously, there remains a need for new and improved compositions for use in the conversion of cellulose into fermentable sugars and for related fields of cellulosic material processing such as pulp and paper, textiles and animal feeds. Cellulases with improved performance decrease the cost of the processes and typically offer substantial environmental benefits when compared to the equivalent chemical and/or physical processes. For example, the production of fuel ethanol from cellulose delivers substantial environmental and economic benefits. When compared to gasoline, using ethanol as a fuel significantly reduces net carbon emissions by fixing the carbon dioxide released during combustion back into the biomass grown as feedstock for further ethanol production. Using agricultural biomass as feedstock can also stimulate rural economies and reduce dependence on foreign petroleum. Producing ethanol from cellulose rather than starch, as for corn ethanol, or sugar has the additional benefit of avoiding competition with the production of foodstuffs for humans and animals. The US Departments of Agriculture and Energy estimate that 30% of transportation fuel use in America, the largest petroleum market in the world, could be displaced by using cellulosic fuel without affecting food harvests (Perlack *et al.*, 2005). Additionally, due to the low energy input required to generate cellulosic biomass, it has been estimated that the use of cellulose ethanol reduces net greenhouse gas production by 88% when compared to gasoline whereas using corn ethanol produces a decrease of only 18% (Farrell *et al.*, 2006).

SUMMARY OF THE INVENTION

[0009] The present invention relates to Family 6 cellulase variants. More specifically, the invention relates to Family 6 cellulase variants that exhibit reduced inhibition by glucose. The present invention also relates to genetic constructs comprising nucleotide sequences encoding for Family 6 cellulase variants, methods for the production of Family 6 cellulase variants from host strains and the use of the Family 6 cellulase variants in the hydrolysis of cellulose.

[0010] It is an object of the invention to provide an improved cellulase with reduced inhibition by glucose.

[0011] This invention relates to a Family 6 cellulase variant comprising one or more of the amino acid substitutions selected from the group consisting of:

a basic, non-polar or proline residue at position 103 (X103H, K, R, A, V, L, P, M),

a valine or isoleucine residue at position 136 (X136V, I),

a tyrosine or lysine residues at position 186 (X186Y, K),

an acidic, glutamine or serine residue at position 365 (X365D, E, Q, S), and

an alanine, phenylalanine, leucine, glutamine or serine residue at position 410 (X410A, F, L, Q, S).

[0012] The positions of the amino acid substitution(s) are determined from sequence alignment of the Family 6 cellulase variant with a *Trichoderma reesei* Cel6A amino acid sequence as defined in SEQ ID NO: 1. The basic amino acid at position 103 may be a histidine, arginine or lysine or the non-polar amino acid at position 103 is an alanine, valine, methionine or leucine. The Family 6 cellulase variant of the present invention exhibits at least about 1.4-fold less inhibition by glucose than the parental Family 6 cellulase from which it is derived. For example, the Family 6 cellulase variant may exhibit from about 1.4-, 1.5-, 1.6-, 1.8-, 2.0-, 2.5-, 3.0-, 3.5-, 4-, 5-, 6-, 7-, 8-, 9-, 10-, 11-, 12-, 13-, 14-, 15-, 16-, 17-, 18-, 19- or 20-fold less inhibition by glucose than the parental Family 6 cellulase from which it is derived.

[0013] In one embodiment of the invention, the Family 6 cellulase variant of the present invention has an amino acid sequence that is at least about 45% to about 100% identical to the amino acids 83-447 of SEQ ID NO: 1. For example, the Family 6 cellulase variant may have an

amino acid sequence that is at least about 55% to about 100% identical to the amino acids 83-447 of SEQ ID NO: 1, or at least about 63% to about 100% identical to the amino acids 83-447 of SEQ ID NO: 1

5 [0014] In still another embodiment, the Family 6 cellulase variant of the present invention has an amino acid sequence that is least about 95% identical to amino acids 83-447 of SEQ ID NO: 1

[0015] The Family 6 cellulase variant may be derived from a parental Family 6 cellulase that comprises one or more naturally-occurring amino acid(s) at the substituted positions corresponding to that of the Family 6 cellulase variant, but that is otherwise identical to the Family 6 cellulase variant, for example a native Family 6 cellulase from *Neocallimastix*
10 *patriciarum*, *Orpinomyces*, or *Thermobifidia fusca*. The parental Family 6 cellulase may contain one or more amino acid substitutions at other positions, given that these substitutions are also present in the corresponding Family 6 cellulase variant.

[0016] This invention also includes a Family 6 cellulase variant as defined above and further comprising one or more of an isoleucine, valine, threonine, tyrosine or glutamine residue at
15 position 134, an isoleucine residue at position 215 and a proline residue at position 413. The Family 6 cellulase variant comprising these mutations may be from a filamentous fungus, such as *Trichoderma reesei*.

[0017] The present invention also relates to Family 6 cellulase variants exhibiting at least 1.4-fold less inhibition by glucose than a parental Family 6 cellulase from which it is derived, the
20 Family 6 cellulase variant being selected from the group consisting of:

TrCel6A-Y103A-S413P (SEQ ID NO: 37);
TrCel6A-Y103H-S413P (SEQ ID NO: 38);
TrCel6A-Y103K-S413P (SEQ ID NO: 39);
TrCel6A-Y103L-S413P (SEQ ID NO: 40);
25 TrCel6A-Y103M-S413P (SEQ ID NO: 41);
TrCel6A-Y103P-S413P (SEQ ID NO: 42);
TrCel6A-Y103R-S413P (SEQ ID NO: 43);
TrCel6A-Y103V-S413P (SEQ ID NO: 44);
TrCel6A-L136I -S413P (SEQ ID NO: 45);
30 TrCel6A-L136V -S413P (SEQ ID NO: 46);
TrCel6A-S186K -S413P (SEQ ID NO: 47);

TrCel6A-S186Y -S413P (SEQ ID NO: 48);
 TrCel6A-G365D-S413P (SEQ ID NO: 49);
 TrCel6A-G365E-S413P (SEQ ID NO: 50);
 TrCel6A-G365Q-S413P (SEQ ID NO: 51);
 5 TrCel6A-G365S-S413P (SEQ ID NO: 52);
 TrCel6A-R410A-S413P (SEQ ID NO: 53);
 TrCel6A-R410F-S413P (SEQ ID NO: 54);
 TrCel6A-R410L-S413P (SEQ ID NO: 55);
 TrCel6A-R410Q-S413P (SEQ ID NO: 56);
 10 TrCel6A-R410S-S413P (SEQ ID NO: 57);
 TrCel6A- M134V -L136I-S413P (SEQ ID NO: 62);
 TrCel6A-L136I-L215I-S413P (SEQ ID NO: 63);
 TrCel6A-M134V-L136I-L215I -S413P (SEQ ID NO: 71);
 HiCel6A-Y107K(SEQ ID NO: 78);
 15 HiCel6A-Y107L (SEQ ID NO: 79);
 HiCel6A-Q139T (SEQ ID NO: 80);
 HiCel6A-L141V (SEQ ID NO: 81);
 HiCel6A-A194Y (SEQ ID NO: 82);
 PcCel6A-Y98K (SEQ ID NO: 83);
 20 PcCel6A-Y98L (SEQ ID NO: 84);
 PcCel6A-L131I (SEQ ID NO: 85);
 PcCel6A-L131V (SEQ ID NO: 86);
 PcCel6A-S182K (SEQ ID NO: 87);
 PcCel6A-S182Y (SEQ ID NO: 88);
 25 PcCel6A-G359Q (SEQ ID NO: 89);
 PcCel6A-R404Q (SEQ ID NO: 90).

[0018] Furthermore, the present invention also relates to genetic constructs for directing
 expression and secretion of the Family 6 cellulase variant from a host microbe including, but not
 30 limited to, strains of *Trichoderma reesei* or *Saccharomyces cerevisiae*..

[0019] The present invention relates to genetic construct comprising a DNA sequence encoding
 A Family 6 cellulase variant having one or more of a basic, non-polar or proline residue at
 position 103, a valine or isoleucine residue at position 136, a tyrosine or lysine residue at

position 186, an acidic, glutamine or serine residue at position 365, or an alanine, phenylalanine, leucine, glutamine or serine residue at position 410, which DNA sequence is operably linked to DNA sequences regulating its expression and secretion from a host microbe. Preferably, the DNA sequences regulating the expression and secretion of the Family 6 cellulase variant are
5 derived from the host microbe used for expression of the isolated cellulase. The host microbe may be a yeast, such as *Saccharomyces cerevisiae*, or a filamentous fungus, such as *Trichoderma reesei*.

[0020] The invention also relates to a genetic construct as defined above, wherein the Family 6 cellulase variant encoded by the genetic construct further comprises one or more of a valine or
10 threonine residue at position 134, an isoleucine residue at position 215 and a proline residue at position 413. Preferably, the DNA sequences regulating the expression and secretion of the Family 6 cellulase variant are derived from a filamentous fungus, including, but not limited to, *Trichoderma reesei*.

[0021] The invention also relates to a genetically modified microbe comprising a genetic
15 construct encoding the Family 6 cellulase variant and capable of expression and secretion of a Family 6 cellulase variant comprising one or more of a basic, non-polar or proline residue at position 103, a valine or isoleucine residue at position 136, a tyrosine or lysine residues at position 186, an acidic, glutamine or serine residue at position 365, or an alanine, phenyl alanine, leucine, glutamine or serine residue at position 410. In one embodiment, the Family 6 cellulase
20 variant further comprises one or more of a valine or threonine residue at position 134, an isoleucine residue at position 215 and a proline residue at position 413. Preferably, the genetically modified microbe is a yeast or filamentous fungus. More preferably, the genetically modified microbe is a species of *Saccharomyces*, *Pichia*, *Hansenula*, *Hypocrea*, *Trichoderma*, *Aspergillus*, *Fusarium*, *Humicola* or *Neurospora*.

[0022] The present invention also relates to the use of a Family 6 cellulase variant comprising
25 one or more of a basic, non-polar or proline residue at position 103, a valine or isoleucine residue at position 136, a tyrosine or lysine residues at position 186, an acidic, glutamine or serine residue at position 365, or an alanine, phenylalanine, leucine, glutamine or serine residue at position 410 for hydrolysis of a cellulosic substrate.

[0023] The invention also relates to a process of producing the Family 6 cellulase variant as
30 defined above, including transformation of a yeast or fungal host with a genetic construct comprising a DNA sequence encoding the Family 6 cellulase variant, selection of recombinant

yeast or fungal strains expressing the Family 6 cellulase variant, and culturing the selected recombinant strains in submerged liquid fermentations under conditions that induce the expression of the Family 6 cellulase variant.

[0024] Family 6 cellulase variants of the present invention comprising one or more of a basic, non-polar or proline residue at position 103, a valine or isoleucine residue at position 136, a tyrosine or lysine residues at position 186, an acidic, glutamine or serine residue at position 365, or an alanine, phenylalanine, leucine, glutamine or serine residue at position 410 display reduced glucose inhibition relative to the parental Family 6 cellulases from which they are derived.

[0025] Family 6 cellulase variants, as described herein, cellulases find use in a variety of applications in industry that require high concentrations of cellulosic substrates and enzymes that can retain high activity in the presence of normally inhibitory concentrations of the glucose produced from such substrates. For example, Family 6 cellulase variants, as described herein, may be used for the purposes of saccharification of lignocellulosic feedstocks for the production of fermentable sugars, or improving the digestibility of feeds in ruminant and non-ruminant animals.

BRIEF DESCRIPTION OF THE DRAWINGS

[0026] **Figure 1** shows an amino acid sequence alignment of 35 fungal Family 6 cellulases to amino acids 83-447 of *Trichoderma reesei* Cel6A (TrCel6A; SEQ ID NO:1). The amino acid numbering for each cellulase is as indicated at the left and right of each sequence. The residues at positions 103, 134, 136, 186, 215, 365, 410, and 413 (relative to TrCel6A) are indicated with an arrow.).

[0027] **Figure 2** depicts plasmid vector YEp352/PGK91-1 Δ NheI- α_{ss} -TrCel6A-S413P directing the expression and secretion of parental and variant TrCel6A cellulases from recombinant *Saccharomyces cerevisiae*.

[0028] **Figure 3** depicts plasmid vector Yep352/91-1 Δ NheI- α_{ss} -6H-Hi(Pc)Cel6A directing the expression and secretion of native and variant HiCel6A or PcCel6A from recombinant *Saccharomyces cerevisiae*.

[0029] **Figure 4** depicts data demonstrating the activity of TrCel6A-S413P on Sigmacell50 which has been treated with a cellulase free of Cel7A and Cel6A. The cellobiohydrolase-free cellulase is added at the start of data acquisition and there is no discernable effect on apparent absorbance. Five

minutes after the start of data collection, TrCel6A-S413P is added and the apparent absorbance begins to decrease. The activity of the Cel6A is proportionate to the slope of the linear portion of the decrease in absorbance. Activity in the absence of glucose and in the presence of the highest concentration of glucose used, 50 g/L, is shown. Data represent averages of 5 replicates and error bars represent one standard deviation.

[0030] **Figure 5** shows the effect of increasing glucose concentration on the relative activity of the TrCel6A-S413P.

[0031] **Figure 6** shows the effect of increasing glucose concentration on the relative activity of the TrCel6A-S413P and TrCel6A-Y103K-S413P.

[0032] **Figure 7** shows the effect of increasing glucose concentration on the relative activity of the TrCel6A-S413P and TrCel6A-M134T-S413P.

[0033] **Figure 8** shows the effect of increasing glucose concentration on the relative activity of the TrCel6A-S413P and TrCel6A-L136I-S413P.

[0034] **Figure 9** shows the effect of increasing glucose concentration on the relative activity of the TrCel6A-S413P and TrCel6A-S186Y-S413P.

[0035] **Figure 10** shows the effect of increasing glucose concentration on the relative activity of the TrCel6A-S413P and TrCel6A-G365Q-S413P.

[0036] **Figure 11** shows the effect of increasing glucose concentration on the relative activity of the TrCel6A-S413P and TrCel6A-R410F-S413P.

[0037] **Figure 12** shows the effect of increasing glucose concentration on the relative activity of the TrCel6A-S413P and TrCel6A-M134V-L136I-S413P.

[0038] **Figure 13** shows the effect of increasing glucose concentration on the relative activity of the TrCel6A-S413P and TrCel6A-L136I-L215I-S413P.

[0039] **Figure 14** shows the effect of increasing glucose concentration on the relative activity of the TrCel6A-S413P and TrCel6A-M134V-L136I-L215I-S413P.

[0040] **Figure 15** shows the activity of TrCel6A-S413P and TrCel6A-L215I-S413P in the presence of 10 g/L glucose.

[0041] **Figure 16** shows SDS-PAGE analysis of parental Family 6 cellulases and Family 6 cellulase variants purified as described in Example 8 (TrCel6A-S413P) and Example 13 (HiCel6A variants, panel A) and PcCel6A variants (panel B). The purified Family 6 cellulases were visualized by Coomassie blue staining.

DETAILED DESCRIPTION OF THE INVENTION

[0042] The present invention relates to Family 6 cellulase variants. More specifically, the invention relates to Family 6 cellulase variants with reduced inhibition by glucose relative to the parental Family 6 cellulase from which it is derived. The present invention also relates to genetic constructs comprising nucleotide sequences encoding for Family 6 cellulase variants, methods for the production of Family 6 cellulase variants from host strains and the use of the Family 6 cellulase variants in the hydrolysis of cellulose.

[0043] The following description is of a preferred embodiment by way of example only and without limitation to the combination of features necessary for carrying the invention into effect.

DEFINITIONS

Family 6 cellulase variants

[0044] Family 6 (previously, Family B) cellulases enzymes are a group of enzymes that hydrolyse the beta-1,4 glucosidic linkages in cellulose with inversion of configuration of the anomeric carbon (Claeysens, M. and Henrissat, B. 1992). Most of the Family 6 cellulases identified thus far are mesophilic. However, this family also includes thermostable cellulases from *Thermobifida fusca* (TfCel6A and TfCel6B of Figure 1) and the alkalophilic cellulases from *Humicola insolens* (HiCel6A and HiCel6B of Figure 1).

[0045] Figure 1 and Table 1 show that there is a high degree of conservation of primary amino acid sequence among most of the cellulases of Family 6. Multiple alignment across 35 currently known Family 6 cellulase amino acid sequences shows that the most naturally occurring Family 6 cellulases of fungal origin show from about 47% to about 100% amino acid sequence identity to

amino acids 83-447 of TrCel6A (Figure 1 and Table 1); Family 6 cellulases of bacterial origin show a much lower degree of amino acid sequence identity to TrCel6A.

[0046] A cellulase is classified as a Family 6 cellulase if it comprises amino acids common to other Family 6 cellulase, including two aspartic acid (D) residues which may serve as catalytic residues. These aspartic acid residues are found at positions 175 and 221 (see Figure 1; based on TrCel6A numbering). By "TrCel6A numbering", it is meant the numbering corresponding to the position of amino acids based on the amino acid sequence of TrCel6A (Table 1; Figure 1; SEQ ID NO:1). As is evident by Figure 1, Family 6 cellulases exhibit a substantial degree of sequence similarity. Therefore, by aligning the amino acids to optimize the sequence similarity between cellulase enzymes, and by using the amino acid numbering of TrCel6A as the basis for numbering, the positions of amino acids within other cellulase enzymes can be determined relative to TrCel6A. Methods to align amino acid sequences are well known and available to those of skill in the art and include BLAST (Basic Local Alignment Search Tool, see URL: blast.ncbi.nlm.nih.gov/Blast.cgi) which is useful for aligning two sequences and CLUSTALW (see URL: ebi.ac.uk/Tools/clustalw2/index.html) for alignment of two or more sequences.

[0047] The conservation of overall three-dimensional structure provides further guidance to the alignment of the primary amino acid sequences of Family 6 cellulases. The topology of Family 6 catalytic domains is a variant of the α/β -barrel with a central β -barrel containing seven parallel β -strands connected by five α -helices. One important difference between Family 6 cellobiohydrolases and endo- β -1,4-glucanases is the length of their N- and C-terminal loops present on each side of the active site and which are responsible for their functional behavior on cellulose. In the cellobiohydrolases, an extensive C-terminal loop forms a tunnel with the N-terminal loop enclosing the active site. This confers the unique property of cellobiohydrolases to attack the ends of crystalline cellulose where the N- and C-terminal loops maintain a single cellulose chain in the active site and facilitate the processive degradation of the substrate. In the endo- β -1,4-glucanases, the C-terminal loop is reduced in length and the N-terminal loop pulls it away from the active site and could also be shorter resulting in a more open active site allowing access to internal β -1,4 glycosidic bonds of cellulose for hydrolysis. The role of these loops in the functional behavior of Family 6 enzymes on cellulose was confirmed by the deletion of fifteen amino acids of the C-terminal loop of the *Cellulomonas fimi* cellobiohydrolase Cel6B in order to mimic the properties of an endo- β -1,4-glucanase (Meinke, A., et al. 1995.). The mutation enhanced the endo- β -1,4-

glucanase activity of the enzyme on soluble cellulose, such as carboxymethylcellulose, and altered its cellobiohydrolase activity on insoluble cellulose.

[0048] For the purpose of this invention, a "Family 6 cellulase" is defined as an enzyme capable of hydrolyzing polysaccharides using an inverting mechanism and characterized by having an α/β -barrel structure with a central β -barrel containing seven parallel β -strands connected by five α -helices an amino acid sequence that is from about 47% to about 100% identical to the amino acids 83 to 447 of SEQ ID NO: 1, representing the Family 6 catalytic domain of TrCel6A. For example, a Family 6 cellulase may have an amino acid sequence that is about 47%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% or 100% identical to the amino acids 83- 447 of SEQ ID NO: 1. One of skill in the art recognizes that the amino acid sequence of a given Family 6 cellulase may be modified by the addition, deletion or substitution of one or more amino acids and still be considered a Family 6 cellulase. Techniques for altering amino acid sequences include, but are not limited to, site-directed mutagenesis, cassette mutagenesis, random mutagenesis, synthetic oligonucleotide construction, cloning and other genetic engineering techniques as would be known by those of skill in the art (Eijsink VG, et al. 2005., which is incorporated herein by reference). Non-limiting examples of Family 6 cellulase that may be modified following the general approach and methodology as outlined herein are provided in Table 1.

Table 1: Family 6 Cellulases

Organism	Protein	SEQ ID NO:	Identity with TrCel6A catalytic domain (amino acids 83-447) (%)
<i>Hypocrea koningii</i>	cellobiohydrolase II (Cbh2)	2	98.9
<i>Trichoderma viride</i> CICC 13038	cellobiohydrolase II (CbhII;Cbh2)	3	98.9
<i>Hypocrea koningii</i> 3.2774	cellobiohydrolase II (Cbh2;CbhII)	4	98.1
<i>Hypocrea koningii</i> AS3.2774	cbh2	5	97.8
<i>Trichoderma parceramosum</i>	cellobiohydrolase II (CbhII)	6	97.8
<i>Aspergillus nidulans</i> FGSC A4	cellobiohydrolase (AN5282.2)	7	72.4
<i>Aspergillus niger</i> CBS 513.88	An12g02220	8	72.4
<i>Aspergillus oryzae</i> RIB 40	AO090038000439	9	67.8
<i>Aspergillus niger</i> CBS 513.88	An08g01760	10	67.7
<i>Acremonium cellulolyticus</i>	cellobiohydrolase II (Acc2)	11	67.3

- 12 -

Y-94			
<i>Talaromyces emersonii</i>	cellobiohydrolase II (CbhII)	12	66.8
<i>Gibberella zeae</i> K59	Cel6 - Cel6	13	66.1
<i>Fusarium oxysporum</i>	endoglucanase B	14	66.1
<i>Neurospora crassa</i> OR74A	NCU09680.1 (64C2.180)	15	65.9
<i>Aspergillus nidulans</i> FGSC A4	AN1273.2	16	65.5
<i>Aspergillus tubingensis</i>	unnamed protein product (fragment)	17	65.5
<i>Magnaporthe grisea</i> 70-15	MG05520.4	18	65.4
<i>Chaetomium thermophilum</i>	unnamed protein product	19	65.1
<i>Chaetomium thermophilum</i> CT2	cellobiohydrolase (Cbh2)	20	65.0
<i>Stilbella annulata</i>	unnamed protein product	21	64.9
<i>Humicola insolens</i>	avicelase 2 (Avi2)	22	63.7
<i>Humicola insolens</i>	cellobiohydrolase (CBHII) - Cel6A	23	63.1
<i>Cochliobolus heterostrophus</i> C4	cellobiohydrolase II (CEL7)	24	59.6
<i>Agaricus bisporus</i> D649	cellobiohydrolase II (Cel3; Cel3A)	25	57.7
<i>Polyporus arcularius</i> 69B-8	cellobiohydrolase II (Cel2)	26	57.1
<i>Lentinula edodes</i> Stamets CS-2	cellulase - Cel6B	27	56.3
<i>Lentinula edodes</i> L54	cellobiohydrolase (CbhII-1)	28	56.0
<i>Malbranchea cinnamomea</i>	unnamed protein product	29	54.9
<i>Phanerochaete chrysosporium</i>	cellobiohydrolase II	30	54.9
<i>Volvariella volvacea</i>	cellobiohydrolase II-I (CbhII-I)	31	53.8
<i>Chrysosporium lucknowense</i>	cellobiohydrolase (EG6; CBH II) - Cel6A	32	49.5
<i>Pleurotus sajor-caju</i>	cellobiohydrolase II	33	47.2
<i>Trametes versicolor</i>	ORF	34	47.0
<i>Neurospora crassa</i> OR74A	NCU03996.1	35	46.8
<i>Magnaporthe grisea</i> 70-15	MG04499.4	36	45.1

[0049] By "Family 6 cellulase variant" or "modified Family 6 cellulase", it is meant a Family 6 cellulase which comprises one or more than one the following amino acid substitutions: the amino acid at position 103 has been replaced by a basic, non-polar or proline residue; the amino acid at position 136 has been replaced by a valine or isoleucine residue; the amino acid at position 186 has been replaced by a tyrosine or lysine residue; the amino acid at position 365 has been replaced by an acidic, glutamine or serine residue; or the amino acid at position 410 has been replaced by an

alanine, phenylalanine, leucine, glutamine or serine residue; said position determined from sequence alignment of said isolated cellulase with a *Trichoderma reesei* Cel6A amino acid sequence as defined in SEQ ID NO:1. It will be understood that the Family 6 cellulase variant may be derived from any Family 6 cellulase. For example, the Family 6 cellulase variant may be derived from a wild-type cellulase or from a cellulase that already contains other amino acid substitutions. In one embodiment of the invention, the Family 6 cellulase variant exhibits reduced inhibition by glucose over the corresponding parental Family 6 cellulase from which it is derived.

[0050] By “wild type” or “native” Family 6 cellulase, it is meant a Family 6 cellulase having an amino acid sequence as encoded by the genome of the organism that naturally produces such Family 6 cellulase without the introduction of any substitutions, deletions, additions or modifications. For example, by wild type TrCel6A, wild type HiCel6A and wild type PcCel6A it is meant the cellulases of SEQ ID NO: 1, SEQ ID NO: 23 and SEQ ID NO: 30 respectively, without any amino acid substitutions.

[0051] By “parental Family 6 cellulase”, it is meant a Family 6 cellulase that exhibits at least 1.4-fold less inhibition by glucose and comprises one or more naturally-occurring amino acid(s) at the mutated positions corresponding to that of the Family 6 cellulase variant, but that is otherwise identical to the Family 6 cellulase variant. The parental Family 6 cellulase does not include those cellulases in which the naturally-occurring amino acid at position 103 is a basic, non-polar or proline residue, the naturally occurring amino acid at position 136 is a valine or isoleucine residue, the naturally occurring amino acid at position 186 is a tyrosine or lysine residue, the naturally occurring amino acid at position 365 is an acidic, glutamine or serine residue, and/or the naturally occurring amino acid at position 410 is an alanine, phenylalanine, leucine, glutamine or serine residue. This definition encompasses parental Family 6 cellulases that contain one or more additional amino acid substitution at other positions that have been introduced by genetic engineering or other techniques, provided that these substitutions are also present in the corresponding Family 6 cellulase variant. For example, the parental cellulase may contain a mutation of the amino acid at position 413 (TrCel6A numbering) to a proline to confer increased thermostability. Non-limiting examples of Family 6 cellulase variants may be found in Table 2.

Table 2: Family 6 cellulase variants

Protein	SEQ ID NO:	Protein	SEQ ID NO:
TrCel6A-Y103A-S413P	37	TrCel6A-S186Y-G365Q-S413P	67
TrCel6A-Y103H-S413P	38	TrCel6A-S186Y-R410Q-S413P	68

TrCel6A-Y103K-S413P	39	TrCel6A-G365Q-R410Q-S413P	69
TrCel6A-Y103L-S413P	40	TrCel6A-M134V-L136I-S413P	62
TrCel6A-Y103M-S413P	41	TrCel6A-L136I-L215I-S413P	63
TrCel6A-Y103P-S413P	42	TrCel6A-M134V-L136I-L215I-S413P	71
TrCel6A-Y103R-S413P	43	TrCel6A-Y103L-S186Y-G365Q-S413P	70
TrCel6A-Y103V-S413P	44	TrCel6A-Y103L-L136I-S186Y-G365Q-S413P	72
TrCel6A-L136I-S413P	45	TrCel6A-Y103L-S186Y-G365Q-R410Q-S413P	75
TrCel6A-L136V-S413P	46	TrCel6A-Y103L-L136I-S186Y-G365Q-R410Q-S413P	77
TrCel6A-S186K-S413P	47	TrCel6A-L136I-S186Y-G365Q-R410Q-S413P	76
TrCel6A-S186Y-S413P	48	TrCel6A-Y103L-L136I-G365Q-R410Q-S413P	74
TrCel6A-G365D-S413P	49	TrCel6A-Y103L-L136I-S186Y-R410Q-S413P	73
TrCel6A-G365E-S413P	50	HiCel6A-Y107K	78
TrCel6A-G365Q-S413P	51	HiCel6A-Y107L	79
TrCel6A-G365S-S413P	52	HiCel6A-Q139T	80
TrCel6A-R410A-S413P	53	HiCel6A-L141V	81
TrCel6A-R410F-S413P	54	HiCel6A-A194Y	82
TrCel6A-R410L-S413P	55	PcCel6A-Y98K	83
TrCel6A-R410Q-S413P	56	PcCel6A-Y98L	84
TrCel6A-R410S-S413P	57	PcCel6A-L131I	85
TrCel6A-Y103L-L136I-S413P	58	PcCel6A-L131V	86
TrCel6A-Y103L-S186Y-S413P	59	PcCel6A-S182K	87
TrCel6A-Y103L-G365Q-S413P	60	PcCel6A-S182Y	88
TrCel6A-Y103L-R410Q-S413P	61	PcCel6A-G359Q	89
TrCel6A-L136I-S186Y-S413P	64	PcCel6A-R404Q	90
TrCel6A-L136I-G365Q-S413P	65		
TrCel6A-L136I-R410Q-S413P	66		

Reduced glucose inhibition

[0052] Glucose inhibition of cellulases is measured by determination of the inhibition constant K_G , defined as the concentration of glucose which reduces the activity of the cellulase by 50%. The value of K_G is not dependent on the nature of product inhibition—i.e., competitive, non-competitive or mixed-type. Cellulases that are less inhibited by glucose will have a higher value for K_G —i.e., it takes a higher concentration of glucose to reduce the enzyme activity by 50%.

[0053] For the purposes of the present invention, a Family 6 cellulase variant exhibits reduced glucose inhibition with respect to the corresponding parental Family 6 cellulase if it has a K_G which is at least 1.4-fold, or at least about 1.8-fold, higher than that of the parental Family 6 cellulase. The K_G is the concentration of glucose which reduces the activity of the isolated and parental Family 6 cellulases by 50% and is determined by the assay detailed in Example 9.

[0054] The Family 6 cellulase variant may have a K_G which is about 1.4-fold higher than that of a corresponding parental Family 6 cellulase, or about 1.8-fold higher. For example, the Family 6 cellulase variant may have a K_G that is at least about 1.4-, 1.5-, 1.6-, 1.8-, 2.0-, 2.5-, 3.0-, 3.5-, 4.0-, 4.5-, 5.0-, 5.5-, 6.0-, 6.5-, 7.0-, 8.0-, 10.0-, 12.0-, 15.0- or 20.0-fold higher than that of the corresponding parental Family 6 cellulase.

[0055] Examples of Family 6 cellulase variants exhibiting reduced inhibition by glucose are shown in Table 2.

Genetic Constructs Encoding Family 6 cellulase variant

[0056] The present invention also relates to genetic constructs comprising a DNA sequence encoding the Family 6 cellulase variant operably linked to regulatory DNA sequences directing the expression and secretion of the Family 6 cellulase variant from a host microbe. By “regulatory DNA sequences” it is meant a promoter and a DNA sequence encoding a secretion signal peptide. The regulatory DNA sequences are preferably functional in a fungal host. The regulatory DNA sequences may be derived from genes that are highly expressed and secreted in the host microbe under industrial fermentation conditions. In a preferred embodiment, the regulatory sequences are derived from any one or more of the *Trichoderma reesei* cellulase or hemicellulase genes.

[0057] The genetic construct may further comprise a selectable marker gene to enable isolation of a genetically modified microbe transformed with the construct as is commonly known to those of

skill in the art. The selectable marker gene may confer resistance to an antibiotic or the ability to grow on medium lacking a specific nutrient to the host organism that otherwise could not grow under these conditions. The present invention is not limited by the choice of selectable marker gene, and one of skill in the art may readily determine an appropriate gene. In a preferred embodiment, the selectable marker gene confers resistance to hygromycin, phleomycin, kanamycin, geneticin, or G418, complements a deficiency of the host microbe in one of the *trp*, *arg*, *leu*, *pyr4*, *pyr*, *ura3*, *ura5*, *his*, or *ade* genes or confers the ability to grow on acetamide as a sole nitrogen source.

[0058] The genetic construct may further comprise other DNA sequences, for example, transcriptional terminators, DNA encoding peptide tags, synthetic sequences to link the various DNA sequences together, origins of replication, and the like. The practice of the present invention is not limited by the presence of any one or more of these other DNA sequences.

Genetically Modified Microbes Producing Family 6 Cellulase Variants

[0059] The Family 6 cellulase variant may be expressed and secreted from a genetically modified microbe produced by transformation of a host microbe with a genetic construct encoding the Family 6 cellulase variant. The host microbe may be a yeast or a filamentous fungus including, but not limited to, a species of *Saccharomyces*, *Pichia*, *Hansenula*, *Trichoderma*, *Hypocrea*, *Aspergillus*, *Fusarium*, *Humicola*, or *Neurospora*. For example, the host microbe may be *Saccharomyces cerevisiae* or an industrial strain of *Trichoderma reesei*. Typically, the host microbe is one which does not contain a gene encoding a Family 6 cellulase or from which the gene(s) encoding any or all Family 6 cellulases have been deleted.

[0060] The genetic construct may be introduced into the host microbe by any number of methods known by one skilled in the art of microbial transformation, including but not limited to, treatment of cells with CaCl_2 , electroporation, biolistic bombardment, PEG-mediated fusion of protoplasts (e.g. White et al., WO 2005/093072, which is incorporated herein by reference). After selecting the recombinant fungal strains expressing the Family 6 cellulase variant, the selected recombinant strains may be cultured in submerged liquid fermentations under conditions that induce the expression of the Family 6 cellulase variant.

The Use of Family 6 Cellulase Variant in the Hydrolysis of Cellulosic Substrates

[0061] The Family 6 cellulase variant of the invention is used for the enzymatic hydrolysis of cellulose. The Family 6 cellulase variants of the present invention are particularly useful for the hydrolysis of cellulose-containing substrates under conditions where an inhibitory level of glucose is present or accumulates. For example, the Family 6 cellulase variants of the present invention may be useful in processes in which the initial starting concentration of cellulose is about 20 g/L and in which about 50% of the cellulose will be converted to glucose or in which the initial starting concentration of cellulose are as high as about 200 g/L and in which about 5% of the cellulose will be converted to glucose. For example, the initial cellulose concentration may be 20, 30, 40, 50, 60, 80, 100, 120, 150 or 200 g/L and about 50%, 33%, 25%, 20%, 17%, 12%, 10%, 8%, 7% or 5% of the cellulose, respectively, is converted to glucose. In the case of using the parental TrCel6A-S413P in such processes, the concentration of glucose will therefore be similar to the K_G of the enzyme and reduce its activity by about 50% (Table 5, Figure 5).

[0062] For example, the Family 6 cellulase variant of the present invention may be used for the enzymatic hydrolysis of the cellulose present in "pretreated lignocellulosic feedstock." A pretreated lignocellulosic feedstock is a material of plant origin that, prior to pretreatment, contains at least 20% cellulose (dry wt) and at least 10% lignin (dry wt), and that has been subjected to physical and/or chemical processes to make the fiber more accessible and/or receptive to the actions of cellulolytic enzymes. After pretreatment, the lignocellulosic feedstock may contain greater than about 20% cellulose and greater than about 10% lignin. In one embodiment, the pretreated lignocellulosic feedstock contains greater than about 20% cellulose and greater than about 12% lignin. Non-limiting examples of pretreatment processes include chemical treatment of a lignocellulosic feedstock with sulfuric or sulfurous acid, or other acids; ammonia, lime, ammonium hydroxide, or other bases; ethanol, butanol, or other organic solvents; or pressurized water (See U.S. Patent Nos. 4,461,648, 5,916,780, 6,090,595, 6,043,392, 4,600,590, Weil et al. (1997) and Öhgren, K., et al. (2005)).

[0063] Lignocellulosic feedstocks that may be used in the invention include, but are not limited to, agricultural residues such as corn stover, wheat straw, barley straw, rice straw, oat straw, canola straw, and soybean stover; fiber process residues such as corn fiber, sugar beet pulp, pulp mill fines and rejects or sugar cane bagasse; forestry residues such as aspen wood, other hardwoods, softwood, and sawdust; or grasses such as switch grass, miscanthus, cord grass, and reed canary grass. The lignocellulosic feedstock may be first subjected to size reduction by methods including, but not limited to, milling, grinding, agitation, shredding, compression/expansion, or other types of

mechanical action. Size reduction by mechanical action can be performed by any type of equipment adapted for the purpose, for example, but not limited to, a hammer mill.

[0064] By the term “enzymatic hydrolysis”, it is meant a process by which cellulase enzymes or mixtures, including those comprising the Family 6 cellulase variant of the present invention, act on cellulose to convert all or a portion thereof to soluble sugars. By the term “cellulase mixture”, it is meant a mixture of enzymes which decomposes cellulose and comprises one or more “endoglucanases”, enzymes that hydrolyze beta-1,4 glycosidic bonds within cellulose chains, “cellobiohydrolases” or “exoglucanases”, enzymes that sequentially cleave cellobiose from either the reducing or non-reducing end of a cellulose chain, and one or more “beta-glucosidases”, enzymes that hydrolyzes cellobiose to glucose. A cellulase mixture may contain other enzymes or proteins in addition to endoglucanases, cellobiohydrolases and beta-glucosidases.

[0065] The enzymatic hydrolysis process preferably converts about 80% to about 100% of the cellulose to soluble sugars, or any range therebetween. For example, the enzymatic hydrolysis process converts about 90% to about 100% of the cellulose to soluble sugars, or any range therebetween. In a preferred embodiment, the enzymatic hydrolysis process converts about 98% to about 100% of the cellulose to soluble sugars, or any range therebetween.

[0066] Preferably, the primary cellulases are produced in one or more submerged liquid culture fermentations and separated from the cells at the end of the fermentation. The cells may be separated from the cellulases by filtration, centrifugation, or other processes familiar to those skilled in the art. The cell-free cellulase-containing fraction may then be concentrated (for example, via ultrafiltration), preserved, and/or stabilized prior to use. Alternatively, the primary cellulases are not separated from the cells, but are added to the enzymatic hydrolysis with the cells.

Decreasing the Glucose Inhibition of Family 6 Cellulases

[0067] The glucose inhibition constants of the Family 6 cellulase variants were determined by incubation of the enzyme in the presence of substrate at various concentrations of glucose. Activity of the cellulase was determined via a turbidometric assay with insoluble acid swollen cellulose as a substrate. By “turbidometric assay” it is meant an assay of the optical density of a suspension of particulates in a fluid, in which the fraction of incident light which is not scattered by the particles, and therefore directly transmitted through the suspension, is measured. “Nephelometry” is a related

technique which measures the fraction of light which is scattered by the particles, typically from an observation angle of 90 degrees from the incident light.

[0068] The effect of amino acid substitutions at positions 103, 136, 186, 365 and 410, in combination with each other or with additional amino acid substitutions at positions 134 and 215 on the glucose tolerance of TrCel6A with substitution of the serine at position 413 with proline, was determined via a comparative study of variants of the parental Family 6 cellulases TrCel6A-S413P, HiCel6A and PcCel6A.

[0069] The absolute values of K_G for the Family 6 cellulase variants and the relative reductions in glucose inhibition over the parental Family 6 cellulase are shown in Table 3, below:

Table 3: Reduced Glucose Inhibition of Family 6 cellulase variants

Amino Acid Substitution	SEQ ID No:	K_G (g/L)	Relative K_G
None (TrCel6A-S413P)	156	9.2	1.00
Y103A	37	25.1	2.73
Y103H	38	21.0	2.28
Y103K	39	39.1	4.26
Y103L	40	32.0	3.48
Y103M	41	26.0	2.83
Y103P	42	26.8	2.91
Y103R	43	24.2	2.63
Y103V	44	26.2	2.85
M134I	157	31.7	3.45
M134Q	158	23.7	2.58
M134T	159	31.2	3.39
M134V	160	24.6	2.67
M134Y	161	23.4	2.54
L136I	45	39.5	4.29
L136V	46	24.9	2.71
S186K	47	19.9	2.16
S186Y	48	23.8	2.59
G365D	49	17.6	1.91
G365E	50	21.1	2.29
G365Q	51	38.7	4.20
G365S	52	29.8	3.23
R410A	53	25.0	2.72
R410F	54	37.0	4.02
R410L	55	49.5	5.38
R410Q	56	37.8	4.11
R410S	57	24.5	2.66
M134V-L136I	62	48.6	5.28
L136I-L215I	63	37.5	4.08

M134V-L136I-L215I	71	61.9	6.73
<i>H. insolens</i> HiCel6A wildtype	23	31.9	1.00
Y107K	78	45.5	1.43
Y107L	79	61.0	1.91
Q139T	80	51.9	1.62
L141V	81	54.6	1.71
A194Y	82	63.2	1.98
<i>P. chrysosporium</i> wildtype	30	20.3	1.00
Y98K	83	65.5	3.22
Y98L	84	31.9	1.57
L131I	85	32.9	1.62
L131V	86	33.3	1.64
S182K	87	39.7	1.95
S182Y	88	46.4	2.28
G359Q	89	35.2	1.73
R404Q	90	47.4	2.33

[0070] By “relative K_G ” it is mean the ratio of the absolute K_G of the Family 6 cellulase variant by the absolute K_G of the parental Family 6 cellulase from which the Family 6 cellulase variant is derived. As such, the Family 6 cellulase variant is said to exhibit “-fold” less inhibition by glucose than the parental Family 6 cellulase from which it is derived. For example, as shown in Table 3 the Y98K variant of the *P. chrysosporium* Family 6 cellulase exhibits a relative K_G of 3.22 and is therefore 3.22-fold less inhibited by glucose than the parental, wild-type *P. chrysosporium* Family 6 cellulase from which the Y98K variant is derived.

EXAMPLES

[0071] The present invention will be further illustrated in the following examples. However, it is to be understood that these examples are for illustrative purposes only and should not be used to limit the scope of the present invention in any manner.

[0072] Example 1 describes the strains and vectors used in the following examples. Examples 2-5 describe the cloning of the TrCel6A-S413P gene and transformation in yeast, the making of error prone-PCR and site-saturation mutagenesis libraries of Cel6A, and the generation of combinatorial mutants. Examples 6 and 7 describe the expression of TrCel6A-S413P variants from microculture and the high-throughput screening to identify Family 6 cellulase variants with reduced inhibition by glucose. Examples 8 and 9 describe the expression and characterization of isolated and parental

Family 6 cellulases with reduced inhibition by glucose. Examples 10-14 describe the construction, expression, purification, and characterization of PcCel6A-His6 and HiCel6A-His6 and their variants. Example 15 describes an immunoassay for measuring the concentration of TrCel6A variants.

Example 1: Strains and Vectors

[0073] *Saccharomyces cerevisiae* strain YDR483W BY4742 [14317] (MAT α his3 Δ 1 leu2 Δ 0 lys2 Δ 0 ura3 Δ 0 Δ kre2) was obtained from ATCC (#4014317). *Humicola insolens* and *Phanerochaete chrysosporium* strains were obtained from ATCC[®] (#22082[™] and #201542[™], respectively). *Escherichia coli* strain DH5 α (F⁻ ϕ 80lacZ Δ M15 Δ (lacZYA-argF)U169 recA1 endA1 hsdR17(r_k⁻, m_k⁺) phoA supE44 thi-1 gyrA96 relA1 λ) was obtained from Invitrogen. The YEp352/PGK91-1 vector was obtained from the National Institute of Health. The YEpFLAG Δ Kpn10-S413P vector is described in U.S. Patent Application 60/841,507. The YEpFLAG-1 vector was obtained from Sigma as a part of the Amino-Terminal Yeast FLAG Expression Kit. The pGEM T-easy vector was obtained from Promega.

Example 2: Cloning of the TrCel6A-S413P gene into the YEp352/PGK91-1 and transformation in yeast

[0074] In order to facilitate cloning using *NheI* and *KpnI* restriction enzymes, the unique *NheI* site at position 1936 of the YEp352/PGK91-1 vector was blunted using the DNA Polymerase I large (Klenow) fragment to generate YEp352/PGK91-1 Δ *NheI*. The TrCel6A-S413P gene was amplified by PCR from YEpFLAG Δ Kpn10-S413P vector (U.S. Application No. 60/841,507) using primers 5'NheCel6A and 3'BglKpnCel6A. In parallel, the yeast α -factor leader sequence was amplified by PCR from the YEpFLAG-1 vector (Sigma) using primers (5'BglAlphaSS and 3'NheAlphaSS) to introduce restriction sites for *BglII* at the 5' end and *NheI* at 3' end of the amplicon.

[0075] The yeast α -factor leader sequence was isolated by *BglII/NheI* digestion and a three piece ligation performed with the TrCel6A-S413P gene (isolated by *NheI/BglII* digestion) and YEp352/PGK91-1 Δ *NheI* vector (isolated by *BglII* digestion). The resulting vector YEp352/PGK91-1 Δ *NheI*- α_{ss} -TrCel6A-S413P (Figure 2) was transformed in yeast strain BY4742 using the procedure described by Gietz, R. D. and Woods, R. A. (2002). Primer sequences are listed below:

5'BglAlphaSS: 5'ACC AAA AGA TCT ATG AGA TTT CCT TCA ATT (SEQ ID NO: 91)

3'NheAlphaSS: 5'TGA GCA GCT AGC CCT TTT ATC CAA AGA TAC (SEQ ID NO: 92)

5'NheCel6A: 5'AAA AGG GCT AGC TGC TCA AGC GTC TGG GGC (SEQ ID NO: 93)

3'BglKpnCel6A: 5'GAG CTC AGA TCT GGT ACC TTA CAG GAA CGA TGG GTT (SEQ ID NO: 94)

Example 3: Making Error Prone-PCR Libraries

[0076] Random mutagenesis libraries were generated using two methods: a Mutazyme[®] II DNA polymerase method and a Mn²⁺/biased dNTP mix method. For the Mutazyme[®] II DNA polymerase method, a series of four independent PCR were performed using 10, 20, 30, and 40 ng of YEp352/PGK91-1Δ*NheI*-α_{ss}-*TrCel6A*-S413P vector and the Mutazyme[®] II DNA polymerase with primers YalphaN21 and 3'PGK-term. The amplification was done for 25 cycles. The four PCR products were pooled and diluted to 10 ng/μL. A second PCR mutagenesis step was performed using 30 ng of pooled PCR product with Mutazyme[®] II DNA polymerase using the same primers for 30 amplification cycles. The YEp352/PGK91-1Δ*NheI*-α_{ss}-*TrCel6A*-S413P vector was digested with *NheI* and *KpnI* and the empty vector fragment was isolated. This linear fragment and the final amplicon were transformed simultaneously and cloned by *in vivo* recombination into yeast strain BY4742 (Butler *et al.*, 2003).

[0077] For the Mn²⁺/biased dNTP mix method, a PCR was performed using 25 ng YEp352/PGK91-1Δ*NheI*-α_{ss}-*TrCel6A*-S413P vector, 200 μM dATP, 200 μM dCTP, 240 μM dGTP, 200 μM dTTP, and 640 μM Mn²⁺ with Taq DNA polymerase (Sigma) with primers YalphaN21 and 3'PGK-term for 30 amplification cycles. The final amplicon was cloned into YEp352/PGK91-1Δ*NheI*-α_{ss}-*TrCel6A*-S413P vector as described above.

YalphaN21: 5'AGC ACA AAT AAC GGG TTA TTG (SEQ ID NO: 95)

3'PGK-term: 5'GCA ACA CCT GGC AAT TCC TTA CC(SEQ ID NO: 96)

Example 4: Making Site-Saturation Mutagenesis Libraries

[0078] Seven amino acid positions in *TrCel6A*-S413P (M134, L136, L215, Y103, S186, G365, and R410), identified during high-throughput screening (Example 7), were chosen for site-saturation mutagenesis in order to find an amino acid which further improves the tolerance to glucose. Site-saturation mutagenesis was performed by megaprimer PCR (two-step PCR reaction) using NNS primers (listed below), the YEp352/PGK91-1Δ*NheI*-α_{ss}-*TrCel6A*-S413P vector as template, and the Platinum[®] Taq DNA Polymerase High Fidelity (Invitrogen). The first-step PCR was done using the

NNS primer and the complementary external primer (YalphaN21 or 3'PGK-term). The purified amplicon served as a megaprimer for the second-step PCR and the other complementary external primers were used to amplify the complete mutated gene. This final amplicon was then cloned into YEp352/PGK91-1 Δ *NheI*- α_{ss} -*TrCel6A*-S413P vector as described in Example 3.

3'M134X: 5'GAG TAT CTA GCC ASN NAA AAG AGG GAA C (SEQ ID NO: 97)
 3'L136X: 5'CTT GTC AAG AGT ATC SNN CCA CAT AAA AG (SEQ ID NO: 98)
 3'L215X: 5'GCT CAA TAA CCA GSN NGG TCC GGA TAT C (SEQ ID NO: 99)
 3'Y103X: 5'CTT CAG AGG CGT ASN NTG CAT TGG CCC (SEQ ID NO: 100)
 3'S186X: 5'CAC CAT CGG CAA TSN NGT ATT CGC CAT TC (SEQ ID NO: 101)
 5'G365X: 5'CAG CAA CAG TGG NNS GAC TGG TGC AAT G (SEQ ID NO: 102)
 5'R410X: 5'GAC AGC AGT GCG CCA NNS TTT GAC CCC CAC TGT GC (SEQ ID NO: 103)

Example 5: Making Combinatorial Mutants.

[0079] Based on the *TrCel6A*-S413P positive variants identified in Example 7 for tolerance to glucose, a set of three multiple mutants and one single mutant were designed over the parent *TrCel6A*-S413P (M134V-L136I, L136I-L215I, M134V-L136I-L215I, and M134V). For the M134V-L136I mutant, the YEp352/PGK91-1 Δ *NheI*- α_{ss} -*TrCel6A*-S413P vector served as a template for a megaprimer PCR two-step reaction using the mutagenic primer 5'M134V-L136I, the external primer 3'PGK-term, and the Platinum[®] Taq DNA Polymerase High Fidelity (Invitrogen). The purified amplicon served as a megaprimer for the second-step PCR and the YalphaN21 external primers are used to amplified the complete mutated gene. This final amplicon and the purified YEp352/PGK91-1 Δ *NheI*- α_{ss} -*TrCel6A*-S413P vector digested with *XhoI* and *NruI* were transformed simultaneously and cloned by *in vivo* recombination into yeast strain BY4742. For the L136I-L215I mutant, the *NheI*-*EcoRV* purified fragment and the purified YEp352/PGK91-1 Δ *NheI*- α_{ss} -*TrCel6A*-S413P-L215I vector (isolated in Example 7) digested with *XhoI* and *NruI* were transformed simultaneously and cloned by *in vivo* recombination into yeast strain BY4742. For the M134V-L136I-L215I mutant, the purified megaprimer described above and the purified YEp352/PGK91-1 Δ *NheI*- α_{ss} -*TrCel6A*-S413P-L215I vector (isolated in Example 7) digested with *XhoI* and *NruI* were transformed simultaneously and cloned by *in vivo* recombination into yeast strain BY4742. Finally, for the M134V mutant, the YEp352/PGK91-1 Δ *NheI*- α_{ss} -*TrCel6A*-S413P vector served as a

template for a megaprimer PCR two-step reaction using the mutagenic primer 5'M134V, the external primer 3'PGK-term, and the Platinum[®] Taq DNA Polymerase High Fidelity (Invitrogen). The purified amplicon served as a megaprimer for the second-step PCR and the YalphaN21 external primers are used to amplified the complete mutated gene. This final amplicon and the purified YEp352/PGK91-1Δ*NheI*-α_{ss}-TrCel6A-S413P vector digested with *XhoI* and *NruI* were transformed simultaneously and cloned by *in vivo* recombination into yeast strain BY4742.

5'M134V-L136I: 5'GTT CCC TCT TTT GTG TGG ATA GAT ACT CTT GAC (SEQ ID NO: 104)

5'M134V: 5'GTT CCC TCT TTT GTG TGG CTA GAT ACT (SEQ ID NO: 105)

Example 6: Expression and Isolation of TrCel6A-S413P and its Variants From Microplate Cultures

[0080] This example describes the selection and expression of TrCel6A-S413P variants from *Saccharomyces cerevisiae* for use in a high-throughput screening assay (Example 7).

[0081] *Saccharomyces cerevisiae* transformants, from Example 2-5, were grown on plates containing synthetic complete medium (SC: 2% agar w/v, 0.17% yeast nitrogen base w/v, 0.078% - Ura drop-out supplement w/v, 2% glucose w/v, 2% casamino acids w/v, 0.5% ammonium sulfate w/v, pH 5.5) and 0.12% Azo-barley-β-glucan (Megazyme) for 4 days at 30°C.

[0082] Colonies showing visible clearing halos, after an overnight incubation at 45°C, were selected for liquid media pre-cultures by toothpick inoculation of 150 μL synthetic complete media (SC: 0.17% yeast nitrogen base w/v, 0.078% -Ura drop-out supplement w/v, 2% glucose w/v, 2% casamino acids w/v, 0.5% ammonium sulfate w/v, pH 5.5) in 96-well microplates. Pre-cultures were grown overnight (16 - 18 hr) at 30°C and 300 rpm to stationary phase. For expression culture inoculation, 25 μL of pre-culture was used to inoculate 1 mL of SC media in deepwell microplates containing one glass bead. Expression cultures were grown for 3 days at 30°C and 250 rpm with humidity control. Plates were centrifuged at 3000 rpm for 5 minutes to pellet cells and supernatant was aspirated for screening assays (Example 7). To the remaining pre-culture, stocks were prepared by the addition of glycerol to a final concentration of 15% and stored at -80°C.

Example 7: Screening of *Trichoderma reesei* Cel6A Gene Libraries for Family 6 cellulase variant with Tolerance to Glucose

[0083] This example describes the screening of *Trichoderma reesei* Cel6A-S413P variants with reduced inhibition by glucose by comparison to the parent TrCel6A-S413P that had been cloned into *Saccharomyces cerevisiae*.

[0084] TrCel6-S413P variants from yeast microcultures as described in Example 6 were tested in a 0.25 mL citrate buffered (pH 5) cellulose hydrolysis assay using a 96-well microplate format. An aliquot of supernatant from each variant was added to a first well containing 30 g/L glucose and to a second well absent of glucose and incubated with cellulose, at a concentration 0.067% w/v, for 19 hours at 50°C. Yeast supernatants were complemented with *Trichoderma reesei* Cel7B and Cel5A (40 mg protein/g cellulose) and 125 IU/g cellulose *A. niger* beta-glucosidase. Contained in each 96-well microplate were six parent TrCel6A-S413P controls used for comparison. Cellulase activity was measured by turbidometry. A $\pm$ glucose activity ratio was calculated for all TrCel6A variants and the parental TrCel6A-S413P by dividing the cellulase activity in the presence of glucose by the cellulase activity in the absence of glucose. The $\pm$ glucose activity ratio for each TrCel6A-S413P variant was compared to the average of six parental TrCel6A-S413P controls on a particular microplate and positives were selected at the 95% confidence level using a t-test. All positive variants were produced again in microculture and re-screened to reduce the number of false positives. Table 4 summarizes the screening results obtained for the EP-PCR library (Example 3) and the seven SSM libraries (Example 4).

Table 4: Screening Results of EP-PCR and SSM Libraries

Library	# of Variants Screened	# of Positives
EP-PCR	4800	43
SSM - M134	100	11
SSM - L136	40	2
SSM - L215	68	0
SSM - Y103	110	16
SSM - S186	83	5
SSM - G365	104	8

SSM – R410	142	7
------------	-----	---

Example 8: Expression and Concentration of TrCel6A-S413P and its Variants From Large Scale Cultures

[0085] Two 500 mL volumes of sterile SC*-Ura media (0.77 g/L –Ura drop out supplement, 1.7 g/L yeast nitrogen base, 5 g/L (NH₄)₂SO₄, 20 g/L casamino acids, 20 g/L glucose) were inoculated with 10 mL of overnight cultures of transformed *Saccharomyces cerevisiae* grown from cells freshly picked from an agar plate. The cultures were then incubated for 96 hours at 30 °C with shaking at 200 rpm.

[0086] After incubation, each pair of 500 mL yeast cultures was pooled, centrifuged for 10 minutes at 9000 rpm and the pellet (containing yeast cells) discarded. The supernatant pH was adjusted to 5.0 and then allowed to cool to 4°C for an hour. Subsequent to cooling, 625 g (NH₄)₂SO₄ was added to bring the yeast supernatant to 93% saturation. Precipitation was allowed to occur over a period of 16 hours at 4°C with constant stirring. The next day the precipitate was centrifuged for 15 minutes at 9000 rpm and the supernatant discarded.

[0087] The pellet was resuspended with pipetting in a total volume of 10 mL of 50 mM sodium phosphate, pH 7.0. Once the pellet was resuspended, the solution was mixed with gentle inversion for 30 minutes. The solution was then centrifuged at 3000 rpm for 3 minutes to pellet any insoluble material. The supernatant was removed carefully with a pipette to prevent disruption of the pellet and retained. The concentration of TrCel6A-S413P in this supernatant was determined by ELISA using TrCel6A-specific antibodies with a standard curve of purified TrCel6A-S413P as described in Example 15, below. Purity of the samples was verified by SDS-PAGE analysis such as that shown for TrCel6A-S413P in Figure 16.

Example 9: Enzymatic Characterization of Family 6 cellulase variants

[0088] Acid swollen cellulose (ASC) was produced from Sigmacell50 using procedures known to those skilled in the art. The ASC was slurried in 150 mM citrate, pH 5.0, to a final concentration of 1.8 g cellulose/L and degassed under vacuum for 5 minutes with constant stirring.

[0089] Activity of the TrCel6A-S413P was monitored as a decrease in the absorbance of the slurry, diluted to 0.6 g cellulose/L either by water or a glucose solution, at 600 nm after the addition of

cellulase. The sample was maintained at 50°C and was stirred constantly during data collection. A twin-beam absorbance apparatus was used to subtract the absorbance of a no-enzyme control slurry in real time for each sample. Immediately after initiation of data acquisition, 50 mg/g cellulose of *Trichoderma* cellulase lacking Cel7A and Cel6A was added to all cuvettes. In the absence of cellobiohydrolases there is no significant change in the size of substrate particles and therefore no difference in the apparent absorbance due to an approximately constant size of the scatterers. During this time, the cellobiohydrolase-free enzyme mixture creates new chain ends, which are sites of enzymatic action by cellobiohydrolases. Five minutes after starting data acquisition, a parental Family 6 cellulase or a Family 6 cellulase variant is added to a total dose of 50 mg/g. Cellobiohydrolases, such as a parental Family 6 cellulase or a Family 6 cellulase variant, effect a decrease in the size of the substrate particles *via* hydrolysis and the initial slope of the absorbance decrease was recorded as a measure of activity. As only the linear portion of the activity trace is used and only cellobiohydrolases, such as a parental Family 6 cellulase or a Family 6 cellulase variant, can reduce the apparent absorbance of the substrate, the assay is specific for the activity of the Cel6A alone and is not influenced by any potential inhibition of the non-cellobiohydrolase enzymes. Moreover, any influence of these enzymes would be the same when comparing a parental Family 6 cellulase to a Family 6 cellulase variant with decreased inhibition by glucose. Data collected during a typical experiment with TrCel6A-S413P are depicted in Figure 4.

[0090] Activity data were collected in the absence of glucose and the presence of 10, 20, 30, 40 and 50 g/L glucose. Five replicate data sets were collected for each of these conditions. Data were plotted as absorbance slope vs. glucose concentration and fit with a model of simple linear inhibition using the Solver function in Microsoft Excel. The best fit values of K_G and 95% confidence intervals were calculated using standard statistical methods. The mean value of the slope measured in the absence of glucose was taken as a measure of specific activity. A type 2, two-tailed t-test was used to compare the K_G and specific activity of each TrCel6A-S413P variant tested to the parameter values of the parent enzyme.

[0091] As shown by the results in Table 5, below, and in Figures 5-15, all of the TrCel6A-S413P variants show at least 1.8-fold and as much as 6.73-fold reductions in glucose inhibition over the parental TrCel6A.

Table 5: Glucose inhibition constants for Family 6 cellulase variants derived from TrCel6A-S413P

Amino Acid Substitution	K _G (g/L)	95% Confidence Intervals	Relative K _G	P Value
None (TrCel6A-S413P)	9.2	7.7 - 11.0	1.00	1.00
Y103A	25.1	21.2 - 30.2	2.73	1.5E-08
Y103H	21.0	18.2 - 24.6	2.28	2.2E-07
Y103K	39.1	35.3 - 45.2	4.26	3.2E-11
Y103L	32.0	28.1 - 36.9	3.48	2.6E-10
Y103M	26.0	18.8 - 36.4	2.83	1.4E-07
Y103P	26.8	22.9 - 32.2	2.91	4.4E-09
Y103R	24.2	19.8 - 29.9	2.63	4.2E-08
Y103V	26.2	21.1 - 32.8	2.85	1.5E-08
M134I	31.7	23.0 - 44.3	3.45	1.7E-04
M134Q	23.7	20.6 - 27.7	2.58	2.3E-08
M134T	31.2	25.1 - 39.4	3.39	1.1E-08
M134V	24.6	20.9 - 29.4	2.67	2.0E-08
M134Y	23.4	18.1 - 30.1	2.54	1.1E-03
L136I	39.5	35.5 - 44.7	4.29	2.8E-11
L136V	24.9	21.7 - 28.5	2.71	8.9E-09
S186K	19.9	14.2 - 27.5	2.16	2.6E-05
S186Y	23.8	17.5 - 33.3	2.59	2.2E-03
G365D	17.6	12.4 - 24.7	1.91	3.5E-04
G365E	21.1	16.8 - 27.3	2.29	3.8E-03
G365Q	38.7	33.1 - 46.2	4.20	4.1E-11
G365S	29.8	23.2 - 38.6	3.23	3.7E-09
R410A	25.0	17.0 - 37.4	2.72	1.0E-06
R410F	37.0	26.9 - 51.7	4.02	5.1E-10
R410L	49.5	27.0 - 97.9	5.38	8.1E-10
R410Q	37.8	28.6 - 51.2	4.11	2.0E-10
R410S	24.5	15.8 - 38.0	2.66	4.1E-06
M134V-L136I	48.6	42.3 - 57.0	5.28	5.8E-12
L136I-L215I	37.5	32.7 - 43.6	4.08	4.7E-11
M134V-L136I-L215I	61.9	50.9 - 76.3	6.73	1.1E-12

Example 10: Generation of HiCel6A and PcCel6A and their cloning into the Yep352/PGK91-1- α_{ss} -6His vector

Construction of the YEp352/PGK91-1- α_{ss} -6His vector

[0092] A DNA adapter containing *SpeI*, *NheI*, *KpnI*, and *EcoRI* restriction sites was prepared by annealing primers AT044 and AT045 together. The adapter contains sequences coding for six histidine residues downstream of the *SpeI* site and upstream of the *NheI* site. The adapter was inserted into a YEp based-plasmid (YEp352/PGK91-1 α_{ss}) containing the *pgk1* promoter, alpha mating factor secretion signal, and *pgk1* terminator sequences to make plasmid YEp352/PGK91-1/ α_{ss} 6HNKE. Specifically, the linker was inserted as a *NheI/EcoRI* fragment into the *NheI* and

EcoRI sites located downstream of the alpha mating factor secretion signal and upstream of the *pgk1* terminator. Primer sequences are shown below:

AT044: 5' CTA GTC ATC ACC ATC ACC ATC ACG CTA GCT GAT CAC TGA GGT ACC G
(SEQ ID NO: 106)

AT045: 5' AAT TCG GTA CCT CAG TGA TCA GCT AGC GTG ATG GTG ATG GTG ATG A
(SEQ ID NO: 107)

Generation of YEp352/PGK91-1- α_{ss} -6H-HiCel6A

[0093] Lyophilized *H. insolens* was resuspended in 300 μ L sterile H₂O and 50 μ L was spreaded onto Emerson YPSS pH 7 agar plate (0.4% Yeast extract, 0.1% K₂HPO₄, 0.05% MgSO₄•7H₂O, 1.5% Glucose, 1.5% Agar). Fungus was incubated for 6 days at 45°C then spores were inoculated in Novo media (as per Barbesgaard USP # 4,435,307): Incubation for 48 hours at 37°C in 100 mL growth phase media (2.4% CSL, 2.4% Glucose, 0.5% Soy oil, pH adjusted to 5.5, 0.5% CaCO₃), then 6 mL of pre-culture was transferred into 100 mL production phase media (0.25% NH₄NO₃, 0.56% KH₂PO₄, 0.44% K₂HPO₄, 0.075% MgSO₄•7H₂O, 2% Sigmacell, pH adjusted to 7, 0.25% CaCO₃) and culture was incubated for up to 4 days prior to biomass harvest. Then, 50 mg of biomass was used to isolate total RNA with the Absolutely RNA[®] Miniprep Kit (Stratagene) according to the manufacturer procedure. Total cDNA was generated from the total RNA using the SuperScript[™]II Reverse Transcriptase (Invitrogen) according to the manufacturer procedure. Gene encoding for HiCel6A was amplified from the cDNA using the following primers:

5'HiCel6A-cDNA 5'CTA TTG CTA GCT GTG CCC CGA CTT GGG GCC AGT GC (SEQ ID NO: 108)

3'HiCel6A-cDNA 5'CTA TTG AAT TCG GTA CCT CAG AAC GGC GGA TTG GCA TTA CGA AG (SEQ ID NO: 109)

[0094] The PCR amplicon was cloned into the pGEM[®]-T Easy vector by TA-cloning according to the manufacturer's recommendations. Plasmid pGEM-HiCel6A was digested with *NheI* and *EcoRI* to release the HiCel6A gene. This fragment was purified and ligated into the *NheI* and *EcoRI* sites of YEp352/PGK91-1/ α_{ss} 6HNKE to obtain YEp352/PGK91-1/ α_{ss} 6H-HiCel6A.

Generation of YEp352/PGK91-1- α_{ss} -6H-PcCel6A

[0095] Lyophilized *P. chrysosporium* was resuspended in 300 μ L sterile H₂O and 50 μ L were spreaded onto PDA plates. Plates were incubated at 24°C for 4 days. Spores for *P. chrysosporium* were inoculated on a cellophane circle on top of a PDA plate and biomass was harvested after 4-6 days at 24°C. Then, 50 mg of biomass was used to isolate total RNA with the Absolutely RNA[®] Miniprep Kit (Stratagene) according to the manufacturer procedure. Total cDNA was generated from the total RNA using the SuperScript[™]II Reverse Transcriptase (Invitrogen) according to the manufacturer procedure. Gene encoding for PcCel6A was amplified from the cDNA using the following primers (which introduced an N-terminal *NheI* site and C-terminal *KpnI* and *EcoRI* sites):

5'PcCel6A-cDNA 5'CTA TTG CTA GCT CGG AGT GGG GAC AGT GCG GTG GC
(SEQ ID NO: 110)

3'PcCel6A-cDNA 5'CTA TTG AAT TCG GTA CCC TAC AGC GGC GGG TTG GCA GCA GAA AC (SEQ ID NO: 111)

[0096] The PCR amplicon was cloned into the pGEM[®]-T Easy vector by TA-cloning according to the manufacturer's recommendations. Plasmid pGEM-PcCel6A was digested with *NheI* and *EcoRI* to release the PcCel6A gene. This fragment was purified and ligated into the *NheI* and *EcoRI* sites of YEp352/PGK91-1/ α_{ss} 6HNKE to obtain YEp352/PGK91-1/ α_{ss} 6H-PcCel6A.

Example 11: Mutagenesis of Vectors with PcCel6A-His6 and HiCel6A-His6 and their Variants

[0097] The HiCel6A and PcCel6A variants were constructed using a two-step PCR method involving megaprimer synthesis followed by PCR-mediated overlap extension (Vallejo *et al.*, 1994). All PCR reactions were carried out using the High Fidelity iProof *Taq* Polymerase (BioRad). The plasmids YEp352/PGK91-1- α_{ss} -6H-HiCel6A and YEp352/PGK91-1- α_{ss} -6H-PcCel6A served as the templates for both *Humicola insolens* and *Phanerochaete chrysosporium*, respectively. Megaprimers upstream of (and including) the mutagenesis site were amplified using external primer Y α N21 in combination with an internal reverse primer (i.e. DK022 for HiCel6A Y107L; refer to Table 6) specific to a given glucose tolerant variant. Similarly, megaprimers downstream of (and including) the mutagenesis site were amplified using the external primer PGKterm together with an internal forward primer (i.e. DK021 for HiCel6A Y107L; refer to Table 6) unique for a particular glucose tolerant variant. The internal primers were designed to introduce

the desired glucose tolerant mutations into the Cel6A homologues. The megaprimers were purified using the Wizard® SV Gel and PCR Clean-Up System (Promega).

Table 6: List of primers used to introduce glucose tolerant mutations within the *H. insolens* and *P. chrysosporium* Cel6A homologues. Amino acids are italicized in cases where the original residue differs from that found in the homologous *T. reesei* sequence. With respect to the primer sequences, underlined nucleotides effect the desired amino acid substitution. External plasmid primers are included.

Primer name	Homolog	Variant	SEQ ID NO:	Sequence
DK021	HiCel6A	Y107L	112	CTCTGGGCCAACAACCTGTACCGCTCTGAGGTC
DK022	HiCel6A	Y107L	113	GACCTCAGAGCGGTACAGGTTGTTGGCCCAGAG
DK023	HiCel6A	Y107K	114	CTCTGGGCCAACAACAAGTACCGCTCTGAGGTC
DK024	HiCel6A	Y107K	115	GACCTCAGAGCGGTACTTGTGTTGGCCCAGAG
DK025	HiCel6A	Q139I	116	GAGGTCCCGAGCTTCACTGGCTCGACCGCAAC
DK026	HiCel6A	Q139I	117	GTTGCGGTCGAGCCAGATGAAGCTCGGGACCTC
DK027	HiCel6A	Q139T	118	GAGGTCCCGAGCTTCACTGGCTCGACCGCAAC
DK028	HiCel6A	Q139T	119	GTTGCGGTCGAGCCAGGTGAAGCTCGGGACCTC
DK029	HiCel6A	L141I	120	CCGAGCTTCCAGTGGATCGACCGCAACGTCACG
DK030	HiCel6A	L141I	121	CGTGACGTTGCGGTCGATCCACTGGAAGCTCGG
DK031	HiCel6A	L141V	122	CCGAGCTTCCAGTGGGTGACCGCAACGTCACG
DK032	HiCel6A	L141V	123	CGTGACGTTGCGGTCGACCCACTGGAAGCTCGG
DK033	HiCel6A	A194Y	124	TCGAACGGCGAGTGGTACATCGCCAACAACGGC
DK034	HiCel6A	A194Y	125	GCCGTTGTTGGCGATGTACCACTCGCCGTTCTGA
DK035	HiCel6A	A194K	126	TCGAACGGCGAGTGGAAAGATCGCCAACAACGGC
DK036	HiCel6A	A194K	127	GCCGTTGTTGGCGATCTTCACTCGCCGTTCTGA
DK037	HiCel6A	G372Q	128	GGCCAGAAGGAATGGCAGCACTGGTGCAATGCC
DK038	HiCel6A	G372Q	129	GGCATTGCACCAAGTGCTGCCATTCTTCTGGCC
DK039	HiCel6A	R417Q	130	GACACGACCGCTGCCAGTACGACTACCACTGC
DK040	HiCel6A	R417Q	131	GCAGTGGTAGTCGTAAGTGGCAGCGGTCGTGTC
DK041	HiCel6A	R417F	132	GACACGACCGCTGCCTTCTACGACTACCACTGC
DK042	HiCel6A	R417F	133	GCAGTGGTAGTCGTAAGTGGCAGCGGTCGTGTC
DK043	PcCel6A	Y98L	134	GATCTTCCTCAGCCCTCTGTACGCGAACGAGGTC
DK044	PcCel6A	Y98L	135	GACCTCGTTCGCGTACAGAGGGCTGAGGAAGATC
DK045	PcCel6A	Y98K	136	GATCTTCCTCAGCCCTAAGTACGCGAACGAGGTC
DK046	PcCel6A	Y98K	137	GACCTCGTTCGCGTACTTAGGGCTGAGGAAGATC
DK047	PcCel6A	T129I	138	GCAAATATCCCCACTTTCACTGGCTGGACTCTGTC
DK048	PcCel6A	T129I	139	GACAGAGTCCAGCCAGATGAAAGTGGGGATATTGTC
DK049	PcCel6A	T129Q	140	GCAAATATCCCCACTTTCCAGTGGCTGGACTCTGTC
DK050	PcCel6A	T129Q	141	GACAGAGTCCAGCCACTGGAAAGTGGGGATATTGTC
DK051	PcCel6A	L131I	142	CCCACCTTTCACGTGGATCGACTCTGTGCGGAAG
DK052	PcCel6A	L131I	143	CTTCGCGACAGAGTCCGATCCACGTGAAAGTGGG
DK053	PcCel6A	L131V	144	CCCACCTTTCACGTGGGTGACTCTGTGCGGAAG
DK054	PcCel6A	L131V	145	CTTCGCGACAGAGTCCGACCCACGTGAAAGTGGG
DK055	PcCel6A	S182Y	146	TCCAACGGAGAGTTCTACATTGCCAACAACGGA
DK056	PcCel6A	S182Y	147	TCCGTTGTTGGCAATGTAGAAGTCTCCGTTGGA
DK057	PcCel6A	S182K	148	TCCAACGGAGAGTTCAAGATTGCCAACAACGGA
DK058	PcCel6A	S182K	149	TCCGTTGTTGGCAATCTTGAAGTCTCCGTTGGA
DK059	PcCel6A	G359Q	150	CATCCGCCAACAGTGGCAGGACTGGTGCAACATC
DK060	PcCel6A	G359Q	151	GATGTTGCACCAGTCCTGCCACTGTTGGCGGATG

DK061	PcCel6A	R404Q	152	CCAACAGCTCCTCGCCCCAGTACGACTCGACTTGTTTC
DK062	PcCel6A	R404Q	153	GAACAAGTCGAGTCGTACTGGGGCGAGGAGCTGTTGG
DK063	PcCel6A	R404F	154	CCAACAGCTCCTCGCCCTTCTACGACTCGACTTGTTTC
DK064	PcCel6A	R404F	155	GAACAAGTCGAGTCGTAGAAGGGCGAGGAGCTGTTGG
YαN21	PGK plasmid	-	95	AGCACAAATAACGGGTTATTG
PGKterm	PGK plasmid	-	96	GCAACACCTGGCAATTCCTTACC

[0098] During the second round of PCR, both megaprimers of a desired construct were allowed to anneal and extend for 10 cycles to generate the final template. The external primers YαN21 and PGKterm were then added for another 25 cycles to amplify the final product, which was subsequently purified using the Wizard® SV Gel and PCR Clean-Up System. Both the purified PCR product and the linearized vector YEp352/PGK91-1α_{ss}-6HNKE (digested with *NheI* + *KpnI*) were transformed and cloned via *in vivo* recombination within the BY4742 yeast strain using the procedure described by Gietz and Woods (2002). For each construct, the vector was isolated from the transformed yeast using a method modified from Hoffman and Winston (Hoffman and Winston, 1987) and transformed in *E. coli* DH5α chemically-competent cells. Plasmids were isolated from the *E. coli* cells using the Wizard® Plus SV Minipreps DNA Purification System (Promega). The integrity of the cloned region of all the variants was confirmed by DNA sequence analysis.

Example 12: Expression and Concentration of PcCel6A-His6 and HiCel6A-His6 and their Variants From Large Scale Cultures

[0099] Two 500 mL volumes of sterile SC*-Ura media (0.77 g/L -Ura drop out supplement, 1.7 g/L yeast nitrogen base, 5 g/L (NH₄)₂SO₄, 20 g/L casamino acids, 20 g/L glucose) were inoculated with 10 mL of overnight cultures of transformed *Saccharomyces cerevisiae* grown from cells freshly picked from an agar plate. The cultures were then incubated for 96 hours at 30 °C with shaking at 200 rpm.

[00100] After incubation, each pair of 500 mL yeast cultures was pooled, centrifuged for 10 minutes at 9000 rpm and the pellet (containing yeast cells) discarded. The supernatant pH was adjusted to 5.0 and then allowed to cool to 4°C for an hour. BSA (0.1 g) was added to help co-precipitate Cel6A. Subsequent to cooling, 559 g (NH₄)₂SO₄ was added to bring the yeast supernatant to 85% saturation. Precipitation was allowed to occur over a period of 16 hours at 4°C with constant stirring. The next day the precipitate was centrifuged for 15 minutes at 9000 rpm and the supernatant discarded.

[00101] The pellet was resuspended with pipetting in a total volume of 50 mL binding buffer (200 mM NaCl, 20 mM sodium phosphate, 30 mM imidazole, pH 7.4). Once the pellet was resuspended, the solution was mixed with gentle inversion for 30 minutes at 4°C. The solution was then filtered with glass fiber filter paper to remove insoluble material before purification as described in Example 13.

Example 13: Purification of PcCel6A-His6 and HiCel6A-His6 and their Variants

[00102] For activity assays (Example 14), the His-tagged parental *Humicola* Cel6A and *Phanerochaete* Cel6A cellulases and their respective variants were purified from culture supernatants using immobilized metal affinity chromatography. Prior to loading proteins onto the His-trap column, the Ni²⁺ resin was equilibrated with binding buffer (200 mM NaCl, 20 mM sodium phosphate, 30 mM imidazole, pH 7.4). The culture supernatant was adjusted to the same salt concentrations and pH as the binding buffer and applied to a 1 mL His-trap column (GE Healthcare) at a flow rate of 0.5 – 1.0 mL/min. The column was then washed with the same binding buffer until OD_{280nm} reached a steady baseline. The bound His-tagged Cel6A was eluted from the column with elution buffer (20 mM sodium phosphate, 500 mM imidazole, pH 7.4).

[00103] The elution buffer was removed from the purified proteins by buffer exchange and 5 mL of the eluted fractions was added to 15 mL of 50 mM citrate pH 5.0 buffer and loaded onto a Centricon Plus-20 (polyethersulfone membrane, nominal molecular weight cut-off of 5 kDa). The column was centrifuged at 1600xg for 10 minutes or until the volume remaining in the column was 2.5 mL. At this point, 17.5 mL of 50 mM citrate buffer pH 5.0 was further added to the column and the centrifugation repeated. The last centrifugation was done in the same fashion until only 1-2 mL remained in the column. The protein concentration of this final supernatant determined using a Biorad protein assay using a *Trichoderma reesei* cellulase of known protein concentration. Purity of the parental and variant PcCel6A and HiCel6A was confirmed by SDS-PAGE analysis (Figure 16).

Example 14: Enzymatic Characterization of PcCel6A-His6 and HiCel6A-His6 and their Variants

[00104] Activity assays were performed as described in Example 9 using the parental PcCel6A-His6 and HiCel6A-His6 cellulases and variants derived therefrom that were purified from yeast culture filtrates as described in Example 13, above. As shown by the results in Table 7, below, all

of the HiCel6A and PcCel6A variants show at least 1.43-fold and as much as 3.22-fold less inhibition by glucose over the corresponding parental Family 6 cellulase control.

Table 7: Glucose inhibition constants for parental PcCel6A and HiCel6A cellulases and Family 6 cellulase variants derived from PcCel6A and HiCel6A.

Amino Acid Substitution	Equivalent TrCel6A position	K _G (g/L)	95% Confidence Intervals	Relative K _G	P Value vs. wildtype
<i>P. chrysosporium</i> wildtype	--	20.3	16.7 - 24.9	1.00	1.00
Y98K	103	65.5	52.9 - 82.7	3.22	2.1E-08
Y98L	103	31.9	26.1 - 39.1	1.57	2.8E-03
L131I	136	32.9	28.7 - 38.0	1.62	7.1E-04
L131V	136	33.3	26.3 - 42.4	1.64	2.1E-03
S182K	186	39.7	32.3 - 49.5	1.95	5.5E-05
S182Y	186	46.4	36.0 - 60.6	2.28	7.4E-06
G359Q	365	35.2	27.0 - 46.2	1.73	1.3E-03
R404Q	410	47.4	31.6 - 74.4	2.33	7.6E-05
<i>H. insolens</i> wildtype	--	31.9	26.4 - 38.8	1.00	1.00
Y107K	103	45.5	35.3 - 59.9	1.43	2.6E-02
Y107L	103	61.0	46.3 - 82.3	1.91	2.1E-04
Q139T	134	51.9	40.6 - 67.8	1.62	2.6E-03
L141V	136	54.6	41.7 - 73.8	1.71	1.5E-03
A194Y	186	63.2	50.2 - 81.7	1.98	5.2E-05

Example 15: Determination of TrCel6A-S413P concentration by ELISA.

[00105] Supernatants and purified standard were diluted in phosphate-buffered saline (PBS), pH 7.2, and incubated overnight at 4°C in microtitre plates (Costar EIA #9018). Following overnight incubation, these plates were washed with PBS containing 0.1% Tween-20 (PBS/Tween) and then incubated in PBS containing 1% bovine serum albumin (PBS/BSA) for 1 hr at room temperature. Blocked microtitre wells were washed with PBS/Tween. Rabbit polyclonal antisera specific for TrCel6A was diluted in PBS/BSA, added to the microtitre plates and incubated for 2 hr at room temperature. Plates were washed and incubated with a goat anti-rabbit antibody coupled to horseradish peroxidase (Sigma #A6154), diluted 1/2000 in PBS/BSA, for 1 hr at room temperature. After washing, 100 µL of tetramethylbenzidine (Sigma #8665) was added to well and incubated for 30 min at room temperature. The absorbance at 360 nm was measured in each well and converted into protein concentration using the TrCel6A standard curve.

PBS contains:

<u>Component</u>	<u>g/L</u>
NaCl	80
KCl	2
Na ₂ HPO ₄	14.4
KH ₂ PO ₄	2.4

REFERENCES

- Barr, B., Hsieh, Y., Ganem, B., and Wilson, D. (1996) Identification of two functionally different classes of exocellulases. *Biochemistry* 35:586.
- Butler, T. and Alcalde, M. (2003) In *Methods in Molecular Biology*, vol. 231: (F. H. Arnold and G. Georgiou, editors), Humana Press Inc. Totowa (New Jersey), pages 17-22.
- Caminal, G., Lopez-Santin, J., and Sola, C. (1985) Kinetic modeling of the enzymatic hydrolysis of pretreated cellulose. *Biotech. Bioeng.* 27:1282.
- Claeysens, M. and Henrissat, B. 1992, *Protein Science* 1: 1293-1297.
- Converse, A., Matsuno, R., Tanaka, M., and Taniguchi, M. (1988) A model of enzyme adsorption and hydrolysis of microcrystalline cellulose with slow deactivation of the adsorbed enzyme. *Biotech. Bioeng.* 32:45.
- Eriksson, T., Karlsson, J., and Tjerneld, F. (2002) A model explaining the declining rate in hydrolysis of lignocellulose substrates with cellobiohydrolase I (Cel7A) and endoglucanase I (Cel7B) of *Trichoderma reesei*. *Appl. Biochem. Biotechnol.* 101:41.
- Farrell, A., Plevin, R., Turner, B., Jones, A., O'Hare, M., and Kammen, D. (2006) Ethanol can contribute to energy and environmental goals. *Science* 311:506.
- Gietz, R.D. and Woods, R.A. (2002) Transformation of yeast by the Liac/ss carrier DNA/PEG method. *In Methods in Enzymology*, 350:87-96.
- Gruno, M., Valjamae, P., Pettersson, G., and Johansson, G. (2004) Inhibition of the *Trichoderma reesei* cellulases by cellobiose is strongly dependent on the nature of the substrate. *Biotech. Bioeng.* 86:503.
- Gusakov, A., and Sinitsyn, A. (1992) A theoretical analysis of cellulase product inhibition: Effect of cellulase binding constant, enzyme/substrate ratio, and β -glucosidase activity on the inhibition pattern. *Biotech. Bioeng.* 40:663.
- Holtzaple, M., Cognata, M., Shu, Y., and Hendrickson, C. (1990) Inhibition of *Trichoderma reesei* cellulase by sugars and solvents. *Biotech. Bioeng.* 36:275.

Hoffman, C. S., and Winston, F. (1987) A ten-minute DNA preparation from yeast efficiently releases autonomous plasmids for transformation of *Escherichia coli*. *Gene*, 57: 267-272.

Lee, Y-H. and Fan, L.T. (1983) Kinetic studies of enzymatic hydrolysis of insoluble cellulose: (II) Analysis of extended hydrolysis times. *Biotech. Bioeng.* 25:939.

5 Meinke, A., et al. 1995. *J. Biol. Chem.* 270:4383-4386.

Nidetzky, B. and Steiner, W. (1993) A new approach for modeling cellulase-cellulose adsorption and the kinetics of the enzymatic hydrolysis of microcrystalline cellulose. *Biotech. Bioeng.* 42:469.

0 Perlack, R., Wright, L., Tuhollow, A., and Graham, R. (2005) Biomass as feedstock for a bioenergy and bioproducts industry: The technical feasibility of a billion-ton annual supply. Oak Ridge National Laboratory.

Steipe, B. (2004) Consensus-based engineering of protein stability: From intrabodies to thermostable enzymes. *Meth. Enz.* 388:176.

5 Teleman, A., Koivula, A., Reinikainen, T., Valkeajärvi, A., Teeri, T.T., Drakenberg, T., and Teleman, O. (1995) Progress-curve analysis shows that glucose inhibits the cellotriose hydrolysis catalysed by cellobiohydrolase II from *Trichoderma reesei*. *Eur. J. Biochem.* 231:250.

Tolan, J., (2002) Iogen's process for producing ethanol from cellulosic biomass, *Clean Tech. and Enviro. Policy*, 3:339.

0 Vallejo, A. N., Pogulis, R. J. and Pease, L. R. (1994) In vitro synthesis of novel genes: mutagenesis and recombination by PCR. *PCR Methods Appl.*, 4:123-130.

Zhang, S., Wolfgang, D., and Wilson, D. (1999) Substrate heterogeneity causes the nonlinear kinetics of insoluble cellulose hydrolysis. *Biotech. Bioeng.* 66:35.

Zhang, S., Irwin, D., and Wilson, D. (2000) Site-directed mutation of noncatalytic residues of *Thermobifida fusca* exocellulase Cel6B. *Eur. J. Biochem.* 267:3101.

5

**THE EMBODIMENTS OF THE INVENTION IN WHICH AN EXCLUSIVE
PROPERTY OR PRIVILEGE IS CLAIMED ARE DEFINED AS FOLLOWS:**

1. A Family 6 cellulase variant comprising one or more than one of the amino acid substitutions selected from the group consisting of:
 - a basic, non-polar or proline residue at position 103 (X103H, K, R, A, V, L, P, M);
 - a valine or isoleucine residue at position 136 (X136V, I);
 - a tyrosine or lysine residues at position 186 (X186Y, K);
 - an acidic, glutamine or serine residue at position 365 (X365D, E, Q, S); and
 - an alanine, phenylalanine, leucine, glutamine or serine residue at position 410 (X410A, F, L, Q, S),said position determined from alignment of a parental Family 6 cellulase with a *Trichoderma reesei* Cel6A amino acid sequence as defined in SEQ ID NO: 1.
2. The Family 6 cellulase variant of claim 1, wherein said Family 6 cellulase variant comprises an amino acid sequence which is at least about 45% to about 100% identical to amino acids 83-447 of SEQ ID NO: 1.
3. The Family 6 cellulase variant of claim 2, wherein said Family 6 cellulase variant comprises an amino acid sequence which is at least about 55% to about 100% identical to amino acids 83-447 of SEQ ID NO: 1.
4. The Family 6 cellulase variant of claim 3, wherein the amino acid sequence of said Family 6 cellulase variant is at least about 63% identical to amino acids 83-447 of SEQ ID NO: 1.
5. The Family 6 cellulase variant of claim 4, wherein the amino acid sequence of said Family 6 cellulase variant is at least about 95% identical to amino acids 83-447 of SEQ ID NO: 1.
6. The Family 6 cellulase variant of claim 1, wherein said basic residue at position 103 is a histidine, arginine or lysine residue.

7. The Family 6 cellulase variant of claim 1, wherein said non-polar residue at position 103 is an alanine, valine, methionine or leucine residue.
8. The Family 6 cellulase variant of claim 1, wherein said acidic residue at position 365 is aspartic acid or glutamic acid.
9. The Family 6 cellulase variant of any one of claims 1 through 8, further comprising an isoleucine valine, threonine, tyrosine or glutamine residue at position 134.
10. The Family 6 cellulase variant of any one of claims 1 through 9, further comprising an isoleucine residue at position 215.
11. The Family 6 cellulase variant of any one of claims 1 through 10, further comprising a proline residue at position 413.
12. The Family 6 cellulase variant of any one of claims 1 through 11, wherein said Family 6 cellulase variant exhibits at least about 1.4- fold less inhibition by glucose than the parental Family 6 cellulase from which it is derived
13. The Family 6 cellulase variant of claim 12, wherein said Family 6 cellulase variant exhibits at least about 1.8- fold less inhibition by glucose than the parental Family 6 cellulase from which it is derived.
14. A genetic construct comprising DNA encoding the Family 6 cellulase variant of any one of claims 1 through 13.
15. The genetic construct of claim 14, further comprising regulatory DNA sequences that direct the expression of secretion of the Family 6 cellulase variant.
16. The genetic construct of claim 15, in which the regulatory DNA sequences are derived from genes that are highly expressed and secreted in a host organism used to produce the Family 6 cellulase variant.
17. A genetically modified microbe comprising the genetic construct of any one of claims 14 through 16.

18. The genetically modified microbe of claim 17, wherein said microbe is a species of yeast or filamentous fungus.
19. The genetically modified microbe of claim 18, wherein said microbe is a species of *Saccharomyces*, *Pichia*, *Trichoderma*, *Hypocrea*, *Aspergillus*, *Fusarium*, *Humicola*, or *Neurospora*.
20. The genetically modified microbe of claim 19, wherein said microbe is *Saccharomyces cerevisiae* or *Trichoderma reesei*.
21. A process for producing the Family 6 cellulase variant of any one of claims 1 through 13 comprising, growing the genetically modified microbe of claim 17 under conditions that induce the expression and secretion of the Family 6 cellulase variant and recovering the Family 6 cellulase variant from the culture medium.
22. A process for the hydrolyzing a cellulosic substrate to glucose comprising, treating the cellulosic substrate with a cellulase mixture so as to effect the conversion of the cellulose to glucose, said cellulase mixture comprising the Family 6 cellulase variant of any one of claims 1 through 13 and one or more endoglucanases, cellobiohydrolases and beta-glucosidases.
23. The process of claim 22, wherein the cellulosic substrate is a pretreated lignocellulosic feedstock containing greater than 20% (dry wt) cellulose and 10% (dry wt) lignin.
24. The process of claim 23, wherein the pretreated lignocellulosic feedstock comprises corn stover, wheat straw, barley straw, rice straw, oat straw, canola straw, soybean stover; corn fiber, sugar beet pulp, pulp mill fines and rejects, sugar cane bagasse, hardwood, softwood, sawdust, switch grass, miscanthus, cord grass, or reed canary grass.
25. A Family 6 cellulase variant selected from the group consisting of
- TrCel6A-Y103A-S413P (SEQ ID NO: 37);
- TrCel6A-Y103H-S413P (SEQ ID NO: 38);
- TrCel6A-Y103K-S413P (SEQ ID NO: 39);
- TrCel6A-Y103L-S413P (SEQ ID NO: 40);

TrCel6A-Y103M-S413P (SEQ ID NO: 41);
TrCel6A-Y103P-S413P (SEQ ID NO: 42);
TrCel6A-Y103R-S413P (SEQ ID NO: 43);
TrCel6A-Y103V-S413P (SEQ ID NO: 44);
TrCel6A-L136I -S413P (SEQ ID NO: 45);
TrCel6A-L136V -S413P (SEQ ID NO: 46);
TrCel6A-S186K -S413P (SEQ ID NO: 47);
TrCel6A-S186Y -S413P (SEQ ID NO: 48);
TrCel6A-G365D-S413P (SEQ ID NO: 49);
TrCel6A-G365E-S413P (SEQ ID NO: 50);
TrCel6A-G365Q-S413P (SEQ ID NO: 51);
TrCel6A-G365S-S413P (SEQ ID NO: 52);
TrCel6A-R410A-S413P (SEQ ID NO: 53);
TrCel6A-R410F-S413P (SEQ ID NO: 54);
TrCel6A-R410L-S413P (SEQ ID NO: 55);
TrCel6A-R410Q-S413P (SEQ ID NO: 56);
TrCel6A-R410S-S413P (SEQ ID NO: 57);
TrCel6A- M134V -L136I-S413P (SEQ ID NO: 62);
TrCel6A-L136I-L215I-S413P (SEQ ID NO: 63);
TrCel6A-M134V-L136I-L215I -S413P (SEQ ID NO: 71);
HiCel6A-Y107K(SEQ ID NO: 78);
HiCel6A-Y107L (SEQ ID NO: 79);
HiCel6A-Q139T (SEQ ID NO: 80);
HiCel6A-L141V (SEQ ID NO: 81);
HiCel6A-A194Y (SEQ ID NO: 82);
PcCel6A-Y98K (SEQ ID NO: 83);
PcCel6A-Y98L (SEQ ID NO: 84);
PcCel6A-L131I (SEQ ID NO: 85);
PcCel6A-L131V (SEQ ID NO: 86);
PcCel6A-S182K (SEQ ID NO: 87);
PcCel6A-S182Y (SEQ ID NO: 88);
PcCel6A-G359Q (SEQ ID NO: 89); and
PcCel6A-R404Q (SEQ ID NO: 90).

FIGURE 1

208	T. reesei - Cel6A	-----EYEDINTILVIEFOSLANVTLNLTGTPCANAQATLECIINAVTQULPFWAVYLDAGHAGLCHPANOQPAQLFANVTQASSPALALGLATWVANTNORNIIT-----SPPSYTQGNVATNEKLYIN-----
	H. koningi - Cbh2	-----EYEDINTILVIEFOSLANVTLNLTGTPCANAQATLECIINAVTQULPFWAVYLDAGHAGLCHPANOQPAQLFANVTQASSPALALGLATWVANTNORNIIT-----SPPSYTQGNVATNEKLYIN-----
	T. viride GICC 13039 - Cbh2	-----EYEDINTILVIEFOSLANVTLNLTGTPCANAQATLECIINAVTQULPFWAVYLDAGHAGLCHPANOQPAQLFANVTQASSPALALGLATWVANTNORNIIT-----SPPSYTQGNVATNEKLYIN-----
	M. koningi 3.2774 - Cbh2	-----EYEDINTILVIEFOSLANVTLNLTGTPCANAQATLECIINAVTQULPFWAVYLDAGHAGLCHPANOQPAQLFANVTQASSPALALGLATWVANTNORNIIT-----SPPSYTQGNVATNEKLYIN-----
	H. koningi AS3.2774 - Cbh2	-----EYEDINTILVIEFOSLANVTLNLTGTPCANAQATLECIINAVTQULPFWAVYLDAGHAGLCHPANOQPAQLFANVTQASSPALALGLATWVANTNORNIIT-----SPPSYTQGNVATNEKLYIN-----
	T. paracremoris - Cbh11	-----EYEDINTILVIEFOSLANVTLNLTGTPCANAQATLECIINAVTQULPFWAVYLDAGHAGLCHPANOQPAQLFANVTQASSPALALGLATWVANTNORNIIT-----SPPSYTQGNVATNEKLYIN-----
	A. nidulans F08C M4 - An5282.2	-----EYEDINTILVIEFOSLANVTLNLTGTPCANAQATLECIINAVTQULPFWAVYLDAGHAGLCHPANOQPAQLFANVTQASSPALALGLATWVANTNORNIIT-----SPPSYTQGNVATNEKLYIN-----
	A. niger CBS 513.88 - An12902220	-----EYEDINTILVIEFOSLANVTLNLTGTPCANAQATLECIINAVTQULPFWAVYLDAGHAGLCHPANOQPAQLFANVTQASSPALALGLATWVANTNORNIIT-----SPPSYTQGNVATNEKLYIN-----
	A. oryzae R18 40 - A009003800439	-----EYEDINTILVIEFOSLANVTLNLTGTPCANAQATLECIINAVTQULPFWAVYLDAGHAGLCHPANOQPAQLFANVTQASSPALALGLATWVANTNORNIIT-----SPPSYTQGNVATNEKLYIN-----
	A. niger CBS 513.88 - An08401760	-----EYEDINTILVIEFOSLANVTLNLTGTPCANAQATLECIINAVTQULPFWAVYLDAGHAGLCHPANOQPAQLFANVTQASSPALALGLATWVANTNORNIIT-----SPPSYTQGNVATNEKLYIN-----
	A. cellulolyticus Y-94 - Acc2	-----EYEDINTILVIEFOSLANVTLNLTGTPCANAQATLECIINAVTQULPFWAVYLDAGHAGLCHPANOQPAQLFANVTQASSPALALGLATWVANTNORNIIT-----SPPSYTQGNVATNEKLYIN-----
	T. merconii - Cbh11	-----EYEDINTILVIEFOSLANVTLNLTGTPCANAQATLECIINAVTQULPFWAVYLDAGHAGLCHPANOQPAQLFANVTQASSPALALGLATWVANTNORNIIT-----SPPSYTQGNVATNEKLYIN-----
	G. zeae MS9 - Cel6	-----EYEDINTILVIEFOSLANVTLNLTGTPCANAQATLECIINAVTQULPFWAVYLDAGHAGLCHPANOQPAQLFANVTQASSPALALGLATWVANTNORNIIT-----SPPSYTQGNVATNEKLYIN-----
	F. oxysporum - endoglucanase B	-----EYEDINTILVIEFOSLANVTLNLTGTPCANAQATLECIINAVTQULPFWAVYLDAGHAGLCHPANOQPAQLFANVTQASSPALALGLATWVANTNORNIIT-----SPPSYTQGNVATNEKLYIN-----
	M. crassa OR74A - MCO09680.1	-----EYEDINTILVIEFOSLANVTLNLTGTPCANAQATLECIINAVTQULPFWAVYLDAGHAGLCHPANOQPAQLFANVTQASSPALALGLATWVANTNORNIIT-----SPPSYTQGNVATNEKLYIN-----
	A. nidulans F08C M4 - An1273.2	-----EYEDINTILVIEFOSLANVTLNLTGTPCANAQATLECIINAVTQULPFWAVYLDAGHAGLCHPANOQPAQLFANVTQASSPALALGLATWVANTNORNIIT-----SPPSYTQGNVATNEKLYIN-----
	A. tubingenensis - unnamed protein	-----EYEDINTILVIEFOSLANVTLNLTGTPCANAQATLECIINAVTQULPFWAVYLDAGHAGLCHPANOQPAQLFANVTQASSPALALGLATWVANTNORNIIT-----SPPSYTQGNVATNEKLYIN-----
	M. grisea 70-15 - MCO5520.4	-----EYEDINTILVIEFOSLANVTLNLTGTPCANAQATLECIINAVTQULPFWAVYLDAGHAGLCHPANOQPAQLFANVTQASSPALALGLATWVANTNORNIIT-----SPPSYTQGNVATNEKLYIN-----
	C. thermophilus - unnamed protein	-----EYEDINTILVIEFOSLANVTLNLTGTPCANAQATLECIINAVTQULPFWAVYLDAGHAGLCHPANOQPAQLFANVTQASSPALALGLATWVANTNORNIIT-----SPPSYTQGNVATNEKLYIN-----
	C. thermophilus C12 - Cbh2	-----EYEDINTILVIEFOSLANVTLNLTGTPCANAQATLECIINAVTQULPFWAVYLDAGHAGLCHPANOQPAQLFANVTQASSPALALGLATWVANTNORNIIT-----SPPSYTQGNVATNEKLYIN-----
	S. amulata - unnamed protein	-----EYEDINTILVIEFOSLANVTLNLTGTPCANAQATLECIINAVTQULPFWAVYLDAGHAGLCHPANOQPAQLFANVTQASSPALALGLATWVANTNORNIIT-----SPPSYTQGNVATNEKLYIN-----
	H. insolens - Avi2	-----EYEDINTILVIEFOSLANVTLNLTGTPCANAQATLECIINAVTQULPFWAVYLDAGHAGLCHPANOQPAQLFANVTQASSPALALGLATWVANTNORNIIT-----SPPSYTQGNVATNEKLYIN-----
	H. insolens - Cel6A	-----EYEDINTILVIEFOSLANVTLNLTGTPCANAQATLECIINAVTQULPFWAVYLDAGHAGLCHPANOQPAQLFANVTQASSPALALGLATWVANTNORNIIT-----SPPSYTQGNVATNEKLYIN-----
	C. heterostrophus C4 - Cel7	-----EYEDINTILVIEFOSLANVTLNLTGTPCANAQATLECIINAVTQULPFWAVYLDAGHAGLCHPANOQPAQLFANVTQASSPALALGLATWVANTNORNIIT-----SPPSYTQGNVATNEKLYIN-----
	A. bisporus D649 - Cel3A	-----EYEDINTILVIEFOSLANVTLNLTGTPCANAQATLECIINAVTQULPFWAVYLDAGHAGLCHPANOQPAQLFANVTQASSPALALGLATWVANTNORNIIT-----SPPSYTQGNVATNEKLYIN-----
	P. arcularius 69B-8 - Cel2	-----EYEDINTILVIEFOSLANVTLNLTGTPCANAQATLECIINAVTQULPFWAVYLDAGHAGLCHPANOQPAQLFANVTQASSPALALGLATWVANTNORNIIT-----SPPSYTQGNVATNEKLYIN-----
	L. scodes Stamets CS-2 - Cel6B	-----EYEDINTILVIEFOSLANVTLNLTGTPCANAQATLECIINAVTQULPFWAVYLDAGHAGLCHPANOQPAQLFANVTQASSPALALGLATWVANTNORNIIT-----SPPSYTQGNVATNEKLYIN-----
	L. scodes 154 - Cbh11-1	-----EYEDINTILVIEFOSLANVTLNLTGTPCANAQATLECIINAVTQULPFWAVYLDAGHAGLCHPANOQPAQLFANVTQASSPALALGLATWVANTNORNIIT-----SPPSYTQGNVATNEKLYIN-----
	M. cinnamomea - unnamed protein	-----EYEDINTILVIEFOSLANVTLNLTGTPCANAQATLECIINAVTQULPFWAVYLDAGHAGLCHPANOQPAQLFANVTQASSPALALGLATWVANTNORNIIT-----SPPSYTQGNVATNEKLYIN-----
	P. chrysosporium - cellobiohydrolase II	-----EYEDINTILVIEFOSLANVTLNLTGTPCANAQATLECIINAVTQULPFWAVYLDAGHAGLCHPANOQPAQLFANVTQASSPALALGLATWVANTNORNIIT-----SPPSYTQGNVATNEKLYIN-----
	V. volucae - Cbh11-1	-----EYEDINTILVIEFOSLANVTLNLTGTPCANAQATLECIINAVTQULPFWAVYLDAGHAGLCHPANOQPAQLFANVTQASSPALALGLATWVANTNORNIIT-----SPPSYTQGNVATNEKLYIN-----
	C. lucinomea - Cel6A	-----EYEDINTILVIEFOSLANVTLNLTGTPCANAQATLECIINAVTQULPFWAVYLDAGHAGLCHPANOQPAQLFANVTQASSPALALGLATWVANTNORNIIT-----SPPSYTQGNVATNEKLYIN-----
	P. reijer-Cajun - cellobiohydrolase II	-----EYEDINTILVIEFOSLANVTLNLTGTPCANAQATLECIINAVTQULPFWAVYLDAGHAGLCHPANOQPAQLFANVTQASSPALALGLATWVANTNORNIIT-----SPPSYTQGNVATNEKLYIN-----
	T. versicolor - ORF	-----EYEDINTILVIEFOSLANVTLNLTGTPCANAQATLECIINAVTQULPFWAVYLDAGHAGLCHPANOQPAQLFANVTQASSPALALGLATWVANTNORNIIT-----SPPSYTQGNVATNEKLYIN-----
	M. crassa OR74A - MCO03996.1	-----EYEDINTILVIEFOSLANVTLNLTGTPCANAQATLECIINAVTQULPFWAVYLDAGHAGLCHPANOQPAQLFANVTQASSPALALGLATWVANTNORNIIT-----SPPSYTQGNVATNEKLYIN-----
	M. grisea 70-15 - MCO4499.4	-----EYEDINTILVIEFOSLANVTLNLTGTPCANAQATLECIINAVTQULPFWAVYLDAGHAGLCHPANOQPAQLFANVTQASSPALALGLATWVANTNORNIIT-----SPPSYTQGNVATNEKLYIN-----

AMENDED SHEET

FIGURE 1 (cont'd)

332

↓

↑

447

FIGURE 1 (cont'd)

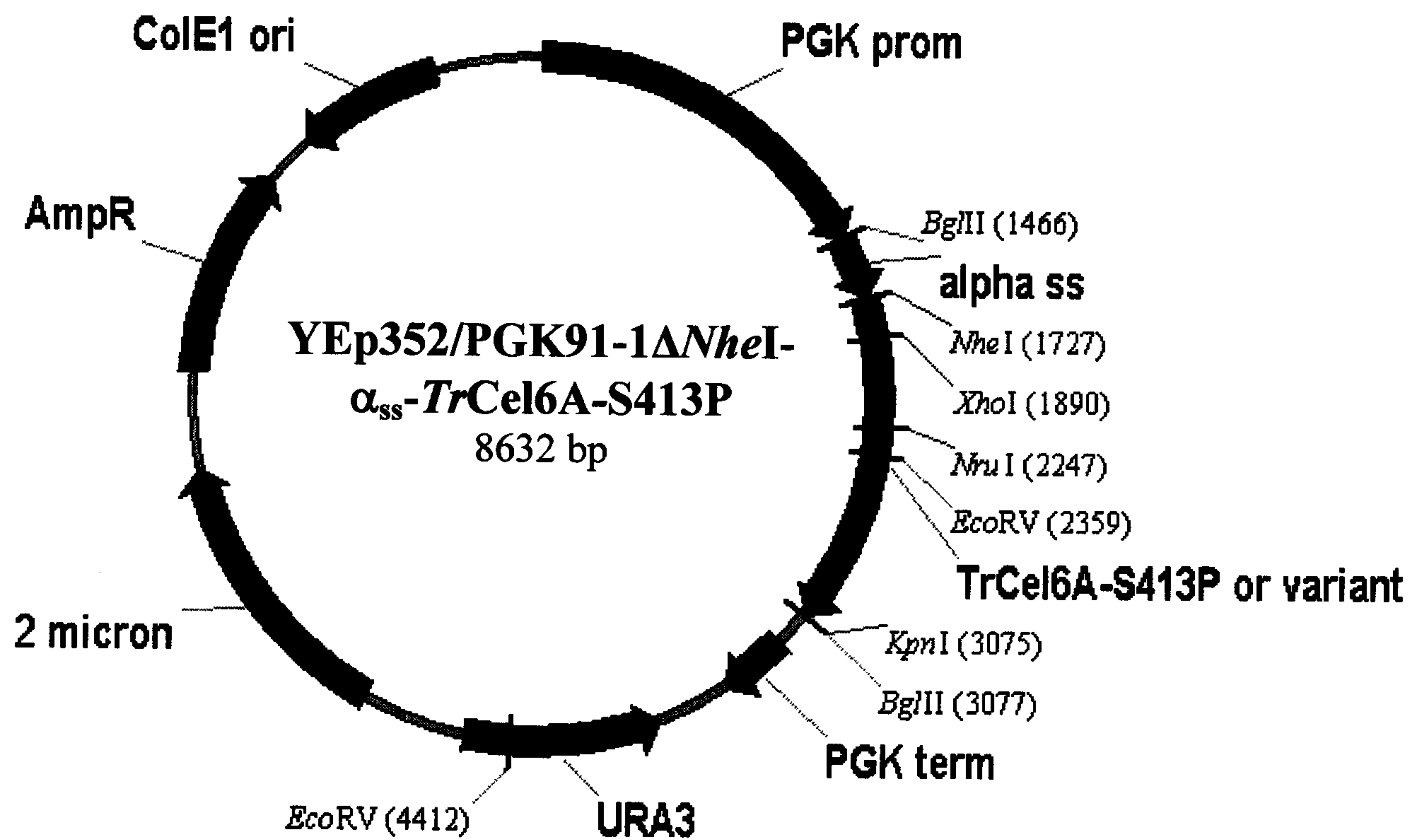
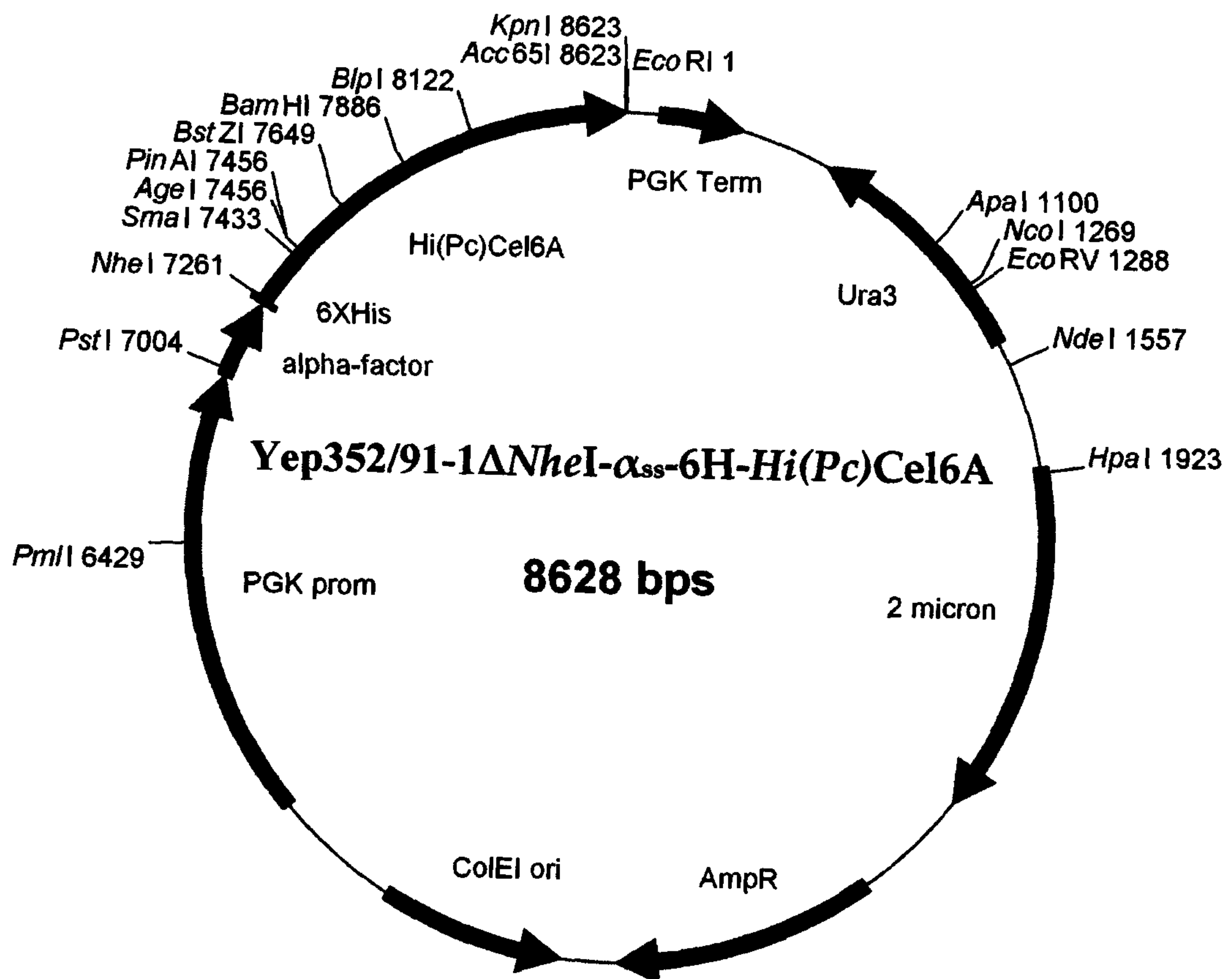
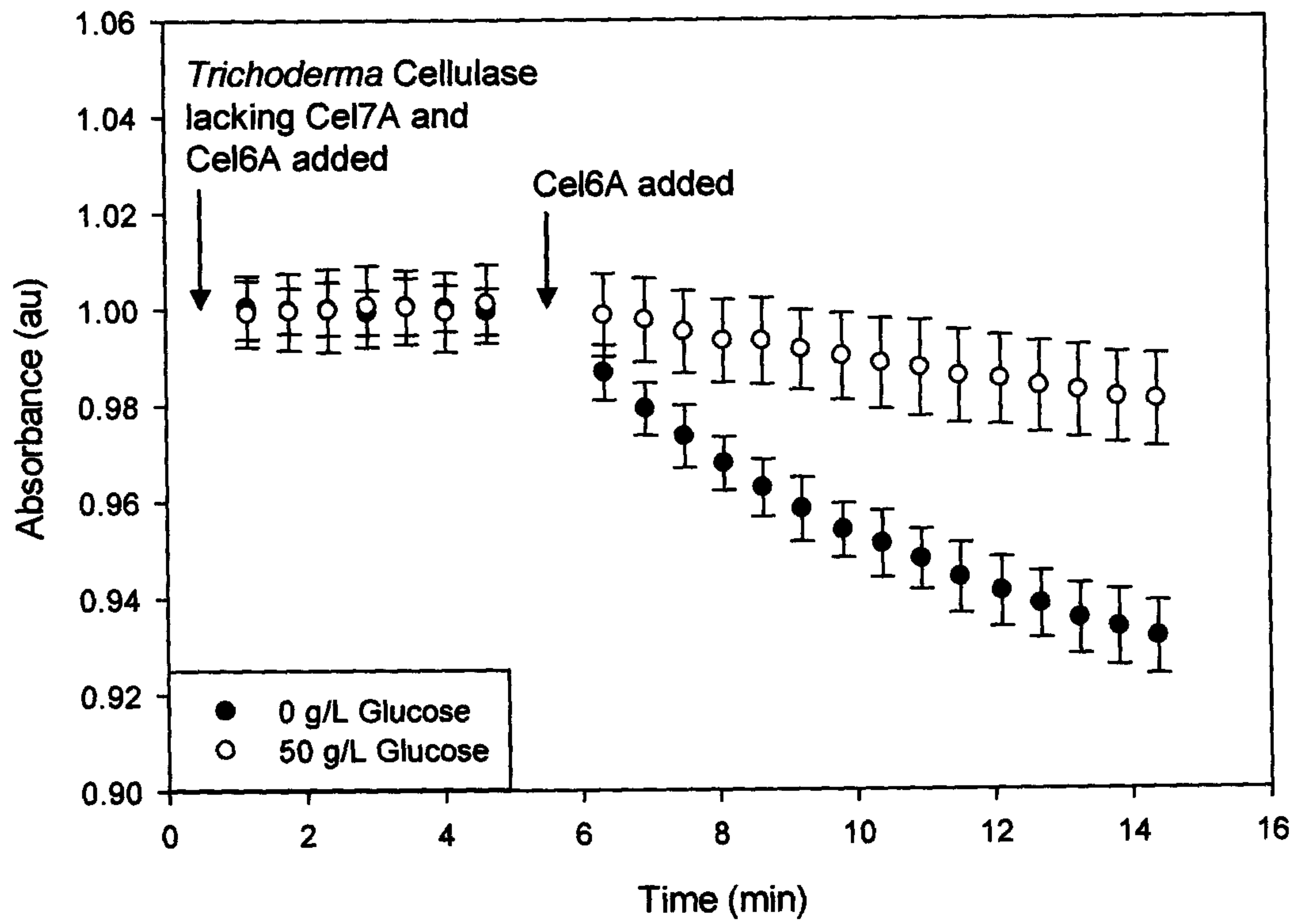
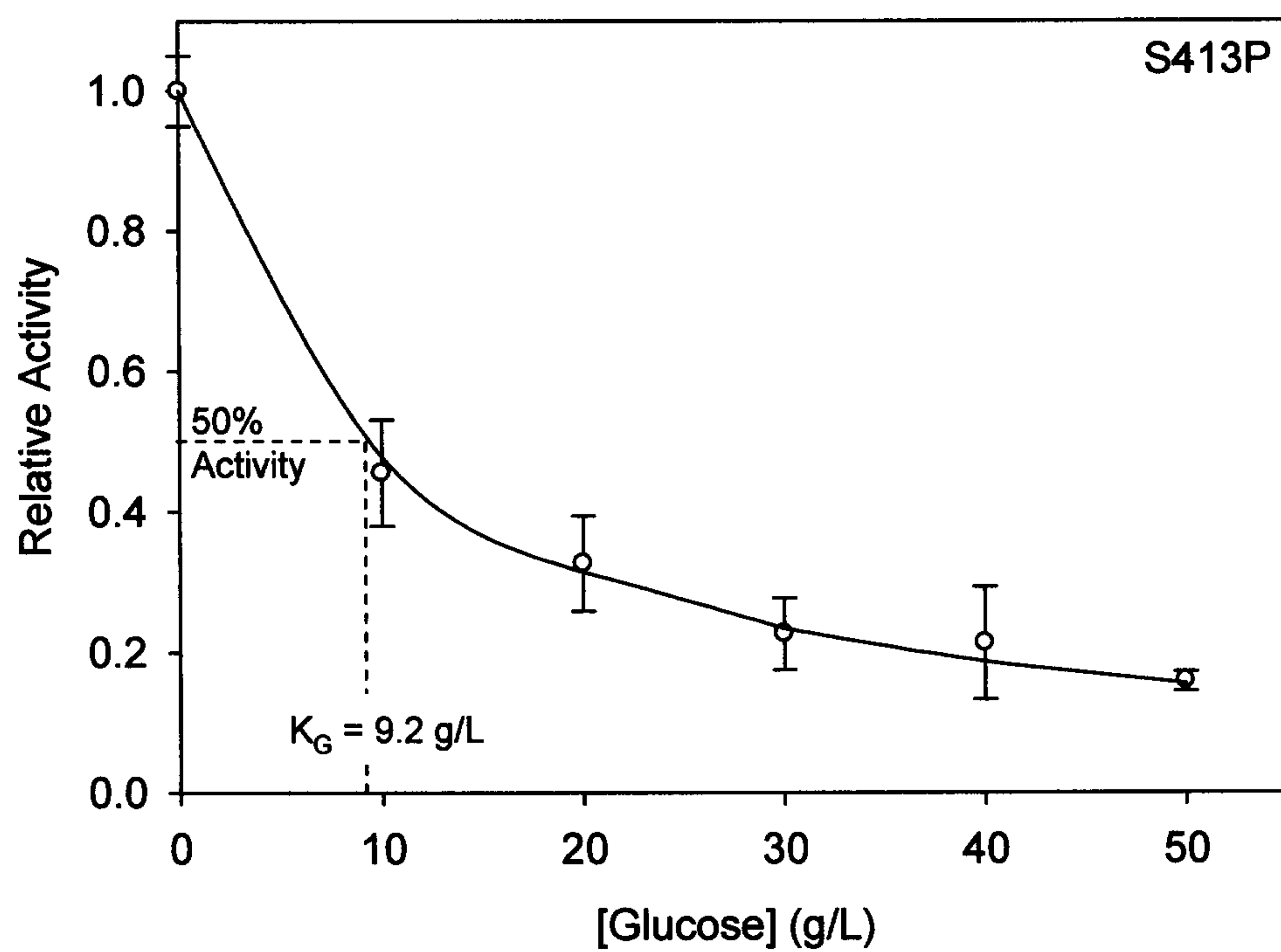
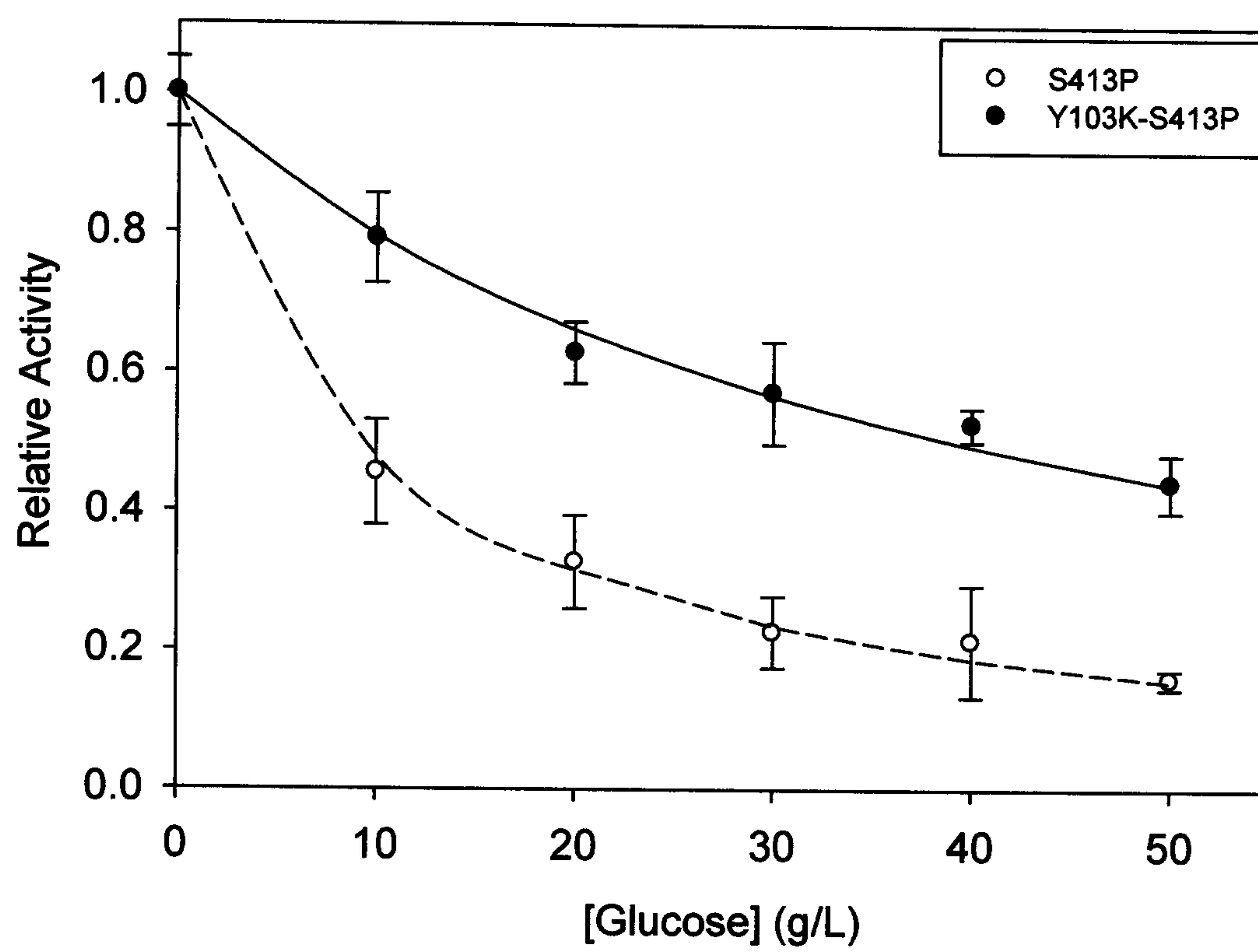


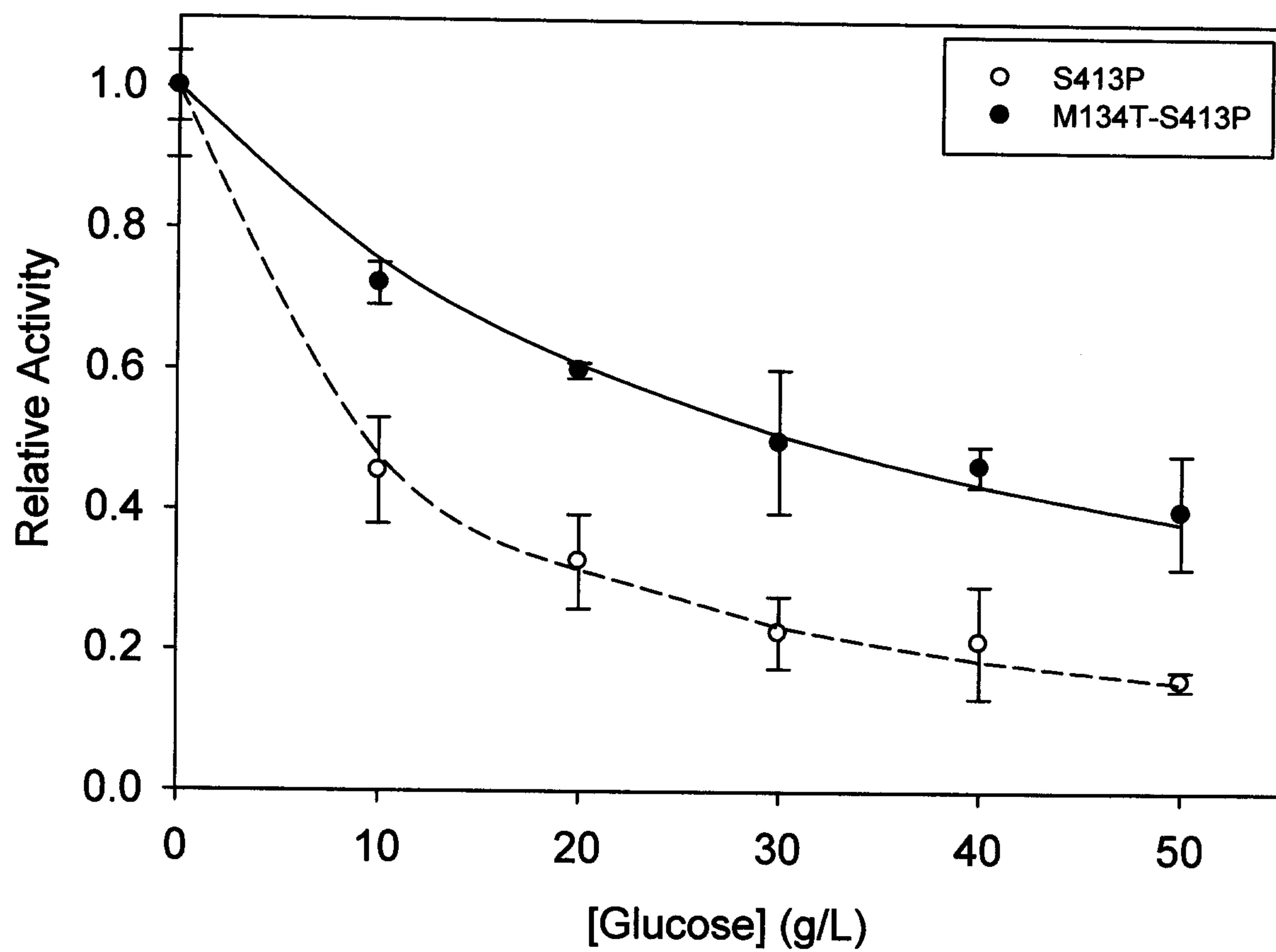
FIG. 2

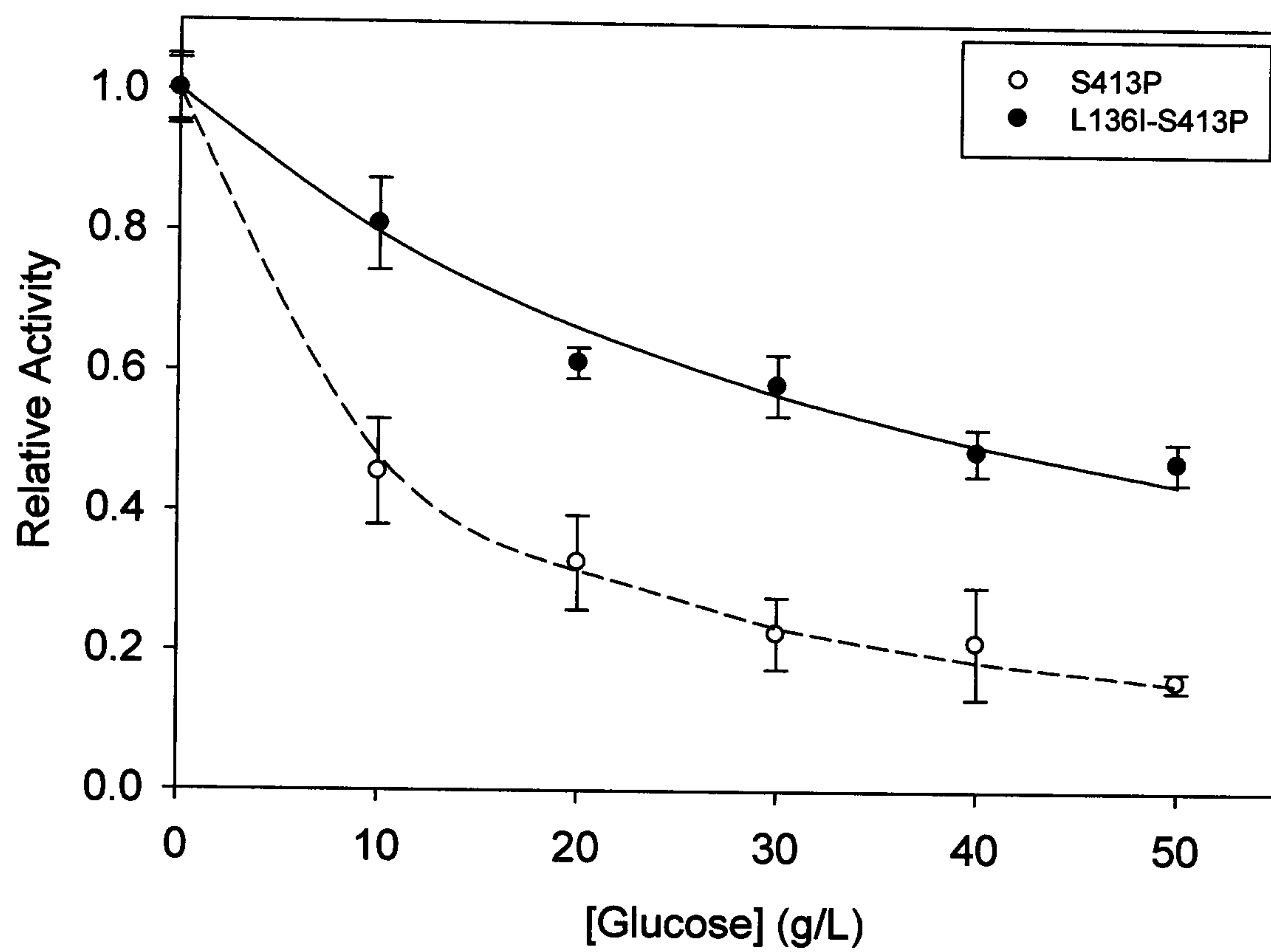
**FIG. 3**

**FIG. 4**

**FIG. 5**

**FIG. 6**

**FIG. 7**

**FIG. 8**

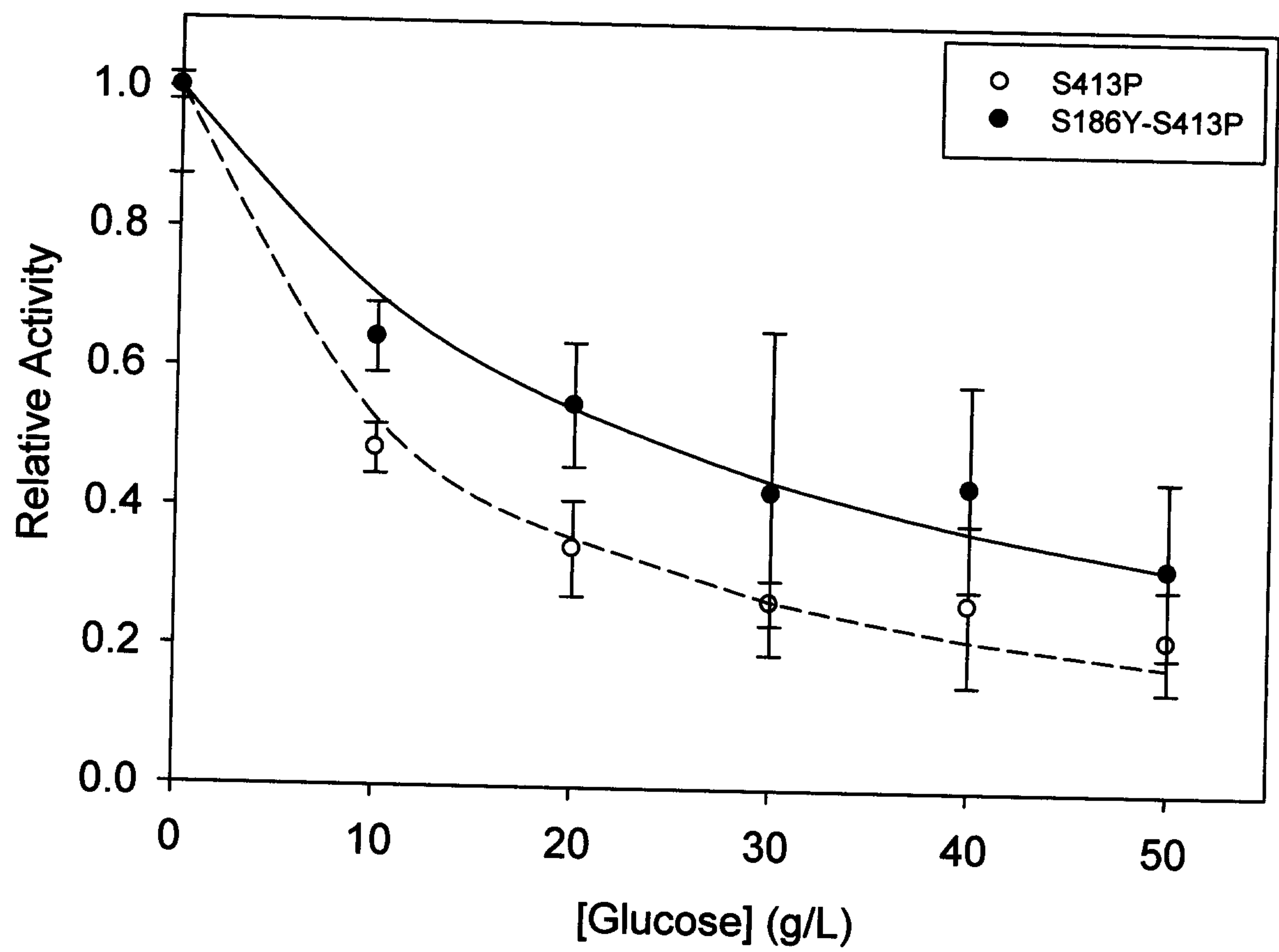
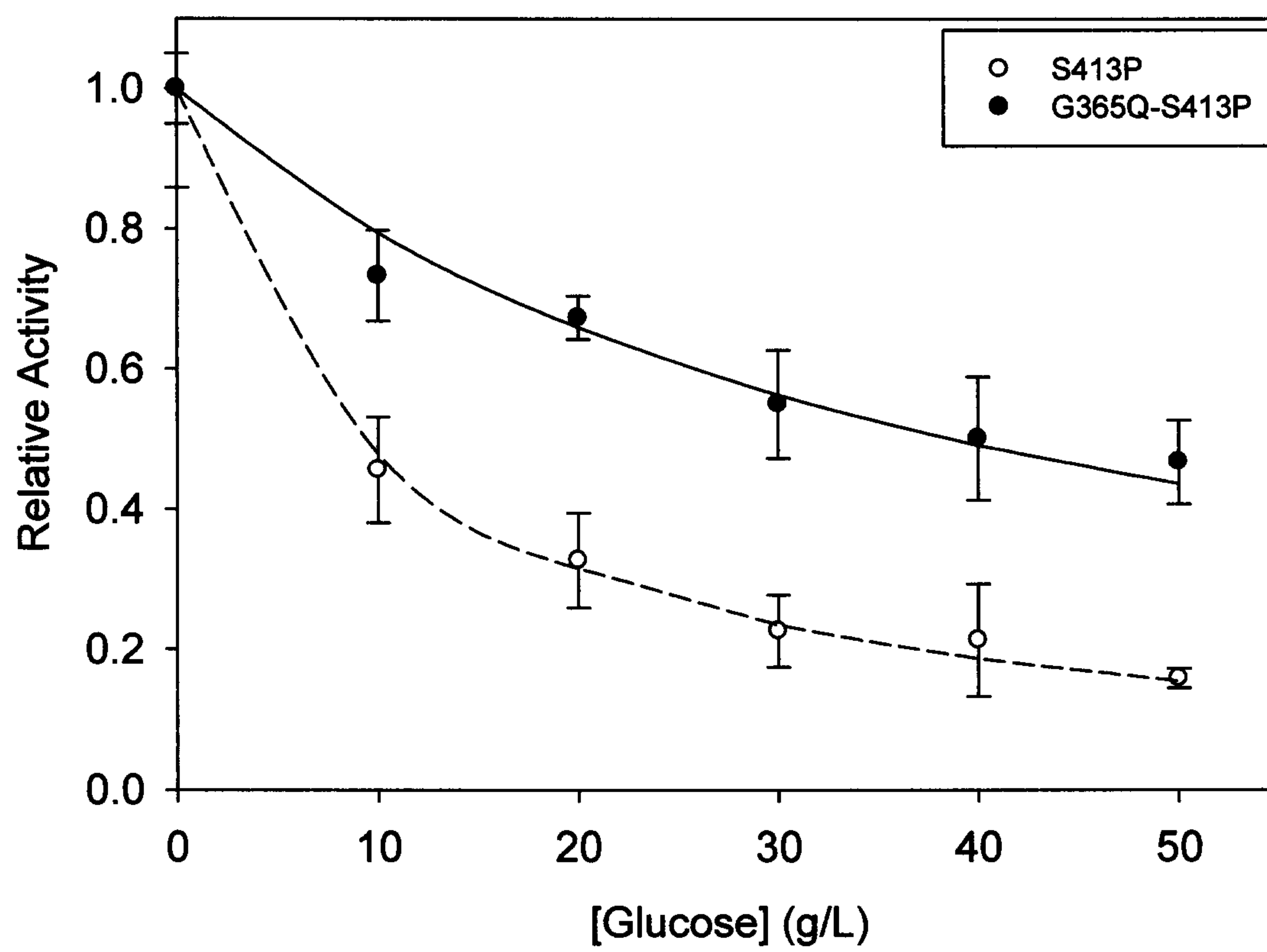


FIG. 9

**FIG. 10**

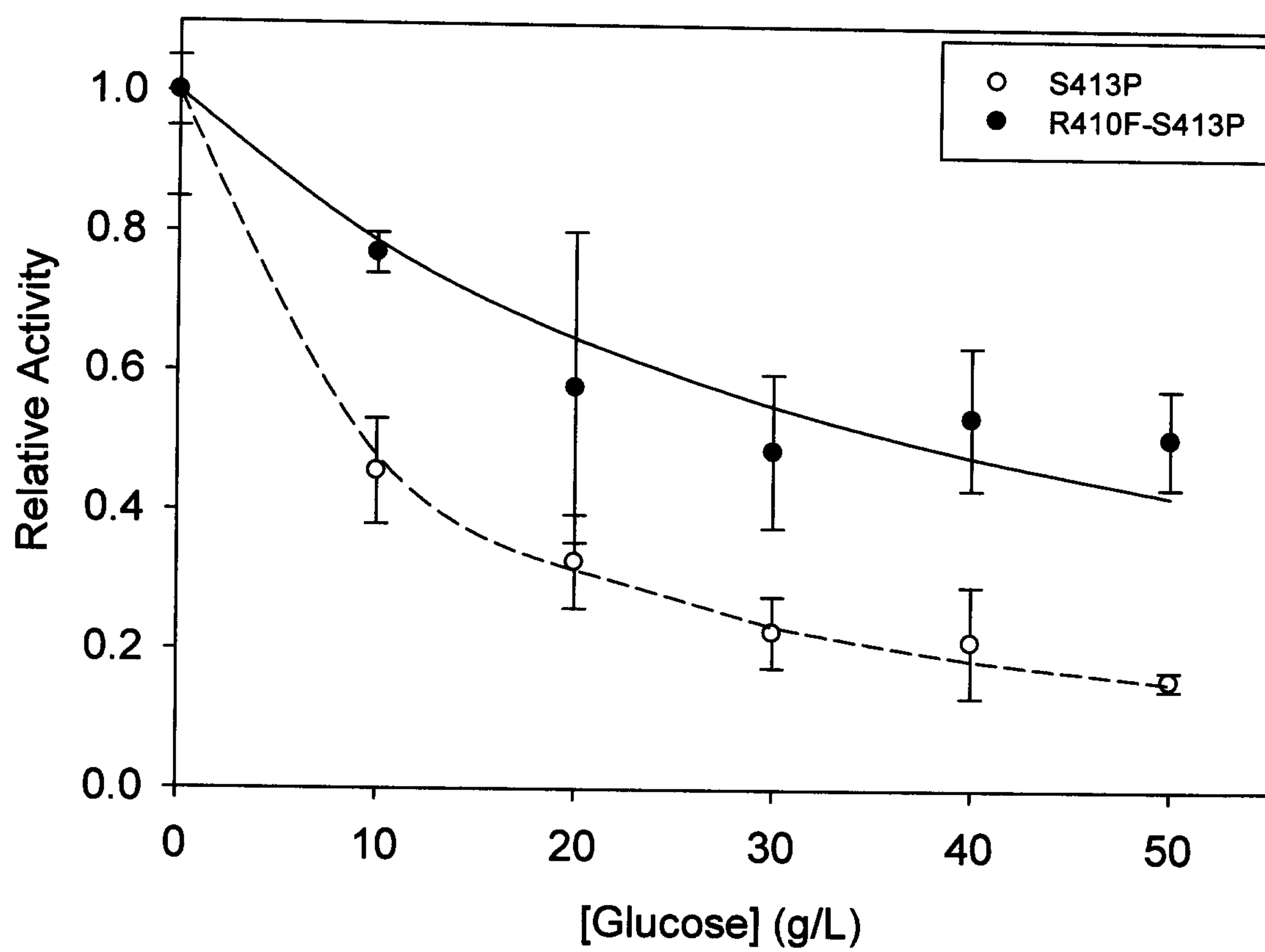


FIG. 11

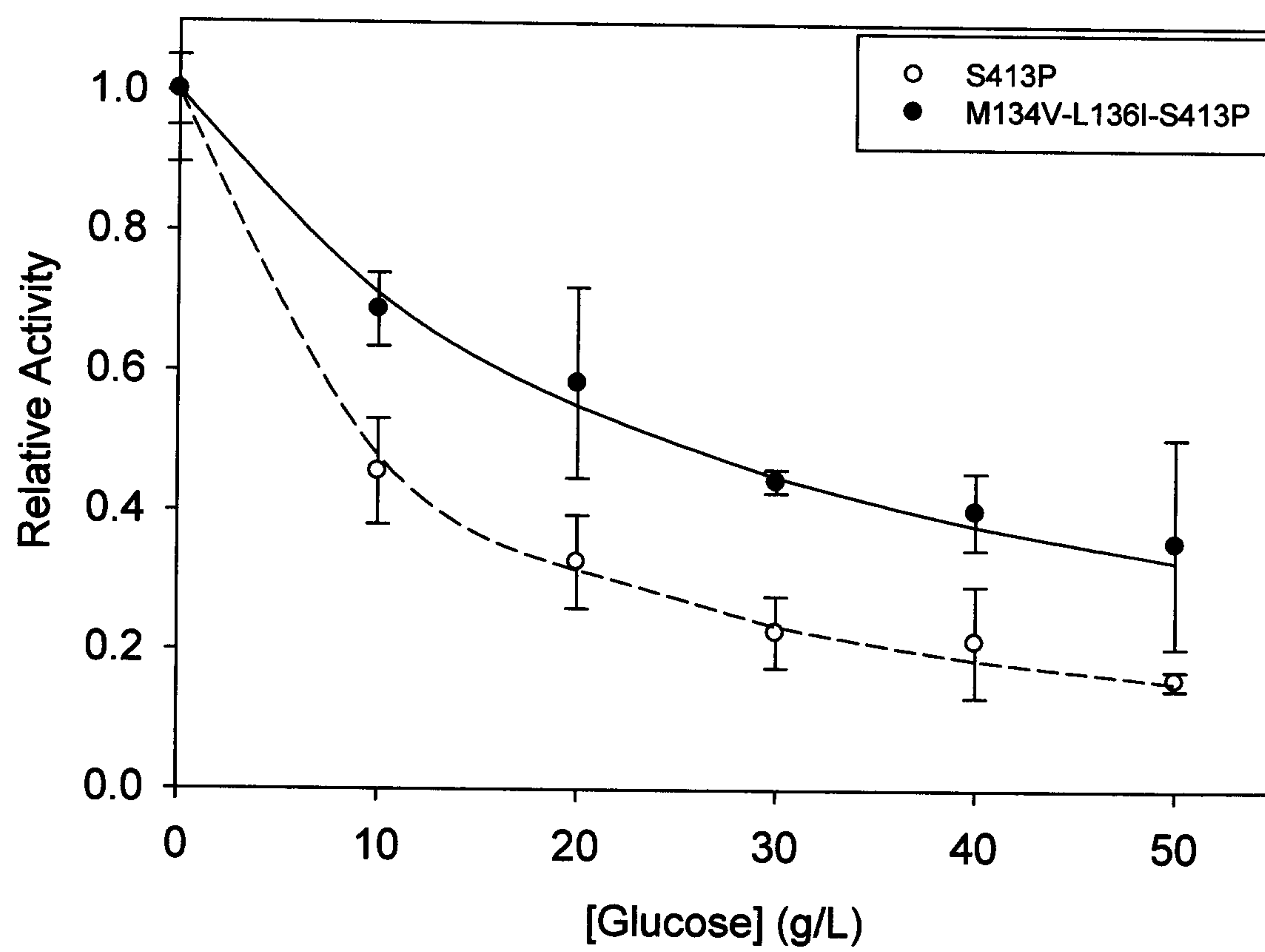
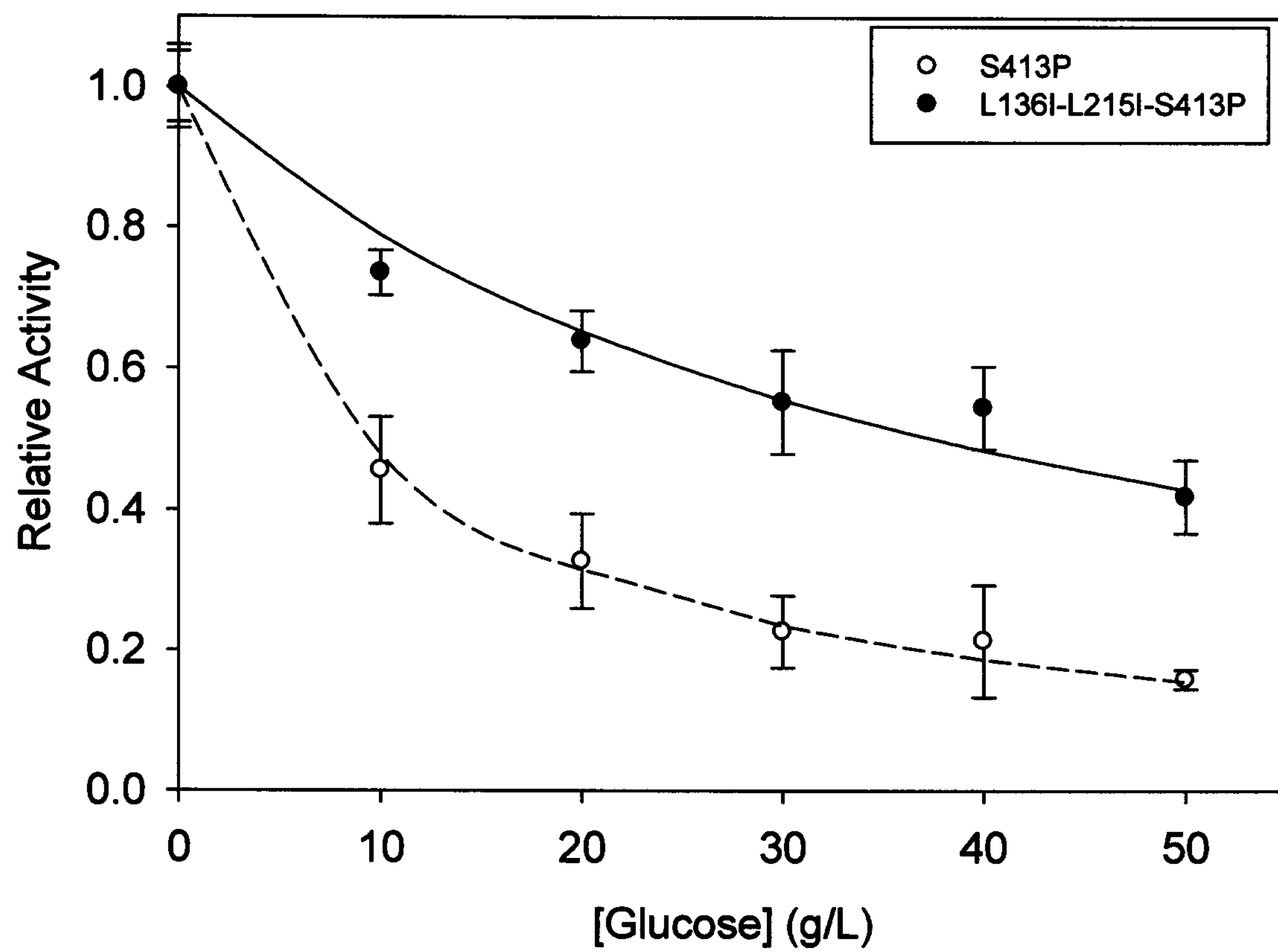
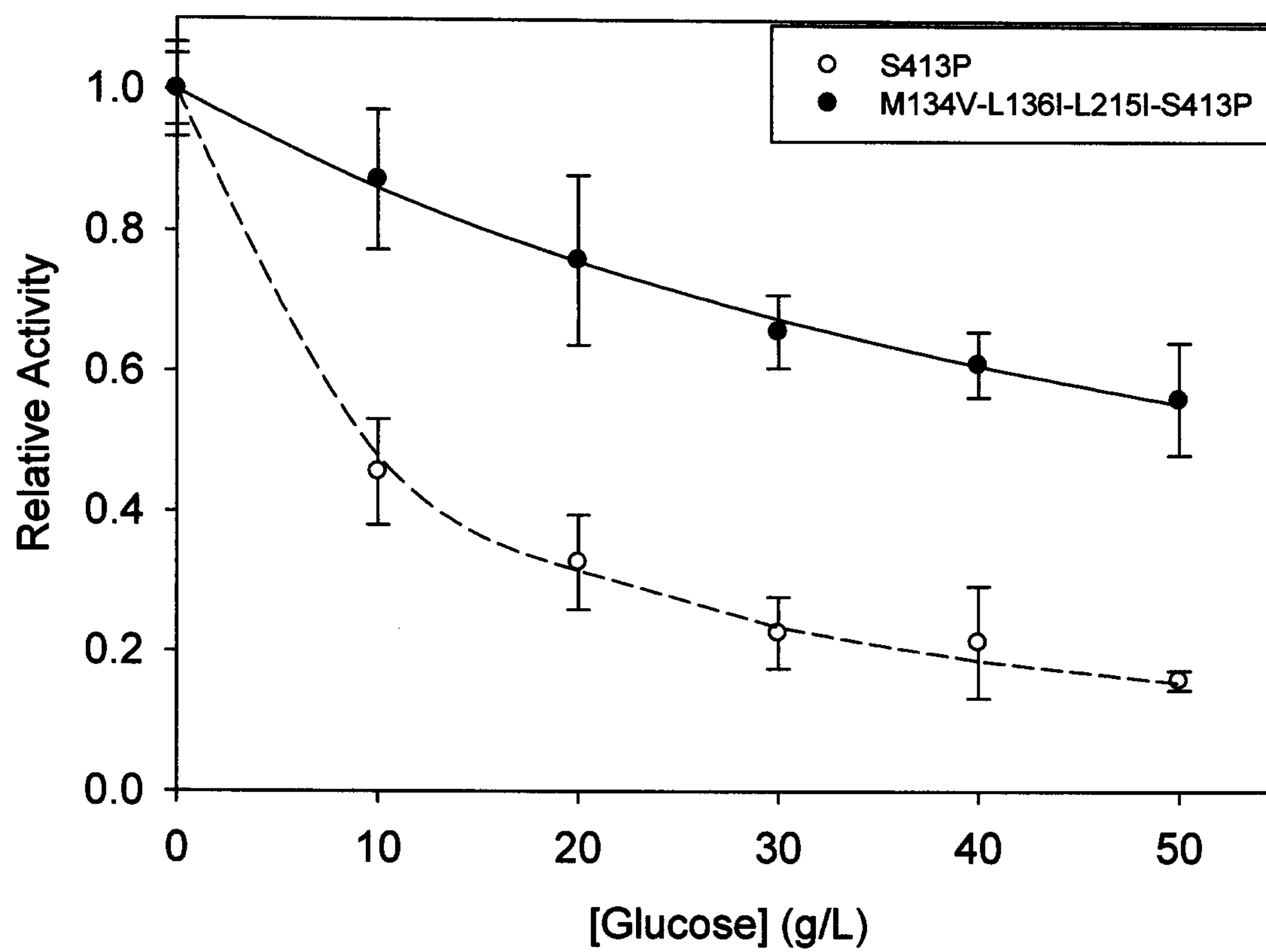
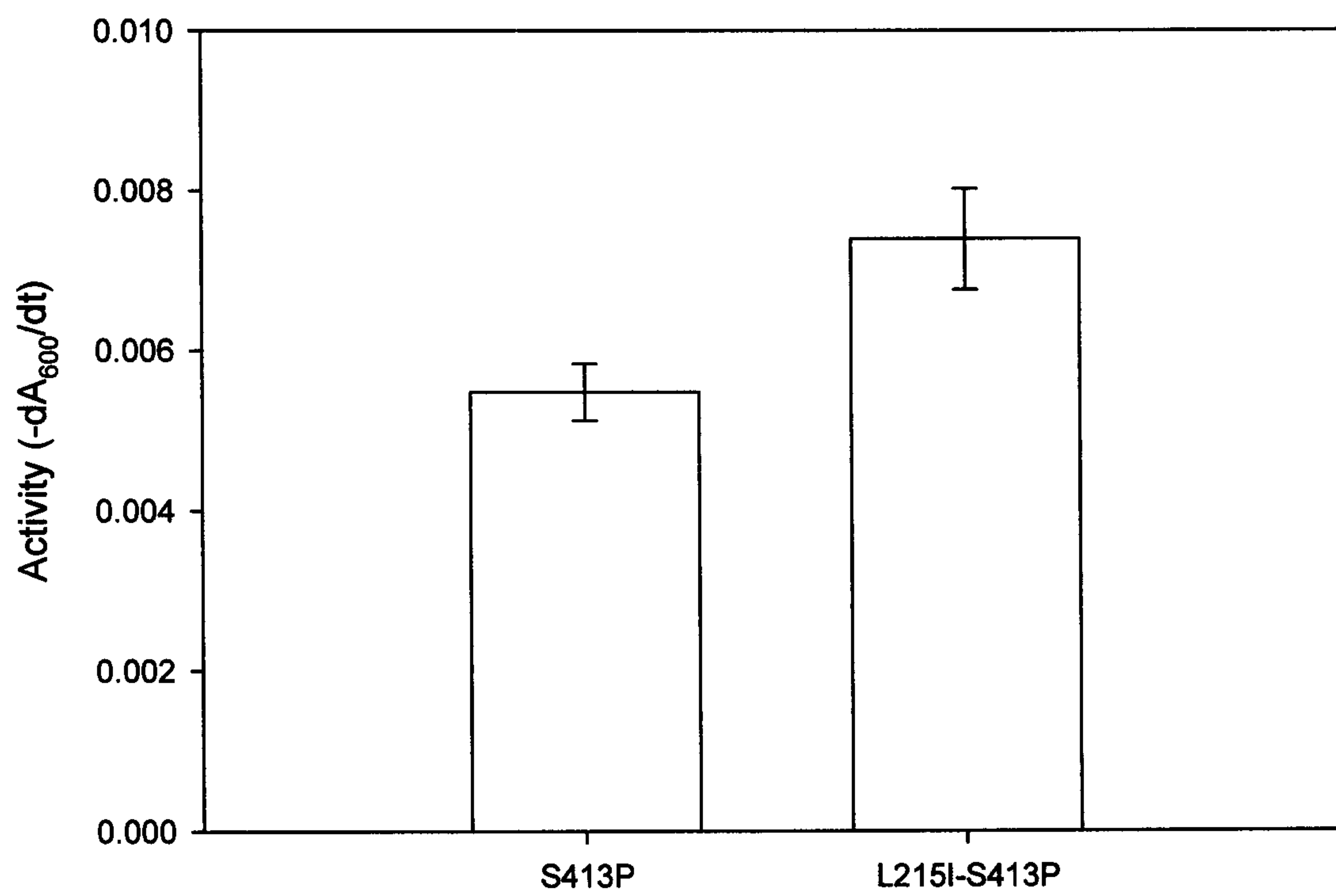


FIG. 12

**FIG. 13**

**FIG. 14**

**FIG. 15**

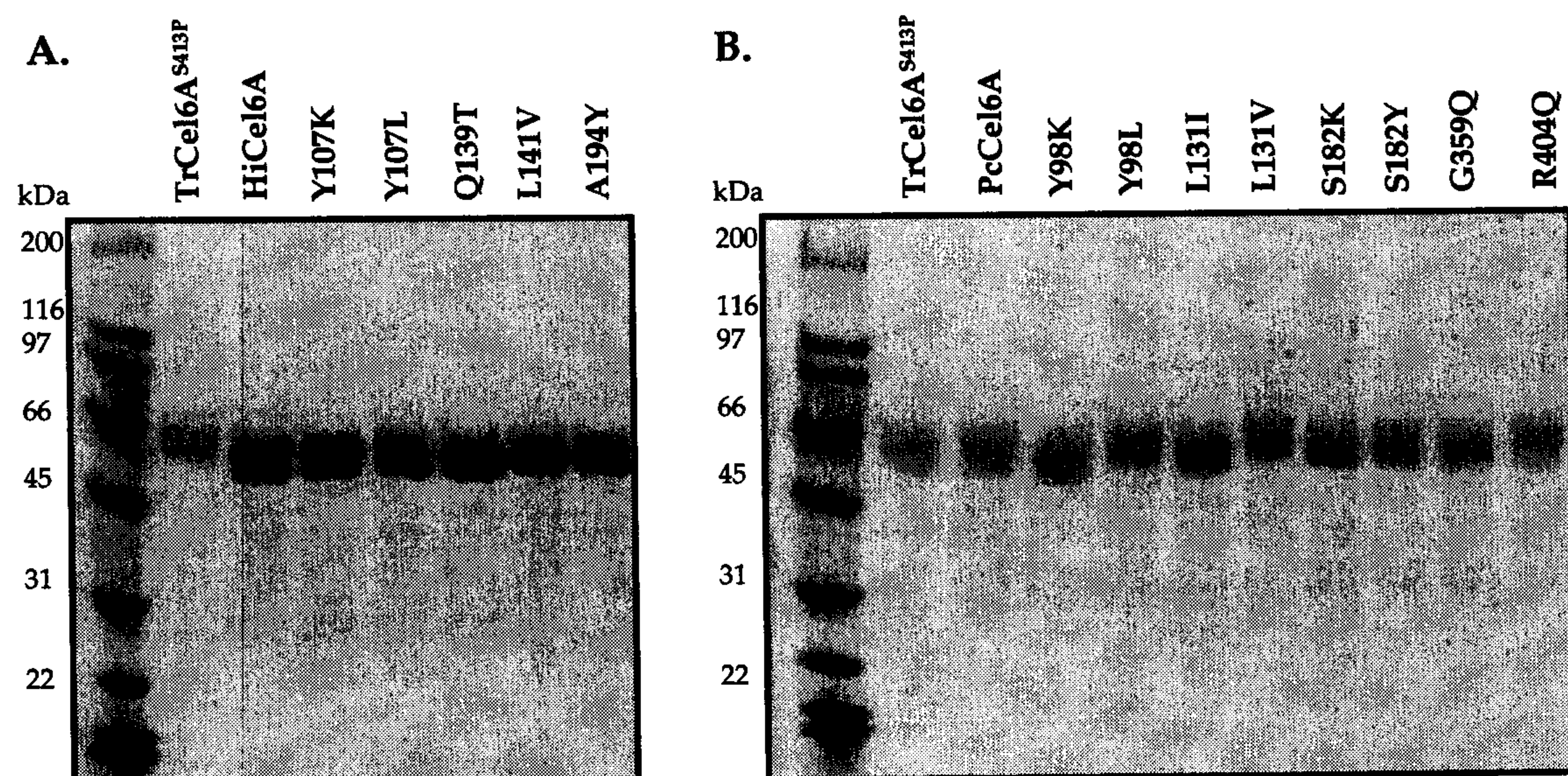


FIG. 16

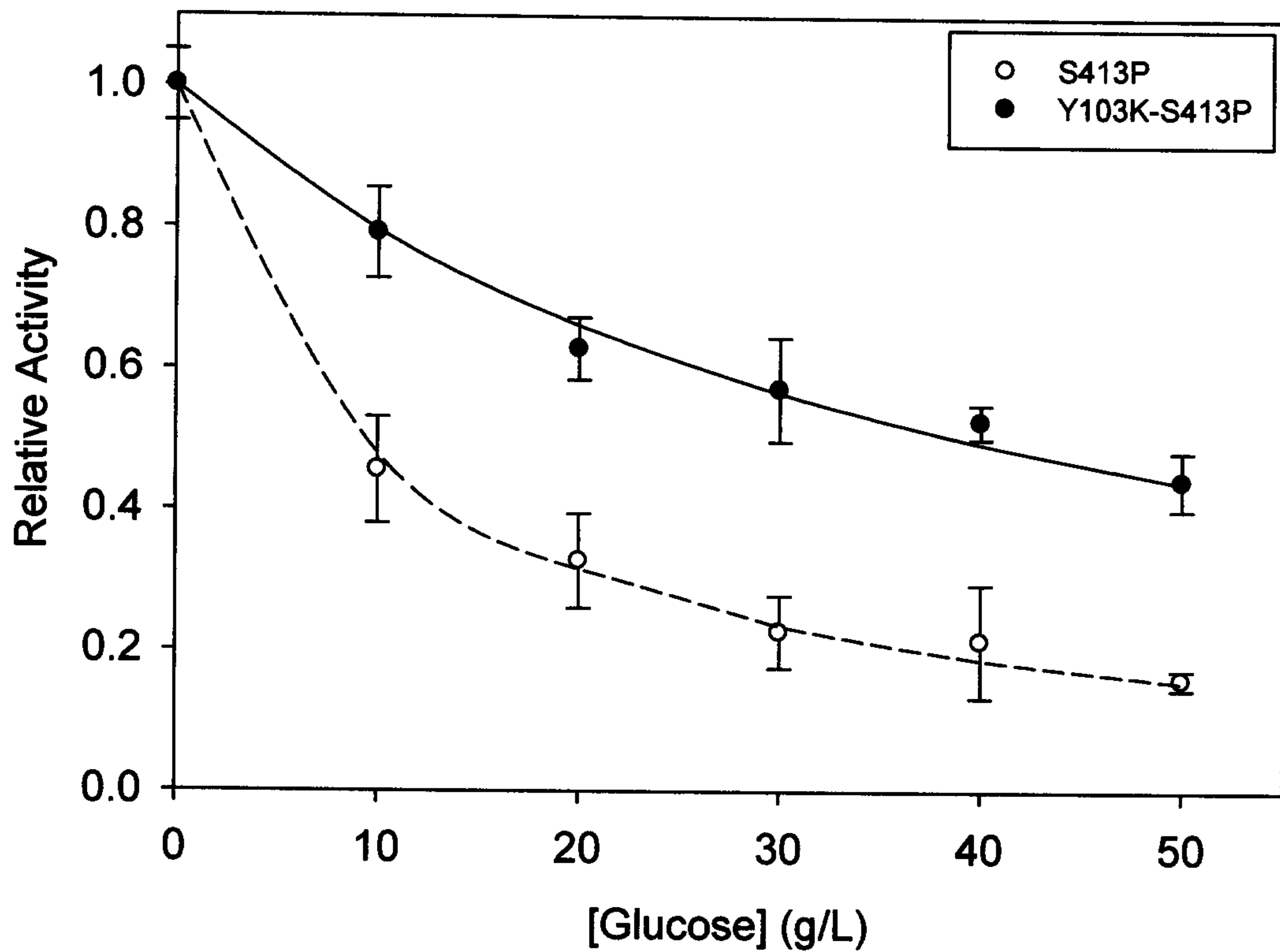


FIG. 6