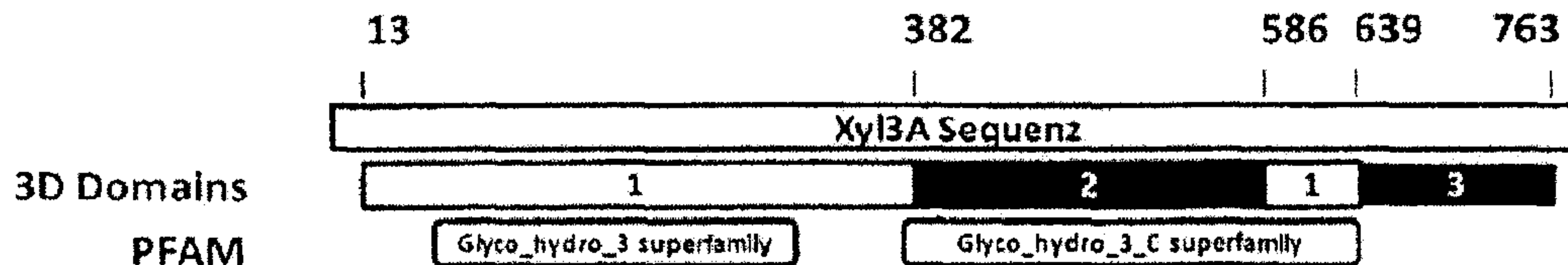




(86) **Date de dépôt PCT/PCT Filing Date:** 2011/10/26
(87) **Date publication PCT/PCT Publication Date:** 2012/05/03
(45) **Date de délivrance/Issue Date:** 2016/12/13
(85) **Entrée phase nationale/National Entry:** 2013/04/25
(86) **N° demande PCT/PCT Application No.:** EP 2011/068734
(87) **N° publication PCT/PCT Publication No.:** 2012/055904
(30) **Priorité/Priority:** 2010/10/26 (DE10 2010 042 910.4)

(51) **Cl.Int./Int.Cl.** **C12N 9/24** (2006.01),
C12N 15/62 (2006.01), **C12N 5/10** (2006.01)
(72) **Inventeurs/Inventors:**
REISINGER, CHRISTOPH, DE;
QOURA, FARAH, DE;
KLIPPEL, BARBARA, DE;
ANTRANIKIAN, GARABED, DE
(73) **Propriétaire/Owner:**
CLARIANT PRODUKTE (DEUTSCHLAND) GMBH, DE
(74) **Agent:** BORDEN LADNER GERVAIS LLP

(54) **Titre : β -PYRANOSIDASE STABLE EN TEMPERATURE**
(54) **Title: TEMPERATURE-STABLE β -PYRANOSIDASE**



(57) **Abrégé/Abstract:**

The invention relates to temperature-stable polypeptides with β -pyranosidase activity. The polypeptide substrates include β -glucopyranosides and β -xylopyranosides. Said polypeptides can be expressed alone or as fusion proteins for example in yeast or bacteria and subsequently purified. The polypeptides according to said invention can be used alone or in a mixture with other enzymes for the degradation of plant raw materials, among others for the enzymatic degradation of biomass containing lignocellulose, in particular hemicellulose and the hemicellulose component xylan. Said enzymes are suitable for use in textile processing, as an additive of detergents, or in the food or feed industry.



ABSTRACT

The invention relates to temperature-stable polypeptides with β -pyranosidase activity. The polypeptide substrates include β -glucopyranosides and β -xylopyranosides. Said polypeptides can be expressed alone or as fusion proteins for example in yeast or bacteria and subsequently purified. The polypeptides according to said invention can be used alone or in a mixture with other enzymes for the degradation of plant raw materials, among others for the enzymatic degradation of biomass containing lignocellulose, in particular hemicellulose and the hemicellulose component xylan. Said enzymes are suitable for use in textile processing, as an additive of detergents, or in the food or feed industry.

Temperature-stable β -Pyranosidase

DESCRIPTION

Background

In view of the decreasing supply of fossil fuels and increasing CO₂ emissions, the exploitation of renewable biomass of plant origin as an alternative source of energy is becoming increasingly important. As a result of the continuously increasing and greatly accelerating demand for energy by industrialised and emerging countries, equivalent alternative energy sources need to be found in addition to a more efficient utilisation of fossil energy supplies (Lyco *et al.*, 2009, J. Biotechnol. 142 (1): 78-86).

The focus of research is shifting to the concept of bio-refining where biomass of plant origin is used as a source of both energy and raw materials. The objective of bio-refining is the full use of renewable raw materials for the production of chemicals, fuels and energy. Dry biomass such as straw and wood obtained from raw and waste materials containing lignocellulose is used as the starting material of lignocellulose bio-refining.

Special interest is dedicated to the utilisation of raw materials containing lignocellulose which originate from agriculture and forestry residues such as straw and wood as these are inexpensive and do not compete with the food and feed industry (Lyco *et al.*, 2009, J. Biotechnol. 142 (1): 78-86; Kumar *et al.*, 2008, J. Ind. Microbiol. Biotechnol. 35 (5): 377-391; Peters, 2007, Adv. Biochem. Eng. Biotechnol. 105: 1-30; Kamm *et al.*, 2006, Biochem. Mol. Biol. Int. 44 (2): 283-292). Lignocellulose is the biopolymer which is available in the largest quantities on earth and consists of cellulose, hemicellulose and lignin. The proportion of cellulose is about 30 - 60 %, that of hemicellulose about 20 - 40 % and the lignin proportion is about 10 - 30 %. In contrast to cellulose which is composed of unbranched glucose units, hemicellulose consists of pentoses and hexoses which may have additional carbohydrate branches. Lignin, on the other hand, is a polymer of phenolic molecules (Peters, 2007, Adv. Biochem. Eng. Biotechnol. 105: 1 - 30).

The use of lignocellulose in bio-refining requires reacting the raw material to obtain useful sugars. For this purpose, lignocellulose is pre-treated by mechanical, thermal and/or chemical methods so as to make the cellulose and hemicellulose more accessible for subsequent hydrolysis (Hendriks & Zeeman, 2009, *Bioresour. Technol.* 100 (1): 10-18). In enzymatic hydrolysis, especially cellulolytic and xylanolytic enzymes play a critical role. These enzymes which belong to the glycoside hydrolases are capable of decomposing the glycosidic bonds in cellulose and hemicellulose.

In addition to the use of cellulolytic and xylanolytic enzymes of mesophilic micro-organisms, the use of thermostable enzymes is of increasing interest as they are suitable for high-temperature processes which increases both the solubility and hence accessibility of the substrate (Turner et al., 2007, *Microb. Cell. Fact.* 6: 9). Moreover, thermostable enzymes are characterised by increased specificity of the substrate and stability vis-à-vis solvents and detergents (Viikari *et al.*, 2007, *Adv. Biochem. Eng. Biotechnol.* 108: 121-145; Antranikian G. (2008) in: *Industrial relevance of thermophiles and their enzymes - Thermophiles - Biology and Technology at High Temperatures*, ed. Robb Fea (CRC Press, Taylor & Francis, Boca Raton), pp 113-160).

Xylan

Xylan, the second-most frequent polysaccharide in nature, is the main component of hemicellulose. It is arranged within the fibril between cellulose and lignin and plays an important role in keeping the micro-fibrils together. Xylans consist of a homopolymer backbone of β -1,4-linked xylopyranose units. This is either linear and unsubstituted, may additionally be acetylated or substituted with arabinosyl and glucuronopyranosyl groups. Therefore, a distinction is made between homoxylan, arabinoxylan, glucuronoxylan and glucuronoarabinoxylan (Saha, 2003, *J. Ind. Microbiol. Biotechnol.* 30 (5): 279-291; Bergquist *et al.*, 2001, *Methods Enzymol* 330: 301-319; Kulkarni *et al.*, 1999, *FEMS Microbiol. Rev.* 23 (4): 411-456). In most cases, xylans are present as complex, highly branched heteropolymers (Collins *et al.*, 2002).

The full hydrolysis of xylans to obtain monosaccharides requires a plurality of enzymes which jointly contribute to degradation (Collins et al., 2005, *FEMS Microbiol. Rev.* 29 (1): 3-23; Bergquist *et al.*, 2001, *Methods Enzymol* 330: 301-319). Endoxylanases cleave the glycosidic bonds within the xylan backbone, mainly forming shorter xylol oligosaccharides, but also xylose, xylobiose and xylotriose (Polizeli *et al.* 2005, *Appl. Microbiol. Biotechnol.* 67 (5): 577-591; Dwivedi *et al.*, 1996, *Appl. Microbiol. Biotechnol.* 45 (1-2): 86-93).

Xylosidases cleave off xylose monomers from the non-reducing end of xylooligosaccharides and xylobiose, whereas xylan is generally not used as a substrate (Collins et al., 2005, FEMS Microbiol. Rev. 29 (1): 3-23; Polizeli *et al.*, 2005, Appl. Microbiol. Biotechnol. 67 (5): 577-591). Additional enzymes such as α -arabinofuranosidases, α -glucuronidases and acetyl xylan esterases are involved in the release of side groups of heterogeneous xylans. α -Arabinofuranosidases separate arabinose from branched arabinoxylans and arabinans, while α -glucuronidases hydrolyse the α -1,2 bonds between the β -xylopyranosyl backbone and the glucuronic acid. Acetylated xylan is hydrolysed by acetyl xylan esterases which separate the acetyl groups of the xylan (Jaeger et al., 2006, Biokatalyse. Angewandte Mikrobiologie, ed. Antranikian G. Springer Verlag, pp 135-160). Xylanolytic enzymes have been identified both in fungi and bacteria and in archaea.

Glycoside hydrolases

Glycoside hydrolases are enzymes which hydrolyse the glycosidic bonds between one or more carbohydrates and a residue that does not contain carbohydrates. They cleave a plurality of α - and β -linked substrates and are distinguished in terms of their substrate specificity. Glycoside hydrolases are classified on the basis of similarities of the amino acid sequence or, respectively, percentage identities of the amino acid sequence into so-called glycoside-hydrolase families (GH families). In addition to a similar amino acid sequence, the members of a family have a similar three-dimensional structure and the same reaction mechanism (Henrissat, 1991, Biochem. J. 280 (Pt 2): 309-316).

A list of the GH families may be found in the CAZy data base („Carbohydrate-Active enZymes") in the Internet (Cantarel *et al.*, 2009). Glycoside hydrolases are divided into two classes according to their reaction mechanism: (1) into enzymes which cleave the glycosidic bonds with reversal of the configuration of the anomeric carbon atom and (2) into enzymes which hydrolyse the glycosidic bonds while maintaining the anomeric configuration (Davies & Henrissat, 1995, Structure 3 (9): 853-859; McCarter & Withers, 1994, Curr. Opin. Struct. Biol. 4 (6): 885-892). Aspartate and/or glutamate residues have been identified as catalytic amino acids in most glycoside hydrolases; however other amino acid residues may also be involved in cleaving the glycosidic bond (Davies & Henrissat, 1995, Structure 3 (9):853-859).

Glycoside hydrolases consist of a catalytic domain and may contain additional domains binding carbohydrates. These are connected to the catalytic domain by a flexible linker and permit the enzymes to bind to the substrate (Shoseyov *et al.*, 2006; Boraston et al., 2004,

Biochem. J. 382 (Pt 3): 769-781). The nomenclature for glycoside hydrolases has been standardised by Henrissat *et al.* (1998, FEBS Lett. 425 (2): 352-354).

Thus the designation of the enzymes and the genes encoding them is made by indicating the substrate in the form of three letters. The designation for the substrate is followed by the number of the GH family the enzyme belongs to and a letter indicating the order in which the enzymes have been identified. An abbreviation for the species is placed first so as to distinguish similar enzymes of different organisms.

Endoxylanases of thermophilic bacteria

A number of endoxylanases of thermophilic bacteria are known which are classified into the glycoside hydrolase families 10, 11 and 43 according to Henrissat *et al.* (1998, FEBS Lett. 425 (2): 352-354), owing to similarities of the amino acid sequences. Bacteria of the species *Caldicellulosiruptor* produce endoxylanases with maximum activities at temperatures of 65-70°C and pH 5.5-6.5. Many endoxylanases are also formed by the anaerobic bacterium *Clostridium*.

In addition to the endoxylanase activity, the enzymes XynC and XynX of *C. thermocellum* display the activity of an endoglucanase (Jung *et al.*, 1998, Biochem. Mol. Biol. Int. 44 (2): 283-292; Hayashi *et al.*, 1997, J. Bacteriol. 179 (13): 4246-4253). The endoxylanases XynC and XynY are known to be located in cellulosomes (Hayashi *et al.*, 1997, J. Bacteriol. 179 (13): 4246-4253; Fontes *et al.*, 1995, Biochem. J. 307 (Pt 1): 151-158).

The most thermo-stable endoxylanases are produced by *Thermotoga maritima* and *T. neapolitana*. These show maximum activities at 85-105°C and half-lives of up to 22 hrs at 90°C and, respectively, 12 hrs at 95°C (Zverlov *et al.*, 1996 Appl. Microbiol. Biotechnol. 45 (1-2): 245-247; Saul *et al.*, 1995, Appl. Environ. Microbiol. 61 (11): 4110-4113; Winterhalter & Liebl, 1995, Appl. Environ. Microbiol. 61 (5): 1810-1815). Additional endoxylanases are formed of thermophilic bacteria from the species *Geobacillus*, *Rhodothermus*, *Thermoanaerobacterium* and *Thermobifida*.

β -Xylosidases of thermophilic bacteria

The β -xylosidases known to date are formed by thermophilic bacteria of the genera *Caldicellulosiruptor*, *Clostridium*, *Geobacillus*, *Thermoanaerobacter* and *Thermo-*

anaerobacterium. They are classified into glycoside hydrolase families 3, 39, 43 and 52 according to Henrissat *et al.* (1998, FEBS Lett. 425 (2): 352-354) on the basis of similarities of the amino acid sequences. In addition to β -xylosidases, the anaerobic bacterium *Clostridium stercorarium* also produces endoxylanases and α -arabinofuranosidases and is hence capable of fully degrading arabinoxylan to obtain xylose and arabinan (Adelsberger *et al.*, 2004, Microbiology 150 (Pt 7): 2257-2266).

The gene encoding the β -xylosidase XynB1 of *G. stearothermophilus* is part of a gene cluster which encodes other enzymes involved in the degradation of xylan. In addition to the gene for the β -xylosidase XynB1, this cluster also contains genes for xylanases and α -glucuronidases (Shulami *et al.*, 1999, J. Bacteriol. 181 (12): 3695-3704). The genes encoding the β -xylosidases of *Thermo anaerobacter brockii* and *Thermoanaerobacterium* sp. JW/SL YS485 are also located directly beside genes encoding a β -glucosidase or, respectively, an acetyl xylan esterase (Breves *et al.*, 1997, Appl. Environ. Microbiol. 63 (10): 3902-3910; Lorenz & Wiegel, 1997, J. Bacteriol. 179(17): 5436-5441).

Industrial applications of cellulolytic and xylanolytic enzymes

The use of cellulolytic and xylanolytic is widespread in the industry. For example, they are used in food and feed production, but cellulases and xylanases are also employed in the paper, pulp and in the textile industry. Their use serves the purpose of increasing yield and improving quality. Moreover, the enzymes are used as an environmentally harmless alternative to conventional methods of treatment.

In the food industry, the cellulolytic and xylanolytic enzymes are used for preparing fruit juices, wine and beer, and in the extraction of oil, preferably olive oil, rapeseed oil and sunflower oil as well as in the baked goods industry. In the preparation of fruit juices, cellulases and hemicellulases are used together with pectinases to increase the yield of the juice and to clarify fruit juices. By treating the flesh of fruit with the enzymes, the formation of juice is enhanced and simultaneously the process time is shortened. By subsequently clarifying the juice with the aid of the enzymes, the viscosity is reduced and the filterability thus improved. However, cellulases and hemicellulases are also used together with pectinases in the preparation of oil, preferably olive oil, rapeseed oil and sunflower oil in order to enhance extraction. These enzymes are also used in the preparation of wines where they contribute to clarification, but also to the extraction of dyes present in the fruit and to improving the aroma of the wine.

Moreover, cellulases are used in breweries to hydrolyse barley glucan so as to facilitate filtration of the beer. Xylanases are also used in the baked goods industry. There, they are used as a flour additive to make the dough easier to process and improve the quality of the baked goods (Beg *et al.*, 2001, Appl. Microbiol. Biotechnol. 56 (3-4): 326-338; Bhat, 2000, Biotechnol. Adv. 18 (5): 355-383; Galante *et al.*, 1998, in: Enzymes, biological control and commercial applications, eds. Harman GE & Kubicek CP, publisher Taylor & Francis, London, Vol. 2, pp. 327-342; Bhat & Bhat, 1997, Biotechnol. Adv. 15 (3-4): 583-620).

Cellulases and xylanases are also used in the feed industry where the enzymes contribute to an increase of the nutritional value and easier digestibility of the feed by digesting the cellulose and hemicellulose in feed products of plant origin (Bhat & Bhat, 1997, Biotechnol. Adv. 15 (3-4): 583-620).

In the paper and pulp industry, cellulases and xylanases are used to modify the wood fibre structure so as to make the pulp easier to process. However, the use of xylanases for bleaching paper is also widespread. In such applications, treatment of the pulp with endoxylanases results in the release of lignin, thus rendering the cell wall of the wood fibres more accessible for bleaching agents, whereby the use of bleaching agents can be significantly reduced (Bhat, 2000, Biotechnol. Adv. 18 (5): 355-383; Buchert *et al.*, 1998, in: *Trichoderma & Gliocladium* - enzymes, biological control and commercial applications, eds. Harman GE & Kubicek CP, publisher Taylor & Francis, London, Vol. 2, pp 343-363).

DISCLOSURE OF THE INVENTION

Definitions

The term "glycoside hydrolase" refers to a protein with enzymatic activity which is capable of hydrolysing glycosidic bonds between carbohydrates or between a carbohydrate and a residue that does not contain a carbohydrate.

The term " β -pyranosidase" refers to a glycoside hydrolase having the activity of a β -glycosidase (E.G. 3.2.1.-) which is capable of catalysing the hydrolysis of a β -glycosidic bond on a pyranose.

The terms " β -glucopyranosidase" and " β -glucosidase" refer to a glycoside hydrolase which is capable of catalysing the hydrolysis of a β -glycosidic bond on a glucose, producing low-molecular glucose oligomers or monomers.

The terms " β -xylopyranosidase" and " β -xylosidase" refer to a glycoside hydrolase which is capable of catalysing the hydrolysis of a β -glycosidic bond on a xylose, producing low-molecular xylose oligomers or monomers.

A "polypeptide" is an oligomer or polymer of amino acid elements which are linked to each other by peptide bonds.

The term "monomer or oligomer component" includes all monomer or oligomer components which may be released from a polymer by enzymatic activity. The term oligomer includes all compounds of at least two components.

"Family 3 glycoside hydrolase" or "glycoside hydrolase of family 3" refers to a polypeptide having glycoside hydrolase activity which is assigned to family 3 according to Henrissat *et al.* (1998, FEBS Lett. 425 (2): 352-354).

"Structural domain 1" of glycoside hydrolase 3 refers to a domain which has at least 50 %, preferably at least 70 %, and more preferably at least 90 % identity of the amino acid sequence with the sequence section comprising amino acid residues 1 to 357 of SEQ ID NO. 4 (barley β -D-glucan exohydrolase isoenzyme Exo I, gene bank reference number AF102868.1, published in: Varghese et al., 1999, Structure, Vol. 7, No. 2, p 179-190).

"Structural domain 2" of glycoside hydrolase 3 refers to a domain having at least 50 %, preferably at least 70 % and more preferably at least 90 % identity of the amino acid sequence with the sequence section comprising the amino acid residues 374 to 559 of SEQ ID. NO. 4 (barley β -D-glucan exohydrolase isoenzyme Exo I, Varghese et al., 1999, Structure, Vol. 7, No. 2, p 179-190).

The term "% identity" in relation to amino acid or nucleotide sequences means the percentage determined by using the following method: The alignment of two amino acid sequences with each other or two nucleic acid sequences with each other is carried out with AlignX. AlignX is a stand-alone application sold by Invitrogen with Vector NTI Advance 10.3.0. The algorithm for the alignments is the ClustalW algorithm (Nucleic Acid Research, 22 (22): 4673-4680, 1994). The parameters used are the value 15/10 for the gap opening penalty and value 6.66/0.1 for the gap extension penalty.

"Similar sequences" are characterised by certain minimum percentage identities among each other, expressed by a number value, for example at least 30 %, at least 50 %, at least 70 % or at least 90 %, and refer to sequences or sequence sections the common evolutionary origin of which may be detected by other criteria such as structural comparisons.

The term "mutation" comprises any kind of nucleotide sequence modification or amino acid sequence modification including deletions, insertions, point mutations, inversions of combinations thereof.

"pNP" represents para-nitrophenyl.

"Thermostability" or "temperature stability" relates to a characteristic of an enzyme which may be determined by dividing an enzyme preparation into two fractions one of which is exposed to a particular temperature and the activity of this fraction after the incubation period is then compared at the particular temperature with the activity of the fraction which was not incubated at the particular temperature. The corresponding value is typically recorded in %.

"pH stability" refers to the characteristic of an enzyme which may be determined by dividing an enzyme preparation into two fractions one of which is exposed to a particular pH value and the activity of this fraction after the incubation period at the particular pH value is then compared with the activity of the fraction which was not incubated at the particular pH value. The corresponding value is typically recorded in %.

Short description of the invention

The present invention relates to polypeptides with β -pyranosidase activity comprising an amino acid sequence having at least 71 %, if possible at least 75 %, if possible at least 80 %, and preferably at least 85 % sequence identity with SEQ ID NO 2. SEQ ID NO. 2 is contained in SEQ ID NO. 1. SEQ ID NO.1 contains a polypeptide which is called *FgXyl3a* in the following. The underlying DNA is represented by SEQ ID NO. 3. It was isolated from the gene bank clone Bgl13 which contains a section of genomic DNA from *Fervidobacterium gondwanense* (*F. gondwanense*).

The invention also relates to polypeptides comprising an amino acid sequence which have at least 70 %, if possible at least 75 %, if possible at least 80 %, and preferably at least 85 % sequence identity with SEQ ID NO. 1. The invention further comprises polypeptides comprising a structural domain 1 of a glycoside hydrolase of family 3 (GHF3) and a

structural domain 2 of a glycosidehydrolase of family 3 (GHF3) wherein at least one of these two domains has at least 70 %, if possible at least 75 %, if possible at least 80 %, and preferably at least 85 % sequence identity with the corresponding domain *FgXyl3a*. Structural domain 1 of *FgXyl3a* consists of the region comprising the amino acid residues 13 to 381 and 586 to 638 of SEQ ID NO. 1; structural domain 2 of *FgXyl3a* consists of the region comprising the amino acid residues 382 to 585 of SEQ ID NO. 1.

The β -pyranosidase activity of the polypeptide according to the present invention is selected from β -xylopyranosidase activity, β -glucopyranosidase activity and of a combination of these two activities, the β -xylopyranosidase activity being preferred.

In a preferred embodiment, the present invention relates to a polypeptide with β -pyranosidase activity, the amino acid sequence of which corresponds to at least one of the above criteria and which is capable of decomposing at least one β -glycosidic bond contained in the following substrates: xylobiose, xylotriose, xylotetraose, xylan.

In a more preferred embodiment, the present invention relates to a polypeptide with β -pyranosidase activity which satisfies at least one of the above criteria and the activity of which vis-à-vis pNP- β -cellobioside is markedly lower than vis-à-vis pNP- β -xylobioside, namely 5 % or less vis-à-vis pNP- β -cellobioside as compared with the activity vis-à-vis pNP- β -xylobioside (pNP representing para-nitrophenyl in this specification).

In a yet more preferred embodiment, the present invention relates to a polypeptide with β -pyranosidase activity which corresponds to at least one of the above criteria and which has maximum activity vis-à-vis pNP- β -xylopyranoside in the acidic, neutral or slightly alkaline range, i.e. in a range from pH 4.0 to pH 8.0, therein, if possible, in the slightly acidic to neutral range, i.e. in a range from pH 5.5 to 7.0, and preferably in a range of pH 6.2 to 6.8.

In a yet more preferred embodiment, the present invention relates to a pH-stable polypeptide with β -pyranosidase activity which corresponds to at least one of the above criteria and the maximum activity of which vis-à-vis pNP- β -xylopyranoside is at least 50 %, if possible at least 60 %, if possible at least 70 %, and preferably at least 80 % after 48 hours of incubation at pH 9.0.

In a yet more preferred embodiment, the present invention relates to a polypeptide with β -pyranosidase activity, which polypeptide is active at high temperature, and which satisfies at least one of the above criteria, and which has maximum activity vis-à-vis pNP- β -

xylopyranoside in a range between 60°C and 100°C, if possible between 70 °C and 95 °C, and preferably between 80 °C and 90 °C.

In a yet more preferred embodiment, the present invention relates to a temperature-stable polypeptide with β -pyranosidase activity which satisfies at least one of the above criteria and the maximum activity of which vis-à-vis pNP- β -xylopyranose after three hours of incubation at 60°C is at least 40 %, preferably at least 50 %.

In a more preferred embodiment, the polypeptide which satisfies at least one of the above criteria may be present as a fusion protein; it is preferably fused to one of the following: a carbohydrate binding domain of another protein, a signal peptide, an affinity tag or a protease cleavage site.

The present invention also relates to a mixture containing the polypeptide with β -pyranosidase activity which satisfies at least one of the above criteria and one or more pectinases and/or one or more endoxylanases and/or one or more β -glucosidases and/or one or more β -glucanases and/or one or more cellobiohydrolases and/or one or more β -xylosidases and/or one or more α -arabinofuranosidases and/or one or more α -glucuronidases and/or one or more acetyl xylan esterases.

The present invention also relates to a nucleic acid which encodes the polypeptide described above, and to a vector containing said nucleic acid.

The present invention also relates to a host cell transformed with the vector described above. The host cell according to the present invention may be a prokaryote or a eukaryote. A eukaryotic host cell is preferably selected from the group consisting of *Saccharomyces cerevisiae*, *Yarrowia lipolytica*, *Schizosaccharomyces lactis*, *Kluyveromyces lactis*, *Pichia methanolytica*, *Pichia pastoris*, *Pichia angusta*, *Hansenula polymorpha*, *Aspergillus niger*, *Chrysosporium lucknowense*, *Trichoderma reesei*, *Penicillium* sp.

In a particularly preferred embodiment, the eukaryotic host cell is a methylotrophic yeast, preferably from the group comprising *Pichia methanolytica*, *Pichia pastoria*, *Pichia angusta*, *Hansenula polymorpha*.

A prokaryotic host cell according to the present invention is preferably selected from the group comprising *Bacillus* sp., *Bacillus subtilis*, *Bacillus licheniformis*, *Bacillus megaterium*, *Thermus thermophilus*, *Pseudomonas fluorescens*, *Fervidobacterium* sp., *Escherichia coli*.

The present invention also relates to a method for purifying the above polypeptide with β -pyranosidase activity comprising the following steps:

- a) obtaining the host cell transformed with the vector described above;
- b) cultivating the host cell under conditions where the polypeptide with β -pyranosidase activity is expressed,
- c) purifying the polypeptide with β -pyranosidase activity.

In a preferred embodiment, step c) of the method for purifying the polypeptide with β -pyranosidase activity described above comprises heat precipitation.

The present invention also relates to the use of the polypeptide with β -pyranosidase activity described above or of the mixture described above containing the polypeptide with β -pyranosidase activity described above for degrading one or more of the following substrates: β -xylopyranoside or β -glucopyranoside.

In a preferred embodiment, the present invention relates to the use described above for the enzymatic degradation of biomass containing lignocellulose and/or for textile processing and/or as an additive to detergents and/or in the food and/or feed industry.

In a preferred embodiment, the present invention relates to the use described above for preparing fruit juices and/or for preparing wine or beverages containing wine, and/or for the preparation of beer or beverages containing beer, and/or for the preparation of oil, preferably olive oil, rapeseed oil or sunflower oil and/or for the preparation of baked goods.

Detailed description of the invention

Polypeptides with β -pyranosidase activity

The present invention relates to polypeptides with β -pyranosidase activity comprising an amino acid sequence having at least 71 %, if possible at least 75 %, if possible at least 80 %, and preferably at least 85 % sequence identity with SEQ ID NO 2. The sequence identity is determined as described in the section "Definitions".

SEQ ID NO. 2 is contained in SEQ ID NO. 1. SEQ ID NO. 1 contains the amino acid sequence of the polypeptide *FgXyl3a*. The DNA underlying the SEQ ID NO. 1, represented

by SEQ ID NO. 3, originates from the gene bank clone Bgl13 which contains a section of genomic DNA from *Fervidobacterium gondwanense* (*F. gondwanense*). Finding the SEQ ID NO. 3 and determining the domain structure are described in detail in example 3.

In short, SEQ ID NO. 2 comprises that region of SEQ ID NO. 1 which contains both structural domains of the glycoside hydrolase and was determined by a sequence comparison with SEQ ID NO. 4 (enzyme BglB of *Thermotoga neapolitana*, Pozzo et al., 2010, J. Mol. Biol., 2: 397 (3), 724-739). SEQ ID NO. 2 thus comprises amino acid residues 13 to 638 of SEQ ID NO. 1.

The present invention also relates to polypeptides comprising an amino acid sequence which has at least 70 %, if possible at least 75 %, if possible at least 80 %, and preferably at least 85 % sequence identity with SEQ ID NO. 1.

The invention further comprises polypeptides comprising a structural domain 1 of a glycoside hydrolase of family 3 (GHF3) and a structural domain 2 of a glycoside hydrolase of family 3 (GHF3), wherein at least one of these two domains has at least 70 %, if possible at least 75 %, if possible at least 80 %, and preferably at least 85 % sequence identity with the corresponding domain *FgXyl3a*, hat. Structural domain 1 of *FgXyl3a* consists of the region comprising amino acid residues 13 to 381 and 586 to 638 of SEQ ID NO. 1; the structural domain 2 of *FgXyl3a* consists of the region comprising the amino acid residues 382 to 585 of SEQ ID NO. 1. The method for determining structural domains is described in example 1. It may be applied generally to polypeptides the sequence identity of which satisfies at least the criteria described herein.

The cloning (example 2), expression (example 3), purification from a cell extract (examples 4 and 5) and determination of the molecular weight (example 6) of a polypeptide with β -pyranosidase activity according to the present invention are described in the examples.

The β -pyranosidase activity of the polypeptide according to the present invention is selected from β -xylopyranosidase activity, β -glucopyranosidase activity and a combination of these two activities, the β -xylopyranosidase activity being preferred. The detection of activity and the determination of the substrate specificity of a polypeptide according to the present invention are described in detail in examples 7 and 8.

In a preferred embodiment, the present invention relates to a polypeptide with β -pyranosidase activity which satisfies at least one of the above criteria and which is capable of hydrolysing

at least one of the β -glycoside bonds contained in the following substrates: xylobiose, xylotriose, xylotetraose, xylan. The determination of the hydrolysis products of a polypeptide with β -pyranosidase activity is described in detail in example 13.

In more preferred embodiment, the present invention relates to a polypeptide with β -pyranosidase activity which satisfies at least one of the above criteria and the activity of which vis-à-vis pNP- β -cellobioside is markedly lower than vis-à-vis pNP- β -xylobioside, namely 5 % or less vis-à-vis pNP- β -cellobioside compared with the activity vis-à-vis pNP- β -xylobioside. A method for determining substrate specificity is described in example 8.

In a yet more preferred embodiment, the present invention relates to a polypeptide with β -pyranosidase activity which satisfies at least one of the above criteria and which has maximum activity vis-à-vis pNP- β -xylopyranoside in the acidic, neutral or slightly basic range, i.e. in a range of pH 4.0 to pH 8.0, therein, if possible, in the slightly acidic to neutral range, i.e. in a range of pH 5.5 to 7.0, and preferably in a range of pH 6.2 to 6.8. A method for determining the influence of the pH value on the activity of a polypeptide with β -pyranosidase activity is described in more detail in example 11.

In a yet more preferred embodiment, the present invention relates to a pH-stable polypeptide with β -pyranosidase activity which satisfies at least one of the above criteria and the maximum activity of which vis-à-vis pNP- β -xylopyranoside after 48 hours of incubation at a pH of 9.0 is at least 50 %, if possible at least 60 %, if possible at least 70 %, and preferably at least 80 %. A method for determining the influence of the pH value on the activity of a polypeptide with β -pyranosidase activity is described in more detail in example 11.

In a yet more preferred embodiment, the present invention relates to a polypeptide with β -pyranosidase activity which is active at high temperature and which satisfies at least one of the above criteria and which has maximum activity vis-à-vis pNP- β -xylopyranoside in a range between 60°C and 100°C, if possible between 70°C and 95°C, and preferably between 80 °C and 90 °C. A method for determining the influence of the temperature on the activity of a polypeptide with β -pyranosidase activity is described in more detail in example 9.

In an yet more preferred embodiment, the present invention relates to a temperature-stable polypeptide with β -pyranosidase activity which satisfies at least one of the above criteria and the maximum activity of which vis-à-vis pNP- β -xylopyranoside after 3 hours of incubation at 60 °C is at least 40 %, preferably at least 50 %. A method for determining the influence of

the temperature stability on the activity of a polypeptide with β -pyranosidase activity is described in more detail in example 10.

Fusion protein

The polypeptide according to the present invention may either occur in the isolated form or it may be fused to one or more additional oligo- or polypeptides. That is, in a further preferred embodiment the polypeptide satisfying at least one of the criteria described above may be present as a fusion protein, wherein it is preferably fused to one of the following: a carbohydrate binding domain of another protein, a signal peptide, an affinity tag or a protease cleavage site.

The fusion protein according to the present invention is not limited to the method by which it has been obtained but comprises fusion proteins of all kinds as long as they contain a component satisfying at least one of the above criteria. One option is to obtain the fusion protein according to the present invention by methods of molecular biology. As a person skilled in the art will know, methods for preparing nucleic acids encoding fusion proteins are standard methods of molecular biology which are described in greater detail, for example, in Sambrook *et al.* (Molecular cloning, a laboratory manual, 2nd edition, 1989, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York). In a preferred embodiment, the fusion protein besides the polypeptide with β -xylosidase activity according to the present invention may be fused to a carbohydrate binding domain of another protein. In a preferred embodiment, the polypeptide with β -xylosidase activity according to the present invention may be fused to a polypeptide selected from one or more of the following group: signal peptide, affinity tags, protease cleavage site. Translational coupling may be used to direct the expressed polypeptide according to the present invention to cellular compartments or organelles or to export it from the host cell. Signal peptides which are also called signal sequences are well known to a person skilled in the art and include the lead sequences of the periplasmatic proteins *OmpA*, *OmpT*, *PelB*, *PhoA*. Signal sequences for the export of proteins are found, for example, in naturally occurring secreted proteins, for example proteins with carbohydrate-modifying characteristics such as cellobiohydrolase I or II, endoglucanase, *AmyE*, and in *S. cerevisiae* Mfa or chicken egg lysozyme. Cleavage sites for proteases which are suitable as tags for recombinantly expressed proteins are well known to a person skilled in the art. Protease cleavage site means a poly- or oligopeptide comprising a peptide bond which may be specifically cleaved by a certain protease, and a recognition sequence which is usually near the cleavage site and is recognised by the corresponding protease. The protease cleavage sites that may be used in accordance with the present invention are not subject to any limitation. They expressly include the cleavage site of the tobacco etch virus (TEV) protease, the cleavage sites of the blood coagulation factors from mammals such as factor Xa or thrombin. The affinity tags that may be used according to the present invention are not

subject to any limitations. As a person skilled in the art knows, affinity tags that have proved to be advantageous for the purification of many polypeptides are described in detail, for example, in Sambrook *et al.* (Molecular cloning, a laboratory manual, 2nd edition, 1989, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York). Expressly included in the definition of the present invention is an oligohistidine tag as described in examples 1 and 5.

Enzyme mixture

A mixture of enzymes may also be used in accordance with the invention. This mixture contains the polypeptide with β -pyranosidase activity which satisfies at least one of the above criteria and comprises one, two or more additional enzymes. These additional enzymes may be selected from the group comprising pectinases, endoxylanases, β -glucosidases, β -glucanases, cellobiohydrolases, β -xylosidases, α -arabinofuranosidases, α -glucuronidases, acetyl xylan esterases. The enzyme mixture according to the invention thus comprises the polypeptide with β -pyranosidase activity which satisfies at least one of the criteria described above and one or more pectinases and/or one or more endoxylanases and/or one or more β -glucosidases and/or one or more β -glucanases and/or one or more cellobiohydrolases and/or one or more β -xylosidases and/or one or more α -arabinofuranosidases and/or one or more α -glucuronidases and/or one or more acetyl xylan esterases.

In a preferred embodiment, the enzyme mixture according to the invention comprises the polypeptide with β -pyranosidase activity which satisfies at least one of the criteria described above and one or more of α -arabinofuranosidases and/or one or more α -glucuronidases and/or one or more acetyl xylan esterases.

Unless specifically stated otherwise, the term "comprises" in the present application is used to indicate that further components may optionally be present in addition to the components listed under "comprises". However, it is considered a special embodiment that the term "comprises" includes the possibility that no other components are present, i.e. the term "comprises" under this special embodiment is understood to mean the same as "consists of".

Nucleic acid and vector

The present invention also relates to a nucleic acid encoding the polypeptides described above. One example for such a nucleic acid is shown by SEQ ID NO. 3. The nucleic acid according to the present invention may be a component of an expression cassette. A person

skilled in the art will be well aware of the typical components of an expression cassette; for example, they are described in more detail in Sambrook *et al.* (Molecular cloning, a laboratory manual, 2nd edition, 1989, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York). In addition to the coding sequence, the expression cassette typically contains at least: one promoter and one terminator. The expression level of a gene encoding a polypeptide according to the present invention may be adjusted by the number of copies of the gene introduced into the host cell; preferably more than one copy is present. The promoter may be adjusted for optimised expression so that it will, for example, respond to the addition of a chemical or the change of one or more physical parameters in such a manner that the induction of a gene may be switched on or off. Examples for inducible promoters include the tetracycline repressor system, the Lac-repressor system (Beneyx, 1999, Curr. Opin. Biotechnol. 10: 411: 422), systems inducible by copper ions (Hottinger et al., 2004, Yeast, 10: 283-296), systems inducible by methanol (Cereghino et al., 2000, FEMS Microbiol. Reviews 24: 45-66) or the temperature-inducible λ PL promoter. Alternatively, derepression of the promoter by reaching a favourable physiological state of the culture may be a useful strategy (promoters PhoA, Trp, Adh2, Fmdh, CBH1 (Price et al., 1990, Methods Enzymol., 185-308-318; Hollenberg, 1995, US 3,589,585). Methods to further increase the yield include the co-expression of one or more proteins involved in the translation, proteins involved in target control, proteins involved in folding (for example chaperones of the Hsp70-family, protein disulfide isomerases), or of proteases assisting correct processing.

The expression cassette may be integrated in a vector which is either propagated episomally in the host cell or is integrated in its genome. Examples of typical vectors are bacterial plasmids, yeast plasmids, centromer-containing linear DNA, constructs of viral origin such as SV40, phage DNA, baculovirus, vaccinia, adenovirus, chicken pox virus, pseudo-rabies and combinations of vectors of bacterial, eukaryotic and viral origin. Integration may be achieved by methods well known to a person skilled in the art of molecular biology, such as homologous recombination, transposition, or by using viral transfection systems. Also included are episomal systems for expression, one or more copies of which are integrated into the genome of the host cell either in a planned or unplanned manner. Moreover all vector systems are included which permit the heterologous expression of a polypeptide according to the invention in a host cell.

Host cell

The present invention also relates to a host cell transformed with the vector described above.

Preferred methods for introducing the vector construct into the host cell include transformation, transfection, conjugation and hybridisation. Transformation may be effected by electroporation, protoplast fusion, lipofection, ballistic bombardment, chemical transformation based on calcium chloride, polyethylene glycol (PEG) or manganese chloride. Other strategies include the application of viral particles. Another option is the use of naturally competent organisms as host cells.

The host cell according to the present invention may be a prokaryote or a eukaryote. A eukaryotic host cell is preferably selected from the groups consisting of *Saccharomyces cerevisiae*, *Yarrowia lipolytica*, *Schizosaccharomyces lactis*, *Kluyveromyces lactis*, *Pichia methanolytica*, *Pichia pastoris*, *Pichia angusta*, *Hansenula polymorpha*, *Aspergillus niger*, *Chrysosporium lucknowense*, *Trichoderma reesei*, *Penicillium* sp..

In a particularly preferred embodiment, the eukaryotic host cell is a methylotrophic yeast, preferably from the group including *Pichia methanolytica*, *Pichia pastoria*, *Pichia angusta*, *Hansenula polymorpha*.

A prokaryotic host cell according to the present invention is preferably selected from the group comprising *Bacillus* sp., *Bacillus subtilis*, *Bacillus licheniformis*, *Bacillus megaterium*, *Thermus thermophilus*, *Pseudomonas fluorescens*, *Fervidobacterium* sp., *Escherichia coli*.

After transformation of the host cell and growth to the desired cell density, the selected inducible promoter may be induced by a chemical activator or by changing one or more physical parameters so as to induce the cultivated host cells to produce the desired polypeptide.

Purification method

The present invention also relates to a method for purifying the above-described polypeptide with β -pyranosidase activity. In a particular embodiment, the polypeptide is not secreted by the host cell. After culturing, the host cells are then isolated and the polypeptide according to the present invention contained therein is isolated. Such a process is described in examples 3 to 5. In a particular embodiment, the polypeptide according to the present invention is produced as a fusion protein with at least one signal peptide which directs the recombinant protein to secretion from the host cell.

In detail, the process comprises the following steps:

- a) obtaining the host cell transformed with a vector as described above;
- b) cultivating the host cell under conditions where the polypeptide with β -pyranosidase activity is expressed,
- c) purifying the polypeptide with β -pyranosidase activity.

The host cell according to a) may be recovered by methods well known to a person skilled in the art, for example as described in the chapter "host cell".

The manner in which step b) is performed is influenced by many factors which are basically known to a person skilled in the art. Thus the growth medium, the growth temperature and other conditions of growth are selected depending on the kind of host cell used. The conditions under which the polypeptide is expressed further depend especially on the selection of the promoter used as described above. Different methods for harvesting host cells from a protein-expressing cell culture are well known to a person skilled in the art and include, for example, centrifugation of the medium containing the host cells and filtration of the medium containing the host cells. Then the cells are typically subjected to lysis, either by chemical or mechanical methods or by a combination of both.

The manner of performing step c) is selected depending on the nature of the expressed polypeptide. In particular embodiments, in which a fusion protein containing the polypeptide according to the present invention and another domain such as an affinity tag is expressed, the characteristics of the affinity tag may also be used to purify the fusion protein. One example for such an affinity tag is a tag comprising several, typically at least six, histidine residues as described in examples 3 to 5.

In a preferred embodiment, step c) of the process for purifying the polypeptide described above comprises heat precipitation. For this purpose, a mixture of different polypeptides comprising the desired polypeptide is brought to a certain temperature which is generally higher than the temperature at which the host cell grew during expression. This mixture may comprise two or more different polypeptides. Such a mixture may also be a lysate which is produced when the cells expressing the desired polypeptides are subjected to lysis, for example with mechanical or chemical methods or a combination of both, or an extract obtained or enriched from said lysate.

Depending on their properties, different proteins precipitate at different elevated temperatures. In this way, an enrichment in the in the cell lysate or the extract recovered

therefrom can be achieved. By further increasing the temperature, the desired protein may optionally be precipitated.

Use

The present invention also relates to the use of the polypeptide with β -pyranosidase activity or the mixture containing the polypeptide with β -pyranosidase activity described above for any purpose for which such enzyme activity is required or desired.

In particular, the present invention describes the use of the polypeptide with β -pyranosidase activity or the mixture containing the polypeptide with β -pyranosidase activity for degrading one or more substrates containing one or more β -xylopyranosidic and/or one or more β -glucopyranosidic bonds, hereinafter called β -xylopyranoside or β -glucopyranoside.

In a preferred embodiment, the present invention relates to the above-described use for the enzymatic degradation of biomass containing lignocellulose.

In a more preferred embodiment, the present invention relates to the above-described use in textile processing. In a particularly preferred embodiment, the present invention relates to the above-described use as an additive to detergents. In a further particularly preferred embodiment, the present invention relates to the above-described use in the food and/or feed industry. Finally, in a further particularly preferred embodiment, the present invention relates to a combination of two or more of these applications.

In a particularly preferred embodiment, the present invention relates to the above-described use for the production of fruit juices. In a more preferred embodiment, the present invention relates to the above-described use for the production of wine or beverages containing wine, and in a further more preferred embodiment, the present invention relates to the above-described use for the production of beer or beverages containing beer. In a further more preferred embodiment, the present invention relates to the above-described use for the production of oil, preferably olive oil, rapeseed oil or sunflower oil. Finally, in a further more preferred embodiment, the present invention relates to the above-described use for the production of baked goods.

Examples

Example 1: Identification and characterisation of β -pyranosidase FgXyl3A of *Fervidobacterium gondwanense*

For identification of cellulolytic and xylanolytic activity, a gene bank of *F. gondwanense* was screened. The gene bank was prepared according to the following method: Genomic DNA was purified from *Fervidobacterium gondwanense* cells using the Qiagen genomic DNA isolation kit (Qiagen GmbH, Hilden). The DNA obtained was degraded into clonable fragments of several kilo bases in length by partial digestion with *Sau3A*I; these fragments were ligated with lambda-ZAP-Express-Predigested (Stratagene™) vector arms and the ligation products then packed into phage particles as instructed by the manufacturer. The primary phage library was amplified in *E. coli* XL1-Blue MRF' cells and the phagemide library cut out by means of helper phages ExAssist and stably established after transfection in *E. coli* XL0LR.

For sample taking, bacteria of the *E. coli* strain XL0LR containing parts of the genome of *F. gondwanense* in the vector pBK-CMV (kanamycin resistance, phagemide vector with Lac-Promoter, Agilent Technologies, Waldbronn) were plated on an LB selection tray and incubated at 37°C over night. The colonies that had grown were transferred to a fresh medium and reincubated at 37°C over night.

For the identification of colonies with β -glucosidase activity, esculin (Sigma-Aldrich, Munich) was used as a substrate. For this purpose, the colonies were detected coated with esculin agar (0.1% (w/v) esculin, 0.01% (w/v) ammonium iron(III) citrate, 50 mM Na-acetate, 1% (w/v) agarose, pH 6.0). It was possible to detect activity from the formation of a brown halo around the colonies which is due to the decomposition of the esculin into glucose and esculetin. Esculetin forms a complex with iron ions which becomes visible by the brown colouring. Clones with enzymatic activity were isolated and cultivated.

Screening with the substrate esculin by means of colorimetric detection methods resulted in the identification of the gene bank clones Bgl13.

In order to determine the open reading frames (ORFs) encoding the active protein, the plasmid of clone Bgl13 was sequenced by means of primer walking using the standard primers T3 and T7.

T3	5'-ATTAACCCTCACTAAAGGGA-3'
T7	5'-TAATACGACTCACTATAGGG-3'

This standard method is described in more detail in Sambrook *et al.* (Molecular cloning, a laboratory manual, 2nd edition, 1989, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York).

The overall size of the inserts was 8479 base pairs (bp). For identification of open reading frames on the insert, the "ORF Finder" from the NCBI data base was used; five open reading frames were determined. So as to be able to allocate a possible function to the proteins encoded by the ORFs, their amino acid sequences were compared with known sequences by means of "blastp" from the NCBI-data base. BLAST represents the "Basic Local Alignment Search Tool" (Altschul *et al.*, 1997, Nucleic Acids Res 25 (17): 3389-3402). The amino acid sequence of ORF 1 displayed similarities with β -xylosidases of the glycoside hydrolase family 3 and, according to the respective nomenclature for glycoside hydrolases, ORF 1 was designated *xyl3A* and the corresponding protein *FgXyl3A* (Henrissat *et al.*, 1998, FEBS Lett. 425 (2): 352-354).

The gene *xyl3A* encoding the β -pyranosidase consists of 2328 bp and encodes the protein *FgXyl3A* with of a size of 775 amino acids. *FgXyl3A* has a calculated molecular weight of 85.9 kilodalton (kDa) and a theoretical isoelectric point (pI) of 5.55. A signal sequence could not be determined by "SignalP 3.0".

The amino acid sequence of *FgXyl3A* is shown in SEQ ID NO.1.

A sequence comparison of *FgXyl3A* with sequences of known β -xylosidases was performed with the aid of "blastp" from the NCBI data base. The amino acid sequence of *FgXyl3A* showed the highest congruence with a putative β -xylosidase of the glycoside hydrolase family 3 of *Thermotoga neapolitana* with an identity of 68 %. The sequence of *FgXyl3A* contains the preserved amino acids which are characteristic for proteins of the glycoside hydrolase family 3 (Zverlov *et al.*, 1997, Microbiology 143 (Pt 11): 3537-3542).

The domain structure of *FgXyl3A* was determined with the aid of "InterProScan" (<http://www.ebi.ac.uk/Tools/InterProScan/>). Accordingly, the protein *FgXyl3A* consists of two domains, namely of an N-terminal catalytic domain and of a C-terminal domain, both domains showing similarity to the glycoside hydrolase family 3 (Fig. 1).

SEQ ID NO. 1 contains a region which comprises the two structural domains determined by a sequence comparison with the protein BglB of *Thermotoga neapolitana* (Pozzo et al., 2010, J. Mol. Biol., 2:397 (3), 724-739). Structural domain 1 contains the amino acid residues Leu13 to Ile381 and Gln586 to Tyr638 of SEQ ID NO.1 and was determined by a sequence comparison by means of InterProScan with SEQ ID NO. 4. A conserved aspartate residue is present there (Asp281) which was identified as a catalytic nucleophil in glycoside hydrolases of the family 3 (Zverlov *et al.*, 1997, Microbiology 143 (Pt 11): 3537-3542; Wulff-Strobel & Wilson, 1995, J. Bacteriol. 177 (20): 5884-5890).

Structural domain 2 contains the amino acid residues Val382 to Val585 of SEQ ID NO.1 and was determined by a sequence comparison by means of InterProScan with SEQ SD NO. 4. SEQ ID NO. 2 is the region of SEQ ID NO. 1 which includes both structural domains and hence comprises the amino acid residues 13 to 638 of SEQ ID NO. 1.

The underlying nucleotide sequence of the full-length protein *FgXyl3a* is given by SEQ ID NO.3.

Example 2: Cloning the β -pyranosidase *FgXyl3A* of *Fervidobacterium gondwanense*

For cloning *xyl3A*, the gene was amplified by means of the following oligonucleotides by a polymerase chain reaction (PCR) according to Sambrook *et al.* (Molecular cloning, a laboratory manual, 2nd edition, 1989. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York):

<i>xyl3A</i> _BamHIF	5'-GGGGATCCGAGATATATAAGGATTCTTC-3'
<i>xyl3A</i> _HindIII.R	5'-GGAAGCTTTTAGAAAGTGTAAGTCTTTG-3'

The oligonucleotides contained a *Bam*HI cleavage site at the 5' terminus and a *Hind*III cleavage site at the 3' terminus. The PCR was performed in a thermal cycler (Gene Amp PGR System 2400, Perkin-Elmer, Massachusetts, USA) and the PCR products purified with the "High Pure PCR Product Purification Kit" (Roche Diagnostics, Mannheim) as instructed by the manufacturer. The amplified gene was ligated into the vector pJET1.2 ("CloneJET™ PCR Cloning Kit" (Fermentas, St. Leon-Rot)) with the aid of the 3'A overhangs formed by high fidelity polymerase in accordance with the instructions of the manufacturer. By restriction with the endonucleases *Bam*HI and *Hind*III, the gene *xyl3A* was cut from the

vector and ligated into the vector pQE-30 cut with the same restriction enzymes (ampicillin resistance, expression vector with a sequence for an N-terminal 6x histidine tag with T5-Promotor, Qiagen, Hilden). Correct cloning of *xyl3A* into the vector pQE-30 was verified by sequencing. For gene expression, the recombinant plasmid pQE-30::*xyl3A* (Fig. 2) was transformed into the *E. coli* strain M15[pREP4] (Nal^s Str^s Rif^s Thi⁻ Lac⁻ Ara⁺ Gal⁺ Mtl⁻ F⁻ RecA⁺, Uvr⁺ Lon⁺ pREP4, Qiagen Hilden). By cloning *xyl3A* into the pQE-30 vector a sequence encoding a His tag was attached to the 5' terminus of the gene. This results in a size of 786 amino acids and a calculated molecular weight of 87.2 kDa for the recombinant protein FgXyl3A.

Example 3: Expression of Fg*xyl3A* in *Escherichia coli* (*E. coli*)

The recombinant clone of *Escherichia coli* (*E. coli*) M15/pQ E-30::*xyl3A* (example 2) was incubated at 37°C. Cell cultivation of the *E. coli* strains was performed in LB-Medium (10 g/l of trypton, 5 g/l of yeast extract, 10 g/l of NaCl, pH 7.0) according to Sambrook *et al.* (Molecular cloning, a laboratory manual, 2nd edition, 1989, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York).

The cultivation of strains with plasmid- or genome-coded resistance against antibiotics was performed under selection pressure by adding the appropriate antibiotic. The concentration of antibiotics used corresponded to the recommendations of Sambrook *et al.* (Molecular cloning, a laboratory manual, 2nd edition, 1989, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York). The cultivation of the cells was carried out aerobically at 37°C on a shaker (Certomat R, B. Braun Biotech International, Melsungen) at 160 rpm. The growth of the cells was determined on the basis of the optical density (OD) at a wavelength of 600 nm in a spectrophotometer (UV-1602, Shimadzu Deutschland, Duisburg). The induction of the gene expression of recombinant *E. coli* strains was performed by adding 1 mM of isopropyl- β -D-thiogalactopyranoside (IPTG) at an OD_{600nm} of 0.6-0.8. This was followed by the incubation of the cells at 37°C and 160 rpm for 4.5 hrs and subsequent centrifugation of the cells in a Sorvall RC 5C Plus centrifuge (Thermo Fisher Scientific, Langenselbold) for 20 minutes at 4221 x g and 4°C, obtaining about 2.0 g of cells (humid weight) of 500 ml culture volume. The pellet was used directly for the preparation of cell-free crude extract (example 4) or stored at -20 °C.

Example 4: Preparation of a crude extract of the recombinant *E. coli* clone M15/pQE-30::xyI3A

For obtaining a crude extract, the cell pellet (example 3) was washed twice in lysis buffer (50 mM of NaH₂PO₄, 300 mM of NaCl, 10 mM of imidazole, pH 8.0), with 20 minutes of centrifugation at 4500 x g and 4°C each time. The pellet was then taken up in a cell lysis buffer, using 5 ml of lysis buffer per 1 g of cells. This was followed by cell lysis by means of ultrasound (Branson Sonifier®, Danbury, Connecticut, USA), wherein 3 cycles were performed for 5 minutes each on ice (output control: 50 %, duty cycle: 5). Cell debris was then separated by centrifugation for 20 minutes at 20200 x g and 4°C and the supernatant transferred into a new vessel. Optionally, heat precipitation of the crude extract was performed at 60°C for 15 minutes and the precipitated protein separated by centrifugation for 20 minutes at 20200 x g and 4°C. The crude extract was stored at 4°C.

The crude extract was tested for β -xylosidase activity by a standard test (example 7) and used for purification (example 5).

Example 5: Purification of FgXyl3A

The multi-step purification of FgXyl3A by (1) heat precipitation, (2) affinity chromatography and (3) gel filtration resulted in the almost complete purification of the protein.

- (1) Heat precipitation: After the first purification step by means of heat precipitation (example 4), a yield of 41.7 % was obtained (Table 1).
- (2) Affinity chromatography: The further purification of FgXyl3A with an attached His tag (examples 2, 3 and 4) was achieved through affinity chromatography by means of Ni²⁺ nitrilotriacetic acid (Ni²⁺-NTA, Qiagen, Hilden). An AKTA™ purifier system (GE Healthcare, Munich) with the following components was used for performing the affinity chromatography: monitor UV-900, pump P-900, in-line mixer M-925, motor valve INV-907 and fraction collector Frac-950. The buffers used were filtered before use (0.45 μ m cellulose mixed ester filter, Whatman, Dassel) and the protein samples centrifuged for 10 minutes at 13000 x g and 4°C. After equilibration of the column matrix with 5 ml column volumes of cell lysis buffer, the crude extract or heat-precipitated crude extract was loaded onto the column at a flow rate of 1 ml/minute. Unbound proteins were removed by washing of the column matrix with 3 column volumes of washing buffer (50 mM of NaH₂PO₄, 300 mM of NaCl, 20 mM

imidazole, pH 8.0) at a flow rate of 2 ml/minutes. The elution of bound proteins was performed by rinsing the column with 5 column volumes of elution buffer (50 mM NaH₂PO₄, 300 mM of NaCl, 250 mM imidazole, pH 8.0) at a flow rate of 2 ml/minute. The fraction size of the samples collected was 2 ml. The fractions containing the target protein were combined, concentrated with the aid of micro-concentrators (Amicon Ultra-15 centrifugal filter unit, 10 kDa cut-off size, Milipore, Schwalbach) and used for further purifications steps or, respectively, subjected to dialysis and used for biochemical characterisation.

This further purification by means of Ni²⁺-NTA agarose resulted in a marked loss in yield which was 1.13 %.

- (3) Gel filtration: For the further purification of *FgXyl3A* with an attached His tag (examples 2, 3 and 4) by means of gel filtration, an AKTA™ Fast Protein Liquid Chromatography (FPLC) plant (GE Healthcare, Munich) with the following components was used: monitor UPC-900, pump P-920, in-line mixer M-925, engine valve INV-907 and fraction collector Frac-950. The purification by means of gel filtration was performed by a HiLoad 16/60 Superdex™ 200 prep grade column (GE Healthcare, Munich). The protein samples were centrifuged for 10 minutes at 13000 x g and 4°C before use and the used buffers filtered (0.45 µm cellulose mixed ester filter, Whatman™, Dassel). At a flow rate of 1 ml/minute the column was loaded with 1 ml of the protein sample maximum and rinsed with 1.5 column volumes of 50 mM Na₂HPO₄ and 150 mM of NaCl (pH 7.0). Fractions of a size of 2 ml were collected. The fractions containing the target protein were combined and concentrated with the aid of micro-concentrators (Amicon Ultra-15 centrifugal filter unit, 10 kDa cut-off size, Milipore, Schwalbach). Subsequent desalting and the change of buffers of protein solutions was achieved by dialysis. For this purpose, dialysis tubing (Membra-Cel MWCO 3500, Serva, Heidelberg) was boiled in the buffer used and then loaded with the protein sample. Dialysis was performed in 50 mM Na-acetate, pH 6.0 and, respectively, pH 6.5 over night at 4°C with stirring against the 100-fold volume of the protein solution. After dialysis, the samples were used for biochemical characterisation. As a result of the last purification step by gel filtration a yield of 0.18 % and a purification factor of 95 was obtained. The specific activity was 20.9 U/mg vis-à-vis pNP-β-xylopyranoside (Table 1).

The purification of histidine-tagged *FgXyl3a* is shown in Fig. 3 and Table 1.

Purification step	Total protein [mg]	Overall activity [U]	Specific activity [U/mg]	Yield (%)	Purification factor [x-fold]
Crude extract	1263.4	273.7	0.22	100	1
Heat precipitation	345.9	114.3	0.33	41.7	1.5
Ni ²⁺ -NTA	3.1	3.1	0.38	1.13	1.7
Gel filtration	0.015	0.32	20.9	0.18	95

The determination of protein concentrations was performed according to Bradford (1976, Anal Biochem 72: 248-254). For this purpose, 1 ml of Bradford reagent (0.007% (w/v) Serva Blau G-250, 5% (v/v) of ethanol, 8.5 % (v/v) of H₃PO₄) was mixed with 10 µl of the protein sample and incubated at room temperature for 5 minutes. The absorbance was determined at the wavelength $\lambda = 595$ nm by spectrophotometry. Bovine serum albumin (BSA) in concentrations of 0.1-1.0 mg/ml was used for preparing calibration lines.

Example 6: Determination of the molecular weight of *FgXyl3A*

The determination of the molecular weight of *FgXyl3A* was performed by means of denaturing and native polyacryl amide gel electrophoresis.

The denaturing separation of proteins was performed by means of sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) according to Laemmli (1970, Nature 227 (5259): 680-685). "Unstained Protein Molecular Weight Marker" or, respectively, "PageRuler™ Unstained Protein Ladder" by Fermentas (St. Leon-Rot) was used as the protein standard. The gel run was carried out in a "Minutenti-Protean Tetra" system (Bio-Rad, Munich) in 25 mM of tris, 192 mM of glycine, 0.1% (w/v) SDS at constant 200 V. The gels were incubated for 30 minutes in Coomassie dye solution (0.25 % (w/v) Serva Blau G-250, 45 % (v/v) ethanol, 10 % (v/v) acetic acid) and then decoloured in a decolouring solution (30 % (v/v) ethanol, 10 % (v/v) acetic acid). The gels were incubated in a drying solution (20 % (v/v) ethanol, 2% (v/v) glycerine) for 10 minutes for the purpose of preservation and dried with the aid of the "DryEase® Minutenti-Gel Drying System" (Invitrogen, Karlsruhe).

The determination of the molecular weight of native proteins and native protein complexes was performed by native polyacryl amide gel electrophoresis. 4-12 % (w/v) tris-glycine gradient gels (Anamed Electrophorese, Groß-Bieberau) were used for this purpose. Before the run, native sample buffer (7.5 % (v/v) glycerine, 0.04 % (w/v) of bromophenol blue, 33 mM tris, pH 6.8) was added to the protein sample. "High Molecular Weight-Marker (GE

Healthcare, Munich) served as the protein standard. The gel run was performed in a "Novex® XCell II™ Minutenti-Cell" facility (Invitrogen, Karlsruhe) for 3 hrs at constant 120 V in 25 mM tris and 192 mM glycine. The gels were incubated in Coomassie dye solution, decoloured and dried for preservation.

Under denaturing conditions, the recombinant protein had a molecular weight of 87 kDa (Fig. 4 A). Under native conditions, a molecular weight of 529,7 kDa was determined (Fig. 4 B). This indicates that the protein *FgXyl3A* might be present as a homohexamer.

Example 7: Detection of the activity of *FgXyl3A* in SDS gels

The detection of enzymatic activity with the aid of zymograms was performed after the electrophoretic separation of the proteins by SDS-PAGE (example 6). After the run, the gels were incubated in 1% (v/v) Triton™ X 100 at room temperature (RT) for 1 hr. The β -xylosidase activity of *FgXyl3A* was detected by incubation of the gels in 0.1 % (w/v) esculin, 0.01 % (w/v) ammonium iron(III) citrate and 50 mM Na-acetate (pH 6.0) for 30-60 minutes at 60 °C.

Example 8: Examination of the enzymatic activity of *FgXyl3A*

For the determination of the substrate spectrum of *FgXyl3A*, the activity of the purified recombinant proteins (example 5) vis-à-vis pNP- β -D-galactopyranoside, pNP- α -D-glucopyranoside, pNP- β -D-glucopyranoside, pNP- β -D-cellobioside and pNP- β -D-xylopyranoside was examined. The substrates were used in a final concentration of 2 mM and the activities were measured by the following standard tests.

(1) Determination of the β -glucosidase activity: The activity of the β -glucosidase was determined in a modified manner according to Park *et al.* (2005, Appl. Microbiol. Biotechnol. 69 (4): 411-422) using pNP- β -D-glucopyranoside (Sigma-Aldrich, Munich) as the substrate. Unless otherwise indicated, a reaction sample contained 2 mM pNP- β -D-glucopyranoside and 50 mM Na-acetate (pH 6.0) in an overall volume of 1 ml. Before adding 0,02 U of the enzyme, the reaction samples were incubated at 80°C for 5 minutes. The reactions were performed at 80°C for 10 minutes and then stopped by adding 100 μ l of 0.1 M Na₂CO₃ and placing the sample on ice. All the measurements were taken in the form of triple determinations. The enzyme was added to the control batches only after Na₂CO₃ had been added and incubation on ice performed. The determination of the released p-nitrophenol was performed at a

wavelength of $\lambda = 410$ nm. The enzymatic activity was calculated by the following formula:

$$\text{Enzymatic activity (U/ml)} = (\Delta E \cdot V_{\text{Tot}}) / (d \cdot e \cdot t \cdot V_E)$$

ΔE is defined as the change in extinction, V_{Tot} as the reaction volume (ml), d as the thickness of the cuvette (cm), e as the molar extinction coefficient ($\text{M}^{-1} \text{cm}^{-1}$), t as the reaction time (minutes) and V_E as the volume of the enzyme sample (ml). $16.56 \text{ M}^{-1} \text{cm}^{-1}$ is used as the molar extinction coefficient. 1 U β -glucosidase activity was defined as the amount of enzyme which releases 1 μmol of p-nitrophenol per minute under standard conditions.

- (2) Determination of the β -xylosidase activity: pNP- β -D-Xylopyranoside (Sigma-Aldrich, Munich) was used as the substrate for determining the activity of β -xylosidase in the modified form according to Park *et al.* (2005, Appl. Microbiol. Biotechnol. 69 (4): 411-422). Unless otherwise described, the standard reaction to determine the β -xylosidase activity was carried out in the presence of 2 mM pNP- β -D-xylopyranoside and 50 mM Na-acetate (pH 6.5) in a total volume of 1 ml. The reactions were started by adding 0.016 U enzyme to the batches pre-incubated at 85 °C for five minutes and then carried out at a temperature of 85°C for 10 minutes. The reactions were stopped by adding 100 μl 0,1 M Na_2CO_3 and placing the samples on ice. All the measurements were taken in the form of triple determinations. Reaction samples to which the enzyme was added only after addition of Na_2CO_3 and incubation on ice served as controls. The spectrophotometric measurement of the released p-nitrophenols was performed at a wavelength of $\lambda = 410$ nm. The enzymatic activity was calculated by the following formula:

$$\text{Enzymatic activity (U/ml)} = (\Delta E \cdot V_{\text{tot}}) / (d \cdot e \cdot t \cdot V_E)$$

ΔE is defined as the change in extinction, V_{Tot} as the reaction volume (ml), d as the thickness of the cuvette (cm), e as the molar extinction coefficient ($\text{M}^{-1} \text{cm}^{-1}$), t as the reaction time (minutes) and V_E as the volume of the enzyme sample (ml). $16.56 \text{ M}^{-1} \text{cm}^{-1}$ was used as the molar extinction coefficient. 1 U β -xylosidase activity was defined as the amount of enzyme which led to a release of 1 μmol of p-nitrophenol per minute under standard conditions.

FgXyl3A showed the highest specific activity with 20.9 U/mg vis-à-vis pNP- β -xylopyranoside (Table 2). In addition, a specific activity of 1.3 and, respectively 9.3 U/mg vis-à-vis pNP- α - and pNP- β -glucopyranoside was determined. The substrate pNP- β -galactopyranoside was also hydrolysed by the β -xylosidase, the specific activity being 2.7 U/mg. Hydrolysis of the disaccharide pNP- β -cellobioside by *Xyl3A* was not detected (Table 2).

Substrate	Specific activity [U/mg]
pNP- β -xylopyranoside	20.9
pNP- β -glucopyranoside	9.3
pNP- α -glucopyranoside	1.3
pNP- β -galactopyranoside	2.7
pNP- β -cellobioside	0

Table 2: Substrate specificity of *FgXyl3A*

Example 9: Influence of the temperature on the activity of *FgXyl3A*

For determination of the temperature profile of *FgXyl3A*, the enzymatic activity of the recombinant proteins (example 5) was measured in a temperature range of 10-115°C with pNP- β -xylopyranoside as the substrate as described in example 8. For this purpose, the reaction batches were pre-incubated at the relevant temperature for 5 minutes before adding the enzyme. Activity was determined by the standard test (example 8).

FgXyl3A showed β -xylosidase activity in a wide temperature range (Fig. 5). More than 50 % of the relative activity was measured at temperatures of 70-95°C, maximum activity being measured at 85°C. A pronounced decrease in activity was observed at temperatures above 95°C; for example, the activity of *FgXyl3A* at 100°C was 23 % of the maximum activity. 18-38 % of the relative activity was measured in a temperature range of 50-65°C.

Example 10: Temperature stability of *FgXyl3A*

For the determination of temperature stability, recombinant *FgXyl3A* (example 5) was incubated at temperatures of 60-90°C for up to 24 hrs. The residual activity was determined under standard conditions (example 8).

Measured by the residual β -xylosidase activity, *FgXyl3A* was stable for a period of 3 hours at a temperature of 60°C. Incubation at 70-90°C, on the other hand, caused deactivation of the enzyme after a few minutes (Fig. 6). By pre-incubation of *FgXyl3A* at 60 °C, 65 % of residual activity was determined after 2 hrs; after 6 hrs of incubation the activity was 30 %. After a longer incubation of 24 hrs, the enzyme exhibited a residual activity of 25 %. At a temperature of 70°C, *FgXyl3A* exhibited 30 % of residual activity after 5 minutes of incubation; and 15 % residual activity after 10 minutes of pre-incubation. The half-life of *FgXyl3A* was 215 minutes at 60°C and 6 minutes at 70°C.

Example 11: Influence of the pH-value on the activity of *FgXyl3A*

For the determination of the pH profile of *FgXyl3A*, the activity of the recombinant enzyme (example 5) in a pH range of 2.0-11.0 with pNP- β -xylopyranoside as the substrate was measured (example 8). The reactions were carried out in the presence of 50 mM of universal buffer (16.7 mM H₃PO₄, 16.7 mM of acetic acid, 16.7 mM of H₃BO₄, Britton & Robinson, 1931) under standard conditions (example 8). The maximum activity of *FgXyl3A* vis-à-vis β -xylosidase was measured at a pH value of 6.5 (Fig. 7). In the pH range of 5.5-7.0, the β -xylosidase exhibited more than 45 % of the relative activity, a pronounced decrease in activity being observed at pH values above 7.0. In the acidic pH range, the enzyme did not exhibit any activity at pH 2.0-3.0 and at pH 4.0 exhibited a maximum activity of 6 %. A relative activity of 36 % was measured at pH 5.0.

Example 12: Determination of the pH stability of *FgXyl3A*

For determination of the pH stability, recombinant *FgXyl3A* (example 5) was incubated in 50 mM of universal buffer with pH-values of 3.0-10.0 for 48 hrs at room temperature. This was followed by the determination of the residual activity under standard conditions (example 8), but in the presence of 50 mM of universal buffer having a pH-value of 6.5.

In the slightly acidic and neutral pH range of 5.0-7.0, *FgXyl3A* exhibited high stability with a residual activity vis-à-vis β -xylosidase of 90-95 % (Fig. 8). In the acidic pH range, on the other hand, a loss in activity was observed. For example, the residual activity was 66 % at pH 3.0 and 4.0. A small loss in activity was observed in the alkaline pH range, where residual activity of 87 % was measured at pH 8.0, 81 % at pH 9.0 and 80 % at pH 10.0.

Example 13: Product analysis by high-performance liquid chromatography (HPLC)

For detection of the hydrolysis products formed, 0.5 % (w/v) of beech wood xylan, xylobiose, xylotriose and xylotetraose were incubated in 50 mM NaH₂PO₄ (pH 6.5) with 0.1 U β -xylosidase for up to 3 hrs at 80 °C. The samples were boiled for 10 minutes to stop the reaction and then centrifuged for 10 minutes at 13000 x g and 4°C. The supernatants of the samples were filtered (0.2 μ m pore size, Pall, Darmstadt) and transferred into sample bottles with micro-inserts (CS Chromatographie Service, Langerwehe). 20 μ l of each sample was injected. The analysis of the products was performed using an Aminex HPX 42-A column (Bio-Rad, Munich), using degassed H₂O (LiCrosolv®, Merck, Darmstadt) as the mobile phase and selecting a flow rate of 0.6 ml/minute at a temperature of 70°C.

After incubation of the enzyme for 3 hrs in the presence of xylobiose, a major part of the substrate was hydrolysed to form xylose (Fig. 9). The incubation of *FgXyl3A* with xylotriose resulted in the formation of xylose and xylobiose, a large proportion part of the substrates being hydrolysed (Fig. 10). If *FgXyl3A* was incubated with the substrate xylotetraose, the hydrolysis products xylose, xylobiose and xylotriose were detected, xylose being the main product (Fig. 11). Longer incubation did not result in an increased degradation of xylotetraose. If the *FgXyl3A* was incubated in the presence of beech wood xylan, small amounts of xylose were detected as the hydrolysis product.

Example 14: Kinetics of *FgXyl3A*

The determination of the kinetic parameters of the β -xylosidase activity of *FgXyl3A* was performed at substrate concentrations of 0.01-6.0 mM pNP- β -D-xylopyranoside. The reaction was performed under standard conditions (example 8). The determination of $v_{\max}$ and K_M was done by non-linear regression according to Michaelis-Menten and yielded a $v_{\max}$ value of 0.53 μ mol/min and a K_M value of 0.06 mM (Fig. 12).

Overview of the sequences in the sequence listing:

SEQ ID NO.1: *FgXyl3A* full-length protein

SEQ ID NO.2: catalytic domain GHF 3 from SEQ ID NO.1 (corresponds to amino acid residues 13 to 638 of SEQ ID NO.1)

SEQ ID NO. 3: DNA sequence encoding SEQ ID NO. 1

SEQ ID NO. 4: barley β -D-glucan-exohydrolase isoenzyme Exo I, full length

Short description of the drawings

Fig. 1: Domain structure of *FgXyl3A*

Fig. 1 shows the domain structure β -xylosidase *FgXyl3A* determined by means of "InterProScan". GHF: glycoside hydrolase family.

Fig. 2: Vector map of the recombinant plasmid pQE-30::*xyl3A*

The *xyl3A* gene encoding the β -xylosidase *FgXyl3A* gene was ligated into the expression vector pQE-30 using the restriction cleavage sites *Bam*HI and *Hind*III. T5: T5 promoter, *bla*: β -lactamase, ColE1: replication origin.

Fig. 3: Purification and detection of the activity of *FgXyl3A*

The purification and activity of *FgXyl3A* was detected by means of SDS-PAGE. The proteins were dyed with Coomassie-Blau; the activity of *FgXyl3A* was detected by esculin. Track 1: Protein standard "PageRuler™ Unstained Protein Ladder" (Fermentas, St. Leon-Rot), Track 2: Crude extract of *E. coli* M15/pQE-30 (control, 35 μ g), track 3: crude extract of *E. coli* M15/pQE-30::*xyl3A* (37 μ g), track 4: heat-precipitated crude extract of *E. coli* M15/pQE-30::*xyl3A* (35 μ g), track 5: Ni²⁺-NTA fraction of *FgXyl3A* (40 μ g), track 6: gel filtration fraction of *FgXyl3A* (1,1 μ g), track 7: activity gel of the gel filtration fraction of *FgXyl3A* (1.1 μ g).

Fig. 4: Determination of the molecular weight of *FgXyl3A*

The molecular weight of *FgXyl3A* was determined by means of denaturing and native polyacryl amide gel electrophoresis. In the Figure the logarithm of the molecular weight is plotted against the running distance of the proteins in the gel (Rf value). Molecular weights of the standard proteins in (A): Recombinant proteins having a size of 200, 150, 120, 100, 85, 70 and 60 kDa. Molecular weights of the standard proteins in (B): thyroglobulin (669 kDa), ferritin (440 kDa), catalase (232 kDa), lactate dehydrogenase (140 kDa), albumin (66 kDa).

Fig. 5: Temperature profile of *FgXyl3A*

The activity of *FgXyl3A* was determined in a temperature range of 10-115°C under standard conditions. Reaction parameters: 2 mM pNP- β -xylopyranoside, 50 mM Na-acetate, pH 6.5; 10 min incubation time.

Fig. 6: Temperature stability of *FgXyl3A*

The temperature stability of *FgXyl3A* was determined after pre-incubation of the enzyme at temperatures of 60-100°C. Reaction parameters: 2 mM pNP- β -xylopyranoside, 50 mM Na-acetate, pH 6.5, 85°C; 10 min incubation time.

Fig. 7: pH-Profile of *FgXyl3A*

The activity of *FgXyl3A* was determined in a pH range of 2.0-11.0 under standard conditions. Reaction parameters: 2 mM pNP- β -xylopyranoside, 50 mM universal buffer, 85°C; 10 min incubation time.

Fig. 8: pH Stability of *FgXyl3A*

The pH stability of *FgXyl3A* was determined after pre-incubation of the enzyme at pH values of 3.0-10.0 under standard conditions. Reaction parameters: 2 mM pNP- β -xylopyranoside, 50 mM universal buffer, 85°C; 10 minutes incubation time.

Fig. 9: Hydrolysis of xylobiose by *FgXyl3A*

The hydrolysis products were analysed after incubation of *FgXyl3A* for 3 hrs in the presence of 0.5% (w/v) xylobiose, 20 mM Na-phosphate buffer, pH 6.5, 80°C by HPLC.

Fig. 10: Hydrolysis of xylotriose by *FgXyl3A*

The hydrolysis products were analysed after incubation of *FgXyl3A* for 3 hrs in the presence of 0.5% (w/v) xylotriose, 20 mM Na-phosphate buffer, pH 6.5, 80°C by HPLC.

Fig. 11: Hydrolysis of xylotetraose by *FgXyl3A*

The hydrolysis products were analysed after incubation of *FgXyl3A* for 3 hrs in the presence of 0.5% (w/v) xylotetraose, 20 mM Na-phosphate buffer, pH 6.5, 80°C by HPLC.

Fig. 12: Kinetics of *FgXyl3A*

The figure shows the Michaelis-Menten plot of *FgXyl3A*. The activity of *FgXyl3A* was determined under standard conditions. Reaction parameters: 0.01-6.0 mM of pNP- β -xylopyranoside, 50 mM of Na-acetate, pH 6,5, 85 °C; 10 minutes incubation time.

CLAIMS

1. A polypeptide with β -pyranosidase activity, said polypeptide comprising an amino acid sequence having at least 71 % sequence identity with SEQ ID NO: 2.
2. A polypeptide with β -pyranosidase activity, said polypeptide comprising an amino acid sequence having at least 75 % sequence identity with SEQ ID NO: 2.
3. A polypeptide with β -pyranosidase activity, said polypeptide comprising an amino acid sequence having at least 80 % sequence identity with SEQ ID NO: 2.
4. A polypeptide with β -pyranosidase activity, said polypeptide comprising an amino acid sequence having at least 85 % sequence identity with SEQ ID NO: 2.
5. A polypeptide with β -pyranosidase activity, said polypeptide comprising an amino acid sequence having at least 70 % sequence identity with SEQ ID NO: 1.
6. A polypeptide with β -pyranosidase activity, said polypeptide comprising an amino acid sequence having at least 75 % sequence identity with SEQ ID NO: 1.
7. A polypeptide with β -pyranosidase activity, said polypeptide comprising an amino acid sequence having at least 80 % sequence identity with SEQ ID NO: 1.
8. A polypeptide with β -pyranosidase activity, said polypeptide comprising an amino acid sequence having at least 85 % sequence identity with SEQ ID NO: 1.
9. A polypeptide with β -pyranosidase activity, said polypeptide comprising the following amino acid sequences:

- (1) an amino acid sequence having at least 70 % sequence identity with the structural domain 1 of *FgXyl3a* consisting of the region comprising the amino acid residues 13 to 381 and 586 to 638 from SEQ ID NO: 1, and
 - (2) an amino acid sequence which is a structural domain 2 of a glycoside hydrolase of the family 3 (GHF3).
10. A polypeptide with β -pyranosidase activity, said polypeptide comprising the following amino acid sequences:
 - (1) an amino acid sequence having at least 75 % sequence identity with the structural domain 1 of *FgXyl3a* consisting of the region comprising the amino acid residues 13 to 381 and 586 to 638 from SEQ ID NO: 1, and
 - (2) an amino acid sequence which is a structural domain 2 of a glycoside hydrolase of the family 3 (GHF3).
11. A polypeptide with β -pyranosidase activity, said polypeptide comprising the following amino acid sequences:
 - (1) an amino acid sequence having least 80 % sequence identity with the structural domain 1 of *FgXyl3a* consisting of the region comprising the amino acid residues 13 to 381 and 586 to 638 from SEQ ID NO: 1, and
 - (2) an amino acid sequence which is a structural domain 2 of a glycoside hydrolase of the family 3 (GHF3).
12. A polypeptide with β -pyranosidase activity, said polypeptide comprising the following amino acid sequences:

- (1) an amino acid sequence having at least 85 % sequence identity with the structural domain 1 of *FgXyl3a* consisting of the region comprising the amino acid residues 13 to 381 and 586 to 638 from SEQ ID NO: 1, and
 - (2) an amino acid sequence which is a structural domain 2 of a glycoside hydrolase of the family 3 (GHF3).
13. A polypeptide with β -pyranosidase activity, said polypeptide comprising the following amino acid sequences:
 - (1) an amino acid sequence which is a structural domain 1 of a glycoside hydrolase of the family 3 (GHF3), and
 - (2) an amino acid sequence having at least 70 % sequence identity with structural domain 2 of *FgXyl3a* consisting of the region comprising the amino acid residues 382 to 585 from SEQ ID NO: 1.
14. A polypeptide with β -pyranosidase activity, said polypeptide comprising the following amino acid sequences:
 - (1) an amino acid sequence which is a structural domain 1 of a glycoside hydrolase of the family 3 (GHF3), and
 - (2) an amino acid sequence having at least 75 % sequence identity with structural domain 2 of *FgXyl3a* consisting of the region comprising the amino acid residues 382 to 585 from SEQ ID NO: 1.
15. A polypeptide with β -pyranosidase activity, said polypeptide comprising the following amino acid sequences:

- (1) an amino acid sequence which is a structural domain 1 of a glycoside hydrolase of the family 3 (GHF3), and
 - (2) an amino acid sequence having at least 80 % sequence identity with structural domain 2 of *FgXyl3a* consisting of the region comprising the amino acid residues 382 to 585 from SEQ ID NO: 1.
16. A polypeptide with β -pyranosidase activity, said polypeptide comprising the following amino acid sequences:
 - (1) an amino acid sequence which is a structural domain 1 of a glycoside hydrolase of the family 3 (GHF3), and
 - (2) an amino acid sequence having at least 85 % sequence identity with structural domain 2 of *FgXyl3a* consisting of the region comprising the amino acid residues 382 to 585 from SEQ ID NO: 1.
17. A polypeptide according to any one of claims 1 to 16, wherein the β -pyranosidase activity is selected from β -xylopyranosidase activity, β -glucopyranosidase activity and a combination of these two activities.
18. A polypeptide according to any one of claims 1 to 17, with β -xylopyranosidase activity.
19. A polypeptide according to any one of claims 1 to 18, capable of hydrolysing at least one of the β -glycoside bonds contained in the following substrates: xylobiose, xylotriose, xylotetraose, and xylan.

20. A polypeptide according to any one of claims 1 to 19, the activity of which vis-à-vis pNP- β -cellobioside is 5 % or less compared with the activity vis-à-vis pNP- β -xylobioside.
21. A polypeptide according to any one of claims 5 to 20, which has maximum enzymatic activity vis-à-vis pNP- β -xylopyranoside in a range from pH 4.0 to pH 8.0.
22. A polypeptide according to any one of claims 5 to 20, which has maximum enzymatic activity vis-à-vis pNP- β -xylopyranoside in a range from pH 5.5 to pH 7.0.
23. A polypeptide according to any one of claims 5 to 20, which has maximum enzymatic activity vis-à-vis pNP- β -xylopyranoside in a range from pH 6.2 to pH 6.8.
24. A polypeptide according to any one of claims 5 to 23, of which the maximum enzymatic activity vis-à-vis pNP- β -xylopyranoside after 48 hours of incubation at pH 9.0 is at least 50 % of the activity of a second fraction of said polypeptide which was not incubated at a pH of 9.0.
25. A polypeptide according to any one of claims 5 to 23, of which the maximum enzymatic activity vis-à-vis pNP- β -xylopyranoside after 48 hours of incubation at pH 9.0 is at least 60 % of the activity of a second fraction of said polypeptide which was not incubated at a pH of 9.0.
26. A polypeptide according to any one of claims 5 to 23, of which the maximum enzymatic activity vis-à-vis pNP- β -xylopyranoside after 48 hours of incubation at pH 9.0 is at least 70 % of the activity of a second fraction of said polypeptide which was not incubated at a pH of 9.0.

27. A polypeptide according to any one of claims 5 to 23, of which the maximum enzymatic activity vis-à-vis pNP- β -xylopyranoside after 48 hours of incubation at pH 9.0 is at least 80 % of the activity of a second fraction of said polypeptide which was not incubated at a pH of 9.0.
28. A polypeptide according to any one of claims 5 to 27, which has maximum enzymatic activity vis-à-vis pNP- β -xylopyranoside in a range between 60°C and 100°C.
29. A polypeptide according to any one of claims 5 to 27, which has maximum enzymatic activity vis-à-vis pNP- β -xylopyranoside in a range between 70°C and 95°C.
30. A polypeptide according to any one of claims 5 to 27, which has maximum enzymatic activity vis-à-vis pNP- β -xylopyranoside in a range between 80°C and 90°C.
31. A polypeptide according to any one of claims 5 to 30, of which the maximum activity vis-à-vis pNP- β -xylopyranoside after three hours of incubation at 60°C is at least 40 % of the activity of a second fraction of this polypeptide which was not incubated at the temperature 60°C.
32. A polypeptide according to any one of claims 5 to 30, of which the maximum activity vis-à-vis pNP- β -xylopyranoside after three hours of incubation at 60°C is at least 50 % of the activity of a second fraction of this polypeptide which was not incubated at the temperature 60°C.
33. A fusion protein of the polypeptide according to any one of claims 1 to 32, and a carbohydrate binding domain of a different protein.

34. A fusion protein of the polypeptide according to any one of claims 1 to 32, and a second polypeptide selected from one or more of the group consisting of a signal peptide, an affinity tag, and a protease cleavage site.
35. A fusion protein comprising the fusion protein according to claim 33 and a second polypeptide selected from one or more of the group consisting of a signal peptide, an affinity tag, and a protease cleavage site.
36. A mixture containing the polypeptide according to any one of claims 1 to 32, and one or more pectinases and/or one or more endoxylanases and/or one or more β -glucosidases and/or one or more β -glucanases and/or one or more cellobiohydrolases and/or one or more β -xylosidases and/or one or more α -arabinofuranosidases and/or one or more α -glucuronidases and/or one or more acetyl xylan esterases.
37. A mixture containing the fusion protein according to any one of claims 33 to 35 and one or more pectinases and/or one or more endoxylanases and/or one or more β -glucosidases and/or one or more β -glucanases and/or one or more cellobiohydrolases and/or one or more β -xylosidases and/or one or more α -arabinofuranosidases and/or one or more α -glucuronidases and/or one or more acetyl xylan esterases.
38. A nucleic acid encoding a polypeptide according to any one of claims 1 to 32.
39. A nucleic acid encoding a fusion protein according to any one of claims 33 to 35.
40. A vector containing the nucleic acid according to claim 38.
41. A vector containing the nucleic acid according to claim 39.
42. A host cell transformed with a vector according to claim 40 or 41.

43. A host cell according to claim 42, said host cell being an eukaryote.
44. A host cell according to claim 43, wherein said eukaryote is selected from the group consisting of *Saccharomyces cerevisiae*, *Yarrowia lipolytica*, *Schizosaccharomyces lactis*, *Kluyveromyces lactis*, *Pichia methanolytica*, *Pichia pastoris*, *Pichia angusta*, *Hansenula polymorpha*, *Aspergillus niger*, *Chrysosporium lucknowense*, *Trichoderma reesei*, and *Penicillium* sp.
45. A host cell according to claim 42 or claim 43, said host cell being a methylotrophic yeast.
46. A host cell according to claim 45, wherein said methylotrophic yeast is selected from the group consisting of *Pichia methanolytica*, *Pichia pastoris*, *Pichia angusta*, and *Hansenula polymorpha*.
47. A host cell according to claim 42, said host cell being a prokaryote.
48. A host cell according to claim 47, wherein said prokaryote is selected from the group consisting of *Bacillus* sp., *Bacillus subtilis*, *Bacillus licheniformis*; *Bacillus megaterium*, *Thermus thermophilus*, *Pseudomonas fluorescens*, *Fervidobacterium* sp., and *Escherichia coli*.
49. A method for purifying a polypeptide with β -pyranosidase activity according to any one of claims 1 to 32 comprising:
 - a) obtaining a host cell transformed with a vector according to claim 40;
 - b) cultivating the host cell under conditions where the polypeptide with β -pyranosidase activity is expressed; and
 - c) purifying the polypeptide with β -pyranosidase activity.

50. A method for purifying a fusion protein with β -pyranosidase activity according to any one of claims 33 to 35 comprising:
- a) obtaining a host cell transformed with a vector according to claim 41;
 - b) cultivating the host cell under conditions where the polypeptide with β -pyranosidase activity is expressed; and
 - c) purifying the polypeptide with β -pyranosidase activity.
51. A method according to claim 49 or 50 where c) comprises heat precipitation.
52. The use of a polypeptide according to any one of claims 1 to 32, or the mixture according to claim 36 for degrading one or more of the following substrates: β -xylopyranoside or β -glucopyranoside.
53. The use of a fusion protein according to any one of claims 33 to 35 or the mixture according to claim 37 for degrading one or more of the following substrates: β -xylopyranoside or β -glucopyranoside.
54. The use according to claim 52 or 53 for the enzymatic degradation of biomass containing lignocellulose and/or for textile processing and/or as an additive to detergents and/or in the food and/or feed industry.
55. The use according to claim 52 or 53 for preparing fruit juices and/or for preparing wine or beverages containing wine, and/or for the preparation of beer or beverages containing beer, and/or for the preparation of oil, and/or for the preparation of baked goods.
56. The use according to claim 55, said oil is olive oil, rapeseed oil or sunflower oil.

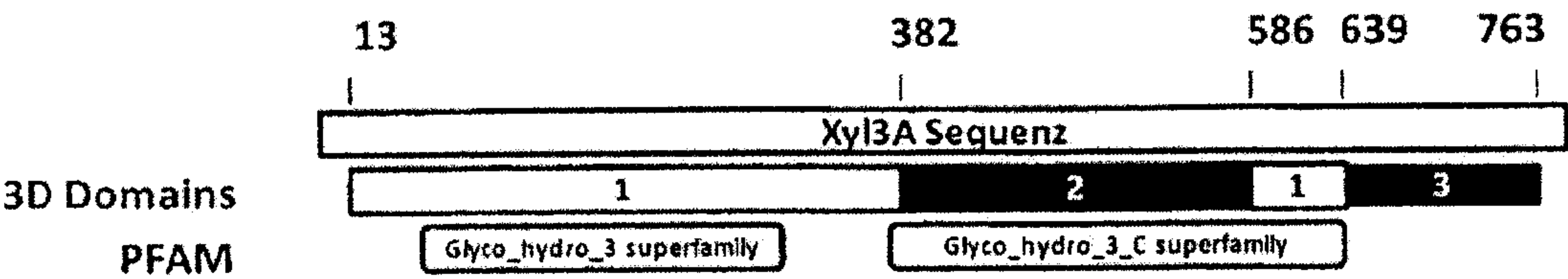


Fig. 1

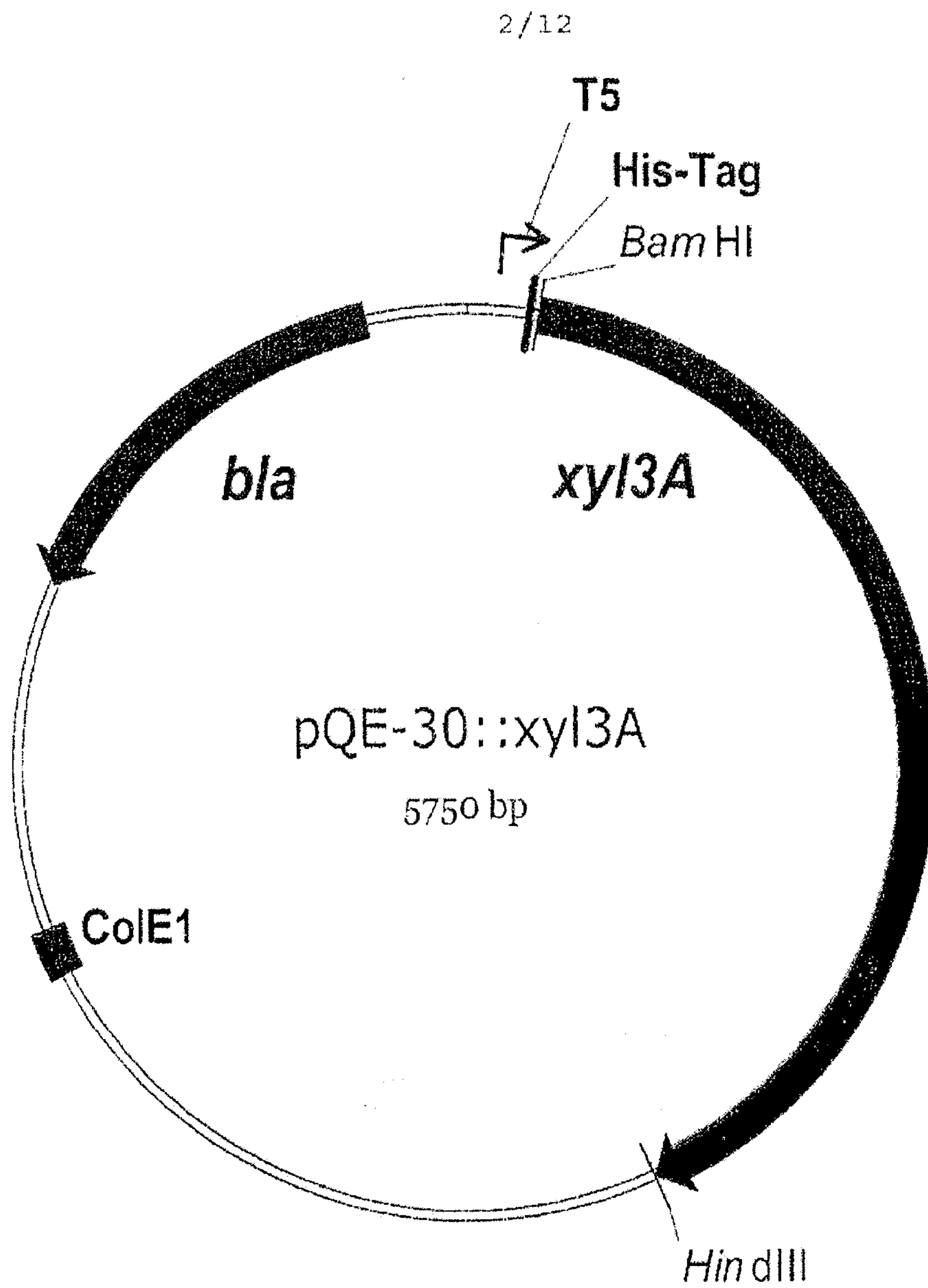


Fig. 2

3/12

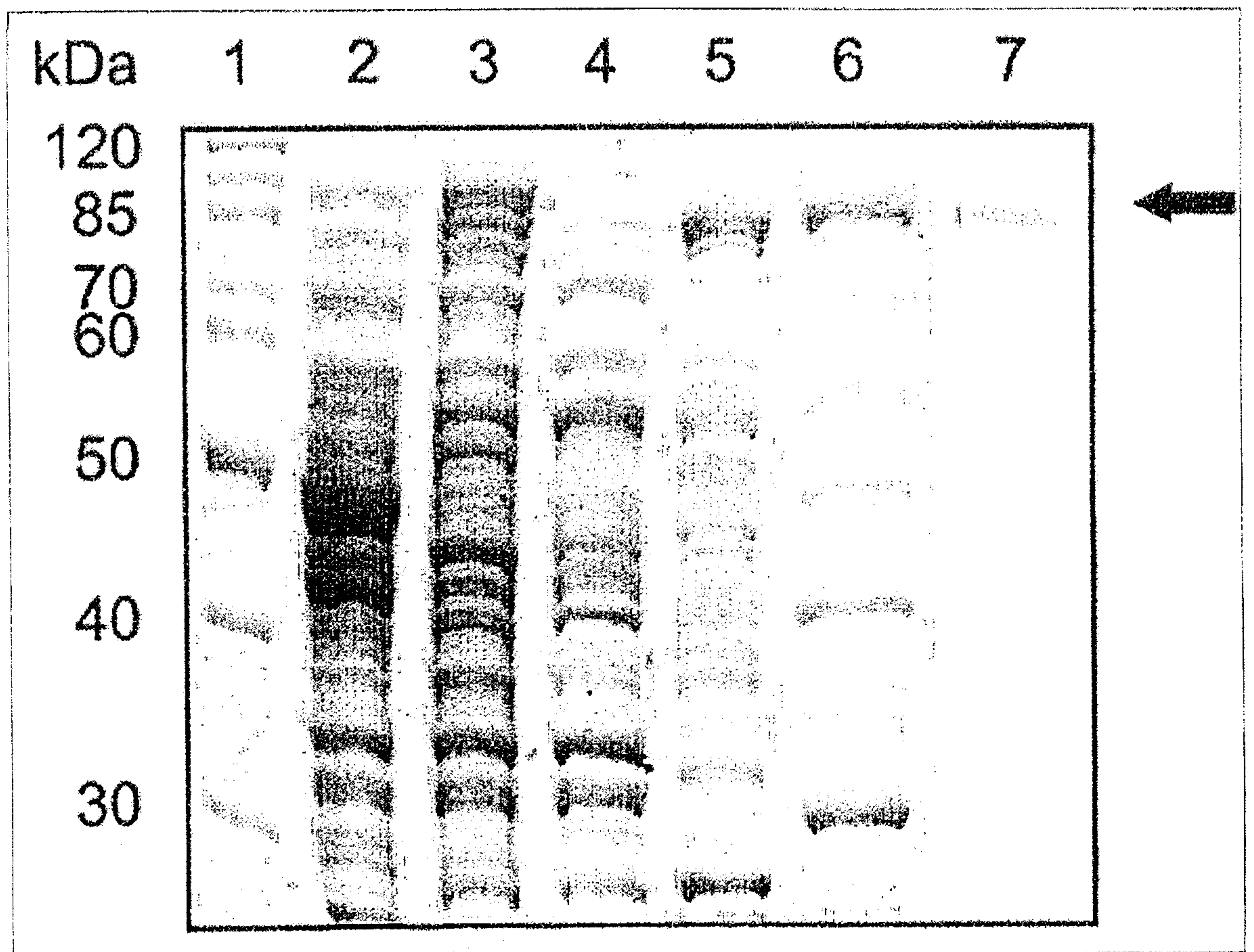


Fig. 3

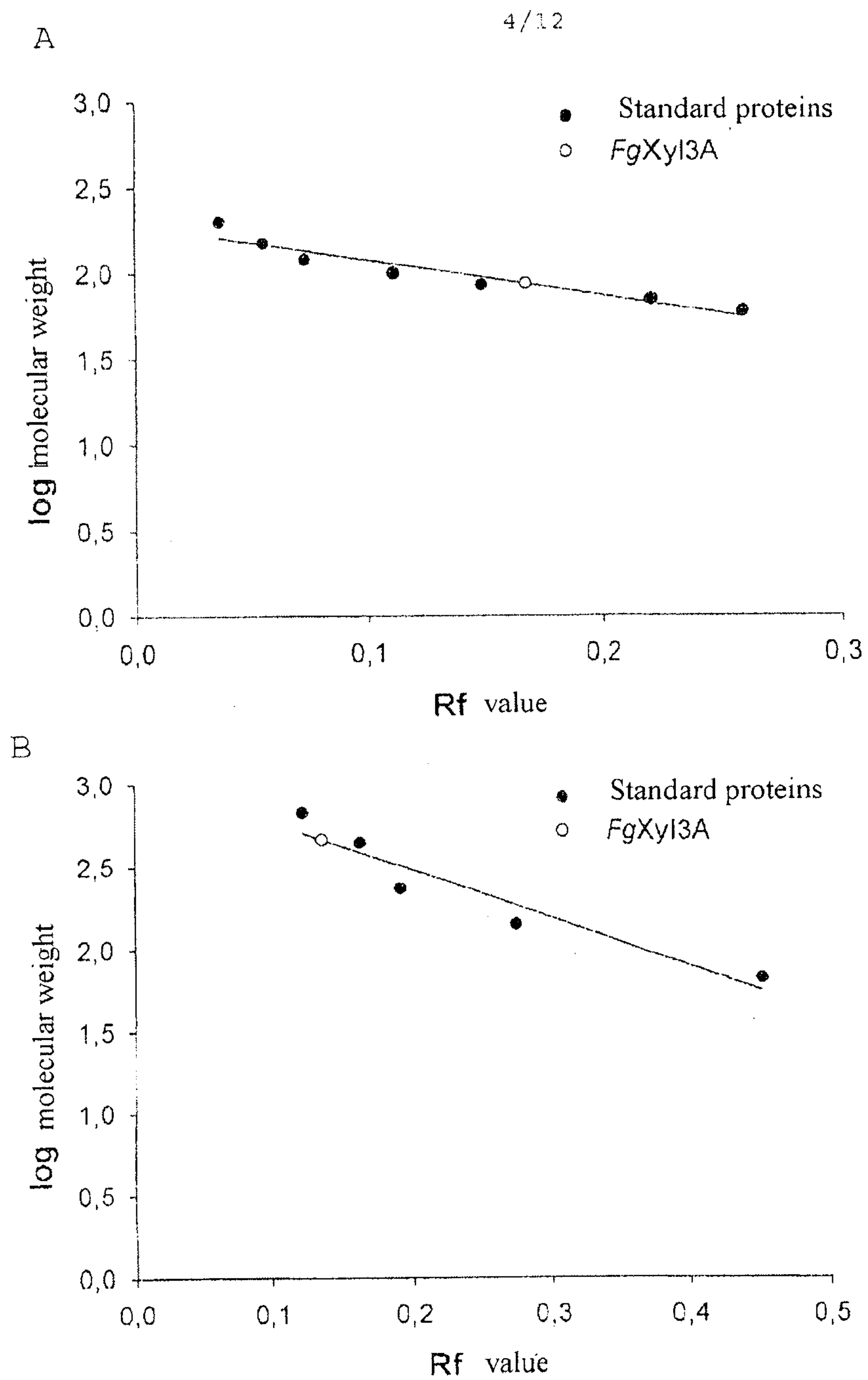


Fig. 4

5/12

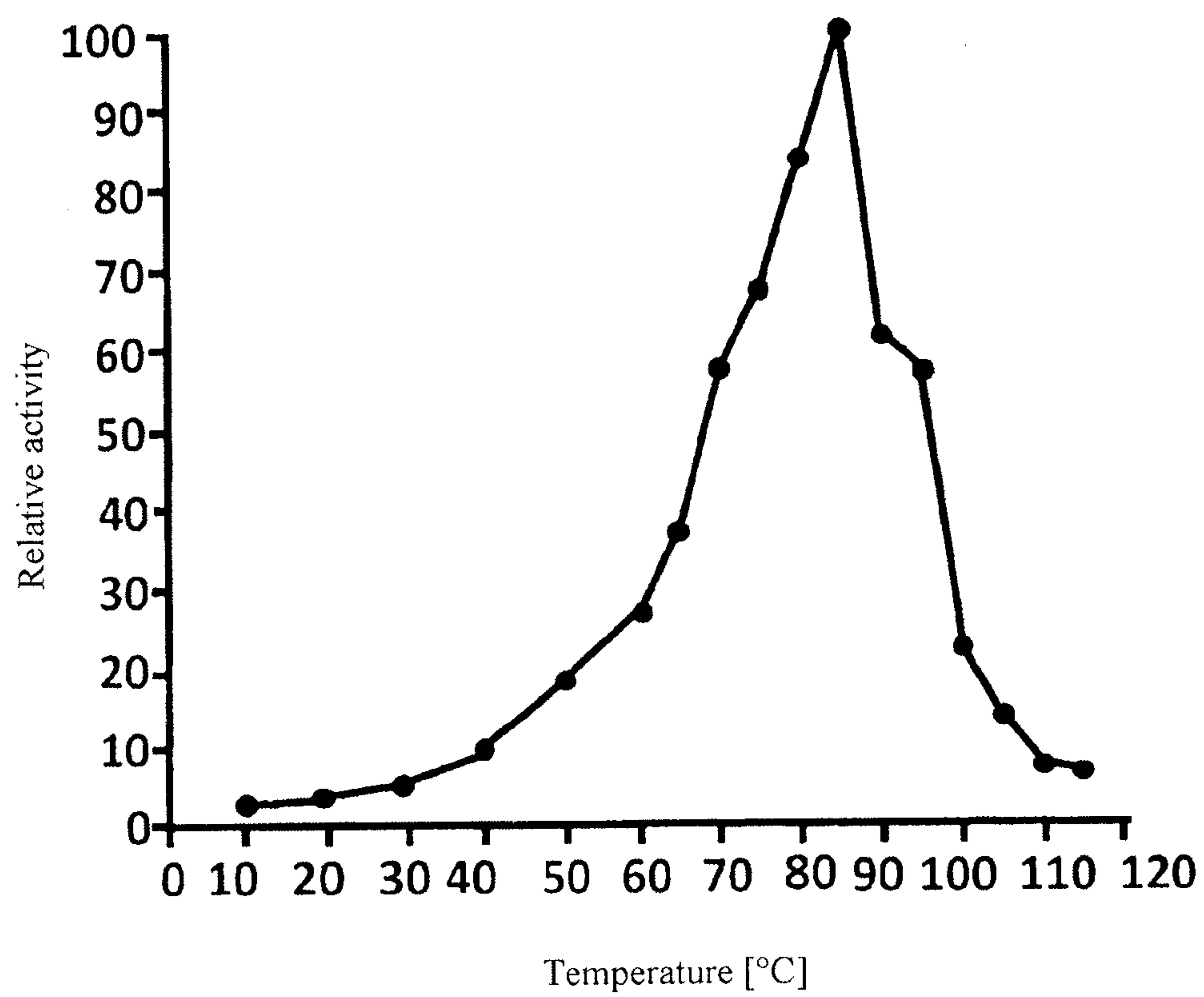


Fig. 5

6/12

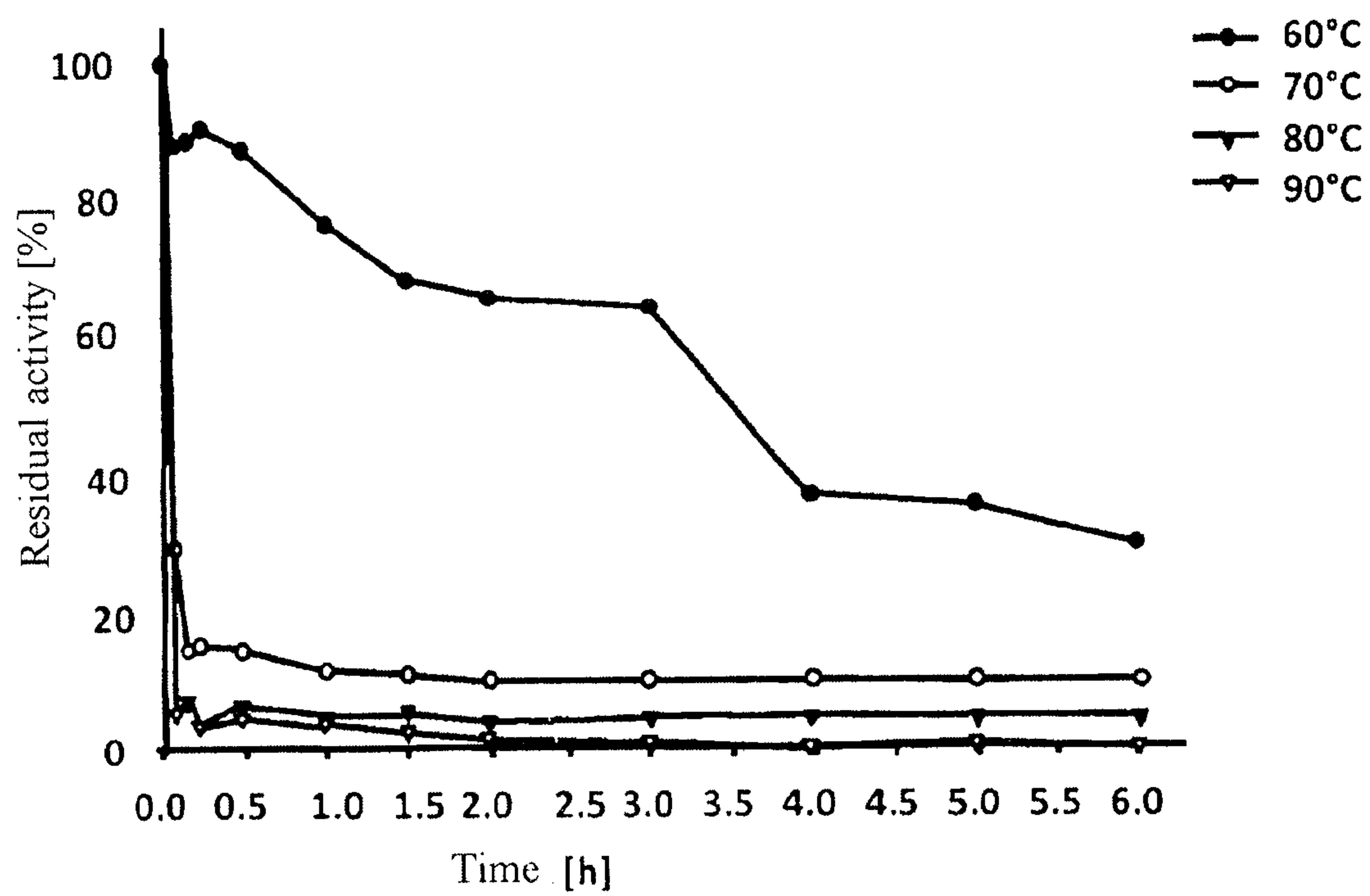


Fig. 6

7/12

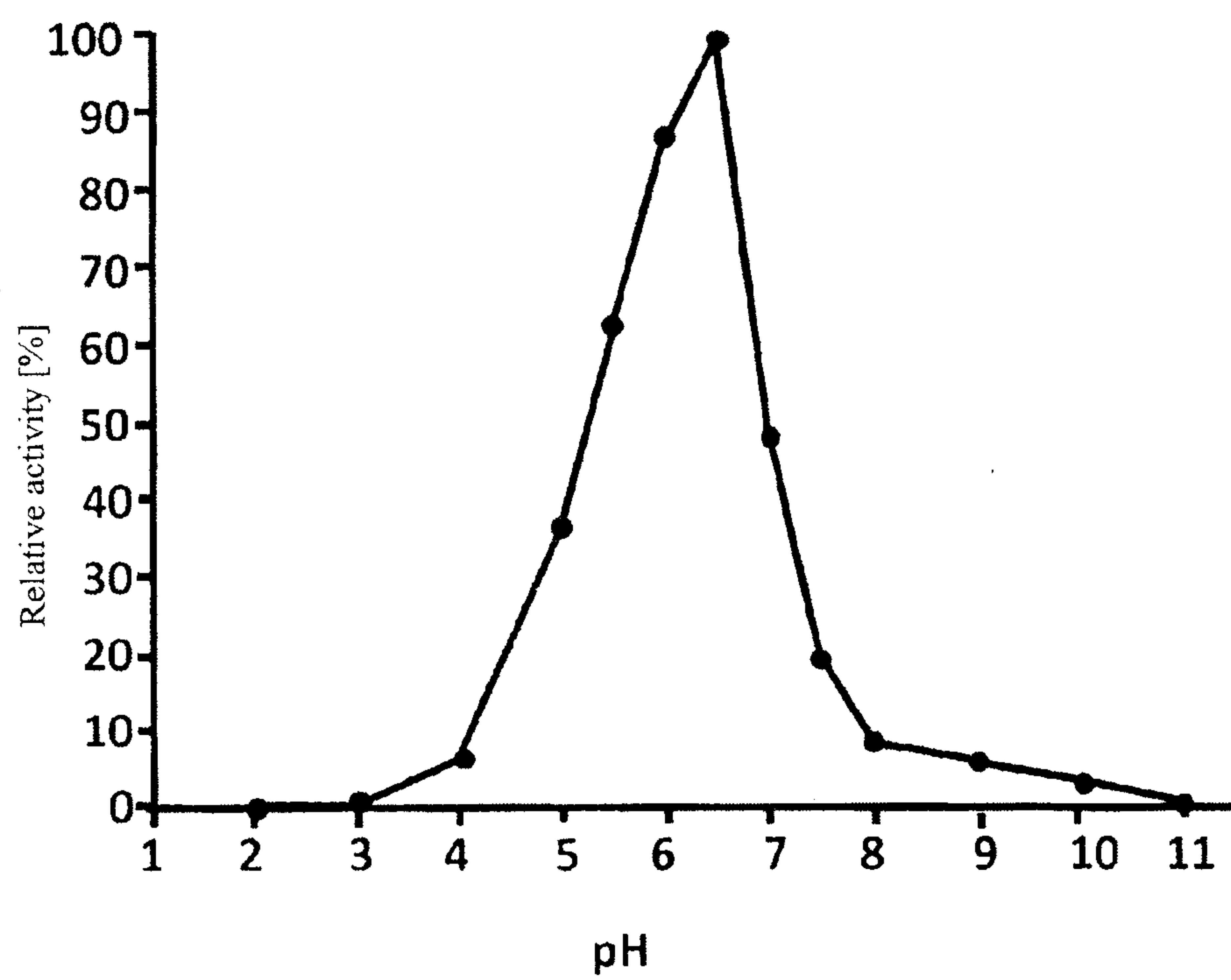


Fig. 7

8/12

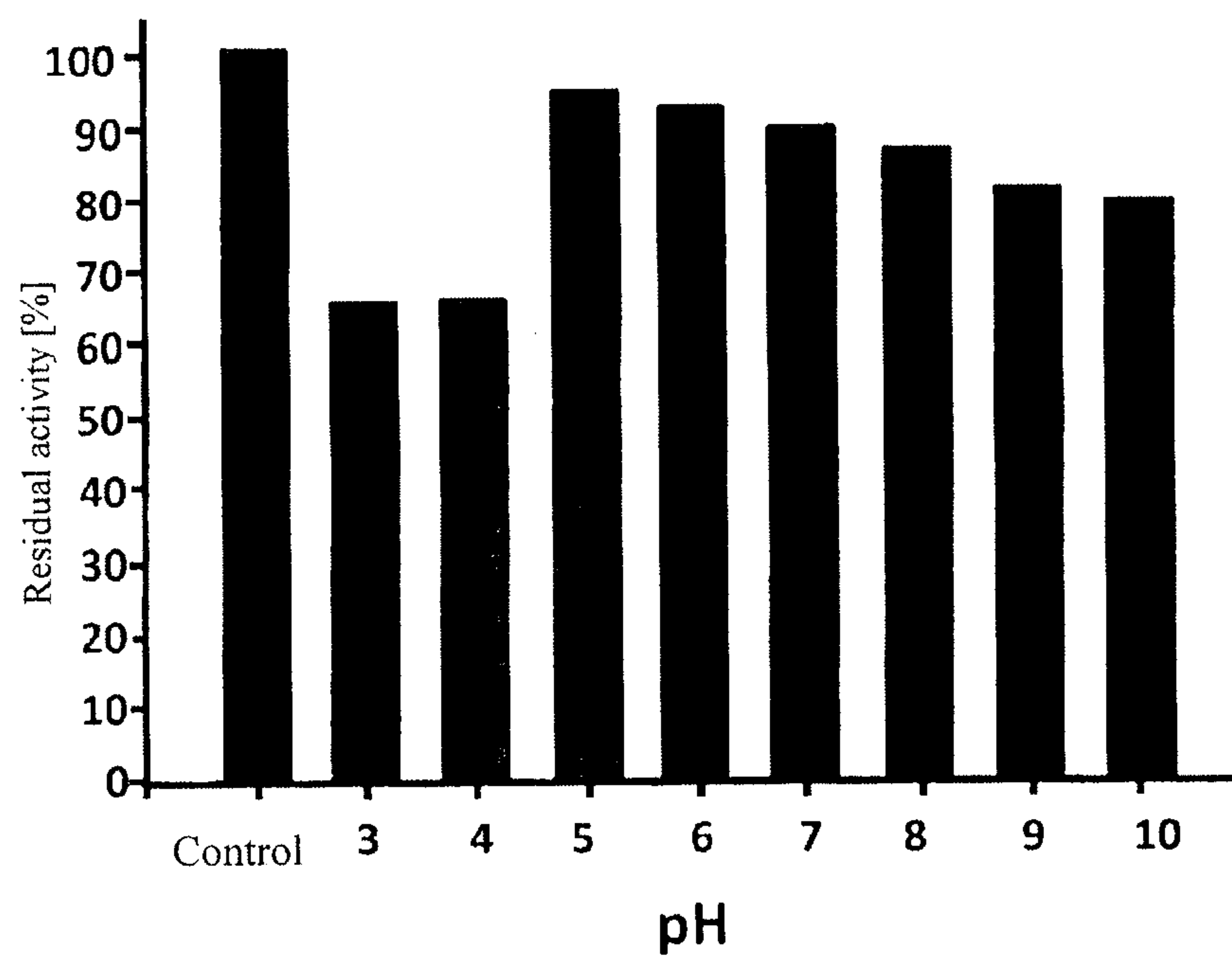


Fig. 8

9/12

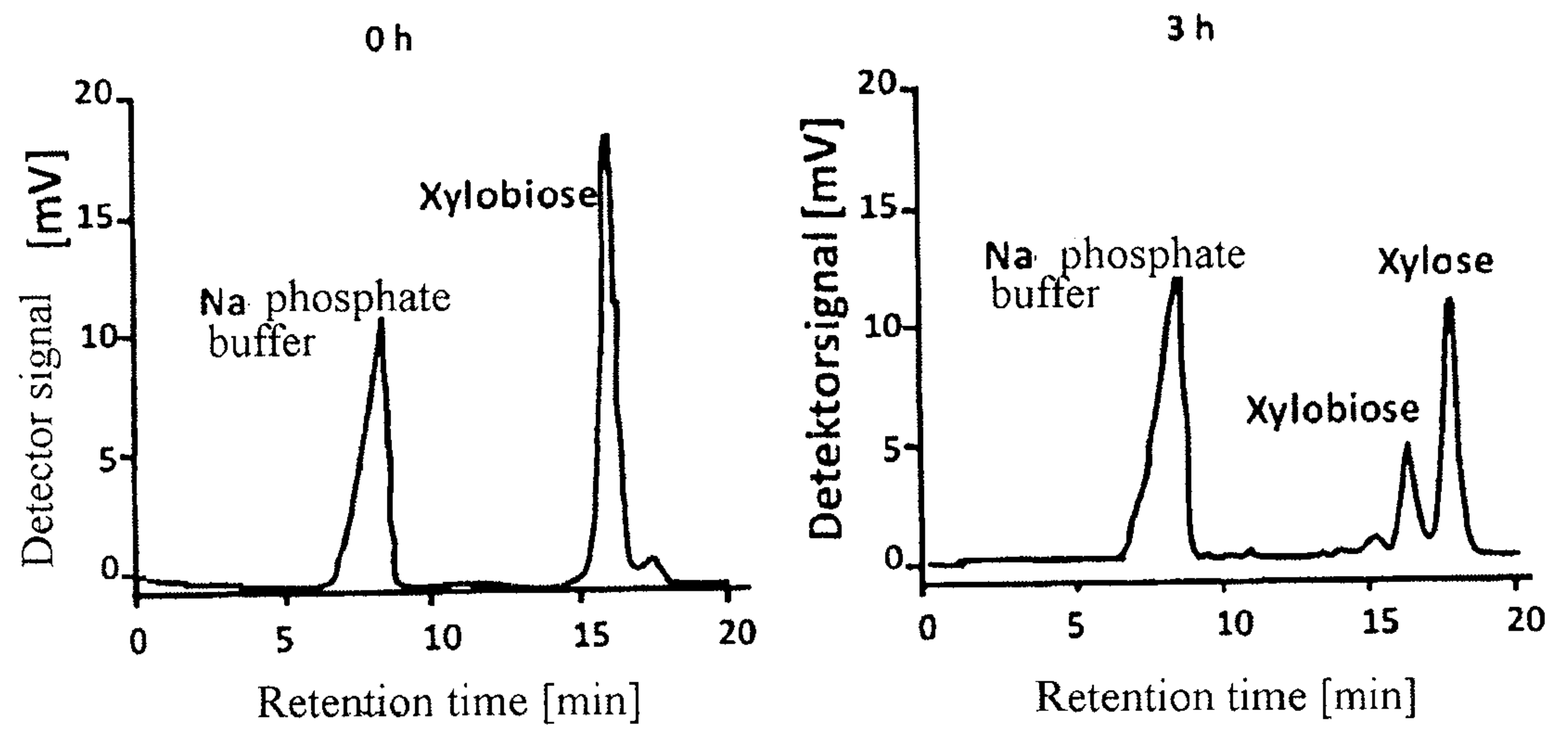


Fig. 9

10/12

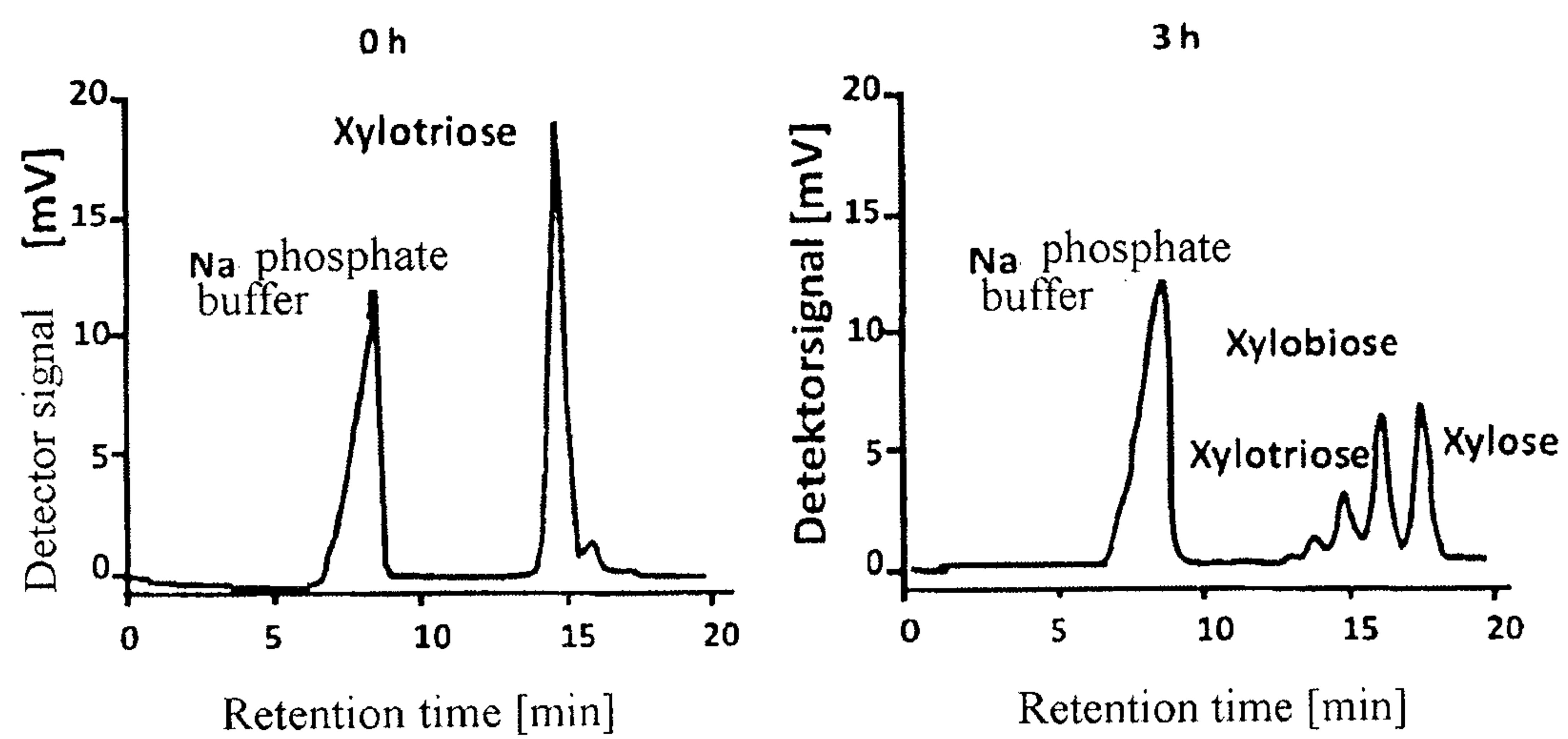


Fig. 10

11/12

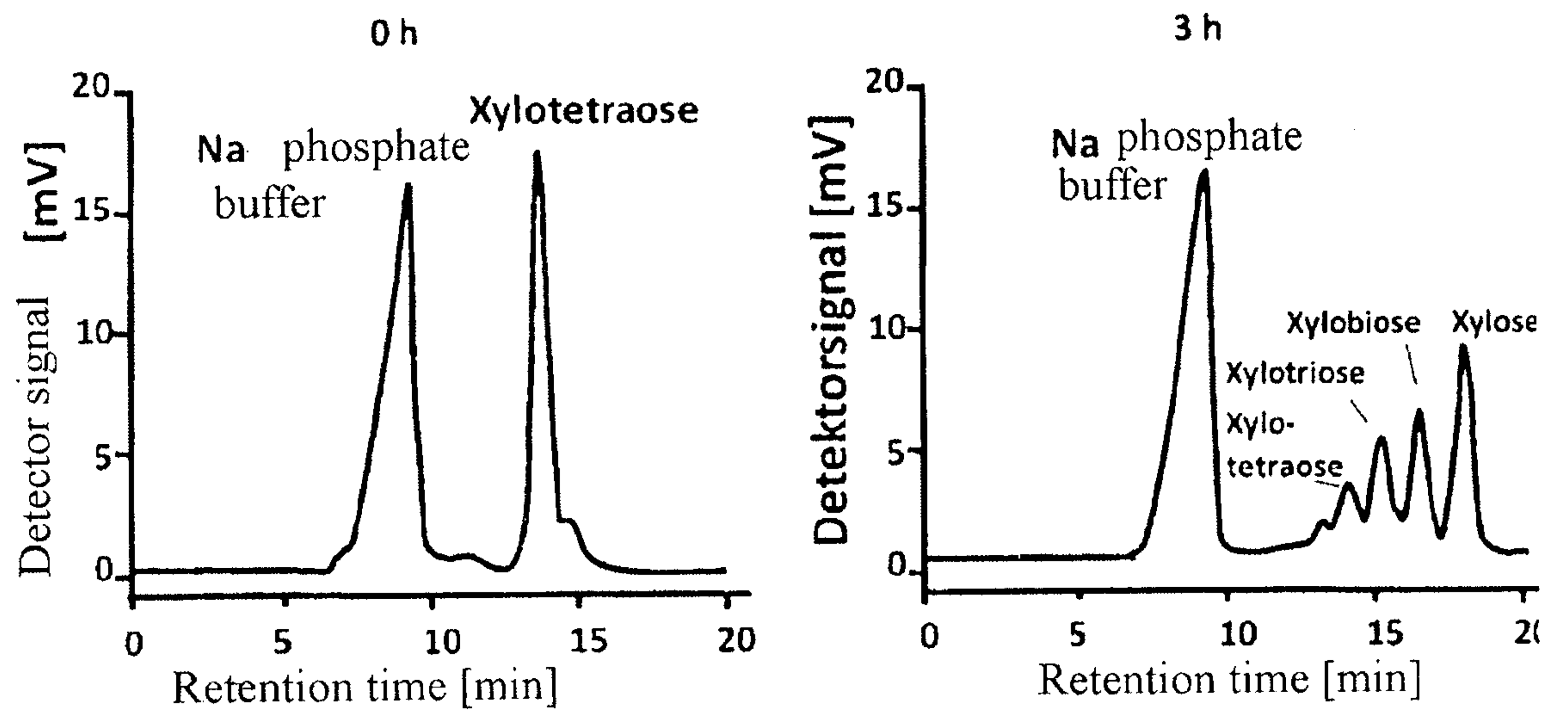


Fig. 11

12/12

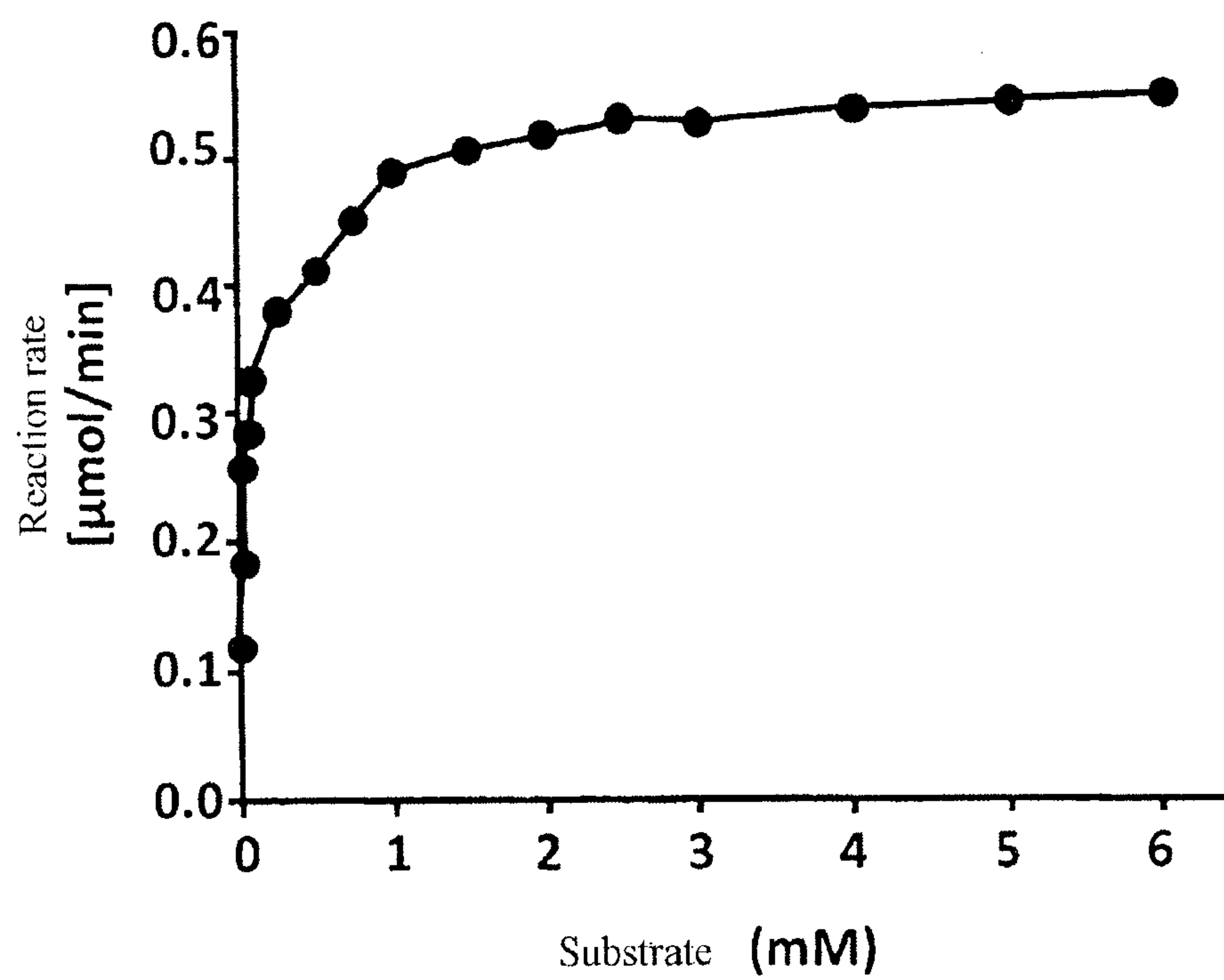


Fig. 12

3D Domains
PFAM

