

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/23180113>

Crystal structure of engineered β -glucosidase from a soil metagenome

ARTICLE in PROTEINS STRUCTURE FUNCTION AND BIOINFORMATICS · NOVEMBER 2008

Impact Factor: 2.63 · DOI: 10.1002/prot.22199 · Source: PubMed

CITATIONS

11

READS

15

8 AUTHORS, INCLUDING:



[Soo-Jin Kim](#)

National Academy of Agricultural Science (So...

92 PUBLICATIONS 513 CITATIONS

SEE PROFILE



[Chang-Muk Lee](#)

RURAL DEVELOPMENT ADMINISTRATION

33 PUBLICATIONS 442 CITATIONS

SEE PROFILE



[Kwang Yeon Hwang](#)

Korea University

125 PUBLICATIONS 2,519 CITATIONS

SEE PROFILE

STRUCTURE NOTE

Crystal structure of engineered β -glucosidase from a soil metagenome

Ki Hyun Nam,^{1†} Soo-Jin Kim,^{2†} Min-Young Kim,² Jae Hee Kim,¹ Yun-Soo Yeo,² Chang-Muk Lee,² Hong-Ki Jun,³ and Kwang Yeon Hwang^{1*}

¹ Division of Biotechnology, College of Life Sciences and Biotechnology, Korea University, Seoul, Korea

² Microbial Genetics Division, National Institute of Agricultural Biotechnology, Rural Development Administration, Suwon, Korea

³ Department of Microbiology, Pusan National University, Pusan, Korea

Key words: protein engineering; crystal structure; β -glucosidase complex with tartaric acid.

INTRODUCTION

Intensive screening of microbial isolates over the last several years has resulted in the identification and commercialization of numerous biomolecules, many of which are the products of microbial secondary metabolism.¹ Soil microorganisms in particular have been identified as valuable sources of naturally occurring industrially relevant antibiotics.² It has been estimated, based on the reassociation kinetics of DNA isolated from various soil samples, that the number of distinct prokaryotic genomes in soil ranges from 2000 to 18,000 genomes per gram of soil.³ Metagenomics is a rapidly developing field that involves the study of complex genomes within different environments and microbial niches.⁴ Several successful attempts to generate novel enzymes with enhanced catalytic activity and thermal stabilities using metagenomics have recently been reported.^{5–7}

The utilization of polysaccharides as an energy, chemical, and carbon source requires the activity of β -glucosidase (Bgl; EC 3.2.1.21).^{8,9} Bgl cleaves β -1,4-glycosidic linkages in disaccharide or glucose-substituted molecules through an acid–base catalytic mechanism, analogous to hydrolysis.¹⁰ The function of Bgl is essential for carbohydrate metabolism in cells, and the use of Bgl in various biotechnological processes, such as food and chemical synthesis, has been explored.^{11,12} Defects in Bgl activity in humans are associated with Gaucher disease, a non-neuronopathic lysosomal storage disorder.¹³ There are two Bgl homologues, BglA and BglB, which have been shown to be members of the GH-1 family of enzymes.

BglA and BglB are involved in the hydrolysis of cellooligosaccharides, but they have different quaternary structures and substrate specificities. BglA is a cellobiase with an unusual octameric structure, while BglB is a monomer that acts as an exo-Bgl, hydrolyzing cellobiose and cellooligosaccharides with a high degree of polymerization.^{14,15}

Recently, we isolated and characterized a novel *bgl* gene from uncultured soil bacteria (*Usbgl*).¹⁶ The gene encoded a protein with a predicted molecular weight of 55 kDa and an amino acid sequence was 56% identical to the family 1-glycosyl hydrolase of *Chloroflexus aurantiacus*. *Usbgl* exhibited substantial glycosyl hydrolase activity in the presence of natural glycosyl substrates, such as sophorose, cellobiose, cellotriose, cellotetraose, salicin, and arbutin, and was able to convert the major ginsenoside Rb1 into the pharmaceutically active minor ginsenoside, Rd. *Usbgl* also exhibited thermostable properties, which indicated that it may be useful in the development of novel biotechnical processes.¹⁶

Here, we report the X-ray crystal structure of engineered *Usbgl*. This type of structural analysis will con-

[†]Ki Hyun Nam and Soo-Jin Kim contributed equally to this work.

Grant sponsor: Korea Science and Engineering Foundation; Grant number: R01-2007-000-20072-0; Grant sponsor: National Institute of Agricultural Biotechnology; Grant number: 05-4-11-16-3.

*Correspondence to: Kwang Yeon Hwang, Division of Biotechnology, College of Life Sciences and Biotechnology, Korea University, Seoul 136-701, Korea.

E-mail: chahong@korea.ac.kr

Received 5 April 2008; Revised 18 June 2008; Accepted 25 June 2008

Published online 19 August 2008 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.22199

tribute not only to our ability to engineer pharmaceutically active forms Bgl, but may also lead to a better understanding of the molecular etiology of Gaucher disease. Our results also provide insight into the development and production of value-added products, including pharmaceuticals.

METHODS

Protein engineering, preparation, crystallization, and activity determination

Recombinant Bgl that lacked a signal peptide (accession no. DQ842022) was amplified from the β -glucosidase gene using the primers 5'-TCGCGGATCCATGACTGAACATGAGCTTCAG-3' (which contained a *Bam*HI site at the 5'-end) and 5'GCAAGCTTAATAGCGGGCGCGGCTAGCCC-3' (which contained a *Hind*III site at the 3'-end). The amplified DNA was ligated into *Bam*HI- and *Hind*III -digested pET28a(+) (Novagen) to create pEGLU. Mutations were introduced using the Quik-Change Multi Site-Directed Mutagenesis kit (Stratagene) and pEGLU (encoding wild-type *Us*Bgl) as the template. Two primers were designed to introduce the following sets of mutations: N-terminal primer (5'-TC GGA TCC AAC GTT AAG AAG TTC CCC GAG GGC TTT CTG TGG GGC-3') introduced the mutations Glu39AsnN, Leu40Val, Gln41Lys, Pro42Lys, and Lys45Glu; C-terminal primer (5'-GC AAG CTT TCA ATC CTC TAG CCC GTT GTT CGC AAT TAC GTC GCG-3') introduced the mutations Arg477Asn, Ala481Glu, and Ala482Asp. BL21(DE3) cells were transformed using recombinant plasmids, and cultures were grown at 37°C in LB medium supplemented with kanamycin (50 μ g/mL). Expression was induced for 18 h at 22°C by the addition of 1 mM IPTG. Cells were subjected to centrifugation at 4000 rpm for 30 min, resuspended in lysis buffer (50 mM Tris-HCl, pH 8.0; buffer A) and then subjected to sonication. The cells were subjected to centrifugation at 14,000 rpm for 30 min, and supernatants were collected and loaded onto a His Trap column (GE-Healthcare) that was pre-equilibrated with buffer A. Recombinant protein was eluted using a linear gradient of 0.5 M imidazole in buffer A, and then further purified by Hi-Load 16/60 Superdex 200 prep-grade chromatography (GE-Healthcare) using buffer B (10 mM Tris-HCl, pH 8.0, 50 mM NaCl, 50 mM KCl, and 10 mM DTT). Purified protein was concentrated to 20 mg/mL using a Centrprep column (Millipore). The initial crystallization of *Us*Bgl was performed with an automated Hydra II crystallization robot at 22°C using the sitting-drop vapor-diffusion method and commercially available kits from Hampton Research. Initial micro-crystals were obtained using 0.8 M Na/K tatrare, Tris-HCl (pH 7.0). Each hanging drop was prepared by mixing 1 μ L each of the protein solution and the reservoir solution, and then the mixture was placed

over 0.5 mL of reservoir solution. Optimized crystals were obtained from 0.8 M Na/K tatrare and 0.1 M Tris-HCl (pH 7.0–7.4) or 0.1 M MES (pH 6.8–7.0). Enzyme assays were carried out at 55°C using saccharide or related reagent as the substrate.¹⁶

X-ray data collection and structure determination

X-ray diffraction data was collected from a cooled crystal using an ADSC Quantum 210 CCD detector at beamline 6C using a Pohang Light Source (South Korea). Crystals were flash-frozen in a liquid nitrogen stream with 25% (v/v) glycerol as a cryo-protectant. The wavelength of the synchrotron X-rays was 1.23986 Å. The raw data was processed and scaled using DENZO and SCALEPACK from the HKL2000 program.¹⁷ Initial phases were obtained using molecular replacement. The PHASER program within the CCP4 program suite was employed using BglA from *T.maritima* (PDB code 1OIM) without water molecules and metal as a model.¹⁸ The structure was refined using simulated annealing, energy minimizations, and individual isotropic B factor refinement from the Crystallography & NMR System (CNS).¹⁹ A subset (5%) of the total number of reflections was randomly selected and multiple cycles of editing and adjustment of the model were carried out into sigma A weighted $3F_o - 2F_c$, $F_o - F_c$ using the program Coot.²⁰ Final refinement was carried out in REFMAC from the CCP4 suite of programs,²¹ and final structures were validated with PROCHECK.²² Structural analyses were calculated using CNS and PDBsum.²³ Graphical representations of the molecule were generated using PyMOL.²⁴

Data deposition

The coordinates and structure factors for *Us*Bgl have been deposited in the RCSB Protein Data Bank, accession code 3CMJ.

RESULTS AND DISCUSSION

Previously, we reported the isolation and purification of *Us*Bgl from uncultured soil bacteria, and demonstrated that it possessed catalytic activity.¹⁶ *Us*Bgl was less stable and had lower solubility than other Bgl family members (data not shown). We identified a putative signal peptide in the N-terminal region of *Us*Bgl, and a likely cleavage site between Ala18 and Lys19, using neural networks and hidden Markov models trained on gram-negative bacterium.²⁵ In the reported crystal structure of *Bacillus cereus* oligo-1,6-glucosidase, the N- and C-termini of the protein form a closed loop, and this structural feature is regarded as critical for maintaining enzyme stability.²⁶ Several other Bgl proteins also display a similar structural feature, in which the N- and C-termini interact through

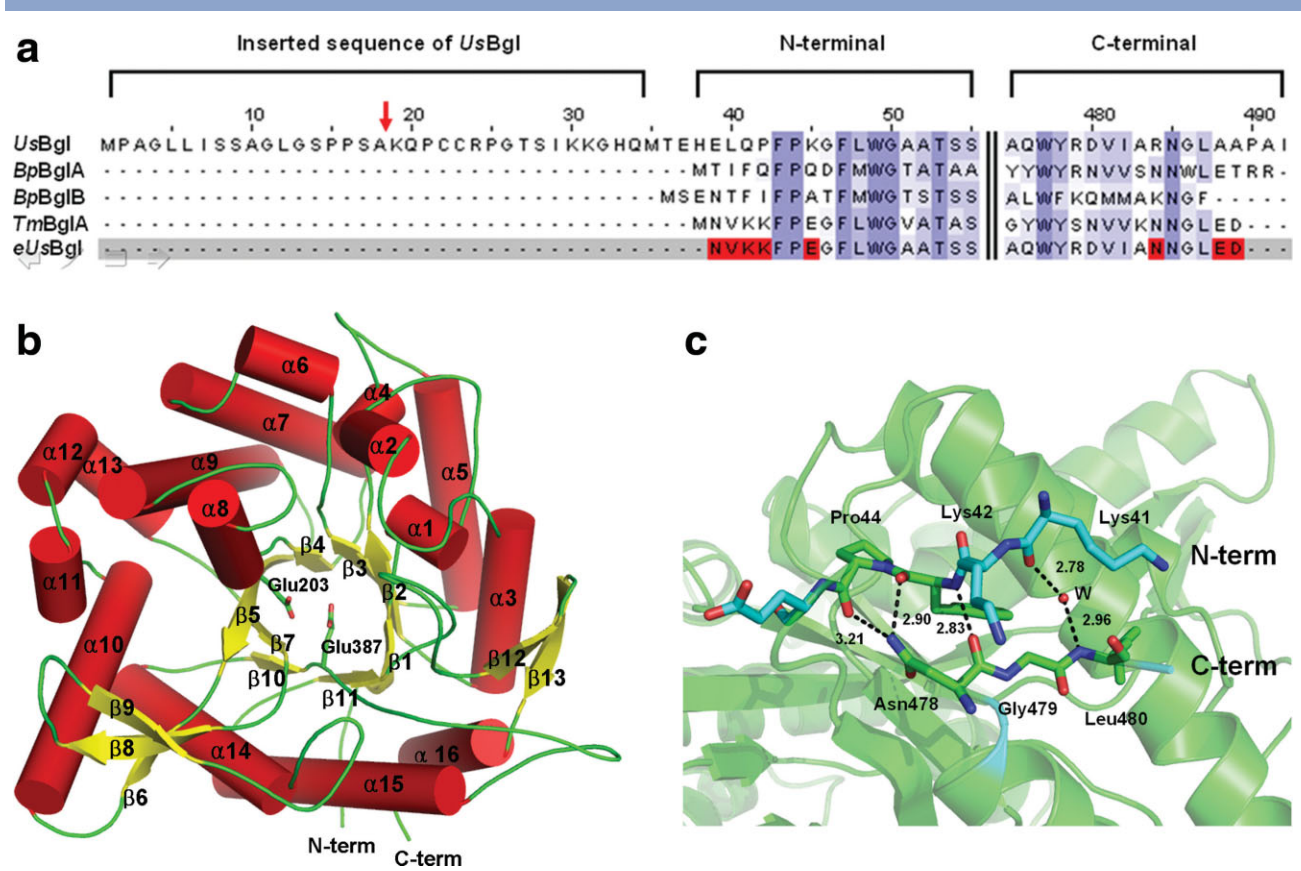


Figure 1
Alignment and overall structure of *UsBgl*. (a) Partial sequence alignment of engineered *UsBgl* with *TmBglA* (SWISS-PROT entry Q08638), *BpBglB* (P22073), and *BpBglA* (P22505). Amino acid numbering is for *UsBgl* (Genbank Accession No. DQ842022); residues in red (e*UsBgl*) correspond to the eight substituted residues of *UsBgl*. Dark blue indicates strictly conserved residues; gray indicates deleted residues. The arrow in *UsBgl* indicates the predicted gram-negative signal peptide sequence. (b) The overall structure of *UsBgl* is a typical TIM barrel fold. Catalytic residues are represented by a ball-and-stick model. (c) Structure of the engineered N- and C-termini (colored in cyan) of *UsBgl*, which associate with each other through hydrogen bonding among the antiparallel β -sheet.

hydrogen bonds. BLAST and PSI-BLAST searches²⁷ using *UsBgl* as a query sequence indicated that *UsBgl* is 47.4%, 44.1%, and 47.0% similar to *Thermotoga maritima* BglA (*TmBglA*), *Bacillus polymyxa* BglA (*BpBglA*), and *Bacillus polymyxa* BglB (*BpBglB*), respectively. To enhance the thermostability of *UsBgl*, we engineered several point mutations in the N- and C-terminal regions of the protein, based on the sequence and structure of *TmBglA*,

which is a highly thermostable protein. These mutations resulted in the following amino acid substitutions of *UsBgl*: Glu39AsnN, Leu40Val, Gln41Lys, Pro42Lys, Lys45Glu (N-terminal), and Arg477Asn, Ala481Glu, Ala482Asp (C-terminal) [Fig. 1(a)]. Engineered *UsBgl*, in which the N- and C-termini were mutated to match those of *TmBglA*, exhibited enhanced protein solubility (data not shown) and catalytic activity (Table I). Moreover, we

Table I
Catalytic Activity of Wild-Type and Engineered (Mutant) *UsBgl*

Substrate	pNPG		Cellobiose		Sophorose	
	Wild	Mutant	Wild	Mutant	Wild	Mutant
K_{cat} (s^{-1})	22.6 ± 5.0	33 ± 2.2	0.71 ± 0.014	1.58 ± 0.042	0.76 ± 0.26	1.46 ± 0.33
K_m (mM)	0.12 ± 0.03	0.097 ± 0.006	4.05 ± 0.68	4.19 ± 0.79	0.95 ± 0.019	0.57 ± 0.058
K_{cat}/K_m ($mM s^{-1}$)	179	336	0.176	0.376	0.80	2.58
Relatively activity	1	1.87	1	2	1	3.2

The activity of purified *UsBgl* was determined using different saccharides in a 50 mM Na-phosphate buffer, pH 6.5.

Table II

Data Collection and Refinement Statistics

Parameters	β -glucosidase
Data collection statistics	
Unit cell parameters (Å)	$a = 70.331$ $b = 71.002$ $c = 86.978$
Space group	$P2_12_12_1$
Wavelength (Å)	1.23986
Resolution range (Å)	20–1.6 (1.66–1.60)
Measured reflections	273,554
Unique reflections	54,735
Completeness	94.1 (66.0)
Redundancy	5.0 (2.3)
Average $I/\sigma(I)$	35.57 (3.55)
R_{merge} (%) ^a	0.065 (0.186)
Refinement statistics	
Reflections in working set	23,215
Reflections in test set	2537
$R_{\text{work}}/R_{\text{free}}$ (%) ^b	16.5/19.7
r.m.s.d. bonds (Å)	0.006
r.m.s.d. angles (°)	1.2
Ramachandran plot (%) ^c	
Most favored	90.9
Additionally allowed	8.6
Generously allowed	0.5
Average B factor (Å ²)	
Protein atom	14.72
Tartric acid	25.69
Metal ion	34.32
Water	26.23

^a $R_{\text{sym}} = \sum_i \sum_h I(h, i) / \langle I(h) \rangle / \sum_h \sum_i I(h, i)$, where $I(h, i)$ is the intensity of the i th measurement of reflection h and $\langle I(h) \rangle$ is the mean value of $I(h, i)$ for all i measurements.

^b $R_{\text{work}} = \sum ||F_{\text{obs}}| - |F_{\text{calc}}|| / \sum |F_{\text{obs}}|$, where F_{obs} and F_{calc} are the observed and calculated structure-factor amplitudes, respectively. R_{free} was calculated as R_{work} using a randomly selected subset (5%) of unique reflections not used for structure refinement.

^cCategories were defined by PROCHECK.

obtained crystals of engineered *UsBgl* under the crystallization conditions discussed earlier, whereas we have not been successful in crystallizing wild-type *UsBgl*.

Engineered *UsBgl* eluted as a monomer during size-exclusion chromatography and crystallized under K/Nat-tartrate conditions. The crystal belonged to the space group $P2_12_12_1$, and there was one molecule per asymmetric unit. Unlike other BglB proteins, we did not observe oligomeric structures in the crystallographic packing arrangement. The structure of *UsBgl* was determined using molecular replacement and was refined to 1.6 Å with an R_{work} of 16.5% and R_{free} of 19.7% (Table II). The overall structure of *UsBgl* had the classic TIM barrel structure, which was nearly identical to the structures of other Bgl proteins^{14,28–30} [Fig. 1(b)]. Each β -strand was hydrogen bonded to two neighboring strands, and within the barrel, a complete ring of hydrogen bonds was present. The engineered N- and C-termini were sealed by hydrogen bonding, in which the N δ of Asn478 interacted with the carbonyl group of Phe43 and Pro44 through a distance of 2.87 and 3.10 Å, respectively. The carbonyl

group of Asn478 interacted with the amino group of Phe43 over a distance of 2.78 Å, and the carbonyl group of Lys42 and amino group of Leu480 interacted with the hydrogen bond network via water molecules (2.76 and 2.87 Å, respectively) [Fig. 1(c)]. These results indicated that the closed structure of the N- and C-termini may play a role in increasing the stability of the engineered protein. However, we cannot rule out the contributions of the deleted signal peptide or the deleted regions of wild-type *UsBgl* to the stability of the engineered protein.

When we compared the structure of *UsBgl* with *TmBglA* (pdb code 1OIM), *BpBglA* (1BGA), and *BpBglB* (2O9T), the rms deviations were 0.68, 0.79, and 0.98, respectively.^{14,29–30} The entrance to the substrate binding pocket of *UsBgl* was formed primarily by the four extended loops that connected the strands and helices at the C-terminal side of the barrel, similar to other GH-1 family members. These long segments at the C-termini of the β -strands define the active centre, with the catalytic residues located at the ends of β 4 (acid-base Glu203) and β 11 (nucleophilic Glu387). These two amino acid residues have been shown to be required for the activity of glycosyl hydrolase.³¹ Several highly conserved residues in β 4 and β 11, including Trp49, Tyr113, Arg114, Phe115, Asn254, Asn330, Ile385, and Tyr432, may also play an important role in substrate recognition and the binding of sugar substrates and inhibitors.

The active site pocket of *UsBgl* bound to tartaric acid contained 13.2 buried vertices using the PDBsum program.²³ This was similar to *BpBglB* (2O9T) (13.3 buried vertices). Among BglA family members, the number of buried vertices of *TmBglA* (1OIM) and *BpBglA* (1BGA) is 16.1 and 9.2, respectively. Typically, BglA proteins have a narrow cavity and active site pocket, whereas the active site pocket of BglB proteins is larger than that of BglA. Thus, while the amino acid sequence and C α structure of *UsBgl* is similar to *TmBglA*, the substrate specificity,¹⁶ oligomeric state in solution based on size-exclusion experiments, and structure are more similar to BglB. Therefore, we propose that *UsBgl* belongs to the BglB family of proteins.

In the tartaric acid-bound form of *UsBgl*, conserved active site residues (Glu203 and Glu387) interact with the bound substrate. Interestingly, we found that tartaric acid adopted a slightly twisted form, and interacted with several residues in the active site pocket of *UsBgl* (Fig. 2). Tartaric acid interacted with the O ϵ atom of Glu203 (acid-base) and the O ϵ atom of Glu387 (nucleophilic) over 2.92 and 2.69 Å, respectively, as well as the conserved O ϵ atom of Gln57, N ϵ atom of His158, N δ atom of Asn202, O η atom of Tyr332, N ϵ atom of Trp434, O ϵ atom of Glu441 and N ϵ atom of Trp442 over a distance of 2.74, 2.97, 2.66, 2.71, 2.78, 2.33, and 3.04 Å, respectively. The mechanism of binding of the hydroxyl group of tartaric acid was similar to that of the glucoside geometry in the structure of the *TmBglA*-glucoside-tetrazole complex (2J7B, to be published).

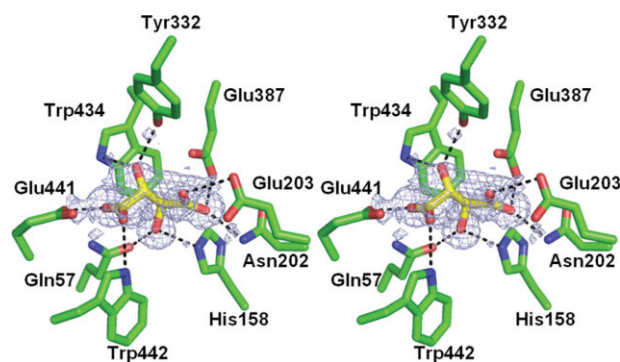


Figure 2

Active site pocket of *UsBgl* bound to tartaric acid. Stereoview of active site and neighboring residues and their interactions with tartaric acid. The electron density map (in light-blue) is shown using the final sigma A-weighted $3F_o - 2F_c$ electron density map, contoured at 1σ .

In summary, we have determined the crystal structure of engineered *UsBgl*, in which the N- and C-terminal regions were modified to enhance the stability of the protein. Our results provide insight into the engineering of enzymes for enhanced protein stability and increased activity. On the basis of biological and structural analysis, we propose that *UsBgl* belongs to the BglB family of GH-1 enzymes. The structure of the tartaric acid-bound form of *UsBgl* has significant implications for drug discovery efforts.

ACKNOWLEDGMENTS

The authors thank Dr. H. S. Lee, K. J. Kim, and K. H. Kim for assistance during data collection at beamline 6C of the Pohang Light Source, Korea.

REFERENCES

- Courtois S, Cappellano CM, Ball M, Francou FX, Normand P, Helynck G, Martinez A, Kolvek SJ, Hopke J, Osburne MS, August PR, Nalin R, Guérineau M, Jeannin P, Simonet P, Pernodet JL. Recombinant environmental libraries provide access to microbial diversity for drug discovery from natural products. *Appl Environ Microbiol* 2003;69:49–55.
- Daniel R. The soil metagenome—a rich resource for the discovery of novel natural products. *Curr Opin Biotechnol* 2004;15:199–204.
- Daniel R. The metagenomics of soil. *Nat Rev Microbiol* 2005;3:470–478.
- Streit WR, Schmitz RA. Metagenomics—the key to the uncultured microbes. *Curr Opin Microbiol* 2004;7:492–498.
- Sandoval C, Curtis H, Congote LF. Enhanced proliferative effects of a baculovirus-produced fusion protein of insulin-like growth factor and α_1 -proteinase inhibitor and improved anti-elastase activity of the inhibitor with glutamate at position 351. *Protein Eng* 2002;15:413–418.
- Kohno M, Enatsu M, Funatsu J, Yoshiizumi M, Kugimiya W. Improvement of the optimum temperature of lipase activity for *Rhizopus niveus* by random mutagenesis and its structural interpretation. *J Biotechnol* 2001;87:203–210.
- Nixon AE, Benkovic SJ. Improvement in the efficiency of formyl transfer of a GAR transformylase hybrid enzyme. *Protein Eng* 2000;13:323–327.
- Schafer ML, King KW. Utilization of cellulose oligosaccharides by *Cellvibrio gilvus*. *J Bacteriol* 1965;89:113–116.
- Halliwell G, Vincent R. The action on cellulose and its derivatives of a purified 1,4- β -glucanase from *Trichoderma koningii*. *Biochem J* 1981;199:409–417.
- Rixon JE, Ferreira LM, Durrant AJ, Laurie JJ, Hazlewood GP, Gilbert HJ. Characterization of the gene *celD* and its encoded product 1,4- β -D-glucan glucosylhydrolase D from *Pseudomonas fluorescens* subsp. *cellulosa*. *Biochem J* 1992;285:947–955.
- Diaz GA, Gelb BD, Risch N, Nygaard TG, Frisch A, Cohen IJ, Miranda CS, Amaral O, Maire I, Poenaru L, Caillaud C, Weizberg M, Mistry P, Desnick RJ. Gaucher disease: the origins of the Ashkenazi Jewish N370S and 84GG acid β -glucosidase mutations. *Am J Hum Genet* 2000;66:1821–1832.
- Riou C, Salmon JM, Valler MJ, Gunata Z, Barre P. Purification, characterization, and substrate specificity of a novel highly glucose-tolerant β -glucosidase from *Aspergillus oryzae*. *Appl Environ Microbiol* 1998;64:3607–3614.
- Li B, Renganathan V. Gene cloning and characterization of a novel cellulose-binding β -glucosidase from *Phanerochaete chrysosporium*. *Appl Environ Microbiol* 1998;64:2748–2754.
- Isorna P, Polaina J, Latorre-García L, Cañada FJ, González B, Sanz-Aparicio J. Crystal structures of *Paenibacillus polymyxa* β -glucosidase B complexes reveal the molecular basis of substrate specificity and give new insights into the catalytic machinery of family I glycosidases. *J Mol Biol* 2007;371:1204–1218.
- Gonzalez-Blasco G, Sanz-Aparicio J, Gonzalez B, Hermoso JA, Polaina J. Directed evolution of β -glucosidase A from *Paenibacillus polymyxa* to thermal resistance. *J Biol Chem* 2000;275:13708–13712.
- Kim SJ, Lee CM, Kim MY, Yeo YS, Yoon SH, Kang HC, Koo BS. Screening and characterization of an enzyme with β -glucosidase activity from environmental DNA. *J Microbiol Biotechnol* 2007;17:905–912.
- Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* 1997;276:307–326.
- McCoy AJ, Grosse-Kunstleve RW, Storoni LC, Read RJ. Likelihood-enhanced fast translation functions. *Acta Crystallogr D Biol Crystallogr* 2005;61:458–464.
- Brünger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL. Crystallography and NMR System: a new software suite for macromolecular structure determination. *Acta Crystallogr D* 1998;54:905–921.
- Paul E, Kevin C. Coot: model-building tools for molecular graphics. *Acta Crystallogr D* 2004;60:2126–2132.
- Winn MD, Murshudov GN, Papiz MZ. Macromolecular TLS refinement in REFMAC at moderate resolutions. *Methods Enzymol* 2003;374:300–321.
- Laskowski PA, McArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Cryst* 1993;26:283–291.
- Laskowski RA. PDBsum: summaries and analyses of PDB structures. *Nucleic Acids Res* 2001;29:221–222.
- DeLano WL. The PyMOL molecular graphics system. Available at: <http://www.pymol.org>.
- Bendtsen JD, Nielsen H, von Heijne G, Brunak S. Improved prediction of signal peptides: SignalP 3.0. *J Mol Biol* 2004;340:783–795.
- Watanabe K, Hata Y, Kizaki H, Katsube Y, Suzuki Y. The refined crystal structure of *Bacillus cereus* oligo-1,6-glucosidase at 2.0 Å resolution: structural characterization of proline-substitution sites for protein thermostabilization. *J Mol Biol* 1997;269:142–153.

27. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 1997;25:3389–3402.
28. Wierenga RK. The TIM-barrel fold: a versatile framework for efficient enzymes. *FEBS Lett* 2001;492:193–198.
29. Zechel DL, Boraston AB, Gloster T, Boraston CM, Macdonald JM, Tilbrook DM, Stick RV, Davies GJ. Iminosugar glycosidase inhibitors: structural and thermodynamic dissection of the binding of isofagomine and 1-deoxynojirimycin to β -glucosidases. *J Am Chem Soc* 2003;125:14313–14323.
30. Sanz-Aparicio J, Hermoso JA, Martínez-Ripoll M, Lequerica JL, Polaina J. Crystal structure of β -glucosidase A from *Bacillus polymyxa*: insights into the catalytic activity in family 1 glycosyl hydrolases. *J Mol Biol* 1998;275:491–502.
31. Akiba T, Nishio M, Matsui I, Harata K. X-ray structure of a membrane-bound β -glycosidase from the hyperthermophilic archaeon *Pyrococcus horikoshii*. *Proteins* 2004;57:422–431.