

# Novel substrate specificity of a membrane-bound $\beta$ -glycosidase from the hyperthermophilic archaeon *Pyrococcus horikoshii*

Ikuo Matsui<sup>a,\*</sup>, Yukihiro Sakai<sup>a</sup>, Eriko Matsui<sup>a</sup>, Hisasi Kikuchi<sup>b</sup>, Yutaka Kawarabayashi<sup>b</sup>, Koichi Honda<sup>a</sup>

<sup>a</sup>National Institute of Bioscience and Human-Technology, Tsukuba, Ibaraki 305, Japan

<sup>b</sup>National Institute of Technology and Evaluation, MITI, Nishihara, Shibuyaku, Tokyo, Japan

Received 14 December 1999; received in revised form 12 January 2000

Edited by Shozo Yamamoto

**Abstract** A  $\beta$ -glycosidase gene homolog of *Pyrococcus horikoshii* (BGPh) was successfully expressed in *Escherichia coli*. The enzyme was localized in a membrane fraction and solubilized with 2.5% Triton X-100 at 85°C for 15 min. The optimum pH was 6.0 and the optimum temperature was over 100°C, respectively. BGPh stability was dependent on the presence of Triton X-100, the enzyme's half-life at 90°C (pH 6.0) was 15 h. BGPh has a novel substrate specificity with  $k_{cat}/K_m$  values high enough for hydrolysis of  $\beta$ -D-Glcp derivatives with long alkyl chain at the reducing end and low enough for the hydrolysis of  $\beta$ -linked glucose dimer more hydrophilic than aryl- or alkyl- $\beta$ -D-Glcp.

© 2000 Federation of European Biochemical Societies.

**Key words:**  $\beta$ -Glycosidase; Thermophilic archaeon; Membrane protein; Thermostable enzyme; *Pyrococcus horikoshii*

## 1. Introduction

The current biotechnological interest in enzymes from thermophilic microorganisms is motivated by their ability to work under conditions normally denaturing for mesophilic enzymes. Whereas conventional enzymes are irreversibly inactivated by heat, the enzymes from these extremophiles show not only great stability but also enhanced activity in the presence of common protein denaturants such as heat, detergents, organic solvents and proteolytic enzymes [1–5]. Thus, these molecules have considerable industrial potential by giving better yields under extreme operational conditions. Among the possible applications of such enzymes from thermophiles, the enzymatic synthesis of glycosides [6,7] and the saccharification of cellulosic materials by  $\beta$ -glycosidases have long been recognized as important and have recently received renewed attention from the pharmaceutical industry [6].

A  $\beta$ -glycosidase gene homolog (ORF ID: PH0366) was

identified from the hyperthermophilic archaeon *Pyrococcus horikoshii* through genome sequencing [8,9]. The gene (BGPh) was successfully expressed in *Escherichia coli* and proved to be a membrane protein. The gene product was purified to homogeneity and the substrate specificity was characterized in detail. Here, we report its novel substrate specificity and enzymatic behavior as a membrane protein.

## 2. Materials and methods

### 2.1. Chemicals

The pET-11a vector was purchased from Stratagene. The pET-15b vector and *E. coli* strain BL21(DE3) were obtained from Novagen. Vent DNA polymerase was purchased from New England Biolabs. Restriction enzymes were purchased from Promega and Toyobo (Osaka, Japan), and were used according to the manufacturers' recommendations. Ultrapure deoxynucleotide solution (dNTPs) was obtained from Pharmacia Biotech. Isopropyl- $\beta$ -D-thiogalactopyranoside (IPTG) was from Takara Shuzo (Otsu, Shiga, Japan).

### 2.2. Cloning and expression of the gene

The genome of *P. horikoshii* was sequenced using the method of Kaneko et al. [10]. Standard cloning techniques were used throughout. The gene encoding  $\beta$ -glycosidase (BGPh) was amplified by the PCR method using the following two primers: TAAGAAGGAGATA-TACATATGCCGCTGAAATTCCCGGAAATGTTTCTCTTTGG-TACC (upper primer, containing an *NdeI* site as underlined); TTTACTGCAGAGAGGATCCCTAATCCTAAAGTTGAAGTTC-TGGTAG (lower primer, containing a *BamHI* site as underlined). The PCR product was cloned into expression vectors pET-11a and pET-15b using *NdeI* and *BamHI* sites. The resulted vectors were designated as pET-11a/BGPh and pET-15b/BGPh, respectively. The absence of additional mutations within the coding region of BGPh was verified by sequencing on an Applied Biosystems 373A DNA sequencer (Taq DyeDeoxy Terminator Cycle Sequencing kit, Perkin-Elmer).

The *E. coli* strain BL21(DE3) was transformed with the pET-11a/BGPh plasmid to express mature BGPh and pET-15b/BGPh plasmid to express His-tagged BGPh. The transformant colony was propagated in 2×YT+ampicillin medium at 37°C and was induced at OD<sub>600</sub> = 1 with 1 mM IPTG for 4 h. The induced cells were collected by centrifugation and stored at –20°C. The frozen cells (7 g) were mixed with 10 ml of 50 mM Tris–HCl buffer (pH 7.5) containing 1 mg of bovine DNase I (Sigma) and incubated at 37°C for 30 min. Triton X-100 was added to the suspension, resulting in a final concentration of 2.5%. Then, the cell suspension was heated at 85°C for 10 min and centrifuged at 5000×g for 20 min. The supernatant was collected and stored at 4°C.

### 2.3. Purification of recombinant protein

The solubilized recombinant BGPh with His-tag (His-BGPh) was subjected to affinity chromatography with Ni-conjugated Sepharose, using a stepwise elution from 5 mM to 1 M imidazole in 20 mM Tris–HCl (pH 8.0) with 0.5 M NaCl solution (His-bind Buffer kit, Novagen) containing 0.1% Triton X-100. His-BGPh was eluted with 100 mM imidazole with 0.1% Triton X-100. The enzyme samples were

**Abbreviations:** BGPh,  $\beta$ -glycosidase from *Pyrococcus horikoshii*; BMPH, a  $\beta$ -mannosidase gene homolog from *P. horikoshii*; BGPF,  $\beta$ -glucosidase from *Pyrococcus furiosus*; BMPf,  $\beta$ -mannosidase from *P. furiosus*; S $\beta$ -gly,  $\beta$ -glycosidase from *Sulfolobus solfataricus*; IPTG, isopropyl- $\beta$ -D-thiogalactopyranoside; His-BGPh, BGPh with His-tag; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis; X-Glu, 5-bromo-4-chloro-3-indolyl- $\beta$ -glucopyranoside; *p*-Nph- $\beta$ -D-Glcp, *p*-nitrophenyl- $\beta$ -D-glucopyranoside; LA- $\beta$ -D-Glcp,  $\beta$ -D-glucopyranosides with long alkyl chains

\*Corresponding author. Fax: (81)-298-546151.  
E-mail: ikmatsui@nibh.go.jp

analyzed by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) [11]; a low molecular weight electrophoresis calibration kit, purchased from Pharmacia Biotech, was also run. The purified protein demonstrated a single band with a molecular weight of 35 kDa measured on SDS–PAGE followed by Coomassie blue staining. The protein concentration was determined using a Coomassie protein assay reagent (Pierce Chemical Company). The His-tagged protein was detected with QIAexpress Detection System (Qiagen) after blotting onto a nitrocellulose membrane (Pharmacia Biotech).

#### 2.4. Cellular localization of the activity

Localization of the BGPh activity in *E. coli* transformant cells (BL21(DE3)/pET-11a/BGPh or BL21(DE3)/pET-15b/BGPh) was examined by fractionation of the cell components. The cell membrane was isolated as follows: 7 g of the induced cells, which were frozen at  $-20^{\circ}\text{C}$ , was thawed and mixed with 10 ml of 50 mM Tris–HCl buffer (pH 7.5). The cell suspension (suspension I) was sonicated with a Sonifier 250 (Branson) for 4 min at an output control level of 4 and at 30% duty cycle. The sonicated sample was centrifuged at  $9000\times g$  for 10 min to remove cell debris, then the supernatant (12 ml) was ultracentrifuged at  $100\,000\times g$  for 1 h to separate the membrane fraction (1 ml) from the supernatant. The enzyme reactions were carried out at  $90^{\circ}\text{C}$  for 15 min in a solution (200  $\mu\text{l}$ ) containing 1.2 mM 5-bromo-4-chloro-3-indolyl- $\beta$ -glucopyranoside (X-Glu) and 5  $\mu\text{l}$  of each fraction, as the enzyme source, in 50 mM phosphate buffer (pH 6) with 0.3 M NaCl. After the reaction, the solution was cooled in ice and diluted with 1 ml of water; the absorbance at 620 nm was immediately measured. As a control, the assay reactions were performed under the same conditions but without X-Glu to subtract the turbidity derived from each fractionated sample.

To analyze the solubilizing effect of Triton X-100, suspension I was also heated with and without 2.5% Triton X-100 at  $85^{\circ}\text{C}$  for 10 min and the supernatant was obtained by centrifugation at  $15\,000\times g$  for 10 min. The activity of the supernatants was measured using X-Glu as shown above.

#### 2.5. Measurement of the kinetic parameters

The enzyme reactions were carried out at  $90^{\circ}\text{C}$  in a solution (200  $\mu\text{l}$ ) containing the substrate and the purified His-BGPh in 50 mM phosphate buffer (pH 6) with 0.1% Triton X-100 and 0.3 M NaCl. For the hydrolysis of *p*-nitrophenyl (*p*-Nph)- $\beta$ -D-saccharides, the reaction was terminated by the addition of 1 M  $\text{Na}_2\text{CO}_3$  (1 ml), then centrifuged at  $15\,000\times g$  for 10 min. The concentration of the *p*-Nph group in the supernatant was quantified by measuring the absorbance at 400 nm. For the hydrolysis of  $\beta$ -D-glucoside, the released glucose was analyzed with a Glucose C-II Test kit (Wako Pure Chemicals, Japan). Initial velocities were obtained directly from the initial slopes of the time course plots. The  $K_m$  and  $k_{\text{cat}}$  values were calculated using the Michaelis–Menten equation and the least squares method [12]. The subsite affinity for a long alkyl chain was determined using the method reported previously [13–15] on the basis of the subsite theory [16].

#### 2.6. Dependence on Triton X-100, optimum temperature, optimum pH and thermostability

To measure the dependence of the activity on Triton X-100, the enzyme reactions were carried out at  $98^{\circ}\text{C}$  for 20 min in 50 mM phosphate buffer (pH 6) containing Triton X-100 varying from 0.1% to 0.00002%, 3 mM *p*-Nph- $\beta$ -D-glucopyranoside (*p*-Nph- $\beta$ -D-Glcp), 57.5  $\mu\text{M}$  of the purified His-BGPh and 0.1 M NaCl. Optical density measurements at  $A_{400}$  were performed as described for the enzyme assays.

The optimum temperature was measured from  $50^{\circ}\text{C}$  to  $100^{\circ}\text{C}$  in 150 mM citrate buffer (pH 5.0) without Triton X-100 and with 1  $\mu\text{l}$  of heated suspension I (BL21(DE3)/pET-11a/BGPh) as enzyme source. Further details of the measurement were described in the dependence on Triton X-100.

The optimum pH was measured at  $90^{\circ}\text{C}$  in 139 mM buffer systems ranging from pH 3.9 to 5.5 in sodium acetate buffer and from 5.5 to 7.99 in phosphate buffer with 1  $\mu\text{l}$  of heated suspension I (BL21(DE3)/pET-11a/BGPh or BL21(DE3)/pET-15b/BGPh). Further details of the measurement were reported in the determination of the optimum temperature.

To measure the thermostability, the His-BGPh solutions (29 nM) in 50 mM phosphate buffer (pH 6.0) containing 100 mM NaCl and 0.1%

Triton X-100 were heated in sealed Eppendorf tubes at  $90^{\circ}\text{C}$  in various increments up to 24 h. The heated enzymes were assayed in phosphate buffer (pH 6.0) at  $90^{\circ}\text{C}$  as described for the determination of optimum temperature.

#### 2.7. Sequence alignment, phylogenetic tree and hydropathy profile

Sequence alignment of  $\beta$ -glycosidases was performed using the GeneWorks program (IntelliGenetics) based on a PAM-250 scoring matrix. The enzymes of interest were: BGPh studied in this paper and  $\beta$ -mannosidase (BMPh) from *P. horikoshii* [8,9],  $\beta$ -glucosidase (BGPh) and  $\beta$ -mannosidase (BMPf) from *Pyrococcus furiosus* [17], and  $\beta$ -glucosidase (S $\beta$ -gly) from *Sulfolobus solfataricus* [18]. Phylogenetic trees for the same sequences were constructed using the GeneWorks program based on the unweighted pair group method with an arithmetic mean [19]. Each hydropathy profile was analyzed with DNASIS-Mac v2.0 software based on the Kyte and Doolittle method [20].

### 3. Results and discussion

#### 3.1. Localization of the activity in *E. coli* membrane

The intracellular localization of His-BGPh was examined (Table 1). The activity was present in the membrane fraction whereas very little activity was detected in the soluble fraction after the ultracentrifugation. The solubilizing efficiency with Triton X-100 was elevated by heating up to  $85^{\circ}\text{C}$ , whereas only 22% of the activity was extracted at room temperature. The best condition for the solubilization was 2.5% Triton X-100 at  $85^{\circ}\text{C}$  for 15 min. The native-type BGPh was also solubilized under the same condition as His-BGPh (data not shown); however, the denaturation with 8 M urea and the renaturation by direct dilution with buffer had no effect on the solubilization of the activity (data not shown).

The activity of His-BGPh was dependent on the concentration of Triton X-100. At 0.00002% Triton X-100, the activity decreased to 10% of that with 0.1% Triton X-100. Furthermore, His-BGPh was stabilized in the presence of 0.1% Triton X-100: the half-life of the activity at  $90^{\circ}\text{C}$  and pH 6.0 was 15 h. These results strongly indicate that BGPh is a thermostable membrane protein.

#### 3.2. The substrate specificity of BGPh

For BGPh both with or without His-tag, the optimum pH was 6.0 and the optimum temperature was over  $100^{\circ}\text{C}$ . The substrate specificity of His-BGPh was examined using *p*-Nph- $\beta$ -D-saccharides and  $\beta$ -D-glucosides as substrates. The specificity is summarized in Table 2 in comparison with that of S $\beta$ -gly [7,21]. His-BGPh hydrolyzed aryl glycosides efficiently, showing  $k_{\text{cat}}/K_m$  values decreasing in the order *p*-Nph- $\beta$ -D-Glcp > *p*-Nph- $\beta$ -D-Galp > *p*-Nph- $\beta$ -D-Xylp > *p*-Nph- $\beta$ -D-Manp.  $\beta$ -Linked glucose dimers tested were poorly hydrolyzed; the order of preference was  $\beta$ 1-3 >  $\beta$ 1-4  $\approx$   $\beta$ 1-6. The  $k_{\text{cat}}$  values of BGPh without His-tag for these  $\beta$ -linked glucose dimers approached  $400\text{ s}^{-1}$ , which is comparable with those of S $\beta$ -gly (Table 2). His-BGPh had approximately 50% of the activity of BGPh due to interference by the His-tag located at the N-terminus (data not shown). Surprisingly, the best substrates for His-BGPh were  $\beta$ -D-glucosides with long alkyl chains (LA- $\beta$ -D-Glcp). The  $K_m$  values decreased according to the elongation of the alkyl chain from  $\text{C}_1$  to  $\text{C}_{12}$ , although the  $k_{\text{cat}}$  value was constant (approximately  $35\text{ s}^{-1}$ ) for each alkyl- $\beta$ -D-Glcp. The  $k_{\text{cat}}$  values of native-type BGPh for LA- $\beta$ -D-Glcp approached  $70\text{ s}^{-1}$ , calculated on the basis of the value of His-BGPh, estimating a 50% decrease in the activity from the inhibitory effect of the His-tag. The value was also appre-

Table 1  
Cellular localization of the activity

Cell fractions	Activity after each treatment ( $A_{620}$ )				
	Sonication	Non-heated	Heated	Non-heated with 2.5% Triton X-100	Heated with 2.5% Triton X-100
Suspension I	0.585	0.585	0.567	0.485	0.428
Supernatant at $9000\times g$	0.112	ND	ND	ND	ND
Supernatant at $15000\times g$	ND	0.008	0.005	0.107	0.255
Supernatant at $100000\times g$	0.010	ND	ND	ND	ND
Fraction precipitated at $100000\times g$	0.478	ND	ND	ND	ND

The transformant *E. coli* BL21(DE3)/pET-15b/BGPh cells were used for this experiment. The enzyme reactions were performed at 90°C and pH 6 for 15 min using X-Glu as substrate, and then  $A_{620}$  was measured as shown in Section 2.

ND: not determined.

ciable, around 30% of that of S $\beta$ -gly (Table 2). The  $K_m$  value of His-BGPh for the hydrolysis of *n*-dodecyl- $\beta$ -D-Glcp (alkyl chain: C<sub>12</sub>) was extremely low, 30  $\mu$ M at 90°C and pH 6.0. Of the substrates examined thus far, the best substrate was *n*-dodecyl- $\beta$ -D-Glcp as shown in Table 2. The  $k_{cat}/K_m$  value of His-BGPh against *n*-dodecyl- $\beta$ -D-Glcp was five times higher than that of *p*-Nph- $\beta$ -D-Glcp and 870 times higher than that of laminaribiose. Even the value for *n*-octyl- $\beta$ -D-Glcp was 0.76 times higher than that of *p*-Nph- $\beta$ -D-Glcp and 128 times higher than that of laminaribiose. The  $k_{cat}/K_m$  value of S $\beta$ -gly against *n*-octyl- $\beta$ -D-Glcp, with the longest alkyl chain so far examined [21], was 0.4-fold higher than that for *p*-Nph- $\beta$ -D-Glcp and 0.48-fold higher than that for laminaribiose. Laminaribiose and cellobiose were not good substrates for the hydrolysis of His-BGPh because of their  $K_m$  values higher than 100 mM. His-BGPh also hydrolyzed cellotriose and cellotetraose with low efficiency: the kinetic parameters were not determined because of the extremely high  $K_m$  value, whereas S $\beta$ -gly was able to hydrolyze these oligosaccharides with high efficiency: the  $k_{cat}/K_m$  values descended in the order: cellotetraose > cellotriose > cellobiose. Thus, the substrate specificity of His-BGPh is different from those of the other  $\beta$ -glycosidases, including S $\beta$ -gly [7,17,21–23]. BGPh has a novel sub-

strate specificity with high efficiency to hydrolyze LA- $\beta$ -D-Glcp and low efficiency to hydrolyze any  $\beta$ -linked glucose dimer which is more hydrophilic than aryl- or alkyl- $\beta$ -D-Glcp. The subsite affinity ( $A(C_{11})$ ) to bind a long alkyl chain (C<sub>11</sub>) was calculated according to the following equation:  $A(C_{11}) = RT \ln[(k_{cat}/K_m) \text{ for } n\text{-dodecyl-}\beta\text{-D-Glcp} / (k_{cat}/K_m) \text{ for methyl-}\beta\text{-D-Glcp}]$ . The affinity was determined to be 4.26 kcal/mol. The value was reasonable when compared with the highest affinity (4.23 kcal/mol) known, that of the recognition of one glucose unit in the subsite structure of *Saccharomycopsis* amylase [13,14]. These facts indicate that the hydrophobicity of the aglycon part of the substrates is strongly recognized by the BGPh molecule and the hydrophobic substrates, including aryl- and LA- $\beta$ -D-Glcp, are hydrolyzed effectively with low  $K_m$  values due to hydrophobic interaction between the aglycon moiety and the BGPh molecule. Thus, BGPh might be useful to synthesize novel  $\beta$ -glycosides, including new biosurfactants, using its transglycosylation activity because of its stability in organic solvents (data not shown).

Henrissat proposed an alternate and complementary classification scheme for glycosyl hydrolases based on amino acid sequence similarities [24–26]. For example, glycosyl hydrolase

Table 2  
Comparison of the kinetic parameters between His-tagged BGPh from *P. horikoshii* and S $\beta$ -glu from *S. solfataricus* strain MT-4 against *p*-Nph- $\beta$ -D-saccharides and  $\beta$ -D-glucosides

Substrates	His-BGPh (90°C and pH 6.0)			S $\beta$ -gly <sup>a</sup> (75°C and pH 6.5)		
	$k_{cat}$ (s <sup>-1</sup> )	$K_m$ (mM)	$k_{cat}/K_m$ (mM <sup>-1</sup> s <sup>-1</sup> )	$k_{cat}$ (s <sup>-1</sup> )	$K_m$ (mM)	$k_{cat}/K_m$ (mM <sup>-1</sup> s <sup>-1</sup> )
Laminaribiose	184	138.23	1.33	908	1.0	908.0
Cellobiose	194	1698.18	0.11	1333	30.0	44.4
Cellotriose	ND <sup>b</sup>	ND	ND	197	3.0	66
Cellotetraose	ND	ND	ND	584	1.7	343
$\beta$ -Gentiobiose	ND	ND	ND	1360	100	14
<i>p</i> -Nph- $\beta$ -D-Glcp	79	0.35	225.67	542	0.5	1084.0
<i>p</i> -Nph- $\beta$ -D-Galp	123	1.30	94.34	1020	4.7	217.0
<i>p</i> -Nph- $\beta$ -D-Xylp	3	0.10	31.83	284	4.0	71.0
<i>p</i> -Nph- $\beta$ -D-Manp	2	0.14	14.60	NH <sup>c</sup>	NH	NH
Salicin	44	1.96	22.20	880	5.0	175.9
Methyl- $\beta$ -D-Glcp (alkyl: C <sub>1</sub> )	35	40.74	0.85	— <sup>d</sup>	—	—
<i>n</i> -Amyl- $\beta$ -D-Glcp (alkyl: C <sub>5</sub> )	31	2.02	15.11	256	1.1	232
<i>n</i> -Hexyl- $\beta$ -D-Glcp (alkyl: C <sub>6</sub> )	33	0.54	60.28	263	1.0	263
<i>n</i> -Octyl- $\beta$ -D-Glcp (alkyl: C <sub>8</sub> )	34	0.20	170.70	313	0.7	434
<i>n</i> -Nonyl- $\beta$ -D-Glcp (alkyl: C <sub>9</sub> )	39	0.08	471.57	—	—	—
<i>n</i> -Decyl- $\beta$ -D-Glcp (alkyl: C <sub>10</sub> )	37	0.08	469.62	—	—	—
<i>n</i> -Undecyl- $\beta$ -D-Glcp (alkyl: C <sub>11</sub> )	43	0.05	944.37	—	—	—
<i>n</i> -Dodecyl- $\beta$ -D-Glcp (alkyl: C <sub>12</sub> )	36	0.03	1152.90	—	—	—

<sup>a</sup>Cited from [7,21].

<sup>b</sup>ND: the parameters were not determined because of too high  $K_m$  values.

<sup>c</sup>NH: the substrate was not hydrolyzed by S $\beta$ -gly.

<sup>d</sup>—: The parameters were not reported in the references.

BG Ph	MPLKFPPEMFL	FGATSSSHQI	EGNNRW----	---NDWYYVE	Q-----IGK	37
BM Ph	M-----KPY	WGVVQSARQF	EMGDPYRRNI	DPRSDWYVW	RDYNIKNL	44
BM Pf	M---FPEKFL	WGVVQSARQF	EMGDKLRRNI	DTNIDWYVW	RDKTNIKGL	47
BG Pf	M--KFPKNFM	FGYSWSGQCF	EMGLP--GSE	-VESDWVWV	HDKENIASGL	45
Sb-gly	M-YSPFNSFR	FGWSQAGFQS	EMGTP--GSE	DPNIDWYKVV	HDPENMAAGL	47
Consensus	M--.FP..E.	FG..QSGFQF	EMG.P.-...	D...DW.VV	.D..NI..GL	50
<div>↑</div>						
BG Ph	LPYRSG-KAC	NHWELYRDDI	QLMTSLGYN	YRFSIEWERL	FF-----	78
BM Ph	VSGDLPEEGI	NNYELYEIDH	RLAKELGLNA	YQLTIIEWERI	FFCPTFNVEV	94
BM Pf	VSGDLPEEGI	NNYELYEKDH	EIARKLGLNA	YRIGIIEWERI	FFWPTTFIDV	97
BG Pf	VSGDLPENGP	AYWHLYKQDH	DIAELKGMDC	IRGGIIEWARI	FFKPTTFDVKV	95
Sb-gly	VSGDLPENGP	GYWGNVYTFH	DNAQKMGSKI	ARLNIEWERL	FFNPLPRPQN	97
Consensus	VSGDLPE.G.	N.WEIL..DH	..A.KLGLNA	YR..IIEWERI	EE.PT...V	100
<div>→</div>						
BG Ph	-----EENK	FNE-----	DAFMHYREII	DLLLTRGITP	105	
BM Ph	EFERD-GYGL	IKKVKIEKEH	LEELDCLANQ	KEVRHYLNVL	RNLKKIGFTT	143
BM Pf	DYSNESYNL	IEDVKITKDT	LEELDEIANK	REVAYYRSVI	NSLSKSGCFKV	147
BG Pf	DVEKD-EEGN	IISVDVPEST	ILHELEKIANM	EALHYRKYIY	SDWKEKGTTF	144
Sb-gly	FDESKQDVTE	VEINENELKR	LLEY---ANK	DALNHVREIF	KDLKSPGLYF	144
Consensus	..E.-....	I..V..E...	L..EL...AN.	..A..HYR..I.	..LK..GT..T.	150
<div>↑</div>						
BG Ph	LVTLLHFTSP	LMF-----	-----MKKG	FLREENLKHV	EKYIEKVA-E	142
BM Ph	FVTLLHQTNP	LMIHDPETTR	GNFQKARAFG	WVDERTIEEF	AKYAAYVAWK	193
BM Pf	IVNLNHFTLP	LMHDPPEAR	ERALTNRKNG	WVNPRTVIEF	AKYAAYIAWK	197
BG Pf	ILNLVHWPIL	LMIHDPPIVR	KLGPDRAPAG	WLDEKTVVEF	VKFAAFVAYH	194
Sb-gly	ILNLVHWPIL	LMHDPPIVR	R-GDFTGPGS	WLSTRTVYEF	ARFSAYIAWK	193
Consensus	IVNL..H..TLP	LM..HDPI..R	.....G	WL..ERTV..EF	AKYAAYVA..K	200
<div>↑</div>						
BG Ph	LLEHVKLVAT	FNEPMVYVMM	GYLTAY--WP	PFIRSPFKAF	KVAANLLKAH	190
BM Ph	FDNVDYVST	FLEPMVTAEL	GYLAPYVGWP	PGILNPSAAK	KVIINQIVAH	243
BM Pf	FGDIDVWST	FNEPMVVEL	GYLAPYSGFP	PGVLNPEAAK	LAILHMINAH	247
BG Pf	LDDLVDWST	MNEPMVYVYQ	GYINLRSQFP	PGYLSFEAAE	KAKFNLCQAH	244
Sb-gly	FDDLVDYEST	MNEPMVVGGL	GYVGKSGFP	PGYLSFELSR	RHMVNIICAH	243
Consensus	FDD..LD..WST	FNEPMV...L	GYL...YSGFP	PG..LSPEAA.	K...N..I..AH	250
<div>↑</div>						
BG Ph	ALAYELLHG-	-----KF	K--VGIVKN	I--PIILPAS	DKERDRKAAE	226
BM Ph	AFAYDSIKKF	-----SS	KP--VGIVLN	II--PAYPLDP	NDSKSVRAAE	282
BM Pf	ALAYRQIKKF	DTEKADKDSK	EPAEVGLIYN	NIGVAYPKDP	NDSKDVKAAE	297
BG Pf	ICAYDAIKEY	-----SE	KS--VGVIYA	F--AWH--DP	LAEEYKDEVE	280
Sb-gly	AFAYDGKSV	-----SK	KP--VGIIYA	N--SSF--QP	LTDKDMEAWE	279
Consensus	A..AYD..IK..	-----S.	KP--VGIIYN	..-.....DP	...KD..AAE	300
<div>↑</div>						
BG Ph	KADNLFNWHF	LDAIWSGKYR	GVF---KTYR	IP-QSDADHI	GNYITASEV	272
BM Ph	NYDLFHNRLF	LEAVNRGNVD	LDITGE-YTK	IPHIKRNDAI	GNYITREVV	331
BM Pf	NDNFFHSGLF	FEAIHKGKLN	IEFDGETFID	APYLKGNDAI	GNYITREVV	347
BG Pf	EIRKDD---	YE-FV-----	-----TILH	S--KGKLDAI	GNYITSLVY	312
Sb-gly	MAENDNRWWF	FDALIRGEIT	RGN--EKIVR	DDLKGRIDAI	GNYITRTVV	327
Consensus	.....F	.EAI..G...	----E....	.P.....DAI	GNYITR.VV	350
<div>↑</div>						
BG Ph	RHT---WNPL	K-----	FFFEVK---L	ADISERKTQM	GWSVYKGLY	307
BM Ph	KYVEPKYEEL	PLITFVGVEG	YGSGNPNLSL	SPDNNPTSDF	GWEVFPQGLY	381
BM Pf	TYQEPMPFSI	PLITFKGVQG	YGACRPGTL	SKDDRFPVSDI	GWELYPGMY	397
BG Pf	GAIDGHLVPL	P-----G	YGFMSERGTF	AKSGRPASDF	GWEMYPEGLE	354
Sb-gly	KRTEKGYVSL	G-----G	YGHGCERNVS	SLAGLPTSDF	GWEFFPEGLY	369
Consensus	...E....L	P-----G	YG.....L	S....P.SDF	GWE.YPEGLY	400
<div>↑</div>						
BG Ph	MALKKA-SFY	GRPLYITENG	IAITLIEWRV	EFIICHLQVY	HKAIELGLLV	356
BM Ph	DSTLEA-AHY	NKEVFITENG	IAADSKDILRP	RYIIDHNEV	KKLIENGILV	430
BM Pf	DSIVEA-HKY	GVVYVITENG	IAADSKDILRP	YYIASHIKMI	EKAFELGVEV	446
BG Pf	NLLKYLNNAY	ELPMIITENG	MADAADRYRP	HYLVSHLKAV	YNAMKEGALV	404
Sb-gly	DVLTKYWNFY	HLVYVITENG	IAADADYQRP	YYLVSHVYQV	HRAINSGALV	419
Consensus	D..L..A-..IV	..P.YITENG	IAAD..D..RP	..YI..SH...V	..KAIE..G..LV	450
<div>↑</div>						
BG Ph	FGYFVMSFMD	NMEWKEGFGP	RFGLVVDYQ	IFERRHRKSA	Y-VYGEIARS	405
BM Ph	GGYFHMALTD	NMEWANGFKI	RFGLYEVDLI	IKERIHHRRS	VEIYKKIME	480
BM Pf	FGYFHMALTD	NMEWALGFRM	RFGLYEVNLI	IKERIHHREKS	VSIFREIVAN	496
BG Pf	FGYLHMSLTD	NMEWAGGFRM	RFGLVYVDFE	IKKRYLRPSA	L-VFREIATQ	453
Sb-gly	FGYLHMSLAD	NMEWASGFSM	RFGLLVVDYN	IKRLVWRPSA	L-VYREIATN	468
Consensus	FGYFHMALTD	NMEW..GE..M	REGL..EVD..	IKER..HE..SA	..-VYREIA..	500
<div>↑</div>						
BG Ph	KEIKDELLKR	YGL---PELQ	L	423		
BM Ph	-GIE-----	-----	-	483		
BM Pf	NGVTKKIEE-	-----ELLR	G	510		
BG Pf	KEIPEELAHL	ADLKFTV--R	K	472		
Sb-gly	GAITDEIEHL	NSVPPVKPLR	H	489		
Consensus	..I..E....	...---.LR	.	521		

Fig. 1. Aligned amino acid sequences of five  $\beta$ -glycosidases from hyperthermophilic archaea. The abbreviations of the sources of the enzymes are: BGPh,  $\beta$ -glycosidase from *P. horikoshii*; BMPH, a  $\beta$ -mannosidase gene homolog from *P. horikoshii* [8,9]; BGPf,  $\beta$ -glucosidase from *P. furiosus* [17]; BMPf,  $\beta$ -mannosidase from *P. furiosus* [17]; S $\beta$ -gly,  $\beta$ -glycosidase from *S. solfataricus* [18]. The conserved residues, identified automatically by the GeneWorks program, are shown in the open boxes. The reversed open triangles indicate the location of the nucleophile (E324) and the putative acid/base catalyst (E155 and H111) with R75 in the spatial proximity of the nucleophile of BGPh. The arrow shows the prominent deletion of more than 30 residues found in BGPh.

←

family 1 is composed of exo-acting,  $\beta$ -specific enzymes with similar amino acid sequences. The five  $\beta$ -glycosidases, including BGPh from the archaea domain (as shown in Fig. 1), belong to family 1. Some family 1 glycosyl hydrolases also have glycosyl transferase activities. The *S. solfataricus*  $\beta$ -glucosidase has been implicated in the glycosylation of membrane lipid components [27]. Similarly, the enzymatic analysis of BMPf predicted its possible role in the synthesis of intracellular components including protein, membrane components or other compounds [17]. Since the localization of BGPh on *E. coli* membrane strongly indicates the intimate interaction of the enzyme and lipid components, the detection of BGPh on the *Pyrococcus* cell surface using antibody against the enzyme must be done to clarify its true function in the *Pyrococcus* cell.

### 3.3. The structural elements responsible for membrane localization and the conservation of residues forming the active site

The sequence alignment among BGPh and four different  $\beta$ -glycosidases, whose biochemical characteristics have been reported [7,17,21–23], is shown in Fig. 1. According to the phylogenetic analysis based on the alignment, the tree has three branches: one corresponding to a  $\beta$ -glycosidase group that includes BGPf and S $\beta$ -gly; another containing BMPH and BMPf, which were close to  $\beta$ -mannosidase. BGPh belongs to the third branch, located some distance from the first two branches. The polypeptide length of BGPh is also approximately 13% shorter than those of the other four  $\beta$ -glycosidases and might be one of the shortest sequences so far reported [8,17,18,28]. As shown in Fig. 1, the residues E155 and H111 of BGPh correspond to E206 and H150 as the putative acid/base catalyst in the S $\beta$ -gly molecule [28,29], whose steric structure has been reported [30]. The residues E324 and R75 of BGPh correspond to E387, the nucleophile, and R79 in the spatial proximity of the nucleophile [28,29]. The complex structure of *Bacillus polymyxa*  $\beta$ -glycosidase with the inhibitor gluconate has been reported [31]. The BGPh residues, Q19, H111, N154, E155, Y267, E324, W362, E369 and W370, are completely conserved (Fig. 1) and correspond to the *B. polymyxa*  $\beta$ -glycosidase residues Q20, H121, N165, E166, Y296, E352, W398, E405 and W406, which form the intimate interaction with the inhibitor [31].

To understand the localization mechanism of BGPh to the membrane, a major structural difference between BGPh and the other soluble  $\beta$ -glycosidases was analyzed using the sequence alignment and the steric structure of S $\beta$ -gly [30]. The S $\beta$ -gly molecule has the classic ( $\beta/\alpha$ )<sub>8</sub> barrel fold first seen in the structure of triose phosphate isomerase [32]. For BGPh, the prominent deletion of more than 30 residues was found after the 78th residue, as indicated in Fig. 1. The deletion region of BGPh corresponds to loops from residues 89 to 125 of S $\beta$ -gly, mainly shielding the helices 3 and 4 from solvent. The hydrophilic loops, which pack against the outer face of the barrel helices 3 and 4, were not present in the BGPh

molecule. The increased hydrophobicity at barrel helices 3 and 4 of BGPh is also indicated by the comparison of the hydrophobicity plots of S $\beta$ -gly (data not shown).

A tetrameric S $\beta$ -gly structure has been reported, in which these loop regions were located at the four edges of regular tetragonal molecular arrangement [30]. Since BGPh as well as S $\beta$ -gly were proved to be tetramer by gel filtration using buffer containing 0.01% Triton X-100 (data not shown), the deletion of these hydrophilic loops probably results in the exposure of helices 3 and 4 to the solvent at the four edges of the tetrameric structure. The exposed hydrophobic areas might interact with lipid components to embed the molecule in the membrane. Furthermore, the exposed hydrophobic areas may lead the hydrophobic substrates to the active site and bind them there. However, further studies using the crystallographic analysis are needed for a more definitive description of the detailed mechanism for recognition of the hydrophobic aglycon part, including a long alkyl-chain.

**Acknowledgements:** We are grateful to Emiko Yamamoto and Toshie Katakura for technical assistance.

### References

- [1] Cowan, D.A. (1992) The Archaeobacteria: Biochemistry and Biotechnology (Danson M.J., Hough, D.W. and Lunt, G.G., Eds.), pp. 149–170, Portland Press, London.
- [2] Tricone, A., Nicolaus, B., Lama, L. and Gambacorta, A. (1993) Indian J. Chem. (Section B) 32, 25–29.
- [3] Hudson, R.C., Schofield, L.R., Coolbear, T., Daniel, R.M. and Morgan, H.W. (1991) Biochem. J. 273, 645–650.
- [4] Kume, S. and Fujio, Y. (1991) J. Gen. Appl. Microbiol. 37, 25–34.
- [5] Tokatlidis, K., Dhurjati, P., Millet, J., Beguin, P. and Rossi, M. (1991) FEBS Lett. 282, 205–208.
- [6] Toone, E.J., Simon, E.S., Bednarski, M.D. and Whitesides, G.M. (1989) Tetrahedron 45, 5365–5422.
- [7] Tricone, A., Improta, R., Nucci, R., Rossi, M. and Gambacorta, A. (1994) Biocatalysis 10, 195–210.
- [8] Kawarabayasi, Y., Sawada, M., Horikoshi, H., Haikawa, Y., Hino, Y., Yamamoto, S., Sekine, M., Baba, S., Kosugi, H., Ho- soyama, A., Nagai, Y., Sakai, M., Ogura, K., Otsuka, R., Nakazawa, H., Takamiya, M., Ohfuku, Y., Funahashi, T., Tanaka, T., Kudoh, Y., Yamazaki, J., Kushida, N., Oguchi, A., Aoki, K., Yoshizawa, T., Nakamura, Y., Robb, F.T., Horikoshi, K., Masuchi, Y., Shizuya, H. and Kikuchi, H. (1998) DNA Res. 5, 55–76.
- [9] Kawarabayasi, Y., Sawada, M., Horikoshi, H., Haikawa, Y., Hino, Y., Yamamoto, S., Sekine, M., Baba, S., Kosugi, H., Ho- soyama, A., Nagai, Y., Sakai, M., Ogura, K., Otsuka, R., Nakazawa, H., Takamiya, M., Ohfuku, Y., Funahashi, T., Tanaka, T., Kudoh, Y., Yamazaki, J., Kushida, N., Oguchi, A., Aoki, K., Yoshizawa, T., Nakamura, Y., Robb, F.T., Horikoshi, K., Masuchi, Y., Shizuya, H. and Kikuchi, H. (1998) DNA Res. 5, 147–155.
- [10] Kaneko, T., Tanaka, A., Sato, S., Kotani, H., Sazuka, T., Miyajima, N., Sugiura, M. and Tabata, S. (1995) DNA Res. 2, 153–166.
- [11] Laemmli, U.K. (1970) Nature 227, 680–682.
- [12] Sakada, M. and Hiromi, K. (1976) J. Biochem. 80, 547–555.

- [13] Matsui, I., Ishikawa, K., Matsui, E., Miyairi, S., Fukui, S. and Honda, K. (1991) *J. Biochem.* 109, 566–569.
- [14] Matsui, I., Ishikawa, K., Matsui, E., Miyairi, S., Fukui, S. and Honda, K. (1992) *FEBS Lett.* 310, 216–218.
- [15] Suganuma, T., Matsuno, R., Ohnishi, M. and Hiromi, K. (1978) *J. Biochem.* 84, 293–316.
- [16] Hiromi, K. (1970) *Biochem. Biophys. Res. Commun.* 40, 1–6.
- [17] Bauer, M.W., Bylina, E.J., Swanson, R.V. and Kelly, R.M. (1996) *J. Biol. Chem.* 271, 23749–23755.
- [18] Cubellis, M.W., Rozzo, C., Montecucchi, P. and Rossi, M. (1990) *Gene* 94, 89–94.
- [19] Nei, M. (1987) *Molecular Evolutionary Genetics*, pp. 293–298, Columbia University Press, New York.
- [20] Kyte, J. and Doolittle, R.F. (1982) *J. Mol. Biol.* 157, 105–132.
- [21] Nucci, R., Moracci, M., Vaccaro, C., Vespa, N. and Rossi, M. (1993) *Biotechnol. Appl. Biochem.* 17, 239–250.
- [22] Kengen, S.W.M., Luesink, E.J., Stams, A.J.M. and Zehnder, A.J.B. (1993) *Eur. J. Biochem.* 213, 305–312.
- [23] Bauer, M.W. and Kelly, R.M. (1998) *Biochemistry* 37, 17170–17178.
- [24] Henrissat, B. (1991) *Biochem. J.* 280, 309–316.
- [25] Henrissat, B. and Bairoch, A. (1993) *Biochem. J.* 293, 781–788.
- [26] Henrissat, B. and Bairoch, A. (1996) *Biochem. J.* 316, 695–696.
- [27] Grogan, D.W. (1991) *Appl. Environ. Microbiol.* 57, 1644–1649.
- [28] Febbraio, F., Barone, R., D'Auria, S., Rocci, M. and Nucci, R. (1997) *Biochemistry* 36, 3068–3075.
- [29] Moracci, M., Trincon, A., Perugino, G., Ciaramella, M. and Rossi, M. (1998) *Biochemistry* 37, 17262–17270.
- [30] Aguilar, C.F., Sanderson, I., Moracci, M., Ciaramella, M., Nucci, R., Rocci, M. and Pearl, L.H. (1997) *J. Mol. Biol.* 271, 789–802.
- [31] Sanz-Aparicio, J., Hermoso, J.A., Martinez-Ripollo, M., Lequerica, J.L. and Polaina, J. (1998) *J. Mol. Biol.* 275, 491–502.
- [32] Banner, D.W., Bloomer, A.C., Petsko, G.A., Phillips, D.C., Pogson, C.I., Wilson, I.A., Corran, P.H., Furth, A.J., Milman, J.D., Offord, R.E., Priddle, J.D. and Waley, S.G. (1975) *Nature* 255, 609–614.