



Title	A Thermophilic Alkalophilic $\alpha$ -Amylase from <i>Bacillus</i> sp. AAH-31 Shows a Novel Domain Organization among Glycoside Hydrolase Family 13 Enzymes
Author(s)	Saburi, Wataru; Morimoto, Naoki; Mukai, Atsushi; Kim, Dae Hoon; Takehana, Toshihiko; Koike, Seiji; Matsui, Hirokazu; Mori, Haruhide
Citation	Bioscience, Biotechnology, and Biochemistry, 77(9), 1867-1873 <a href="https://doi.org/10.1271/bbb.130284">https://doi.org/10.1271/bbb.130284</a>
Issue Date	2013-09
Doc URL	<a href="http://hdl.handle.net/2115/67547">http://hdl.handle.net/2115/67547</a>
Rights	This is an Accepted Manuscript of an article published by Taylor & Francis in Bioscience biotechnology and biochemistry on 2013 Sep, available online: <a href="http://www.tandfonline.com/10.1271/bbb.130284">http://www.tandfonline.com/10.1271/bbb.130284</a>
Type	article (author version)
File Information	130710_amyL(HUSCAP).pdf



[Instructions for use](#)

1 Running title: Primary Structure of *Bacillus* sp. AAH-31  $\alpha$ -Amylase

2

3 **A Thermophilic Alkalophilic  $\alpha$ -Amylase from *Bacillus* sp. AAH-31 Shows a Novel**  
4 **Domain Organization among Glycoside Hydrolase Family 13 Enzymes**

5

6 Wataru SABURI,<sup>1,\*</sup> Naoki MORIMOTO,<sup>2,\*</sup> Atsushi MUKAI,<sup>1</sup> Dae Hoon KIM,<sup>1</sup> Toshihiko  
7 TAKEHANA,<sup>2</sup> Seiji KOIKE,<sup>2</sup> Hirokazu MATSUI,<sup>1</sup> and Haruhide MORI<sup>1,†</sup>

8 <sup>1</sup>*Research Faculty of Agriculture, Hokkaido University, N-9, W-9, Sapporo 060-8589,*  
9 *Japan*

10 <sup>2</sup>*Adeka Corporation, 7-2-34 Higashi-ogu, Arakawa-ku, Tokyo 116-8553, Japan*

11

12 Received April 3, 2013; Accepted May 30, 2013

13

14 <sup>†</sup>To whom correspondence should be addressed. Tel/Fax: +81-11-706-2497; E-mail:  
15 hmori@chem.agr.hokudai.ac.jp

16 \*These authors contributed equally to this work.

17

18 *Abbreviations:* GH, glycoside hydrolase; AmyL, *Bacillus* sp. AAH-31  $\alpha$ -amylase; PVDF,  
19 polyvinylidene difluoride; CD, cyclodextrin; TVAI, *Thermoactinomyces vulgaris* R-47  
20  $\alpha$ -amylase I; TVAII, *T. vulgaris* R-47  $\alpha$ -amylase II; CBM20, carbohydrate-binding  
21 module family 20; rAmyL, recombinant AmyL; pNPG5, *p*-nitrophenyl  
22  $\alpha$ -maltopentaoside; GA SBD, starch-binding domain of *Aspergillus niger* glucoamylase

23

24

1  $\alpha$ -Amylases (EC 3.2.1.1) hydrolyze internal  $\alpha$ -1,4-glucosidic linkages of starch and  
2 related glucans. *Bacillus* sp. AAH-31 produces an alkalophilic thermophilic  $\alpha$ -amylase  
3 (AmyL) of higher molecular mass, 91 kDa, than typical bacterial  $\alpha$ -amylases. In this  
4 study, the AmyL gene was cloned to determine its primary structure, and the  
5 recombinant enzyme, produced in *Escherichia coli*, was characterized. AmyL shows no  
6 hydrolytic activity towards pullulan, but the central region of AmyL (Gly395-Asp684)  
7 was similar to neopullulanase-like  $\alpha$ -amylases. In contrast to known neopullulanase-like  
8  $\alpha$ -amylases, the N-terminal region (Gln29-Phe102) of AmyL was similar to  
9 carbohydrate-binding module family 20 (CBM20), which is involved in the binding of  
10 enzymes to starch granules. Recombinant AmyL showed more than 95% of its  
11 maximum activity in a pH range of 8.2-10.5, and was stable below 65°C and from pH  
12 6.4 to 11.9. The  $k_{cat}$  values for soluble starch,  $\gamma$ -cyclodextrin, and maltotriose were 103  
13  $s^{-1}$ , 67.6  $s^{-1}$ , and 5.33  $s^{-1}$ , respectively, and the  $K_m$  values were 0.100 mg/mL, 0.348 mM,  
14 and 2.06 mM, respectively. Recombinant AmyL did not bind to starch granules. But the  
15 substitution of Trp45 and Trp84, conserved in site 1 of CBM20, with Ala reduced  
16 affinity to soluble starch, while the mutations did not affect affinity for oligosaccharides.  
17 Substitution of Trp61, conserved in site 2 of CBM20, with Ala enhanced hydrolytic  
18 activity towards soluble starch, indicating that site 2 of AmyL does not contribute to  
19 binding to soluble long-chain substrates.

20

21 **Key words:**  $\alpha$ -amylase; neopullulanase; glycoside hydrolase family 13;  
22 carbohydrate-binding module family 20; starch-binding domain

23

1  $\alpha$ -Amylases (EC 3.2.1.1), widely distributed in living organisms including animals,  
2 plants, and microorganisms, catalyze the hydrolysis of internal  $\alpha$ -1,4-glucosidic  
3 linkages of starch, glycogen, and related glucans with net retention of the anomeric  
4 configuration. Fungal and bacterial  $\alpha$ -amylases are utilized in various industries,  
5 including the production of isomerized sugar (liquefaction and saccharification of starch  
6 granules), bakery applications, and the de-sizing of textiles and paper. Alkalophilic  
7 liquefying enzymes are particularly useful as additive compounds for dishwashing and  
8 laundry detergents to remove food residues on dishes and food stains from clothes.<sup>1)</sup>

9 According to a sequence-based classification of glycoside hydrolases (GHs),<sup>2)</sup>  
10  $\alpha$ -amylases are classified into GH families 13, 57, and 119, while most enzymes are  
11 found in GH family 13. GH family 13 is the largest of the GH families. It is composed  
12 of various retaining GHs and glycosyltransferases that act on  $\alpha$ -glucans such as starch,  
13 maltooligosaccharides, and sucrose.<sup>3)</sup> Enzymes of this family show three-dimensional  
14 structures similar to enzymes belonging to GH families 70 and 77, and hence these  
15 families are categorized into clan GH-H, regarded as the  $\alpha$ -amylase family.

16 GH family 13 enzymes share three domains (A, B, and C).<sup>3)</sup> Domain A is the  
17 catalytic domain formed by a  $(\beta/\alpha)_8$ -barrel. Domain B is a long loop connecting the  
18 third  $\beta$ -strand and the third  $\alpha$ -helix, and a substrate binding cleft is formed at the  
19 interface of domains A and B. Domain C, following domain A, is formed by an  
20 antiparallel  $\beta$ -sheet. GH family 13 enzymes have been further classified into 40  
21 subfamilies based on amino acid sequences.<sup>4)</sup> Most subfamilies contain enzymes that  
22 utilize the same substrate and produce the same product.  $\alpha$ -Amylases are found in  
23 subfamilies 1, 2, 5, 6, 7, 15, 19, 20, 24, 27, 28, and 32.

24 *Bacillus* sp. AAH-31, first isolated from Japanese paddy soil, produces extracellular  
25 liquefying  $\alpha$ -amylase (AmyL), which is highly stable at high pH and temperatures.<sup>5)</sup>  
26 Many alkaline  $\alpha$ -amylases have been identified in various alkalophilic bacteria,<sup>6-11)</sup> but  
27 AmyL is more stable at high temperature than the known alkalophilic enzymes.  
28 Furthermore, this enzyme is stable in the presence of chelating reagents, including

1 EDTA, nitrilotriacetic acid, and sodium tripolyphosphate. These properties of AmyL are  
2 useful in automatic dishwashing detergents, because an automatic dishwashing machine  
3 generally operates at high pH at temperatures above 60°C. On SDS-PAGE analysis, the  
4 molecular mass of AmyL was estimated to be 91 kDa, considerably higher than those of  
5 the typical bacterial  $\alpha$ -amylases of GH family 13 (approximately 60 kDa) and close to  
6 enzymes possessing starch-binding domains.<sup>11,12)</sup> In the present study, the gene  
7 encoding AmyL was cloned to elucidate its primary structure, and the recombinant  
8 enzyme, produced in *Escherichia coli*, was characterized. Furthermore, a site-directed  
9 mutational study at the starch-binding site, which was identified through sequence  
10 analysis, was carried out.

11

## 12 **Materials and Methods**

13 *Analysis of N-terminal and partial internal amino acid sequences.* Ten  $\mu\text{g}$  of AmyL  
14 purified from the culture supernatant of *Bacillus* sp. AAH-31<sup>5)</sup> was further separated by  
15 SDS-PAGE and transferred to a polyvinylidene difluoride (PVDF) membrane  
16 (Sequi-Blot PVDF Membrane, Bio-Rad, Hercules, CA) by electroblotting with a  
17 semi-dry blotting apparatus.<sup>13)</sup> The band of AmyL, detected by staining with Coomassie  
18 Brilliant Blue, was cut off from the membrane and subjected to N-terminal sequence  
19 analysis with a Procise 492HT protein sequencer (Perkin Elmer, Waltham, MA). To  
20 analyze partial internal amino acid sequences of AmyL, the enzyme was digested by  
21 endoproteinase Asp-N (Takara Bio, Otsu, Japan). Twenty-five  $\mu\text{g}$  of AmyL was  
22 transferred to a PVDF membrane as described above, and the band of AmyL was  
23 incubated in 100  $\mu\text{L}$  of 50 mM sodium phosphate buffer (pH 8.0) containing 0.65 ng/ $\mu\text{L}$   
24 of endoproteinase Asp-N for 4 h at 37°C. Ten  $\mu\text{L}$  of the resulting supernatant was  
25 subjected to separation by HPLC under the following conditions: column, Capcell Pak  
26 C8 ( $\phi 4.6 \times 250$  mm, Shiseido, Tokyo); column temperature, room temperature; elution,  
27 0-60% acetonitrile; flow rate, 0.5 mL/min; detection, absorbance at 280 nm. The  
28 N-terminal amino acid sequences of the purified peptides were analyzed as described

1 above.

2

3 *Gene cloning of AmyL and construction of its expression plasmid.* To obtain partial  
4 *AmyL*, PCR was carried out with LA Taq DNA polymerase (Takara Bio), the genomic  
5 DNA of *Bacillus* sp. AAH-31, and primers designed from the N-terminal amino acid  
6 sequence and an internal amino acid sequence (P2, described below) of *AmyL*, S:  
7 5'-CARGAYTAYGARAAAYATHGT-3' (sense), and R: 5'  
8 -TCYTCIACRTAYTGYTCRTC-3' (antisense). The amplified DNA fragment was  
9 cloned into the *EcoRV* site of the pBluescript II SK (+) vector (Stratagene, La Jolla,  
10 CA), and sequenced with an ABI Prism 310 Genetic Analyzer (Applied Biosystems,  
11 Foster City, CA). Two µg of the genomic DNA digested with *HindIII* (Takara Bio) was  
12 cyclized through self-ligation with T4 DNA Ligase (Takara Bio), and used as the  
13 template for inverse PCR. The regions upstream and downstream of the section  
14 obtained by the PCR, as described above, were amplified with cyclized genomic DNA  
15 as template. Primers, HS1: 5'-ACCATCGGCTCCAGTTTG-3' (sense) and HR1:  
16 5'-TGGATGATTGTTTGAAGTCCAG-3' (antisense), were used for the upstream region,  
17 and primers, HS2: 5'-ACCAGTGGCAGATTGAGC-3' (sense) and HR2:  
18 5'-GGCACTTTCTACTTTGGAAAG-3' (antisense), for the downstream region. The  
19 further downstream region was amplified by inverse PCR, in which the genomic DNA  
20 cyclized after digestion by *BamHI* was used as template. The primers used for this PCR  
21 were BS: 5'-TGTCCCAAAGTGAGGACCAT-3' (sense) and BR:  
22 5'-CTCTGTTCCGTAATAGATGGC-3' (antisense).

23 *AmyL* was inserted into pET-21a vector (Novagen, Darmstadt, Germany) to  
24 overexpress it in *E. coli*. The DNA fragment amplified by PCR was cloned into pET-21a  
25 via the *NdeI* and *XhoI* sites, using Primestart HS DNA polymerase (Takara Bio), the  
26 genomic DNA of *Bacillus* sp. AAH-31, and primers  
27 5'-ACGCATATGCAGGATTATGAGAACATTGTCCTAAGA-3' (sense, *NdeI* site  
28 underlined) and 5'-GTTCCCTCGAGTTTGTTAATGTTGCCAGGACC-3'

1 (antisense, *Xho*I site underlined).

2

3 *Preparation of recombinant AmyL.* The transformant of *E. coli* BL21 (DE3)  
4 harboring the expression plasmid for *AmyL* was cultured at 37°C in 300 mL of  
5 Luria-Bertani medium supplemented with 100 µg/mL of ampicillin until  $A_{600}$  reached  
6 0.5, and 0.3 mL of 0.1 M isopropyl 1-thio-β-D-galactopyranoside (Wako Pure Chemical  
7 Industries, Osaka, Japan) was added to induce production of the recombinant protein.  
8 The induction culture was carried out at the same temperature for 3 h. Bacterial cells  
9 harvested by centrifugation were suspended in 10 mM Tris-HCl buffer (pH 7.0)  
10 containing 0.5 M NaCl and disrupted by sonication. After the cell debris was removed  
11 by centrifugation, the cell-free extract was applied to a Ni-chelating column by  
12 Chelating Sepharose Fast Flow (GE Healthcare, Uppsala, Sweden). The column was  
13 thoroughly washed with 10 mM Tris-HCl buffer (pH 7.0) containing 0.5 M NaCl, and  
14 the adsorbed protein was eluted with 0.15 M imidazole in the same buffer. Highly  
15 purified fractions, confirmed by SDS-PAGE, were collected and dialyzed against 10 mM  
16 Tris-HCl buffer (pH 7.0). The protein concentration was determined from the  
17 concentrations of amino acids after complete acid hydrolysis of the purified enzyme  
18 with 6 N HCl at 110°C for 24 h. Amino acids were measured by the ninhydrin  
19 colorimetric method using a JLC-500/V (JEOL, Tokyo).<sup>14)</sup>

20

21 *Standard enzyme assay.* A reaction mixture (100 µL) consisting of an appropriate  
22 concentration of the enzyme, 100 mM Tris-HCl buffer (pH 9.5), 5 mg/mL of soluble  
23 starch (Nacalai Tesque, Kyoto, Japan), and 0.025 mg/mL of Triton X-100 was incubated  
24 at 50°C for 10 min, and the resulting reducing sugar was measured by the  
25 dinitrosalicylic acid method,<sup>15)</sup> in which 0-5 mM maltose (Nacalai Tesque) was used as  
26 the standard reducing sugar. One U of enzyme activity was defined as the amount of  
27 enzyme producing 1 µmol of reducing sugar per min.

28

1     *Effects of pH and temperature on the activity and stability of recombinant AmyL.*

2     Enzyme activities at various pH values were measured to determine the optimum pH of  
3     AmyL. One hundred mM Britton-Robinson buffer (pH 6.0-11.0, composed of a mixture  
4     of 100 mM acetate, phosphate, and glycine, pH adjusted with 5 N NaOH) was used as  
5     the reaction buffer. The other reaction conditions were same as for the standard enzyme  
6     assay.

7     To evaluate pH and temperature stabilities, the enzyme solution was incubated at  
8     various pHs and temperatures, and residual activity was measured by the standard  
9     enzyme assay. For the pH studies, the enzyme solution was incubated in 100 mM  
10    Britton-Robinson buffer (pH 4-13) at 4°C for 24 h. For heat treatment, the enzyme  
11    solution was incubated at 30-80°C for 15 min, and then cooled down immediately on  
12    ice. The enzyme was considered to be stable in those ranges of pH and temperature over  
13    which the enzyme maintained more than 90% of its original activity.

14  
15    *Kinetic parameters.* The reaction velocities for the hydrolysis of various  
16    concentrations of soluble starch,  $\gamma$ -cyclodextrin ( $\gamma$ -CD), and maltotriose were measured.  
17    In the hydrolysis of soluble starch, a reaction mixture (500  $\mu$ L) consisting of 1.07-2.14  
18    nM enzyme, 100 mM Tris-HCl buffer (pH 9.5), 0.05-0.3 mg/mL of soluble starch, and  
19    0.025 mg/mL of Triton X-100 was incubated at 50°C. In the reaction to  $\gamma$ -CD, 0.1-1.0  
20    mM  $\gamma$ -CD (Nacalai Tesque) was added to the reaction mixture. Fifty  $\mu$ L of the reaction  
21    mixture was removed every 2 min up to 10 min, and the reducing sugar produced was  
22    measured by the copper bicinchoninate acid method,<sup>16)</sup> in which 0-300  $\mu$ M maltose was  
23    used as the standard. Initial reaction velocities, determined from the slopes of the  
24    progress curves, at various substrate concentrations were fitted to the Michaelis-Menten  
25    equation by nonlinear regression with Grafit version 7.0.2 (Erithacus Software, West  
26    Sussex, UK). Data shown are mean  $\pm$  standard deviation for three independent  
27    experiments.

28    For the analysis of maltotriose as substrate, a reaction mixture (100  $\mu$ L) consisting of



1 49.5-94.5 nM enzyme, 100 mM Tris-HCl buffer (pH 9.5), 1-8 mM maltotriose (Wako  
2 Pure Chemical Industries), and 0.025 mg/mL of Triton X-100 was incubated at 50°C for  
3 10 min, and heated at 100°C for 10 min to terminate the reaction. After the addition of  
4 200  $\mu$ L of 2 M Tris-HCl buffer (pH 7.0), the D-glucose released was measured by the  
5 glucose-oxidase peroxidase method.<sup>17)</sup> Kinetic parameters were determined as described  
6 above.

7  
8 *Estimation of the molecular mass of recombinant AmyL.* To determine the molecular  
9 mass of recombinant AmyL, gel filtration column chromatography was carried out. A  
10 Superose 12 column ( $\phi$ 10  $\times$  300 mm, GE Healthcare) was used, and column  
11 chromatography was performed in 10 mM Tris-HCl buffer (pH 8.5) containing 100 mM  
12 KCl at 0.5 mL/min. The eluted protein was detected by the absorbance at 280 nm. A set  
13 of gel filtration standards (Bio-Rad) was used to calibrate the molecular mass.

14  
15 *Analysis of the ction pattern of recombinant AmyL on p-nitrophenyl*  
16  *$\alpha$ -maltopentaoside (pNPG5).* One mL of a reaction mixture consisting of 5.6 nM  
17 recombinant AmyL, 1 mM pNPG5 (Calbiochem, Bad Soden, Germany), and 20 mM  
18 Tris-HCl buffer (pH 9.5) was incubated at 50°C. At 2, 5, 10, and 15 min, 100  $\mu$ L of the  
19 reaction mixture was removed and mixed with 10  $\mu$ L of 5 M acetic acid to terminate the  
20 reaction. The reaction products were analyzed by HPLC under the following conditions:  
21 injection volume, 10  $\mu$ L; column, YMC C18 Pack Pro ( $\phi$ 6.0 mm  $\times$  250 mm, YMC,  
22 Kyoto, Japan); column temperature, 40°C; eluent, 10% acetonitrile; flow rate, 1  
23 mL/min; detection, absorbance at 313 nm.

24  
25 *Preparation of W45A, W61A, and W84A AmyL mutants.* Site-directed mutagenesis  
26 was carried out with a Primestar Mutagenesis Basal Kit (Takara Bio). The expression  
27 plasmid for wild-type AmyL was used as template. The sequences of the primers used in  
28 the PCRs were as follows: for W45A,

1 5'-CTTGACGCGAGTTCAAACAATCATCCA-3' (sense, nucleotides mutated in  
2 boldface) and 5'-TGA ACTCGCGTCAAGGGGGGCCAAGCT-3' (antisense); for  
3 W61A, 5'-GGA ACTGCGAAAAGCAATCCCATCCCT-3' (sense) and  
4 5'-GCTTTTCGCGAGTTCCATCTGACTCATC-3' (antisense); for W84A,  
5 5'-GGGCAAGCGCTGCCGGGAGAGAATCTC-3' (sense) and  
6 5'-CGGCAGCGCTTGCCCATCCATCACATA-3' (antisense). The expression plasmid  
7 of each of the mutant enzymes was introduced into *E. coli* BL21 (DE3), and the  
8 enzymes were prepared in the same fashion as the wild type.

9

## 10 **Results**

### 11 *Cloning and sequence analysis of AmyL*

12 The N-terminal and partial internal amino acid sequences of AmyL purified from the  
13 culture supernatant of *Bacillus* sp. AAH-31 were analyzed. The N-terminal sequence of  
14 AmyL was QDYENIVLRGSLAPLDW, and the sequences of the peptides derived from  
15 AmyL by Asn-N endoprotease digestion were P1, DWAK; P2, DEQYVE; and P3,  
16 DWSSNNHPLTK. From PCR using primers, based on the sequence of the N-terminal  
17 and P2 and genomic DNA as template, a 488 bp DNA fragment was obtained (Fig. 1).  
18 The sequence of P3 was confirmed in the deduced amino acid sequence of the DNA  
19 fragment. To obtain the regions upstream and downstream of the DNA fragment  
20 obtained, inverse PCRs were carried out successively, and a 3,533-bp portion of the  
21 DNA sequence, including the full-length of *AmyL* (2,466 bp) encoding 821 amino acid  
22 residues, was determined. The nucleotide sequence has been deposited in the DDBJ  
23 database under accession no. AB775769 (Fig. 1). No ATG codon was found in the  
24 upstream region encoding the N-terminal amino acid sequence of AmyL. The TTG  
25 codon found 7 bp downstream of a putative ribosome binding site (5'-AGGAGG-3')  
26 with good complementarity to the 3'-terminal of *Bacillus subtilis* 16S rRNA<sup>18)</sup> was  
27 regarded as the initiation codon of the *AmyL* gene. There were putative -35 and -10  
28 regions, sequences of which are 5'-TTGTCA-3' and 5'-TATAAT-3', respectively,

Fig. 1

1 upstream of the ribosome binding site, and these regions were separated by 18 bp. An  
2 inverted-repeat sequence was found 6 bp downstream of the stop codon.

3 Twenty-eight amino acid residues at the N-terminal, which were not found in the  
4 native enzyme, are predicted to be a signal peptide by the Signal-P 4.0 program.<sup>19)</sup> The  
5 deduced amino acid sequence of mature AmyL is comprised of 793 amino acid residues,  
6 and the molecular mass estimated from the sequence, 91,274 Da, coincided well with  
7 the molecular mass of 91 kDa for native AmyL.<sup>5)</sup> The P1 sequence was also found in  
8 the deduced sequence of AmyL (Asp354-Lys357). A BLAST search of the mature  
9 AmyL sequence revealed that Gly395-Asp684 was assigned as the catalytic domain of  
10  $\alpha$ -amylase (PF00128),<sup>20)</sup> and the four conserved sequences of GH family 13 enzymes  
11 were found in this region, as follows: region I, 464-YDFVPNH-470; region II,  
12 542-GFRLDYAKG-550; region III, 571-FIFGEIWD-578; and region IV,  
13 635-FLDNHD-640. Asp546, Glu575, and Asp640 are predicted to be the catalytic  
14 nucleophile, the general acid/base catalyst, and the second catalytic aspartic acid  
15 fulfilling multiple functions essential for catalysis, including elevating the  $pK_a$  of the  
16 general acid/base catalyst and promoting substrate distortion.<sup>21)</sup> The catalytic domain of  
17 AmyL showed high sequence identity to neopullulanase-like  $\alpha$ -amylases belonging to  
18 subfamily 20 of GH family 13,<sup>4)</sup> although AmyL has no hydrolytic activity towards  
19 pullulan, unlike the other enzymes in this subfamily. The sequence identities of the  
20 catalytic domain of AmyL to *Geobacillus stearothermophilus* neopullulanase,  
21 *Thermoanaerobacter pseudethanolicus* cyclodextrinase, *Thermoactinomyces vulgaris*  
22 R-47  $\alpha$ -amylase I (TVAI), *T. vulgaris* R-47  $\alpha$ -amylase II (TVAIL), and *Bacillus*  
23 *acidpullulyticus* maltogenic amylase were 37%, 37%, 33%, 36%, and 37%, respectively.

24 In contrast to neopullulanase-like  $\alpha$ -amylases, the N-terminal region of AmyL  
25 (Gln29-Phe102) showed similarity to Carbohydrate-Binding Module family 20  
26 (CBM20, PF00686), involved in binding to starch granules.<sup>22)</sup> CBM20 is formed by a  
27  $\beta$ -sandwich fold, and most CBM20s have two sugar-binding sites, site 1 and site 2. Two  
28 Trp residues, Trp45 and Trp84, of AmyL corresponded to the Trp residues conserved in

1 site 1 of CBM20, which formed a carbohydrate-binding platform (Fig. 2). Lys and Asn,  
2 forming direct hydrogen bonding interactions,<sup>23)</sup> were also found at the equivalent  
3 positions of AmyL (Lys77 and Asn89). Trp61 of AmyL corresponded to the Trp residue  
4 conserved in site 2. Other amino acid residues responsible for binding to the  
5 carbohydrate in site 2 were not easily predicted, because this site is structurally more  
6 diverse than site 1. When the sequence of AmyL was compared with the starch-binding  
7 domain of *Aspergillus niger* glucoamylase (GA SBD), the best characterized enzyme in  
8 CBM family 20,<sup>23)</sup> no other amino acid residues in site 2 that contribute to the  
9 interaction with the carbohydrate were found at the corresponding positions of AmyL.

10 Pro234-Thr301 and Tyr714-Gly780 were assigned to the  
11 chitobiose/ $\beta$ -hexosaminidase C-terminal domain (PF13290) and a domain of unknown  
12 function (DUF3459, PF11941), respectively, and their functions are not yet understood .

13

#### 14 *Production and characterization of recombinant AmyL*

15 Recombinant AmyL with a His-tag at the C-terminal (rAmyL) was produced in *E.*  
16 *coli* BL21 (DE3), and purified to homogeneity by Ni-chelating column chromatography.  
17 Sixteen mg of the purified enzyme, with a specific activity of 60.9 U/mg, was obtained  
18 from a 300-mL culture. The molecular mass of rAmyL was estimated to be 107 kDa,  
19 close to the value obtained by SDS-PAGE (91 kDa), by gel filtration column  
20 chromatography, indicating that rAmyL is a monomer in solution. rAmyL retained more  
21 than 90% of its maximum activity in a broad pH range, 7.8-10.5, as observed for the  
22 native enzyme (Fig. 3).<sup>5)</sup> At high pH values, rAmy was less active than the native  
23 enzyme. It retained more than 90% of its original activity below 65°C (for 15 min at pH  
24 9.5) and from pH 6.4 to 11.9 (for 24 h at 4°C). At 70°C and higher temperatures, rAmy  
25 was more rapidly inactivated than the native enzyme. The native enzyme is stable in a  
26 pH range of 6.4-10.3, and thus rAmy was more stable at alkaline pH values.

27 The  $k_{cat}$  values of rAmyL for soluble starch,  $\gamma$ -CD, and maltotriose were  $103 \pm 9 \text{ s}^{-1}$ ,  
28  $67.6 \pm 2.0 \text{ s}^{-1}$ , and  $5.33 \pm 0.43 \text{ s}^{-1}$ , respectively, and the  $K_m$  values were  $0.100 \pm 0.028$

Fig. 2

Fig. 3

1 mg/mL,  $0.348 \pm 0.019$  mM, and  $2.06 \pm 0.33$  mM, respectively.

2 To determine the predominant productive binding mode, the product distribution  
3 from the hydrolysis of pNPG5 was analyzed by measuring the releasing velocity of each  
4 aglycone having a *p*-nitrophenyl group. The releasing velocity of *p*-nitrophenyl  
5  $\alpha$ -glucoside was the most prominent at 53% of the entire reaction velocity (the sum of  
6 the reaction velocities of the aglycones), followed by *p*-nitrophenyl  $\alpha$ -maltoside (34%).  
7 This result indicates that pNPG5 is hydrolyzed mainly through two productive binding  
8 modes covering subsites -4 to +2 and subsites -3 to +3.

9 A binding assay using raw cornstarch was carried out because a CBM20-like region  
10 was identified in the deduced amino acid sequence of AmyL. No decrease in enzyme  
11 activity of the supernatant was observed after incubation of 0.94-9.4  $\mu$ M AmyL at 4°C  
12 for 1 h with gentle shaking in the presence of 100 mg/mL of starch granules (data not  
13 shown), and thus AmyL did not adsorb onto the starch granules.

14

15 *Site-directed mutagenesis to three conserved Trp residues in the N-terminal*

16 *CBM20-like sequence of AmyL*

17 To confirm the function of the CBM-like region in the process of binding to soluble  
18 substrates, three conserved Trp residues (Trp45 and Trp84 from site 1 and Trp61 from  
19 site 2) were replaced with Ala. All the mutant enzymes were prepared as for wild-type  
20 AmyL, and the kinetic parameters for soluble starch,  $\gamma$ -CD, and maltotriose were  
21 measured (Table 1). The kinetic parameters for  $\gamma$ -CD were only minimally impacted by  
22 the mutations. Using maltotriose as substrate, W45A and W84A showed 1.7 to 1.9-fold  
23 lower  $K_m$  values, and W61A had a  $K_m$  similar to the wild type. In contrast to the  
24 reactions with oligosaccharides, W45A and W84A showed 2.2 to 2.4-fold lower  $k_{cat}$   
25 values and 2.5-fold higher  $K_m$  values for soluble starch than the wild type. In contrast to  
26 the other two mutated AmyLs, W61A had a 0.8-fold and a 2.3-fold lower  $k_{cat}$  and  $K_m$  for  
27 soluble starch, respectively, resulting in a 1.9-fold higher  $k_{cat}/K_m$  than the wild type.

28

Table 1

1

## 2 **Discussion**

3 In this study, we analyzed the primary structure of AmyL, the enzymatic  
4 characteristics of recombinant AmyL, and the function of the N-terminal domain. The  
5 amino acid sequence of the central region of AmyL showed high similarity to the  
6 catalytic domains of neopullulanase-like  $\alpha$ -amylases of GH family 13 subfamily 20 (Fig.  
7 1). The calcium ion bound to the conserved binding site of typical  $\alpha$ -amylases is not  
8 observed in neopullulanase-like  $\alpha$ -amylases.<sup>24)</sup> Similarly to neopullulanase-like  
9  $\alpha$ -amylases, AmyL also presumably does not have a calcium ion bound to the conserved  
10 binding site. The absence of a conserved calcium site in AmyL is thought to be the  
11 reason this enzyme showed high tolerance to chelating reagents. The amino acid  
12 residues responsible for thermostability and high resistance at alkaline pH were not  
13 easily predicted from the amino acid sequence of AmyL. Pro551 and Pro642, which  
14 were not found at the corresponding positions of the related enzymes, were predicted to  
15 be situated at the loops connecting the 4th  $\beta$ -strand and the 4th  $\alpha$ -helix, and the 7th  
16  $\beta$ -strand and the 7th  $\alpha$ -helix, respectively. These residues might be important for high  
17 thermostability, because Pro residues situated at loops connecting secondary structures  
18 were postulated to be important for high thermostability.<sup>25)</sup>

19 In contrast to AmyL, which is a monomer in solution, many neopullulanase-like  
20  $\alpha$ -amylases are homodimers. The N-terminal domain of neopullulanase-like  $\alpha$ -amylase  
21 is involved in the formation of the active site together with domain A of the other  
22 subunit.<sup>26-28)</sup> The active-site clefts of dimeric neopullulanase-like  $\alpha$ -amylases are narrow  
23 and deep, suitable for binding CDs.<sup>26)</sup> The N-terminal domain of TVAI, also a monomer  
24 similar to AmyL, interacts strongly with domains A and B, unlike other  
25 neopullulanase-like  $\alpha$ -amylases, to create the globular structure.<sup>24)</sup> These structural  
26 differences are responsible for the variations in substrate specificity. TVAI has higher  
27 activities towards starch and pullulan, and lower activities towards  $\alpha$ - and  $\beta$ -CDs than  
28 TVAII (a typical neopullulanase-like  $\alpha$ -amylase).<sup>29,30)</sup> AmyL shows high activity

1 towards starch and  $\gamma$ -CD, and low activity towards  $\alpha$ - and  $\beta$ -CDs,<sup>5)</sup> similarly to TVAI.  
2 Although the N-terminal region of AmyL does not show any similarity to that of TVAI,  
3 AmyL is also a monomeric enzyme like TVAI, and the active-site cleft of AmyL might  
4 be shallow and more suitable for binding to starch, as was observed for TVAI.  
5 TVAI has CBM34, which contains two maltooligosaccharide binding sites, at the  
6 N-terminal, in contrast to dimeric neopullulanase-like  $\alpha$ -amylases.<sup>31)</sup> It shows hydrolytic  
7 activity towards raw starch, and hence N-terminal CBM34 is considered to behave like  
8 a starch-binding domain. The N-terminal region of AmyL showed similarity to the  
9 portion of CBM20 involved in binding to starch granules. No adsorption of AmyL to  
10 starch granules was observed, but our site-directed mutational study revealed that the  
11 two Trp residues (Trp45 and Trp84) conserved in site 1 of CBM20 contribute to binding  
12 to a soluble long-chain substrate (Table 1), suggesting that the CBM20-like region of  
13 AmyL plays a role as substrate-binding domain. Both AmyL and TVAI have N-terminal  
14 sugar-binding domains followed by the catalytic domains, but AmyL has a  
15 substrate-binding domain structurally different from that of TVAI. To our knowledge,  
16 there is no report that enzymes belonging to GH family 13 subfamily20 have N-terminal  
17 CBM20, and hence AmyL is an enzyme having a novel domain organization. The  
18 W61A AmyL mutant showed higher catalytic efficiency for the hydrolysis of soluble  
19 starch than the wild type, and thus Trp61 of AmyL, corresponding to the conserved Trp  
20 residue in site 2 of CBM20, is not thought to be involved in sugar binding. Mutant  
21 AmyL showed a lower  $K_m$  value for soluble starch than the wild type, while the  $k_{cat}$   
22 value of this mutant enzyme was similar to that of the wild type, suggesting that Trp61  
23 disturbs the formation of the Michaelis complex. AmyL does not have a Tyr residue at  
24 the position corresponding to Tyr556 of GA SBD, which interacts mostly with  
25 carbohydrates,<sup>23)</sup> and site 2 of AmyL presumably does not act as a sugar-binding site. In  
26 the complex structure of *Bacillus cereus*  $\beta$ -amylase and maltose, maltose was observed  
27 to be bound only at site 1 of CBM20 at the C-terminal of the enzyme,<sup>32)</sup> and hence site 2  
28 of CBM20 of this enzyme does not appear to possess binding affinity for carbohydrates.

1 Neopullulanase-like  $\alpha$ -amylases hydrolyze pullulan, producing panose, but AmyL  
2 shows no hydrolytic activity towards this polysaccharide.<sup>5)</sup> This is an essential  
3 difference in substrate specificity between AmyL and neopullulanase-like  $\alpha$ -amylases.  
4 On comparison of the amino acid sequences of AmyL and neopullulanase-like  
5  $\alpha$ -amylases for which the three-dimensional structures have been solved, the loop  
6 connecting the 8th  $\beta$ -strand and the 8th  $\alpha$ -helix in AmyL is longer than that in  
7 neopullulanase-like  $\alpha$ -amylases, in which the conserved Asp on this loop is involved in  
8 the formation of subsite -2 (Fig. 4).<sup>31)</sup> In neopullulanase-like  $\alpha$ -amylases, short loop  
9 connecting the 8th  $\beta$ -strand and the 8th  $\alpha$ -helix appears to make space for binding to  
10 pullulan, and thus this long loop of AmyL might result in steric hindrance that disturbs  
11 the process of binding to pullulan.

Fig. 4

12 In this study, we found that AmyL has a catalytic domain similar to  
13 neopullulanase-like  $\alpha$ -amylases, although it has no hydrolytic activity towards pullulan.  
14 In addition, we found that the N-terminal CBM20-like domain contributed to the  
15 process of binding to soluble long-chain substrates. Mutational analysis revealed that  
16 two Trp residues conserved at site 1 of CBM20 were important for high activity towards  
17 soluble long-chain substrates. Structural insight into the substrate binding mode and  
18 specificity should be pursued through the three-dimensional structure analysis of AmyL.

19

## 20 **Acknowledgments**

21 We thank Mr. Tomohiro Hirose of the Instrumental Analysis Division, Equipment  
22 Management Center, Creative Research Institution, Hokkaido University for amino acid  
23 analysis.

24

25

26

## 27 **References**

28 1) Ito S, Kobayashi T, Ara K, Ozaki K, Kawai S, and Hatada Y, *Extremophiles*, **2**,



- 1 185-190 (1998).
- 2 2) Cantarel B, Coutinho PM, Rancurel C, Bernard T, Lombard V, and Henrissat B,  
3 *Nucleic Acids Res.*, **37**, D233-D238 (2009).
- 4 3) MacGregor EA, Janecek S, and Svensson B, *Biochim. Biophys. Acta*, **1546**, 1-20  
5 (2001).
- 6 4) Stam MR, Danchin EG, Rancurel C, Coutinho PM, and Henrissat B, *Protein Eng.*  
7 *Des. Sel.*, **19**, 555-562 (2006).
- 8 5) Kim DH, Morimoto N, Saburi W, Mukai A, Imoto K, Takehana T, Koike S, Mori H,  
9 and Matsui H, *Biosci. Biotechnol. Biochem.*, **76**, 1378-1383 (2012).
- 10 6) Horikoshi K, *Agric. Biol. Chem.*, **35**, 1783-1791 (1971).
- 11 7) Hayashi T, Akiba T, and Horikoshi K, *Agric. Biol. Chem.*, **52**, 443-448 (1988).
- 12 8) Boyer EW and Ingle MB, *J. Bacteriol.*, **110**, 992-1000 (1972).
- 13 9) Igarashi K, Hatada Y, Hagihara H, Saeki K, Takaiwa M, Uemura T, Ara K, Ozaki K,  
14 Kawai S, Kobayashi T, and Ito S, *Appl. Environ. Microbiol.*, **64**, 3282-3289 (1998).
- 15 10) Hagihara H, Igarashi K, Hayashi Y, Endo K, Ikawa-Kitayama K, Ozaki K, Kawai S,  
16 and Ito S, *Appl. Environ. Microbiol.*, **67**, 1744-1750 (2001).
- 17 11) Murakami S, Nishimoto H, Toyama Y, Shimamoto E, Takenaka S, Kaulpiboon J,  
18 Prousoontorn M, Limpaseni T, Pongsawasdi P, and Aoki K, *Biosci. Biotechnol.*  
19 *Biochem.*, **71**, 2393-2401 (2007).
- 20 12) Shirokizawa O, Akiba T, and Horikoshi K, *FEMS Microbiol. Lett.*, **58**, 131-135  
21 (1990).
- 22 13) Pluskal MG, Przekop MB, Kavonian MR, Vecoli C, and Hicks DA, *BioTechniques*,  
23 **4**, 272-283 (1986).
- 24 14) Moore S and Stein WH, *J. Biol. Chem.*, **176**, 367-388 (1948).
- 25 15) Miller GL, *Anal. Chem.*, **31**, 426-428 (1959).
- 26 16) McFeeters RF, *Anal. Biochem.*, **103**, 302-306 (1980).
- 27 17) Miwa I, Okuda J, Maeda K, and Okuda G, *Clin. Chim. Acta*, **37**, 538-540 (1972).
- 28 18) Murray CL and Rabinowitz JC, *J. Biol. Chem.*, **257**, 1053-1062 (1982).

- 1 19) Petersen TN, Brunak S, von Heijne G, and Nielsen H, *Nat. Methods*, **8**, 785-786  
2 (2011).
- 3 20) Punta M, Coggill PC, Eberhardt RY, Mistry J, Tate J, Boursnell C, Pang N, Forslund  
4 K, Ceric G, Clements J, Heger A, Holm L, Sonnhammer EL, Eddy SR, Bateman A,  
5 and Finn RD, *Nucleic Acids Res.*, **40**, D290-301 (2012).
- 6 21) Uitdehaag JCM, Mosi R, Kalk KH, van der Veen BA, Dikhuizen L, Withers SG, and  
7 Dijkstra BW, *Nat. Struct. Biol.*, **6**, 432-436 (1999).
- 8 22) Christiansen C, Abou Hachem M, Janecek S, Viksø-Nielsen A, Blennow A, and  
9 Svensson B, *FEBS J.*, **276**, 5006-5029 (2009).
- 10 23) Sormachi K, Le Gal-Coëffet MF, Williamson G, Archer DB, and Williamson MP,  
11 *Structure*, **5**, 647-661 (1997).
- 12 24) Kamitori S, Abe A, Ohtaki A, Kaji A, Tonzuka T, and Sakano Y, *J. Mol. Biol.*, **318**,  
13 443-453 (2002).
- 14 25) Watanabe K, Chishiro K, Kitamura K, and Suzuki Y, *J. Biol. Chem.*, **266**,  
15 24287-24294 (1991).
- 16 26) Kim JS, Cha SS, Kim HJ, Kim TJ, Ha NC, Oh ST, Cho HS, Cho MJ, Kim MJ, Lee  
17 HS, Kim JW, Choi KY, Park KH, and Oh BH, *J. Biol. Chem.*, **274**, 26279-26286  
18 (1999).
- 19 27) Kamitori S, Kondo S, Okuyama K, Yokota T, Shimura Y, Tonzuka T, and Sakano Y,  
20 *J. Mol. Biol.*, **287**, 907-921 (1999).
- 21 28) Hondoh H, Kuriki T, and Matsuura Y, *J. Mol. Biol.*, **326**, 177-188 (2003).
- 22 29) Ibuka A, Tonzuka T, Matsuzawa H, and Sakai H, *J. Biochem.*, **123**, 275-282  
23 (1998).
- 24 30) Shimura Y, Tonzuka T, and Sakano Y, *J. Appl. Glycosci.*, **45**, 393-395 (1998).
- 25 31) Abe A, Tonzuka T, Sakano Y, and Kamitori S, *J. Mol. Biol.*, **335**, 811-822 (2004).
- 26 32) Mikami B, Adachi M, Kage T, Sarikaya E, Nanmori T, Shinke R, and Utsumi S,  
27 *Biochemistry*, **38**, 7050-7061 (1999).
- 28

1 **Figure legends**

2

3 **Fig. 1.** DNA and Amino Acid Sequences of AmyL.

4 Numbers of amino acids and nucleotides are shown to the left and right, respectively.  
5 Asterisk indicates the stop codon. The N-terminal and partial internal amino acid  
6 sequences of AmyL purified from the culture supernatant of *Bacillus* sp. AAH-31 are  
7 shown in bold face. The ribosome binding site (RBS), inverted repeat, and -35 and -10  
8 regions are underlined. *Hind*III and *Bam*HI sites are boxed. Conserved regions of GH  
9 family 13 enzymes are indicated by double underlining. Arrows indicate primers used to  
10 obtain *AmyL*.

11

12 **Fig. 2.** Comparison of the N-Terminal Sequence of AmyL with CBM20.

13 The amino acid sequences of the N-terminal region of AmyL and CBM20s, the  
14 three-dimensional structures of which are known, were aligned with the Clustal W  
15 program (<http://clustalw.ddbj.nig.ac.jp/>). Amino acid residues conserved in sites 1 and 2  
16 are shown in boldface, and the circle and triangle indicate site 1 and site 2, respectively.  
17 AnGA, *A. niger* glucoamylase (Genbank ID: CAK38411.1); BcBA, *B. cereus*  $\beta$ -amylase  
18 (BAA34650.1); Bc251CGT, *B. circulans* 251 cyclodextrin glucanotransferase (CGTase,  
19 CAA55023.1); Bc8CGT, *B. circulans* 8 CGTase (CAA48401.1); GsCGT, *G.*  
20 *stearothermophilus* NO. 2 CGTase (CAA41770.1); TtCGT, *Thermoanaerobacterium*  
21 *thermosulfurigenes* EM1 CGTase (AAB00845.1); Novamyl, *G. stearothermophilus*  
22 C599 maltogenic  $\alpha$ -amylase (ABA38666.1); G4-Amy, *Pseudomonas stutzeri* MO-19  
23 maltotetraose-forming  $\alpha$ -amylase (AAA25707.1).

24

25 **Fig. 3.** Comparison of Physicochemical Properties between Native and Recombinant  
26 AmyL.

27 The pH activity curve (a), pH stability (b), and thermal stability (c) of the native (●)  
28 and recombinant (○) AmyL were compared. Data for the native enzyme are from

1 reference 5.

2

3 **Fig. 4.** Comparison of Partial Amino Acid Sequences between  $\beta$ -Strand 8 and  $\alpha$ -Helix 8  
4 of AmyL and Related Enzymes.

5 The amino acid sequences of AmyL and related enzymes were aligned with the  
6 Clustal W program. Amino acid residues completely conserved in the enzymes are  
7 shown in bold face. TVAI, *T. vulgaris* R-47  $\alpha$ -amylase I (Genbank ID: BAA02471.1);  
8 TVAII, *T. vulgaris*  $\alpha$ -amylase II (BAA02473.1); NPU, *G. stearothermophilus*  
9 neopullulanase (AAA22622.1); TMA, *Thermus* sp. IM6501 (AAC15072.1); BCD,  
10 *Bacillus* sp. cyclomaltodextrinase (AAA92925.1).

CCGGCTTAAaAGCGCTGGGTATCAATGATCCTGACTATGGCCACTATCTGGCAGAACGCTTTCCTGCTGAACGGC -676  
 GCAAACGTTCTATCGTTTTATGATGAAACTTTTGGGTGCTGGATGAACTCAAAGAAAATTACAGATTACTGCTTT -601  
 TAACCAACGGCTCACCTGACTTGCACCGGAGAACTTTCCACCTTCCCGACCTGACACCATACTTTGAACATA -526  
 TAGTCATTTCCGGTATTTCCGTAAGGAAGCCGTGATCCGCTCTATCTTTGAACATGCGCTGAACCTTGATGGCTC -451  
 TAGTTCCTGATGAAGCCATCATGGTAGGTGATAACTTGTGACGGATATCCTCGGTTCCTCACGGGCGAGGATGA -376  
 AAAATGTGTGGATCAACAGGAATGGAAAACCCCTCATCCGGAAGTGATTCCAGACTTTGAGATCAAAGTTTAA -301  
 CCGAACTCCCGCCACTGGTCCAAAAGTTGCGTAATGATGTTAGAAAAGTGTAGAAAACCTGAAAAGCCGGCTATCA -226  
 AAGAGGATGGCTGCTTTTTTATTTGATTTTGTTCAAAACAAAGTTTATTTTCAAATTTTATATCCAAATTC -151  
 TTGTCAACATTTCTCTTGATCAAGTATAATAATGAACATCGGTTTCATAAAAAATTAAGACAAGGAGTTTAGAA -76  
 -35 -10 RBS -1  
 TTGAAGAAATATTTATCTATTATCTTATCACTTTCTTTGTTCTATCTTTTTTTGTGGTACCGTTTGCACAAAAT 75  
 1 L K K Y L S I I L S L S F V L S F F V V P F A Q N  
 ACGGAGGCTCAGGATTATGAGAACATTCTCTTAAGAGGTAGCTTGGCCCCCTTGAAGTGGAGTTCAAACAATCAT 150  
 26 T E A S Q D Y E N I V L R G S L A P L D W S S N N H  
 CCATTAACAAAAGATGAGTCAAGTGAACCTTGGAAAAGCAATCCCATCCCTTCCAGGAGGACAACGCCTGGAA 225  
 51 P<sup>HS2</sup>L T K D E S D G T W K S N P I P L P G Q R L E  
 TTCAAATATGTGATGGATGGCAATGGCTGCCGGGAGAGAATCTCGTCTTTGATATTCCCAAACCGGCAACTAC 300  
 76 F K Y V M D G Q W L P G E N L V F D I P Q T G N Y  
 ATCTTTATCTTTCACCGGACAACCGCGTAAAGTGGATGTCATTCTGGTTGAAGCTGATGGTAAAGTCAAGCTG 375  
 101 I F I F H P D N Q R K V D V I L V E A D G K V T L  
 CTCTTGACCGTGCCGGACAACACACCGCTCAGCATTTGTGCCACCATCCGCTCCAGTTTGAACAATTTAATAT 450  
 126 L L T V P D N T P S R I V P<sup>HS1</sup>T I G S S L N N F N Y  
 TCCGTGACAAAGCTTTCCAAAGTAGAAAGTGCCGAGAACCAGTGGCAGATTGAGCTTTCCGGTGAACCGGCGAG 525  
 151 S V T K L S K V E S A<sup>HR2</sup>E<sup>HS2</sup>N O W Q I E L S G E P G Q  
 GAGTTTTCATATCTTTATGCGCTGGGGATGAACGAGTACGTAGAGACAGAGATGCCCCCGTACAGCCACTTTC 600  
 176 E F S Y L Y A L G D E Q Y V E D<sup>R</sup>R D A P R T A T F  
 TTAGAAGATCATCTCGTTATGAAGATGTGGTAGAATCCTGGAAAAGCTGTGCCCATTTGCCAAAATGTGTCTCAC 675  
 201 L E D H L V I E D V V E S W K A V P I A K N V S H  
 AATTTAATCATGAGCCCTTTATCCCGGTTCCGAAGACGACGTGAACATCACTGTTTATGTTGATCACTTACGGC 750  
 226 N F N H E P F I P G S Q D D V N I T V H V D H Y G  
 ACAGTAGATTCCAGGGCCATTTACTTTACCCTGACGGCTCCTCGCGCTTGGAAAAAGAGGCGATGCCGCCAAT 825  
 251 T V D S G A I Y F T T D G S S P L G K R G D A A N  
 GGCATGTGGTGCCGTTACAAGTGACATCTCCTCTGAAAACAGTGACGGAAACCATTCGTTCTGTGTGACGGGA 900  
 276 G N V P L Q V T S S E N S D G T I R S V L T G  
 ACAATTTCCAAAGCAGCCGATTATACACCGTGAAGTACCGCATAGATGTCTGGGACAGTCAAAGTAAACGATGAC 975  
 301 T I P K Q P D Y T P V K Y R I D V W D S Q S N D D  
 CACTCCCAATTTGCTGACAATAACAGCCTTGTGCCGGAACAGGCCACTGAGTTTGGCTTACTACGTGGAAGAATTT 1050  
 326 H S Q F A D N N S L V P E Q A T E F A Y Y A E F  
 AAATCTCCAGACTGGCCAAAGATGCGGTGATCTACATCTTTGTGGATCGTTTTAAAGACGGCGATCCATCC 1125  
 351 K S P D W A K D A V I Y H I F V D R F K D D P S  
 AATAACGAACCTGTGATGAATCGTTCCCTATGAGAAGCATTGAAAGGCTGGATGGGCGGAGACCTGGAAAGGT 1200  
 376 N N E P V D E S L P Y A E R L K G W M G G D L E G  
 GTCAAACAGAAGCTGGACTATTTGGAGGAAGTGGCGTTAACGTCATTTGGCTCTCTCCCGTATTTGAAGGCCA 1275  
 401 V K Q K L D Y L E E L G V N V I W L S P V F E G P  
 TACTCGCATGGCTACCATCCACAGACTTTAAACAAGTGGATCCCTCGCTTCGGAGACAAGGAACGTTGCAATCT 1350  
 426 Y S H G Y H P T D F K Q V D P R F G D K E L Q Q S  
 CTGATTGAGGAGCCATCAACGGGACATAAAGTGAATGATGACTTTGTCCAAATCATTCTAATCCCGCAT 1425  
 451 L I E E A H Q R D M K V I Y D E V P N H S S N R H  
 CCGTCTTTCAAGCGCTTTCAACCGGGAAGGACAGCCCTTACTATTACTGGTACACCTTTACCAACTGGCCT 1500  
 476 P F F D A F N R G K D S P Y Y Y W Y T F T N W P  
 TACGAATAAAGACATTTATGGAATCGATGAACCTCCCAATTAACAATGATTATCCCGAAACCCGTGAATAT 1575  
 501 Y E Y K T F Y G I D E L P Q L N N D Y P E T R E Y  
 ATGCTGATGATGGTGGCCTTTTGGTTGTAGAGTTGGACTTTGACGGATTGAGATTAGACTYAGCCAAAGG 1650  
 526 M L Y D V V P F W L L E L D F D G F R L D Y A K G  
 CCAAGCTACAGTTTCTGGGTAGACTTCCGTACGCGCTCAAAAAATGAAGCCAGATGCCCTTTATATTCCGGCGAA 1725  
 551 P S Y S F W V D F R H A V K K M K P D A F I F G E  
 ATTTGGGACAACCGGGACAAGATTAATTCCTATGCGGCAAGCTGGATGGGGCTTTGGATTTCCCGCTGCAAAAGC 1800  
 576 I W D N R D K I N S Y A G K L D G A L D F P L Q S  
 GCGTTAGTGGACATTTGCTTATAATCACCAATGTCCAGGGTGGCAATACGATTAGAGACAATCGAACACT 1875  
 601 A L V D T F A Y N H P M S R V A N T I R D N L N T  
 TACCATCTGAATACATGATGGTGGCTTCTGGATAACCATGATCTCCCGCTTCTTCCAGTCCAGAGGG 1950  
 626 Y H P E Y M M V T F L D N H D L P R F L F Q S R G  
 GATAGTGCCAAACTGAAACTGGCCGCAACAGCCAGTTTACTTTACCGGCGGATCCCGCCATCTATTACGGAA 2025  
 651 D S A K L K L A A T A Q F T L P G I P A I Y Y G T  
 GAGATAGGCTTTGTCCAAAGTGAGGACCATAATCAATACACAGATTGGAGAGACCGGTGGTTCCGTTGAAATGATG 2100  
 676 E<sup>BR1</sup>I G L S Q S E D H N Q Y T D W R D R W F R E M M  
 CCTTGGGATGAAAAGCAACAGGACTTGAATTTGAAGCTTACTACAAACAGCTTATCGATCTCAGACACCGGGAA 2175  
 701 P W D E K Q Q D L K L K A Y Y K Q L I D L R H R E  
 GAAGCCTTCAAACAGGAGTACAATGAACTGTACGTTGATCAAGACCTGTTGTCTTTGAACGTAACACAAA 2250  
 726 E A F K T G E Y N E L Y V D Q D L F V F E R K H K  
 AACAAACACTTGTGTCATCAACAAAGGCAACACCAAGAGGTGGCTATTGACGAGTTGCTTAACTTA 2325  
 751 N K H F I V I I N K G N T T Q E V A I D E L L N L  
 AGACGTACCAACGGCATTTCCTTCAAGACGCCCTGAACAAAGGAAAGTGATCAAGCAAATCGCCAGGAGAA 2400  
 776 R R T N G I S L Q D A L N K G K V I K A N R Q G E  
 CTTAAAGTGAAGAGTTTCTCCAGCTTCCAGCCACTTTTACGGGTCTGGGCAACATTAACAAATAATTTGGGAA 2475  
 801 L K V K S S P A S A T I L R V L G N I N K \*  
 CTCCCATGTACTCTAAAACATGGGAGGTTTTTTGAGCTTGGAAATCCATGAACCTTAAACCAACCATAGTGAA 2550  
 TATCATGGATATGTGAGTGTCTTATAAACATTTGCTGGGCAAAATTCCTCAGGGGTGATCGAAAGTGGAGAT 2625  
 GTATCTCATGTTGCTTGGGCTGATATCAAAGGAAGTCTCTTGTCTCAAACATCAAAGGAGTGTCTATCTTAAA 2700  
 TCGTTCCGGGAGGCACCAATGTGCTCGTTGAAGTGTCCGGATTACCCCTTATCGTCCAGGAGGAAAAAACGT 2775  
 CCCCATTGGCCCCCATGATTCCACATCCATGAATTTGGAACATGTGAAATCGGAAATCTTAAGGATCC 2846

Fig. 1, Saburi et al.

AmyL 29: -----QDYENIVLRGSLAPL-DWSSN-NHPLTKDES-----DGT  
 AnGA 543: -----FDLTATT--TYGENIYLVGSISQLGDWETSDGIALSADKY---TSSDPL  
 BcBA 448: ----TPVMQTI VVKNV P-TTIGDTVYITGNRAELGSWDTK-QYPIQLYYD----SHSND  
 Bc251CGT 612: ----QVSVRFV VNNAT-TALGQNVYLTGSVSELGNWDP--AKAIGPMYN-QVVYQYPN  
 Bc8CGT 617: ----DQVTVRFV VNNAS-TTLGQONLYLTGNVAELGNWSTG-STAIGPAFN-QVIHQYPT  
 GsCGT 609: ----DQVSVRFV VNNAT-TNLGQNIYIVGNVYELGNWDT--SKAIGPMFN-QVVYSYPT  
 TtCGT 609: ----NQICVRFV VNNAS-TVYGENVYLTGNVAELGNWDT--SKAIGPMFN-QVVYQYPT  
 Novamyl 613: ----TQTSVVFTVKSAPPTNLGDKIYLTGNIPELGNWSTDTSGAVNNAQGPLLAPNYPD  
 G4-Amy 445: GEPGALVSVSFRCDNGA-TQMGD SVYAVGNVSQLGNWSPAAALRLTDTSG-----YPT

AmyL 61: <sup>▲</sup>WKSNIPIPLPGGQRLEF<sup>●</sup>KYVM----DGQ---WLPGENLVFDIPQTG-----NYIFIFH-  
 AnGA 587: WYVT-VTLPAGESFEYK<sup>●</sup>FIR---IESDDSV<sup>●●</sup>EWESDPNREYTVPQACGTSTATVTDTW--  
 BcBA 499: WRGN-VVLP<sup>●</sup>PAERNIEFKAFIK--SKDGTVK<sup>●●</sup>SWQTIQQSWNPVPLKT----TSHTSSW--  
 Bc251CGT 662: WYYD-VSVPAGKTIEFK<sup>●</sup>FLK---KQGS-TVTW<sup>●●</sup>E<sup>●</sup>GGSNHTFTAPSSG---TATINVNW--  
 Bc8CGT 669: WYYD-VSVPAGKQLEFK<sup>●</sup>FFK---KNGS-TITW<sup>●●</sup>ESGSNHTFTTPASG---TATVTVNW--  
 GsCGT 660: WYID-VSVPEGKTIEFK<sup>●</sup>FIK---KDSQGNVTW<sup>●●</sup>ESGSNHVYTTPTNT---TGKIIVDW--  
 TtCGT 660: WYYD-VSVPAGTTIQFK<sup>●</sup>FIK---KNGN-TITW<sup>●●</sup>E<sup>●</sup>GGSNHTYTPSSS---TGTIVVNW--  
 Novamyl 668: WFYV-FSVPAGKTIQFK<sup>●</sup>FFI---KRDGTIQW<sup>●●</sup>ENGSNHVATTPPTGA---TGNITVTWQN  
 G4-Amy 497: WKGS-IALPAGQNEEWK<sup>●</sup>CLIRNEANATQVRQW<sup>●●</sup>GGANNSLTPSEGA---TTVGRL----

Fig. 2, Saburi et al.

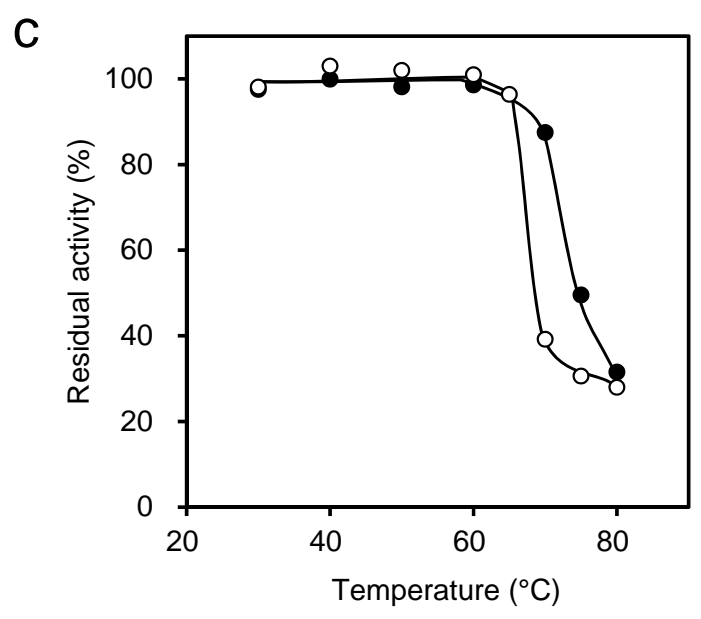
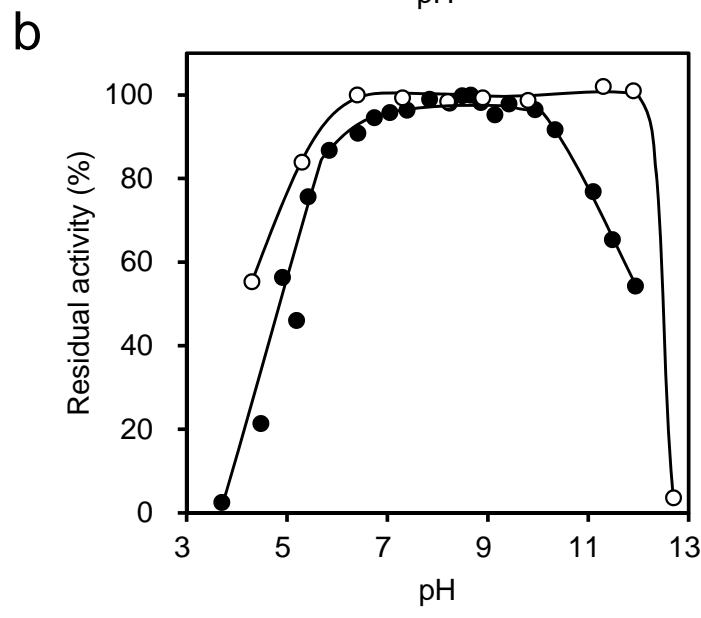
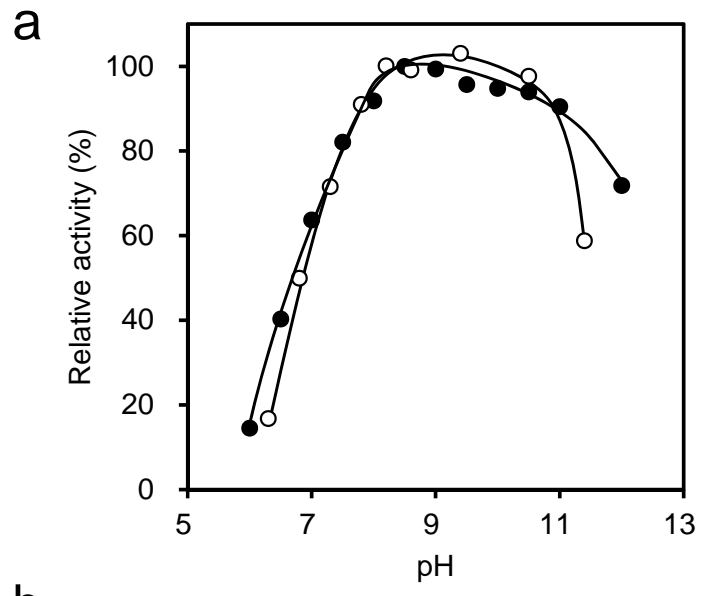


Fig. 3., Saburi, et al.

	$\beta$ -strand 8 ➔	$\alpha$ -strand 8 ▬
AmyL 668:	IPAIYYGTEIGLSQSEDHNQYTDWRDRWFREMPWDEKQQDLKPKAYYKQLIDL	RHRE
TVA1 529:	TPTIYYGDEYGMQGGADP-----	DNRRSFDWSQATPSNSAVALTQKLITIRNQY
TVA2 449:	TPLIYYGDEIGMAGATDP-----	DCRRPMIWEEKEQNRGLFEFYKELIRLRHRL
NPU 452:	SPCIYYGDEIGMTGGNDP-----	ECRKCMVWDPMQQNKELHQHVKQLIALRKQY
TMA 452:	SPCIYYGDEIGMTGGNDP-----	ECRKCMVWDPEKQNKELYEHVKQLIALRKQY
BCD 449:	TPCIYYGDEVGLDGGHDP-----	GCRKCMEWDETKHDKDLFAFYQTVIRLRQAH

Fig. 4, Saburi et al.



**Table 1.** Comparison of Kinetic Parameters between Trp Mutated AmyL and the Wild Type

Enzyme	Soluble starch			$\gamma$ -CD			Maltotriose		
	$k_{\text{cat}}$ ( $\text{s}^{-1}$ )	$K_{\text{m}}$ ( $\text{mg}\cdot\text{mL}^{-1}$ )	$k_{\text{cat}}/K_{\text{m}}$ ( $\text{s}^{-1}\cdot\text{mg}^{-1}\cdot\text{mL}$ )	$k_{\text{cat}}$ ( $\text{s}^{-1}$ )	$K_{\text{m}}$ ( $\text{mM}$ )	$k_{\text{cat}}/K_{\text{m}}$ ( $\text{s}^{-1}\cdot\text{mM}^{-1}$ )	$k_{\text{cat}}$ ( $\text{s}^{-1}$ )	$K_{\text{m}}$ ( $\text{mM}$ )	$k_{\text{cat}}/K_{\text{m}}$ ( $\text{s}^{-1}\cdot\text{mM}^{-1}$ )
Wild type	103 $\pm$ 9	0.100 $\pm$ 0.028	1030	67.6 $\pm$ 2.0	0.348 $\pm$ 0.019	194	5.33 $\pm$ 0.43	2.06 $\pm$ 0.33	2.59
W45A	42.9 $\pm$ 4.1	0.251 $\pm$ 0.028	171	76.9 $\pm$ 6.0	0.503 $\pm$ 0.032	153	4.26 $\pm$ 0.20	1.21 $\pm$ 0.18	3.52
W61A	84.6 $\pm$ 7.0	0.0439 $\pm$ 0.0151	1930	69.5 $\pm$ 2.2	0.353 $\pm$ 0.039	197	4.71 $\pm$ 0.32	1.82 $\pm$ 0.44	2.59
W84A	47.7 $\pm$ 6.8	0.248 $\pm$ 0.026	192	73.0 $\pm$ 7.3	0.381 $\pm$ 0.038	192	4.05 $\pm$ 0.10	1.10 $\pm$ 0.09	3.68

Data are mean  $\pm$  standard deviation for three independent experiments.