HOKKAIDO UNIVERSITY

Title	Structure-function relationship of GH13_31 α-glucosidase from Bacillus sp. AHU2216 [an abstract of entire text]
Author(s)	Auiewiriyanukul, Waraporn
Description	この博士論文全文の閲覧方法については、以下のサイトをご参照ください。 https://www.lib.hokudai.ac.jp/dissertations/copy-guides/
Degree Grantor	北海道大学
Degree Name	博士(農学)
Dissertation Number	甲第13326号
Issue Date	2018-09-25
Doc URL	https://hdl.handle.net/2115/75528
Type	doctoral thesis
File Information	Waraporn_Auiewiriyanukul_summary.pdf

Hokkaido University Collection of Scholarly and Academic Papers : HUSCAP

博 士 論 文 の 要 約

博士の専攻分野の名称： 博　士（農学）　　氏名　Waraporn Auiewiriyanukul

学 位 論 文 題 名

Structure-function relationship of GH13_31 α-glucosidase from *Bacillus* sp. AHU2216

（*Bacillus* sp. AHU2216 株由来 GH13_31α-グルコシダーゼの構造と機能に関する研究）

α-Glucosidase (AGase; EC 3.2.1.20) catalyzes the hydrolysis of the non-reducing end α-glucosidic linkage of the substrates to produce α-glucose. AGase is involved in the amylolytic pathway of various organisms including microorganisms, plants, and animals. This enzyme shows a wide variety of substrate specificity depending on the enzyme origins, and is classified into three groups based on substrate specificity: group I enzymes prefer to hydrolyze heterogeneous substrates, (e.g., sucrose and aryl α-glucosides) to the homogeneous substrates (e.g. maltose). Group II and III enzymes are specific to homogeneous substrates rather than heterogeneous substrates. Group III enzymes have high activity to polysaccharides. In the classification of glycoside hydrolases based on amino acid sequence, AGase is mainly classified into glycoside hydrase families (GH) 13 and GH31. Roughly, GH13 includes group I enzymes, and GH31 includes group II and III enzymes. The origins of GH13 AGases are predominantly found in the bacteria, yeast (*Saccharomyces cerevisiae*), and insects. On the other hand, AGases of plants, animals, fungi, yeast (*Schizosaccharomyces pombe*), and some species of bacteria are the members of GH31.

GH13 is one of the largest families and contains various amylolytic enzymes such as α-amylase (EC 3.2.1.1), cyclodextrin glucanotransferase (EC 2.4.1.19), and so on. This family is further divided into 42 subfamilies based on the phylogenetic analysis. GH13 enzymes are multidomain proteins containing three common domains A, B, and C. Domain A is a catalytic domain formed by a $(\beta/\alpha)_8$-barrel fold. Domain B is the variable length of a large loop between the third β-strand and third α-helix. Domain C, connecting to the domain A, folds into antiparallel β-strands of the Greek key structure. Domain C is suggested to stabilize the catalytic domain and help substrate-binding. The substrate binding sites of GH13 enzymes are divided into two classes: cleft-shaped substrate-binding sites for endo-type enzymes and pocket-shaped substrate-binding sites for exo-type enzymes. The active site of GH13 enzymes is made up of several subsites. Each subsite is capable of binding to one glucosyl residue of the substrate. The subsites themselves are composed of amino acid residues situated on loops

that connect the C-terminal end of β-strand to the N-terminal end of the adjacent α-helix of the catalytic domain A. GH13 enzymes share seven highly conserved regions (CSR); CSRs include the classic conserved regions I, II, III, and IV on β-strand 3, 4, 5, and 7 of domain A, respectively. GH13 enzymes have catalytic triad corresponding to As206, Glu230, and Asp297 in Taka-amylase A, which are situated on the β-strand 4, 5, and 7 of the domain A, respectively, and act as nucleophile, general acid/base, and transition state stabilizer, respectively. For all GH13 enzymes, substrate binding mode at subsite -1 is similar to each other, but those at subsites +1 and +2 varies and give a wide variety of specificity of the enzymes.

GH13 AGases fall into subfamily 17 (GH13_17), GH13_23, GH13_30, GH13_31, and GH13_40. GH13 AGases have high amino acid sequence similarities to oligo-1,6-glucosidase (EC 3.2.1.10), dextran glucosidase (3.2.1.70), sucrose isomerase, isomaltulose synthase, and trehalose synthase. GH13 AGase-type enzymes have two extra α-helices on β→α loop 8 of domain A, which form domain B′. Domain B′ interacts with the domain B. Because of the presence of domain B′, the enzymes have the pocket-shaped substrate-binding sites. In addition to the common residues found in GH13 enzymes [His (region I), Arg (region II), and His (region IV)], Asp and Arg on β→α loop 2 of domain A and domain B′, respectively, form salt-bridge to bind 4-OH of the glucosyl residue at the non-reducing end of the substrates. Although some enzymes have hydrolytic activities to sucrose and/or trehalose, GH13 AGases are roughly divided into two groups based on the α-glucosidic linkage specificity: α-(1→4)-glucosidic linkage-specific enzyme and α-(1→6)-glucosidic specific enzyme. Comparison of the amino acid sequences of these enzymes suggests that the amino acid next to the catalytic nucleophile is differently conserved depending on the glucosidic linkage specificity. For α-(1→4)-glucoside specific enzymes, Ala or Thr is found, whereas Val is conserved in the α-(1→6)-glucoside specific enzymes. In *Streptococcus mutans* dextran glucosidase, Val195 next to the catalytic nucleophile makes hydrophobic contact with the α-(1→6)-linked substrate. This Val residue was presumably made steric hindrance upon binding to the α-(1→4)-linked substrate. Compared with α-(1→6)-glucoside-specificity, the structural basis for the specificity to α-(1→4)-glucoside is still less understood. Only the crystal structure of the substrate complex of *Halomonas* sp. AGase (HaG) belonging to GH13_23 has been determined.

The substrate chain-length specificity of GH13 AGases is diverse. Most GH13 AGases prefer trisaccharide as substrate. The length of β→α loop 4 of domain A differ depending on the substrate-chain length specificity. In dextran glucosidases, short β→α loop 4 provides the space for long-chain substrates. On the other hand, in disaccharide specific HaG, long β→α loop 4 covers a major part of the active site entrance and obstructs the formation of +2 subsite

to avoid the long-chain substrate binding.

In addition to hydrolytic activity, AGase possesses transglucosylation activity. This activity is utilized in the industrial productions of oligosaccharides and glucosides. For example, isomaltooligosaccharides and panose (α-D-glucopyranosyl-(1→6)-α-D-glucopyranosyl-(1→4)-D-glucose), which have prebiotic property and are utilized as food ingredients, are synthesized from maltooligosaccharides through α-(1→6)-glucosyl transfer catalyzed by the fungal AGase. Finding glycoside hydrolases with high transglycosylation activity to native substrates is important to establish an efficient enzymatic synthesis of glycosides.

In this study, AGase from a soil isolate, *Bacillus* sp. AHU2216 BspAG13_31A, is focused as the model to understand the relationship between the structure and function of GH13_31 AGases. Structural basis for the regioselectivity of α-(1→4)-glucosidic linkage and high transglucosylation activity of BspAG13_31A was analyzed.

Cell-free extract of *Bacillus* sp. AHU2216 isolated from soil near a hot spring in Hiratanai, Hokkaido, Japan showed transglucosylation activity producing maltotriose (G3; maltooligosaccharides of degree of polymerization "n" are shown as Gn) from G2 even at a low G2 concentration (15 mM), while other AGases catalyze predominantly hydrolysis at a similar substrate concentration (20 mM). The partial 16S rDNA sequence (1,276 bp), amplified by PCR, was 99.6% identical to the corresponding region of *Bacillus megaterium* WSH_002. The genome sequence of *B. megaterium* WSH_002 contains the gene BMWSH_3792, which encodes a putative AGase belonging to GH13_31. Although its product is recorded as oligo-1,6-glucosidase in GenBank (GenBank accession number, AEN90673.1), its deduced amino acid sequence possesses an Ala next to the catalytic nucleophile as frequently observed in α-(1→4)-glucoside-acting AGase without carrying three residues essential for α-(1→6)-linkage specificity. Thus, the protein was predicted to be α-(1→4)-specific AGase, and the corresponding gene in *Bacillus* sp. AHU2216 was cloned. The gene product BspAG13_31A was composed of 555 amino acid residues and 98.0% identical with the amino acid sequence of BMWSH_3792.

The recombinant BspAG13_31A was expressed in the *E. coli* BL21 (DE3) transformant in 1 L of LB medium containing 100 μg/mL ampicillin and purified to homogeneity by Ni-affinity column chromatography. 167 mg of purified enzyme with 193 U/mg was obtained. Purified enzyme showed a single band (61.7 kDa) on SDS-PAGE. Recombinant BspAG13_31A showed the highest activity to 4 mM G2 at pH 7.3 and at 35°C. The enzyme retained the activity in a pH ranges of 6.8–8.7 at 37°C for 20 min and 5.9–10.5 at 4°C for 24 h. BspAG13_31A showed more than 95% of residual activity after 20 min incubation below 37°C. Substrates of BspAG13_31A was screened based on degrading activity to various

α-glucosides (1 mM). BspAG13_31A showed high reaction velocity only towards maltooligosaccharides (MOS; 9.21–94.3 s^{-1}) and *p*-nitrophenyl α-glucoside (pNPG; 5.36 s^{-1}). This enzyme showed very low velocity to oligosaccharides linked by other linkages, and it had high specificity to α-(1→4)-glucosidic linkage. The velocity to G2 was the highest among the MOS. *p*-Nitrophenyl α-maltoside was formed in the reaction with pNPG.

Velocities for pNP and D-glucose liberations from pNPG were measured at various pNPG concentrations. The transglucosylation ratio r_{TG} followed a saturation curve, and K_{TG} value, giving 50% or r_{TG}, was determined to be 2.48 mM. In the reaction with G2, reaction equation for the reaction in which both hydrolysis and transglucosylation simultaneously occurred was fitted to velocities of D-glucose liberation. The K_{TG} value was determined to be 21.6 mM.

Initial reaction velocities of BspAG13_31A to 10 mM MOS (G2–G6) were measured, and r_{TG} values were determined by high performance anion exchange chromatography with pulsed amperometric detection. The r_{TG} for G2 was the highest (30%), and decreased to 7.0% for G5 as the chain-length of MOS increased. Because of the low r_{TG} values, v_h for 1–25 mM G3–G7, calculated from v_{glc}, followed the Michaelis-Menten equation. The k_{cat}/K_m values for MOS decreased with an increasing degree of polymerization of the substrates, mainly because of increase in the K_m value.

Crystal structures of BspAG13_31A wild type in apo-form and E256Q in complex with G2, G3, and G4 were determined at 2.5 Å, 1.7 Å, 1.7 Å and 1.6 Å resolutions, respectively. E256Q showed typical behavior of mutant enzymes of general acid/base catalyst: reaction rate of E256Q to 1 mM pNPG with good leaving group was 0.3% of that of wild type, but its activity to G2 with poor leaving group was less than 1.6×10^{-3}% of that of wild type. All complex-structures were highly similar to that of apo-form with root-mean-square deviation (RMSD) of 0.18–0.21 Å for the 555 Cαs. Overall structure of BspAG13_31A consists of the catalytic domain A (residues 1–104, 172–362 and 455–476), a loop-rich domain B (residues 105–171), domain B′ (residues 363–454) and domain C (residues 477–555). Domain A is a $(\beta/\alpha)_8$-barrel, and domain B and B′ were protruding loops connecting third β-strand and third α-helix (β→α loop 3) of the $(\beta/\alpha)_8$-barrel and β→α loop 8, respectively. Domain C is formed by two antiparallel β-sheets. A deep active-site pocket was composed of domains A, B, and B′. The β→α loop 6 (Phe282–Pro295) included a disordered part (Gly291–Lys294 of the apo-form and Asp289–Glu293 of the complexes) because of poor electron density and took different conformations between the apo-form and the complexes. In the apo-form of BspAG13_31A, Trp288 on this loop was placed to prevent substrates from binding at subsites +1 and +2, while in the substrate complexes, this loop showed a different conformation suitable for binding substrates. These results showed flexible structural

movement of β→α loop 6. The conformational change of this loop along with the reaction process may be important for traglucosylation by stabilizing the glucosyl enzyme intermediate and preventing water molecules from binding and acting as substrate.

An electron-density blob, which appeared to be an ion and was coordinated by oxygen atoms of Asp21 OD1, Asn23 OD1, Asp25 OD1, Asp29 OD1, Ile27, and a water molecule, was found in all the structures. Considering the crystallization conditions, the coordination bonds, and refinement results, this blob was assigned as Ca^{2+}. This position is the β→α loop 1 calcium-binding site, which is observed in some GH13 enzymes in addition to other GH13_31 enzymes.

The F_o-F_c map clearly revealed the existence of a substrate molecule, although the electron density for part of the glucose ring at subsite +3 was poor in the complex with G4. The ligands were accommodated in subsite -1 to subsite +3. In all three complexes, the substrates spanned from subsite -1, which was located at the bottom of the active site pocket. In all three complexes, equivalent D-glucosyl residue in each subsite showed similar orientations. The torsion angles, Φ (5-O, 1-C, 4′ -O, and 4′ -C) and Ψ (1-C, 4′ -O, 4′ -C, and 5′ -C), of bound MOS were as follows: (Φ, Ψ) = (40°–43°, −150°–−155°) for residues in subsites -1 and +1, (57°–65°, −164°–−168°) for residues in subsites +1 and +2, and (105°, −113°) for residues in subsites +2 and +3. Only the torsion angles of residues at subsites +2 and +3 were similar to the D-glucosyl residues in α-(1→4)-glucan: (Φ, Ψ) = (102°, −120°). The torsion angles of the residues between -1 and +1 of BspAG13_31A were similar to those of the MOS-complexes of other GH13 enzymes. At subsites +1 and +2, however, the glucosyl residues of the complexes were accommodated to form a hydrogen bond between the two 6-OH groups, while α-(1→4)-glucan forms the hydrogen bonds between 2-OH and 3-OH of the adjacent D-glucosyl residues. The torsion angles of the residues between +1 and +2 subsites of BspAG13_31A were similar to those of G7 bound to amylosucrase, (Φ, Ψ) = (67°, −148°), while those of G4 bound to neopullulanse [(Φ, Ψ) = (92°, −112°)] were similar to the D-glucosyl residues in α-(1→4)-glucan.

The subsite -1 of BspAG13_31A consists of the catalytic nucleophile Asp199, general acid/base catalyst Glu256, transition state stabilizer Asp327, and residues for substrate binding, Asp60, Tyr63, His103, Arg197, His326, and Arg411. These residues in subsite -1 were well-conserved in GH13 AGases and other glucoside-acting enzymes. At subsite +1, the side-chains of His203, Gln256 (Glu256 in wild type), and Asn258 were in suitable locations for hydrogen bonds with 2-O, 3-O, and 2-O of the D-glucosyl residue, respectively. These are thought to be important for specificity to α-(1→4)-linkage. The D-glucosyl residue in subsite +2 has only two water-mediated hydrogen bonds except possible hydrogen bond between 4-O

and Asn258 side chain. At subsite +3, the sugar residue is fixed by three direct (Phe282, Met285, and Gln328) and one water-mediated hydrogen bonds. The torsion angles of glucose residues in subsites +1 and +2 were less stable than the global minimum, because of steric hindrance by Asn258 in subsite +2. Asn258 is predicted to be the reason for high disaccharide specificity of this enzyme.