Structural and functional studies of novel α-L-glucosidase from *Cecembia lonarensis*

(*Cecembia lonarensis*由来新規α-L-グルコシダーゼの構造と機能に関する研究)

Hokkaido University Graduate School of Agriculture
Frontiers in Biosciences　Doctor Course

Kang, Hye-Jin

Contents

List of Abbreviations

BiAfcB	1,3-1,4-α-L-fucosidase from *Bifidobacterium longum* subsp. *infantis*
BpGH29	α-L-galactosidase from *Bacteroides plebeius* DSM 17135
BSA	bovine serum albumin
CBB	coomassie brilliant blue
ClAgl29A	α-L-glucosidase from *Cecembia lonarensis* LW9 (EKB48091.1)
ClAgl29B	α-L-glucosidase from *Cecembia lonarensis* LW9 (EKB48090.1)
ESI	electrospray ionization
GH	glycoside hydrolase
β-L-Glc-F	β-L-glucopyranosyl fluoride
β-L-Fuc-N_3	α-L-fucosyl-(1→2)-β-L-fucosyl azide
HLH	helix-loop-helix
HMBC	heteronuclear multiple-bond correlation
HPAEC-PAD	high-performance anion exchange chromatography with pulsed amperometric detection
HPLC	high-performance liquid chromatography
HSQC	heteronuclear single-quantum coherence
HSQC-TOCSY	HSQC-total correlation spectroscopy
IPTG	isopropyl 1-thio-β-D-galactopyranoside
LB	Lysogeny broth
LBK	LB medium with 50 μg/mL kanamycin
Le^a	$Lewis^a$
Le^x	$Lewis^x$
MS	mass spectrometry
MW	molecular weight
NMR	nuclear magnetic resonance
PCR	polymerase chain reaction
PEG	polyethylene glycol
*p*NP	*p*-nitrophenol
*p*NP α-L-Ara	*p*-nitrophenyl α-L-arabinopyranoside
*p*NP β-L-Ara	*p*-nitrophenyl β-L-arabinopyranoside
*p*NP α-D-Gal	*p*-nitrophenyl α-D-galactopyranoside
*p*NP α-L-Gal	*p*-nitrophenyl α-L-galactopyranoside
*p*NP β-D-Gal	*p*-nitrophenyl β-D-galactopyranoside

*p*NP α-D-GalNAc	*p*-nitrophenyl *N*-acetyl α-D-galactosaminide
*p*NP α-D-Glc	*p*-nitrophenyl α-D-glucopyranoside
*p*NP α-L-Glc	*p*-nitrophenyl α-L-glucopyranoside
*p*NP β-D-Glc	*p*-nitrophenyl β-L-glucopyranoside
*p*NP β-D-GlcNAc	*p*-nitrophenyl *N*-acetyl β-D-glucosaminide
*p*NP α-L-Fuc	*p*-nitrophenyl α-L-fucopyranoside
*p*NP β-D-Lac	*p*-nitrophenyl β-D-galactosyl-(1→4)-β-D-glucopyranoside
*p*NP α-D-Xyl	*p*-nitrophenyl α-D-xylopyranoside
*p*NP β-D-Xyl	*p*-nitrophenyl β-D-xylopyranoside
*p*NP 2-O-α-L-Glc α-L-Glc	*p*-nitrophenyl α-L-glucopyranosyl-(1→2)-α-L-glucopyranoside
*p*NP 3-O-α-L-Glc α-L-Glc	*p*-nitrophenyl α-L-glucopyranosyl-(1→3)-α-L-glucopyranoside
*p*NP 6-O-α-L-Glc α-L-Glc	*p*-nitrophenyl α-L-glucopyranosyl-(1→6)-α-L-glucopyranoside
RMSD	root mean square deviation
SDS	sodium dodecyl sulfate
SDS-PAGE	SDS-polyacrylamide gel electrophoresis
TLC	thin-layer chromatography
TmaFuc	α-L-fucosidase from *Thermotoga maritima*
Tris	tris(hydroxymethyl)aminomethane

Abstract

Natural L-glucose has not been found, and nature is believed to have a homochirality of D-glucose. However, an L-glucose-utilizing bacterium in soil has been reported, and its unique catabolic pathway of L-glucose has been elucidated. Moreover, our group's discovery of two α-L-glucosidases, ClAgl29A and ClAgl29B, in the *Cecembia lonarensis* genome has opened the possibility of α-L-glucoside in nature. This dissertation presents an X-ray crystallographic analysis of ClAgl29A and ClAgl29B, shedding light on their unique substrate specificity and the structural differences from α-L-fucosidases, which are their homologs and abundant in nature. It also explores the specificities of ClAgl29A and ClAgl29B for substrates other than α-L-glucoside. The glycosidases catalyze the formation of glycosidic linkages, often referred to as transglycosylation, which is used for the industrial production of oligosaccharides. For efficient α-L-glucoside synthesis using α-L-glucosidases, several mutant enzymes are prepared and evaluated.

1. Structural Characterizations of α-L-Glucosidases

The crystal structures of ClAgl29A and ClAgl29B were determined. Their structure comprised catalytic TIM $(\beta/\alpha)_8$ barrel and β-sandwich domains. The ClAgl29B complex obtained with β-L-glucose suggested that α-L-glucosidase had gained the α-L-glucoside recognition mechanism during the process of molecular evolution. A significant finding was a substantial difference in the recognition sites of the equatorial hydroxy group at carbon 4 of L-glucose (O4) compared to the evolutionarily related α-L-fucosidases. While α-L-fucosidases recognized axial O4 of L-fucose with two His residues located at the C-terminus of β-strand 1 and β-strand 2 of the TIM barrel, respectively, α-L-glucosidase recognized O4 of L-glucose with different orientation from L-fucose with an Asp located on the long loop connecting β-strand 1 and α-helix 1 of the TIM barrel. Recombinant proteins of five amino acid sequences with α-L-glucosidase-like sequences, identified by the structural analysis, were prepared. However, several of these showed only weak α-L-glucosidase activity.

2. Substrate Specificity of α-L-Glucosidases

This section evaluated the substrate specificity of the two α-L-glucosidases for *p*-nitrophenyl α-L-rhamnoside, α-L-6-deoxy-1-glucoside, and α-L-xyloside. Both enzymes hydrolyzed α-L-6-deoxy-1-glucoside a little more efficiently than α-L-glucoside. They had little activity toward α-L-rhamnoside, widely found in nature. Together with the previous results of weak substrate specificity of ClAgl29A and ClAgl29B for α-L-fucoside, these enzymes were found not to act on glycosides that are abundant in nature but on glycosides whose existence is undetermined.

3. α-L-Glucoside Synthesis Using α-L-Glucosidases

To investigate the cause of the low transglucosylation activity of ClAgl29B, ClAgl29B having a point mutation of S132D, Q275L, and S132D/Q275L, respectively, which mimic ClAgl29A, were produced. The mutant enzymes exhibited greater transglucosylation activity than the wild type,

indicating the involvement of Ser132 and Gln275 in directing the reaction toward transglucosylation or hydrolysis. The mutant ClAgl29A and ClAgl29B having a mutation at the catalytic amino acid residue catalyzed α-L-glucosylation of D-xylosides and D-glucosides, using β-L-glucosyl fluoride as a substrate. The reactions formed α-L-Glc-(1→4)-D-Glc and α-L-Glc-(1→4)-D-Xyl.

Chapter 1. Introduction

1. L-Carbohydrates

Carbohydrates are essential in all living organisms, mediating diverse biological functions. They are originally highly hydrated carbon-based molecules with a chemical formula of $C_n(H_2O)_m$. Today, carbohydrates include polyalcohol aldehydes, ketones, and acids, as well as polyalcohols themselves and their derivatives. Monosaccharides, the building blocks of oligosaccharides and polysaccharides and not further simplified into simple molecules by hydrolysis, generally have homochirality. Glucose, fructose, mannose, and galactose, which are abundant in nature, are believed to have D-enantiomeric homochirality. Several monosaccharides, including L-fucose, L-rhamnose, and L-arabinose, have L-enantiomeric homochirality. An L-enantiomeric monosaccharide is generally rare, but it has been found in important constituents of clinically useful agents such as antibiotics.

Some studies have shown that monosaccharides without homochirality exist in nature. Natural arabinose has L-enantiomers for the most part, while D-arabinose is found in polysaccharides of Mycobacteria (Brennan, P. J., 2003) and in glycosides of the genus Aloe (Kuzuya, H., 2001). In addition, bacterial endo-D-arabinase, degrading the unique glycan structures of *Mycobacterium tuberculosis* strain H37Rv (Kotani, S., 1971), and similar activity in *Cellulomonas* (Mcneil, M. R., 1994) and *Mycobacterium smegmatis* (Xin, Y., 1999) have been identified. The findings suggest the potential existence of both enantiomeric sugars and their associated degrading enzymes in nature besides the case of arabinose. The study of a bacterium, *Paracoccus laeviglucosivorans*, in soil that uses L-glucose suggests the possibility that both enantiomers of glucose exist (Shimizu, T., 2012). The authors also argued that the discovery of a bacterium utilizing L-glucose is not necessarily related to the presence of L-glucose in nature, as this bacterium may use a scyllo-inositol metabolizing enzyme to break down L-glucose.

2. Catalytic mechanisms of glycosidases

Glycosidases, also known as glycoside hydrolases, are enzymes that catalyze the hydrolysis of a glycosidic bond. Depending on the stereochemistry of the product with its substrates, glycosidases have been classified into either retaining or inverting glycosidase (Sinnott, M. L., 1990).

In the case of retaining glycosidases, which retain the anomeric configuration during the catalysis, side chains of two amino acid residues are involved in the hydrolysis of glycosidic bonds, one acts as a general acid/base catalyst and another nucleophile catalyst. The catalytic reaction pathway can be divided into glycosylation and deglycosylation steps. In the glycosylation step, the general acid/base catalyst transfers the proton to the scissile *O*-glycosidic bond. Simultaneously, the nucleophile catalyst attacks an anomeric carbon with an electron-deficient state and forms a glycosyl-enzyme intermediate. In the deglycosylation step, the intermediate is hydrolyzed by a catalytic water

molecule activated by the general acid/base catalyst (Davies, G. J., 1995) (Fig. 1-1a). Since the water attack occurs from the same side as the anomeric configuration of the substrate, the anomeric configuration of the product is retained with respect to the substrate. The catalytic mechanism of the retaining glycosidases is called the double displacement mechanism because it involves two displacement reactions in catalysis. As a side reaction, retaining glycosidases can also catalyze the synthesis of glycosides, called transglycosylation, if a hydroxy group of alcohol degrades the glycosyl-enzyme intermediate instead of a water molecule.

In contrast, inverting glycosidases operate by a single displacement reaction. The general acid and the general base catalysts are poised in the catalytic center, where the acid catalyst donates a proton to the glycoside bond to be cleaved, while the water molecules activated by the base catalyst attack the anomeric carbon. Since a water molecule attacks from the opposite side of an anomeric configuration of the substrate, the anomeric form of the product is inverted relative to the substrate (Fig. 1-1b) (McCarter, J. D., 1994).

3. Glycoside hydrolase family 29

Glycoside hydrolases (GHs) are a large and diverse group of enzymes involved in the hydrolysis of glycosidic bonds in biological systems. Some of them are also involved in the synthesis of polysaccharides and oligosaccharides. GHs are classified into families based on amino acid sequence similarity, and each is assigned a family number (Cantarel, B. L., 2009; Drula, E., 2022). The classification is compiled in the Carbohydrate-Active enZymes (CAZy) database (http://www.cazy.org/). The CAZy database organizes GHs into more than 180 families. Since the classification is based on sequence similarity and prevalent functions in organisms, the CAZy database has become a primary means for annotating uncharacterized nucleic acids and proteins. GH family 29 (GH29) is the largest α-L-fucosidase family with over 5000 entries. GH29 can be further divided into the subfamilies GH29-A and GH29-B by sequence similarity and substrate specificity (Sakurama, H., 2012a). The GH29-A subfamily contains enzymes with relatively broader substrate specificities (α-1,2/ α-1,3/ α-1,4/ α-1,6-L-fucosyl linkages) and is highly effective against chromogenic substrates such as *p*-nitrophenyl α-L-fucopyranoside (*p*NP α-L-Fuc), whereas GH29-B is strict for α-1,3/4-L-fucosyl linkages with little or no activity on *p*NP α-L-Fuc (Ashida, H., 2009). The GH29-A subfamily includes α-L-fucosidase from *Thermotoga maritima* (TmaFuc), one of the well-studied α-L-fucosidase (Sulzenbacher, G., 2004). This subfamily includes an α-L-galactosidase from *Bacteroides plebeius* DSM 17135, BpGH29 (Robb, C. S., 2022). Among the GH29-B subfamily, 1,3-1,4-α-L-fucosidase from *Bifidobacterium longum* subsp. *infantis* (BiAfcB), which is associated with the hydrolysis of α-L-fucosyl residue in Lewis^x (Le^x) and Lewis^a (Le^a), has been studied well for its structure-function relationship (Sela D.A., 2012; Sakurama H., 2012b). Sakurama *et al*.

discussed the difference in substrate specificity between GH29-A and GH29-B based on the structural comparison between BiAfcB and TmaFuc (Sakurama H., 2012a). BiAfcB has a binding site for the β-D-galactosyl residue of Le^x and Le^a, but TmaFuc does not. Since the binding of the β-D-Gal of Le^x or Le^a induces the conformational change to the active form, GH29-B enzymes have little or no activity on *p*NP α-L-Fuc.

Recently, our research group found the novel α-L-glucosidases in GH29 based on the diversity of amino acid residues involved in substrate recognition in α-L-fucosidases (Shishiuchi, R., 2022). The difference between α-L-glucoside and α-L-fucoside is the orientation of the hydroxyl group at C4 (OH-4) and the presence of a hydroxyl group attached to the carbon at position 6 (OH-6). Of them, the diversification of OH-4 recognition mechanism within GH29 resulted in the discovery of α-L-glucosidase in this family. The catalytic domain of GH29 enzymes displays a TIM $(\beta/\alpha)_8$-barrel fold. In TmaFuc, two imidazole side chains of His34 and His128 at the C-terminal end of β-strands 1 and 2 of the catalytic domain, respectively, stabilize the axial OH-4 of α-L-fucoside through hydrogen bonds. BiAfcB also has the corresponding His36 and His85. While the conservation of the His128 (TmaFuc) position is relatively high, variations (Ala, Arg, Asn, Asp, Cys, Gln, Glu, Ile, Lys, Phe, Ser, Thr, Tyr, and Val) are observed at the His34 (TmaFuc) position for Alpha-L-fucosidase (PF01120) in the Pfam database (https://pfam.xfam.org/family/PF01120). Among them, our group focused on two amino acid sequences (EKB48090.1 and EKB48091.1) of *Cecembia lonarensis* LW9. These sequences possess Asp114 (EKB48090.1) and Asp103 (EKB48091.1) instead of His. Since EKB48090.1 and EKB48091.1, prepared as recombinant enzymes in *Escherichia coli*, showed substrate specificity to *p*-nitrophenyl α-L-glucopyranoside (*p*NP α-L-Glc) rather than *p*NP α-L-Fuc, they were designated ClAgl29B (EKB48090.1) and ClAgl29A (EKB48091.1), respectively, as novel α-L-glucosidases.

4. Oligosaccharides

Carbohydrates consisting of 2 to 10 monosaccharides linked via glycosidic linkages are called oligosaccharides. In some cases, carbohydrates with more than 10 monosaccharides, seen in complex carbohydrates, are also called oligosaccharides, and the boundary between oligosaccharides and polysaccharides is not well defined. Carbohydrates consisting of 10 to 100 monosaccharides are often referred to as megalosaccharides. Most naturally occurring oligosaccharides are found in the plant kingdom. In the animal kingdom, lactose and its derived oligosaccharides in milk, and trehalose, which is commonly found in insects, are only a few examples. Oligosaccharides such as fructooligosaccharides, galactooligosaccharides, and xylooligosaccharides have the function of regulating the intestinal environment by stimulating the growth and activity of beneficial bacteria such as lactic acid bacteria and bifidobacteria in the gut, contributing to immune system regulation,

etc. (Kong, C., 2021). Recent studies further suggest that human milk oligosaccharides, blood-type oligosaccharides, and heparan sulfated oligosaccharides may prevent gastrointestinal symptom upset, one of the sequelae of COVID-19, by directly inhibiting the entry of residual SARS-CoV-2 (Chutipongtanate, S., 2022). In addition, a healthy microbiome brought about by consuming these oligosaccharides has been suggested to prevent the severity of COVID-19. These beneficial oligosaccharides are composed of D-enantiomeric monosaccharides. It remains unclear whether oligosaccharides of L-enantiomeric monosaccharides have similar or superior functions to D-oligosaccharides. On the other hand, interesting functions are beginning to be discovered for monosaccharides with L-enantiomer. For example, L-sorbose shows multiple anti-tumor activities (Xu, H. L., 2023), and L-glucose fluorescent probes are used for cancer diagnosis (Anastasiou, I. A., 2021). In addition, L-glucose has been perceived as sweet, similar to D-glucose (Dubovski, N., 2022). Oligosaccharides composed of L-sugars may have hidden functions that turn conventional knowledge upside down, like the mirror-image relationship between D-enantiomers and L-enantiomers.

5. Glycosynthase

Glycosynthase is a mutant retaining glycosidase, of which the catalytic nucleophile residue (generally Glu or Asp) is replaced by a smaller amino acid without charge, such as alanine, glycine, or serine. The mutant enzyme has no or strongly reduced hydrolytic activity (Ducros, V., 2003) but catalyzes the glycosyl transfer reaction to an appropriate acceptor molecule when using an activated donor sugar, such as glycosyl fluoride or glycosyl azide, with an opposite anomer configuration to a proper substrate. This reaction mimics the deglycosylation step of a retaining glycosidase (Fig. 1-1c), and the substrate, which has an inverted anomeric configuration relative to the substrate, mimics the glycosyl-enzyme intermediate in the reaction of the wild-type enzyme. An acceptor molecule bound to the pseudo-intermediate with the catalytic nucleophile mutant facilitates the leaving of fluorides and azides (Malet, C., 1998). The glycosynthase methodology has successfully converted a number of retaining glycosidases into glycosynthases. Glycosynthases originated from GH29-A α-fucosidases from *Sulfolobus solfataricus* and *T. maritima* (Cobucci-Ponzano, B., 2009) using β-L-fucosyl azide as a donor substrate have been reported. BiAfcB, belonging to GH29-B, has been converted to glycosynthase. The glycosynthase catalyzed the transfer of α-L-fucosyl residue from β-L-fucosyl fluoride to β-D-galactosyl-(1→3)-*N*-acetyl-D-glucosamine and β-D-galactosyl-(1→4)-*N*-acetyl-D-glucosamine to produce Le^a and Le^x, respectively (Sakurama, H., 2012b).

6. Purpose of the doctoral thesis research

As already mentioned above, our research group found novel α-L-glucosidases, ClAgl29A and

ClAgl29B, in GH29, a family of α-L-fucosidase. However, how these enzymes recognize α-L-glucoside remains ambiguous. My first goal is to elucidate the α-L-glucoside recognition mechanism of these enzymes through their X-ray crystallographic analysis. The α-L-glucosidases were found among the putative α-L-fucosidases belonging to GH29 with a variation of His involved in recognizing the axial OH-4 of α-L-fucoside. ClAgl29A and ClAgl29B have Asp corresponding to His34 of TmaFuc, but there are several sequences with other amino acid substitutions. I also aim to clarify the substrate specificity of these sequences (Chapter 2).

In Chapter 3, the substrate specificities of ClAgl29A and ClAgl29B toward *p*NP α-L-rhamnopyranoside (*p*NP α-L-Rha), *p*NP α-L-quinovopyranoside (*p*NP α-L-Qui), and *p*NP α-L-xylopyranoside (*p*NP α-L-Xyl) were evaluated. The presence of α-L-glucoside is still unknown in nature. On the other hand, naturally occurring α-L-rhamnoside, which are structurally similar to α-L-glucoside, are abundant. It has been suggested that α-L-quinovoside (6-deoxy α-L-glucoside) may be a cell wall component in some bacteria. α-L-Xyloside has not been reported in nature, but the pyranose form of this pentose is structurally similar to α-L-glucoside, lacking the CH_2OH at position 5. The purpose of this chapter is to evaluate the specificity of α-L-glucosidases to these α-L-glycosides to gain insight into their original substrates.

In Chapter 4, I aim to identify the amino acid residues involved in the differences in the efficiency of transglycosylations catalyzed by ClAgl29A and ClAgl29B. Both enzymes have α-L-transglucosylation activity in addition to α-L-glucoside hydrolysis. Both enzymes have similar active site pockets, but ClAgl29A has higher α-L-transglucosylation activity than ClAgl29B. The purpose of this study was to elucidate the structural factors causing this difference.

As mentioned above, ClAgl29A and ClAgl29B synthesize α-L-glucoside through the transglycosylation. In this reaction, as the reaction proceeds, the concentration of α-L-glucoside synthesized decreases due to their inherent α-L-glucoside hydrolysis activity. This decrease can be resolved using glycosynthases that have lost their hydrolytic activity. In Chapter 5, I converted ClAgl29A and ClAgl29B to glycosynthases to produce efficient α-L-glucoside synthases.

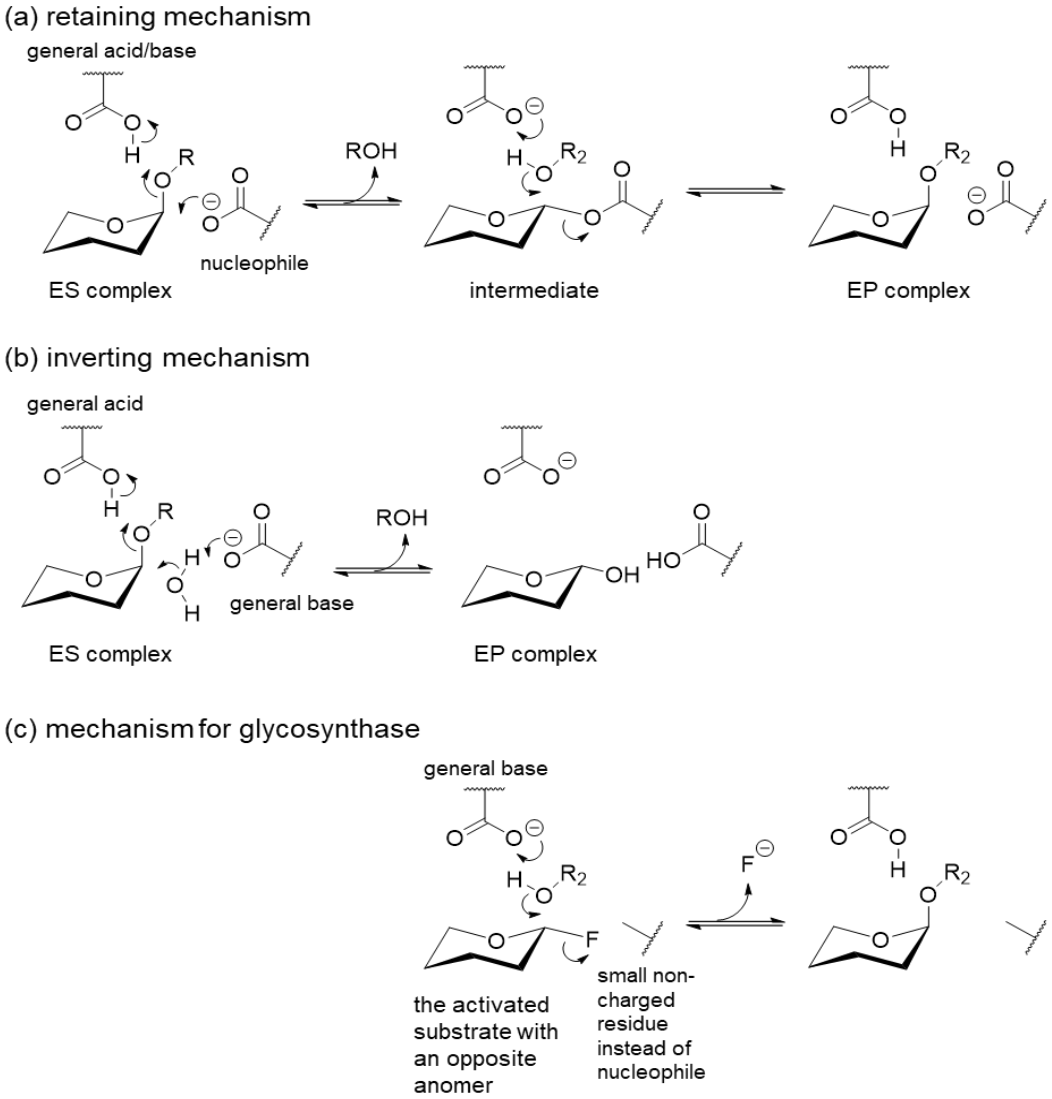


Figure 1-1. Glycosidase mechanisms. (a) The mechanism for retaining glycosidases (double displacement mechanism). When R_2 is H, the reaction is hydrolysis, and when R_2 is not H, the reaction is transglycosylation. (b) The mechanism for inverting glycosidases (single displacement mechanism) (c) The mechanism for glycosynthase.

Chapter 2. Crystal structures of ClAgl29A and ClAgl29B and characterization of GH29 enzymes from *Coraliomargarita akajimensis* JCM 23193

1. Introduction

Glucose is a common monosaccharide in nature and a major energy source for living organisms and is believed to have a homochirality of D-enantiomer. Meanwhile, the discovery of L-glucose-utilizing bacteria in *Paracoccus* species 43P (Shimizu, T., 2012) and the elucidation of their metabolic pathways suggested the presence of natural L-glucose. Since most carbohydrates exist as oligosaccharides, polysaccharides, and glycosides rather than monosaccharides, it is conceivable that L-glucose also exists as L-glucosides and that enzymes exist to hydrolyze them.

Organisms have diverse GHs to hydrolyze a variety of glycosides. The diverse GHs can be classified into 189 GH families based on their sequence similarities (Drula, E., 2022). Because classification is based on sequence similarity, the function is conserved within a family. On the other hand, substrate specificity diverges within a family in many cases. Previous work by our group, based on the diversification of amino acid residues involved in substrate recognition, led to the discovery of novel α-L-glucosidases with different substrate specificities within GH29, a family of α-L-fucosidase (Shishiuchi, R., 2022). Since the major difference between α-L-glucoside and α-L-fucoside is OH-4 orientation (Fig. 2-1A), two His residues involved in OH-4 recognition were of interest: His34 and His128 in TmaFuc and His36 and His85 in BiAfcB, which are located at the β-strands 1 and 2 of the catalytic TIM $(\beta/\alpha)_8$ barrel fold domain, respectively (Fig. 2-1B). Sequence comparisons revealed that His at β-strand 1 is relatively less conserved than that at β-strand 2 (Fig. 2-1C). Among proteins having amino acid residues other than His at β-strand 1, two proteins, ClAgl29A and ClAgl29B, with His substituted for Asp, were found as α-L-glucosidase (Shishiuchi, R., 2022).

ClAgl29A and ClAgl29B apparently showed higher substrate specificity for *p*NP α-L-Glc than for *p*NP α-L-Fuc, with k_{cat}/K_m values for *p*NP α-L-Glc hydrolysis being 4.2 (ClAgl29A) and 12 (ClAgl29B) times higher than for *p*NPα-L-Fuc hydrolysis (Shishiuchi, R., 2022). In this chapter, the three-dimensional structures of ClAgl29A and ClAgl29B were determined by X-ray crystallography to elucidate the structural factors involved in the substrate specificity of α-L-glucosidase, which is distinctly different from α-L-fucosidase, also classified as GH29.

Coralliomargarita akajimensis, isolated from seawater at Majanohama, Akajima, Japan, is an obligately aerobic Gram-negative bacterium and belongs to the phylum “Verrucomicrobia”. Its genome encodes eleven GH29 enzymes. Among them, five sequences, Caka0074 (Phe39), Caka0506 (Cys49), Caka0522 (Val56), Caka0523 (Val76), and Caka0536 (Cys91), have a non-His residue (amino acid residue in parentheses) at the equivalent to the His residue at the end of the β-strand 1, which corresponds to His34 and His36 of TmaFuc and BiAfcB, respectively (Fig. 2-2). Since α-L-

glucosidases have Asp at this position, the substrate specificity of the enzymes substituted with other residues are interesting. This chapter also evaluated the substrate specificity of these five GH29 enzymes, Caka0074, Caka0506, Caka0522, Caka0523, and Caka0536.

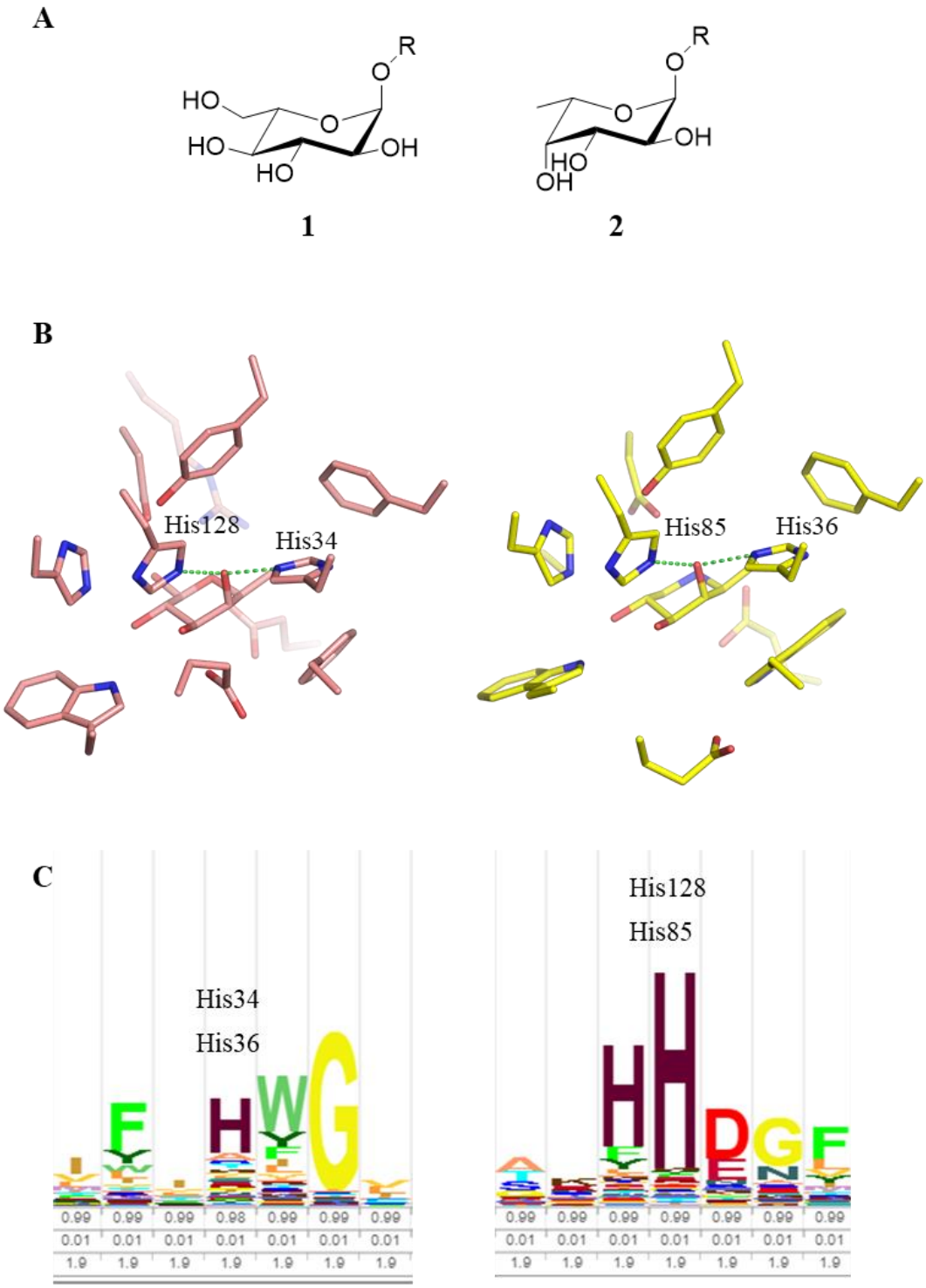


Fig. 2-1. Recognition of OH-4 of α-L-fucoside by α-L-fucosidases.

(A) Structures of α-L-glucoside (1) and α-L-fucoside (2). (B) α-L-Fucosyl residue binding site of TmaFuc (left; PDB ID: 2zwy) and BiAfcB (right; PDB ID: 3ues). Green dashed lines indicate hydrogen bonds. (C) Conservation of His residues associated with the axial O4-recognition.

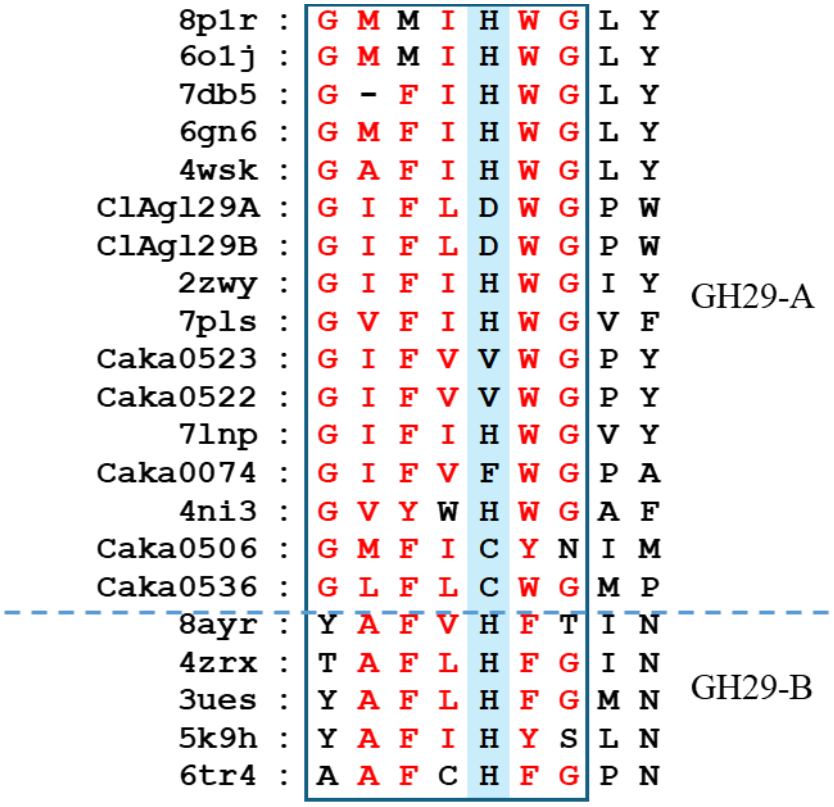


Fig. 2-2. Multiple amino acid sequence alignment around β-strand 1 of GH29 enzymes. Enzymes are indicated by the designated names in this study or PDB IDs. The residue represented by blue corresponds to His34 (TmaFuc; PDB ID: 2zwy) and His36 (BiAfcB; PDB ID: 3ues). Sequences were aligned by MAFFT version 7 online version (https://mafft.cbrc.jp/alignment/server/).

2. Materials and methods

2.1. Materials

C. akajimensis JCM 23193 was provided by RIKEN BioResource Center (Tokyo, Japan) through the National BioResource Project of MEXT, Japan. L-Glucose was purchased from TCI (Tokyo, Japan). pNP α-L-Fuc was obtained from Sigma-Aldrich (St. Louis, MA, USA). All other chemicals and solvents used were of analytical grade available commercially.

2.2. X-ray crystal structure analysis of ClAgl29A and ClAgl29B

2.2.1. Crystallization

The His-tagged ClAgl29A and ClAgl29B were expressed in *E. coli* Rosetta (DE3), and purified using Ni-affinity, anion exchange and gel filtration column chromatographies according to the report (Shishiuchi, R., 2022). The purified enzymes were dialyzed against 10 mM sodium phosphate buffer (NaPB) (pH 7.5) and concentrated using Amicon Ultra-15 and Ultra-0.5 centrifugal filter units (30,000 NMWL) (Merck, Darmstadt, Germany).

All crystallizations were performed with the hanging drop vapor diffusion method at 20°C for 2~4 weeks. The 48-well IWAKI micro plates (AGC Techno Glass; Shizuoka, Japan) containing 200 μL of reservoir solution, the 12 mm siliconized glass cover slides (Hampton Research; Journey Aliso Viejo, CA, USA), and the high vacuum sealing compound HIVAC-G (Shin-etsu Silicone; Tokyo, Japan)) were used for the crystallization. The crystals of ClAgl29A were obtained in a drop composed of 2 μL of ClAgl29A (A_{280}=20, 145 μM), 2 μL of reservoir solution [0.1 M citrate-NaOH buffer (pH 3.5) and 16% (w/v) polyethylene glycol (PEG) 3,350] and 1 μL of 1 mM L-fucose. The crystals of ClAgl29B were obtained in a drop composed of 2 μL of ClAgl29B (A_{280}=10, 74.8 μM), 2 μL of reservoir solution [5% (v/v) 2-propanol, 0.1 M citrate-NaOH buffer (pH 3.5) and 5% (w/v) PEG 20,000] and with or without 1 μL of 1 mM L-glucose.

2.2.2. Data collection

The composition of cryoprotectant for the ClAgl29A crystals was 0.1 M citrate-NaOH buffer (pH 3.5), 18% (w/v) PEG 3,350, 10% (v/v) glycerol and 1 mM L-fucose, and that for the ClAgl29B crystals was 5% 2-propanol, 0.1 M citrate-NaOH buffer (pH 3.5), 6% PEG 20,000, 20% glycerol and 1 or 30 mM L-glucose. All crystals were soaked into the cryoprotectant for 10 seconds and flash-frozen in liquid nitrogen. The X-ray diffraction data were collected at the X06SA beamline of the Swiss Light Source, Paul Scherrer Institute (Villigen, Switzerland) for ClAgl29A and at the AR-NE3A beamline of the Photon Factory, KEK (Tsukuba, Japan) for ClAgl29B using an EIGER 16M detector (Dectris; Daïtwil, Switzerland) and Pilatus 2M-F pixel detector (Dectris), respectively, at a wavelength of 1.00000 Å. The datasets were collected from a single crystal under

a stream of nitrogen at a temperature of 100 K. The diffraction dataset was indexed, integrated, and scaled using the XDS program (Kabsch, W., 2010).

2.2.3. Structure determination and refinement

The structure of ClAgl29A was determined by the molecular replacement method with AutoMR Wizard in the PHENIX program package (Liebschner, D., 2019; McCoy, A. J., 2007) using the protein coordinates of TmaFuc (PDB ID: 2xsd) as a search model. The refinement process was carried out using the programs Refmac5 (Murshudov, G. N., 2011) and phenix.refine (Afonine, P. V., 2012) inconjunction with interactive fitting and rebuilding based on 2m*F*o−D*F*c and m*F*o−D*F*c electron densities using COOT (https://www.ccp4.ac.uk/) (Emsley, P., 2004; Emsley, P., 2010). The final structure was defined as the convergence of the R_{work} and R_{free} values. The data processing and refinement statistics are presented in Table 2-1. The graphical representations were prepared using the pymol Molecular Graphics Systems Version 2.0 (Schrödinger, LLC, New York, NY, USA).

2.3. Construction of expression plasmid vectors for Caka_0074, _0506, _0522, _0523, and _0536

The DNA regions encoding the presumptive mature proteins of Caka0074, 0506, 0522, 0523, and 0536 were amplified by polymerase chain reaction (PCR) using genomic DNA of *C. akajimensis* JCM 23193 as a template. Their signal sequences were predicted by SignalP 5.0 (https://services.healthtech.dtu.dk/services/SignalP-5.0/) (Almagro Armenteros, J. J., 2019). The primers used for the PCR are listed in Table 2-2. PCR was performed using PrimeSTAR Max DNA polymerase (Takara Bio, Shiga, Japan), and the amplification reaction was carried out with the following amplification conditions: 94°C for 2 min and 25 cycles of 98°C for 10 sec; 55°C for 15 sec; 72°C for 30 sec. The amplification of DNA fragments was confirmed by agarose gel electrophoresis. The plasmid DNA of pET28a was amplified by PCR using KOD One DNA polymerase (Toyobo, Osaka, Japan) and a pair of primers (pet28-HindIIIs and pet28-NdeIa) (Table 2-2). The amplification cycle was as follows: 94°C for 2 min and 30 cycles of 98°C for 10 sec; 65°C for 15 sec; 68°C for 30 sec. The reaction mixture was treated with DpnI, and agarose gel electrophoresis was performed to confirm DNA amplification. The amplified pET28a DNA (100 ng) and each insert DNA fragment (Caka_0074, 140 ng; Caka_0506, 200 ng; Caka_0522, 200 ng; Caka_0523, 180 ng; Caka_0536, 125 ng) were mixed and added 1 μL of 5 x TEDA [0.5 M tris(hydroxymethyl)aminomethane (Tris)-HCl (pH 7.5), 50 mM MgCl2, 50 mM DTT, 0.25 g/mL PEG8000, and 0.01 U/μL T5 exonuclease (New England Bio Lab, Ipswich, MA, USA)] (Xia, Y., 2019) and kept at 30°C for 40 min. The reaction mixture was used for transformation of *E. coli* DH5α and the expression plasmid DNAs, pET28a-Caka_0074, pET28a-Caka_0506, pET28a-Caka_0522,

pET28a-Caka_0523, and pET28a-Caka_0536, were prepared.

2.4. DNA sequencing analysis

A cycle sequencing reaction was performed using BigDyeTM Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA, USA). The reaction solution (10 μL), including an appropriate oligonucleotide primer (0.3 μM), template DNA (ca. 40 ng/μL), and BigDye™ Terminator v3.1 Ready Reaction Mix diluted one-eighth using BigDyeTM Terminator v1.1 & v3.1 5× Sequencing Buffer was held at 96°C for 2 min, and then the temperature step (96°C for 10 sec, 50°C for 5 sec, and 60°C for 4 min) was repeated 25 times. The products were purified by ethanol precipitation after adding one-tenth volume of 125 mM EDTA (pH 8.0) and 3 M sodium acetate (pH 5.2), respectively. The precipitates were dissolved in 15 μL of HiDi formamide (Applied Biosystems, Foster City, CA, USA), heated at 96°C for 4 min, and quickly chilled on ice. ABI PRISM 3100 Genetic analyzer (Applied Biosystems) was used to analyze DNA sequence.

2.5. Production and purification of Caka0074, 0506, 0522, 0523, and 0536

E. coli BL21 (DE3) (Merck) was transformed by pET28a-Caka_0506 and Caka_0523 and Rosetta-gami 2 (DE3) (Merck) was done by pET28a-Caka_0074, Caka_0522, and Caka_0536. BL21 (DE3) transformants were grown in Lysogeny broth (LB) medium [10 g/L tryptone, 5 g/L yeast extract, and 5 g/L NaCl (Nacalai Tesque, Kyoto, Japan)] with 50 μg/mL kanamycin (LBK), and Rosetta-gami 2 (DE3) transformants were in LBK supplemented with 30 μg/mL chloramphenicol. Each transformant was grown in LB medium with appropriate antibiotics at 37°C until OD_{600} reached about 0.5, and cultured at 18°C (Rosetta-gami 2 (DE3) transformants) and 30°C (BL21 (DE3) transformants) overnight without any inducer for Rosetta-gami 2 (DE3) transformants and with 0.1 mM isopropyl 1-thio-β-D-galactopyranoside (IPTG) to express the recombinant proteins. The cells were harvested by centrifuging at 5,800 × g at 4°C for 15 min. The collected cells were suspended in the extraction buffer (20 mM sodium phosphate buffer (pH 7.5) and 500 mM NaCl) and were disrupted using SONIFIER 250 ultrasonication (Branson, Danbury, CT, USA) by repeating sonication in the condition (output 2, duty cycle 40, 5 sec) 10 times. The lysed cells were centrifuged at 13,000 × g at 4°C for 10 min, and the supernatants were collected as crude extract.

The recombinant proteins were purified by Amicon® Pro Purification System (NMWL, 10 kDa; Merck). The crude extract (500 μL) was applied onto the column, and the column was washed with 20 mM sodium phosphate buffer (pH 7.5) and 300 mM NaCl containing 20 mM imidazole (pH 7.5). Proteins were eluted by raising the imidazole concentration (20 to 250 mM). The purity of proteins was examined by sodium dodecyl sulfate- polyacrylamide gel electrophoresis (SDS-PAGE). The eluted fractions were dialyzed against 20 mM sodium phosphate buffer (pH 7.5) and concentrated

with a 10 kDa cutoff Amicon Ultra-15 centrifugal filter. The concentration of *C. akajimensis*-derived proteins was computed from the theoretical extinction coefficient at 280 nm (ProtoParam tool, https://web.expasy.org/protparam/) derived from the amino acid sequence.

2.6. SDS-PAGE

SDS-PAGE was performed as described (Laemmli, U. K., 1970) with Mini-Protean III (Bio-rad, Richmond, CA, USA). Protein samples were mixed with the same volume of sample buffer consisting of 4% (w/v) SDS, 0.02% (w/v) bromophenol blue, 20% (v/v) glycerol, 10% (v/v) 2-mercaptoethanol, and 0.125 M Tris-HCl buffer (pH 6.8) and incubated at 100°C for 3 min. Polyacrylamide gels with an acrylamide concentration of 12%T [12% (w/v) acrylamide, 0.32% (w/v) *N*, *N*′-methylenebisacrylamide, 0.38 M Tris-HCl (pH 8.8), 0.05% (w/v) ammonium peroxodisulfate, 0.05% (v/v) *N*,*N*,*N*′,*N*′-tetramethylethylenediamine] were prepared as separation gel and 4%T polyacrylamide gels [4% (w/v) acrylamide, 0.1% (w/v) *N*, *N*′-methylenebisacrylamide, 0.125 M Tris-HCl (pH 6.8), 0.05% (w/v) ammonium peroxodisulfate, 0.05% (v/v) *N*,*N*,*N*′,*N*′-tetramethylethylenediamine] as stacking gel. Gel thickness was set at 0.75 mm. The samples and size marker (Protein Molecular Weight Marker (Broad) (Takara Bio)) were electrophoresed at 12 mA in stacking gel and 25 mA in separating gel. To prevent heat generation in the gels, the power was set to 3 W or less. The proteins in the gel were stained using a staining solution consisting of 60 mg/L Coomassie Brilliant Blue G-250 and 35 mM HCl (Lawrence, A. M., 2009).

2.7. Measurement of reaction velocity

Reaction rates on 2 mM *p*NP α-L-Glc, *p*NP α-L-Fuc, and *p*NP α-L-galactopyranoside (*p*NP α-L-Gal) were measured in 40 mM sodium citrate buffer (pH 5.0) at 35°C for 10 min. The reaction was stopped by adding two volumes of 1 M sodium carbonate. The amount of *p*NP released was measured at the absorption of 400 nm using a 1 cm cuvette with a molar extinction coefficient of 5,560 M^{-1} cm^{-1}.

Table 2-1. Summary of Data Collection and Refinement Statistics[a]

	ClAgl29A	ClAgl29B	ClAgl29B/L-Glc
PDB ID	7XSF	7XSG	7XSH
Data collection			
Beamline	SLS X06SA	PF AR-NE3A	PF AR-NE3A
Wavelength	1.00000	1.00000	1.00000
Space group	$P12_1$ 1	$P2_12_12_1$	$P2_12_12_1$
Unit cell			
a, b, c (Å)	74.58 95.52 83.55 β = 97.33°	73.13 123.41 166.60	72.39 121.58 166.37
Resolution range (Å)	47.81-2.01 (2.13-2.01)	45.42-1.61 (1.71-1.61)	46.60-1.71 (1.81-1.71)
Total no. of reflections	539,809 (86,368)	1,292,170 (201,560)	1,058,673 (157,759)
No. of unique reflections	76,908 (12,196)	194,275 (30,357)	159,024 (25,291)
Multiplicity	7.0 (7.1)	6.7 (6.6)	6.7 (6.2)
Completeness (%)	98.9 (97.3)	99.5 (97.0)	99.8 (99.1)
mean *I/σ(I)*	14.80 (2.45)	22.17 (4.51)	17.44 (2.37)
R_{mean}	9.0 (96.2)	6.9 (45.9)	10.1 (94.2)
R_{sym}	8.3 (89.1)	6.3 (42.3)	9.3 (86.3)
Wilson B-factor	35.12	14.47	19.14
CC1/2	0.999 (0.90)	0.999 (0.942)	0.999 (0.810)
Refinement			
R_{work}	20.5	17.2	17.9
R_{free}	24.3	18.8	20.2
Number of atoms			
Macromolecules	8634	8985	8938
Ligands	15	40	76
Solvent	203	1008	
Non-hydrogen atoms	9146	10069	9932
RMSD values from ideal			
Bonds	0.007	0.009	0.009
Angles	1.25	1.43	1.39
Ramachandran (%)			
Favored	96.56	97.23	96.88
Allowed	3.44	2.77	3.12
Outliers	0.00	0.00	0.00
Rotamer outliers	0.76	0.82	1.13
Clashscore	2.19	1.80	1.42

[a]Values in parentheses are for the highest-resolution shell. RMSD, and root-mean-square deviation.

Table 2-2. Oligonucleotide sequence used in this study

Name	Sequence (5′→3′)	Purpose
Caka_0074_F_pET28a	CCTGGTGCCGCGCGGCAGCCATATGACTGATATCAATGCTGCAGA	Cloning for Caka_0074
Caka_0074_R_pET28a	GGTGCTCGAGTGCGGCCGCAAGCTTTCAGCTAAAACCAAGTCTGAAGAC	
Caka_0506_F_pET28a	CCTGGTGCCGCGCGGCAGCCATATGATTCAATCAGTGAAAGAACAGCAGC	Cloning for Caka_0506
Caka_0506_R_pET28a	GGTGCTCGAGTGCGGCCGCAAGCTTTCATTTGTGTGTGTCCTTTCTAGG	
Caka_0522_F_pET28a	CCTGGTGCCGCGCGGCAGCCATATGGATTCGCATCCTTCCATGGC	Cloning for Caka_0522
Caka_0522_R_pET28a	GGTGCTCGAGTGCGGCCGCAAGCTTCTACTTGACCGGAATAAGGCG	
Caka_0523_F_pET28a	CCTGGTGCCGCGCGGCAGCCATATGGACAGTGAGGCACCTACGAT	Cloning for Caka_0523
Caka_0523_R_pET28a	GGTGCTCGAGTGCGGCCGCAAGCTTCTACTGTATTGGTTCAAGTCGAATCC	
Caka_0536_F_pET28a	CCTGGTGCCGCGCGGCAGCCATATGCAGAACCACATGCCGGGG	Cloning for Caka_0536
Caka_0536_R_pET28a	GGTGCTCGAGTGCGGCCGCAAGCTTTTACTTCCAGTGGGTAAATAGGTACG	
pET28-HindIIIs	AAGCTTGCGGCCGCACTCGAGCACC	Amplification of pET28a DNA
pET28-NdeIa	CATATGGCTGCCGCGCGGCACCAGG	Amplification of pET28a DNA
Caka_0522_F_seq	GCATGGCCTGAAGATGGGAA	Sequencing for Caka_0522
Caka_0522_R_seq	TGCATCGACCAAACATCAAAGAC	
Caka_0523_F_seq	ACTGAGTGGTTTCATCCCAAGT	Sequencing for Caka_0523
Caka_0523_R_seq	AGAGAGAGTGAATCCTTCCCCT	
Caka_0536_F_seq	TGGCTATGCTATCTGGGATTCG	Sequencing for Caka_0536
Caka_0536_R_seq	CATTCACCTCCCACCACCAA	

The underlined sequences are homologous to those of the plasmid DNA.

3. Results

3.1 Crystal structures of ClAgl29A and ClAgl29B

3.1.1. Overall structure of ClAgl29A and ClAgl29B

The crystal structures of ClAgl29A and ClAgl29B were determined at 2.0 Å (ClAgl29A) and 1.6 Å (ClAgl29B) resolutions, respectively. All residues were built based on the electron density except for residues 358–375 in ClAgl29A and residues 21–31 and 377–383 in ClAgl29B. Each protein consisted of two monomers in an asymmetric unit. As is the case with a majority of GH29 proteins, the structure of each protein comprised two domains: a catalytic TIM $(\beta/\alpha)_8$ barrel domain (residues 24–480, ClAgl29A; 32–485, ClAgl29B) and a β-sandwich domain (residues 481–573, ClAgl29A; 486–587, ClAgl29B) (Fig. 2-3). ClAgl29A and ClAgl29B structures were closely similar, with the root mean square deviation (RMSD), calculated between Cα-atoms of matched residues at best 3D superposition, 1.11 Å. PDBeFold (https://www.ebi.ac.uk/msd-srv/ssm/) detected TmaFuc (PDB ID: 1HL9) (Sulzenbacher, G., 2004), α-L-fucosidase of *Paenibacillus thiaminolyticus* (6GN6) (Kovalová, T., 2019) and α-L-fucosidase of *Bacteroides thetaiotaomicron* (BT2970, 6HZY) (Coyle, T., 2019) as similar 3D structures of ClAgl29B with the top high Q-scores, 0.43, 0.38, 0.37, respectively. ClAgl29A and ClAgl29B had no obvious hydrogen bonds between β-strand 5 and β-strand 6 of the inner β-barrel similar to the other structure-known GH29 enzymes. GH29 enzymes do not have the obvious α-helix 5, and ClAgl29B had a helix-loop-helix (HLH) structure (residues 362–396) instead of α5 (Fig. 2-3). This HLH structure existed independently of the barrel and appeared not to act to protect the inner β-barrel from the solvent, as did the α-helix of the common TIM $(\beta/\alpha)_8$ barrel. The corresponding part of ClAgl29A could not be determined due to the unclear electron density. As observed in TmaFuc and BT2970, ClAgl29A and ClAgl29B had no distinguishable α-helix 6, exposing a part of the inner β-barrel to the solvent. Comparison with other enzymes revealed that ClAgl29A and ClAgl29B have specific structures at the β→α loop 3 and N-terminus (Fig. 2-3). The β→α loop 3, longer than that of other GH29 enzymes, consisted of a small antiparallel β-sheet and a 3_{10} helix (232–289 for ClAgl29A and 244–301 for ClAgl29B). At the N-terminus, the structure containing a large α-helix (46–74 for ClAgl29A and 57–84 for ClAgl29B) coiled around the TIM $(\beta/\alpha)_8$ barrel. A sodium ion bound to ClAgl29A and ClAgl29B similarly. ClAgl29A held sodium ion via O of Asp85, O of His 87, and Oδ1 of Asp386, and ClAgl29B held via O of Lys95, O of His97, and Oδ1 of Asp397.

3.1.2. Active-site pocket structure of ClAgl29B

ClAgl29A and ClAgl29B were co-crystallized with L-fucose and L-glucose, respectively, and the crystal was also soaked into 1 mM L-fucose and L-glucose, respectively, before flash-frozen, but electron densities from ligands were not observed (Fig. 2-4 A and B). On the other hand, an

electron density of the monosaccharide at the active pocket was found in the crystal structure derived from the ClAgl29B crystal soaked into 30 mM L-glucose, and it fitted well to β-L-glucose (Fig. 2-4 C). No structural rearrangements were observed between enzymes with and without ligand binding. Structural comparison with other GH29 enzymes (Cobucci-Ponzano, B., 2003, Sulzenbacher, G., 2004, Lammerts van Bueren, A., 2010) indicated that Asp327 (ClAgl29B) and Asp315 (ClAgl29A) were the residues providing the catalytic nucleophile. The Oδ1 atom of Asp327 was at 2.8 Å to the O1 of β-L-glucose and 3.0 Å to the Nη1 atom of Arg361. The carboxy groups of Glu403 (ClAgl29B) and Glu391 (ClAgl29A) at pseudo β-strand 6 were located 5.5 Å from the respective catalytic nucleophiles at the end of β-strand 4 and should act as a general acid/base catalyst. ClAgl29B tightly entrapped β-L-glucose, where each hydroxy group was in direct contact with one or more amino acid side chains (Fig. 2-5), indicating that the enzyme explicitly recognizes L-glucose. C6 of L-glucose was surrounded by aromatic side chains derived from Phe112, Phe195, Tyr241, Trp325, and Phe425. O6 could form a hydrogen bond with Sγ of Cys418 at the β-strand 6 of the $(\beta/\alpha)_8$ barrel. O2 was hydrogen-bonded to Nε1 of Trp140 and Nε2 of His199. O3 was stabilized through hydrogen bonds with the Oδ1 of Asp139 and Nε2 of His198. Asp139 is also involved in the recognition of the equatorial O4 of L-glucose. The Oδ2 atom of Asp139 from O4 of L-glucose was 2.5 Å. Tyr137 and Trp330 were located within 4 Å of the ligand and were part of the active site pocket formation. Asp114, which led to the discovery of α-L-glucosidase, was not in direct contact with the substrate but its Oδ2 was located within hydrogen bonding distance of the Oδ2 of Asp139 that recognized O4 of L-glucose. The amino acid residues of the active site pocket of ClAgl29A are similar to those of ClAgl29B; therefore, the ligand-bound state of ClAgl29A and ClAgl29B is considered the same.

3.1.3 Structural comparison of the active-site structures

The structure of the active-site pocket of ClAgl29B was compared to that of β-L-fucose-bound TmaFuc (PDB ID: 1ODU), which had the best overall structural similarity (Fig. 2-6). These active-site pocket structures were similar, except for the recognition site for O4. As mentioned above, Asp139 is responsible for stabilizing equatorial O4, together with O3 of β-L-glucose in ClAgl29B. The corresponding residue of Glu66 in TmaFuc was hydrogen-bonded to O3 of β-L-fucose. In TmaFuc, His34 was responsible for stabilizing axial O4 of β-L-fucose, whereas the corresponding Asp114 in ClAgl29B was far from equatorial O4 of L-glucose, where its Oδ2 was positioned 3.8 Å from the equatorial O4. Oδ2 of Asp114 was hydrogen-bonded to Oδ2 of Asp139, which recognizes O4 of L-glucose. The role of Asp114 appears to affect the orientation of the side chain of Asp139 to be able to interact with O4, together with the side chain of Phe195 (Fig. 2-7). TmaFuc had Thr125 corresponding to Phe195, and its side chain was sufficiently small not to interfere with

His34 to get closer to α-L-fucoside. Minor variations in the spatial positions of Phe290 in TmaFuc and Phe425 in ClAgl29B were likely to affect substrate specificity (Fig. 2-8). In GH29, the residue at this position was conserved as Phe or Trp and was responsible for excluding equatorial O4 from the active-site pocket (Sulzenbacher, G., 2004). ClAgl29B had Phe425 at this site, but its side chain had different spatial arrangements and did not hinder the binding of O4 to L-glucose. The spatial variation of these aromatic residues can be attributed to a difference in these main chain conformations: the φ and ψ angles are −92.0 and −48.5 of Phe290 in TmaFuc and −92.9 and 107.1 of Phe425 in ClAgl29B. Their disparate ψ angles seemed to be attributed to a hydrogen bond between the O atom of the aromatic residues. In TmaFuc, the O atom of Phe290 was close to the Nδ2 atom of Asn324 at β7. Alternatively, the O atom of Phe425 in ClAgl29B hydrogen bonds with the N atom of Asp139 in the long loop following β-strand 1.

3.2 Functional analysis of proteins with variations in the His34 site of TmaFuc

Five GH29 proteins, Caka0074 (Phe39), Caka0506 (Cys49), Caka0522 (Val56), Caka0523 (Val76), and Caka0536 (Cys91), which have a non-His residue (amino acid residue in parentheses) at the equivalent to His34 of TmaFuc, from *C. akajimensis* JCM 23193 were produced as the N-terminal His_6-tagged recombinant proteins. In the SignalP analysis, Caka0074 has no signal sequence, Caka0506 up to Ser19, Caka0522 up to Gly46, Caka0523 up to Gly26, and Caka0536 up to Gly15 were predicted as the secretory signal sequences and cleaved. First, *E. coli* BL21 (DE3) was used as the host, but recombinant Caka0074, Caka0522, and Caka0536 were produced in the insoluble fraction. These recombinant proteins were then produced in soluble form using Rosetta-gami 2 (DE3). The recombinant proteins were purified using Ni-affinity column chromatography. Each purified enzyme was detected near the estimated size on SDS-PAGE, all of which were in the position of molecular weight smaller than the theoretical molecular weight (Caka0074, 55,835; Caka0506, 54,176; Caka0522, 72,576; Caka0523, 69,526; and Caka0536, 57,622) (Fig. 2-9).

Their hydrolytic activities toward 2 mM *p*NP α-L-Glc, *p*NP-α-L-Fuc, and *p*NP-α-L-Gal were determined. Five enzymes showed low activity toward these substrates (Table 2-3). Caka0506 preferred *p*NP α-L-Fuc, Caka0522 preferred *p*NP α-L-Glc and *p*NP α-L-Gal, while Caka0536 accepted all substrates (*p*NP α-L-Glc, *p*NP α-L-Fuc, and *p*NP α-L-Gal).

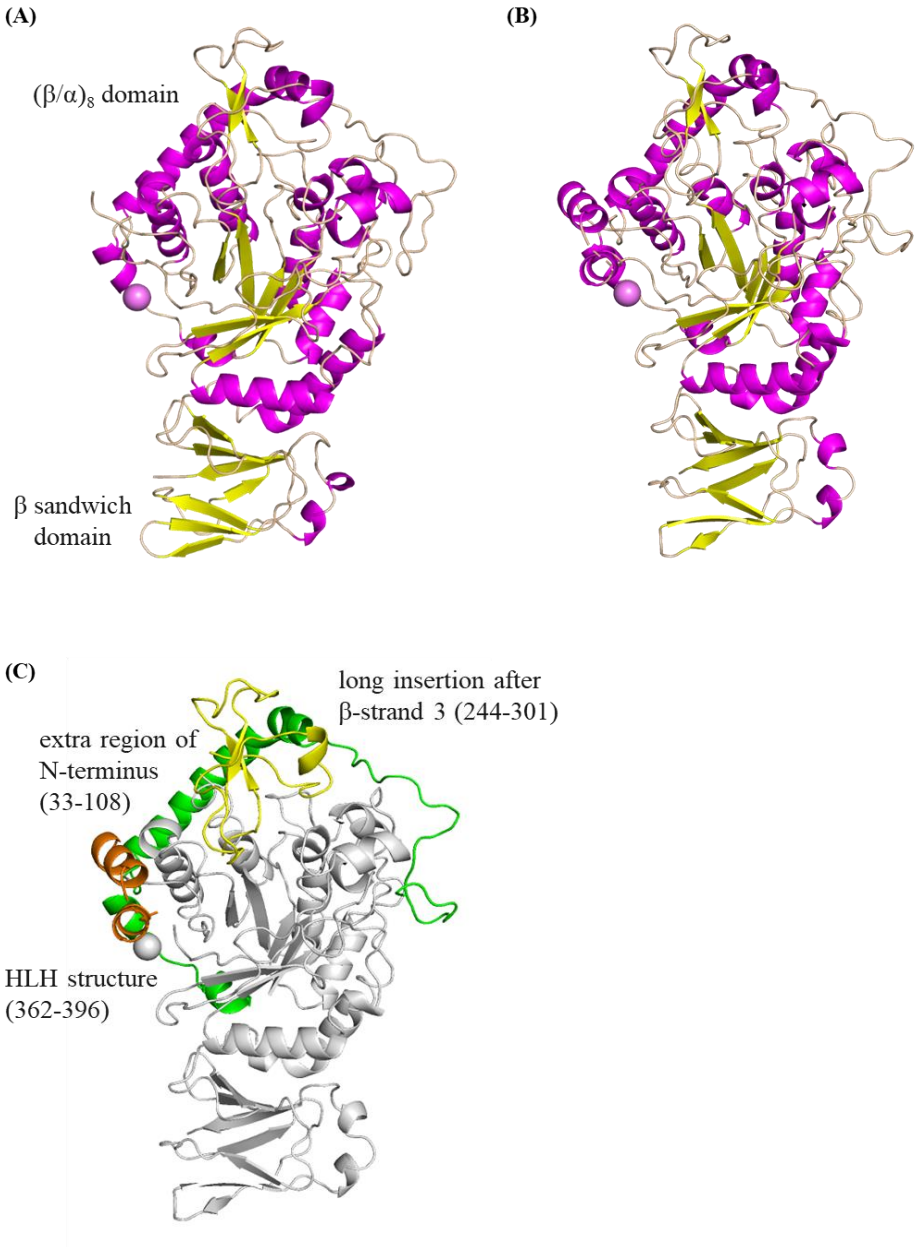


Fig. 2-3. Three-dimensional structures of ClAgl29A and ClAgl29B.

Cartoon representations of ClAgl29A (A) and ClAgl29B (B). The sphere models show the bound sodium ions. The characteristic structural elements found in ClAgl29B are indicated in green, the additional region of the N-terminus (33-108); yellow, the long insertion after β-strand 3 (244-301); orange, the helix-loop-helix structure (362-396) (C).

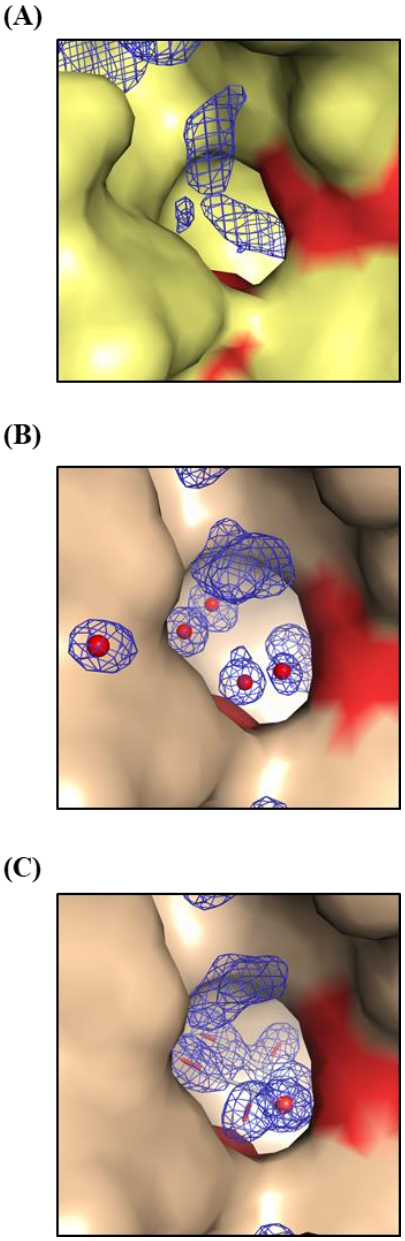


Fig. 2-4. Electron density maps observed at the active pocket of ClAgl29A and ClAgl29B.

The 2m*F*o-D*F*c map (σ = 1.5) at the active pocket observed in the crystal structures of ClAgl29A (A), ClAgl29B (B), and ClAgl29B bound with β-L-glucose (C) are shown as blue mesh. The catalytic nucleophile and acid/base residues are shown in red. Water molecules and β-L-glucose are shown as red spheres and the white and red line model, respectively.

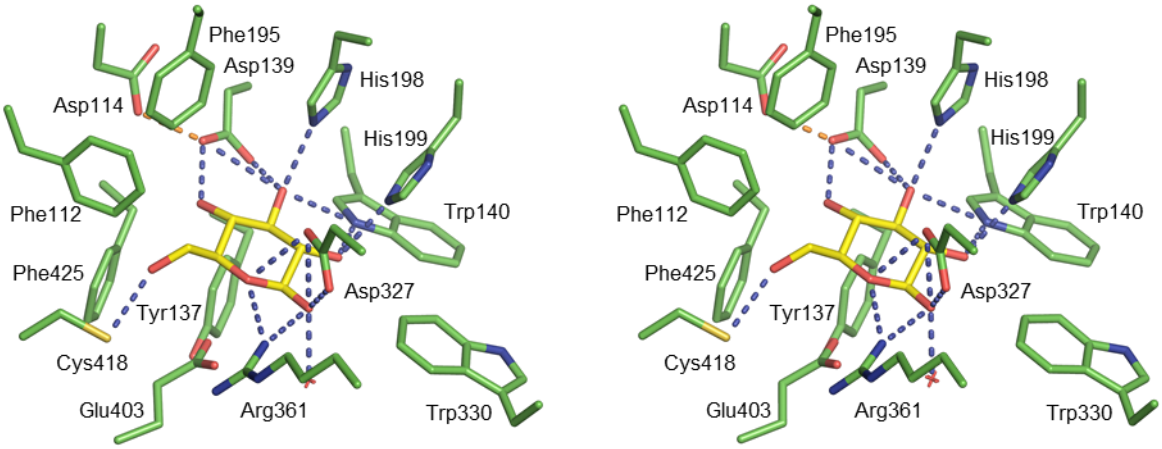


Fig. 2-5. Stereo view of the β-L-glucose - ClAgl29B interactions.

Hydrogen bonds are shown as blue dotted lines. The interaction between Asp139 and Asp114 is shown as orange dotted line.

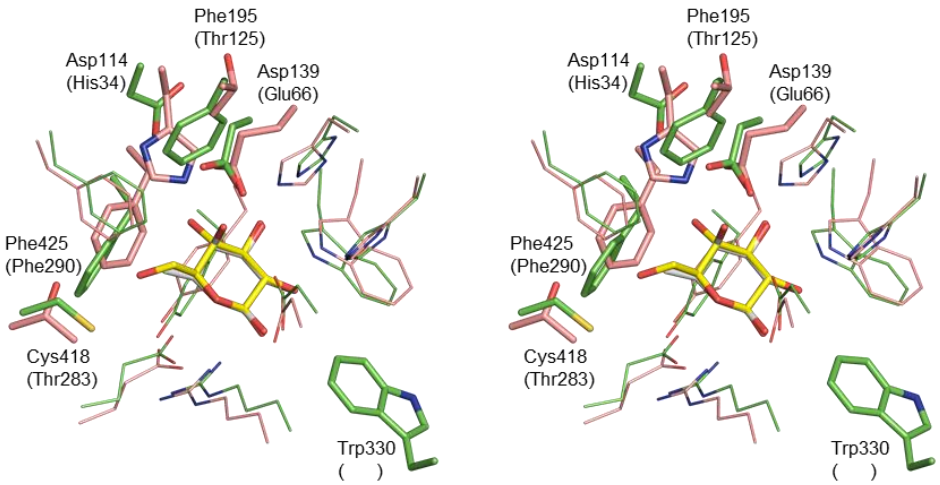


Fig. 2-6. Comparison of the active-site pockets of ClAgl29B and TmaFuc.

ClAgl29B is shown in green and TmaFuc in wheat. β-L-Glucose and β-L-fucose are shown in yellow and gray, respectively. The different sites in the two enzymes are indicated by thicker sticks and residue numbers are assigned. Residue numbers in parentheses are those of TmaFuc.

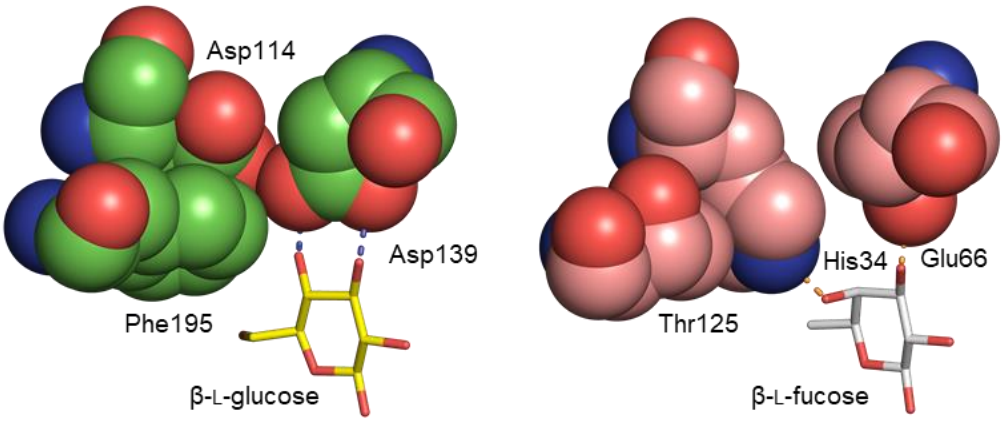


Fig. 2-7. Comparison of the O4 recognition sites of ClAgl29B and TmaFuc.

ClAgl29B (left) is shown in green and TmaFuc (right) in light orange.

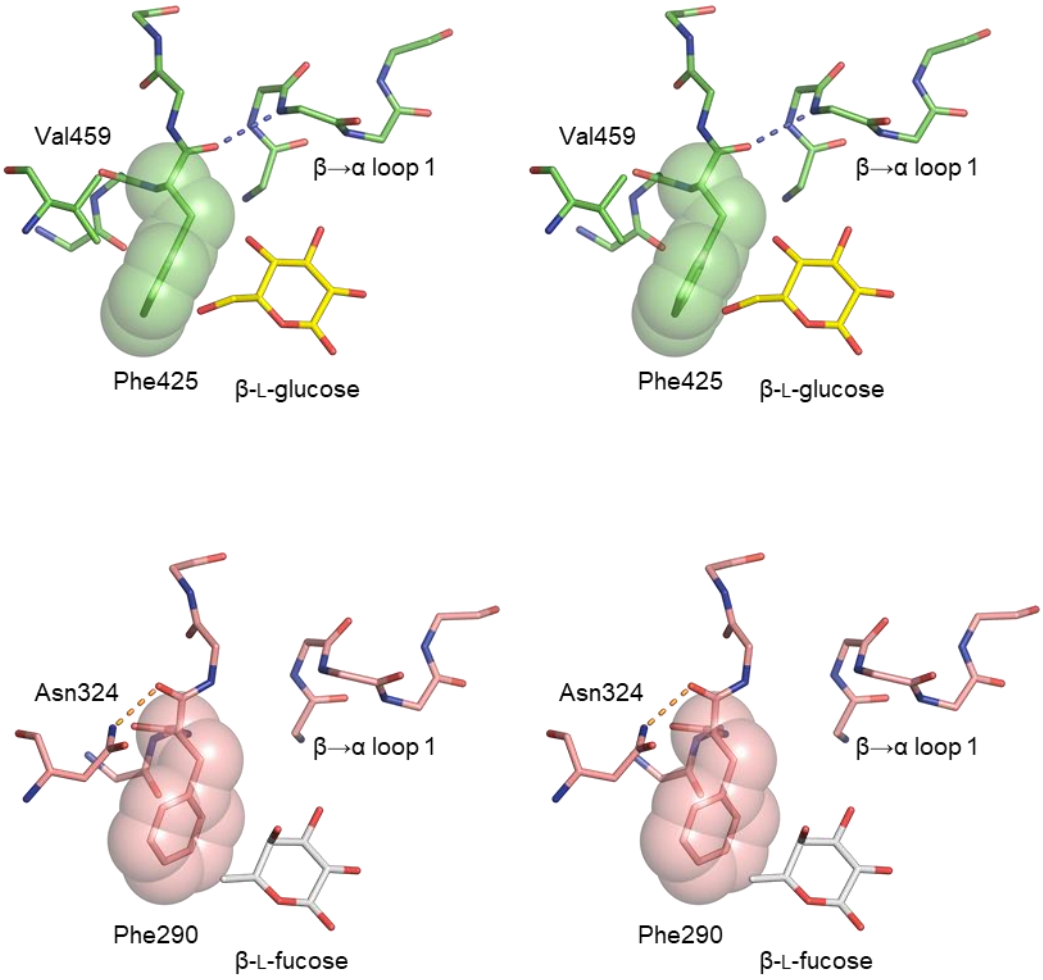


Fig. 2-8. Differences in the orientation of Phe side chains located near O4 in ClAgl29B and TmaFuc.

The stereo diagrams above and below are the structures of ClAgl29B and TmaFuc, respectively. D otted lines indicate hydrogen bonds.

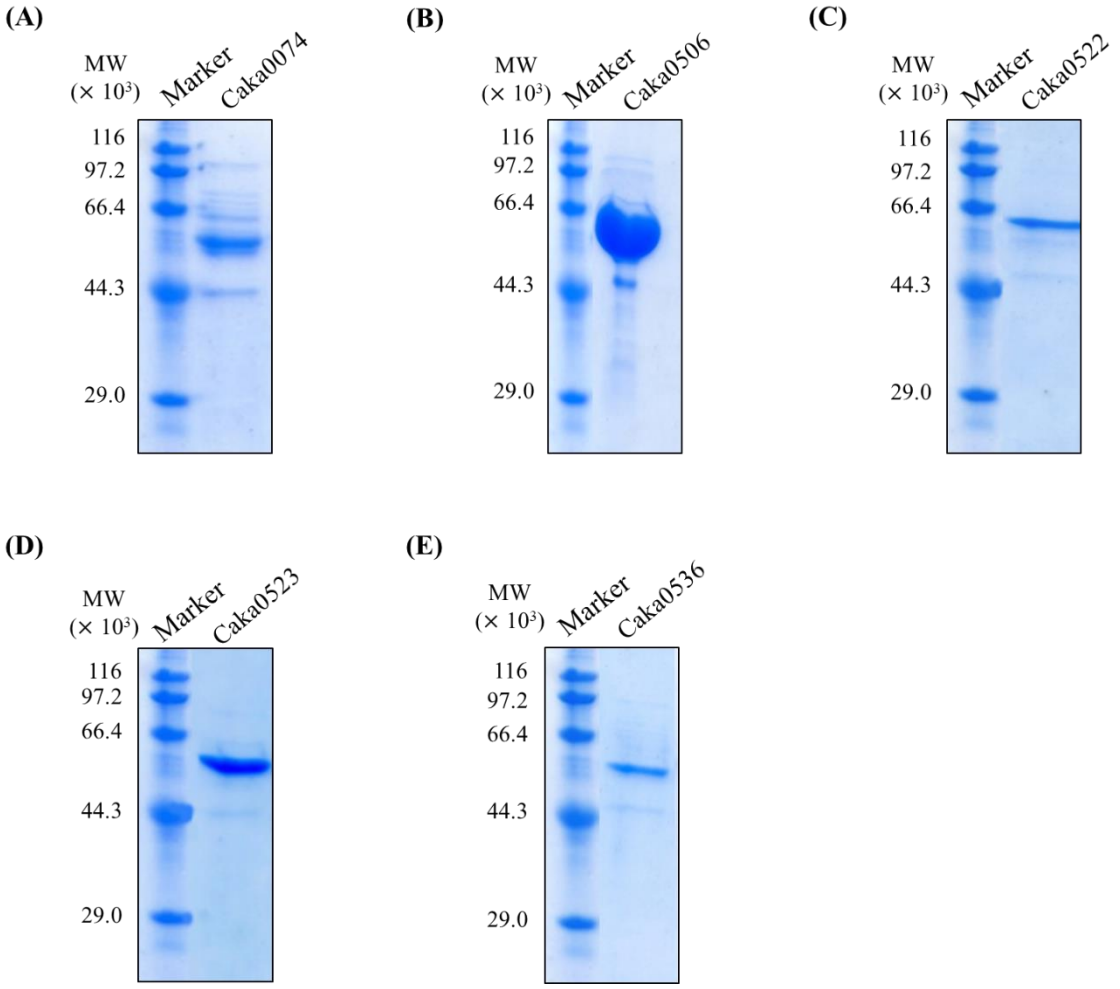


Fig. 2-9. SDS-PAGE of purified recombinant Caka0074, Caka0506, Caka0522, Caka0523, and Caka0536.

Caka0074 (A_{280}=0.363, 4 μL) (A), Caka0506 (A_{280}=1.48, 4 μL) (B), Caka0522 (A_{280}=0.439, 4 μL) (C), Caka0523 (A_{280}=1.05, 4 μL) (D), and Caka0536 (A_{280}=0.484, 4 μL) (E) were analyzed by SDS-PAGE. Lanes Marker are protein size marker. The molecular masses of the standard are shown on the left.

Table 2-3. The reaction velocity of GH29 enzymes from *C. akajimensis* JCM 23193

	velocity (nmol/min/mg)				
Substrate	**Caka0074**	**Caka0506**	**Caka0522**	**Caka0523**	**Caka0536**
*p*NP α-L-Fuc	1.4 ± 0.3	12 ± 0.4	2.1 ± 0.6	1.3 ± 0.8	3.1 ± 1.5
*p*NP α-L-Glc	1.3 ± 0.6	7.0 ± 0.2	8.2 ± 0.1	< 0.12	6.0 ± 0.3
*p*NP α-L-Gal	1.4 ± 0.3	5.5 ± 1.9	7.5 ± 0.3	< 0.12	4.5 ± 0.8

4. Discussion

Our research group has identified an α-L-glucosidases, ClAgl29A and ClAgl29B in GH29, a family of α-L-fucosidases, based on the substitutions of Asp103 (ClAgl29A) and Asp114 (ClAgl29B) for one of the two conserved His residues that recognize the axial O4 of α-L-fucoside. In this chapter, X-ray crystallographic analysis was performed to clarify the substrate recognition mechanism of these α-L-glucosidases. The obtained structure of ClAgl29B complex with β-L-glucose showed that the O4 recognition site was distinct from that of the known GH29 α-L-fucosidases (Fig. 2-6). The equatorial O4 of β-L-glucose was recognized by Asp139 in the long loop structure following β-strand 1 (Fig. 2-5). GH29-A enzymes conserve the Glu residue, which recognizes O3 of α-L-fucosyl residue, corresponding to Asp139 (Fig. 2-10). The carboxy side chain of Asp139 was also within the hydrogen bonding distance of the O3 of β-L-glucose (Figs. 2-5 and 2-7), but its orientation differed from that of the Glu residue. The carboxy side chain of Asp114, instead of the conserved His, and the aromatic ring side chain of Phe195 at β-strand 2 are responsible for this orientation (Fig. 2-7). ClAgl29A possessed this equatorial O4 recognition machinery, but the other GH29 members did not (Fig. 2-10). The GH29 enzyme has a Phe or Trp residue in the loop following β-strand 7 to prevent the binding of equatorial O4 (Fig. 2-10; Sulzenbacher, G., 2004), and ClAgl29B had an equivalent Phe425, but the orientation of the side chain was different, allowing the binding of equatorial O4 of β-L-glucose. The difference in orientation seems to be attributed to which atoms the O atoms form hydrogen bonds with. The O atom of Phe425 of ClAgl29B formed a hydrogen bond with the N atom of Asp139 in the long β→α loop 1. The O atoms of Phe413 of ClAgl29A likewise formed a hydrogen bond with the N atom of the corresponding Asp103. On the other hand, the equivalent O atoms of α-L-fucosidases formed a hydrogen bond with the side chain of the conserved Asn in β-strand 8 (Fig. 2-10). The corresponding residues were Val447 and Val459 in the ClAgl29A and ClAgl29B, respectively, and their side chains had no atoms to form a hydrogen bond. Compared to the divergence of the OH-4 recognition mechanisms of the substrates shown so far, no significant difference was found in the recognition of O6. The environment around C6 was hydrophobic for both α-L-glucosidase and α-L-fucosidase, and the O6 of β-L-glucose seemed able to form hydrogen bond with the Sγ of Cys418. In GH29, this position is less conserved, and Cys and Asp that could form hydrogen bond to the substrate are also located (Fig. 2-10). Considering that TmaFuc with Thr283 at this position has α-L-galactosidase activity, the distinction of O6 in GH29 is not likely strict. Based on the above, the divergence of α-L-glucosidase and α-L-fucosidase in GH29 is most likely due to the divergence of the O4 recognition mechanism. The precise recognition mechanism of α-L-glucosidase disclosed in this study prompts us to imagine the presence of natural α-L-glucoside. In addition, a BLAST search using ClAgl29B as the query sequence revealed that the top 45 amino acid sequences of the hits might be endowed with the substrate recognition mechanism of α-L-

glucosidases (Table 2-4). These sequences were derived from the Bacteroidetes, Verrucomicrobia, Kiritimatiellaeota, Planctomycetes, Terrabacteria, and Armatimonadetes bacterial phyla, and Asgard archaea. On the other hand, the recognition of moderate O6 indicates that 6-deoxy α-L-glucoside (α-L-quinovoside) could be a possible substrate. The evaluation of α-L-quinovoside hydrolysis activity will be discussed in Chapter 3.

The relationship between the diversity at the position corresponding to His34 in TmaFuc (Fig. 2-1) and Asp114 in ClAgl29B and their function was interesting. However, the hydrolytic activities on *p*NP α-L-Fuc, *p*NP α-L-Glc, and *p*NP α-L-Gal of the recombinant Caka0074, Caka0506, Caka0522, Caka0523, and Caka0536 were too low to clearly elucidate their function in this study. Enzymes classified as GH29-B have strict substrate specificity because aglycon binding is critical for catalytic activity, and they are known not to act on *p*NP glycosides (Ashida, H., 2009, Sakurama H., 2012a). Phylogenetic analysis revealed that the five Caka proteins are classified as GH29-A (Fig. 2-11), suggesting that substrate specificity of aglycones is not the cause of their low catalytic activity. Multiple sequence alignment indicated that these Caka proteins conserved catalytic amino acid residues, but several amino-acid residues associated with the substrate recognition were diverged (Fig. 2-12). In other words, these enzymes may have unknown substrate specificity. *C. akajimensis* JCM 23193 is isolated from sea water. It is assumed that this bacterium has many enzymes involved in the degradation of fucoidan and its derivatives, such as sulfated polysaccharides, contained in seaweeds. This bacterium encodes 11 GH29 enzymes in its genome, suggesting that it has a variety of enzymes involved in the degradation of carbohydrates.

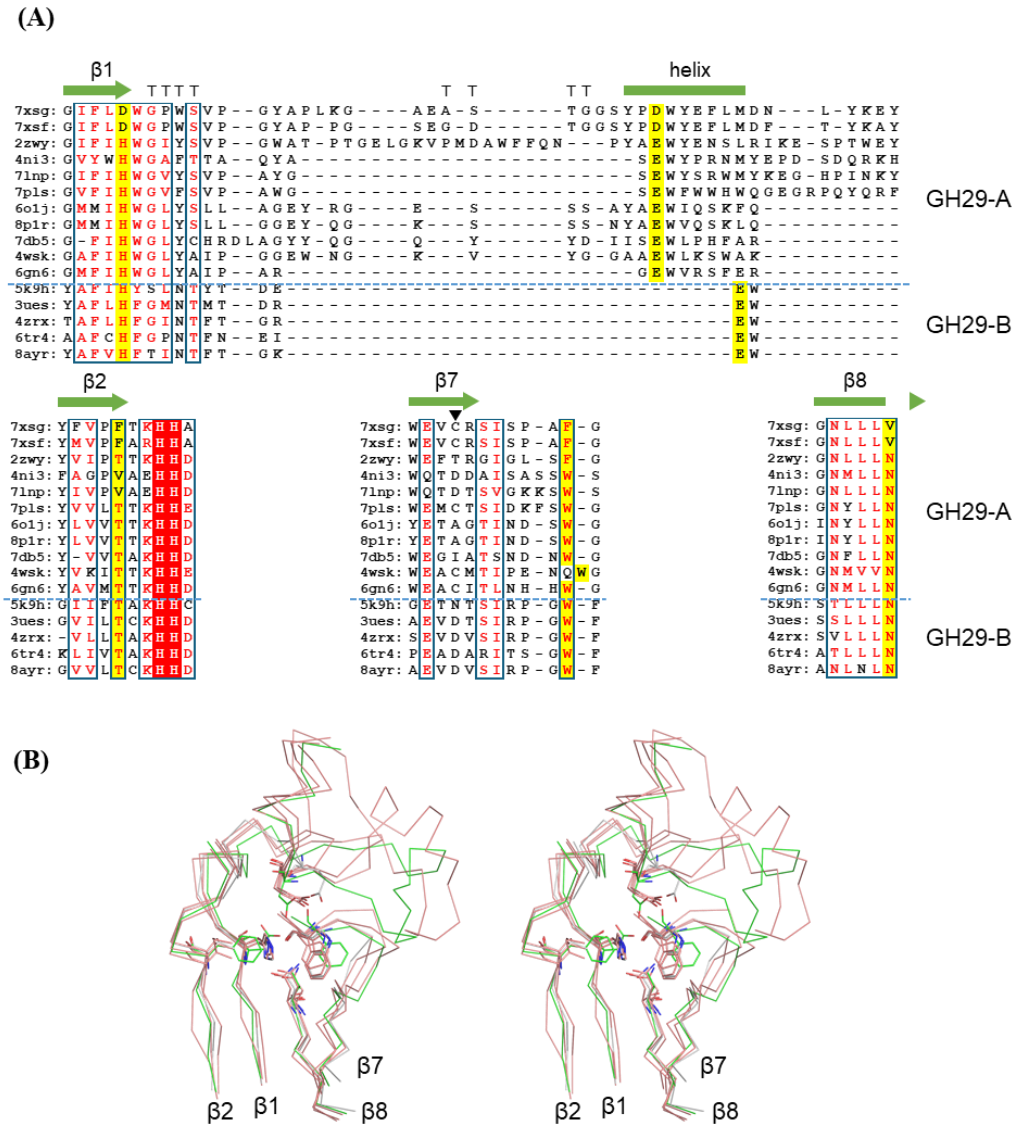


Fig. 2-10. Partial multiple amino-acid sequence alignments of structure-known GH29 enzymes (A) and comparison of 3D structures of important sites for α-L-glucoside recognition (B).

(A) An amino-acid sequence alignment was prepared based on the multiple structural alignment using MUSTANG. Amino acid residues that are essential for α-L-glucosidase to recognize the equatorial O4 and their corresponding residues are highlighted in yellow. The inverted triangle marks a site for Cys418 of which possibly hydrogen bonds with O6 of β-L-glucose in ClAgl29B. (B) The spatial arrangement of the highlighted amino acid residues in (A) are shown. The α-L-glucosidase (7xsg) is shown in green, the GH29-A enzymes (2wzy, 4ni3, and 4wsk) in wheat, and the GH29-B enzyme (3ues) in gray.

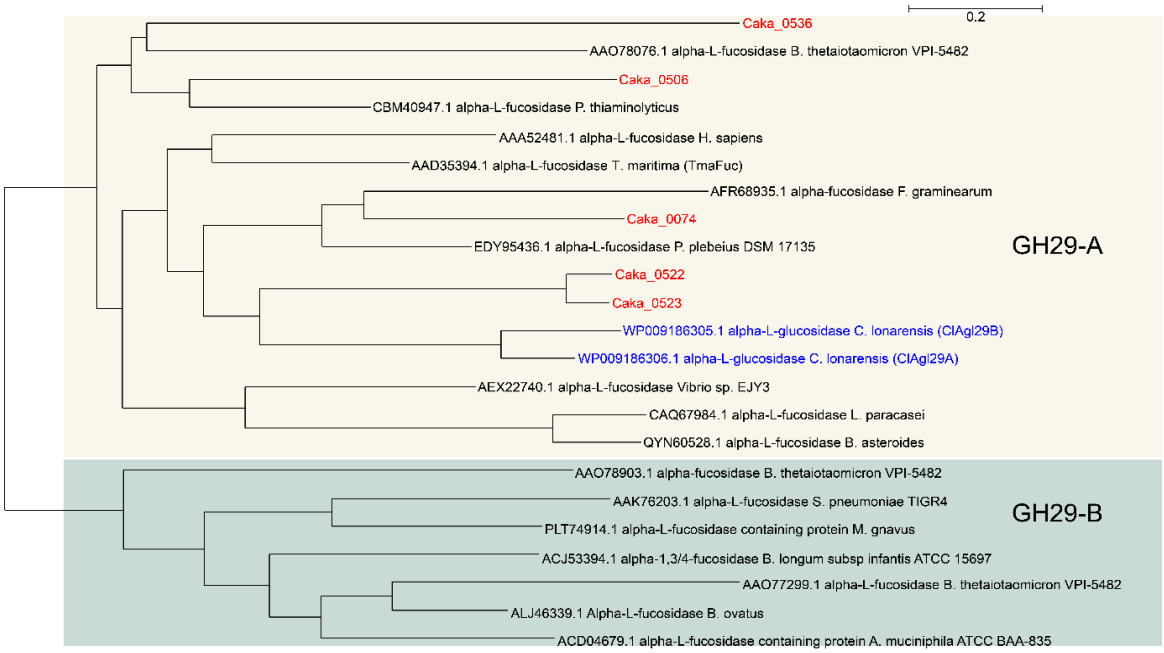


Fig. 2-11. Neighbor-joining phylogenetic tree of GH29 3D-structure-known enzymes and five GH29 sequences by *C. akajimensis*.

Sequence alignment and phylogenetic tree were conducted by MAFFT version 7 online version (https://mafft.cbrc.jp/alignment/server/).

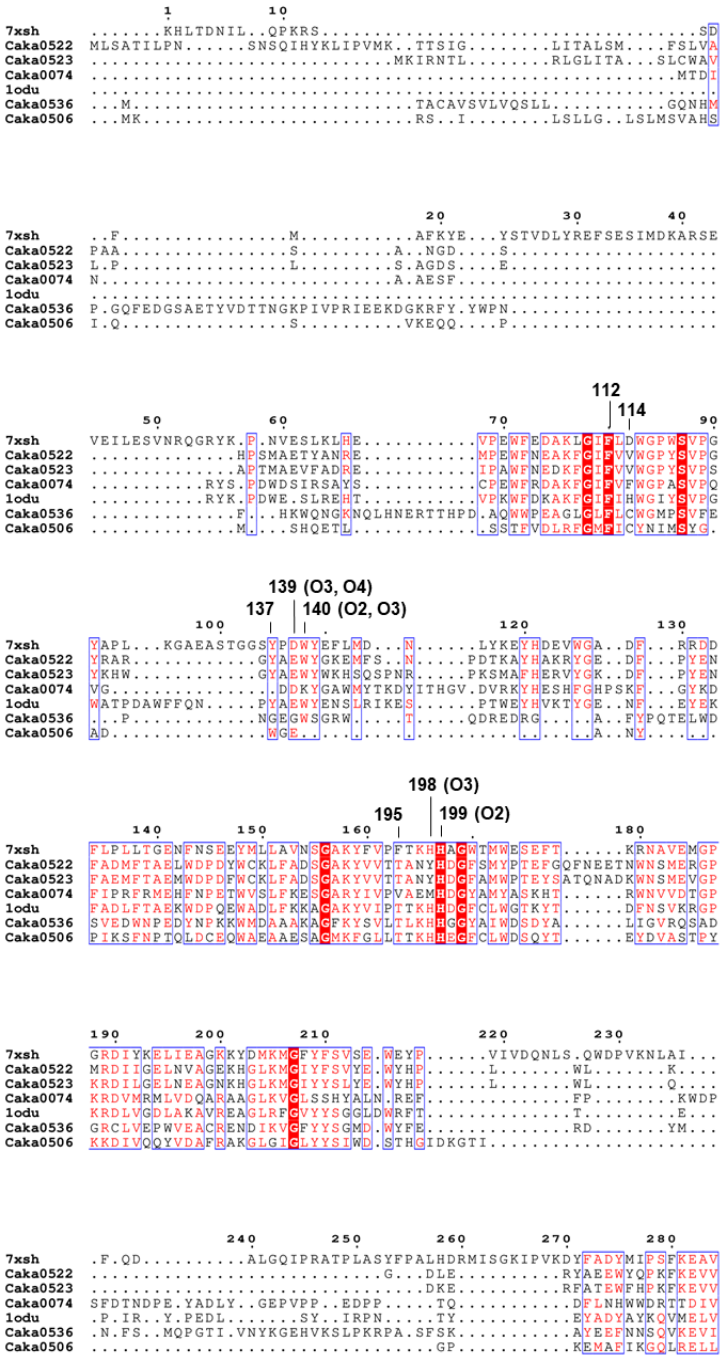


Fig. 2-12. Multiple protein structural alignment of ClAgl29B, TmaFuc, and five GH29 proteins from *C. akajimensis* JCM 23193.

Continued on next page.

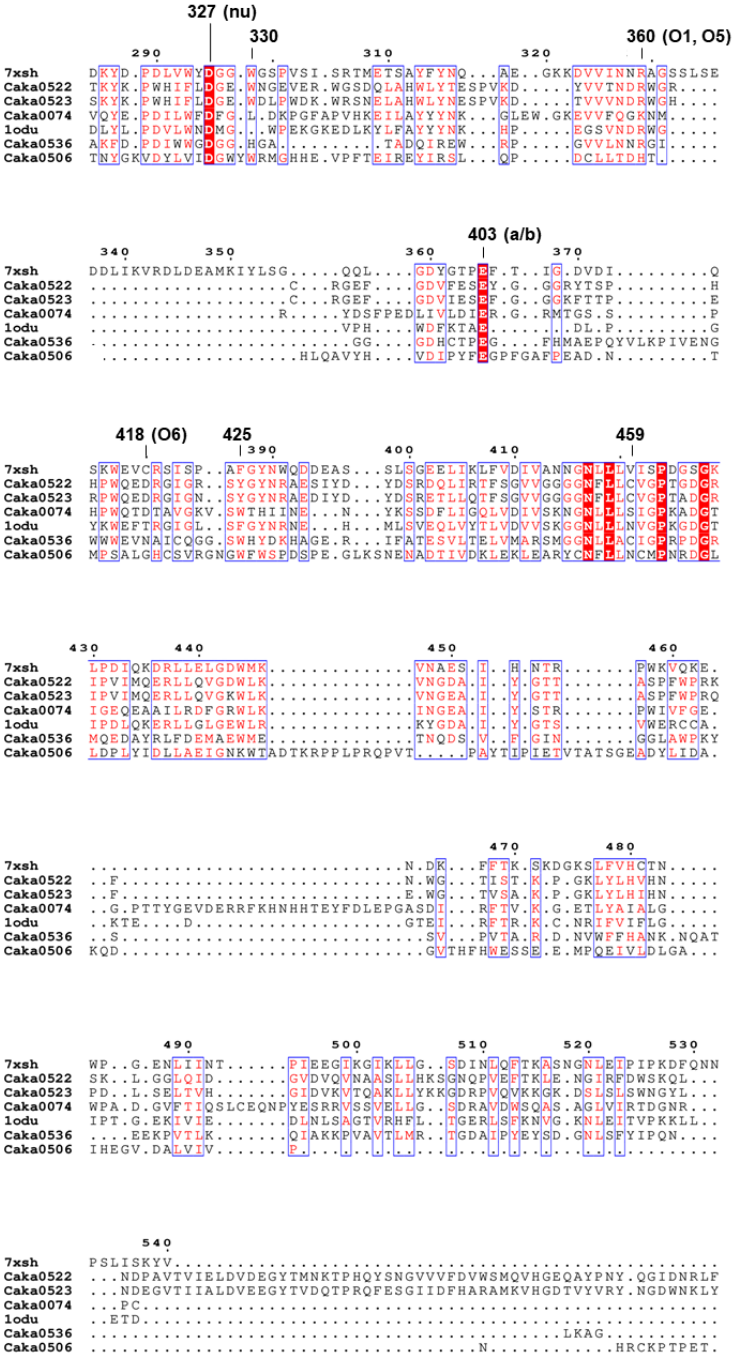


Fig. 2-12. Multiple protein structural alignment of ClAgl29B, TmaFuc, and five GH29 proteins from *C. akajimensis* JCM 23193.

Continued on next page.

Fig. 2-12. Multiple protein structural alignment of ClAgl29B, TmaFuc, and five GH29 proteins from *C. akajimensis* JCM 23193.

The multiple protein structural alignment was constructed using MUSTANG. The 3D structures of ClAgl29B and TmaFuc used were chains A of 7xsh and 1odu, respectively. The predicted Caka proteins structures were obtained from AlphaFold Protein Structure Database. In the β-L-glucose complex of ClAgl29B, amino acid residues located within 4 Å of the ligand are numbered. The residue numbers are for ClAgl29B, and the accompanying parentheses indicate the partner with whom a hydrogen bond is formed. Nu and a/b represent catalytic nucleophile and general acid/base residues, respectively.

Table 2-4. List of putative α-L-glucosidases

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. ident	Acc. Len	Accession
alpha-L-fucosidase	*Cecembia lonarensis*	1164	1164	100%	0	100	586	WP_009186305.1
alpha-L-fucosidase	*Cyclobacteriaceae bacterium*	775	775	99%	0	63.72	589	WKZ61350.1
alpha-L-fucosidase	*Cytophagales bacterium*	682	682	97%	0	57.04	569	MCA6380965.1
alpha-L-fucosidase	*Alistipes sp.*	401	401	97%	5.00E-130	38.6	530	MBQ3247166.1
alpha-L-fucosidase	*Bacteroidales bacterium*	364	364	96%	2.00E-115	37.72	551	MCF8225137.1
alpha-L-fucosidase	candidate division KSB1 bacterium	402	402	96%	3.00E-129	39.93	633	NQT27003.1
alpha-L-fucosidase	*Prevotellaceae bacterium*	389	389	96%	2.00E-124	38.87	577	MDR0559157.1
alpha-L-fucosidase	*Phycisphaerae bacterium*	349	349	96%	6.00E-107	35.82	766	MBN1488282.1
alpha-L-fucosidase	*Cytophagales bacterium*	774	774	96%	0	65.93	573	MCA6380964.1
alpha-L-fucosidase	*Cecembia lonarensis*	687	687	96%	0	60.04	573	WP_009186306.1
alpha-L-fucosidase	*Bacteroidia bacterium*	430	430	96%	3.00E-141	40.14	542	MCX6332918.1
alpha-L-fucosidase	*Bacteroidales bacterium*	400	400	96%	2.00E-129	39.18	556	MBL7110955.1
alpha-L-fucosidase	*Prolixibacteraceae bacterium*	400	400	96%	2.00E-129	38.09	542	MDX9881157.1
alpha-L-fucosidase	*Bacteroidales bacterium*	389	389	96%	4.00E-125	38.63	548	HJZ40941.1
hypothetical protein AMS26_19480	*Bacteroides sp. SM23_62*	389	389	96%	9.00E-125	38.46	554	KPL11806.1
alpha-L-fucosidase	*Bacteroidota bacterium*	731	731	96%	0	63.69	579	MBM3442356.1
alpha-L-fucosidase	*Cyclobacteriaceae bacterium*	700	700	96%	0	60.88	571	WKZ61351.1

alpha-L-fucosidase	*Alistipes sp.*	394	394	96%	8.00E-127	37.64	556	MBQ7341979.1
alpha-L-fucosidase	*Alistipes sp.*	388	388	96%	1.00E-124	37.45	556	MBQ3260723.1
alpha-L-fucosidase	*Bacteroides sp.*	370	370	96%	1.00E-117	38.26	550	HER08986.1
alpha-L-fucosidase	*Parabacteroides*	363	363	96%	1.00E-114	37.59	564	WP_122357195.1
alpha-L-fucosidase	*Verrucomicrobiota bacterium*	343	343	96%	2.00E-105	35.47	689	MCX6924306.1
alpha-L-fucosidase	*Bacteroidota bacterium*	335	335	96%	4.00E-104	35.26	541	MDX2190639.1
alpha-L-fucosidase	*Verrucomicrobiota bacterium*	336	427	96%	1.00E-98	36.01	1218	MCX6926123.1
alpha-L-fucosidase	*Bacteroides sp.*	379	379	96%	6.00E-121	38.49	557	MDX2432295.1
alpha-L-fucosidase	*Tannerella sp.*	373	373	96%	2.00E-118	38.45	565	MDR3252040.1
alpha-L-fucosidase	*Mangrovibacterium sp.*	398	398	95%	1.00E-128	38.32	544	MDD4191703.1
alpha-L-fucosidase	*Candidatus Thermoplasmatota*	352	352	95%	2.00E-110	36.35	571	MDD5502387.1
alpha-L-fucosidase	*Planctomycetota bacterium*	347	347	95%	3.00E-106	35.96	755	MCY3017553.1
alpha-L-fucosidase	*Prevotellaceae bacterium*	357	357	95%	2.00E-112	39.5	573	MDR3245229.1
alpha-L-fucosidase	*Prevotellaceae bacterium*	348	348	95%	2.00E-108	36.72	598	MDR1898005.1
alpha-L-fucosidase	*Opitutales bacterium*	343	343	95%	6.00E-107	36.33	568	MBE6412545.1
alpha-L-fucosidase	*Tannerella sp.*	368	368	95%	1.00E-116	38.56	579	MDR1455543.1
alpha-L-fucosidase	*Bacteroides sp. GM023*	383	383	94%	8.00E-123	38.12	540	WP_191031480.1
alpha-L-fucosidase	*Tannerellaceae bacterium*	382	382	94%	2.00E-122	37.91	537	MDR2810484.1
alpha-L-fucosidase	*Tannerella sp.*	369	369	94%	5.00E-117	37.08	568	MDR1813610.1
alpha-L-fucosidase	*Parabacteroides faecis*	366	366	94%	5.00E-116	38.28	564	WP_258914886.1
alpha-L-fucosidase	*Bacteroidales bacterium*	345	345	94%	8.00E-108	37.08	554	MDR3094571.1

alpha-L-fucosidase	*Bacteroidales bacterium*	361	361	94%	4.00E-114	37.41	559	HYW96709.1
alpha-L-fucosidase	*Opitutales bacterium*	325	325	94%	7.00E-100	35.44	588	MBQ6534138.1
alpha-L-fucosidase	*Tannerellaceae bacterium*	373	373	94%	1.00E-118	38.71	576	MDR0537976.1
alpha-L-fucosidase	*Tannerella sp.*	377	377	94%	3.00E-120	37.99	545	MDR3266865.1
alpha-L-fucosidase	*Tannerella sp.*	370	370	94%	6.00E-118	38.18	539	MDR0845399.1
alpha-L-fucosidase	*Tannerella sp.*	375	375	93%	1.00E-119	38.29	541	MDR2680875.1
alpha-L-fucosidase	*Cytophagales bacterium*	682	682	97%	0	57.04	569	MCA6380965.1

Chapter 3. Substrate specificity of GH29 L-glucosidases from *C. lonarensis*

1. Introduction

C. lonarensis α-L-glucosidases, ClAgl29A and ClAgl29B, belonging to GH29, a family of α-L-fucosidases, exhibit the substrate specificity for α-L-glucoside. The structural analysis in Chapter 2 revealed that ClAgl29A and ClAgl29B have a structure that excludes α-L-fucoside and accepts α-L-glucoside and that seem to have evolved to recognize α-L-glucoside, suggesting that the α-L-glucoside might exist in the growing environment of *C. lonarensis*. On the other hand, given that the existence of α-L-glucoside has not yet been demonstrated, it is possible that *C. lonarensis* α-L-glucosidases may retain substrate selectivity for other α-L-glycosides as substrates, rather than α-L-glucoside. Therefore, evaluating the hydrolysis activities of ClAgl29A and ClAgl29B toward several α-L-glycosides would be demanded to solve the uncertainties for the substrate specificity.

In nature, α-L-fucoside (6-deoxy α-L-galactoside) and α-L-rhamonoside (6-deoxy α-L-mannoside), which are similar to α-L-glucoside (Fig. 3-1), are abundant. α-L-Fucosides are widely found in living organisms as a non-reducing terminal sugar of complex carbohydrates and is a component of the polysaccharide fucoidan found in brown algae. α-L-Rhamonosides are found in a glycon part of various glycosides, and also in the gums and polysaccharides of plants and microorganisms. α-L-Quinovoside (6-deoxy α-L-glucoside) (Fig. 3-1), or more precisely α-L-quinovosyl, has been suggested to be present as a component sugar of the *O*-antigens of pathogenic bacteria, including *Providencia stuartii O44*, which is associated with urinary tract infections (Kocharova, N. A., 2005) and *Yersinia pseudotuberculosis*, which is a cause of pseudotuberculosis (De Castro, C., 2013). Our previous study has demonstrated that the two enzymes do not use α-L-fucoside as a substrate (Shishiuchi, R., 2022), but their substrate specificity for α-L-rhamnoside and α-L-quinovoside remains unknown.

In this chapter, *p*NP α-L-rhamnoside (*p*NP α-L-Rha), *p*NP α-L-quinovoside (*p*NP α-L-Qui), and *p*NP α-L-xyloside (*p*NP α-L-Xyl) were chemically synthesized, and substrate specificities of ClAgl29A and ClAgl29B for these α-L-glycosides were evaluated. Although α-L-xyloside (Fig. 3-1) has not been known to exist in nature, the substrate specificity toward this glycoside as an analog of α-L-glucoside was of interest. In addition, the bindings of α-L-rhamnoside, α-L-quinovoside, and α-L-xyloside to ClAgl29B were discussed through molecular docking simulation of methyl of α-L-glucoside (Me α-L-Glc), methyl of α-L-quinovoside (Me α-L-Qui), methyl of α-L-rhamnoside (Me α-L-Rha), and methyl of α-L-xyloside (Me α-L-Xyl).

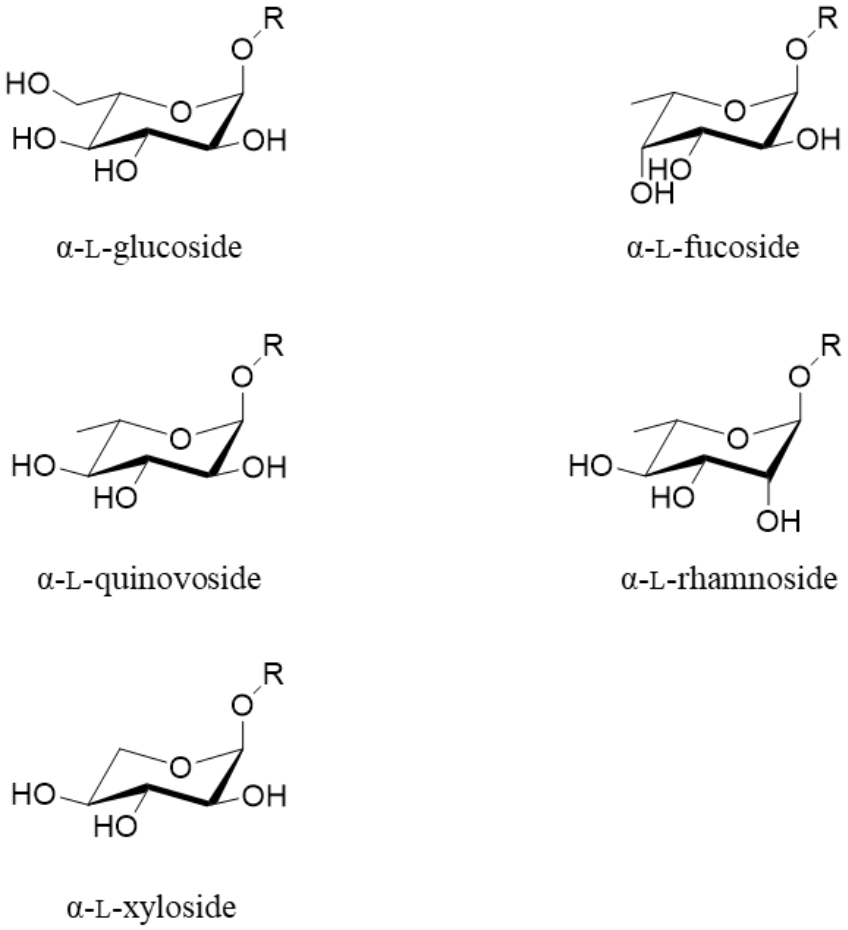


Fig. 3-1. α-L-Glycosides featured in this chapter.

2. Materials and methods

2.1 Materials

L-xylose was purchased from Combi-Blocks (San Diego, CA, USA) and L-rhamnose was obtained from Nacalai Tesque (Kyoto, Japan). All the chemicals and solvents used were of analytical grade.

2.2 Production and purification of ClAgl29A and ClAgl29B

The N-terminal His_6-tagged ClAgl29A and ClAgl29B were produced in *E. coli* Rosetta (DE3) and purified as described in Chapter 2.2.5.

2.3 Synthesis of *p*-nitrophenyl α-L-glycosides

【*p*NP α-L-Rha】

1,2,3,4-tetra-*O*-acetyl L-rhamnose: To a solution of -rhamnose (1.00 g, 6.1 mmol) in dry pyridine (10 mL) at 0°C under a nitrogen atmosphere was slowly added acetic anhydride (10 mL, 106 mmol). The reaction mixture was stirred at 0°C for 1 h before adding a catalytic amount of 4-dimethylaminopyridine (DMAP; 73 mg, 0.61 mmol). As the reaction mixture was allowed to reach room temperature. After 6 h, the transparent yellow mixture was slowly poured into rapidly stirred ice water (50 mL), giving a sticky solid. The acetylated -rhamnose was extracted in ethyl acetate, washed with saturated sodium bicarbonate solution, and dried with anhydrous sodium sulfate. The solvent was evaporated and co-evaporated with toluene (100 mL) several times, and tetraacetylated L-rhamnose was obtained as a yellow syrup.

2,3,4-tetra-*O*-acetyl L-rhamnosyl trichloroacetimidate (Andersen, S. M., 2015): A solution of 1,2,3,4-tetra-*O*-acetyl-L-rhamnose (5.8 mmol, 1.94 g) and 3-(dimethylamino)-1-propylamine (DMAPA; 5 equiv. 2.96 g, 29 mmol) in dichloromethane (30 mL) was stirred at room temperature. The reaction is followed by TLC analysis until complete conversion. After 3 h, trichloroacetonitrile (10 equiv. 5.8 mL, 58 mmol) and 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU, 0.2 equiv. 0.29 mL) were added to the reaction and the reaction change to a yellow color. The reaction was stirred for another 45 min. The reaction mixture was diluted in dichloromethane, and the organic phase was washed with 1 M hydrochloric acid and saturated aqueous bicarbonate. The organic phase was dried over anhydrous sodium sulfate, filtered, and concentrated under reduced pressure. The crude product was co-evaporated with toluene several times.

***p*NP α-L-Rha**: A solution of 2,3,4-tetra-*O*-acetyl L-rhamnosyl trichloroacetimidate (5.8 mmol, 2.5 g), *p*-nitrophenol (10 equiv. 0.8 g, 58 mmol), and boron trifluoride diethyl ether complex (0.2 mL, Fujifilm Wako) in dichloromethane (15 mL) was stirred at room temperature for 5 h. The reaction mixture was diluted in dichloromethane, and the organic phase was washed with saturated aqueous bicarbonate and brine. The organic phase was dried over anhydrous sodium sulfate and

concentrated under reduced pressure. Flash column chromatography (hexane/ethyl acetate = 3/2) of the crude products provided pure *p*-nitrophenyl per *O*-acetyl α-L-rhamnoside. The acetylated glycosides were deacetylated with sodium methoxide in methanol. For NMR analysis of *p*NP α-L-Rha, the following *p*NP α-L-Qui and *p*NP α-L-Xyl described below were insoluble in deuterium oxide or methanol-d4. The compounds were washed several times with methanol-d4 and dissolved in dimethyl sulfoxide-d6 for NMR measurement. However, the H/D substitutions were insufficient, and the ^{1}H signals derived from the hydroxyl groups were detected. ^{1}H NMR (500 MHz, DMSO-d6) δ (ppm): 8.18 (d, *J* = 9.3 Hz, 2H, PNP), 7.21 (d, *J* = 9.3 Hz, 2H, PNP), 5.54 (d, *J* = 1.7 Hz, 1H, H-1), 5.18 (d, *J* = 5.2 Hz, 1H, OH-2), 4.95 (d, *J* = 5.0 Hz, 1H, OH-4), 4.78 (d, *J* = 4.8 Hz, 1H, OH-3), 3.86 (t, *J* = 4.6 Hz, 1H, H-2), 3.64 (ddd, *J* = 3.5, 5.7 and 9.1 Hz, 1H, H-3), 3.38 (qd, *J* = 6.1 and 9.6 Hz, 1H, H-2), 3.30 (td, *J* = 5.6 and 9.3 Hz, 1H, H-4), 1.08 (d, *J* = 6.1 Hz, 3H, H-6); ^{13}C NMR (126 MHz, DMSO-d6) δ (ppm): 161.4 (PNP), 141.8 (PNP), 125.9 (PNP), 116.8 (PNP), 98.6 (C-1), 71.8 (C-4), 70.4 (C-3), 70.2 (C-5), 70.0 (C-2), 18.0 (C-6); $^{1}J_{CH}$ = 172.8 Hz.

【*p*NP α-L-Qui】

1,2,3,4-tetra-*O*-acetyl L-quinovose: A solution of L-rhamnose (2.00 g, 6.1 mmol) in molybdic acid (100 mg) in water (20 mL) was exposed to microwave irradiation using a microwave oven operating at 500 W for 5 min (Hricoviniova, Z., 2009). The ion-exchange resin DOWEX 1 × 4 in the HCO^{3-} form removed the catalyst. The reaction mixture was filtered, evaporated, and co-evaporated with toluene several times. The resulting mixed syrup of L-rhamnose and L-quinovose was acetylated according to the method shown in the previous section. Flash column chromatography used a mixed hexane and ethyl acetate solvent in a 2:1 ratio to isolate per *O*-acetyl L-quinovose from the crude products.

***p*NP α-L-Qui** (Apparu M, 1981): 1,2,3,4-tetra-*O*-acetyl L-quinovose (6 mmol, 2 mg) and *p*-nitrophenol (7.5 mmol, 1.04 g) were dissolved in stannic chloride (Tin(IV) chloride, ca. 1.0 mol/L in dichloromethane, TCI; 10 mL) was stirred at 60°C for 5.5 h. The reaction mixture was diluted in dichloromethane, poured into an ice-cold saturated bicarbonate solution, further neutralized with solid sodium bicarbonate, and filtered on a Celite bed. The organic phase was washed with saturated aqueous bicarbonate and brine. The organic phase was dried over anhydrous magnesium sulfate and concentrated under reduced pressure. Flash column chromatography (hexane/ethyl acetate = 5/2) of the crude products provided pure *p*-nitrophenyl per *O*-acetyl α-L-quinovoside. The acetylated glycosides were deacetylated with sodium methoxide in methanol. ^{1}H NMR (500 MHz, DMSO-d6) δ (ppm): 8.19 (d, J = 9.3 Hz, 2H, PNP), 7.22 (d, J = 9.3 Hz, 2H, PNP), 5.58 (d, J = 3.6 Hz, 1H, H-1), 5.24 (d, J = 6.2 Hz, 1H, OH-2), 5.14 (d, J = 5.7 Hz, 1H, OH-4), 5.01 (d, J = 4.9 Hz, 1H, OH-3), 3.58 (m, 1H, H-3), 3.45 (m, 1H, H-5), 3.44 (m, 1H, H-2), 2.92 (td, J = 5.6 and 9.2 Hz, 1H, H-4),

1.06 (d, J = 6.3 Hz, 3H, H-6); ^{13}C NMR (126 MHz, DMSO-d6) δ (ppm): 162.3 (PNP), 141.8 (PNP), 126.0 (PNP), 116.9 (PNP), 97.6 (C-1), 75.9 (C-3), 75.5 (C-4), 71.6 (C-2), 69.2 (C-5), 18.0 (C-6); $^{1}J_{CH}$ = 173.1 Hz.

【*p*NP α-L-Xyl】

L-Xylose was acetylated according to the method described above. The addition of the *p*-nitrophenyl group to the obtained 1,2,3,4-tetra-*O*-acetyl L-xylose followed the synthetic method of *p*-nitrophenyl per *O*-acetyl α-L-quinovoside using Tin(IV) chloride as a catalyst. Flash column chromatography (hexane/ethyl acetate = 5/2) of the crude products provided pure p-nitrophenyl per *O*-acetyl α-L-xyloside. The acetylated glycosides were deacetylated with sodium methoxide in methanol. Two sugar ring conformations, $^{1}C_4$ and $^{4}C_1$, were detected in an H-1 integral ratio of 1.18:0.84. ^{1}H NMR (500 MHz, DMSO-d6) δ (ppm): 8.19 (d, J = 9.3 Hz, PNP), 7.22 (d, J = 9.3 Hz, PNP), 5.61 (d, J = 3.5 Hz, H-1, 1C4), 5.46 (d, J = 4.1 Hz, H-1, 4C1), 5.26 (d, J = 6.3 Hz, 1H, OH-2, 1C4), 5.21 (d, J = 5.3 Hz, OH-2, 4C1), 5.13 (d, J = 5.3 Hz, OH-4, 1C4), 5.06 (d, J = 4.9 Hz, OH-3, 1C4), 5.04 (d, J = 4.4 Hz, OH-4, 4C1), 4.97 (d, J = 4.8 Hz, OH-3, 4C1), 3.82 (m, H-2, 4C1), 3.70 (m, H-3, 4C1), 3.69 (m, H-5a, 4C1), 3.68 (m, H-4, 4C1), 3.57 (m, H-3, 1C4), 3.53 (m, H-5a, 1C4), 3.45 (m, H-2, 1C4), 3.40 (m, H4, 1C4), 3.32 (dd, J = 6.8 and 10.9 Hz, H-5b, 4C1), 3.26 (t, J = 10.7 Hz, H-5b, 1C4); 13C NMR (126 MHz, DMSO-d6) δ (ppm): 162.1 (PNP), 141.9 (PNP), 126.0 (PNP), 116.9 (PNP), 98.7 (C-1, 4C1), 97.7 (C-1, 1C4), 73.3 (C-3, 1C4), 71.4 (C-2, 1C4), 71.0 (C-3, 4C1), 69.5 (C-4, 1C4), 69.0 (C-2, 4C1), 67.1 (C-4, 4C1), 64.6 (C-5, 4C1), 63.2 (C-5, 1C4); $^{1}J_{CH}$ (1C4) = 177.0 Hz, $^{1}J_{CH}$ (4C1) = 170.9 Hz.

2.4 Enzyme reaction

Reaction activities toward 2 mM *p*NP L-Qui, *p*NP L-Rha, and *p*NP L-Xyl were determined in 40 mM sodium acetate buffer (pH 5.5) at 35°C for 10 min. The reaction was terminated by adding two volumes of 1 M sodium carbonate and the concentration of *p*NP released was measured as described in the method in 2.7 of Chapter 2.

To determine kinetic parameters, the initial velocities for hydrolysis of various concentrations of substrates (0.025 – 2.0 mM for *p*NP L-Qui and 0.25 – 3.0 mM for *p*NP L-Xyl) were measured in 40 mM sodium acetate buffer (pH 5.5) at 35°C for 10 min. The concentrations of ClAgl29A used were 50.4 nM for *p*NP L-Qui and 75.7 nM for *p*NP L-Xyl and those of ClAgl29B were 37.2 nM for *p*NP L-Qui and 62.0 nM for *p*NP L-Xyl. GraFit 7.02 (Erithacus Software) was used to fit the reaction data to the reaction rate equations and to calculate the rate parameters.

2.5 Molecular docking simulation

A molecular docking study was performed using the GNINA program. The structure data file of Me α-L-Rha was obtained from PDB database with SYBYL Mol2 file format (RAO.mol2). For those of Me α-L-Glc, Me α-L-Qui and Me α-L-Xyl, these molecules were drawn in ChemDraw Professional 19.0 (PerkinElmer Informatics) and saved as structure-data file format (.sdf). The structure of ClAgl29B (accession code: 7XSH) complex with ligand β-L-glucose (β-L-Glc) was collected from PDB. The hetero atoms, including ligand β-L-Glc and water, were removed by PyMol Molecular Graphic System 2.5.0 (Schrödinger LLC, NY, USA) before initiating the docking study. The β-L-Glc binding site of 7XSH was defined as the ligand binding site and docked.

3. Results

3.1 Production and purification of ClAgl29A and ClAgl29B

The N-terminal His_6-tagged ClAgl29A and ClAgl29B were produced in *E. coli* Rosetta (DE3) and purified using Ni-affinity column chromatography, obtaining 9.65 U of ClAgl29A, and 3.04 U of ClAgl29B from 500 mL culture broth. The ClAgl29A yielded 3.29 mg with a specific activity of 0.561 U/mg and the yield of ClAgl29B was 6.25 mg with a specific activity of 2.54 U/mg. SDS-PAGE analysis of purified ClAgl29A and ClAgl29B gave a single band, and as shown previously, they showed a migration at a smaller molecular weight than the expected molecular weight (Fig. 3-2).

3.2 Kinetic studies for *p*NP α-L-glycosides

First, the hydrolytic rates of ClAgl29A and ClAgl29B on 2 mM *p*NP α-L-Rha, *p*NP α-L-Qui, and *p*NP α-L-Xyl were examined (Table 3-1). The enzymes hardly hydrolyzed *p*NP α-L-Rha and its hydrolytic rate was 1/860 of *p*NP α-L-Qui in ClAgl29A and 1/1200 in ClAgl29B.

Next, kinetic parameters for *p*NP α-L-Qui and *p*NP α-L-Xyl were determined. The initial reaction rates at various substrate concentrations of *p*NP α-L-Qui and *p*NP α-L-Xyl followed the Michaelis-Menten equation (Fig. 3-3) with kinetic parameters in Table 3-2. A comparison of k_{cat}/K_m values showed that ClAgl29A and ClAgl29B have higher substrate specificity for *p*NP α-L-Qui than for *p*NP α-L-Glc. The k_{cat}/K_m value for *p*NP α-L-Qui was 7.4 times higher than that for *p*NP α-L-Glc in ClAgl29A, and the k_{cat}/K_m value for *p*NP α-L-Qui was 2.8 times higher than that for *p*NP α-L-Glc in ClAgl29B. The higher k_{cat}/K_m values for *p*NP α-L-Qui in both enzymes relied on the lower K_m value. The K_m value of *p*NP α-L-Qui was 7.5-fold lower than that of *p*NP α-L-Glc for ClAgl29A and 4.3-fold lower for ClAgl29B. The substrate specificity of both enzymes for *p*NP α-L-Xyl was found to be poor. The k_{cat}/K_m value for *p*NP α-L-Glc was 3.5 times higher than that for *p*NP α-L-Xyl in ClAgl29A, and the k_{cat}/K_m value for *p*NP α-L-Glc was 5.6 times higher than that for *p*NP α-L-Xyl in ClAgl29B.

3.3 Molecular docking studies

The molecular docking simulations of Me α-L-Glc, Me α-L-Qui, Me α-L-Rha, and Me α-L-Xyl on ClAgl29B were performed using the GNINA program. The calculated affinities are summarized in Table 3-3. In all ligands, the mode with the highest affinity among the nine results returned by the GNINA program showed similar to the β-L-Glc bound to ClAgl29B by X-ray crystallography, but with some differences (Fig 3-4, Table 3-4). Although the residues that could form hydrogen bonds with the ligand in Me α-L-Glc, Me α-L-Qui, and Me α-L-Xyl bonds were similar, the catalytic nucleophilic residue Asp327 side chain was predicted to form a hydrogen bond with the O6 of Me

α-L-Glc, whereas with β-L-Glc a hydrogen bond was formed with the O1 (Fig 3-4, Table 3-4). The orientation of O6 of Me α-L-Glc was changed from that of β-L-Glc, and the hydrogen bond with Sγ of Cys418 was eliminated. In the docking simulation with Me α-L-Rha, Nε1 of Trp140 did not form a hydrogen bond with the substrate, but instead, the Oη of Tyr241 and Oδ2 of the catalytic nucleophilic residue Asp327 can form a hydrogen bond with O2.

In the active site of ClAgl29B with L-glucose, the equatorial OH-2 can form hydrogen bonds with Nε1 of Trp140 and Nε2 of His199 (Fig. 3-4, Table 3-4). These non-covalent bonds could neutralize the electron-withdrawing property of the O2 atom of the substrate and contribute to the stabilization of the oxocarbenium ion-like transition state. Me α-L-Rha with the axial OH-2 would not benefit from this hydrogen bonding acceleration effect and would be less likely to be a substrate (Fig. 3-4, Table 3-4). The bulky side chain of Tyr241 in the active site appeared to hinder the interaction with the axial O2 of Me α-L-Rha in ClAgl29B (Fig. 3-5). The active-site pocket of ClAgl29A possesses the corresponding Trp128 and His187 at the respective sites and they would have identical mechanisms for their substrate recognition.

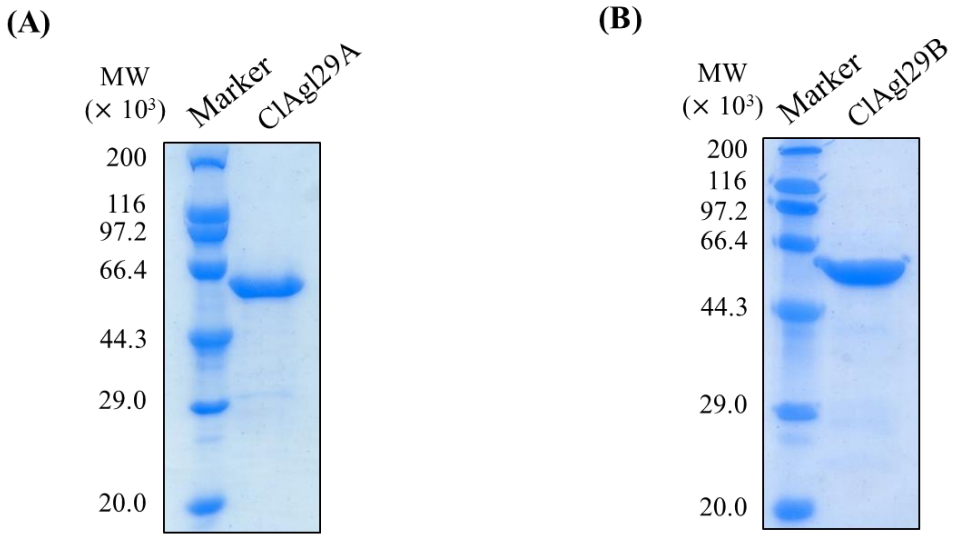


Fig. 3-2. SDS-PAGE of purified recombinant ClAgl29A and ClAgl29B. Lane marker, size marker; lane ClAgl29A, ClAgl29A/pET28a purified by Ni-affinity chromatography (A_{280}=0.73, 1.4 μg); lane ClAgl29B, ClAgl29B/pET28a purified by Ni-affinity chromatography (A_{280}=2.49, 3.7 μg). The molecular weight of the standards is shown on the left.

Table 3-1. Hydrolytic rates of ClAgl29A and ClAgl29B on *p*NP α-L-Rha, *p*NP α-L-Qui, and *p*NP α-L-Xyl

	μmol min^{-1} mg^{-1}
ClAgl29A	
*p*NP α-L-Rha	0.00414
*p*NP α-L-Qui	3.55 ± 0.30
*p*NP α-L-Xyl	0.958 ± 0.043
ClAgl29B	
*p*NP α-L-Rha	0.00216
*p*NP α-L-Qui	2.56 ± 0.41
*p*NP α-L-Xyl	0.658 ± 0.03

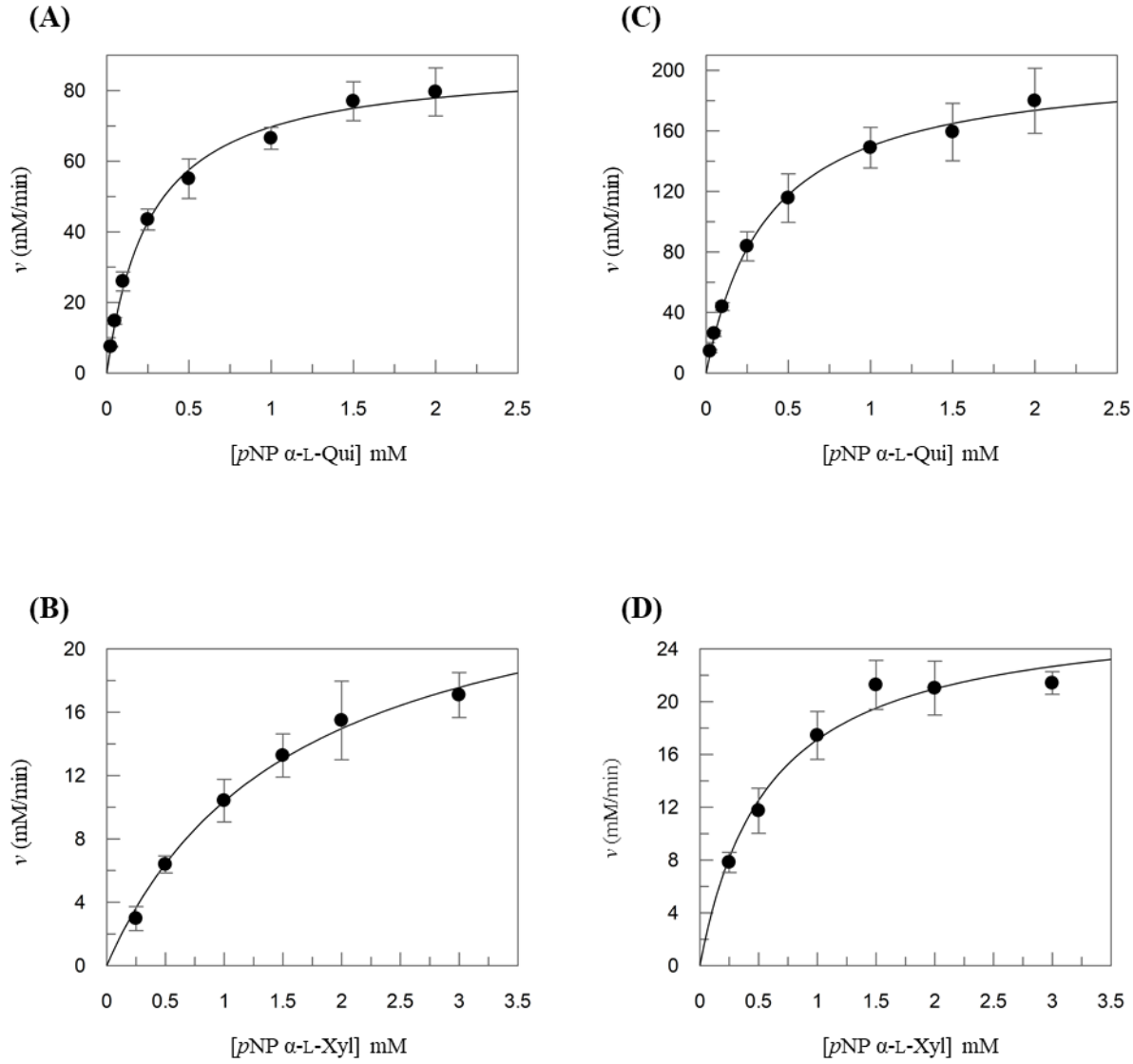


Fig. 3-3. [S]–*v* plots of ClAgl29A and ClAgl29B for *p*NP α-L-Qui and *p*NP α-L-Xyl.

Hydrolysis of *p*NP α-L-Qui (A) and *p*NP α-L-Xyl (B) with 50.4 nM and 75.7 nM of ClAgl29A. Hydrolysis of *p*NP α-L-Qui (C) and *p*NP α-L-Xyl (D) with 37.2 nM and 62.0 nM of ClAgl29B.

Table 3-2. Kinetic parameters of ClAgl29A and ClAgl29B

	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m (s^{-1} mM^{-1})
ClAgl29A			
*p*NP α-L-Qui	0.268 ± 0.015	1.97 ± 0.09	7.35 ± 0.47
*p*NP α-L-Xyl	1.63 ± 0.29	0.452 ± 0.048	0.282 ± 0.048
*p*NP α-L-Glc [a]	2.01 ± 0.26	2.00 ± 0.11	1.00 ± 0.08
*p*NP α-L-Fuc [a]	0.279 ± 0.005	0.0658 ± 0.0008	0.236 ± 0.005
ClAgl29B			
*p*NP α-L-Qui	0.417 ± 0.029	4.59 ± 0.37	12.2 ± 0.7
*p*NP α-L-Xyl	0.587 ± 0.078	0.452 ± 0.0.018	0.781 ± 0.135
*p*NP α-L-Glc [a]	1.79 ± 0.18	7.76 ± 0.41	4.34 ± 0.22
*p*NP α-L-Fuc [a]	0.240 ± 0.011	0.0844 ± 0.0019	0.352 ± 0.007

a) Data was from Shishiuchi et al. (2022)

Table 3-3. Calculated Affinity for the binding of methyl α-L-glycosides to ClAgl29B in the GNINA program

	Affinity (kcal/mol)
Me α-L-Glc	−6.45
Me α-L-Qui	−6.29
Me α-L-Xyl	−5.55
Me α-L-Rha	−4.90

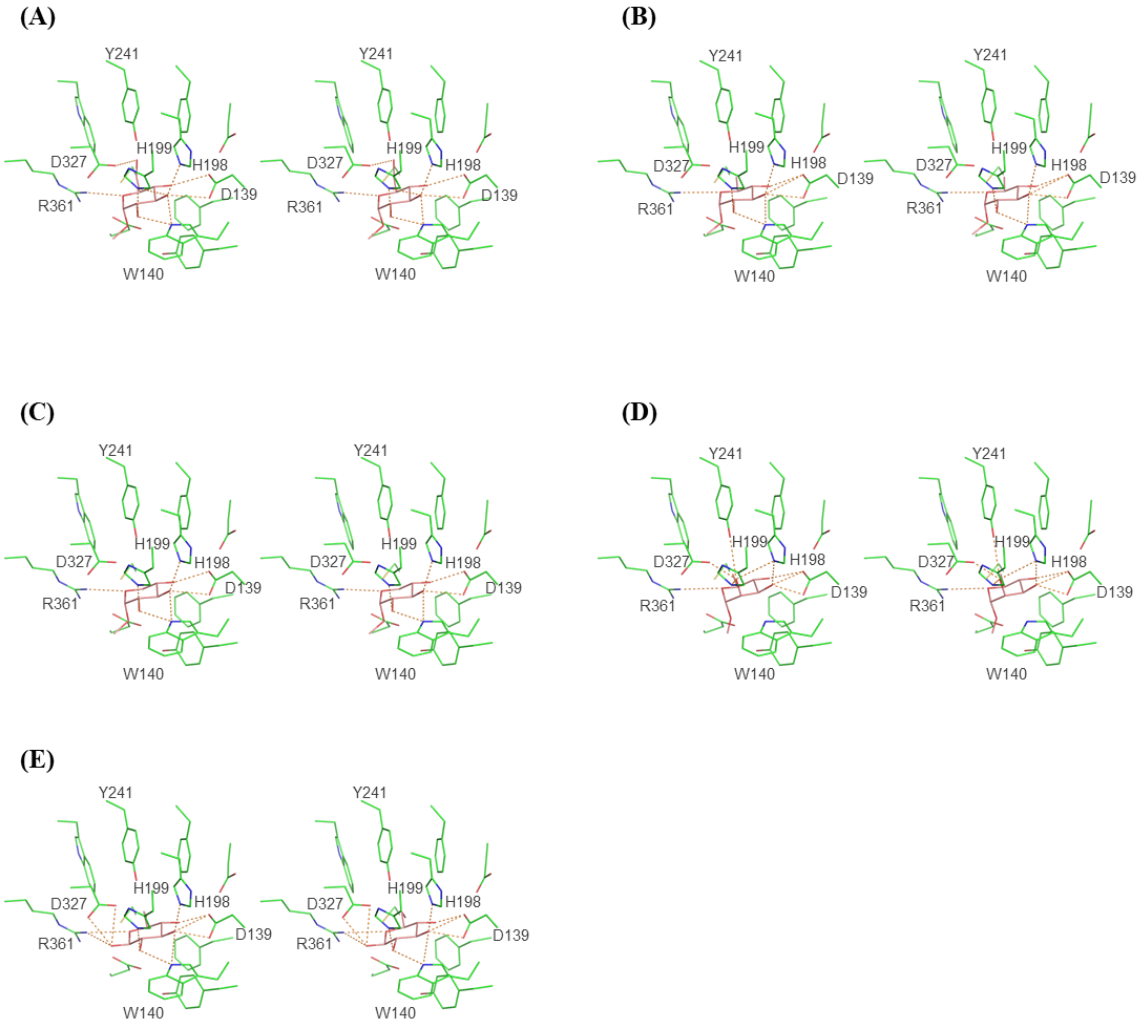


Fig. 3-4. Docked Me α-L-glycosides to ClAgl29B.

A, Me α-L-Glc; B, Me α-L-Qui; C, Me α-L-Xyl; D, Me α-L-Rha; E, β-L-Glc binding in the crystal structure of ClAgl29B (PDB accession, 7XSH). Amino acid residues within 4 Å of the ligand are shown, and predicted hydrogen bonds are indicated by dashed lines.

Table 3-4. Atoms within hydrogen bonding distance of ligand by docking simulation analysis

Residue atom	Ligand atom	Residue atom	Ligand atom
β-L-Glc (7XSH) [a]		**Me α-L-Xyl**	
Asp139 Oδ1	O3, O4	Asp139 Oδ1	O3
Asp139 Oδ2	O4	Asp139 Oδ2	O4
Trp140 Nε1	O2, O3	Trp140 Nε1	O2, O3
His198 Nε2	O3	His198 Nε2	O3
His199 Nε2	O2	His199 Nε2	O2
Asp327 (Nu) [b] Oδ1	O1	Arg361 Nη1	O5
Asp327 (Nu) [b] Oδ2	O1		
Arg361 Nη1	O1, O5		
Cys418 Sγ	O6		
Me α-L-Glc		**Me α-L-Rha**	
Asp139 Oδ1	O3	Asp139 Oδ1	O3
Asp139 Oδ2	O4	Asp139 Oδ2	O3, O4
Trp140 Nε1	O2, O3	His198 Nε2	O2, O3
His198 Nε2	O3	His199 Nε2	O2
His199 Nε2	O2, O3	Tyr241 Oη1	O2
Asp327 (Nu) [b] Oδ2	O6	Asp327 (Nu) [b] Oδ2	O2
Arg361 Nη1	O5	Arg361 Nη1	O5
Me α-L-Qui			
Asp139 Oδ1	O3		
Asp139 Oδ2	O3, O4		
Trp140 Nε1	O2, O3		
His198 Nε2	O3		
His199 Nε2	O2		
Arg361 Nη1	O5		

a) Results obtained by X-ray crystallography.

b) Asp327 provides the catalytic nucleophile (Nu).

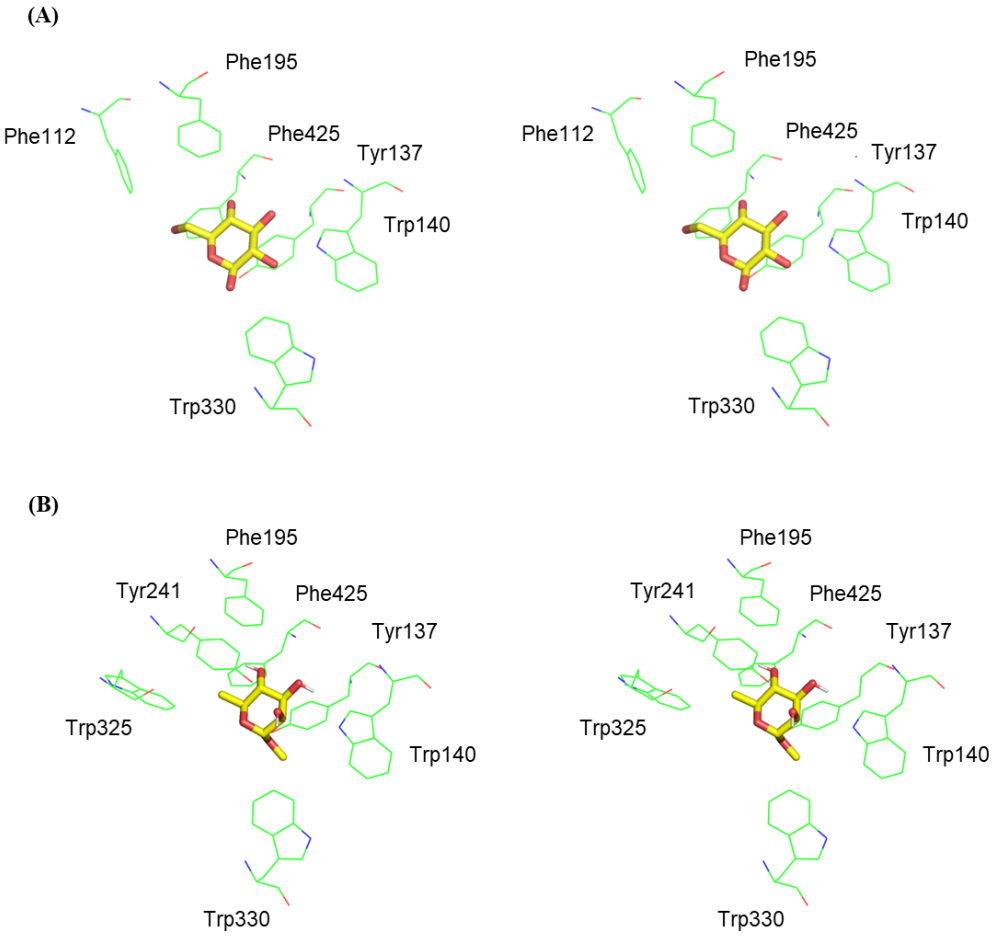


Fig. 3-5. Non-polar interaction of ClAgl29B with L-glucose (A) and Me α-L-Rha (B) in the active-site pocket. The non-polar interaction of L-glucose binding ClAgl29B (A) and the Me α-L-Rha binding ClAgl29B (B) is shown. Color coding is as follows; blue, N atom; red, O atom; yellow, C atom of L-glucose (A) and Me α-L-Rha (B); green, C atom of ClAgl29B.

4. Discussion

ClAgl29A and ClAgl29B are homologs of α-L-fucosidase, but they show hydrolytic activity toward α-L-glucoside rather than α-L-fucoside. On the other hand, the presence of α-L-glucoside has not yet been proven, leaving the possibility that these enzymes have substrate selectivity for α-L-glycosides other than α-L-glucoside. In this chapter, *p*NP derivatives of α-L-rhamnoside, α-L-quinovoside, and α-L-xyloside, which are structurally similar to α-L-glucoside, were chemically synthesized, and their substrate specificities were evaluated. ClAgl29A and ClAgl29B can hydrolyze *p*NP α-L-Qui and *p*NP α-L-Xyl, but they have little to no ability to hydrolyze *p*NP α-L-Rha (Table 3-1). These enzymes also exhibit low substrate specificity for α-L-fucoside (Shishiuchi, R., 2022). This means that ClAgl29A and ClAgl29B are unique enzymes that act on rare glycosides, such as α-L-glucoside, α-L-quinovoside, and α-L-xyloside, and do not act on α-L-rhamnoside and α-L-fucoside, which are commonly found in nature. ClAgl29A and ClAgl29B exhibited higher substrate specificity for *p*NP α-L-Qui than *p*NP α-L-Glc. Unlike L-glucose, L-quinovose has been believed to be present in the form of α-L-quinovosyl in the *O*-antigens of the pathogenic bacteria, including *P. stuartii* O44 (Ovchinnikova, O. G., 2013) and *Y. pseudotuberculosis* (Gorshkova, R. P., 1983) (Fig. 3-6). The structure of the *O*-antigens of *P. stuartii* and *Y. pseudotuberculosis* has been studied because of their pathogenicity. Hence, it could be possible that the lipopolysaccharides of *C. lonarensis*, of which the genome encodes ClAgl29A and ClAgl29B, and other Gram-negative bacteria also contain an α-L-quinovosyl moiety and ClAgl29A and ClAgl29B are involved its metabolism. In the metabolism of polysaccharides in bacteria, a group of enzymes involved in the metabolism is often encoded in the genome adjacent to each other, i.e., polysaccharide utilization loci (PUL) can be found in the genome. Around the ClAgl29A and ClAgl29B genes, which are encoded in tandem in the genome, genes that may be involved in monosaccharide uptake and monosaccharide metabolism are encoded, but not glycosidases that may be involved in the hydrolysis of polysaccharide.

ClAgl29A and ClAgl29B exhibited low activity toward *p*NP α-L-Rha while high activity toward *p*NP α-L-Qui, indicating that the enzymes strictly excluded the axial O2. The molecular docking simulations were performed to understand the low activity toward *p*NP α-L-Rha. The simpler Me α-L-Rha was used for the simulation. The calculated affinity of Me α-L-Rha for ClAgl29B was lower than that of Me α-L-Glc or Me α-L-Qui (Table 3-3), suggesting that it is less likely to bind to the enzyme. Even if Me α-L-Rha does bind to the enzyme, the O2 of Me α-L-Rha is close enough to form a hydrogen bond with the carboxy group of Asp327, which acts as a nucleophilic catalyst, and this may prevent the carboxy group from acting as the catalytic nucleophile (Fig 3-4 D). It is also suggested that the hydrogen bonding of the O2 of Me α-L-Rha to Oη of Tyr241, hydrogen bonded to the catalytic carboxy group of Asp327, may harm the catalytic activity.

Docking simulations of Me α-L-Glc and Me α-L-Qui on ClAgl29B showed that these ligands bind

to the enzyme similarly (Fig. 3-4). The calculated affinity indicated that Me α-L-Glc binds more strongly than Me α-L-Qui. On the other hand, the K_m of *p*NP α-L-Qui of ClAgl29B was smaller than that of *p*NP α-L-Glc, which resulted in a k_{cat}/K_m greater for *p*NP α-L-Qui than for *p*NP α-L-Glc. The X-ray crystallographic and docking studies showed that the C6 of α-L-glucoside and α-L-quinovoside should be in a hydrophobic environment surrounded by Phe112, Phe195, Trp325, and Phe425 (Fig. 3-7). The O6 of α-L-glucoside poorly suited to this environment will likely cause the higher K_m value of *p*NP α-L-Glc. Docking simulations could not evaluate entropic interactions, and a hydrogen bond between the O6 of Me α-L-Glc and the side chain of Asp327, suggesting that the affinity of Me α-L-Glc may be highly valued. In addition, the small k_{cat} on *p*NP α-L-Xyl could be due to the lack of binding free energy from hydrophobic interactions and van der Waals attractions by Phe112, Phe195, Trp325, and Phe425. In the crystal structure, the hydrogen bond between the O6 of β-L-Glc and the Asp327 was not seen (Fig. 3-7). The O6 formed a hydrogen bond with the Sγ of Cys418, and the Asp327 side chain formed a hydrogen bond with the O1 of β-L-Glc. Since the Asp327 side chain is separated from the O1 of α-L-glucoside, this side chain is probably hydrogen bonded to O6 of α-L-glucoside in the complex of the enzyme and α-L-glucoside. Hydrogen bonding to the Asp327 side chain, which acts as a nucleophilic catalyst, and the O6 of substrate would not have a significant effect on catalytic activity since there is no significant difference in k_{cat} between *p*NP α-L-Qui and *p*NP α-L-Glc. Since ClAgl29A has the same α-L-glucosyl binding site as ClAgl29B, these behaviors should also be true for ClAgl29A.

This study showed that ClAgl29A and ClAgl29B act on α-L-quinovosides rather than α-L-glucosides and structurally prefer α-L-quinovosides. These may refute the possibility of natural α-L-glucoside presence. Nevertheless, homologs of these enzymes that act more on α-L-glucoside may exist, and further study of α-L-glucosidase is warranted.

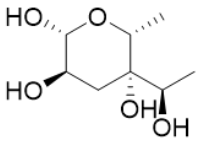
(A)

β-D-Gal*p*A
1
↓
4
[3)-α-L-Fuc*p*-(1→3)-α-D-Glc*p*-(1→4)-α-L-Qui*p*-(1→3)-α-D-Glc*p*NAc-(1→4)-α-D-Gal*p*NAc-(1→]$_n$

(B)

β-D-YerA
1
↓
4
[3)-α-D-Gal-(1→4)-α-L-Qui-(1→3)-β-D-GlcNAc-(1→]$_n$

(C)



yerisiniose A
(D-YerA)

Fig. 3-6. The structures of the O-antigen repeating units.

A, the repeating unit of *P. stuartii* O-antigen. B, the repeating unit of *Y. pseudotuberculosis* O-antigen. C, the structure of yerisiniose A (D-YerA).

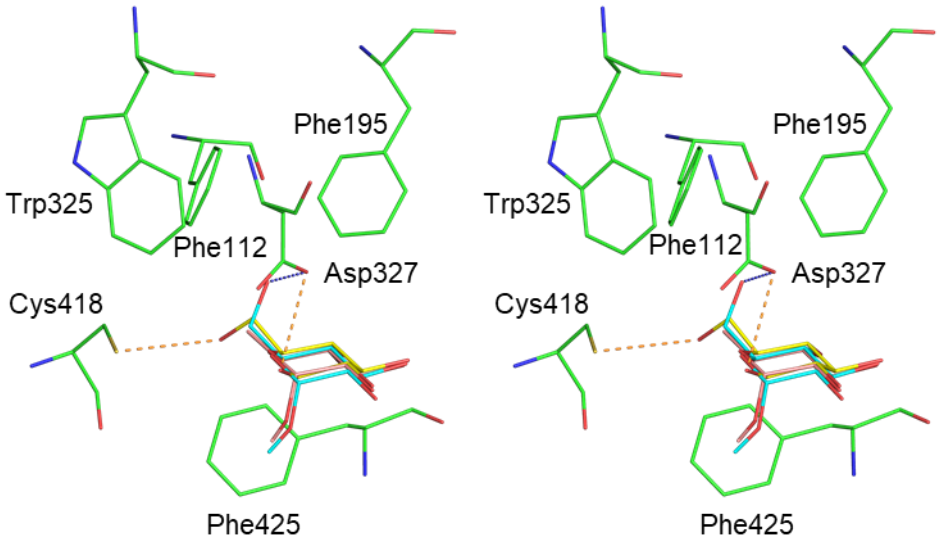


Fig. 3-7. Superimposition of Me α-L-Glc and Me α-L-Qui docked on ClAgl29B, and β-L-Glc in the crystal structure (7XSH).

Amino acid residues around the C6 and the O6 are shown. The ligands are color-coded cyan for Me α-L-Glc, wheat for Me α-L-Qui, and yellow for β-L-Glc. Hydrogen bonds between O6 of Me α-L-Glc and ClAgl29B are shown in blue, and hydrogen bonds between O6 of β-L-Glc and the enzyme in orange.

Chapter 4. Elucidation of key structural factors that distinguish between hydrolysis and transglycosylation in ClAgl29B

1. Introduction

Retaining glycosidases, including α-L-glucosidases, catalyze the hydrolysis of the glycosidic linkage and transglycosylation, in which the glycosyl residue transfers to an appropriate acceptor molecule. *C. lonarensis* α-L-glucosidases, ClAgl29A and ClAgl29B can catalyze an α-L-transglucosylation using *p*NP α-L-Glc as substrate and produce *p*-nitrophenyl α-L-glucopyranosyl-(1→3)-α-L-glucopyranoside (*p*NP 3-*O*-α-L-Glc α-L-Glc), *p*-nitrophenyl α-L-glucopyranosyl-(1→2)-α-L-glucopyranoside (*p*NP 2-*O*-α-L-Glc α-L-Glc), and *p*-nitrophenyl α-L-glucopyranosyl-(1→6)-α-L-glucopyranoside (*p*NP 6-*O*-α-L-Glc α-L-Glc) (Shishiuchi, R., 2022) (Fig. 4-1). Although ClAgl29A and ClAgl29B have the identical α-L-glucosyl binding site, ClAgl29A had significantly higher transglucosylation activity than ClAgl29B. Using 30 mM *p*NP α-L-Glc as substrate, 6.9 mM of transfer products accumulated with ClAgl29A compared to 2.3 mM with ClAgl29B. In addition, differences in the type of product were also observed between the two enzymes. Both enzymes produced *p*NP 3-*O*-α-L-Glc α-L-Glc as the main product, but ClAgl29A produced about twice as much *p*NP 2-*O*-α-L-Glc α-L-Glc as *p*NP 6-*O*-α-L-Glc α-L-Glc, whereas ClAgl29B produced the same amount of *p*NP 2-*O*-α-L-Glc α-L-Glc and *p*NP 6-*O*-α-L-Glc α-L-Glc.

The difference in transglycosylation between ClAgl29A and ClAgl29B may be due to the substitution of these two amino acid residues at the subsite +1. In the transglycosylation of glycosidases, a glycosyl-enzyme intermediates and an acceptor molecule react to form a new glycosidic linkage. In this reaction, the acceptor molecule is considered to bind to subsite +1 (Fig. 4-2), and the difference in affinity between the enzyme and the acceptor due to the difference in the structure of subsite +1 is thought to affect the property of the transglycosylation. Thus, the properties of the mutant enzymes of ClAgl29B, S132D, Q275L, and S132D/Q275L were evaluated, introducing point mutation(s) into ClAgl29B to mimic ClAgl29A.

ClAgl29A and ClAgl29B are the first enzymes to catalyze α-L-glucosylation, and it would be valuable to elucidate the structural factors that govern the transglucosylation properties. In this chapter, ClAgl29B mutant enzymes were prepared and evaluated for their expected increase in the α-L-transglucosylation to understand their structural factors.

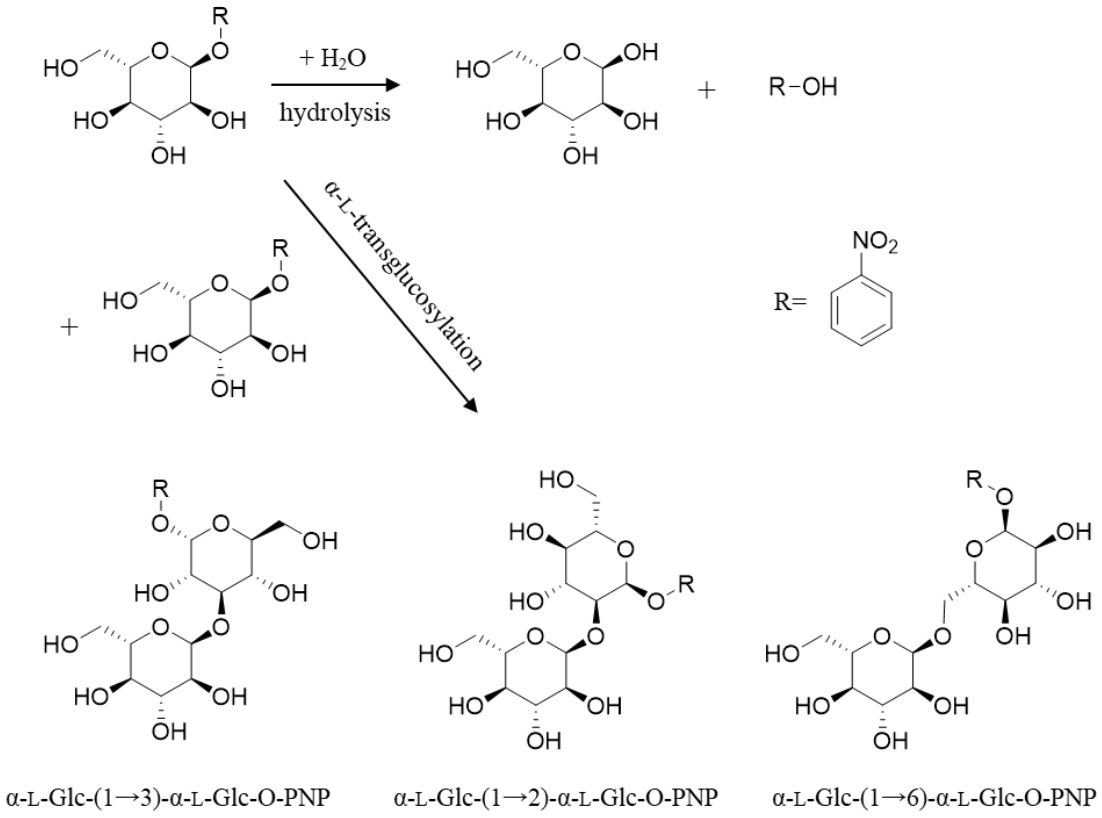


Fig. 4-1. Hydrolysis and transglycosylation reactions catalyzed by ClAgl29A and ClAgl29B.

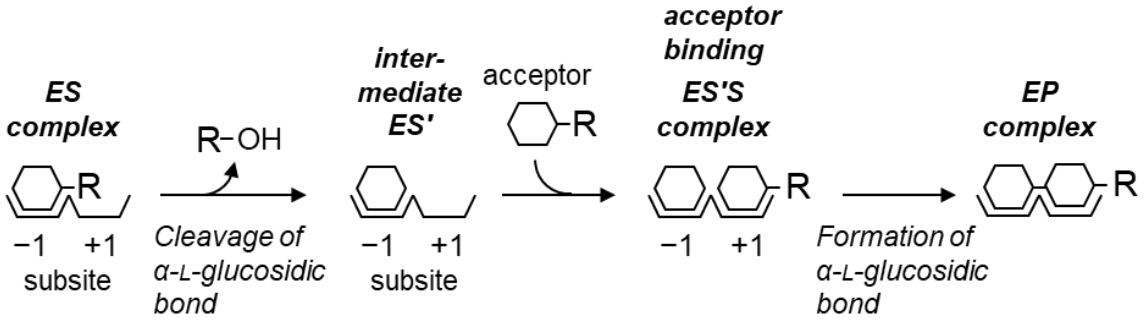


Fig. 4-2. Schematic representation of transglycosylation catalyzed by a retaining glycosidase.

2. Materials and methods

2.1 Materials

*p*NP α-L-Fuc was purchased from Sigma-Aldrich (St. Louis, MA). *p*NP α-L-Glc used was chemically synthesized from L-glucose in our laboratory. All other chemicals used were analytical grade.

2.2 Construction of the plasmid DNAs for production of S132D, Q275L, and S132D/Q275L

The base substitution was introduced into the DNA based on the method in PrimSTAR Mutagenesis Basal Kit (TAKARA Bio, Japan). The ClAgl29B/pET28a plasmid DNA (100 pg/μL), for the production of the wild-type ClAgl29B (Shishiuchi, R., 2022), was used as a template to introduce a mutation. The ClAgl29B/pET28a plasmid DNA (100 pg/μL) to produce the wild-type ClAgl29B (Shishiuchi, R., 2022) was used as a template to introduce a mutation. The primers used for PCR amplification are listed in Table 4-1. The amplification reaction was carried out with the following amplification conditions: After denaturation at 94°C for 2 min, temperature steps of 98°C for 10 s, 55°C for 15 s, and 72°C for 1 min were performed 30 times. The amplified DNA was chemically introduced into *E. coli* DH5α and transformants were selected on LB agar medium containing kanamycin. Plasmid DNA propagated in *E. coli* was extracted and purified by alkaline SDS and PEG/NaCl. DNA sequences were analyzed according to Chapter 2, Section 2.4 to confirm the introduction of mutations and the absence of other unexpected mutations.

2.3 Production and purification of ClAgl29B mutant enzymes

E. coli Rosetta (DE3) was transformed with the prepared ClAgl29B/pET28a derivatives having appropriate mutation(s). Transformants were grown in LB medium supplemented with 50 μg/mL kanamycin and 30 μg/mL chloramphenicol at 18°C, then 0.1 mM IPTG was added, and recombinant protein expression was induced by incubation at 18°C overnight. The cells were harvested by centrifuging at 8,000 × *g* at 4°C for 15 min. The cells were lysed in 20 mM sodium phosphate buffer (pH 7.5) containing 300 mM NaCl by SONIFIER 250 ultrasonication (output 2 and duty cycle 40). The debris was removed by centrifugation at 5,800 × *g* at 4°C for 15 min and the supernatant obtained was subjected to the Ni-affinity chromatography.

Five mL of Chelating Sepharose Fast Flow (Cytiva) equilibrated with 20 mM sodium phosphate buffer (pH 7.5) / 300 mM NaCl (buffer A) containing 10 mM imidazole was packed into a glass column with a diameter of 1.5 cm. The supernatant was loaded onto the column and unbound and weakly bound proteins were washed by buffer A containing 30 mM imidazole. The bound proteins were eluted with a linear gradient of 30-300 mM imidazole in buffer A. Active fractions were collected and dialyzed against 10 mM sodium phosphate buffer (pH 7.5). Amicon® Ultra-15

Centrifugal Filter Unit (NMWL (normal molecular weight limit) 30,000, Merck Millipore) was used for concentrating the proteins. All purification procedures were carried out at 4°C.

2.4 Protein concentration

Protein concentrations were estimated based on the protein hydrolysate (6 M hydrochloric acid, 110°C, 24 h) using an L-8900 amino acid analyzer (Hitachi High-Tech, Tokyo, Japan) equipped with a ninhydrin-detection system. The molar concentration of each protein was calculated from the average of each amino acid concentration quantified. The specific extinction coefficient ($E_{280\ nm,1cm}^{1\%}$) were 25.3 for S132D, 24.8 for Q275L, and 24.0 for S132D/Q275L. For conversion from mol to gram, the molar mass (g/mol) values calculated with the ProtParam tool in ExPASY (https://web.expasy.org/protparam/), 67,146 (S132D), 67,103 (Q275L), and 67,131 (S132D/Q275L), were used.

2.5 Enzyme assay

A mixture of 20 μL of 100 mM sodium acetate buffer (pH 5.5) and 20 μL of 5 mM *p*NP α-L-Glc was incubated for 3 min at 35°C, then 10 μL of the appropriately diluted enzyme was added, and allowed to react for 10 min at 35°C. The enzyme was diluted with 40 mM sodium acetate buffer (pH 5.5) containing 0.5 mg/mL bovine serum albumin (BSA). The reaction was stopped by mixing with 2 volumes of 1 M sodium carbonate. The concentration of *p*NP released was determined by measuring absorbance at 400 nm. The molar absorbance coefficient was 5,560 $M^{-1}cm^{-1}$. One α-L-glucosidase activity unit was defined as the amount of enzyme capable of releasing 1 μmol of *p*NP per minute.

2.6 Biochemical assays

2.6.1 Optimum pH and pH stability

For the determination of the optimum pH for hydrolysis of 2 mM *p*NP α-L-Glc, Britton-Robinson buffer (from pH 2.5 to 11.5) and 40 mM glycine-NaOH buffer (from pH 11.5 to 12.0) was used, the reaction temperature was set at 35°C, and the reaction time was 10 min. Britton-Robinson buffer consisted of a mixture of 40 mM phosphoric acid, 40 mM acetic acid, and 40 mM boric acid and was titrated to the desired pH with 200 mM sodium hydroxide. The enzyme concentrations used were 0.07 μM for S132D, 0.10 μM for Q275L, and 0.124 μM for S132D/Q275L. The enzyme was diluted with 40 mM sodium acetate buffer (pH 5.5) containing 0.5 mg/mL BSA. Measurements of *p*NP concentration released were made according to the methods described in 2.5 of this chapter.

The pH stability range of the enzyme was determined based on the residual activity of the

enzymes (1.07 μM for S132D, 1.52 μM for Q275L, and 0.8 μM for S132D/Q275L) previously kept in Britton-Robinson buffer (from pH 2.5 to 11.5) and 100 mM glycine-NaOH (from pH 11.5 to 12.0) containing 0.5 mg/mL BSA at 4°C for 24 h. The residual activity was measured according to the methods described in 2.5 of this chapter.

2.6.2 Optimum temperature and temperature stability

To evaluate the optimum reaction temperature, following the method described in 2.5 of this chapter, but temperatures were set to from 30°C to 70°C, the reaction rates were measured. The enzyme reaction for the optimal temperature was examined in various temperatures (30°C to 70°C). However, sodium acetate buffer pH 6.0 was used for Q275L and sodium acetate buffer pH 6.5 for S132D and S132D/Q275L. The enzyme concentrations used were 0.036 μM for S132D, 0.061 μM for Q275L, and 0.062 μM for S132D/Q275L.

To evaluate temperature stability range, the enzymes diluted in 100 mM sodium acetate buffer, pH 6.0 for Q275L and pH 6.5 for S132D and S132D/Q275L, containing 0.5 mg/mL BSA were kept at each temperature varying from 30°C to 70°C for 15 min, then cooled on ice, and the residual activity was according to the methods described in 2.5 of this chapter. Enzyme concentration was set as 0.35 μM of S132D, 0.5 μM of Q275L, and 0.62 μM of S132D/Q275L.

2.6.3 Kinetic study of ClAgl29B mutant enzymes

The initial velocities for hydrolysis of various concentration of *p*NP α-L-Glc (0.25, 0.5, 1.0, 1.5, 2.0, 3.0, 4.0, 5.0 mM) and those of *p*NP α-L-Fuc (0.05, 0.1, 0.2, 0.3, 0.5, 0.75, 1.0, 1.25, 1.5 mM) were measured under standard conditions, but using sodium acetate buffer, pH 6.0, for Q275L and pH 6.5, for S132D and S132D/Q275L. Concentrations of Q275L used were 12 nM for *p*NP α-L-Glc and 710 nM for *p*NP α-L-Fuc, those of S132D 15 nM for *p*NP α-L-Glc and 101 μM for *p*NP α-L-Fuc, S132D/Q275L 16 nM for *p*NP α-L-Glc and 800 μM for *p*NP α-L-Fuc. Kinetic parameters, k_{cat} (sec^{-1}) and K_m (mM), were obtained by fitting data from the reactions to the Michaelis–Menten equation using the computer program KaleidaGraph v. 3.6 (Synergy Software; Reading, PA).

2.7 Analysis of transglucosylation products by thin-layer chromatography (TLC)

*p*NP α-L-Glc (37.5 mM) in 50 mM sodium acetate buffer (pH 6.0 for Q275L and pH 6.5 for S132D and S132D/Q275L) (80 μL) was incubated at 35°C for 3 min, then 20 μL of the enzyme solution diluted with 40 mM sodium acetate (pH 6.0 and 6.5) containing 0.5 mg/mL BSA was added, and continued to keep at 35°C for 24 h. The final concentration of each enzyme in the reaction mixture was 0.266 μM for S132D, 0.275 μM for Q275L, and 0.266 μM for S132D/Q275L. At the indicated time, a part of the reaction mixture was taken and heated at 100°C for 3 min to terminate the reaction.

One μL of each sample solution was spotted onto a TLC plate (Silica gel 60 F254, Merck). One μL each of 1 mM *p*NP α-L-Glc and 1 mM L-Glc as standard sugar was spotted similarly. A mixture of ethyl acetate/methanol/water (70/20/10) was used as the developing solvent. The carbohydrates were visualized by spraying a reagent consisting of 0.03% (w/v) α-naphthol and 5% (v/v) sulfuric acid dissolved in 95% (v/v) methanol and then heating.

2.8 Analysis of transglucosylation rate

*p*NP α-L-Glc (37.5 mM) in 50 mM sodium acetate buffer (pH 6.0 for Q275L and pH 6.5 for S132D and S132D/Q275L) (80 μL) was incubated at 35°C for 3 min, then 20 μL of the enzyme solution diluted with 40 mM sodium acetate (pH 6.0 and 6.5) containing 0.5 mg/mL BSA was added, and kept at 35°C. Twenty μL of the reaction solution was taken at 3, 6, 9, and 12 min and treated at 100°C for 2 min to stop the reaction. The final concentration of each enzyme in the reaction mixture was 0.107 μM for S132D, 0.202 μM for Q275L, and 0.266 μM for S132D/Q275L. The collected reaction solution was diluted 10 times with deionized water.

To quantify the amount of released *p*NP, the absorbance at 400 nm of the diluted reaction solution was mixed with twice the amount of 1 M sodium carbonate. The *p*NP concentration was estimated from the molar absorbance coefficient of the *p*NP solution, 5,560 $M^{-1}cm^{-1}$. The *p*NP release rate was calculated from a regression line obtained from a plot with reaction time on the horizontal axis and *p*NP concentration on the vertical axis, and this rate was defined as the total reaction rate (v_{total}), including the hydrolysis and transglucosylation reaction rates.

The concentration of L-Glc was determined by a high-performance anion-exchange chromatography-pulsed amperometric detection (Nanospace SI-2/3016; Osaka soda) system (HPAEC-PAD) equipped with a CarboPac PA1 analytical column (Φ4 × 250 mm, Dionex) in tandem with a CarboPac PA1 guard column (Φ4 × 50 mm, Dionex) and an auto-sampler (Nanospace SI-2/3023; Osaka soda). To the diluted reaction sample, the one-ninth volume of 100 μM sorbitol was added as an internal standard and then transferred to an analytical tube (silicone/PTFE septum-topped cap; GL Science). The volume of sample injected into the column was 20 μL. The sample was separated by isocratic elution using 160 mM sodium hydroxide solution, which was prepared from super special grade 50% sodium hydroxide solution (Fujifilm Wako Pure Chemical Corporation). The flow rate was set at 0.8 mL/min. L-Glucose from 10 μM to 400 μM was used to generate the standard curve. S-MicroChrom Version 5.02 (SMC21; Osaka soda) was used for data analysis. The L-glucose release rate was calculated from a regression line obtained from a plot with reaction time on the horizontal axis and L-glucose concentration on the vertical axis, and this rate was defined as the hydrolysis rate (v_{hydr}). The transglucosylation rate (v_{tg}) was calculated as the difference between the total reaction rate and the hydrolysis rate.

$v_{tg} = v_{total} - v_{hydr}$

2.9 Analysis of transglucosylation products by reverse-phase high-performance liquid chromatography (RP-HPLC)

*p*NP α-L-Glc (37.5 mM) in 50 mM sodium acetate buffer (pH 6.0 for Q275L and pH 6.5 for S132D and S132D/Q275L) (280 μL) was incubated at 35°C for 3 min, then 70 μL of the enzyme solution diluted with 40 mM sodium acetate (pH 6.0 and 6.5) containing 0.5 mg/mL BSA was added, and continued to keep at 35°C by 48 h. The final concentration of each enzyme in the reaction mixture was 0.28 μM for S132D, 2.0 μM for Q275L, and 2.0 μM for S132D/Q275L. At the indicated time, 30 μL of the reaction mixture was taken and heated at 100°C for 3 min to terminate the reaction. After the reaction was stopped by heat treatment, the sample was diluted 20-fold with deionized water and desalted using Amberlite MB-4 (Organo). Analysis of transglucosylation products was performed by RP-HPLC using a Cosmosil $5C_{18}$-PAQ analytical column (4.6 ID x 150 mm; Nacalai Tesque, Kyoto, Japan). Twenty μL of the sample was injected into the column. The column temperature was set at 40°C and the mobile phase was methanol and water (20/80, v/v) at a flow rate of 1.0 mL/min on a Jasco system (Tokyo, Japan). *p*NP glucosides were detected at 312 nm, and their concentration was calculated from a calibration curve prepared using 50 – 2000 μM *p*NP α-D-Glc.

2.10 Docking simulation study

Docking simulations were performed using GNINA, and the receptor, S132D/Q275L, was created using the mutagenesis function of PyMOL. *p*NP 3-*O*-α-L-Glc α-L-Glc and *p*NP 6-*O*-α-L-Glc α-L-Glc used as ligands were drawn by Chemdraw and saved as structure-data file (.sdf).

Table 4-1. Oligonucleotide sequence used in Chapter 4

Name	Sequence (5′→3′)	Purpose
ClAglB_S132D_F	GAAGCAGATACTGGAGGATCCTACCCG	For mutation of S132D
ClAglB_S132D_R	TCCAGTATCTGCTTCAGCGCCTTTTAA	
ClAglB_Q275L_F	TTAGGTCTGATTCCAAGGGCAACGCCA	For mutation of Q275L
ClAglB_Q275L_R	TGGAATCAGACCTAAAGCGTCCTGGAA	
Cl29b_seq1	TGGTCGGTTCCAGGATATGC	Sequencing for ClAgl29B variants
Cl29b_seq2	TCAGTTTCTGAATGGGAATACCC	
Cl29b_seq3	GTCGCTCAATAAGTCCCGCT	
ClAglB_Seq_F	GAAATGGGGCCTGGAAGAGATA	
ClAglB_Seq_R	TTCTGGTGTACCATAATCCCCT	

3. Results

3.1 Comparison of substrate binding sites between ClAgl29A and ClAgl29B

Since the structure of subsite +1 was not determined directly from the three-dimensional structure of ClAgl29B with β-L-glucose bound to the α-L-glucosyl binding site, it was inferred from the superposition with the complex structure of TmaFuc with α-L-fucosyl-(1→2)-β-L-fucosyl azide (β-L-Fuc-N_3) (Cobucci-Ponzano, B., 2009; PDB ID, 2wsp) (Fig. 4-3). In this structural comparison, the amino acid residues located around β-L-Fuc-N_3 residue were inferred to be those involved in the formation of the α-glucosidase subsite +1. The subsite +1 of ClAgl29A was assumed to consist of Asp120, Thr121, Tyr125, Trp128, Leu263, Trp318, Arg349, and Glu391. That of ClAgl29B was closely similar, but two amino acid substitutions were found: ClAgl29B had Ser132 and Gln275 instead of Asp120 and Leu263, respectively.

3.2 Production and purification of ClAgl29B mutant enzymes

The N-terminal His_6-tagged recombinant proteins of S132D, Q275L, and S132D/Q275L were produced in *E. coli* Rosetta (DE3) transformants and purified by Ni^{2+} affinity column chromatography. SDS-PAGE showed a single band for each mutant enzyme (Fig. 4-4). In SDS-PAGE, the purified enzymes moved faster than their expected molecular weight, similar to the wild-type ClAgl29B and ClAgl29A. The estimated molecular weights are 67.146 for S132D, 67.103 for Q275L, and 67.131 for S132D/Q275. From 500 mL of culture, S132D with 1.1 U/mg of specific activity yielded 4.48 mg; Q275L with 0.53 U/mg yielded 6.11 mg; and S132D/Q275L with 0.93 U/mg yielded 2.68 mg. Compared to the specific activity of wild-type ClAgl29B, 3.2 U/mg, these specific activities decreased to 37, 17, and 29%, respectively. Meanwhile, they were close to that of ClAgl29A (0.69 U/mg).

3.3 Effect of pH and temperature on enzyme activity and stability

The optimum pH in hydrolysis of 2 mM *p*NP α-L-Glc was pH 6.0 for Q275L and pH 6.5 for S132D and S132D/Q275L (Fig. 4-5 A, E, and I), and these were slightly higher than that of wild-type ClAgl29B, whose optimum pH was 5.5. The pH stability of ClAgl29B mutant enzymes was examined using Britton-Robinson buffer and glycine-NaOH buffer at 4°C for 24 h. The pH stability of the mutant enzymes was evaluated from the residual activity after 24 hours at 4°C at each pH. The pH stable range was defined as the pH range in which 80% or more of the activity remained, and the pH stable range for S132D was from pH 6.5 to 9.5, for Q275L from pH 7.0 to 11.5, and for S132D/Q275L from pH 6.0 to 10.5 (Fig. 4-5 B, F, and J). Compared to the wild-type enzyme, the stability of the acidic side was decreased. The temperature at which the ClAgl29B mutant enzyme showed the maximum reaction rate in the hydrolysis reaction of 2 mM *p*NP α-L-Glc was 50°C for

every enzyme (Fig. 4-5 C, G, and K). The temperature range in which S132D and Q275L showed more than 80% residual activity after 15 min of holding at each temperature was less than 55°C and less than 50°C for S132D/Q275L. The double mutant enzymes were found to be more temperature-susceptible than their respective mutants (Fig. 4-5 D, H, and L).

3.4 Kinetic studies on hydrolysis of *p*NP α-L-fucoside and glucoside

Kinetic constants for hydrolysis of *p*NP α-L-Glc and *p*NP α-L-Fuc were determined. The initial reaction rates measured at various substrate concentrations were fit to the Michaelis–Menten equation (Fig. 4-6). The k_{cat} and K_m values obtained by nonlinear regression to the Michaelis–Menten equation and the k_{cat}/K_m values are summarized in Table 4-2. All mutant enzymes exhibited a higher k_{cat}/K_m for *p*NP α-L-Glc than for *p*NP α-L-Fuc, indicating their continued specificity for α-L-glucoside. The k_{cat} values for *p*NP α-L-Glc of each mutant enzyme were reduced compared to that of the wild-type enzyme. The K_m for *p*NP α-L-Glc of S132D was similar to that of the wild-type ClAgl29B, while those of Q275L and S132D/Q275L were smaller than that of ClAgl29B. The K_m for *p*NP α-L-Fuc of S132D was larger, while those of the other mutant enzymes were smaller than that of ClAgl29B.

3.5 Analysis of transglucosylation products by TLC

ClAgl29B mutant enzymes (18 μg/mL) were incubated with 30 mM *p*NP α-L-Glc at 35°C, and the reaction mixtures were analyzed by TLC over time. All enzymes produced four transglucosylation products: spots A, B, C, and D, in ascending order of mobility on TLC, with spot B producing the least amount, and spot D being observed late in the reaction (Fig. 4-7). S132D and S132D/Q275L produced roughly equal amounts of spot A and spot C, whereas Q275L produced more of spot A.

3.6 Analysis of an initial transglucosylation rate

The release rates of *p*NP (v_{total}) and L-glucose (v_{hydro}) when 30 mM *p*NP α-L-Glc was used as a substrate were determined (Fig. 4-8), and the transglucosylation rate (v_{tg}) was calculated (Table 4-3). The ratio of the transglucosylation reaction rate to the total reaction rate, the transglucosylation ratio, was calculated to be 80.9% for S132D, 90.2% for Q275L, and 87.4% for S132D/Q275L. Compared to the 32.0% transglucosylation raito of the wild-type ClAgl29B, the introduction of the mutations substantially increased the transglucosylation ratio.

3.7 Analysis of transglucosylation products by RP-HPLC

The changes in the synthesis of each transglucosylation product of the mutant enzymes over time were examined (Fig. 4-9; Fig. 4-10). The enzyme concentration of each mutant enzyme was set to

the amount at which the substrate, *p*NP α-L-Glc, was almost consumed after 48 h of reaction. Comparison with the elution position in RP-HPLC showed that the transglucosylation products of the respective mutant enzymes were unchanged from those of the wild-type enzyme, and were *p*NP 3-*O*-α-L-Glc α-L-Glc, *p*NP 2-*O*-α-L-Glc α-L-Glc, and *p*NP 6-*O*-α-L-Glc α-L-Glc (Shishiuchi, R., 2022). The proportion of each *p*NP disaccharide in the total products differed from that of the wild-type ClAgl29B and varied among the mutant enzymes. All mutant enzymes produced *p*NP 3-*O*-α-L-Glc α-L-Glc as the main product, as did the wild-type enzyme. Whereas the wild-type enzyme produced approximately the same amount of *p*NP 2-*O*-α-L-Glc α-L-Glc and *p*NP 6-*O*-α-L-Glc α-L-Glc, S132D produced an increased amount of *p*NP 6-*O*-α-L-Glc α-L-Glc, which was almost equivalent to *p*NP 3-*O*-α-L-Glc α-L-Glc. In Q275L, *p*NP 6-*O*-α-L-Glc α-L-Glc was produced in almost equal amounts as *p*NP 2-*O*-α-L-Glc α-L-Glc. In S132D/Q275L, *p*NP 6-*O*-α-L-Glc α-L-Glc was produced in increased amounts, but the amount produced was clearly lower than that of *p*NP 3-*O*-α-L-Glc α-L-Glc. In the reaction of Q275L, *p*NP 3-*O*-α-L-Glc α-L-Glc reached the maximum of 7.9 mM at 6 h and decreased significantly after 8 h. The reaction of S132D/Q275L also showed the decrease in the concentrations of *p*NP 3-*O*-α-L-Glc α-L-Glc after 6 h and its maximum was 4.79 mM at 4 h. Such a decrease was not observed in the wild-type ClAgl29B and S132D. In Q275L, the concentrations of *p*NP 6-*O*-α-L-Glc α-L-Glc and *p*NP 2-*O*-α-L-Glc α-L-Glc were 7.9 mM and 0.8 mM, respectively, at 6 h of reaction when the concentration of *p*NP 3-*O*-α-L-Glc α-L-Glc was maximal. The concentration of each product at 48 h was 2.25 mM for *p*NP 3-*O*-α-L-Glc α-L-Glc, 0.90 mM for *p*NP 6-*O*-α-L-Glc α-L-Glc, and 0.5 mM for *p*NP 2-*O*-α-L-Glc α-L-Glc. In S132D/Q275L, the concentrations of *p*NP 6-*O*-α-L-Glc α-L-Glc and *p*NP 2-*O*-α-L-Glc α-L-Glc were 4.6 mM and 1.6 mM, respectively, at 4 h of reaction when the concentration of *p*NP 3-*O*-α-L-Glc α-L-Glc was maximal. The concentration of each product at 48 h was 2.25 mM for *p*NP 3-*O*-α-L-Glc α-L-Glc, 0.90 mM for *p*NP 6-*O*-α-L-Glc α-L-Glc, and 0.5 mM for *p*NP 2-*O*-α-L-Glc α-L-Glc.

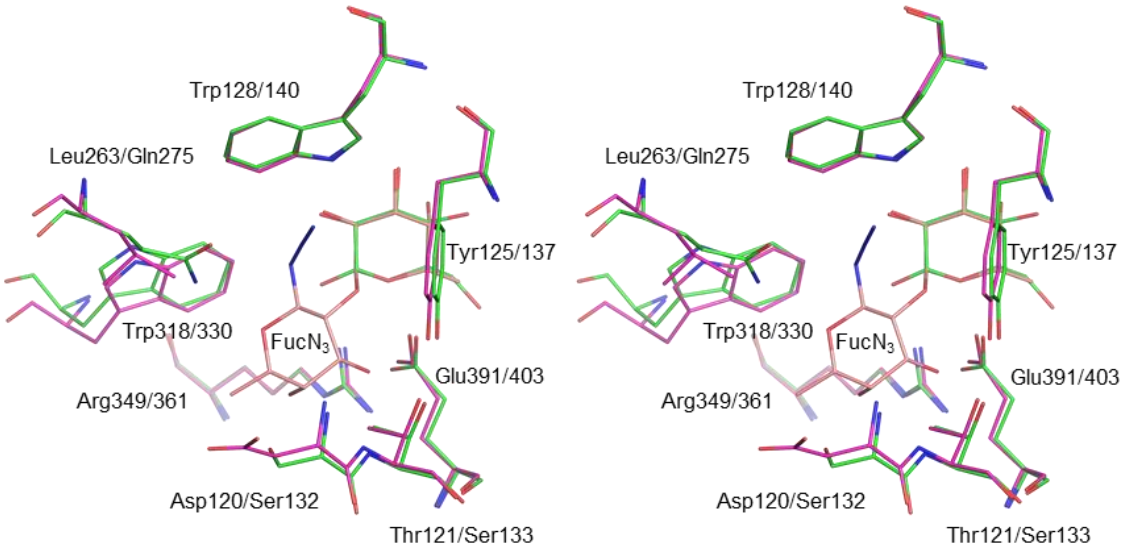


Fig. 4-3. Structural comparison of subsite +1 between ClAgl29A and ClAgl29B.

The structures of ClAgl29A (magenta) and ClAgl29B (green) superimposed on the structure of TmaFuc in complex with α-L-Fuc-(1-2)-β-L-Fuc-N_3 (pdb id 2wsp, wheat). Amino acid residues within 4 Å from β-L-Fuc-N_3 residue ($FucN_3$) are shown. Residue names are shown with ClAgl29A on the left and ClAgl29B on the right.

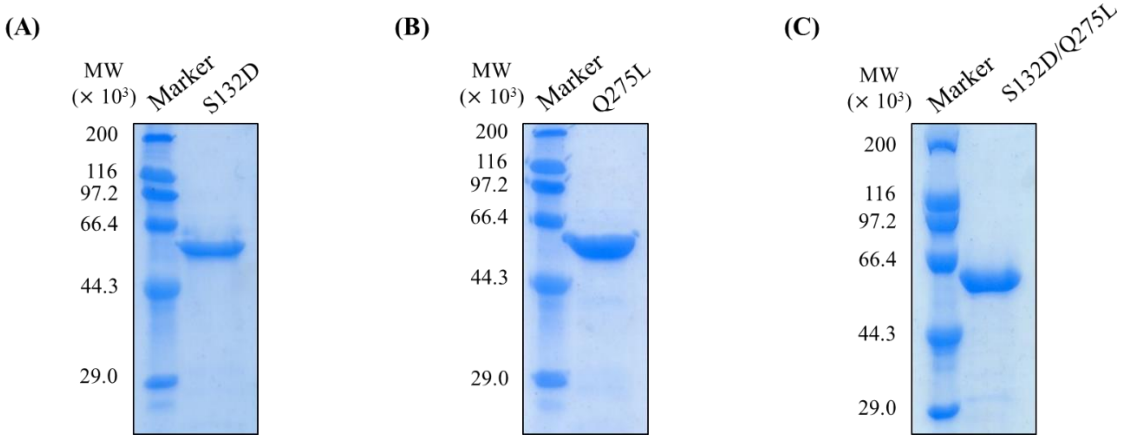
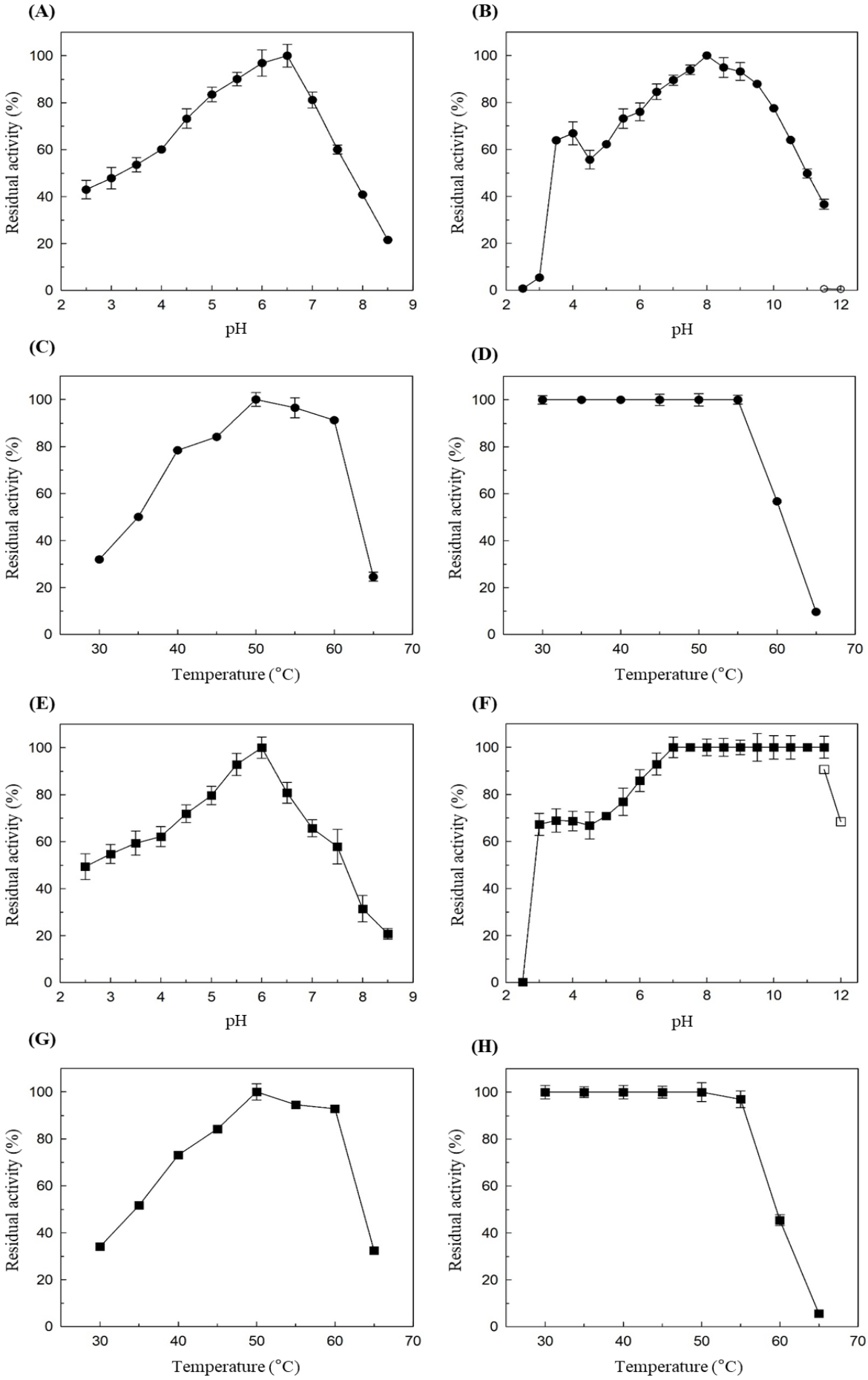


Fig. 4-4. SDS-PAGE of purified recombinant ClAgl29B mutant enzymes. Lane marker, size marker; lane S132D, ClAgl29B S132D purified by Ni-affinity chromatography (A_{280}=0.91, 1.3 μg); lane Q275L, ClAgl29B Q275L purified by Ni-affinity chromatography (A_{280}=2.49, 4.0 μg); lane S132D/Q275L, ClAgl29B S132D/Q275L purified by Ni-affinity chromatography (A_{280}=1.04, 1.7 μg). The molecular weight of the standards is shown on the left.


(A)
Residual activity (%)
pH
(B)
Residual activity (%)
pH
(C)
Residual activity (%)
Temperature (°C)
(D)
Residual activity (%)
Temperature (°C)
(E)
Residual activity (%)
pH
(F)
Residual activity (%)
pH
(G)
Residual activity (%)
Temperature (°C)
(H)
Residual activity (%)
Temperature (°C)

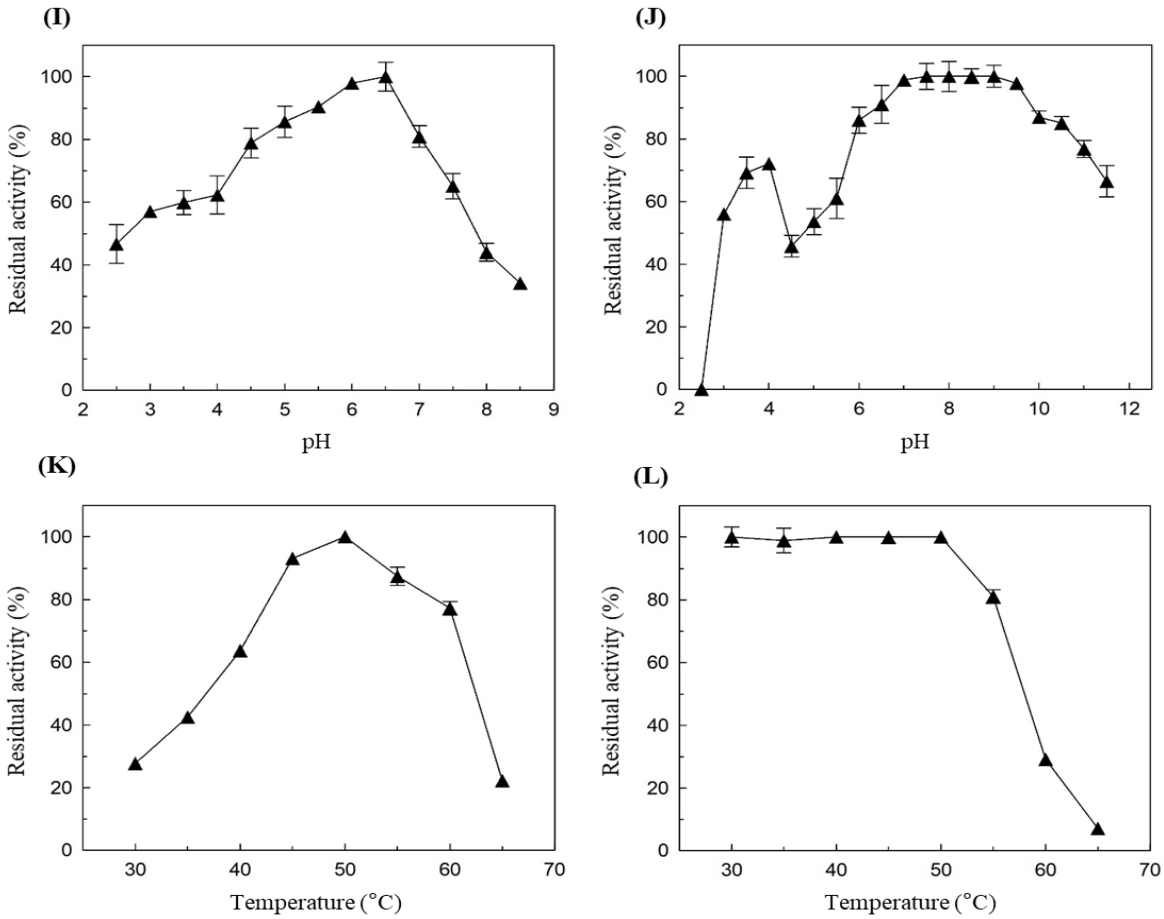


Fig. 4-5. Enzymatic properties. Circle (●); ClAgl29B S132D, Square (■); ClAgl29B Q275L, Triangle (▲); ClAgl29B S132D/Q275L. **A, E, and I,** pH activity curves. The hydrolytic velocity at various pH ranges at 35℃ was measured using the reaction mixture containing 2 mM *p*NP-α-L-Glc and enzyme (ClAgl29B S132D; 0.07 μM, Q275L; 0.10 μM, and S132D/Q275L; 0.124 μM in Britton-Robinson buffer (pH 3.0-8.0). The velocity at pH 6.0 for Q275L and pH 6.5 for S132D and S132D/Q275L was defined as 100%. **B, F, and J,** pH stability curves under conditions at 4℃ for 24 h. The enzymes (ClAgl29B S132D; 0.024 μM, Q275L; 0.051 μM, S132D/Q275L; 0.04 μM) were kept at at 4℃ for 24 h in a buffer (Britton-Robinson buffer (pH 2.5-11.5) and 100 mM glycine-NaOH buffer (pH 11.5-12) supplemented with 0.05% BSA. **C, G, and K,** Thermal activity curves. The hydrolytic velocity was measured at various temperature conditions (30-65℃) using the reaction mixture consisting of 2 mM *p*NP-α-L-Glc and enzyme (0.036 μM for ClAgl29B S132D, 0.061 μM for Q275L, and 0.062 μM for S132D/Q275L) in 40 mM sodium acetate buffer (pH 6.0 and 6.5). The velocity at 50℃ were defined as 100%. **D, H, and L,** Thermal stability curves. The enzymes (0.07 μM for ClAgl29B S132D, 0.1 μM for Q275L, and 0.124 μM for S132D/Q275L) were treated at various temperatures (30-70℃) for 15 min, and hydrolytic velocity was measured at 35℃. The hydrolytic velocity without heat treatment was defined as 100%.

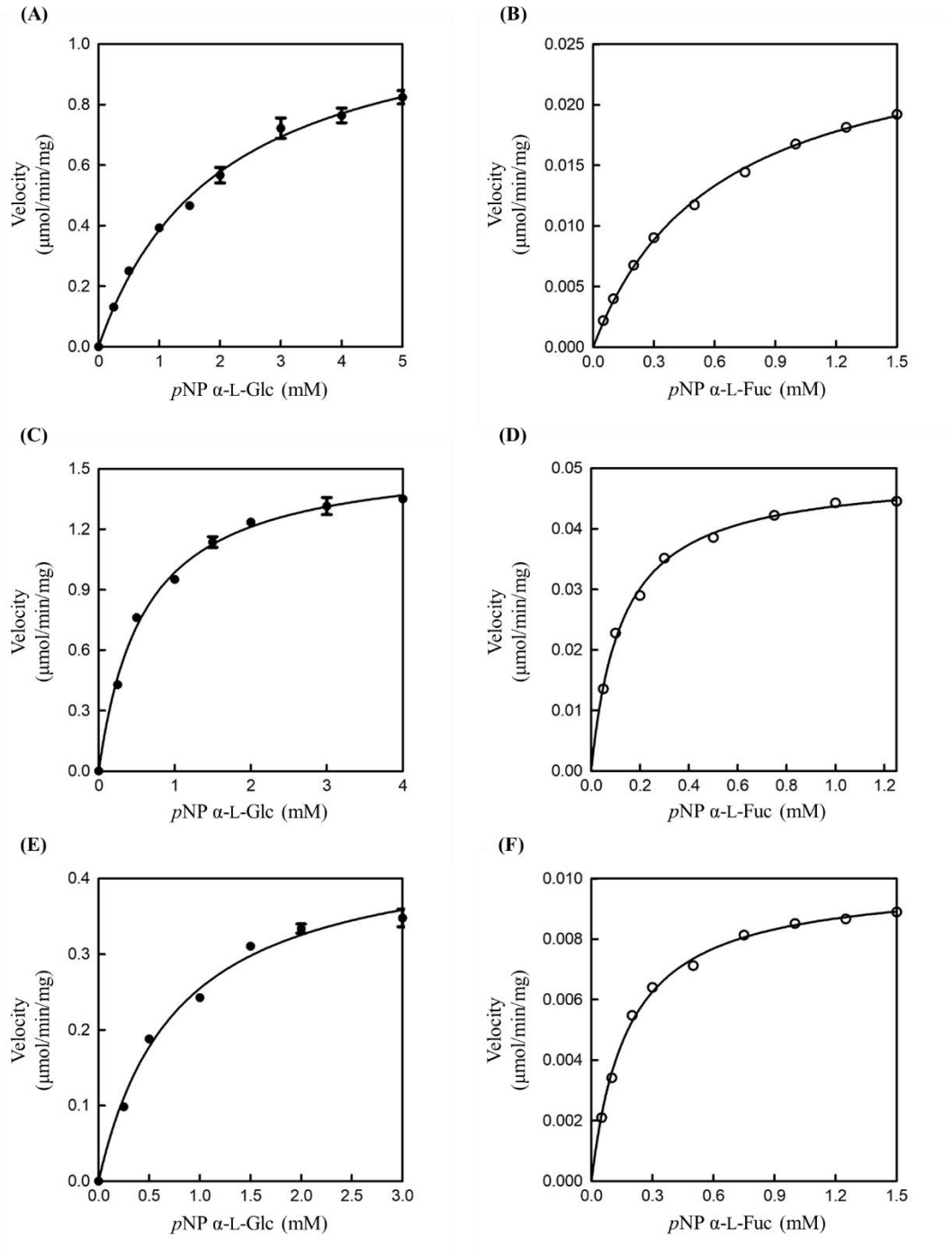


Fig. 4-6. Relationship between substrate concentration and hydrolysis velocity. [S]-*v* plots of ClAgl29B S132D for *p*NP α-L-Glc (**A**) and *p*NP α-L-Fuc (**B**), ClAgl29B Q275L for *p*NP α-L-Glc (**C**) and *p*NP α-L-Fuc (**D**), and ClAgl29B S132D/Q275L for *p*NP α-L-Glc (**E**) and *p*NP α-L-Fuc (**F**). The reaction velocity was measured using sodium acetate buffer (pH 6.0 and 6.5), enzyme and each substrate at 35℃. Enzyme and substrate concentration are shown in Materials and methods. The curve shown is a fit of the data according to the Michaelis-Menten equation.

Table 4-2. Kinetic parameters of S132D, Q275L, and S132D/Q275L

Substrate	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m ($s^{-1}mM^{-1}$)
S132D			
*p*NP α-L-Glc	1.99 ± 0.17	3.62 ± 0.40	1.81 ± 0.06
*p*NP α-L-Fuc	0.610 ± 0.029	0.084 ± 0.001	0.138 ± 0.004
Q275L			
*p*NP α-L-Glc	0.596 ± 0.11	1.73 ± 0.21	2.90 ± 0.12
*p*NP α-L-Fuc	0.127 ± 0.007	0.0540 ± 0.001	0.415 ± 0.004
S132D/Q275L			
*p*NP α-L-Glc	0.690 ± 0.11	1.79 ± 0.22	2.59 ± 0.10
*p*NP α-L-Fuc	0.178 ± 0.009	0.0420 ± 0.0010	0.234 ± 0.006
ClAgl29A[1)]			
*p*NP α-L-Glc	2.01 ± 0.26	2.00 ± 0.40	1.00 ± 0.08
*p*NP α-L-Fuc	0.279 ± 0.029	0.0658 ± 0.0008	0.236 ± 0.005
ClAgl29B[1)]			
*p*NP α-L-Glc	1.79 ± 0.17	7.76 ± 0.41	4.34 ± 0.22
*p*NP α-L-Fuc	0.240 ± 0.029	0.0844 ± 0.0019	0.352 ± 0.007

1) Data was from Shishiuchi *et al*. (2022)

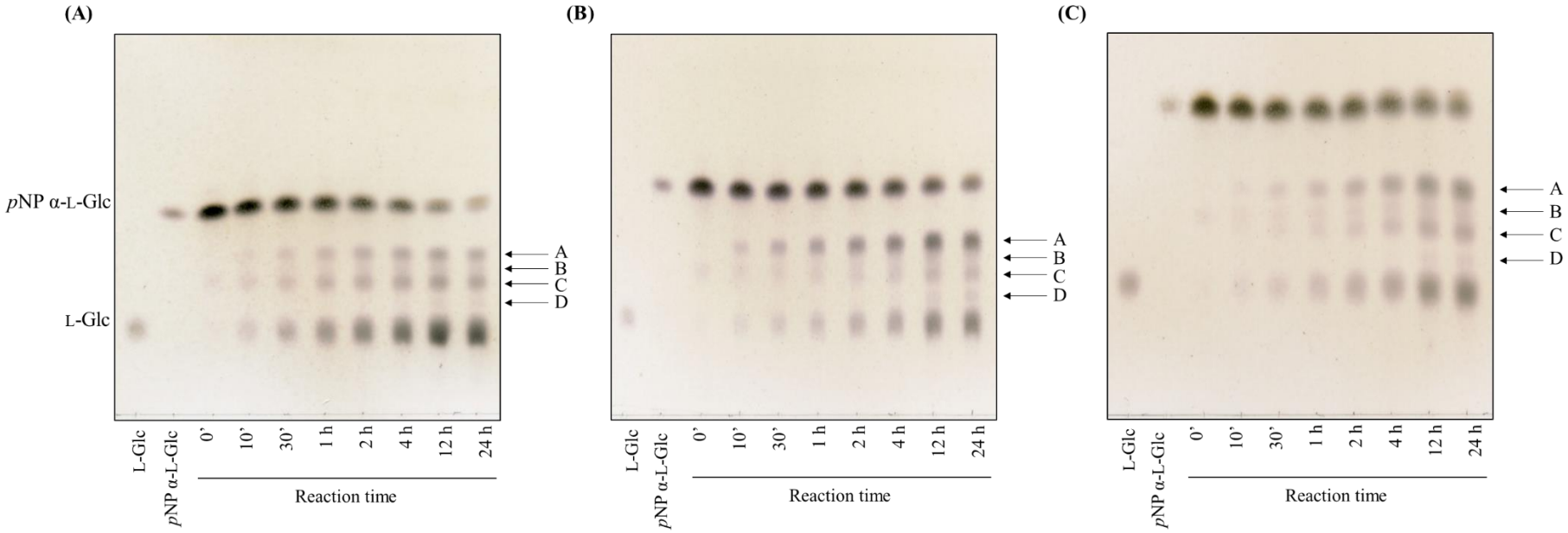


Fig. 4-7. TLC analysis of the reaction products of S132D (A), Q275L (B), and S132D/Q275L (C). The reaction was performed in 30 mM *p*NP α-L-Glc and enzyme (ClAgl29B S132D, 0.266 μM; ClAgl29B Q275L, 0.275 μM; ClAgl29B S132D/Q275L, 0.266 μM) in 40 mM sodium acetate buffer (pH 6.0 and pH 6.5) at 35℃. The mobile phase consisted of ethyl acetate/methanol/water (70/20/10, v/v/v). Arrows indicated transglucosylation products.

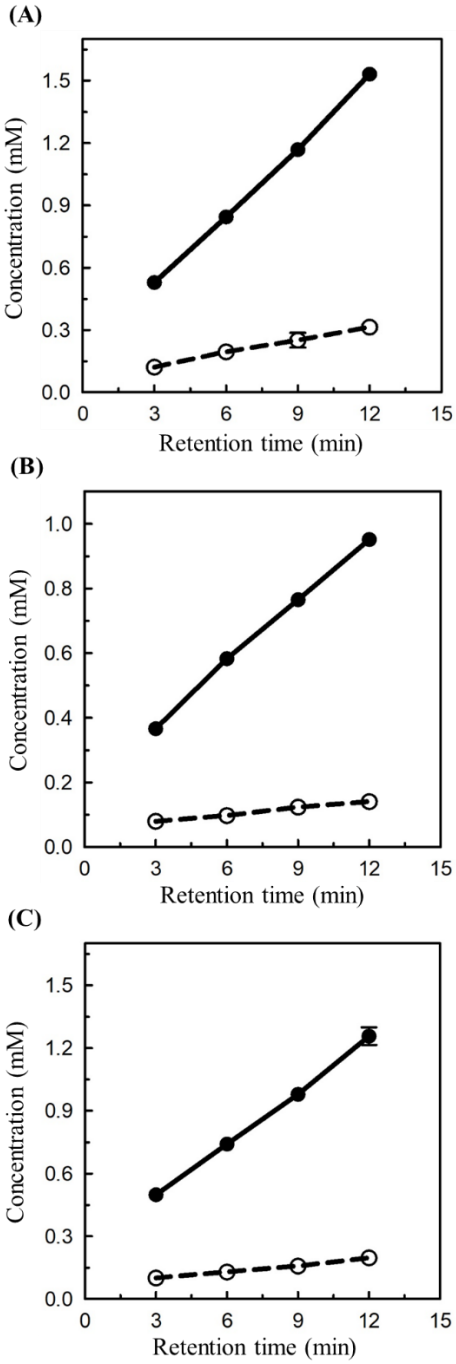


Fig. 4-8. Relationship between reaction time and concentrations of *p*NP (●) and L-Glc (○) for S132D (A), Q275L (B), and S132D/Q275L (C). The reaction was performed in 30 mM *p*NP α-L-Glc and enzyme (ClAgl29B S132D, 0.107 μM; ClAgl29B Q275L, 0.202 μM; ClAgl29B S132D/Q275L, 0.266 μM) in 40 mM sodium acetate buffer (pH 6.0 and pH 6.5) at 35℃ for 3, 6, 9, and 12 min. The concentration of *p*NP was quantified by colorimetric method at 400 nm. The concentration of L-Glc was quantified by HPAEC-PAD. Total reaction velocity and hydrolytic velocity were calculated from the slope of each linear regression line.

Table 4-3. Total, hydrolysis, and the calculated transglucosylation rates using 30 mM *p*NP α-L-Glc as a substrate of S132D, Q275L, and S132D/Q275L

	v_{total}	v_{hydr}	v_{tg}	v_{tg}/v_{total} (×100)
		μmol min^{-1} mg^{-1}		%
S132D	15.5	2.95	12.6	81.0
Q275L	4.71	0.460	4.25	90.2
S132D/Q275L	4.95	0.620	4.33	87.5

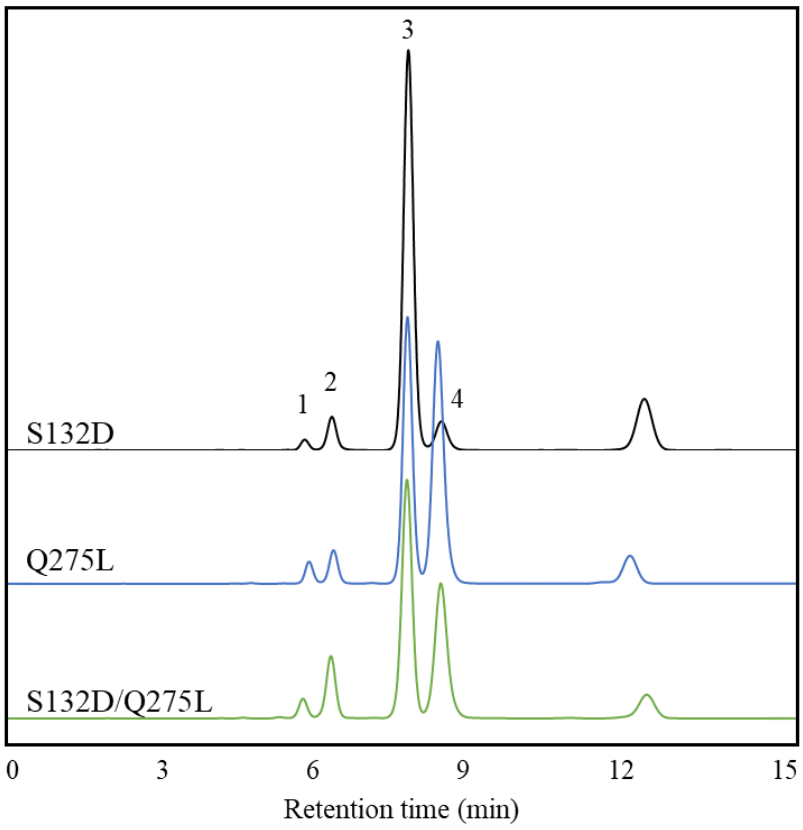


Fig. 4-9. Reversed-phase-HPLC analysis of the reaction mixtures by S132D, Q275L, and S132D/Q275L. The reaction was performed in 30 mM *p*NP α-L-Glc and enzyme (ClAgl29B S132D, 0.28 μM; ClAgl29B Q275L, 2.0 μM; ClAgl29B S132D/Q275L, 2.0 μM) in 40 mM sodium acetate buffer (pH 6.0 and 6.5) at 35℃ for 6 h. (1) *p*NP 2-O-α-L-Glc α-L-Glc; (2) *p*NP 6-O-α-L-Glc α-L-Glc; (3) *p*NP α-L-Glc; (4) *p*NP 3-O-α-L-Glc α-L-Glc.

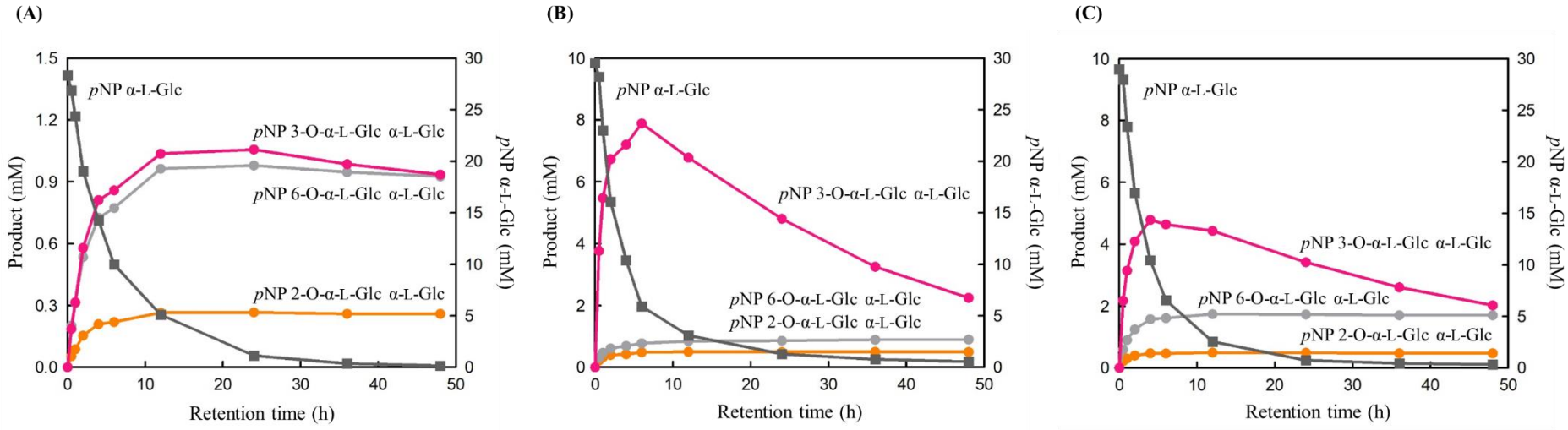


Fig. 4-10. The time course of the concentration of the transglucosylation products by S132D (A), Q275L (B), and S132D/Q275L (C). The reactions started with 30 mM *p*NP α-L-Glc and enzyme (ClAgl29B S132D, 0.28 μM; ClAgl29B Q275L, 2.0 μM; ClAgl29B S132D/Q275L, 2.0 μM) in 40 mM sodium acetate buffer (pH 6.0 and 6.5) at 35℃ for 48 h. The carbohydrates bound to *p*NP were quantified by RP-HPLC.

4. Discussion

The objective of this chapter was to modify the transglucosylation activity of ClAgl29B by site-directed mutagenesis. Since ClAgl29A was known to have higher transglucosylation activity than ClAgl29B, this project attempted to achieve the goal by replacing the subsite +1 of ClAgl29B with an ClAgl29A-like subsite +1. Structural comparison among ClAgl29A, ClAgl29B, and TmaFuc showed that the structures of subsite +1 of two α-L-glucosidases were conserved, but two substitutions were found in Ser132 and Gln275, which were Asp120 and Leu263, respectively (Fig. 4-3).

The specific activity of three mutant enzymes was decreased. The specific activity of ClAgl29A was lower than that of ClAgl29B, suggesting that the introduction of the mutation brought the specific activity closer to that of ClAgl29A. However, it is necessary to confirm the increase in specific activity when the corresponding residues of ClAgl29A are changed to ClAgl29B-like residues. The optimum pH of the mutant enzymes for hydrolysis of 2 mM *p*NP α-L-Glc shifted from 0.5 to 1.0 unit toward the alkaline side. Replacement of uncharged residues to acidic residues, S132D, and of hydrophilic residues to hydrophobic residues, Q275L, could affect the dissociation state of the catalytic group, even though the substituted residues were 8 to 9 Å away from the catalytic residues. It is undeniable that the increased sensitivity of the mutant enzymes to pH on the acidic side may affect the optimum pH. The measurements were performed at 4°C for 24 h in this study, and the pH-stable range should be examined under conditions similar to the enzyme reaction. Kinetic parameter analysis of the mutant enzyme toward *p*NP α-L-Glc obtained in the low substrate concentration range showed a decrease in K_m in Q275L and S132D/Q275L. This may be due to the substitution of Gln275 for the more hydrophobic Leu, which increased the adhesiveness of the *p*NP group to the enzyme. On the other hand, there was also a decrease in k_{cat} in these mutant enzymes and, in addition, an increase in transglucosylations, suggesting an increase in nonproductive binding of the α-L-glucosyl group of *p*NP α-L-Glc to subsite +1.

Notably, S132D, Q275L, and S132D/Q275L showed high transglucosylation ratios of 80.9, 90.2, and 87.4%, respectively. In the present study, the mutation was introduced with reference to the structure of ClAgl29A, but the transglucosylation ratio of the mutant enzymes was superior to that of ClAgl29A. It is possible that the more hydrophobic environment of Q275L makes it easier for the sugar and *p*NP group to bind to the enzyme. However, docking simulation study with the S132D/Q275L mutant enzyme and *p*NP 3-*O*-α-L-Glc α-L-Glc or *p*NP 6-*O*-α-L-Glc α-L-Glc (Fig. 4-11) was insufficient to provide a clear rationale. Similarly, no clear reason can be given for the high α1-6 transfer ability of S132D, which ClAgl29A does not have. In the simulations, the participation of this residue appears to be small. In protein engineering, it may not be strange to go beyond the properties of the parent. For example, *in vitro* recombination of the thymidine kinase genes of herpes simplex virus types 1 and 2 using a DNA family shuffling has yielded enzymes with activity 16,000-fold greater than that of the parent.

The levels of *p*NP 3-*O*-α-L-Glc α-L-Glc decreased in Q275L and S132D/Q275L. This decrease may be attributed to either hydrolysis or consumption as an acceptor of the following transglucosylation. The fact that the concentration of the *p*NP trisaccharide did not increase when *p*NP 3-*O*-α-L-Glc α-L-Glc decreased suggests that hydrolysis is the more likely cause. The accumulation of sufficient amounts of *p*NP 3-*O*-α-L-Glc α-L-Glc can serve as a substrate for hydrolysis.

The mutant enzymes constructed in this chapter produced α-L-glucooligosaccharides more efficiently than previously reported ClAgl29A and ClAgl29B. It is a well-known fact that oligosaccharides composed of D-series monosaccharides benefit human health by regulating the intestinal environment and the immune system (Zhang, N., 2022; Jeurink, P. V., 2013). On the other hand, the functionality of α-L-glucooligosaccharides is still unknown. The development of an efficient enzymatic production method for α-L-glucooligosaccharides in this study may help elucidate these carbohydrates' physiological functions.

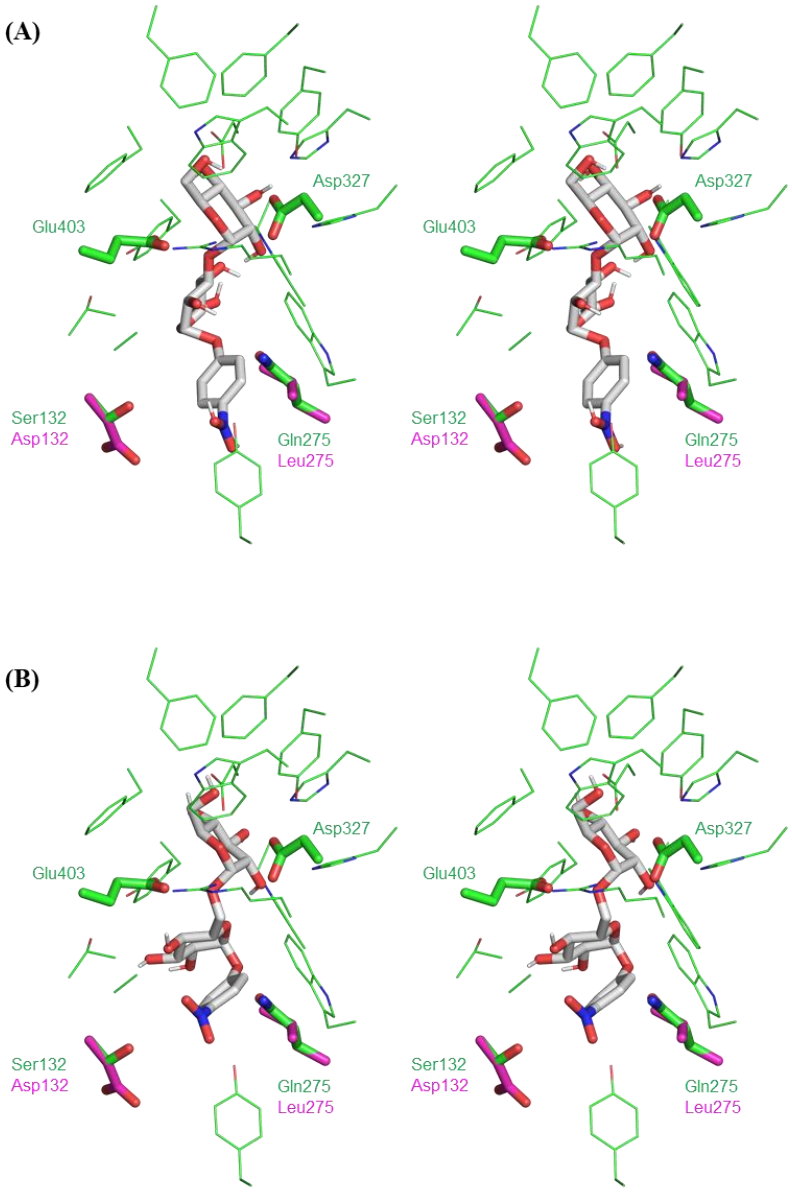


Fig. 4-11. Docking simulations of S132D/Q275L and *p*NP 3-*O*-α-L-Glc α-L-Glc (A) and *p*NP 6-*O*-α-L-Glc α-L-Glc (B).

The relevant residues, Ser132 and Gln275 of the wild-type ClAgl29B, and Asp132 and Leu275 of S132D/Q275L, are shown by stick model. Asp327 and Glu403 provide the catalytic nucleophile and catalytic general acid/base, respectively.

Chapter 5. Production of α-L-oligosaccharides by α-L-glucosidase glycosynthase

1. Introduction

Long since, carbohydrates have been known to be used just as structural support and the energy source of the cell (Varki, A., 2015). However, in recent years, in addition to these basic roles, various functions of carbohydrates have been elucidated. Prebiotic oligosaccharides, such as fructooligosaccharides and galactooligosaccharides, stimulate the activity and growth of beneficial bacteria in our intestines and bring us good health. Carbohydrates, especially glycoconjugates, are also significantly contributing to the pharmaceutical field. Bacteria and viruses infect cells by binding to specific sugar chains on the cell surface. By preventing binding to these sugar chains, infectious diseases such as influenza that target specific tissues and organs can be prevented and treated (Harms, G., 1996). The addition of glycans and modification of glycan structures are also used to confer resistance to protease degradation, targeting specific organs and improving the solubility of compounds (Niu, C., 2016). For these reasons, there has been a great deal of interest in carbohydrate-based pharmaceuticals and developing techniques for synthesizing oligosaccharides. Glycosyltransferases are responsible for synthesizing glycosidic linkages *in vivo*, but these enzymes are not suitable for mass production of the glycosidic linkages *in vitro* due to the difficulty in obtaining the sugar nucleotides as substrates and generally small K_m. On the other hand, glycosyltransferases have recently begun to be used in the production of oligosaccharides using metabolic engineering. Glycosidases have been used in industrial oligosaccharide production. Among glycosidases, glycosidases with a retaining mechanism catalyze transglycosylation to form glycosidic bonds in addition to hydrolysis (Fig. 5-1 A). Many glycosidases, which have their own substrate specificities, have been found, and they can be used to produce a wide variety of glycosidic linkages. Furthermore, substrates of glycosidases are relatively inexpensive.

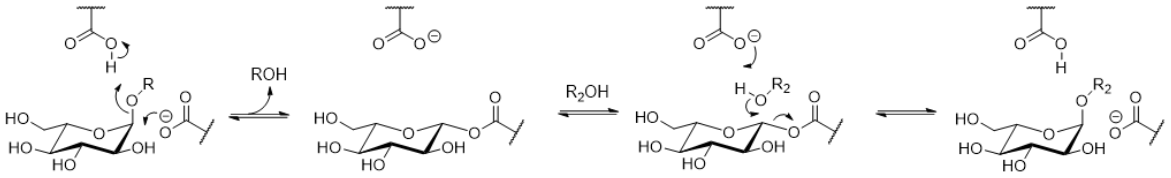
Among the transglycosylation of glycosidases, that of GH29 α-L-fucosidase is one of the most well-studied enzymes. This is because α-L-fucosylation can diversify the functions of oligosaccharides, of which glycoconjugates play many biological roles in cell-cell interactions, signal transduction, immune responses, pathogen adhesion processes, early embryogenesis, and apoptosis (Pekdemir, B., 2024). Among them, α-L-fucosylation of human milk oligosaccharides has attracted much attention. In general, α-L-fucosidases have broad acceptor specificity and can catalyze transfucosylation of diverse mono-, di-, and oligosaccharides or particularly *p*NP monosaccharides acceptors (Wan, L., 2020). As mentioned previous chapters, ClAgl29A and ClAgl29B catalyze α-L-transglucosylation and produce α-(1→3)-, α-(1→2)-, and α-(1→6)-L-glucosidic linkages, using *p*NP α-L-Glc as substrate. ClAgl29B has a lower transglucosylation ability than ClAgl29A, but its S132D, Q275L, and S132D/Q275 mutant enzymes show high transglucosylation activity.

The glycosynthase technology has emerged as a highly effective tool for synthesizing oligosaccharides and polysaccharides. Glycosynthases are mutant glycosidases that replace the nucleophilic catalytic residues of glycoside hydrolase with non-nucleophilic amino acids (Fig. 5-1 B).

Glycosynthase lacks hydrolytic activity but is capable of efficiently catalyzing the formation of glycosidic bonds in high yields when using glycosyl fluoride with an anomeric configuration opposite to the original substrate as a donor substrate (Mackenzie, L. F., 1998). In GH29, 1,3-1,4-α-L-fucosidase from *B. longum* subsp. *infantis* (BiAfcB) has been successfully converted into glycosynthase (Sakurama H., 2012b). The glycosynthase catalyzed the formation of Lewis^{x} and Lewis^{a} using β-L-fucosyl fluoride as a donor substrate. On the other hand, since the half-life of this compound is as short as 20 min, glycosynthase using more stable β-L-fucosyl azide as the activating donor has also been reported (Cobucci-Ponzano, B., 2009). Furthermore, the activated donor and its transglycosylation product can both act as acceptors, resulting in the autocondensation of the donor or the elongation of the transglycosylation product, such as oligosaccharides and polysaccharides (Trincone, A., 2005).

In this Chapter, the catalytic nucleophile mutant enzymes of ClAgl29A, ClAgl29B, S132D-ClAgl29B, and Q275L-ClAgl29B, of which Asp327 was replaced by Ser, were produced, and their glycosynthase ability was evaluated. As mentioned above, ClAgl29A, and ClAgl29B and its derivatives catalyzed transglycosylation, but the transglucosylation products were hydrolyzed late in the reaction. Glycosynthase prevents this hydrolysis. In addition, since *p*NP α-L-Glc is used as a substrate in these reactions, *p*NP is included in the product. The HF produced by glycosynthase has a low boiling point and can be easily removed by evaporation.

(A)



(B)

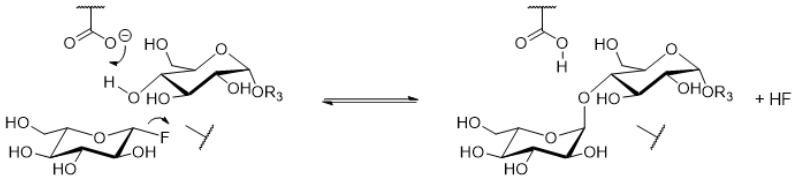


Fig. 5-1 Catalytic mechanism of wild-type α-L-glucosidase (A) and α-L-glucosynthase (B)

R2 = H, hydrolysis; R2 ≠ H, trasnglucosylation

2. Materials and methods

2.1 Materials

*p*NP α-D-glucopyranoside and *p*NP β-D-glucopyranoside were purchased from Tokyo Chemical Industry Co., Ltd. *p*NP α-D-xylopyranoside, *p*NP β-D-xylopyranoside, and *p*NP β-D-cellobioside were purchased from Biosynth Carbosynth. All other chemicals used were analytical grade.

2.2 Construction of expression plasmids

A base substitution was introduced by following the method of PrimSTAR Mutagenesis Basal Kit described in Section 2.2 of Chapter 4, with a slight modification that the PCR products were treated with DpnI prior to the transformation of *E. coli*. Codon substitution of Asp327 of ClAgl29B for Ser was performed with oligonucleotides 5'-TGGTATAGCGGGTGGGGAAGTCCT-3' (sense) and 5'-CCCACCGCTATACCAAACAAGGTCGGGG-3' (antisense). The ClAgl29B/pET28a plasmid DNA and its derivative harboring the mutation(s), S132D, Q275L, and S132D/Q275L (see 2.2 in Chapter 4), were used as a template. The plasmid DNA for the production of D315S of ClAgl29A was constructed by Shishuchi.

2.3 Production and purification of enzymes

The mutant ClAgl29A, D315S of ClAgl29A (D315S_A), and mutant ClAgl29B, D327S of ClAgl29B (D327S_B), S132D/D327S_B, and Q275L/D327S_B were produced in *E. coli* Rosetta (DE3) as His_6-tagged proteins. The enzymes were purified by Ni-affinity chromatography as described in Section 2.3 of Chapter 4. To prepare the enzymes for crystallization, anion exchange and gel-filtration column chromatographies were performed.

DEAE Sepharose Fast Flow (Cytiva) was packed in a glass column (Φ 2.5 cm × 5 cm, 25 cm^3) and equilibrated with 10 mM sodium phosphate buffer (pH 7.5). Enzymes collected after Ni-affinity column chromatography were dialyzed in 10 mM sodium phosphate buffer (pH 7.5) and loaded onto the anion-exchange column. Bound proteins were eluted by a linear concentration increase (from 0 to 700 mM) of sodium chloride in an equilibration buffer with a flow rate of 25 mL/h. The collected fractions, determined from SDS-PAGE results, were concentrated using Amicon Ultra 30,000 (Millipore).

The concentrated protein was loaded onto a Sephacryl S-200 HR (Cytiva) column (Φ 2.5 cm × 92 cm, 450 cm^3) filled with 10 mM sodium phosphate buffer (pH 7.5). A flow rate was set at 35 mL/h. The fractions pooled were determined by SDS-PAGE analysis.

The protein concentration of purified enzymes was estimated by amino acid analysis as described in Section 2.4 of Chapter 4. The specific extinction coefficients ($E^{1\%}_{280\ nm,1cm}$) calculated were 22.5 for D315S_A, 21.6 for D327S_B, and 26.1 for Q275L/D327S_B. The molar mass (g/mol) values used were 66,035 (D315S_A), 66,090 (D327S_B), 67,118 (S132D/D327S_B), and 67,075 (Q275L/D327S_B).

2.4 Synthesis β-L-glucopyranosyl fluoride (β-Glc-F)

To a mixture of 5.52 g acetyl bromide, 255.3 mg zinc bromide, and 10 mL of dichloromethane (DCM) was added 1 g L-glucose and stirred at room temperature for 26 h. Under ice-cold conditions, the reaction mixture was diluted with DCM and was washed with a 10% (w/v) sodium bicarbonate aqueous solution. The organic layer was dried with sodium sulfate and evaporated. The obtained α-L-bromoglucose tetraacetate (1.2 g) was dissolved in one mL of anhydrous acetonitrile, and silver(I) fluoride (1.2 g) was added. The suspension was stirred at room temperature for 1 h and filtered over a Celite pad to remove the insoluble solids. The Celite was washed by 10 mL of acetonitrile and concentrated *in vacuo*. The brown syrup diluted in DCM was washed with water, dried with sodium sulfate, and evaporated. Flash silica gel column chromatography (Hexane/Ethyl acetate=2/1, v/v) of the crude products gave pure β-L-fluoroglucose tetraacetate. ^{1}H NMR (500 MHz, $CDCl_3$): δ 5.36 (dd, J = 6.0, 52.0 Hz, 1H, H-1), 5.22 (m, 1H, H-4), 5.20 (m, 1H, H-3), 5.13-5.08 (m, 1H, H-2), 4.27 (dd, J = 4.7, 12.3 Hz, 1H, H-6a), 4.22 (dd, J = 2.7, 12.3 Hz, 1H, H-6b), 3.92-3.89 (m, 1H, H-5), 2.11 (s, 3H, OAc), 2.10 (s, 3H, OAc), 2.05 (s, 3H, OAc), 2.04 (s, 3H, OAc).

Deacetylation of β-L-fluoroglucose tetraacetate was performed by Zemplén deacetylation procedure. To a flask containing β-L-fluoroglucose tetraacetate 10 mL of absolute methanol was added and stirred at 0°C until the substrate was fully dissolved. Sodium methoxide (28% in methanol, 50 μL) was added in one portion and stirred at room temperature for over 3 h. The reaction flask was cooled in an ice-water bath, and 3 g of silica gel was added. The resulting suspension was stirred for 10 min, then concentrated *in vacuo*. To the flask 10 mL of ethyl acetate/methanol solution (7/3, v/v) was added, and stirred for 5 min. The mixture was filtered and rinsed with ethyl acetate/methanol (7/3, v/v) solution (10 mL). The filtrate was concentrated under a vacuum, giving 132 mg of β-Glc-F. The synthesized β-Glc-F was dissolved with methanol and kept in a -80°C freezer. The concentration of β-Glc-F was determined by the phenol-sulfuric acid method; briefly, 5 μL of diluted β-Glc-F was mixed with 50 μL of 2.5% (w/v) phenol, 250 μL of sulfuric acid was added, and vortexed well. The reaction mixture was incubated at room temperature for 10 min and measured at absorbance 490 nm. The calibration curve was made using 0 to 5 mM of D-glucose.

2.5 Glycosynthase reaction with various acceptors

The solution consisting of β-Glc-F (100 mM, 10 μL) and each acceptor in 200 mM sodium acetate buffer (pH 5.5) (25 mM, 12.5 mM, or 250 mM; 20 μL) was incubated at 35°C for 3 min. Then, 10 μL of each enzyme solution (5.75 μM for D315S_A, 5.30 μM for D327S_B, and 5.80 μM for Q275L/D327S_B) was added and incubated at 35°C for 20 h. The reaction was terminated by heating at 100°C for 2 min. The compounds used for acceptor tests were summarized in Table 5-1. TLC analysis was performed as described in Section 2.7 of Chapter 4, except for the solvent system used, ethyl acetate/methanol/water (7/2/1, v/v/v) for *p*NP derivatives and acetonitrile and water (80/20, v/v) for

oligosaccharides. The reaction products of the *p*NP derivatives were also detected by UV irradiation at 254 nm.

2.6 Estimation of the yield of the product

A solution consisting of 10 μL of 100 mM β-Glc-F and 20 μL of acceptor (*p*NP α-D-Glc, *p*NP β-D-Glc, *p*NP α-D-Xyl, *p*NP β-D-Xyl, and *p*NP β-D-Cel) in 80 mM sodium acetate buffer (pH 5.5) was kept at 35°C for 3 min. Then, 10 μL of the enzyme solution (25.7 μM for D315S _A, 28.2 μM for D327S_B, and 33.4 μM for Q275L/D327S_B) was mixed and the reaction mixture was incubated at 35°C for 20 h. The reaction was terminated by heating at 100°C for 2 min. The reaction solution was centrifuged at 13,000 × *g* for 10 min at 4°C, and the supernatant was subjected to RP-HPLC. For analysis of the reaction product using *p*NP α-D-Xyl as acceptor, the supernatant obtained was diluted to 1/5 and for *p*NP β-D-Xyl and *p*NP β-D-Cel, diluted to 1/2 with Milli-Q before being subjected to RP-HPLC. In the RP-HPLC analysis, Cosmosil 5C_{18}-PAQ analytical column (4.6 ID × 150 mm, Nacalai Tesque) was used, and the temperature was set at 40°C. The mobile phase was methanol/water (20:80, v/v) at a flow rate of 1 mL/min. *p*NP sugars were detected by measuring the absorbance at 312 nm and the transglucosylation ratio (%) was determined by dividing the obtained peak area of a product by the original concentration of the acceptor.

2.7 Structure determination of the glycosynthase products

The reaction mixture containing β-Glc-F (20 mM), each acceptor (*p*NP α-D-Glc, *p*NP β-D-Glc, *p*NP α-D-Xyl, *p*NP β-D-Xyl, or *p*NP β-D-Cel; 10 mM), and each enzyme (6.42 μM for D315S_A, 7.05 μM for D327S_B, and 5.78 μM for Q275L/D327S_B) in 40 mM sodium acetate buffer, pH 5.5 was incubated at 35°C for 20 h. After the reaction was stopped by heating at 100°C for 3 min, the solution was desalted using Amberlite MB-4 and dried under reduced pressure in a CVE-100 centrifugal evaporator (EYELA). The sample dissolved with 1 mL of 20% methanol (v/v) was loaded onto RP-HPLC. The sample (50 μL) was injected into COSMOSIL 5C_{18}-PAQ analytical column (10 ID × 250 mm) with a mobile phase of 20% methanol (v/v) at a flow rate of 3.0 mL/min. Separation was performed at room temperature. pNP glycosides were detected at an absorbance of 312 nm.

The ESI-MS analysis was conducted by the Instrumental Analysis Division of the Global Facility Center, Hokkaido University Research Institute for Innovative Science and Technology. Cations and anions were detected using an Exactive mass spectrometer (Thermo Fisher Scientific). The products were purified and dried up as described above and dissolved in 20 μL absolute methanol.

For NMR analysis, purified products were dried by a centrifugal evaporator and dissolved in D_2O (Kanto Chemical) or DMSO-d6 (Sigma-Aldrich). The *p*NP glycosides from *p*NP α-D-Glc and *p*NP β-D-Glc were dissolved in D_2O adjusting to 7 mg/40 μL. The *p*NP glycoside from *p*NP β-D-Cel was adjusted to 10 mg/40 μL in D_2O. For the *p*NP glycosides of *p*NP α-D-Xyl and *p*NP β-D-Xyl, the samples

were dissolved in methanol-d4 (Kanto Chemical) and evaporated several times for H/D substitution, then dissolved in DMSO-d6 at 10 mg/40 μL. A 1.7 mm micro NMR tube was used for measurement. ^{1}H-NMR (500 MHz) and ^{13}C-NMR (125 MHz), COSY, HSQC-TOCSY, HMBC, and H2BC spectra were recorded using a Bruker AVANCE I spectrometer (Bruker Corporation). Measurements were requested from the GC-MS&NMR Laboratory of the Faculty of Agriculture, Hokkaido University.

2.8 X-ray crystal structure analysis of S132D/D327S_B and Q275L/D327S_B

2.8.1 Crystallization

All crystallizations were performed with the hanging drop vapor diffusion method at 20°C for approximately 2 weeks according to Section 2.1 of Chapter 2. Each crystal was obtained by slightly modifying the conditions for the wild-type ClAgl29B as follows.

The crystals of S132D/D327S_B and Q275L/D327S_B were obtained in a drop composed of 2 μL of reservoir solution [5% (v/v) 2-propanol, 0.1 M citrate-NaOH buffer (pH 3.5), and 6% (w/v) PEG 20,000] and 1 μL of S132D/D327S_B (A_{280}=23.8), and a drop composed of 1 μL of reservoir solution [5% (v/v) 2-propanol, 0.1 M citrate-NaOH buffer (pH 3.5), and 5% (w/v) PEG 20,000] and 1 μL of Q275L/D327S_B (A_{280}=31.7), respectively.

2.8.2 Data collection

All co-crystals were soaked into a cryoprotectant [5% 2-propanol, 0.1 M citrate-NaOH buffer (pH 3.5), 7% PEG 20,000, 10% glycerol, and 30 mM *p*NP 3-*O*-α-L-Glc α-L-Glc] for 10 seconds and flash-frozen in liquid nitrogen. The X-ray diffraction data were collected at beamline BL-5A at Photon Factory, KEK at a wavelength of 1.0000 Å using Pilatus3 S6M detector (Dectris). The dataset was collected from a single crystal under a stream of nitrogen at 100 K. The diffraction dataset was indexed, integrated, and scaled using the XDS program.

2.8.3 Structure determination and refinement

Their structures were determined and refined using the programs Refmac5 in conjunction with interactive fitting and rebuilding based on 2m*F*o−D*F*c and m*F*o−D*F*c electron densities using COOT using the crystal structure of ClAgl29B wild-type (PDB ID: 7XSG) as an initial model. For the refinement of *p*NP 3-*O*-α-L-Glc α-L-Glc, a new restraints cif file for α-L-glucose (designated as Z8U.cif) was created by reversing the chiral of C1 of that for β-L-glucose (Z8T.cif), and Z8U.cif and a restraints file for *p*NP (NPO.cif) were merged using the program Acedrg (Long, F., 2017). The final structure was defined as the convergence of the R_{work} and R_{free} values. The data processing and refinement statistics are presented in Table 5-2. The graphical representations were prepared using the Pymol Molecular Graphics Systems Version 2.0 (Schrödinger, LLC, New York, NY, USA).

Table 5-1. Summary of compounds used in the search for acceptors for the glycosynthase reaction

10 mM (final concentration)	
*p*NP α-D-glucopyranoside (*p*NP α-D-Glc)	*p*NP β-D-glucopyranoside (*p*NP β-D-Glc)
*p*NP α-D-galactopyranoside (*p*NP α-D-Gal)	*p*NP β-D-galactopyranoside (*p*NP β-D-Gal)
*p*NP α-D-xylopyranoside (*p*NP α-D-Xyl)	*p*NP β-D-xylopyranoside (*p*NP β-D-Xyl)
*p*NP β-D-xylopyranoside (*p*NP β-D-Xyl)	*p*NP α-L-arabinopyranoside (*p*NP α-L-Ara)
*p*NP β-D-cellobioside (*p*NP β-D-Cel)	
D-lyxose (D-Lyx)	D-xylose (D-Xyl)

5 mM (final concentration)	
*p*NP *N*-acetyl-β-D-glucosaminide (*p*NP β-D-GlcNAc)	*p*NP β-D-galactosyl-(1→4)-β-D-glucopyranoside (*p*NP β-D-Lac)
*p*NP α-L-glucopyranoside (*p*NP α-L-Glc)	

100 mM (final concentration)	
D-glucose (D-Glc)	L-glucose (L-Glc)
maltose (Mal)	sucrose (Suc)
D-fucose (D-Fuc)	D-fructose (D-Fru)
N-acetyl-D-glucosamine (D-GlcNAc)	L-sorbose (L-Sor)
D-galactose (D-Gal)	D-cellobiose (D-Cel)
D-xylobiose (D-Xyl2)	

Table 5-2. Summary of data collection and refinement statistics

	Q275L/D327S_B with *p*NP 3-*O*-α-L-Glc α-L-Glc	S132D/D327S_B with *p*NP 3-*O*-α-L-Glc α-L-Glc
PDB ID	Not deposited	Not deposited
Data collection		
Beamline	PF BL-5A	PF BL-5A
Wavelength	1.00000	1.00000
Space group	$P2_12_12_1$	$P2_12_12_1$
Unit cell		
a, *b*, *c* (Å)	72.81 122.66 166.89	72.36 121.49 166.71
Resolution range (Å)	44.20–1.60 (1.66–1.60)	44.07–1.50 (1.55–1.50)
Total No. of reflections	1313608 (123695)	1565703 (157501)
No. of unique reflections	194172 (18903)	233621 (22731)
Multiplicity	6.8 (6.5)	6.7 (6.8)
Completeness (%)	98.8 (97.1)	99.6 (97.9)
Mean $I/\sigma(I)$	12.91 (1.16)	14.71 (1.01)
R_{mean}	10.5 (143.5)	9.5 (189.9)
R_{sym}	9.7 (132.1)	8.8 (175.6)
CC1/2	0.999 (0.592)	0.999 (0.501)
Refinement		
R_{work}	19.4	19.2
R_{free}	21.3	21.1
Number of atoms		
Macromolecules	8901	8896
Ligands	132	132
Solvent	590	694
RMSD values from ideal		
Bonds	0.011	0.011
Angles	1.59	1.63
Ramachandran (%)		
Favored	96.83	96.83
Allowed	3.17	3.08
Outliers	0.00	0.00
Rotamer outliers (%)	0.82	0.82
Clashscore	1.64	2.09

3. Results

3.1 Evaluation of acceptor specificity of glycosynthases

D315S_A and D327S_B, which were enzymes with a mutation at catalytic nucleophile residue of ClAgl29A and ClAgl29B were prepared. Additionally, a catalytic nucleophile mutant enzyme of a ClAgl29B mutant enzyme, Q275L/D327S (ClAgl29B), that modified transglucosylation properties in Chapter 4 was also prepared (Q275L/D327S_B). Purified recombinant enzymes of 8.65 mg (D315S _A), 12.8 mg (D327S_B), and 7.99 mg (Q275L/D327S_B) were obtained from 1.5 L of recombinant *E. coli* culture (Fig. 5-2). The specific activities of the purified enzymes were all below the detection limit of 0.00257 U/mg.

The reaction mixtures containing β-Glc-F, various *p*NP-L-glycosides, and respective D315S_A, D327S_B, and Q275L/D327S_B incubated for 20 h was analyzed by TLC (Fig. 5-3). As a negative control, a reaction without an enzyme was performed. The reaction of D315S_A used *p*NP α-D-Xyl, *p*NP β-D-Xyl, and *p*NP β-D-Cel and produced α-L-glucosyl compounds as prominent spots on TLC. In addition, the reaction of D315S_A produced small amounts of α-L-glucosylated *p*NP α-D-Glc and *p*NP β-D-Glc to the extent that they were detectable with UV irradiation. Small amounts of α-L-glucosyl *p*NP α-D-Glc and *p*NP β-D-Glc were produced to the extent that they were detectable with UV irradiation. In the reaction in D327S_B, the same *p*NP compounds as in D315S_A acted as an acceptor, but the amount of product was markedly smaller than in D315S_A. On the other hand, in the reaction of Q275L/D327S_B, of which parent Q275L_B showed higher transglucosylation activity than the wild-type ClAgl29B, the amount of product was increased compared to that in the reaction of D327S_B.

Similarly, glycosynthase reaction products were analyzed by TLC when monosaccharides and disaccharides were used as acceptors (Fig. 5-4). In the reaction of D315S_A, only D-fructose functioned as an acceptor among the tested carbohydrates. D327S_B also used D-fructose as an acceptor, but two kinds of transfer products can be detected. D327S_B further gave the transfer products using D-Glc, D-Xyl, D-GlcNAc, L-Sor, and D-Cel as an acceptor. The glycosynthase reaction in Q275L/D327S_B was able to use compounds similar to those in D327S_B as acceptors, but the amount of product obtained was less than in D327S_B. In addition, Q275L/D327S_B did not have L-Sor as an acceptor.

3.2 Analysis of transglycosylation yield

Glycosynthase reaction products of D315S_A, D327S_B, and Q275L/D327S_B were quantified when using 20 mM β-Glc-F as substrate and 10 mM *p*NP α-D-Glc, *p*NP β-D-Glc, *p*NP α-D-Xyl, *p*NP β-D-Xyl, or *p*NP β-D-Cel as acceptor. These are the compounds whose products were detectable in the previous TLC analysis. The product concentrations in the reaction solution after 20 h are summarized in Table 5-3, 5-4, 5-5.

In D315S_A, it demonstrated the highest transglucosylation activity, reaching 91% yield for 20 h for the *p*NP α-D-Xyl acceptor and 71% yield for the *p*NP β-D-Xyl. In the reactions with *p*NP α-D-Glc,

*p*NP β-D-Glc, *p*NP β-D-Cel, *p*NP α-D-Xyl, and *p*NP β-D-Xyl as acceptors, five, three, three, three, and two products were detected by RP-HPLC, respectively (Fig. 5-5). In D327S_B, the highest transglucosylation activity was observed to reach 67% yield for 20 h for the *p*NP-β-D-Cel acceptor. In the reactions with *p*NP α-D-Glc, *p*NP β-D-Glc, *p*NP β-D-Cel, *p*NP α-D-Xyl, and *p*NP β-D-Xyl as acceptors, five, two, two, three, and one products were detected by RP-HPLC, respectively (Fig. 5-5). In D327S/Q275L_B, in the reactions with *p*NP α-D-Glc, *p*NP β-D-Glc, *p*NP β-D-Cel, *p*NP α-D-Xyl, and *p*NP β-D-Xyl as acceptors, five, two, two, three, and one products were detected by RP-HPLC, respectively (Fig. 5-5).

3.3 Structure elucidation of L-oligosaccharides

The column retention time in RP-HPLC showed that the main product produced from each acceptor was the same in all mutant enzymes. The products were thus prepared using D315S_A and D327S_B and subjected to ESI-MS (Fig. 5-6) and NMR analysis (Fig. 5-7 – 5-16). *p*NP α-D-Glc was converted to *p*-nitrophenyl α-L-glucopyranosyl-(1→4)-α-D-glucopyranoside (α-L-Glc*p*-(1→4)-α-D-Glc*p*-O-PNP) (Fig. 5-7; Table 5-6). Positive mode mass spectrometry found *m*/*z* 486.12 (*m*/*z* calculated was 486.12 $[M+Na]^+$) and negative mode mass spectrometry found *m*/*z* 462.12 (*m*/*z* calculated was 462.12 $[M-H]^-$) (Fig. 5-6 A). The α-(1→4) linkage was determined by the HMBC-correlation between H1′ (δ 4.84 ppm, J = 3.8 Hz) of the glucose residue and C4 (δ 79.3 ppm) of the glucose residue, and that between H4 (δ 3.47 ppm) of the glucose residue and C1′ (δ 99.7 ppm) of the glucose residue.

For *p*NP β-D-Glc as an acceptor, the transfer product was *p*-nitrophenyl α-L-glucopyranosyl-(1→4)-β-D-glucopyranoside (α-L-Glc*p*-(1→4)-β-D-Glc*p*-O-PNP) (Fig. 5-8; Table 5-7), of which mass spectrometry found 486.12 (calculated 486.12 $[M+Na]^+$), and found 462.13 (calculated 462.12 $[M-H]^-$) (Fig. 5-6 B). The HMBC-correlation between H1′ (δ 5.09 ppm, J = 3.8 Hz) of the glucose residue and C4 (δ 78.3 ppm) of the glucose residue, and that between H4 (δ 3.7 ppm) of the glucose residue and C1′ (δ 100.4 ppm) of the glucose residue revealed the α-(1→4) linkage.

For *p*NP α-D-Xyl as acceptor, *p*-nitrophenyl α-L-glucopyranosyl-(1→4)-α-D-xylopyranoside (α-L-Glc*p*-(1→4)-α-D-Xyl*p*-O-PNP) *m*/*z* calculated was 456.11 $[M+Na]^+$, and found was 456.11: calculated was 462.12 $[M-H]^-$, and found was 462.13 (Fig. 5-6 C). The correlation was determined by the HMBC-correlation between H1′ (δ 5.05 ppm, J = 3.7 Hz) of the glucose residue and C4 (δ 82.4 ppm) of the glucose residue, and that between H4 (δ 3.73 ppm) of the glucose residue and C1′ (δ 99.9 ppm) of the glucose residue (Fig. 5-9; Table 5-8).

For *p*NP β-D-Xyl as acceptor, *p*-nitrophenyl α-L-glucopyranosyl-(1→3)-β-D-xylopyranoside (α-L-Glc*p*-(1→3)-β-D-Xyl*p*-O-PNP) *m*/*z* calculated was 456.11 $[M+Na]^+$, and found was 456.11: calculated was 462.12 $[M-H]^-$, and found was 462.13 (Fig. 5-6 D). The HMBC-correlation between H1′ (δ 5.05 ppm, J = 3.8 Hz) of the glucose residue and C3 (δ 84.3 ppm) of the glucose residue, and that between H3 (δ 3.47 ppm) of the glucose residue and C1′ (δ 99.7 ppm) of the glucose residue revealed the α-

(1→3) linkage (Fig. 5-10; Table 5-9).

For *p*NP β-D-Cel as acceptor, *p*-nitrophenyl α-L-glucopyranosyl-(1→4)-β-D-glucopyranosyl-(1→4)-β-D-glucopyranoside (α-L-Glc*p*-(1→4)-β-D-Glc*p*-(1→4)-β-D-Glc*p*-O-PNP) *m*/*z* calculated was 648.18 $[M+Na]^+$, and found was 648.17: calculated was 624.18 $[M-H]^-$, and found was 624.18 (Fig. 5-6 E). The α-(1→4) linkage was determined by the HMBC-correlation between H4′ (δ 3.62 ppm, J = 9.4 Hz) of the glucose residue and C1″ (δ 100.4 ppm) of the glucose residue, and that between H1″ (δ 5.04 ppm) of the glucose residue and C4′ (δ 78.3 ppm) of the glucose residue (Fig. 5-11; Table 5-10).

3.4 Crystal structure analysis of S132D/D327S_B and Q275L/D327S_B in a complex with *p*NP 3-*O*-α-L-Glc α-L-Glc

Crystal structures of S132D/D327S_B and Q275L/D327S_B in a complex of *p*NP 3-*O*-α-L-Glc α-L-Glc were determined to understand the structure of the subsite +1 and the effect of the mutations of S132D and Q275L, which could improve the transglucosylation ability (Chapter 4). The enzymes purified by Ni-affinity column chromatography were further purified by anion exchange and gel filtration chromatographies. For both enzymes, the latter peak in the final gel filtration column chromatography was collected and used for their crystallization (Fig. 5-17 – 5-19). To obtain the structure complex with *p*NP 3-*O*-α-L-Glc α-L-Glc, the ligand compound was added to the cryoprotectant when protein crystals were frozen. Structures of S132D/D327S_B and Q275L/D327S_B were determined at 1.5 and 1.6 Å, respectively. All residues were built based on electron density, except residues 128–132 and 378–386 in both enzymes. The crystal structures of both enzymes were similar to that of the wild-type enzyme with the root-mean-square deviations, 0.390 Å for S132D/D327S_B and 0.248 Å for Q275L/D327S_B. In the structure of S132D/D327S_B, the electron density was unclear and the structure of Ser132 could not be determined. In both S132D/D327S_B and Q275L/D327S_B, the side chain of Ser327 was present in two conformations (Fig. 5-20).

In both enzymes, citrate was bound to the active-site pocket (Fig. 5-21, Fig. 5-22). The bonding mode was consistent for both, and several hydrogen bonds were involved in the binding. Salt bridges were formed between Nε2 of His198 and O5 of citrate and between Nε2 of His199 and O2 of citrate. Citrate should be derived from the crystallization solution.

*p*NP 3-*O*-α-L-Glc α-L-Glc bound to the identical site away from the active pocket in both S132D/D327S_B and Q275L/D327S_B (Fig. 5-23). The catalytic TIM barrel folds of the GH29 enzymes, including ClAgl29A and ClAgl29B, characteristically lack a prominent α6. The *p*NP 3-*O*-α-L-Glc α-L-Glc binding site consisted of a loop connecting β6 and β7, β7 and α7 (Fig. 5-24). The binding site was not pocket-like, but rather shaped like a surface binding site (Fig. 5-24). In common with the two enzymes, several hydrogen bonds were involved in the binding (Fig. 5-25). O6 and O3 of α-L-glucosyl residue at the non-reducing end of *p*NP 3-*O*-α-L-Glc α-L-Glc (Ring *a*) formed hydrogen bonds with O and Oε2 of Glu441, respectively. O6 of the L-glucose ring *b* formed a hydrogen bond with the

N atom of Ile406. In addition, Nδ2 of Asn452 formed a hydrogen bond with O2 of the *p*NP group. Hydrogen bonds mediated by water molecules were also observed. In common with both enzymes, a water-mediated hydrogen bond was observed between O3 of the L-glucose ring *b* and Nη2 of Arg419. Several another water-mediated hydrogen bonds were observed in only S132D/D327S.

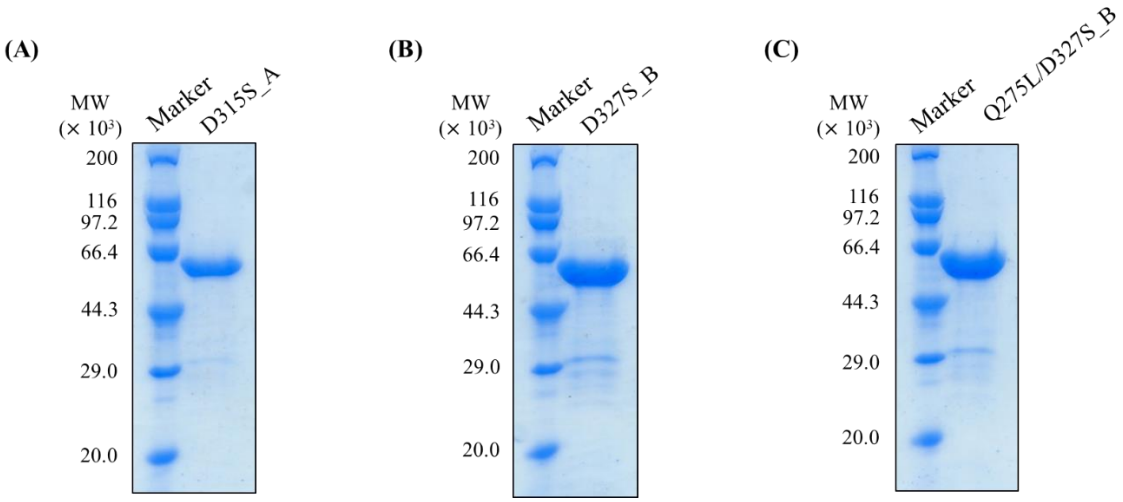


Fig. 5-2. SDS-PAGE of purified recombinant ClAgl29A D315S, ClAg29B D327S, and Q275L/D327S. Lane marker, size marker; lane D315S_A (A_{280}=0.70, 1.2 µg); lane D327S_B (A_{280}=1.24, 2.1 µg); and lane Q275L/D327S_B (A_{280}=1.94, 2.8 µg).

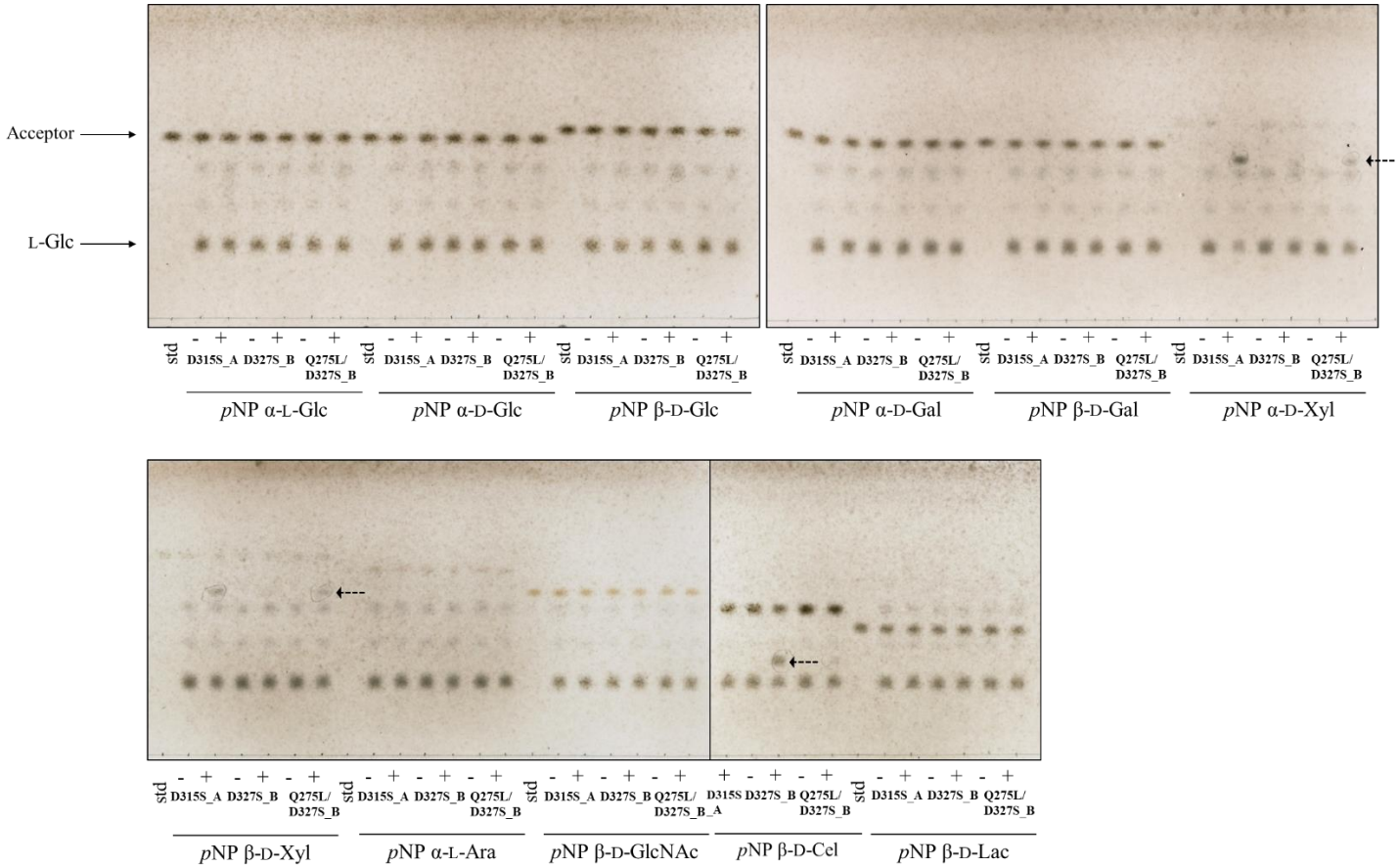


Fig. 5-3. TLC analysis of the reaction mixture of D315S_A, D327S_B, and Q275L/D327S_B with β-L-Glc-F. The reaction was performed in 20 mM β-L-Glc-F, 5 or 10 mM acceptors, enzyme (D315S, 1.15 μM; D327S, 1.06 μM; Q275L/D327S, 1.16 μM) in 40 mM sodium acetate buffer (5.5) at 35℃. The acceptors used in study are shown below the figure. Lane std, 5 or 10 mM acceptor; lane minus, the reaction mixture without enzyme; lane plus, reaction mixture. Glycosynthase reaction products are marked with black arrow. The developing solvent consisted of ethyl acetate/methanol/water (70/20/10, v/v/v).

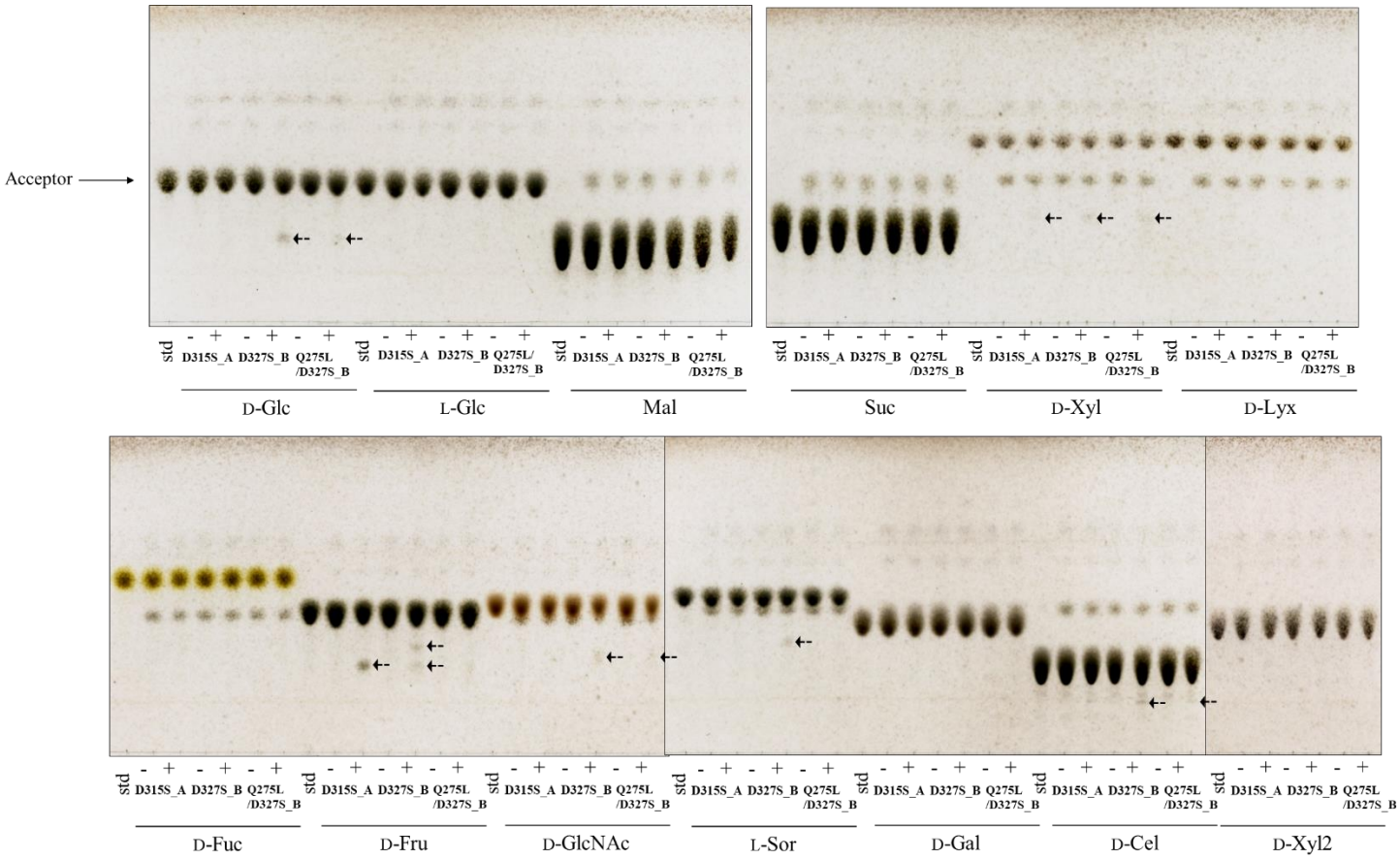


Fig. 5-4. TLC analysis of the reaction mixture of D315S_A, D327S_B, and Q275L/D327S_B with β-L-Glc-F. The enzyme reaction was performed in 20 mM β-L-Glc-F, 100 mM acceptors, enzyme (D315S, 1.15 μM; D327S, 1.06 μM; Q275L/D327S, 1.16 μM) in 40 mM sodium acetate buffer (5.5) at 35℃. Lane std, 20 mM acceptor; lane minus, the reaction mixture without enzyme; lane plus, reaction mixture. The acceptors used in study are shown below the figure. Glycosynthase reaction products are marked with black arrow. The developing solvent was acetonitrile/water (80/20, v/v).

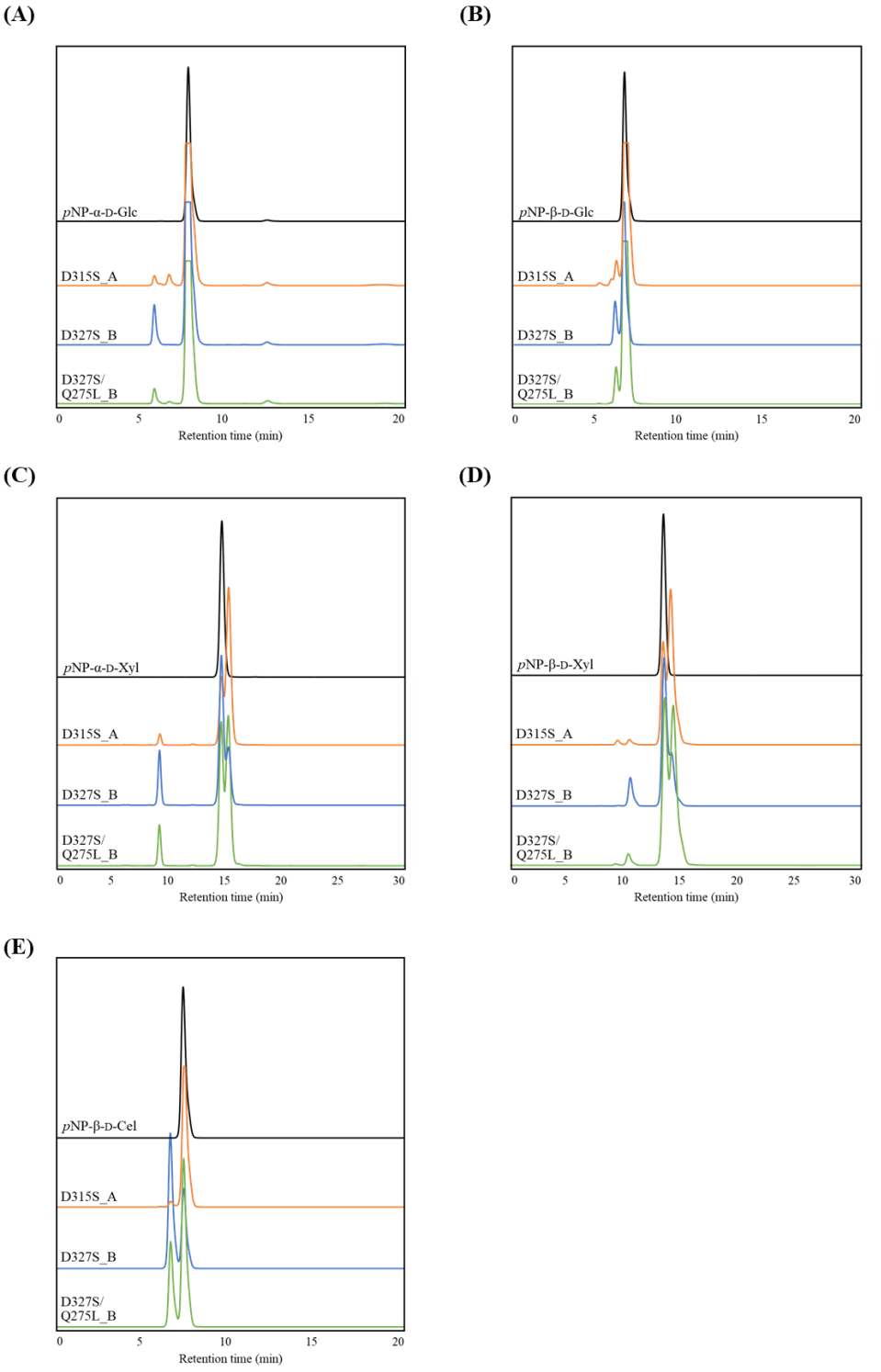


Fig. 5-5. RP-HPLC chromatograms of the reaction mixture of D315S_A, D327S_B, and Q275L/D327S_B with *p*NP α-D-Glc (A), *p*NP β-D-Glc (B), *p*NP α-D-Xyl (C), *p*NP β-D-Xyl (E), and *p*NP β-D-Cel (E). The Reaction was performed in 20 mM β-L-Glc-F, 10 mM acceptors (*p*NP α-D-Glc, *p*NP β-D-Glc, *p*NP α-D-Xyl, *p*NP β-D-Xyl, and *p*NP β-D-Cel), enzymes (D315S, 5.14 μM; D327S, 5.64 μM; Q275L/D327S, 6.68 μM) in 40 mM sodium acetate buffer (5.5) at 35℃. The mobile phase was methanol/water (20/80, v/v) at a flow rate of 1 mL/min.

Table 5-3. Glycosynthase reaction by D315S_A with β-L-Glc-F

The reaction was performed in 20 mM β-L-Glc-F, 10 mM each acceptor (*p*NP α-D-Glc, *p*NP β-D-Glc, *p*NP α-D-Xyl, *p*NP β-D-Xyl, or *p*NP β-D-Cel), 6.42 μM D315S_A, and 40 mM sodium acetate buffer (pH 5.5) at 35°C for 20 h. The ratio was calculated as the concentration of the product to the concentration of the acceptor (10 mM). "trace" indicates that concentration of product is below the quantitative range, which was 40 μM or more.

		Product				
*p*NP α-D-Glc		1	2	3 (trace)	4	5
	Concentration (mM)	0.22 ± 0.0002	0.25 ± 0.012	0.007 ±0.016	0.10 ±0.0004	0.09 ±0.009
	Ratio (%)	2.181 ± 0.0125	2.455 ± 0.0169	0.069± 0.0004	1.051±0.0059	0.855±0.0093

*p*NP β-D-Glc		1	2	3
	Concentration (mM)	0.066 ± 0.0025	0.11 ± 0.007	0.542 ± 0.016
	Ratio (%)	0.663 ± 0.0246	1.072 ± 0.0706	5.417 ± 0.1639

*p*NP α-D-Xyl		1	2	3
	Concentration (mM)	0.3594 ± 0.015	0.0264 ± 0.0011	8.716 ± 0.436
	Ratio (%)	3.594 ± 0.1526	0.264 ± 0.0111	87.16 ± 4.364

*p*NP β-D-Xyl		1	2	3
	Concentration (mM)	0.1302 ± 0.007	0.1767 ± 0.008	6.745 ± 0.293
	Ratio (%)	1.303 ± 0.0647	1.767 ± 0.0499	67.447 ± 2.927

*p*NP β-D-Cello		1 (trace)	2
	Concentration (mM)	0.0266 ± 0.0003	0.3269 ± 0.0016
	Ratio (%)	0.266 ± 0.003	3.269 ± 0.016

Table 5-4. Glycosynthase reaction by D327S_B with β-L-Glc-F

The reaction was performed in 20 mM β-L-Glc-F, 10 mM each acceptor (*p*NP α-D-Glc, *p*NP β-D-Glc, *p*NP α-D-Xyl, *p*NP β-D-Xyl, or *p*NP β-D-Cel), 7.05 μM D327S_B, and 40 mM sodium acetate buffer (pH 5.5) at 35°C for 20 h. The ratio was calculated as the concentration of the product to the concentration of the acceptor (10 mM). "trace" indicates that concentration of product is below the quantitative range, which was 40 μM or more.

		Product				
*p*NP α-D-Glc		1	2 (trace)	3 (trace)	4	5
	Concentration (mM)	0.764 ± 0.0027	0.0067 ± 0.0008	0.0073 ± 0.0004	0.0983 ± 0.006	0.0689 ± 0.001
	Ratio (%)	7.636 ± 0.027	0.0676 ± 0.0002	0.0734 ± 0.0004	0.995 ± 0.007	0.691 ± 0.011

*p*NP β-D-Glc		1	2	3
	Concentration (mM)	0.1 ± 0.0003	-	1.778 ± 0.035
	Ratio (%)	0.096 ± 0.003	-	17.78 ± 0.351

*p*NP α-D-Xyl		1	2	3
	Concentration (mM)	1.357 ± 0.022	-	2.921 ± 0.034
	Ratio (%)	13.568 ± 0.22	-	29.22 ± 0.344

*p*NP β-D-Xyl		1	2	3
	Concentration (mM)	0.0151 ± 0.0009	1.131 ± 0.051	2.301 ± 0.085
	Ratio (%)	0.151 ± 0.0087	11.314 ± 0.511	23.01 ± 0.848

*p*NP β-D-Cello		1	2
	Concentration (mM)	-	6.73 ± 0.12
	Ratio (%)	-	67.296 ± 1.198

Table 5-5. Glycosynthase reaction by Q275L/D327S_B with β-L-Glc-F

The reaction was performed in 20 mM β-L-Glc-F, 10 mM each acceptor (*p*NP α-D-Glc, *p*NP β-D-Glc, *p*NP α-D-Xyl, *p*NP β-D-Xyl, or *p*NP β-D-Cel), 5.78 μM Q275L/D327S_B, and 40 mM sodium acetate buffer (pH 5.5) at 35°C for 20 h. The ratio was calculated as the concentration of the product to the concentration of the acceptor (10 mM). "trace" indicates that concentration of product is below the quantitative range, which was 40 μM or more.

		Product				
*p*NP α-D-Glc		1	2	3 (trace)	4	5
	Concentration (mM)	0.292 ± 0.002	0.053 ± 0.0003	0.005 ± 0.0001	0.107 ± 0.0011	0.024 ± 0.0005
	Ratio (%)	2.958 ± 0.017	0.537 ± 0.004	0.047 ± 0.0004	1.08 ± 0.011	0.239 ± 0.005

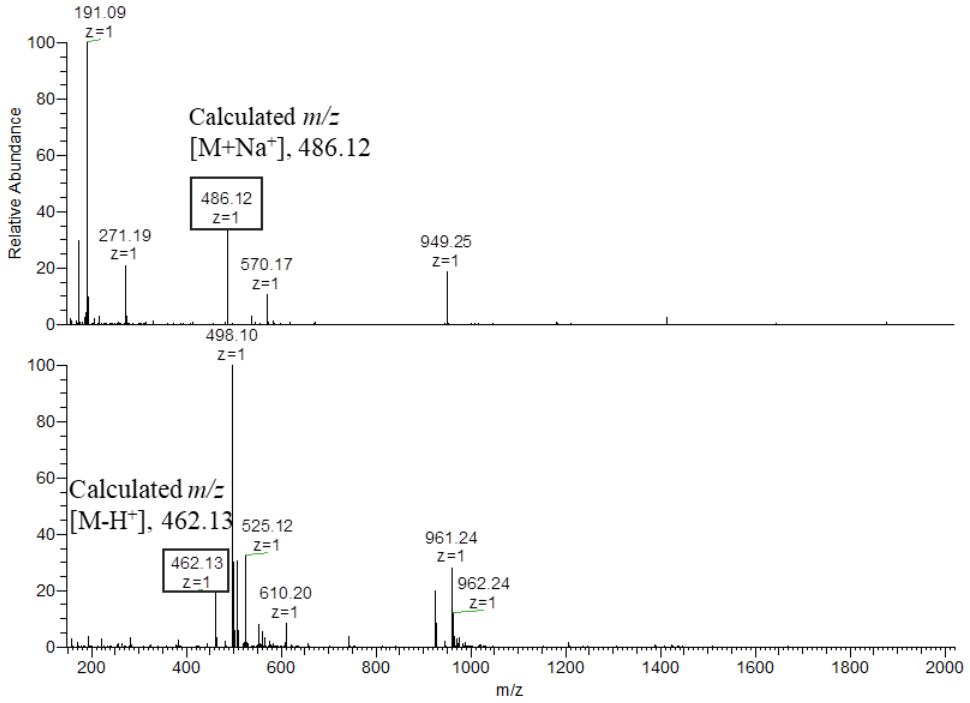
*p*NP β-D-Glc		1	2	3
	Concentration (mM)	0.015 ± 0.0004	-	0.63 ± 0.0049
	Ratio (%)	0.148 ± 0.004	-	6.298 ± 0.049

*p*NP α-D-Xyl		1	2	3
	Concentration (mM)	0.992 0.0272	-	6.415 0.165
	Ratio (%)	9.916 ± 0.272	-	64.149 ± 1.645

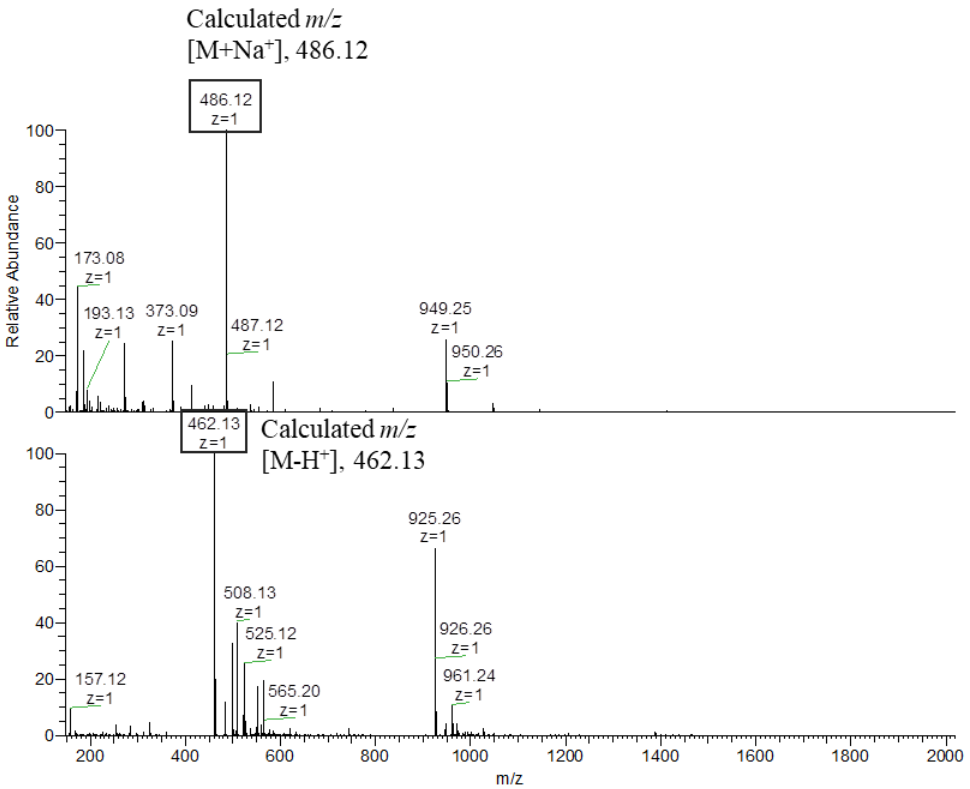
*p*NP β-D-Xyl		1	2	3
	Concentration (mM)	0.0254 0.0051	0.292 0.0139	5.616 0.1445
	Ratio (%)	0.254 ± 0.0512	2.919 ± 0.139	56.158 ± 1.445

*p*NP β-D-Cello		1	2
	Concentration (mM)	-	3.392 0.0788
	Ratio (%)	-	33.915 ± 0.788

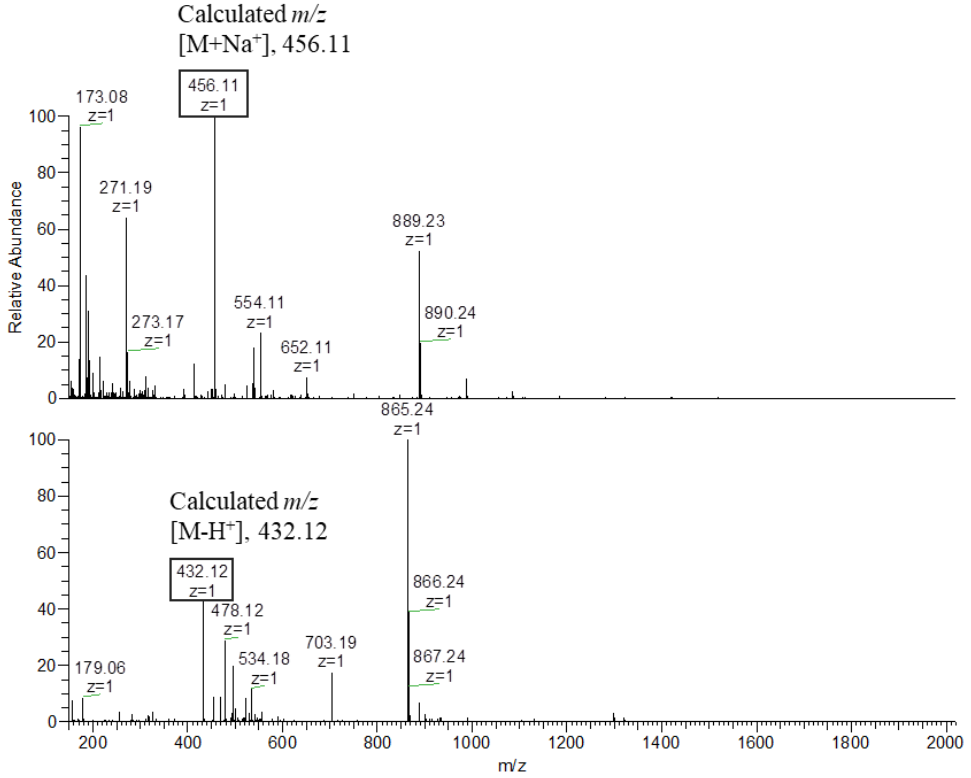
(A)


191.09
z=1
Calculated *m/z*
[M+Na⁺], 486.12
486.12
z=1
271.19
z=1
570.17
z=1
949.25
z=1
498.10
z=1
Calculated *m/z*
[M-H⁺], 462.13
525.12
z=1
462.13
z=1
610.20
z=1
961.24
z=1
962.24
z=1
Relative Abundance
m/z

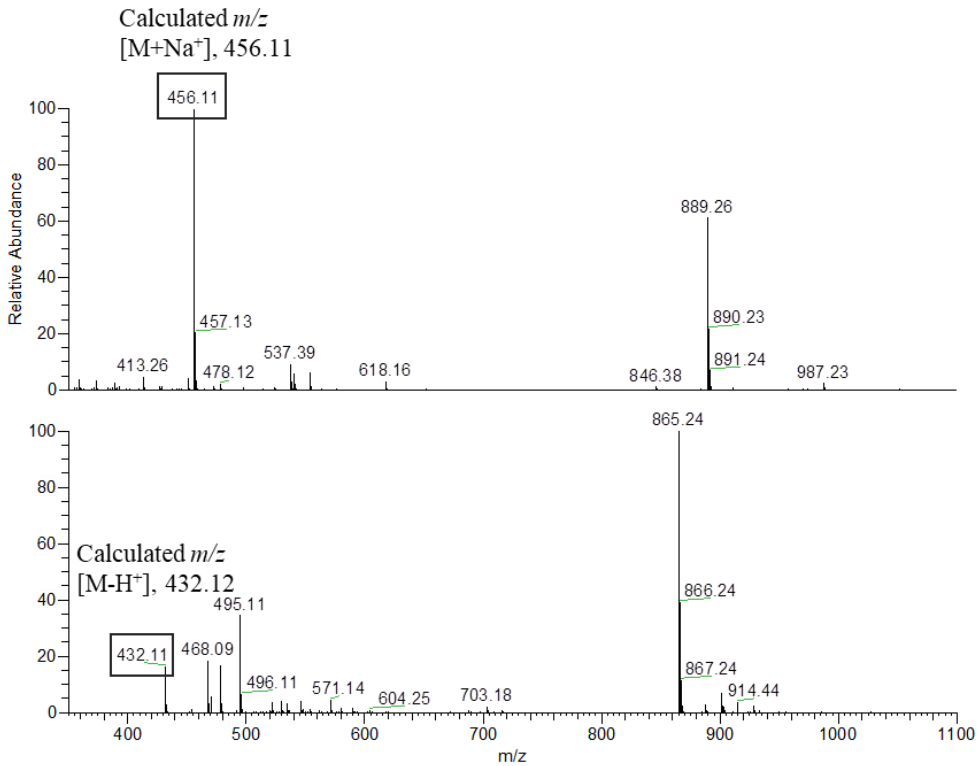
(B)


Calculated *m/z*
[M+Na⁺], 486.12
486.12
z=1
173.08
z=1
193.13
z=1
373.09
z=1
487.12
z=1
949.25
z=1
950.26
z=1
462.13
z=1
Calculated *m/z*
[M-H⁺], 462.13
925.26
z=1
508.13
z=1
525.12
z=1
926.26
z=1
157.12
z=1
565.20
z=1
961.24
z=1
Relative Abundance
m/z

(C)


Calculated m/z
[M+Na+], 456.11
456.11
z=1
173.08
z=1
271.19
z=1
273.17
z=1
554.11
z=1
652.11
z=1
889.23
z=1
890.24
z=1
Relative Abundance
865.24
z=1
Calculated m/z
[M-H+], 432.12
432.12
z=1
478.12
z=1
534.18
z=1
703.19
z=1
179.06
z=1
866.24
z=1
867.24
z=1
m/z

(D)


Calculated m/z
[M+Na+], 456.11
456.11
457.13
413.26
478.12
537.39
618.16
889.26
890.23
891.24
846.38
987.23
Relative Abundance
865.24
Calculated m/z
[M-H+], 432.12
432.11
468.09
495.11
496.11
571.14
604.25
703.18
866.24
867.24
914.44
m/z

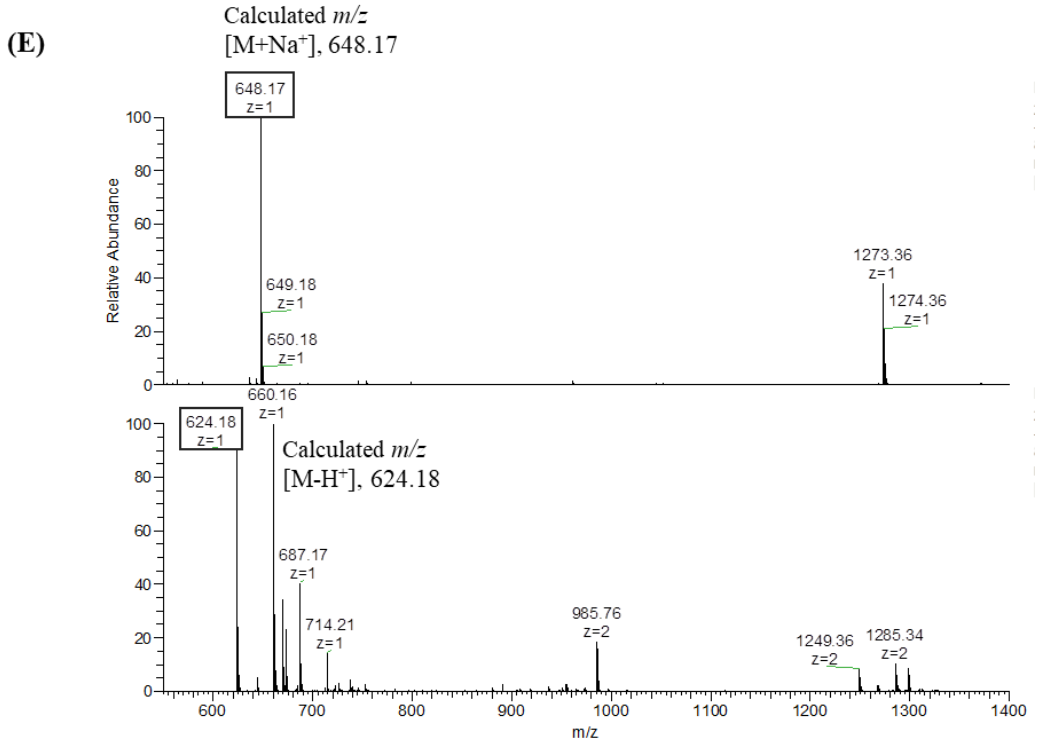


Fig. 5-6. ESI-MS spectrum of glycosynthase reaction products by D315S_A and D327S_B. A; product of *p*NP α-D-Glc with D327S_B, **B**; product of *p*NP β-D-Glc with D327S_B, **C**; product of *p*NP α-D-Xyl with D315S_A, **D**; product of *p*NP β-D-Xyl with D315S_A, **E**; product of *p*NP β-D-Cel with D327S_B. Upper and lower spectrums were obtained by positive and negative mode detections, respectively.

Table 5-6. ^{13}C NMR spectrum and ^{1}H NMR spectrum assignment for α-L-Glcp-(1 → 4)-α-D-Glcp-O-*p*NP

	δC (ppm)	δH (ppm)		J (Hz)	
1	97.4	5.64	d	3.6	
2	71.1	3.48	m		
3	71.8	3.77	t	8.9	
4	79.3	3.47	m		
5	72.5	3.49	m		
6	60.1		m		
1'	99.7	4.84	d	3.8	
2'	72	3.21	dd	3.8	9.7
3'	73.2	3.37	t	9.2	
4'	70.2	3.08	t	9.4	
5'	72.9	3.77	m		
6'	60.8		m		
*p*NP	162.1				
*p*NP	125.8	8.19	d	9.29	
*p*NP	116.6	7.27	d	9.29	
*p*NP	141.8				

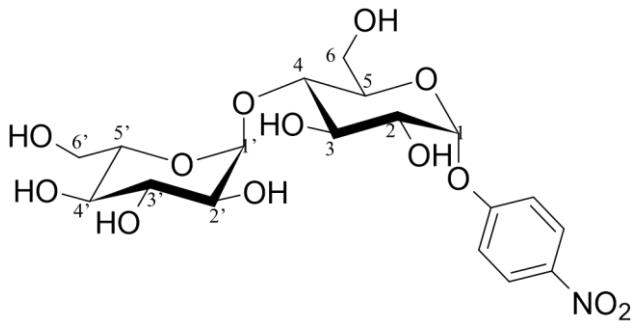


Fig. 5-7. Structure of α-L-Glcp-(1 → 4)-α-D-Glcp-O-*p*NP and the number of each carbon

Table 5-7. ^{13}C NMR spectrum and ^{1}H NMR spectrum assignment for α-L-Glcp-(1 → 4)-β-D-Glcp-O-*p*NP

	δC (ppm)	δH (ppm)		J (Hz)	
1	100.2	5.17	d	7.8	
2	73.9	3.66	t	8.2	
3	75.2	3.8	t	9.4	
4	78.3	3.7	t	9.3	
5	76.3	3.83	m		
6	60.9		m		
1'	100.4	5.09	d	3.8	
2'	72.3	3.64	dd	3.8	9.9
3'	73.7	3.78	t	9.6	
4'	70.2	3.54	t	9.6	
5'	72.8	4.02	m		
6'	61.1		m		
*p*NP	162.6				
*p*NP	126.8	8.04	d	9.18	
*p*NP	117.8	7.11	d	9.17	
*p*NP	143.0				

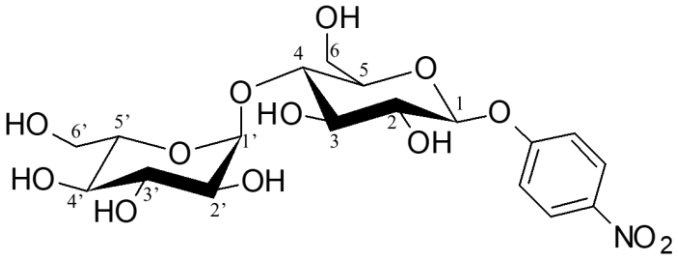


Fig. 5-8. Structure of α-L-Glcp-(1 → 4)-β-D-Glcp-O-*p*NP and the number of each carbon

Table 5-8. ^{13}C NMR spectrum and ^{1}H NMR spectrum assignment for α-L-Glcp-(1 → 4)-α-D-Xyl-O-*p*NP

	δC (ppm)	δH (ppm)		J (Hz)	
1	97.06	5.69	s		
2	70.65	3.72	m		
3	67.82	3.57	m		
4	82.35	3.73	m		
5	63.04		m	10.1	
1'	99.99	5.05	d	3.7	
2'	72.51	3.26	dd	3.4	9.5
3'	73.53	3.48	m		
4'	70.25	3.13	t	9.4	
5'	72.81	3.78	m		
6'	60.84	11.1	d	11.1	
*p*NP	161.8				
*p*NP	125.8	8.19	d	9.18	
*p*NP	116.9	7.25	d	9.21	
*p*NP	141.9				

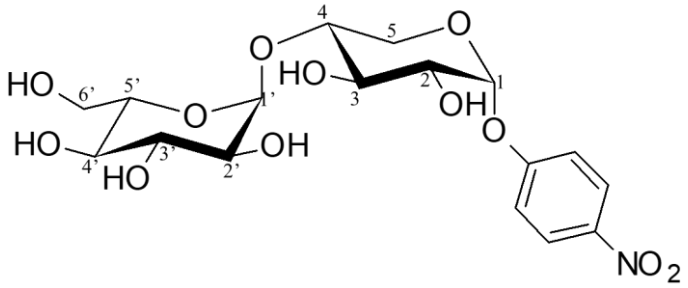


Fig. 5-9. Structure of α-L-Glcp-(1 → 4)-α-D-Xylp-O-*p*NP and the number of each carbon

Table 5-9. ^{13}C NMR spectrum and ^{1}H NMR spectrum assignment for α-L-Glcp-(1 → 3)-β-D-Xyl-O-*p*NP

	δC (ppm)	δH (ppm)	J (Hz)		
1	99.8	5.18	d	7.3	
2	72	3.57	m		
3	84.3	3.47	m		
4	67.7	3.57	m		
5	65.4	3.83	dd	5.1	11.3
1'	99.7	5.05	d	3.8	
2'	72.4	3.24	d	3.8	9.6
3'	73.4	3.47	m		
4'	70.1	3.11	t	9.4	
5'	72.3	3.76	ddd	1.8	5.8
6'	60.7	3.64	dd		
*p*NP	162				
*p*NP	125.8	8.19	d	9.2	
*p*NP	116.6	7.2	d	9.2	
*p*NP	141.9				

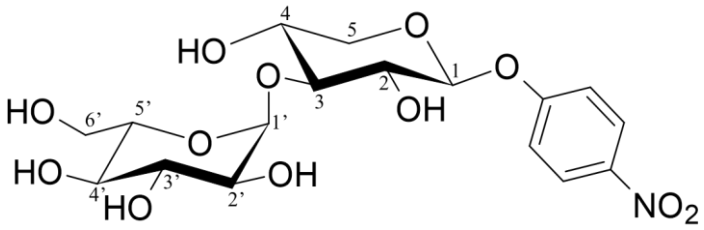


Fig. 5-10. Structure of α-L-Glcp-(1 → 3)-β-D-Xylp-O-*p*NP and the number of each carbon

Table 5-10. ^{13}C NMR spectrum and ^{1}H NMR spectrum assignment for α-L-Glcp-(1 → 4)-β-D-Glcp-(1 → 4)-β-D-Glcp-O-*p*NP

	δC (ppm)	δH (ppm)		J (Hz)	
1	100.29	5.17	d	7.76	
2	73.34	2.08	m		
3	74.92	3.78	t	8.75	
4	79.3	7.74	t	9.40	
5	75.91	3.81	m		
6	60.8		m		
1'	103.35	4.57	d	3.89	
2'	74.37	3.40	t	3.88	9.79
3'	75.22	3.70	t	9.40	
4'	78.34	3.62	t	9.6	
5'	76.1	3.65	m		
6'	60.92		m		
1’	100.36	5.04	d	7.95	
2’	72.33	3.61	dd	9.33	
3’	73.77	3.74	t	7.96	
4’	70.14	3.52	t	9.79	
5’	72.75	4.04	m		
6’	61		m		
1 (*p*NP)	162.6				
2, 6 (*p*NP)	126.9	8.08	d	9.18	
3, 5 (*p*NP)	117.3	7.13	d	9.21	
4 (*p*NP)	143.1				

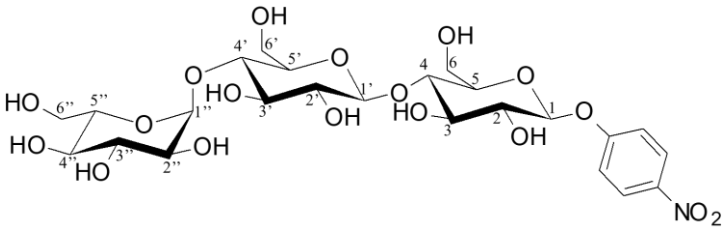


Fig. 5-11. Structure of α-L-Glcp-(1 → 4)-β-D-Glcp-(1 → 4)-β-D-Glcp-O-*p*NP and the number of each carbon

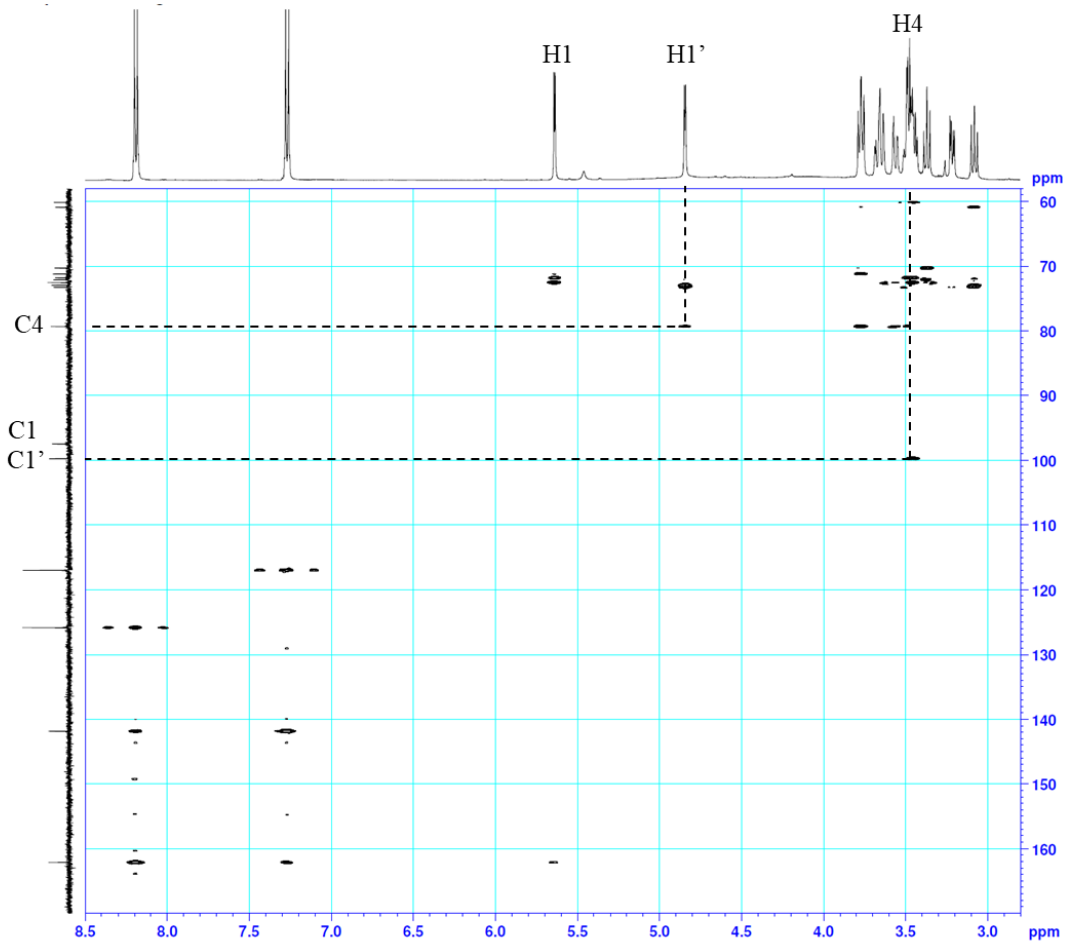


Fig. 5-12. HMBC spectrum of α-L-Glcp-(1 → 4)-α-D-Glcp-O-*p*NP. The dashed lines indicate the correlation between C and H in HMBC.

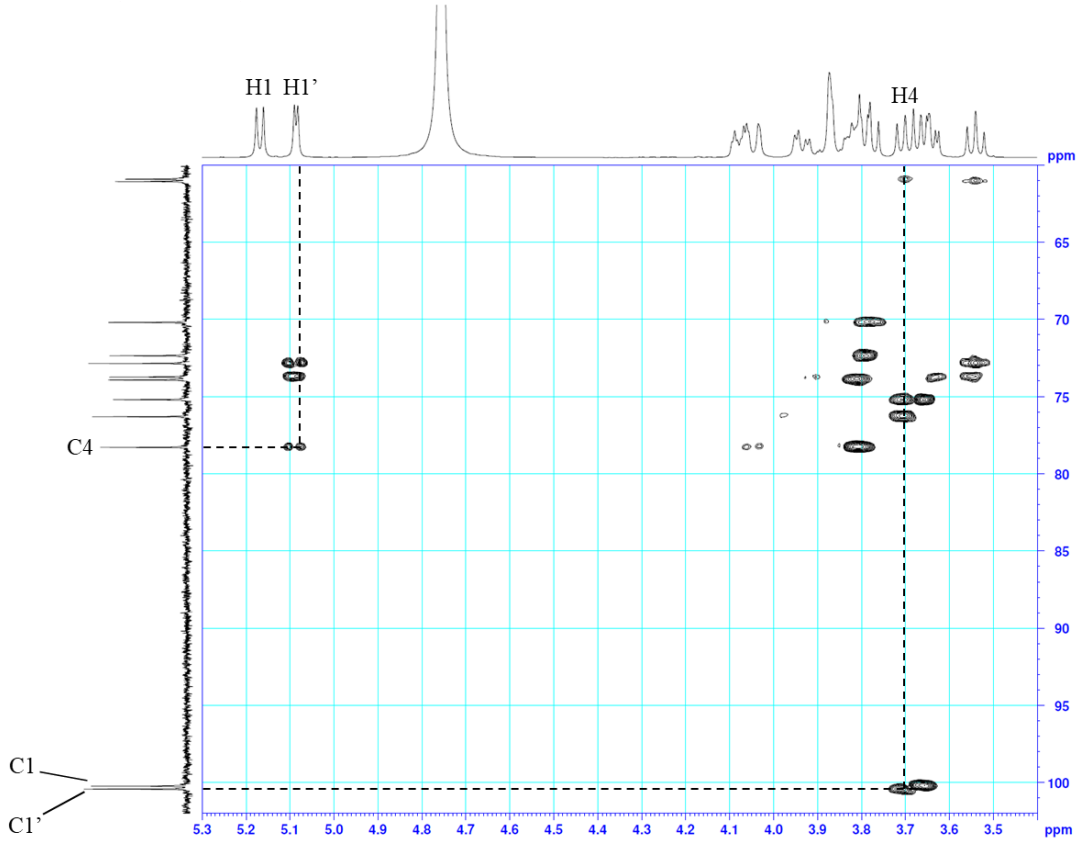


Fig. 5-13. HMBC spectrum of α-L-Glcp-(1 → 4)-β-D-Glcp-O-*p*NP. The dashed lines indicate the correlation between C and H in HMBC.

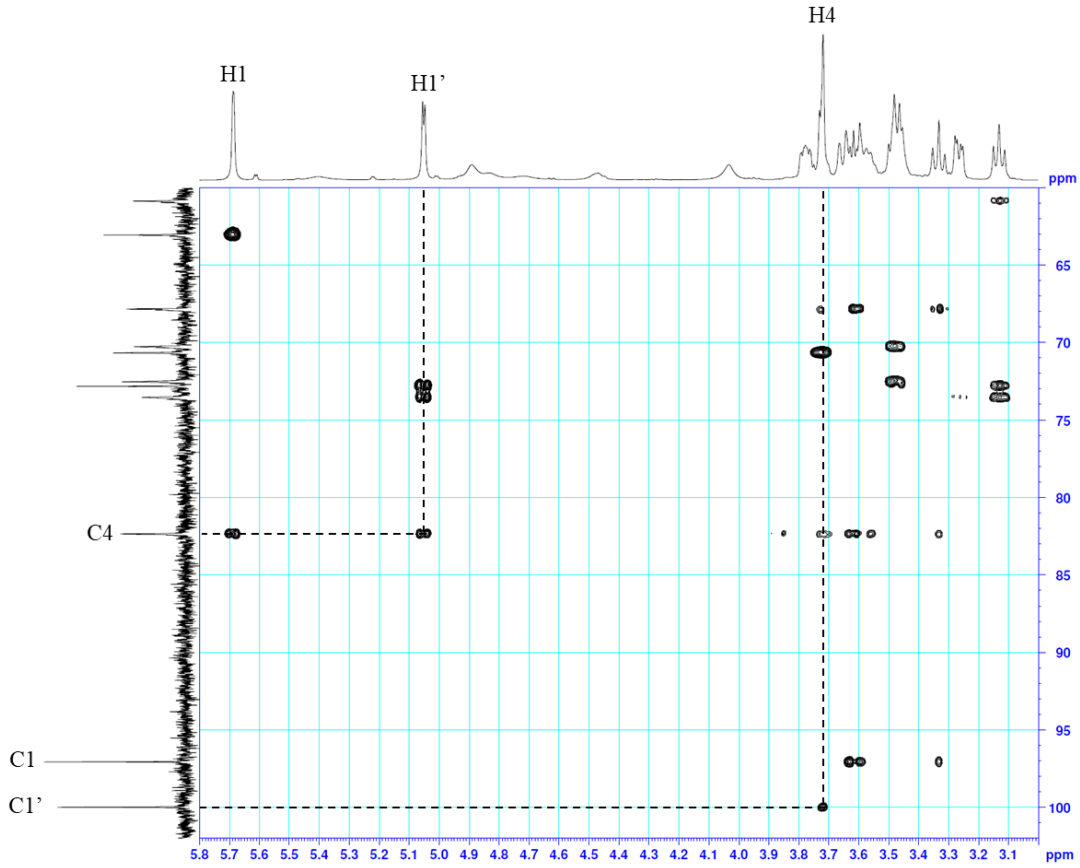


Fig. 5-14. HMBC spectrum of α-L-Glcp-(1 → 4)-α-D-Xylp-O-*p*NP. The dashed lines indicate the correlation between C and H in HMBC.

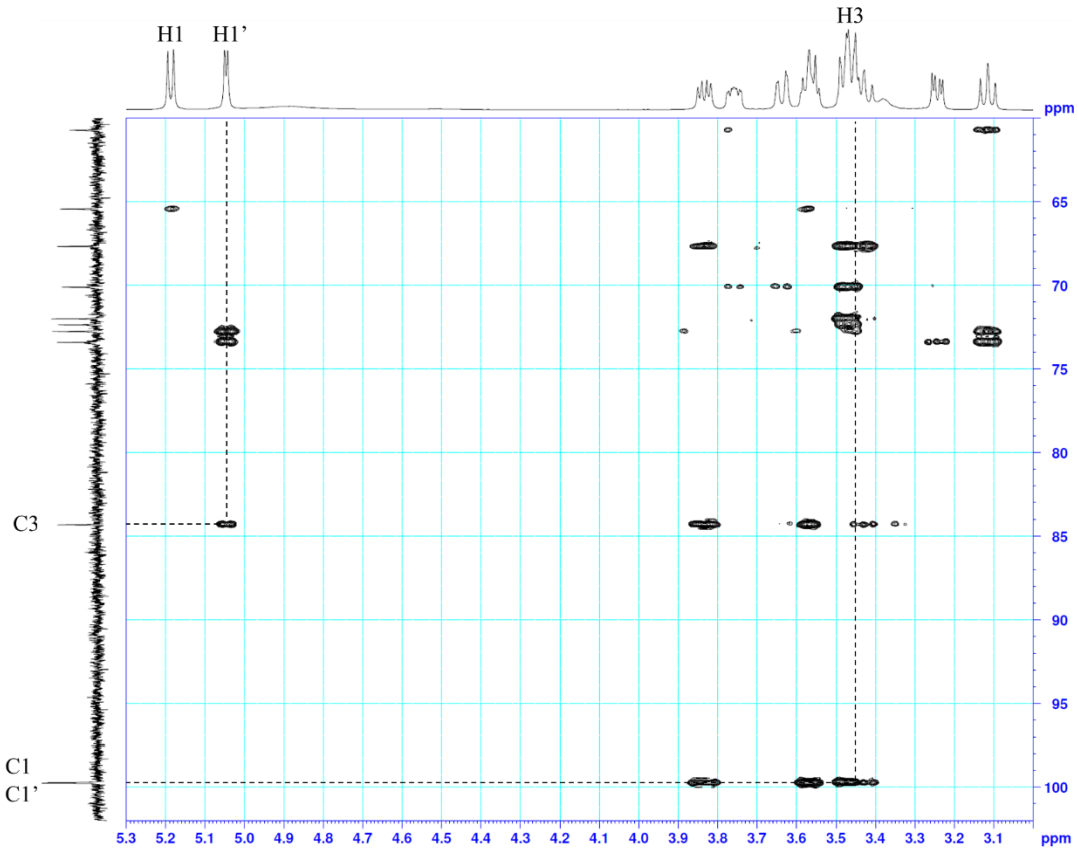


Fig. 5-15. HMBC spectrum of α-L-Glcp-(1 → 3)-β-D-Xylp-O-*p*NP. The dashed lines indicate the correlation between C and H in HMBC.

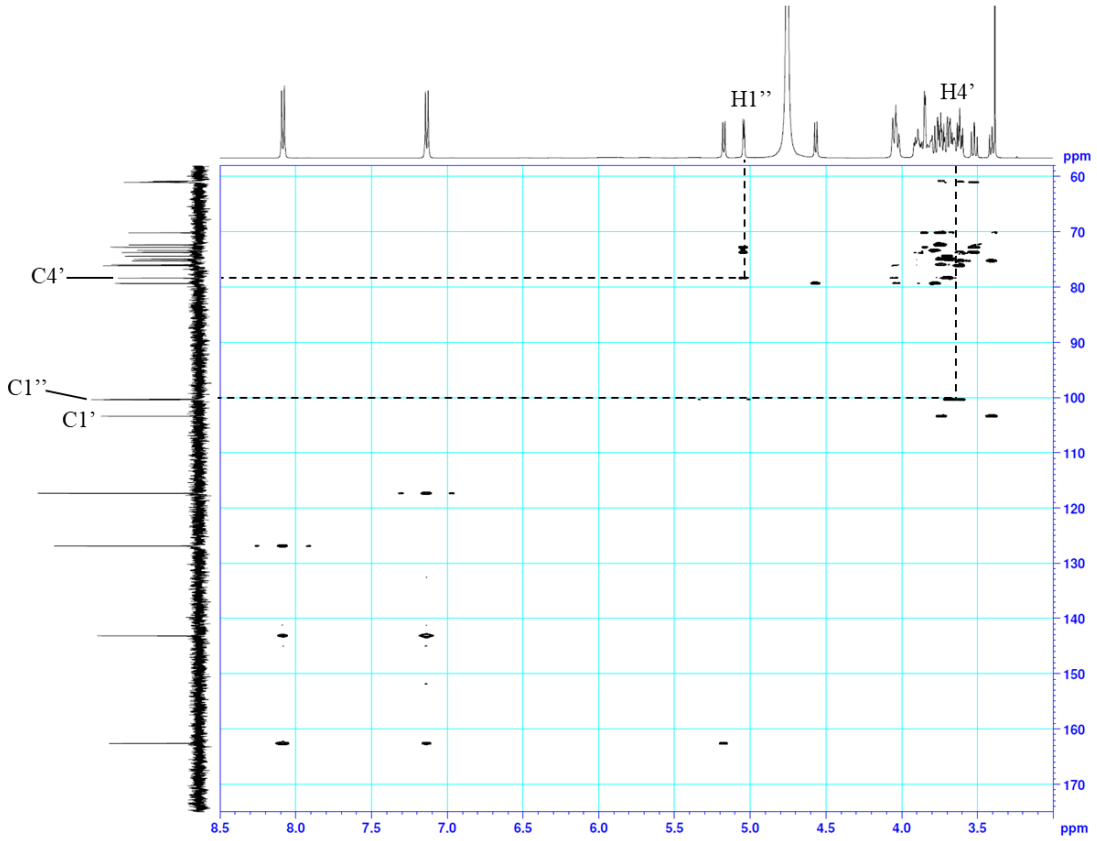


Fig. 5-16. HMBC spectrum of α-L-Glcp-(1 → 4)-β-D-Glcp-(1 → 4)-β-D-Glcp-O-*p*NP. The dashed lines indicate the correlation between C and H in HMBC.

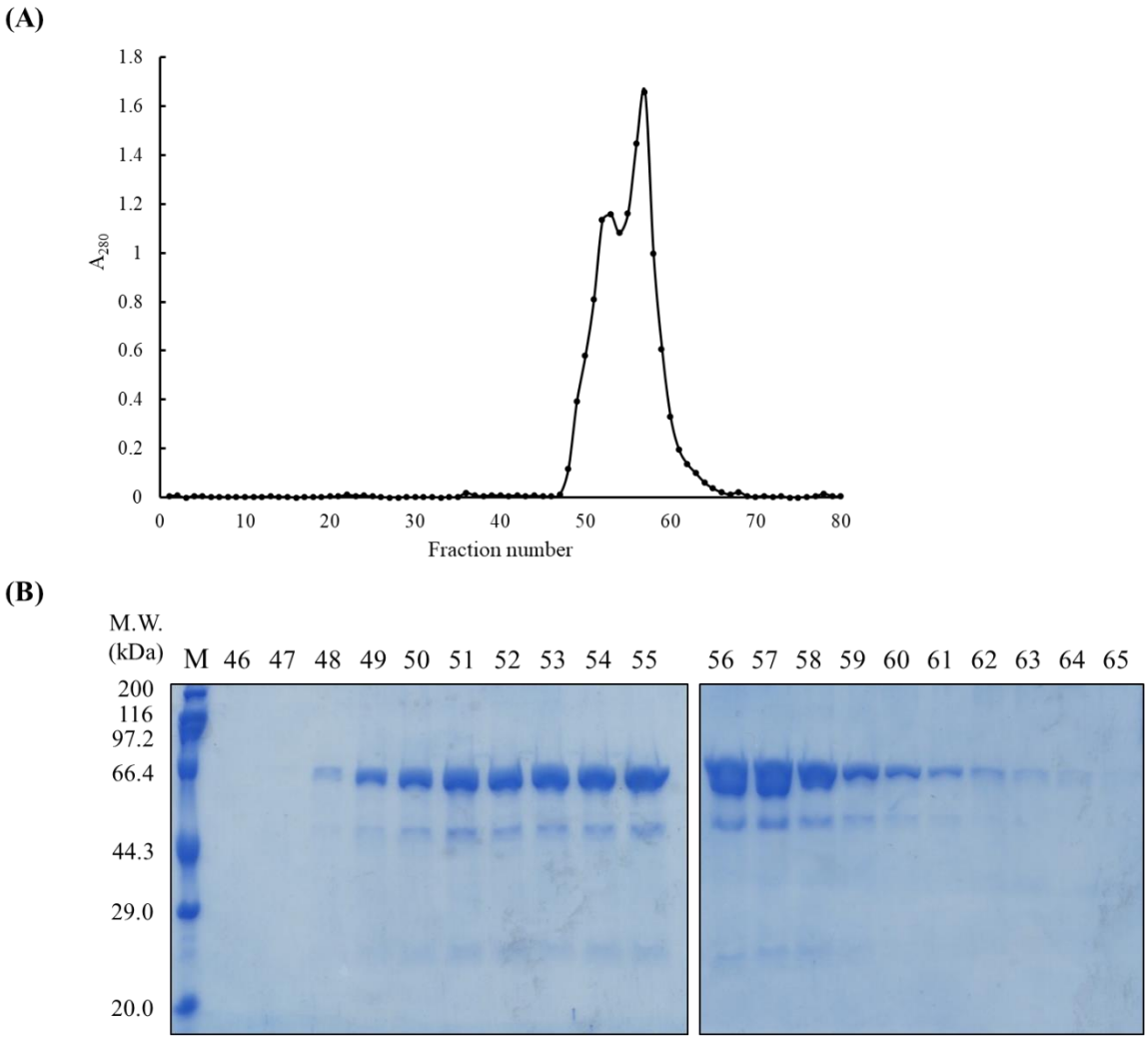


Fig. 5-17. Purification of S132D/D327S_B for crystallization by gel filtration column chromatography. A, Chromatogram of gel filtration column chromatography. Blue; A_{280}. **B**, SDS-PAGE of each fraction. The numbers shown in the upper part indicate the fraction numbers. Fractions 55-63 were collected.

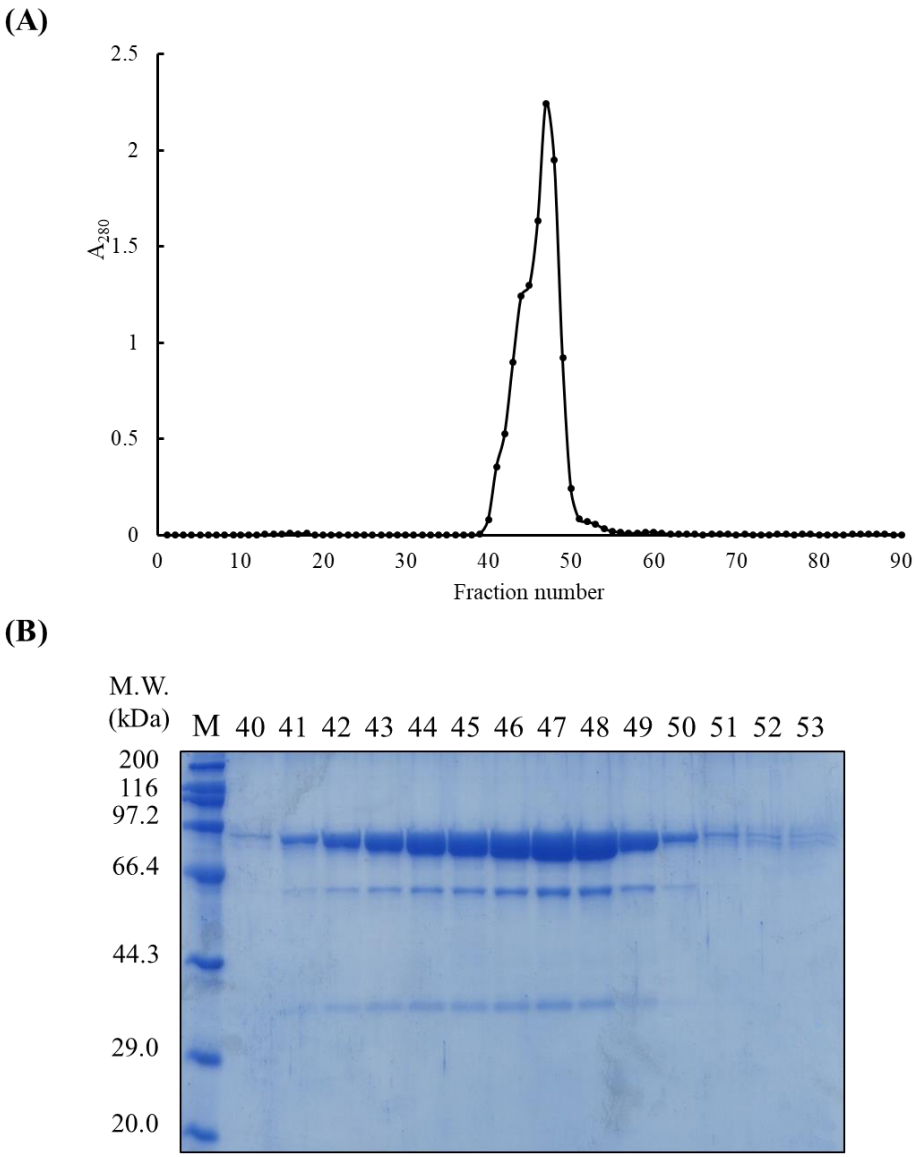


Fig. 5-18. Purification of Q275L/D327S_B for crystallization by gel filtration column chromatography. A, Chromatogram of gel filtration column chromatography. Blue; A_{280}. **B**, SDS-PAGE of each fraction. The numbers shown in the upper part indicate the fraction numbers. Fractions 45-50 were collected.

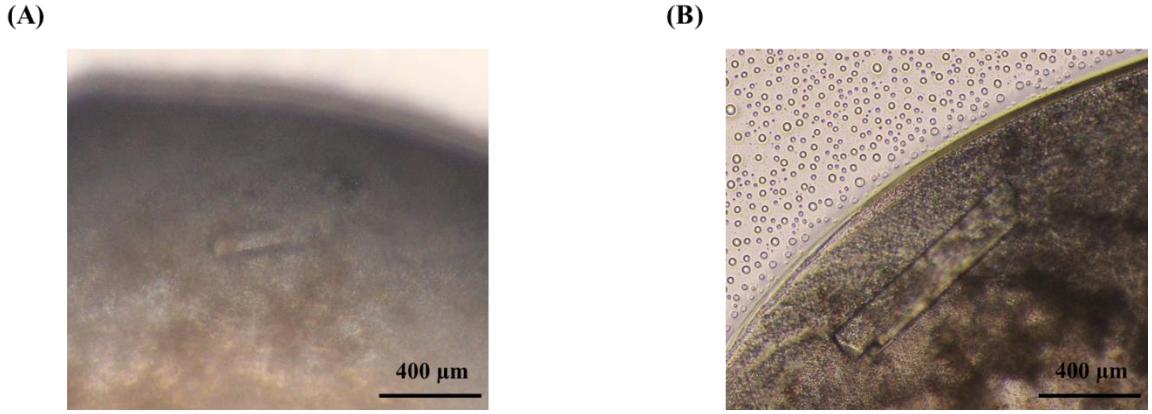


Fig. 5-19. Crystals used to analyze the complex structure of S132D/D327S_B (A) and Q275L/D327S_B (B). A, Cystral of ClAgl29B S132D/D327S, obtained in a drop consisting of 1 µL enzyme (24 a.u; 280 nm) and 2 µL of reservoir solution [5% (w/v) 2-propanol, 0.1 M citrate-NaOH (pH 3.5), and 6% (w/v) PEG 20.000 at 20°C for over 2 weeks. Soaking was carried out using cryoprotectant (5%w-propanol, 0.1 M citrate-NaOH (pH 3.5), 7% PEG 20,000 and 30 mM *p*NP 3-O-α-L-Glc α-L-Glc). **B**, Cystral of ClAgl29B Q275L/D327S, obtained in a drop consisting of 1 µL enzyme (32 a.u; 280 nm) and 1 µL of reservoir solution [5% (w/v) 2-propanol, 0.1 M citrate-NaOH (pH 3.5), and 5% (w/v) PEG 20.000 at 20°C for over 2 weeks. Soaking was carried out using cryoprotectant (5%w-propanol, 0.1 M citrate-NaOH (pH 3.5), 7% PEG 20,000 and 30 mM *p*NP 3-O-α-L-Glc α-L-Glc).

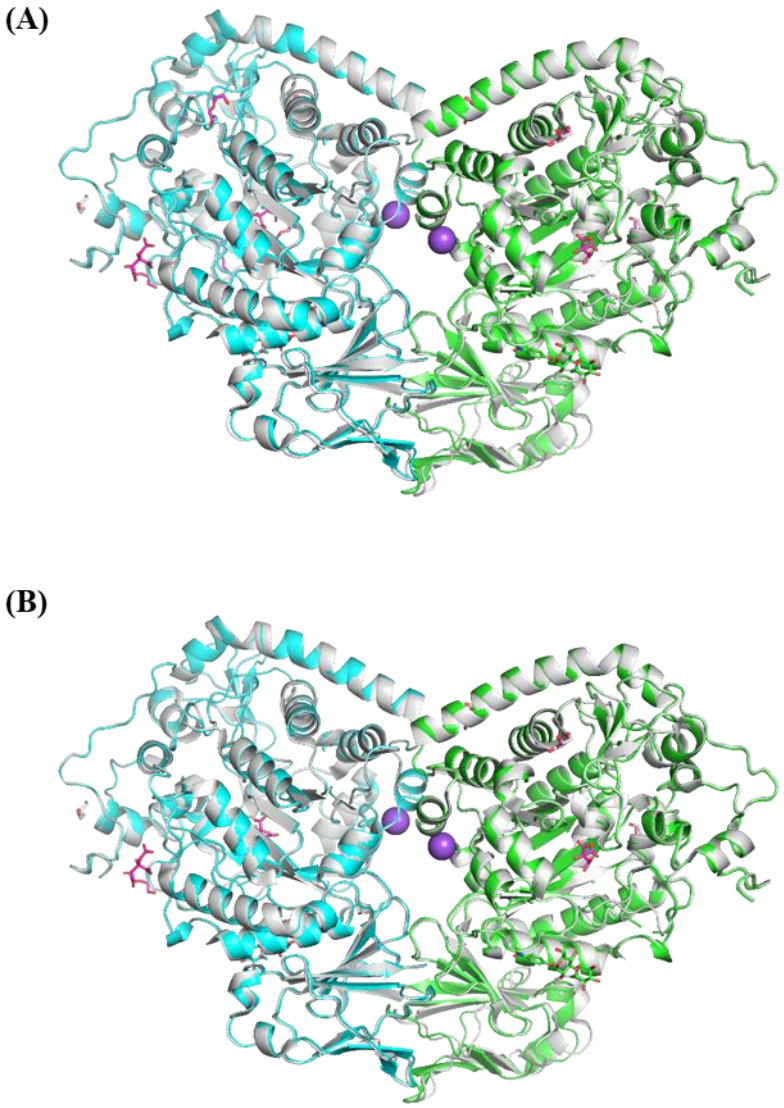


Fig. 5-20. Crystal structures of S132D/D327S_B (A) and Q275L/D327S_B (B).

The structures were superimposed with the wild-type ClAgl29B (gray color, PDB accession, 7xsg). Respective chains A and B are colored green and cyan.

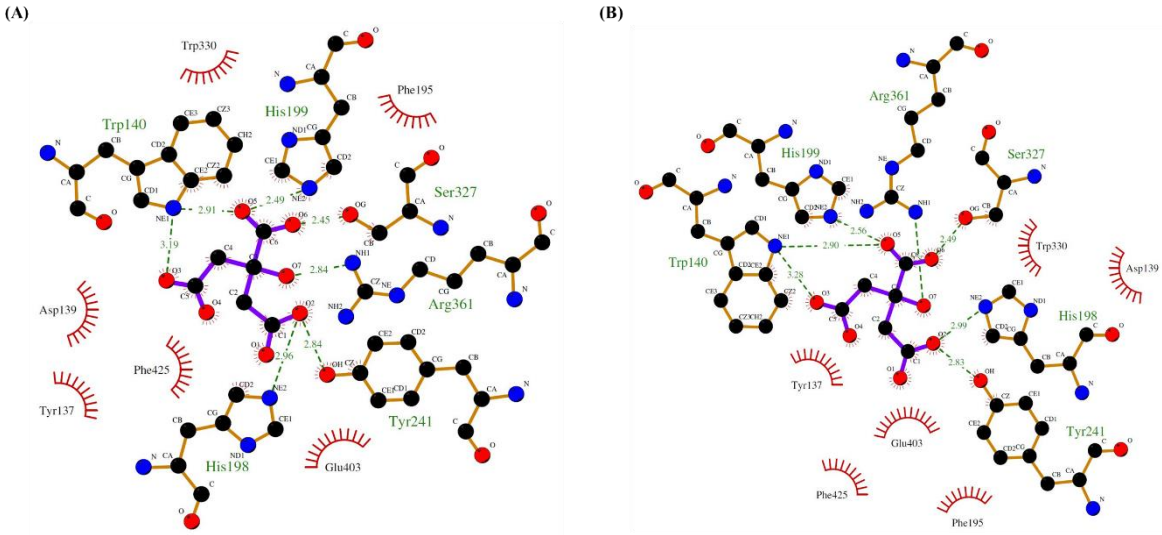


Fig. 5-21. Active–site pocket of S132D/D327S_B (A) and Q275L/D327S_B (B) bound citrate. Diagram of the citrate-ClAgl29B S132D/D327S (A) and ClAgl29B Q275L/D327S (B) interactions. Hydrogen bonds are shown as green dotted lines, while the spoked arcs represent protein residues making nonbonded contact with citrate. The diagram was drawn by LigPlot^{+}.

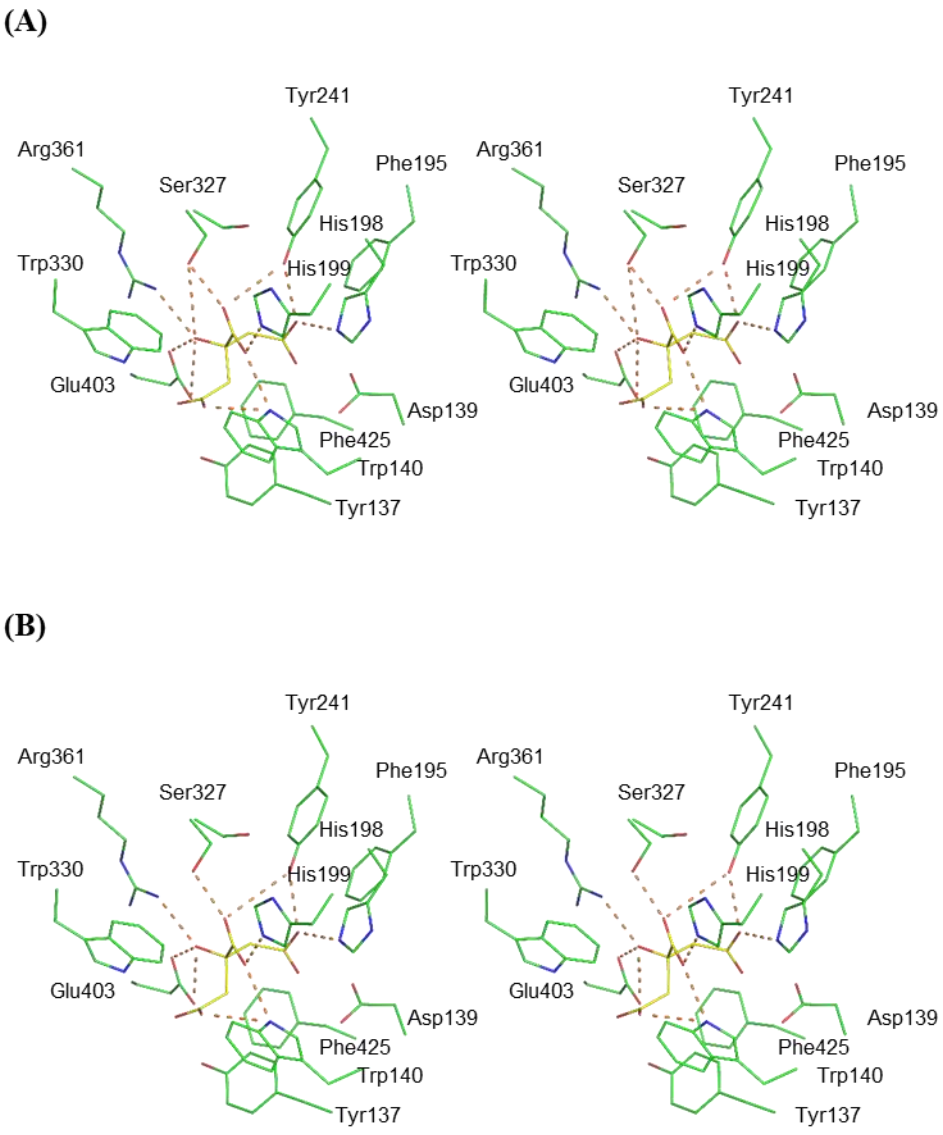


Fig. 5-22. Stereo view of the citrate binding to S132D/D327S_B (A) and Q275L/D327S_B (B).

Residues located within 4 Å of the citrate molecule (yellow) are indicated. Polar contacts at hydrogen bond distances are indicated by dashed lines. Schematic representation is shown in Fig. 5-21.

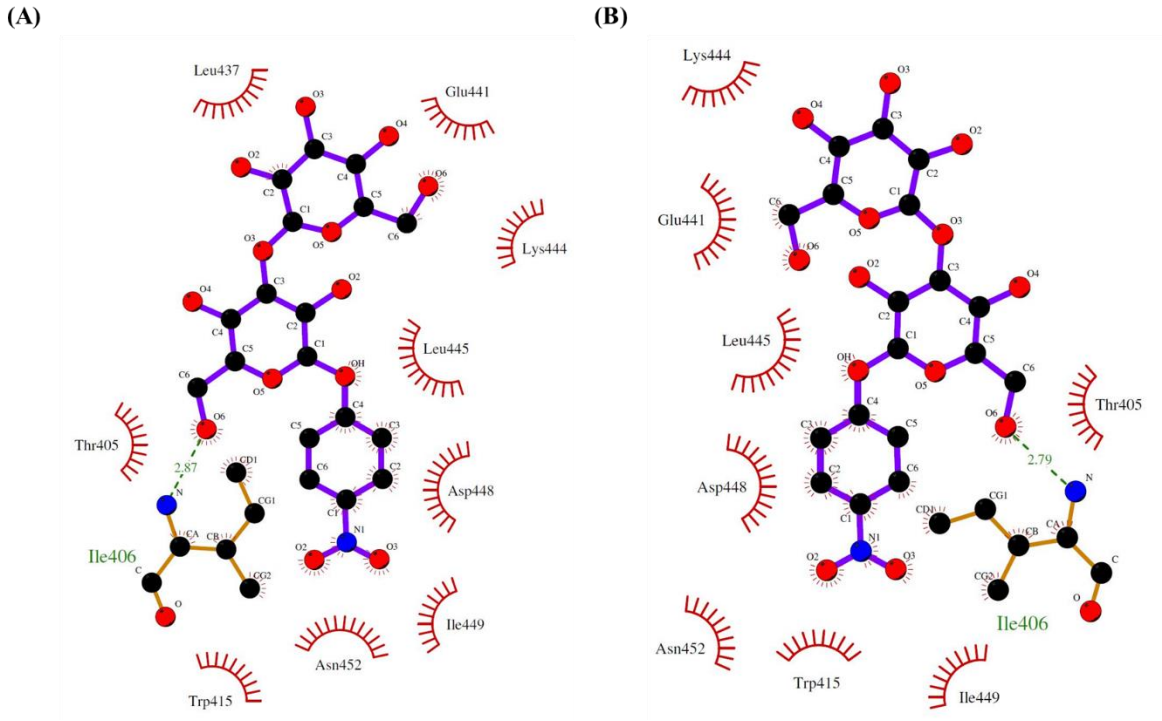


Fig. 5-23. Active–site pocket of S132D/D327S_B (A) and Q275L/D327S_B (B) bound *p*NP 3-*O*-α-L-Glc α-L-Glc. Diagram of *p*NP 3-O-α-L-Glc α-L-Glc bound to ClAgl29B S132D/D327S (A) and ClAgl29B Q275L/D327S (B) interactions. Hydrogen bonds are shown as green dotted lines, while the spoked arcs represent protein residues making nonbonded contact with citrate. The diagram was drawn by LigPlot^{+}.

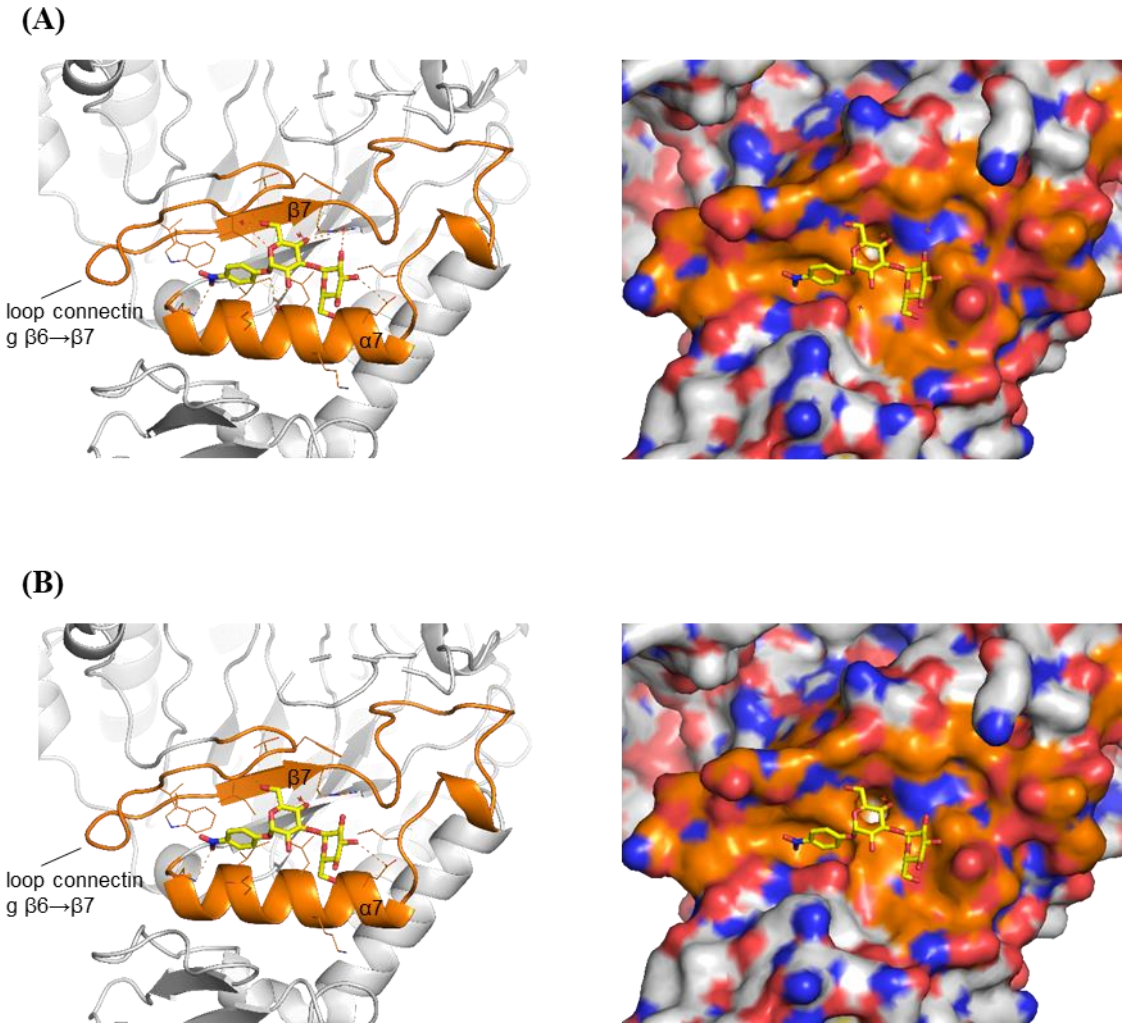


Fig. 5-24. The *p*NP 3-*O*-α-L-Glc α-L-Glc binding to S132D/D327S_B (A) and Q275L/D327S_B (B).

Cartoon models of *p*NP 3-*O*-α-L-Glc α-L-Glc binding site (left). Surface representations of the *p*NP 3-*O*-α-L-Glc α-L-Glc binding site (right).

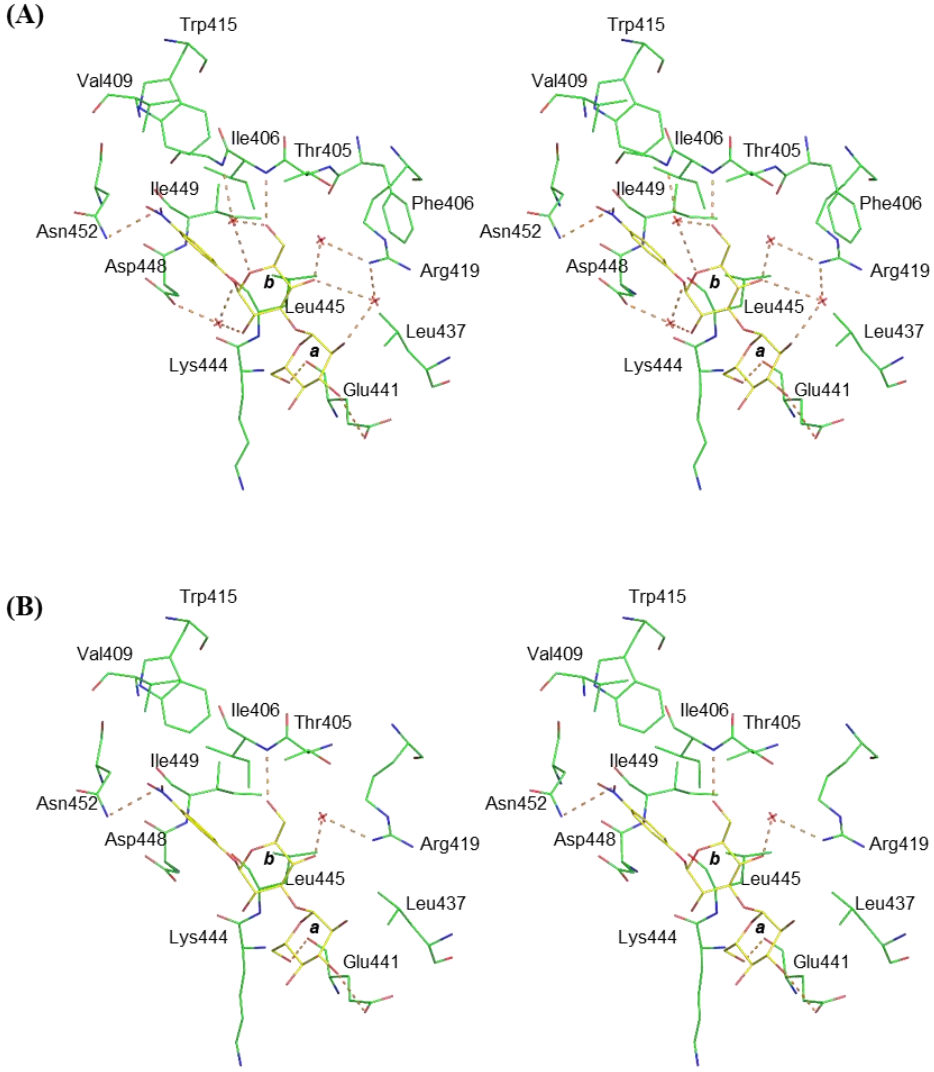


Fig. 5-25. The *p*NP 3-*O*-α-L-Glc α-L-Glc binding to S132D/D327S_B (A) and Q275L/D327S_B (B).

Residues located within 4.5 Å of *p*NP 3-*O*-α-L-Glc α-L-Glc (yellow) are indicated. Asterisks represent water molecule. Hydrogen bond distances are indicated by dashed lines.

4. Discussion

In this chapter, ClAgl29A and ClAgl29B successfully converted to glycosynthase (D315_A and D327S_B), which can catalyze α-L-glucosylation of *p*NP α-D-Glc, *p*NP β-D-Glc, *p*NP α-D-Xyl, *p*NP β-D-Xyl, and *p*NP β-D-Cel. D315_A and D327S_B tended not to accept *p*NP α-D-glucoside well as an acceptor among them, perhaps indicating that ClAgl29A and ClAgl29B have the subsites +1 structure to accept α-D-glucosyl groups. In glycosynthase derived from GH29 α-L-fucosidase, low yields have been a concern due to the instability of β-L-fucosyl fluoride. Moracci and his research group increased the yield, using the more stable β-L-fucosyl azide as an activating sugar donor (Cobucci-Ponzano, B., 2009). However, this study found that glycosynthase derived from α-L-glucosidase using the β-L-Glc-F as a donor gave sufficient yield. The efficiency of D315_A was higher than that of D327S_B, and Q275L/D327S_B showed improved efficiency.

The transglucosylation of ClAgl29A and ClAgl29B using *p*NP α-L-Glc as a substrate produced *p*NP 3-*O*-α-L-Glc α-L-Glc, *p*NP 2-*O*-α-L-Glc α-L-Glc, and *p*NP 6-*O*-α-L-Glc α-L-Glc (Shishiuchi, R., 2022). In Chapter 4, the Q275L_B mutant enzyme exhibited increased transglucosylation efficiency. However, D315_A, D327S_B, and Q275L/D327S_B did not use *p*NP α-L-Glc (Fig. 5-3). This may be because the α-L-glucosyl group of *p*NP α-L-Glc binds more strongly to subsite −1 than β-L-Glc-F, preventing a proper glycosynthase reaction.

D-Glucose, D-xylose, D-fructose, L-sorbose and D-cellobiose can be acted as acceptor of glycosynthases, but α-L-glucosylation efficiency was poor in D315S_A and Q275L/D327S_B, which may indicate that Gln275 is important for binding these monosaccharides and disaccharides to the enzyme.

The structure analysis of ClAgl29B complex with β-L-Glc in Chapter 2 allowed us to understand the structure of α-L-glucosyl binding site, subsite −1, of α-L-glcosidases. However, structure of their plus subsite remained obscure. Thus, it was designed to analyze the crystal structures of S132D/D327S_B and Q275L/D327S_B in a complex of *p*NP 3-*O*-α-L-Glc α-L-Glc. In addition, it was hoped that structural analysis of these mutants would also reveal the mechanism by which the S132D and Q275L mutations, shown in Chapter 4, enhance transglucosylation. Contrary to expectations, citrate was bound to the active site pocket of the obtained crystal structures of S132D/D327S_B and Q275L/D327S_B. The binding was tight, involving several hydrogen bonds and even two salt bridges, suggesting that citrate acts as an inhibitor in the D327S_B. Ser327 formed

hydrogen bonds with the one of the carboxylates in citrate. In the wild-type enzyme, electrical repulsion with Asp327 suggests that citrate is unlikely to enter the active site pocket. In fact, crystal structure analyses of the wild-type enzymes have not yielded a structure in which citrate is bound to the active site pocket.

In both crystal structures, *p*NP 3-*O*-α-L-Glc α-L-Glc bonded to the surface site consisted of a loop connecting β6 and β7, β7 and α7 (Fig. 5-24). This binding site may be a second substrate binding site since several hydrogen bonds and two salt bridges are involved, and similar binding modes were obtained in both enzymes. On the other hand, from the appearance of the binding, the *p*NP group seems to be more tightly bound than sugar residues (Fig. 5-25), surrounded by hydrophobic residues, and it is possible that this binding site is not carbohydrate-specific. A comparison of the structure corresponding to the surface binding site indicated that it was not common in GH29 (Fig. 5-26). ClAgl29A has a similar structure, but even TmaFuc, which has relatively high sequence and structural similarity, has no similar structure.

The present study used the nucleophilic catalytic mutant enzyme to obtain a structure binding of *p*NP 3-*O*-α-L-Glc α-L-Glc at the active site pocket. However, the nucleophilic catalytic mutant may not have been appropriate for this purpose. For example, a mutant enzyme of nucleophilic catalyst of barley α-amylase, which employs the retaining mechanism as same as α-L-glucosidases, has not yielded a crystal structure with productive binding of maltoheptaose (Robert, X., 2005). In the future, it may be necessary to prepare an inactive enzyme with mutations in other residues, such as general acid-base catalytic residues, in order to investigate the plus subsite by X-ray structural analysis of the substrate complex.

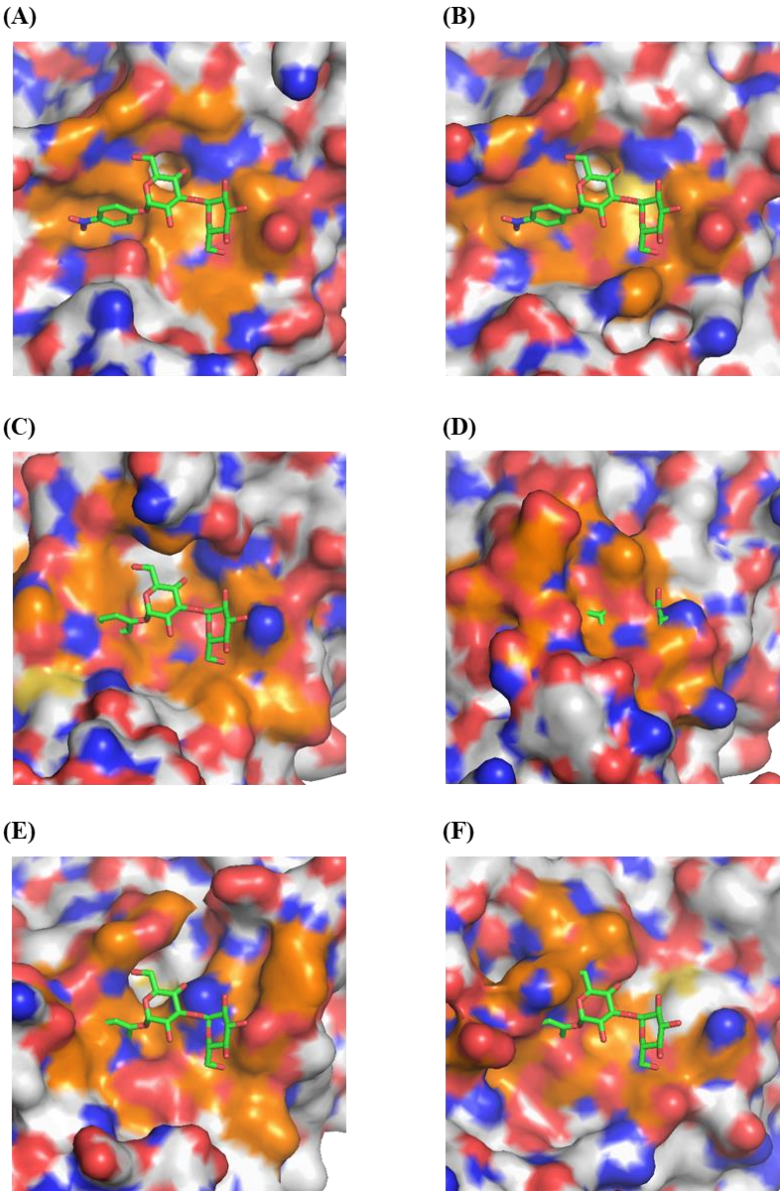


Fig. 5-26. Comparison of the surface binding site of *p*NP 3-*O*-α-L-Glc α-L-Glc found in S132D/D327S_B and Q275L/D327S_B.

A, S132D/D327S_B; B, ClAgl29A (7XSF); C, TmaFuc (2WZY); D, α-L-fucosidase from *Fusarium graminearum* (4NI3); E, α-L-fucosidase from *Bacteroides thetaiotaomicron* (4WSK); and *Bi*AfcB (3UES). Alphanumeric characters in parentheses indicate PDB accession numbers. Residues 4 Å from the ligand are colored orange.

Chapter 6. General Discussion

GH29 is a family of retaining α-L-fucosidase found in various organisms, including bacteria, plants, and animals. The GH29 can be further divided into two subfamilies based on differences in substrate specificity, GH29-A and GH29-B (Sakurama, H., 2012b). ClAgl29A and ClAgl29B were found in GH29-A. Their three-dimensional structures disclosed in Chapter 2 are composed of two domains, N-terminal catalytic $(\beta/\alpha)_8$ TIM barrel and C-terminal β-sandwich domains, as in other GH29 enzymes. The catalytic center is also similar to that of GH29-A, but the O4 recognition structures of ClAgl29A and ClAgl29B differ (Fig. 6-1). Hereafter, the residue number of ClAgl29B will be used to discuss the substrate recognition mechanism. In this study, Asp139, Asp114, Phe195, Phe425, and Phe459 are found to be important for the substrate specificity of α-L-glucosidase. Asp139 is obviously essential as it provides a carboxy group that hydrogen bonds directly to the equatorial O4 of α-L-glucoside. In addition, Asp114 and Phe195, which are involved in determining the orientation of the Asp139 side chain, are also important. These residues are substituted for other residues in GH29 α-L-fucosidase. GH29 α-L-fucosidases conserve a Phe or Trp in β→α loop 7 that acts to exclude the equatorial O4 and to exhibit specificity for the axial O4. The α-L-glucosidase also has a corresponding Phe425, but the orientation of this side chain differs from that of the corresponding aromatic ring side chain of the α-L-fucosidase. This is probably due to the difference in hydrogen bonding partners of the main chain O atom: it forms a hydrogen bond with the Nδ2 atom of conserved Asn at β8 in α-L-fucosidase, whereas with the main chain N atom of Asp139 at β→α loop 1 in ClAgl29B. ClAgl29B has Val459 instead of the conserved Asn. Thus, the divergence of Asn in β8 is also important for distinguishing α-L-glucosidase from α-L-fucosidase.

Five enzymes from *C. akajimensis* were analyzed, focusing on the divergence of His that recognizes O4 in α-L-fucosidase. They do not show significant α-L-fucosidase, α-L-glucosidase, and α-L-galactosidase activities, suggesting that GH29 may harbor enzymes with unknown substrate specificity.

Chapter 3 demonstrated the substrate specificity of ClAgl29A and ClAgl29B for *p*NP α-L-Qui rather than *p*NP α-L-Glc. The three-dimensional structure of ClAgl29B complex with β-L-glucose showed that the C6 and O6 of β-L-glucose are in a hydrophobic environment (Fig. 6-2). The structural comparison of ClAgl29B and α-L-fucosidases, TmaFuc and BiAfcB, which are representative enzymes in GH29-A and GH29-B, respectively, also possess the hydrophobic environment around C6. Since ClAgl29B lacks respective His34 and His36 of TmaFuc and BiAfcB, its hydrophobicity around C6 seems to be higher than α-L-fucosidases (Fig. 6-2). The highly polar O6 of *p*NP α-L-Glc may be disliked by the environment, which may explain the substrate specificity for *p*NP α-L-Qui. In this complex structure, O6 of β-L-glucose can form a hydrogen bond with the Sγ atom of Cys418, which could neutralize the polarity of O6. The findings from the kinetic study suggest that this

hydrogen bond may not be particularly crucial. This chapter also demonstrated that ClAgl29A and ClAgl29B have little activity on *p*NP α-L-Rha. Together with their low substrate specificity for *p*NP α-L-Fuc, the enzymes do not use α-L-rhamnosides or α-L-fucosides, which are widely distributed in nature, as substrates. ClAgl29A and ClAgl29B, in their unique enzymatic substrate specificity, act on glycosides that have not yet been identified or are rarely present, adding a new dimension to our understanding of GH29 enzymes. Since ClAgl29A and ClAgl29B show higher substrate specificity for α-L-quinovoside than for α-L-glucoside, they may rather be named a novel α-L-quinovosidase. In addition, the natural substrate may not be α-L-glucoside, the existence of which is still unknown, but α-L-quinovoside, which is suggested to exist as a building block of O-antigen. It will be necessary to proceed with the substrate specificity analysis of the homologs and to name them carefully.

Chapter 4 demonstrated that the mutant enzymes of ClAgl29B, S132D, Q275L, and S132D/Q275L, made by mimicking ClAgl29A with higher transglucosylation, exhibit transglucosylation activity that exceeds that of the parent. To interpret these results, Chapter 5 aimed to obtain the crystal structures of S132D/D327S and Q275L/D327S, whose catalytic nucleophile residue Asp327 is mutated, complex with, and the transglucosylation product, *p*NP 3-*O*-α-L-Glc α-L-Glc. However, it was not possible to get a picture of the active site pocket in *p*NP 3-*O*-α-L-Glc α-L-Glc, since citrate was bound to the active center of the obtained crystal structure, and *p*NP 3-*O*-α-L-Glc α-L-Glc was bound to the presumed surface substrate binding site away from the active site in both mutant enzymes. Here, the docking of *p*NP 3-*O*-α-L-Glc α-L-Glc on Q275L/D327 is simulated using GNINA (Fig. 6-3). The side chain of Leu275 is predicted to be at subsite +1 and *p*NP biding site. The substitution from Gln would increase the hydrophobicity of this site, making it easier for *p*NP α-L-Glc to act as an acceptor and increasing its transglucosylation ratio. The decreased K_m values for *p*NP α-L-Glc and *p*NP α-L-Fuc in Q275L and S132D/Q275L may also be due to the increased hydrophobicity at subsite +1, which increases the binding of the *p*NP group. The increase in binding strength did not lead to activation of the ES complex, i.e., an increase in k_{cat}, which may also support that the substitution of Leu for Gln enhances the hydrophobic environment.

Chapter 5 also demonstrated the conversion of ClAgl29A and ClAgl29B into glycosynthases. These mutant enzymes were able to catalyze the α-L-glucosylation of several *p*NP D-glycosides using β-L-Glc-F as a donor substrate. In particular, they efficiently catalyze α-L-glucosylation of *p*NP α-D-Xyl and *p*NP β-D-Xyl, and synthesized *p*NP 4-*O*-α-L-glucosyl α-D-xyloside and *p*NP 3-*O*-α-L-glucosyl β-D-xyloside, respectively. These enzymes also converted *p*NP β-D-Cel to *p*NP 4-*O*-α-L-glucosyl β-D-cellobiose. The compound that is a good acceptor suggests that it is more likely to bind to the subsite +1 of ClAgl29A and ClAgl29B, and α-L-Glc-(1→3)-D-Xyl, α-L-Glc-(1→4)-D-Xyl, and α-L-Glc-(1→4)-β-D-Glc-(1→4)-D-Glc may exist in nature. It will be necessary to evaluate the hydrolytic activity of these synthesized sugars with wild-type enzymes. As the wild-type ClAgl29A

and ClAgl29B exhibit the substrate specificity for α-L-quinovoside, glycosynthase reactions with β-L-quinovosyl fluoride are also interesting.

The functionality of α-L-glucoside remains to be elucidated. The efficient synthesis of α-L-glucosylated glycosides by the glycosynthase developed in this study is expected to elucidate novel functions of α-L-glucosides.

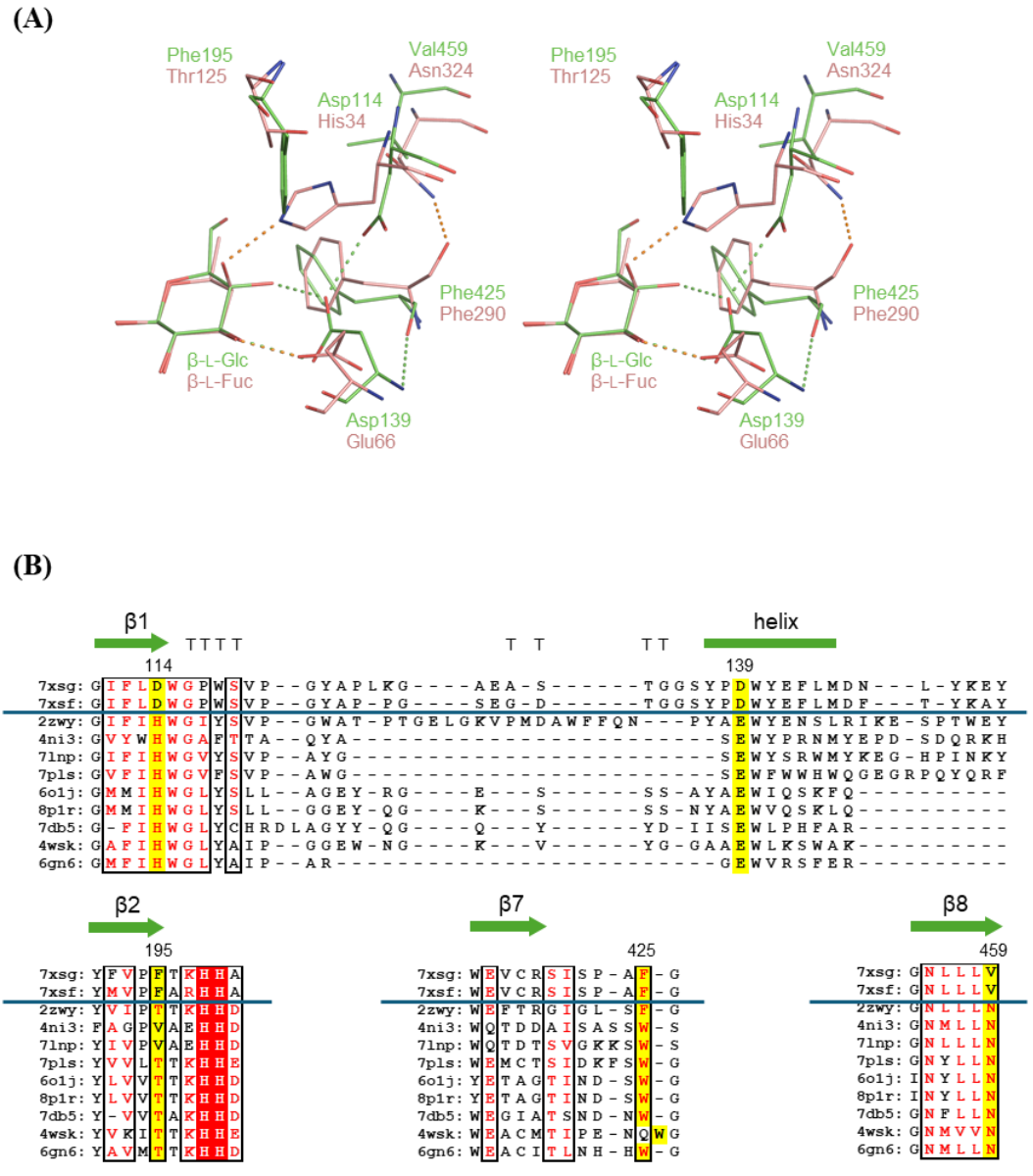


Fig. 6-1. Structural factors for α-L-glucosidase to recognize equatorial O4.

(A) Structural comparison of ClAgl29B (PDB accession, 7XSH) and TmaFuc (PDB accession, 1ODU). Residues involved in O4 recognition in each enzyme and their surrounding residues are shown in comparison. (B) Multiple amino acid sequence alignment of structure-known GH29-A enzymes. The horizontal line distinguishes α-L-glucosidase from α-L-fucosidase; above the line is α-L-glucosidase. Residues involved in the recognition of equatorial O4 and their corresponding residues are highlighted in yellow. Secondary structure elements and numbers shown above the alignment are those of ClAgl29B.

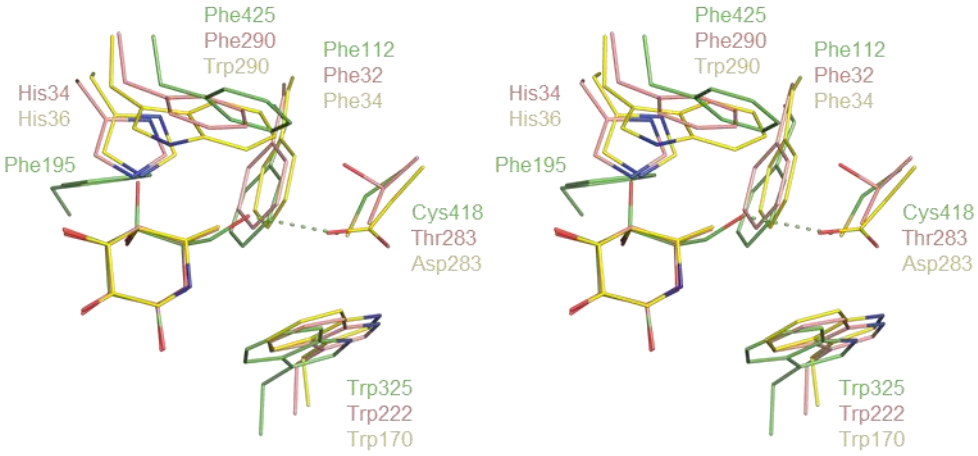


Fig. 6-2. Structure of C6 and O6 binding site of ClAgl29B (green; PDB accession, 7XSH) and corresponding residues of TmaFuc (wheat; 1ODU) and BiAfcB (yellow; 3UES).

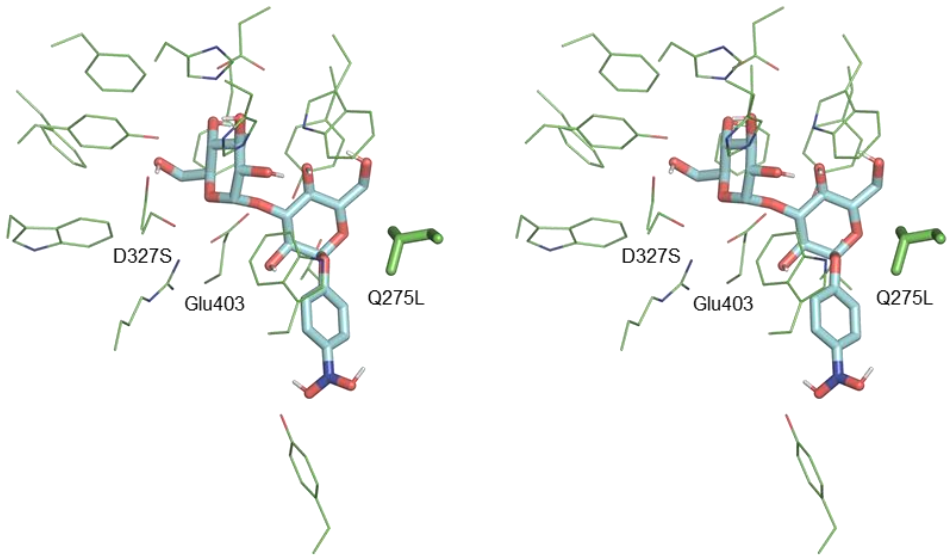


Fig. 6-3. Molecular docking simulation of *p*NP 3-*O*-α-L-Glc α-L-Glc on Q275L/D327_B.

A molecular docking simulation was performed by GNINA. Glu403 is a residue providing a general acid/base catalyst.

References

Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., and Adams, P. D. (2012). Towards automated crystallographic structure refinement with phenix. refine. *Acta Crystallog*. *D*, **68**, 352-367.

Almagro Armenteros, J. J., Tsirigos, K. D., Sønderby, C. K., Petersen, T. N., Winther, O., Brunak, S., and Nielsen, H. (2019). SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat*. *Biotechnol*., **37**, 420-423.

Anastasiou, I. A., Eleftheriadou, I., Tentolouris, A., Mourouzis, I., Pantos, C., and Tentolouris, N. (2021). The use of L-glucose in cancer diagnosis: results from in vitro and in vivo studies. *Curr*. *Med*. *Chem*., **28**, 6110-6122.

Andersen, S. M., Heuckendorff, M., and Jensen, H. H. (2015). 3-(Dimethylamino)-1-propylamine: a cheap and versatile reagent for removal of byproducts in carbohydrate chemistry. *Org*. *Lett*., **17**, 944-947.

Ashida, H., Miyake, A., Kiyohara, M., Wada, J., Yoshida, E., Kumagai, H., and Yamamoto, K. (2009). Two distinct α-L-fucosidases from *Bifidobacterium bifidum* are essential for the utilization of fucosylated milk oligosaccharides and glycoconjugates. *Glycobiology*, **19**, 1010-1017.

Brennan, P. J. (2003) Structure, function, and biogenesis of the cell wall of *Mycobacterium tuberculosis*. *Tuberculosis*, **83**, 91-97.

Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V., and Henrissat, B. (2009). The Carbohydrate-Active EnZymes database (CAZy): an expert resource for glycogenomics. *Nucleic Acids Res*., **37**, D233-D238.

Chutipongtanate, S., Morrow, A. L., and Newburg, D. S. (2022). Human milk oligosaccharides: Potential applications in COVID-19. *Biomedicines*, **10**, 346.

Cobucci-Ponzano, B., Trincone, A., Giordano, A., Rossi, M., and Moracci, M. (2003). Identification of the catalytic nucleophile of the family 29 α-L-fucosidase from *Sulfolobus solfataricus* via chemical rescue of an inactive mutant. *Biochemistry*, **42**, 9525-9531.

Cobucci-Ponzano, B., Conte, F., Rossi, M., and Moracci, M. (2008). The α-L-fucosidase from *Sulfolobus solfataricus*. *Extremophiles*, **12**, 61-68.

Cobucci-Ponzano, B., Conte, F., Bedini, E., Corsaro, M. M., Parrilli, M., Sulzenbacher, G., and

Moracci, M. (2009). β-Glycosyl azides as substrates for α-glycosynthases: Preparation of efficient α-L-fucosynthases. *Chem*. *Biol*., **16**, 1097-1108.

Coyle, T., Wu, L., Debowski, A. W., Davies, G. J., and Stubbs, K. A. (2019) Synthetic and crystallographic insight into exploiting sp^2 hybridization in the development of α-L-fucosidase inhibitors. *Chembiochem*, **20**, 1365-1368.

Davies, G. J., and Henrissat, B. (1995). Structures and mechanisms of glycosyl hydrolases. *Structure*, **3**, 853-859.

De Castro, C., Kenyon, J. J., Cunneen, M. M., Molinaro, A., Holst, O., Skurnik, M., and Reeves, P. R. (2013). The O-specific polysaccharide structure and gene cluster of serotype O: 12 of the *Yersinia pseudotuberculosis* complex, and the identification of a novel L-quinovose biosynthesis gene. *Glycobiology*, **23**, 346-353.

Drula, E., Garron, M. L., Dogan,ul S., Lombard, V., Henrissat, B., and Terrapon, N. (2022). The carbohydrate-active enzyme database: functions and literature. *Nucleic Acids Res*., **50**, D571-D577.

Dubovski, N., Shoshan-Galeczki, Y. B., Malach, E., and Niv, M. Y. (2022). Taste and chirality: L-glucose sweetness is mediated by TAS1R2/TAS2R3 receptor. *Food Chem*., **373**, 131393.

Ducros, V. M., Tarling, C. A., Zechel, D. L., Brzozowski, A. M., Frandsen, T. P., von Ossowski, I., and Davies, G. J. (2003). Anatomy of glycosynthesis: structure and kinetics of the *Humicola insolens* Cel7B E197A and E197S glycosynthase mutants. *Chem*. *Biol*., **10**, 619-628.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallog*. *D*, **60**, 2126-2132.

Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010). Features and development of Coot. *Acta Crystallog*. *D*, **66**, 486-501.

Gorshkova, R. P., Zubkov, V. A., Isakov, V. V., and IuS, O. (1983). Study of a lipopolysaccharide from *Yersinia pseudotuberculosis* of the IVA serotype. *Bioorg. Khim.*, **9**, 1401-1407.

Harms, G., Reuter, G., Corfield, A. P., and Schauer, R. (1996). Binding specificity of influenza C-virus to variably O-acetylated glycoconjugates and its use for histochemical detection of N-acetyl-9-O-acetylneuraminic acid in mammalian tissues. *Glycoconj. J*., **13**, 621-630.

Jeurink, P. V., van Esch, B. C., Rijnierse, A., Garssen, J., and Knippels, L. M. (2013). Mechanisms underlying immune effects of dietary oligosaccharides. *Am. J. Clin. Nutr.*, **98**, 572S-577S.

Kabsch, W. (2010). xds. *Acta Crystallog*. *D*, **66**, 125-132.

Kuzuya, H., Tamai, I., Beppu, H., Shimpo, K., and Chihara, T. (2001) Determination of aloenin, barbaloin and isobarbaloin in *Aloe* species by micellar electrokinetic chromatography. *J*. *Chromatogr*. *B*. *Biomed*. *Sci*. *Appl*., **752**, 91-97.

Kocharova, N. A., Ovchinnikova, O. G., Toukach, F. V., Torzewska, A., Shashkov, A. S., Knirel, Y. A., and Rozalski, A. (2005). The *O*-polysaccharide from the lipopolysaccharide of *Providencia stuartii* O44 contains L-quinovose, a 6-deoxy sugar rarely occurring in bacterial polysaccharides. *Carbohydr. Res*., **340**, 1419-1423.

Kong, C., Akkerman, R., Klostermann, C. E., Beukema, M., Oerlemans, M. M., Schols, H. A., and De Vos, P. (2021). Distinct fermentation of human milk oligosaccharides 3-FL and LNT2 and GOS/inulin by infant gut microbiota and impact on adhesion of *Lactobacillus plantarum* WCFS1 to gut epithelial cells. *Food Funct*., **12**, 12513-12525.

Kotani, S., Kato, T., Matsuda, T., Kato, K., and Misaki, A. (1971). Chemical structure of the antigenic determinants of cell wall polysaccharide of *Mycobacterium tuberculosis* strain H37Rv. *Biken J*., **14**, 379-387.

Kovalová, T., Koval, T., Benešová, E., Vodicková, P., Spiwok, V., Lipovová, P., and Dohnálek, J. (2019) Active site complementation and hexameric arrangement in the GH family 29; a structure-function study of α-L-fucosidase isoenzyme 1 from *Paenibacillus thiaminolyticus*. *Glycobiology*, **29**, 59-73.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, **227**, 680-685.

Lammerts van Bueren, A., Fayers-Kerr, J., Luo, B., Zhang, Y., Sollogoub, M., Blériot, Y., and Davies, G. J. (2010). Analysis of the reaction coordinate of α-L-fucosidases: A combined structural and quantum mechanical approach. *J. Am. Chem. Soc.*, **132**, 1804-1806.

Lawrence, A. M., and Besir, H. (2009). Staining of proteins in gels with Coomassie G-250 without organic solvent and acetic acid. *J. Vis. Exp*., **30**, 1350.

Liebschner, D., Afonine, P. V., Baker, M. L., Bunkóczi, G., Chen, V. B., Croll, T. I., and Adams, P. D. (2019). Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallog*. *D*, **75**, 861-877.

Long, F., Nicholls, R. A., Emsley, P., Gražulis, S., Merkys, A., Vaitkus, A., and Murshudov, G. N.

(2017). AceDRG: a stereochemical description generator for ligands. *Acta Crystallog*. *D*, **73**, 112-122.

Mackenzie, L. F., Wang, Q., Warren, R. A. J., and Withers, S. G. (1998). Glycosynthases: mutant glycosidases for oligosaccharide synthesis. *J. Am. Chem. Soc.*, **120**, 5583-5584.

Malet, C., and Planas, A. (1998). From β-glucanase to β-glucansynthase: glycosyl transfer to α-glycosyl fluorides catalyzed by a mutant endoglucanase lacking its catalytic nucleophile. *FEBS Lett.*, **440**, 208-212.

McCarter, J. D., and Withers, G. S. (1994). Mechanisms of enzymatic glycoside hydrolysis. *Curr. Opin. Struct. Biol*., **4**, 885-892.

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007). Phaser crystallographic software. *J. Appl. Crystallogr.*, **40**, 658-674.

Mcneil, M. R., Robuck, K. G., Harter, M., and Brennan, P. J. (1994). Enzymatic evidence for the presence of a critical terminal hexa-arabinoside in the cell walls of *Mycobacterium tuberculosis*. *Glycobiology,* **4**, 165–174.

Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., and Vagin, A. A. (2011). REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallog*. *D*, **67**, 355-367.

Niu, C., Luo, H., Shi, P., Huang, H., Wang, Y., Yang, P., and Yao, B. (2016). N-glycosylation improves the pepsin resistance of histidine acid phosphatase phytases by enhancing their stability at acidic pHs and reducing pepsin's accessibility to its cleavage sites. *Appl. Environ. Microbiol*., **82**, 1004-1014.

Ovchinnikova, O. G., Rozalski, A., Liu, B., and Knirel, Y. A. (2013). *O*-antigens of bacteria of the genus Providencia: structure, serology, genetics, and biosynthesis. *Biochemistry*, **78**, 798-817.

Pekdemir, B., and Karav, S. (2024). Exploring the diverse biological significance and roles of fucosylated oligosaccharides. *Front. mol. biosci*., **11**, 1403727.

Robb, C. S., Hobbs, J. K., Pluvinage, B., Reintjes, G., Klassen, L., Monteith, S., and Boraston, A. B. (2022). Metabolism of a hybrid algal galactan by members of the human gut microbiome. *Nat. Chem. Biol*., **18**, 501-510.

Robert, X., Haser, R., Mori, H., Svensson, B., and Aghajari, N. (2005). Oligosaccharide binding to barley α-amylase 1. *J. Biol. Chem*., **280**, 32968-32978.

Sakurama, H., Tsutsumi, E., Ashida, H., Katayama, T., Yamamoto, K., and Kumagai, H. (2012a). Differences in the substrate specificities and active-site structures of two α-L-fucosidases (glycoside hydrolase family 29) from *Bacteroides thetaiotaomicron*. *Biosci. Biotechnol. Biochem*., **76**, 1022-1024.

Sakurama, H., Fushinobu, S., Hidaka, M., Yoshida, E., Honda, Y., Ashida, H., and Katayama, T. (2012b). 1, 3-1, 4-α-L-fucosynthase that specifically introduces Lewis a/x antigens into type-1/2 chains. *J. Biol. Chem.*, **287**, 16709-16719.

Sela, D. A., Garrido, D., Lerno, L., Wu, S., Tan, K., Eom, H. J., and Mills, D. A. (2012). *Bifidobacterium longum* subsp. *infantis* ATCC 15697 α-fucosidases are active on fucosylated human milk oligosaccharides. *Appl. Environ. Microbiol*., **78**, 795-803.

Shimizu, T., Takaya, N., and Nakamura, A. (2012). An L-glucose catabolic pathway in *Paracoccus* species 43P. *J. Biol. Chem*., **287**, 40448-40456.

Shishiuchi, R., Kang, H., Tagami, T., Ueda, Y., Lang, W., Kimura, A., and Okuyama, M. (2022). Discovery of α-L-glucosidase raises the possibility of α-L-glucosides in nature. *ACS Omega*, **7**, 47411-47423.

Sinnott, M. L. (1990). Catalytic mechanism of enzymic glycosyl transfer. *Chem*. *Rev*., **90**, 1171-1202.

Sulzenbacher, G., Bignon, C., Nishimura, T., Tarling, C. A., Withers, S. G., Henrissat, B., and Bourne, Y. (2004). Crystal structure of *Thermotoga maritima* α-L-fucosidase: Insights into the catalytic mechanism and the molecular basis for fucosidosis. *J. Biol. Chem*., **279**, 13119-13128.

Trincone, A., Giordano, A., Perugino, G., Rossi, M., and Moracci, M. (2005). Highly productive autocondensation and transglycosylation reactions with *Sulfolobus solfataricus* glycosynthase. *ChemBioChem*, **6**, 1431-1437.

Varki, A., Cummings, R. D., Esko, J. D., Stanley, P., Hart, G. W., Aebi, M., and Seeberger, P. H. (2015). Essentials of Glycobiology. 4th ed. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press.

Wan, L., Zhu, Y., Zhang, W., and Mu, W. (2020). α-L-Fucosidases and their applications for the production of fucosylated human milk oligosaccharides. *Appl Microbiol Biotechnol*., **104**, 5619–5631.

Xia, Y., Li, K., Li, J., Wang, T., Gu, L., and Xun, L. (2019). T5 exonuclease-dependent assembly

offers a low-cost method for efficient cloning and site-directed mutagenesis. *Nucleic Acids Res*., **47**, e15.

Xin, Y., Huang Y., and McNeil, M.R. (1999). The presence of an endogenous endo-D-arabinase in Mycobacterium smegmatis and characterization of its oligoarabinoside product. *Biochim. Biophys. Acta*, **1473**, 267-271.

Xu, H. L., Zhou, X., Chen, S., Xu, S., Li, Z., Nakanishi, H., and Gao, X. D. (2023). Rare sugar L-sorbose exerts antitumor activity by impairing glucose metabolism. *Commun*. *Biol*., **6**, 259.

Hricoviniova, Z. (2009). The influence of microwave irradiation on stereospecific Mo(VI)-catalyzed transformation of deoxysugars. *Tetrahedron-Asymmetr.*, **20**, 1239-1242.

Zhang, N., Jin, M., Wang, K., Zhang, Z., Shah, N. P., and Wei, H. (2022). Functional oligosaccharide fermentation in the gut: Improving intestinal health and its determinant factors-A review. *Carbohydr. Polym*., **284**, 119043.