ACS OMEGA

http://pubs.acs.org/journal/acsodf

Article

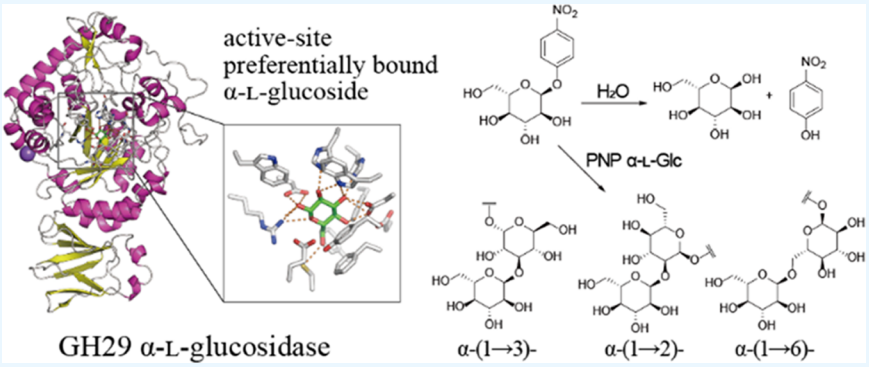
Discovery of α-L-Glucosidase Raises the Possibility of α-L-Glucosides in Nature

Rikako Shishiuchi,‡ Hyejin Kang,‡ Takayoshi Tagami,* Yoshitaka Ueda, Weeranuch Lang, Atsuo Kimura, and Masayuki Okuyama*

Cite This: *ACS Omega* 2022, 7, 47411−47423

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information



ABSTRACT: Glucose, a common monosaccharide in nature, is dominated by the D-enantiomer. Meanwhile, the discovery of L-glucose-utilizing bacteria and the elucidation of their metabolic pathways 10 years ago suggests that L-glucose exists naturally. Most carbohydrates exist as glycosides rather than monosaccharides; therefore, we expected that nature also contains L-glucosides. Sequence analysis within glycoside hydrolase family 29 led us to identify two α-L-glucosidases, ClAgl29A and ClAgl29B, derived from *Cecembia lonarensis* LW9. ClAgl29A and ClAgl29B exhibited higher K_m, k_{cat}, and k_{cat}/K_m values for *p*-nitrophenyl α-L-glucoside than that for *p*-nitrophenyl α-L-fucoside. Structural analysis of ClAgl29B in complex with L-glucose showed that these enzymes have an active-site pocket that preferentially binds α-L-glucoside, but excludes α-L-fucoside. These results suggest that ClAgl29A and ClAgl29B evolved to hydrolyze α-L-glucoside, implying the existence of α-L-glucoside in nature. Furthermore, α-L-glucosidic linkages (α-L-Glc-(1 → 3)-L-Glc, α-L-Glc-(1 → 2)-L-Glc, and α-L-Glc-(1 → 6)-L-Glc) were synthesized by the transglucosylation activity of ClAgl29A and ClAgl29B. We believe that this study will lead to new research on α-L-glucosides, including determining the physiological effects on humans, and the discovery of novel α-L-glucoside-related enzymes.

INTRODUCTION

Carbohydrates are believed to be homochiral in nature. Primary carbohydrates including glucose, fructose, mannose, and galactose are present as D-sugars. Some L-sugars are also present in nature, including L-fucose, L-rhamnose, and L-arabinose. Both D- and L-sugars are present in nature in some cases. For example, L-arabinose is widespread as a component of polysaccharides in the gum and hemicellulose of plants, whereas D-arabinose is found in *Mycobacterium* polysaccharides and in *Aloe* sp. glycoside. The possibility that both enantiomers of glucose exist is suggested. The presence of an L-glucose-utilizing bacterium in soil (*Paracoccus laeviglucosivorans*)[1] may indicate a break in glucose homochirality. However, it is possible that the bacterium only uses a *syllo*-inositol metabolic enzyme to catabolize L-glucose because L-glucose dehydrogenase, an initial enzyme for L-glucose catabolism, is more specific for *syllo*-inositol than L-glucose. The authors implicitly mention that L-glucose-utilizing bacteria are not directly related to the presence of L-glucose in nature.[1]

Carbohydrates often exist as oligosaccharides, polysaccharides, and glycosides through the formation of glycosidic linkages, and L-glucose is possibly present as L-glucoside. If L-glucoside is present, an enzyme could be present to break it down into monosaccharides. Organisms express a wide variety of glycoside hydrolases. Glycoside hydrolases are classified into glycoside hydrolase families (GH) based on their amino acid sequence similarities and are summarized in the carbohydrate-active enzyme database (CAZy, https://www.cazy.org/).[2,3] Enzymes belonging to the same GH can be regarded as evolving from a common ancestor, and their structures and

Received: October 30, 2022
Accepted: November 29, 2022
Published: December 9, 2022

catalytic mechanisms are generally conserved. Meanwhile, there are frequent cases of divergent specificities within a GH. We may therefore be able to find enzymes with novel substrate specificities by examining molecular evolution within a GH.

We hypothesized that α-L-glucosidase may exist in GH29, which consists of α-L-fucosidase, α-1,3/1,4-L-fucosidase, α-1,2-L-fucosidase, and α-L-galactosidase. α-L-Glucoside differs from α-L-fucoside (6-deoxy L-galactoside) and α-L-galactoside in the orientation of the hydroxyl group at C4 (OH-4) (Figure 1).

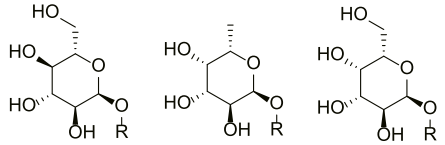


Figure 1. Structures of α-L-glucoside (1), α-L-fucoside (2), and α-L-galactoside (3).

We thought that if we could find the diversification of OH4 and OH6 recognition mechanisms within GH29, we would be able to find α-L-glucosidase in this family. We should also note the presence of an enzyme showing specificity for both α-L-fucoside and α-L-glucoside, such as GH1 β-glucosidase with high k_{cat}/K_m values for β-D-glucoside and β-D-fucoside.[4] The catalytic domain of GH29 enzymes displays an atypical TIM $(\beta/\alpha)_8$-barrel fold and a missing part of the α-helix.[5,6] Two His residues at the loop connecting β-strand 1 and α-helix 1 (β → α loop 1) and β → α loop 2 of the catalytic domain are involved in the stabilization of the axial OH-4 of α-L-fucoside through hydrogen bonds. *Thermotoga maritima* α-L-fucosidase (TmaFuc) is a representative GH29 enzyme with His34 and His128 responsible for its stabilization.[5] TmaFuc has Phe32, His34, and Phe290 surrounding the C6 methyl group of L-fucose.

In this study, divergence of the residue corresponding to His34 of TmaFuc led to the discovery of α-L-glucosidases. Determination of the three-dimensional (3D) structure of α-L-glucosidases revealed an elaborate mechanism for their recognition, which is dissimilar to that of α-L-fucosidases. The synthesis of α-L-glucosidic linkages was further examined by exploiting transglycosylation of α-L-glucosidases. This is the first enzymatic synthesis of an α-L-glucosidic linkage. This study illustrates one aspect of the breakdown of homochirality in glucose and contributes to the development of a new field of glycoscience.

RESULTS AND DISCUSSION

Candidates for α-L-Glucosidase. GH29 is a family of α-L-fucosidases, but we assumed that the family hides a diversity of recognition mechanisms for C4 and C6. This idea prompted us to search for α-L-glucosidase in GH29. We found two unique sequences encoded by the *Cecembia lonarensis* LW9 genome: GenBank accession numbers EKB48090.1 and EKB48091.1. GH29 α-L-fucosidases are equipped with His34 and His128 residues in TmaFuc to stabilize the axial OH-4 of α-L-fucoside. These His residues are invariant in α-L-fucosidases but are not completely conserved in GH29 proteins. The conservation of the His128 position was relatively high, whereas several variations (Ala, Arg, Asn, Asp, Cys, Gln, Glu, Ile, Lys, Phe, Ser, Thr, Tyr, and Val) are observed at the His34 position according to the HMM logo[7] for Alpha_L_fuc (PF01120) in the Pfam database (https://pfam.xfam.org/family/PF01120). The position of His34 was replaced with Asp114 and Asp103 in EKB48090.1 and EKB48091.1, respectively. We anticipated that these Asp residues are involved in recognizing the equatorial OH-4 of α-L-glucoside. A conserved Asp residue is responsible for stabilizing equatorial OH-4 of α-D-glucoside through a hydrogen bond in GH13 and GH31 α-D-glucosidases.[8] There is a proposed evolutionary relationship between GH27 and GH29,[5] and between GH27 and GH13 or GH31,[9−11] and we wondered if the expected α-L-glucosidase might have a substrate recognition machinery comparable to that of α-D-glucosidase.

EKB48090.1 and EKB48091.1 consist of 586 and 573 amino acid residues, respectively. They were predicted to have a signal peptide (Met1−Gly22) and a lipoprotein signal peptide (Met1−Ser18), respectively, by SignalP-5.0. Pairwise alignment of these proteins showed 56.1% identity and 70.3% similarity and contained a 19.8% gap. Sequences belonging to GH29-A,[12] such as α-L-galactosidase from *Bacteroides plebeius* DSM 17153 (EDY95436.1) and TmaFuc (AE001712.1), were the top hits with similar sequences based on a FASTA search of these sequences using PDB as the target database. EKB48090.1 and EKB48091.1 are encoded in tandem in the genome: B879_03287 and B879_03288, respectively. The neighboring genes of B879_03288 encoded putative carbohydrate-metabolism enzymes, including a putative D-xylose transporter, a GntR family transcriptional regulator, a putative 1,5-D-anhydrofructose reductase, and a D-xylose isomerase-like TIM barrel protein. The putative 1,5-D-anhydrofructose reductase displays sequence similarity (31% identity) to L-glucose/*scyllo*-inositol dehydrogenase (EC 1.1.1.370), the first enzyme responsible for L-glucose oxidation.[1] We chose to study EKB48091.1 (denoted as ClAgl29A) and EKB48090.1

Table 1. Kinetic Constants of ClAgl29A and ClAgl29B

	k (s^{-1})[a]	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m (s^{-1} mM^{-1})
ClAgl29A				
PNP α-L-Glc	0.980	2.01 ± 0.26	2.00 ± 0.11	1.00 ± 0.08
PNP α-L-Fuc	0.0626	0.279 ± 0.005	0.0658 ± 0.0008	0.236 ± 0.005
PNP α-L-Gal	1.42×10^{-3}			
ClAgl29B				
PNP α-L-Glc	3.60	1.79 ± 0.18	7.76 ± 0.41	4.34 ± 0.22
PNP α-L-Fuc	0.108	0.240 ± 0.011	0.0844 ± 0.0019	0.352 ± 0.007
PNP α-L-Gal	7.83×10^{-3}			

[a]The initial velocities, measured in McIlvaine buffer (pH 5.5) at 35 °C, were fitted with the Michaelis−Menten equation to obtain kinetic parameters. The rate constants were calculated from each hydrolytic rate at 2 mM substrate.

https://doi.org/10.1021/acsomega.2c06991
ACS Omega 2022, 7, 47411−47423

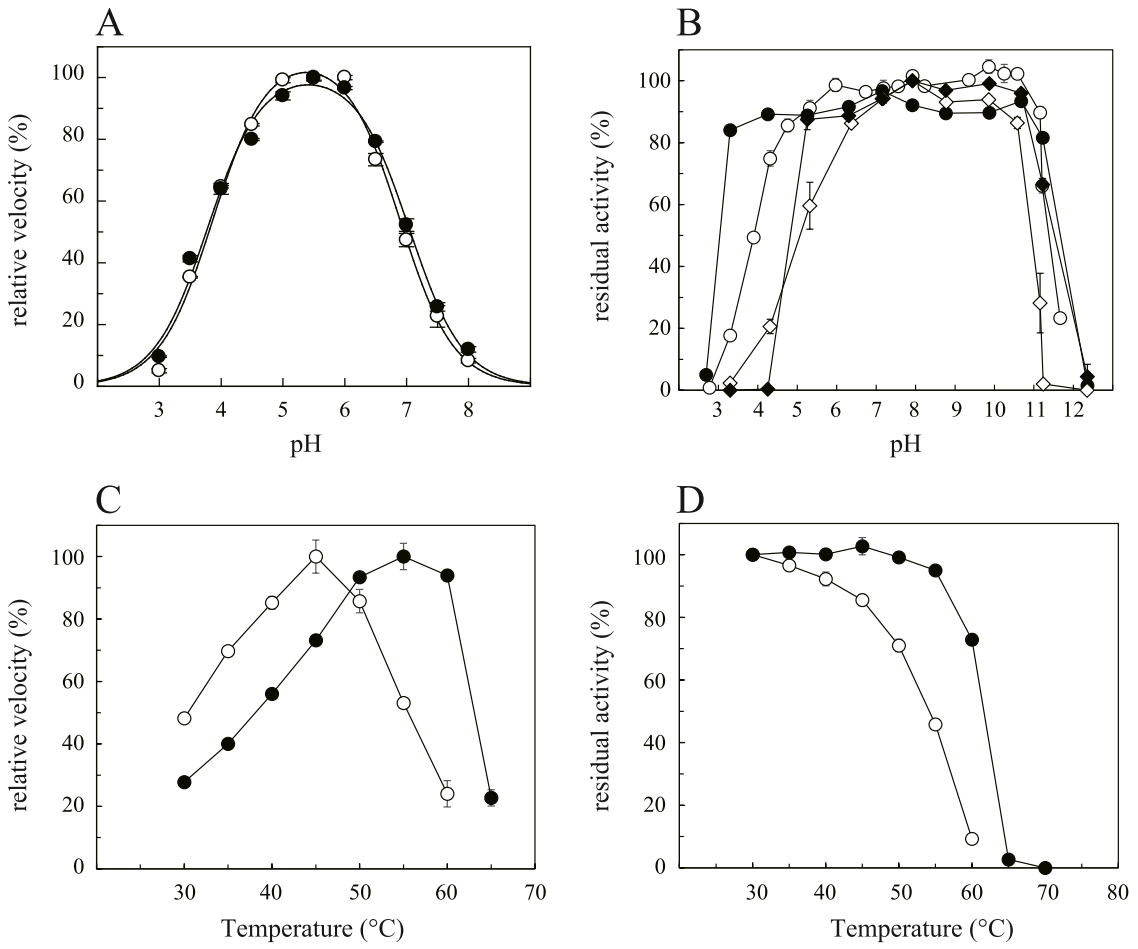


Figure 2. Effects of pH and temperature on the activity and stability of ClAgl29A and ClAgl29B. The relative rate and residual activity are expressed as percentages of the maximum values. The open and closed symbols represent the results for ClAgl29A and ClAgl29B, respectively. Two mM PNP α-L-Glc was used as a substrate in all of the experiments. Error bars represent the standard deviation of three independent experiments. (A) pH rate profile: data were fitted to the following equation: Relative rate = (Limit × log(pH−pK_{a1})/(log(2pH−pK_{a1}−pK_{a2}) + log(pH−pK_{a1}) +1)). (B) pH stability: residual activity was examined after incubation for 24 h at 4 °C (circle) or 3 h at 35 °C (diamond) at various pH values. (C) Temperature−rate profile. (D) Temperature stability: residual activity was examined after incubation for 15 min at various temperatures.

(denoted as ClAgl29B) as potential candidates for α-L-glucosidase based on these sequence features.

Characterization of ClAgl29A and ClAgl29B. Recombinant ClAgl29A and ClAgl29B were produced in *Escherichia coli* without the signal peptides and their enzymatic properties were analyzed. The N-terminal Cys20 and Ser21 residues of mature ClAgl29B protein were deleted to prevent membrane anchoring. The hydrolytic rate constants were determined using 2 mM *p*-nitrophenyl α-L-fucoside (PNP α-L-Fuc), PNP α-L-glucoside (PNP α-L-Glc), and PNP α-L-galactoside (PNP α-L-Gal) for ClAgl29A and ClAgl29B (Table 1).

The values of TmaFuc were also determined as a control. ClAgl29A and ClAgl29B hydrolyzed PNP α-L-Glc 16 times faster and 33 times faster than PNP α-L-Fuc, respectively. These enzymes could hardly hydrolyze PNP α-L-Gal: the rate constant for the hydrolysis of PNP α-L-Gal was 1.5×10^{-3}-fold (ClAgl29A) and 2.2×10^{-3}-fold (ClAgl29B), compared to that for PNP α-L-Glc. In contrast, TmaFuc hydrolyzed PNP α-L-Fuc, but the rate of hydrolysis of PNP α-L-Glc was below the detection limit. This enzyme also hydrolyzes PNP α-L-Gal at 30% of the reaction rate for PNP α-L-Fuc. As discussed later, these two enzymes catalyze transglycosylation at 30 mM PNP α-L-Glc. If the reaction scheme includes transglycosylation, the rate equation will differ from the Michaelis−Menten kinetic equation. However, the reaction rate followed the Michaelis−Menten equation up to at least 6 mM PNP α-L-Glc (Figure S1), and thus we estimated Michaelis−Menten kinetic parameters. The k_{cat}/K_m values showed a specificity of these enzymes that favored PNP α-L-Glc over PNP α-L-Fuc (Table 1). ClAgl29A and ClAgl29B displayed higher k_{cat} and K_m values toward PNP α-L-Glc than PNP α-L-Fuc. Kinetic analysis indicated that ClAgl29A and ClAgl29B may have evolved as α-L-glucosidases which weakly bind α-L-glucoside to maximize the reaction rate, as described by Fersht.[13]

The pH effects on ClAgl29A and ClAgl29B were similar and the optimum pH for the hydrolysis of 2 mM PNP α-L-Glc of both enzymes was 5.5 (Figure 2). The pH stability was evaluated by maintaining the enzymes at various pH values for 24 h at 4 °C or 3 h at 35 °C and measuring their residual activity. The pH ranges with over 80% residual activity were wider at lower temperatures. The ranges showing >80% residual activity after 24 h were 4.8−11.2 for ClAgl29A and 3.3−11.2 for ClAgl29B; those after 3 h were 6.4−10.6 for ClAgl29A and 5.3−10.7 for ClAgl29B. The extensive stability in the alkaline range may be related to the nature of *C. lonarensis* LW9, which was originally isolated from a haloalkaline lake and can grow under alkaline conditions.[14] ClAgl29A and ClAgl29B contain the predicted secretion signal sequence

https://doi.org/10.1021/acsomega.2c06991
ACS Omega 2022, 7, 47411−47423

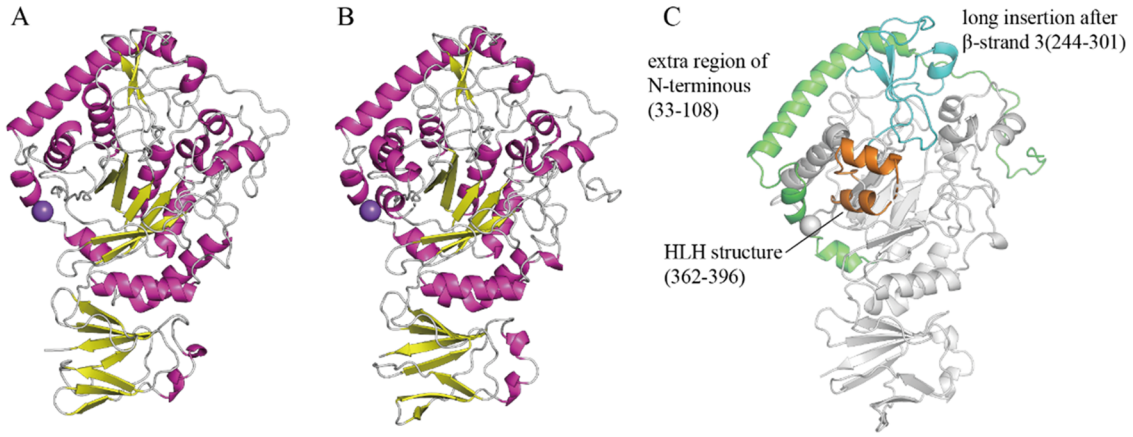


Figure 3. Three-dimensional structures of ClAgl29A and ClAgl29B. Cartoon representations of ClAgl29A (A) and ClAgl29B (B). The sphere models show the bound sodium ions. (C) The characteristic structural elements found in ClAgl29B are indicated in green for the additional region of the N-terminus (33−108),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

at their N-terminus and must be exposed to an alkaline environment. It is unclear why these enzymes seldom function in the alkaline range. ClAgl29A was more susceptible to heat than ClAgl29B. The optimal temperature for ClAgl29A and ClAgl29B was 45 and 55 °C, respectively (Figure 1). ClAgl29B maintained its activity at 55 °C for 15 min, whereas ClAgl29A lost 55% of its activity. ClAgl29B showed 80% residual activity after 15 min incubation at 60 °C, while ClAgl29A lost 90% of its activity. The determined structures of the two enzymes (stated below) are similar, and it is difficult to clearly explain the difference in temperature stability based on structural factors.

Overall Structures of the α-L-Glucosidases. The crystal structures of ClAgl29A and ClAgl29B were determined at 2.0 Å (ClAgl29A) and 1.6 Å (ClAgl29B) resolution. Co-crystals of ClAgl29A with L-fucose were used for the structural analysis, but no electron density was observed in its active site. All residues were built based on electron density, except for residues 23 and 357−379 in ClAgl29A and residues 21−32 and 378−386 in ClAgl29B. Both asymmetric units contained two monomers, and this observation agreed with the results of gel filtration analysis, indicating that both ClAgl29A and ClAgl29B form dimers in solution, even though PISA[15] suggested that a dimer composed of another crystal symmetry mate is a possible biological assembly in both cases.

The structure of each protein comprised two domains: an atypical catalytic TIM $(\beta/\alpha)_8$ barrel domain (residues 24−480, ClAgl29A; residues 32−492, ClAgl29B) and a β-sandwich domain (residues 481−573, ClAgl29A; residues 493−586, ClAgl29B) (Figure 3A,B) which is similar to the majority of GH29 proteins. ClAgl29A and ClAgl29B structures were similar, with the root-mean-square deviation (RMSD) calculated between the Cα-atoms of matched residues at the best 3D superposition, 0.376 Å. We searched PDBeFold (https://www.ebi.ac.uk/msd-srv/ssm/) for similar 3D structures of ClAgl29B and found that TmaFuc (PDB, 1HL9),[5] α-L-fucosidase of *Paenibacillus thiaminolyticus* (PDB, 6GN6),[16] and α-L-fucosidase of *Bacteroides thetaiotaomicron* (BT2970; PDB, 2XIB)[17] were homologous enzymes with the highest Q-scores of 0.45, 0.40, and 0.39, respectively. The TIM $(\beta/\alpha)_8$-barrel fold of GH29 is unorthodox.[5,6] There is a discontinuous barrel with an insufficient hydrogen bond between the fifth and sixth strands (β5 and β6) of the inner β-barrel and a lack of a typical helix structure between β5 and β6. The fifth α-helix (α5) is a feature of the GH29 TIM $(\beta/\alpha)_8$-barrel fold. ClAgl29A and ClAgl29B have no obvious hydrogen bonds between β5 and β6. ClAgl29B has a helix-loop-helix (HLH) structure (residues 362−396) instead of α5 (Figure 3C). α-Helices in a TIM $(\beta/\alpha)_8$ barrel generally surround the inner β-barrel to protect it from the solvent. However, the HLH structure exists independently of the barrel. We could not determine the corresponding part of ClAgl29A because of unclear electron density. ClAgl29A and ClAgl29B had no distinguishable α6, and a part of the inner barrel was exposed to the solvent; this was similar to that of TmaFuc and BT2970.[5,17] ClAgl29A and ClAgl29B had singular structures at the $\beta \rightarrow \alpha$ loop 3 and the N-terminus (Figure 3C). The $\beta \rightarrow \alpha$ loop 3 is longer than that of other GH29 enzymes and consists of a small antiparallel β-sheet and a small helix (residues 232−289 for ClAgl29A and residues 244−301 for ClAgl29B). There was a structure containing a large α-helix (residues 46−74 for ClAgl29A and residues 57−85 for ClAgl29B) coiled around the TIM $(\beta/\alpha)_8$-barrel at the N-terminus. ClAgl29A and ClAgl29B showed Na^+-binding at the same sites: ClAgl29A holds the ion via the O of Asp85, O of His 87, and Oδ1 of Asp386, whereas ClAgl29B holds the ion via the O of Lys96, O of His98, and Oδ1 of Asp398.

Active-Site Structure of α-L-Glucosidase. The ClAgl29B complex obtained with β-L-glucose by a soaking experiment was probably due to mutarotation in the solution, and was refined to 1.7 Å. The structure of ClAgl29B in complex with β-L-glucose allowed us to identify the active site, which was in a small pocket formed by loops connecting β-strands and α-helices of the $(\beta/\alpha)_8$-barrel domain (Figure 4). No structural rearrangements were observed between enzymes with and without ligand binding. Asp327 and Asp315 from ClAgl29B ClAgl29A, respectively, were the residues providing the catalytic nucleophile in comparison with other GH29 enzymes.[5,18,19] 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. Although these interactions should not be observed in the complex with a substrate, it might show that the Oδ1 atom is close enough to attack the anomeric carbon of the substrate and that the side chain of Arg361 affects the dissociation state of the catalytic nucleophile. The carboxy groups of Glu403 (ClAgl29B) and Glu391 (ClAgl29A) at pseudo β6 were located 5.5 Å from the catalytic nucleophile at the end of β4 and should act as a general acid/base catalyst. ClAgl29B tightly

https://doi.org/10.1021/acsomega.2c06991
ACS Omega 2022, 7, 47411−47423

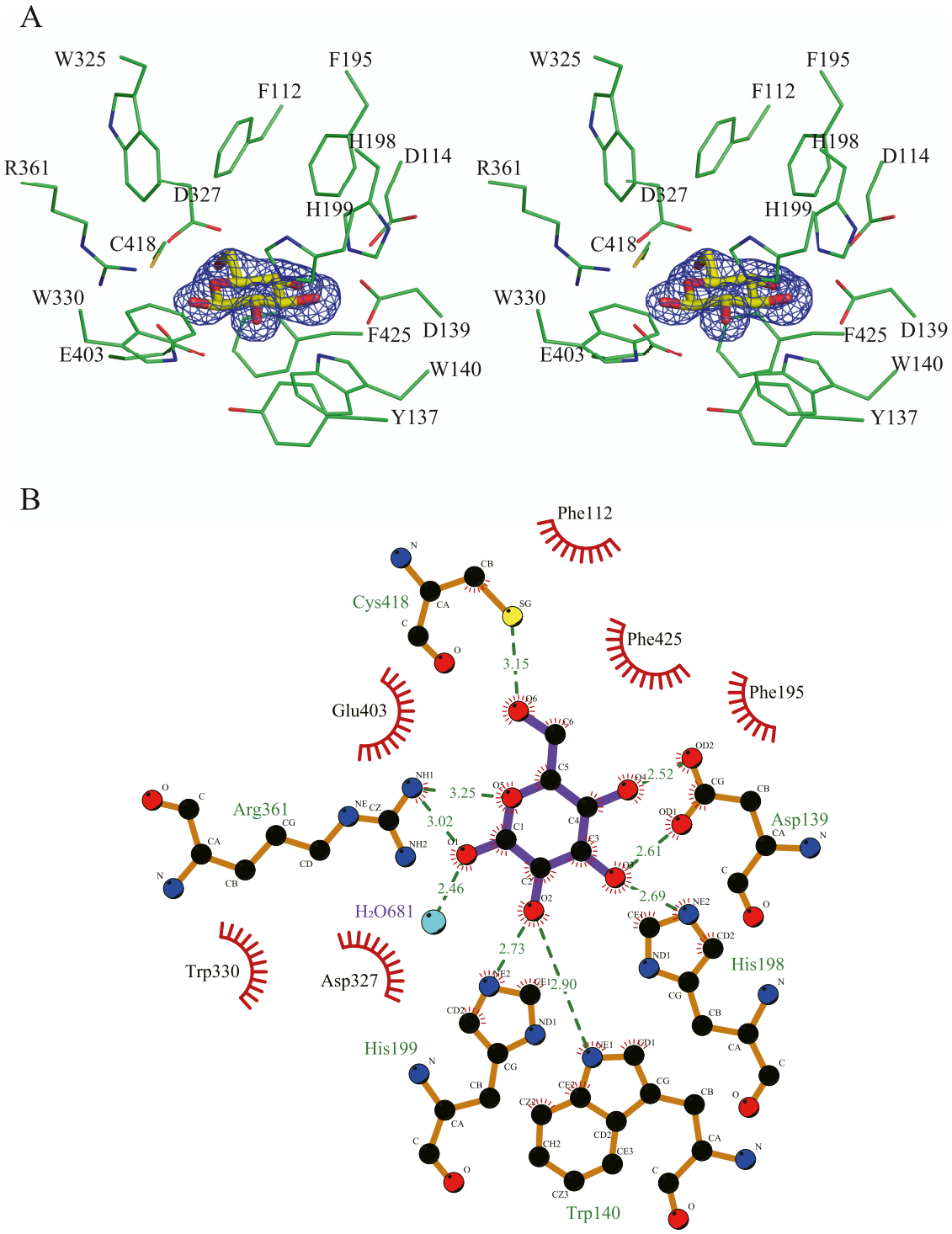


Figure 4. Active-site pocket of ClAgl29B bound β-L-glucose. (A) Electron density observed in the active-site pocket fit to β-L-glucose. The contour level of the polder map is 5σ. Stick models show amino acid residues located within 4 Å of the ligand. (B) Diagram of the β-L-glucose−ClAgl29B interaction. Hydrogen bonds are shown as green dotted lines, while the spoked arcs represent protein residues making nonbonded contact with β-L-glucose. The diagram was drawn by LigPlot+.[40]

entrapped the sugar, where each hydroxy group was in direct contact with one or more amino acid side chains (Figure 4B), 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 can 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. We initially predicted that Asp114 (corresponding to His34 hydrogen bonding to the axial O4 of α-L-fucoside in TmaFuc) stabilized the equatorial O4 of α-L-glucoside. However, the side chain of Asp114 in ClAgl29B was positioned 3.8 Å from the equatorial O4, which is slightly too far away to form a

https://doi.org/10.1021/acsomega.2c06991
ACS Omega 2022, 7, 47411−47423

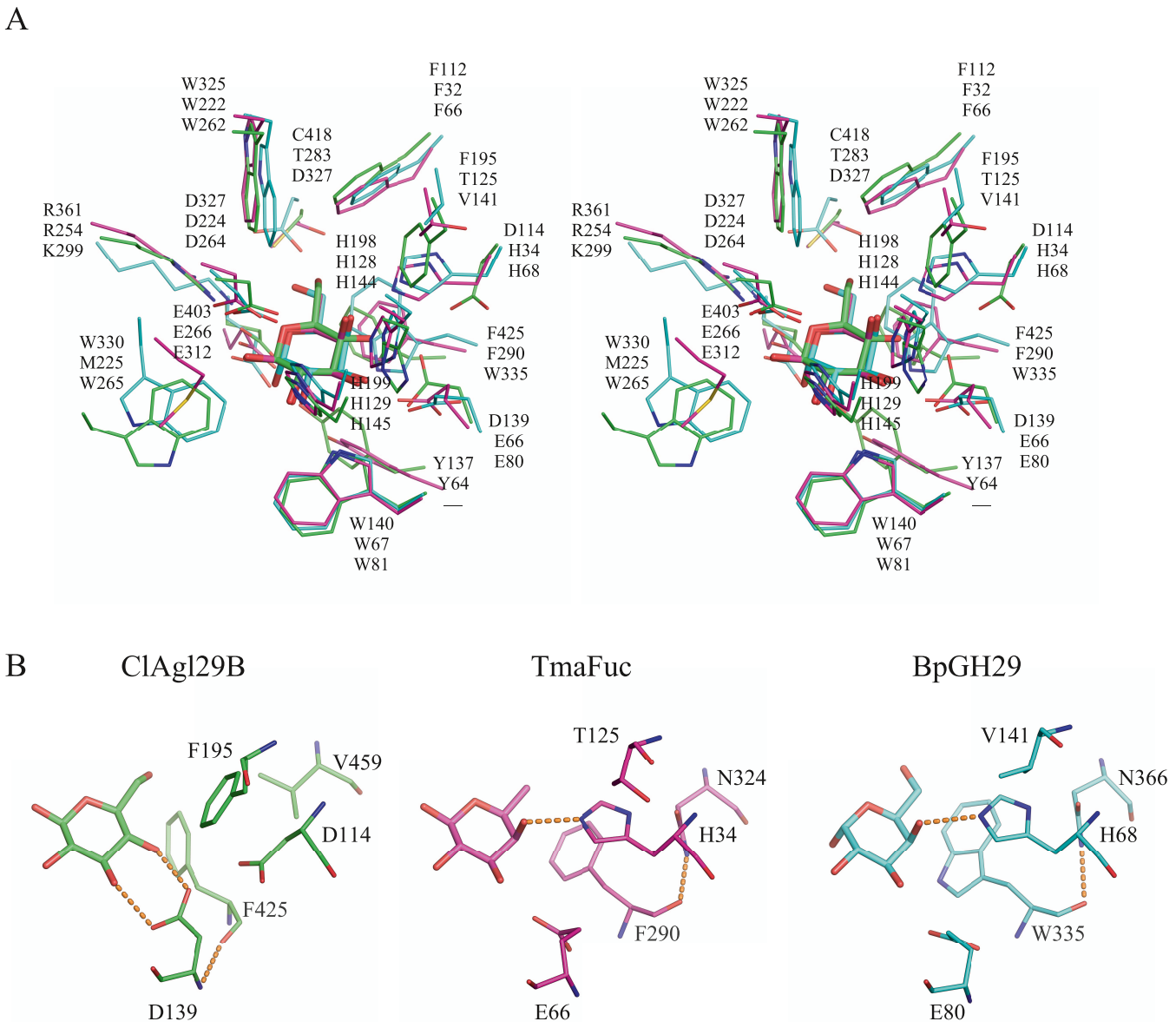


Figure 5. Comparison of the active-site pocket of ClAgl29B with TmaFuc and BpGH29. The active sites of ClAgl29B (green), TmaFuc (magenta), and BpGH29 (cyan) were aligned using PyMOL molecular graphic system version 2.0 (Schrödinger, LLC). (A) Stereo images of the residues involved in ligand binding (β-L-glucose for ClAgl29B, β-L-fucose for TmaFuc, and α-L-galactose for BpGH29). The upper, middle, and lower residue numbers are for ClAgl29B, TmaFuc, and BpGH29, respectively. (B) Structural comparison of the hydroxy group at C4. Orange dashed lines represent hydrogen bonds.

hydrogen bond. Trp325 and Tyr137 were located within 4 Å of the ligand and were part of the active-site pocket formation. 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 to be the same.

Comparison of the Active-Site Structures of α-L-Glucosidase and Other GH29 Enzymes. The active-site structure of ClAgl29B with L-glucose bound was compared to other GH29 enzymes to elucidate the specificity for α-L-glucoside (Figure 5A). We selected TmaFuc (PDB, 1ODU)[5] as a representative of GH29-A and α-L-galactosidase BpGH29 from *B. plebeius* (PDB, 7LK7)[20] as reference structures. The active-site pocket structures of these enzymes were similar, except for the recognition machinery for O6 and O4. No differences were apparent in the binding site of C6 among ClAgl29B, TmaFuc, and BpGH29, which have five aromatic side chains from β1, 2, 3, and 4 of the $(\beta/\alpha)_8$-barrel fold, as previously noted. BpGH29 has an Asp327 side chain that fixes O6, which is nearly identical to the Cys418 side chain of ClAgl29B. The corresponding residue in TmaFuc is Thr283 and could be hydrogen-bonded, which in line with weak α-L-galactosidase activity and structural features. The structures of ClAgl29B around the O4 atom of L-glucose were distinctly dissimilar to those of other GH29 enzymes (Figure 5B) and would be responsible for substrate specificity to α-L-glucoside. As mentioned above, Asp139 is responsible for the stabilization of equatorial O4, together with O3 of L-glucose in ClAgl29B. The corresponding residues were Glu66 in TmaFuc and Glu80 in BpGH29, and their Oε1 was hydrogen-bonded to O3 of L-fucose and L-galactose, respectively. His34 in TmaFuc and His68 in BpGH29 were responsible for stabilizing the axial O4 of L-fucose and L-galactose, whereas the corresponding Asp114

https://doi.org/10.1021/acsomega.2c06991
ACS Omega 2022, 7, 47411−47423

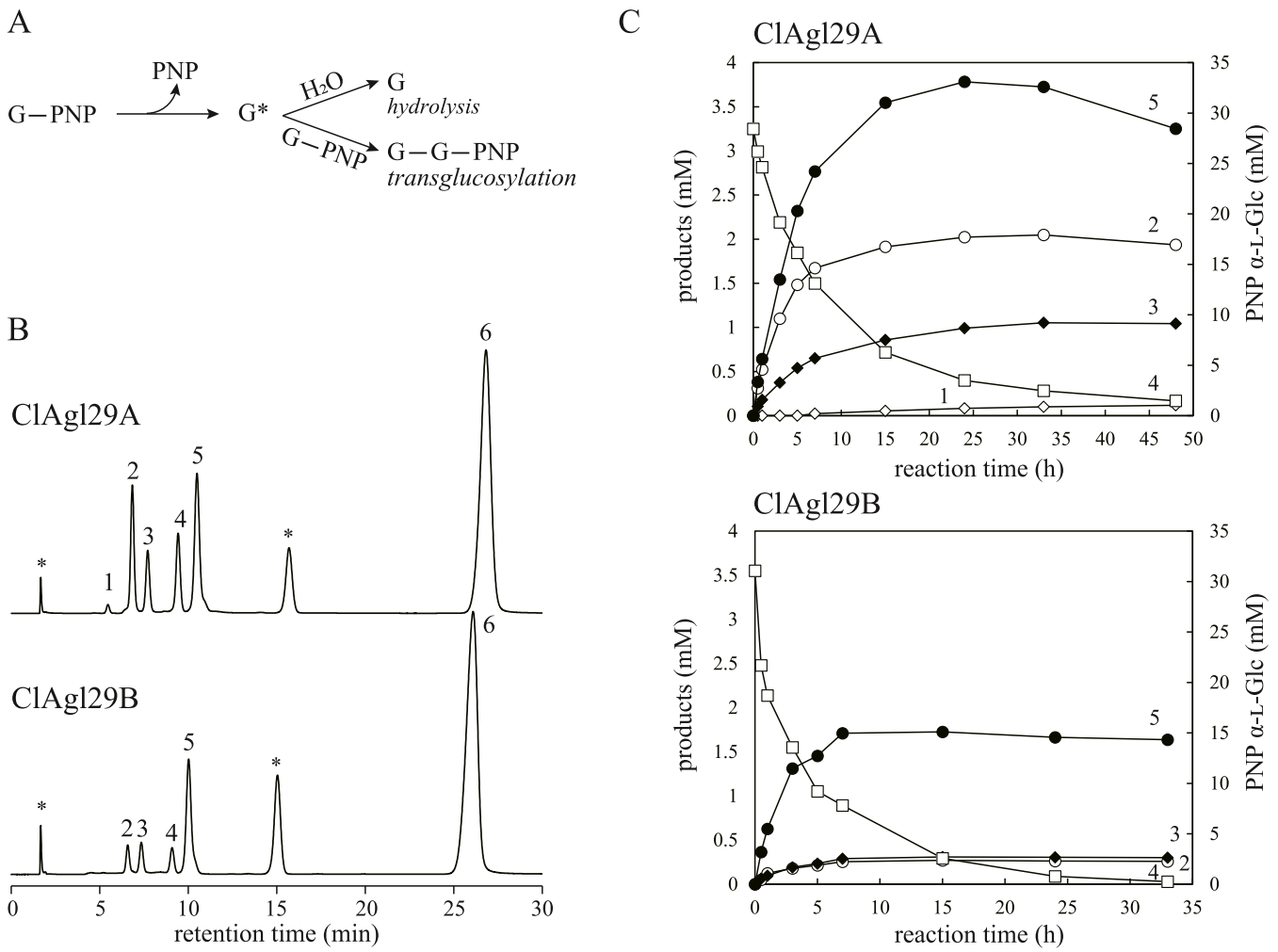


Figure 6. Transglucosylation properties of ClAgl29A and ClAgl29B. (A) Schematic representation of hydrolysis and transglucosylation by α-L-glucosidase. G* indicates the covalent intermediate of enzyme and glucose. (B) Reversed-phase-HPLC analysis of the reaction mixtures containing 30 mM PNP α-L-Glc and ClAgl29A and ClAgl29B in McIlvaine buffer, pH 5.5, that were incubated at 35 °C for 33 h. (1) PNP trisaccharide; (2) α-L-Glc*p*-(1 → 2)-α-L-Glc*p*-O-PNP; (3) α-L-Glc*p*-(1 → 6)-α-L-Glc*p*-O-PNP; (4) PNP α-L-Glc; (5) α-L-Glc*p*-(1 → 3)-α-L-Glc*p*-O-PNP; (6) PNP. The area of the peaks marked with asterisks was unchanged before and after the reaction. (C) Time-dependent change of the concentrations of transglucosylation products and PNP α-L-Glc. The concentrations were calculated using a calibration curve prepared from chromatographic peak areas of PNP α-D-glucoside standards with known concentrations. Closed circle, α-L-Glc*p*-(1 → 3)-α-L-Glc*p*-O-PNP; open circle, α-L-Glc*p*-(1 → 2)-α-L-Glc*p*-O-PNP; closed diamond, α-L-Glc*p*-(1 → 6)-α-L-Glc*p*-O-PNP; open diamond, PNP trisaccharide; and open square, PNP α-L-Glc. The numbers in the diagram correspond to the indications in (B).

in ClAgl29B was separated from the substrate. The bulky side chain of Phe195 appeared to prevent its interaction with the axial O4 of α-L-fucoside and L-galactoside in ClAgl29B. The Phe residue is replaced by Thr125 and Val141 in TmaFuc and BpGH29, respectively, and they are sufficiently small side chains that do not interfere with α-L-fucoside and L-galactoside binding. Minor variations in the spatial positions of Phe290 in TmaFuc, Trp335 in BpGH29, and Phe425 in ClAgl29B are likely to affect substrate preferences. Phe290 undergoes a slight conformational change upon ligand binding and packing against C4, C5, and C6 of L-fucose, and interferes with binding of the equatorial hydroxy group at C4.[5] BpGH29 and some α-L-fucosidases possess the corresponding Trp which also seems to play a role in excluding equatorial O4 from the active-site pocket. ClAgl29B has the counterpart Phe425, which has different spatial arrangements and does not hinder the binding of O4 to L-glucose. The spatial variation of these aromatic residues can be attributed to a difference in the main chain conformation: the φ and ψ angles are −92.0 and −48.5 of Phe290 in TmaFuc; −96.2 and −49.9 of Trp335 in BpGH29; and −92.9 and 107.1 of Phe425 in ClAgl29B. The disparate ψ angles of TmaFuc and BpGH29 are likely attributed to a hydrogen bond between the O atom of the aromatic residues and the Nδ2 atom of Asn residues at β7 and Asn324 in TmaFuc and Asn366 in BpGH29, while ClAgl29B has Val459, which cannot form a similar hydrogen bond. Alternatively, the O atom of Phe425 in ClAgl29B hydrogen bonds with the N atom of Asp139 in the small helix in the long insertion following β1. A subtle difference was also found at the pocket entrance. ClAgl29B and BpGH29 had Trp330 and Trp265 near the entrance, whereas TmaFuc had Met225 at the corresponding position. This difference resulted in a relatively narrower entrance in ClAgl29B and BpGH29, although it did not affect substrate specificity. We have discussed the comparison between ClAgl29B and other GH29 enzymes, where ClAgl29A and ClAgl29B have identical substrate recognition mechanisms.

https://doi.org/10.1021/acsomega.2c06991
ACS Omega 2022, 7, 47411−47423

Table 2. Enzyme Activity of GH29 Enzymes from *C. akajimensis*[a]

	caka_0074	caka_0506	caka_0522	caka_0523	caka_0536
PNP α-L-Fuc	1.4 ± 0.3	12 ± 0.4	2.1 ± 0.6	1.3 ± 0.8	3.1 ± 1.5
PNP α-L-Glc	1.3 ± 0.6	7.0 ± 0.2	8.2 ± 0.1	<0.12	6.0 ± 0.3
PNP α-L-Gal	1.4 ± 0.3	5.5 ± 1.9	7.5 ± 0.3	<0.12	4.5 ± 0.8

[a]Hydrolysis rates were determined using 2 mM substrate (PNP α-L-Fuc, PNP α-L-Glc, and PNP α-L-Gal). The unit of the values is nmol/min/mg.

Transglucosylation of ClAgl29A and ClAgl29B. Some glycoside hydrolases with a retaining mechanism have transglycosylation ability to form new glycosidic linkage in addition to hydrolysis. We investigated whether ClAgl29A and ClAgl29B can be used for the synthesis of novel α-L-glucosyl compounds via transglycosylation. The reaction of α-L-glucosidase with PNP α-L-Glc substrate results in cleavage of the glucosidic linkage, PNP release, and the glucosyl group forming a covalent intermediate with the enzyme. The reaction of the intermediate with a water molecule results in hydrolysis and L-glucose is released. Meanwhile, transglucosylation involves the reaction of the intermediate with the hydroxy group of another PNP α-L-Glc resulting in the formation of an α-L-glucosidic bond (Figure 6A). The transglucosylation ability was evaluated by measuring the ratio of the transglucosylation rate (v_{TG}) to the total reaction rate (PNP release rate, v_T) in the reaction of ClAgl29A and ClAgl29B with 30 mM PNP α-L-Glc as the substrate. The transglucosylation ratio (%) ((v_{TG}/v_T) × 100) was calculated from the equation (($v_T - v_H$)/v_T) × 100 since [v_T = L-glucose release rate (hydrolysis rate, v_H) + v_{TG}] was established. The v_T estimated from the released PNP was 1.05 μmol/min/mg (ClAgl29A) and 4.12 μmol/min/mg (ClAgl29B). The L-glucose release rate was 0.0361 μmol/min/mg (ClAgl29A) and 2.81 μmol/min/mg (ClAgl29B). The transglucosylation ratios calculated from these values were 60.3% (ClAgl29A) and 32.0% (ClAgl29B).

Next, we determined the structures of the transglucosylation products of PNP α-L-Glc. Both enzymes synthesized three types of PNP disaccharides: *p*-nitrophenyl (α-L-glucopyranosyl)-(1 → 3)-α-L-glucopyranoside (α-L-Glc*p*-(1 → 3)-α-L-Glc*p*-O-PNP), *p*-nitrophenyl (α-L-glucopyranosyl)-(1 → 2)-α-L-glucopyranoside (α-L-Glc*p*-(1 → 2)-α-L-Glc*p*-O-PNP), and *p*-nitrophenyl (α-L-glucopyranosyl)-(1 → 6)-α-L-glucopyranoside (α-L-Glc*p*-(1 → 6)-α-L-Glc*p*-O-PNP). ClAgl29B produced a small amount of PNP trisaccharide (Figure 6B). The formation of PNP trisaccharide can be attributed to one of the produced PNP disaccharides acting as an acceptor. The broad regioselectivity of transglucosylation of α-L-glucosidases is similar to that of TmaFuc, which can produce α-L-Fuc-(1 → 3)-α-L-Fuc, α-L-Fuc-(1 → 2)-β-D-Gal, and α-L-Fuc-(1 → 6)-β-D-Gal.[21] ClAgl29A produced the highest amount of α-L-Glc*p*-(1 → 3)-α-L-Glc*p*-O-PNP, followed by α-L-Glc*p*-(1 → 2)-α-L-Glc*p*-O-PNP, and α-L-Glc*p*-(1 → 6)-α-L-Glc*p*-O-PNP based on the analysis of temporal changes in the concentration of the products (Figure 6C). ClAgl29B mainly produces α-L-Glc*p*-(1 → 3)-α-L-Glc*p*-O-PNP. It also yielded almost equal amounts of α-L-Glc*p*-(1 → 2)-α-L-Glc*p*-O-PNP and α-L-Glc*p*-(1 → 6)-α-L-Glc*p*-O-PNP at approximately one-third of the concentration of α-L-Glc*p*-(1 → 3)-α-L-Glc*p*-O-PNP. The amount of transglucosylation product (PNP disaccharide) was higher in ClAgl29A (6.2 mM) than in ClAgl29B (2.2 mM) after 33 h of reaction. At this point, 92 and 99% of PNP α-L-Glc were consumed in ClAgl29A and ClAgl29B, resulting in theoretical yields of 45 and 15%, respectively. The difference in the amounts produced by the two enzymes may reflect their different transglucosylation ratio.

The differences in the transglucosylation abilities of ClAgl29A and ClAgl29B were predicted by comparing the architecture of their subsite +1 with the structure of TmaFuc-bound α-L-fucopyranosyl-(1 → 2)-1-azide-1-deoxy-β-L-fucopyranose (PDB, 2WSP).[22] Subsite +1 of ClAgl29A was presumed to consist of Thr121, Tyr125, Trp128, Leu263, Trp318, and Arg349. Subsite +1 of ClAgl29B was quite similar except that Gln275 replaced the Leu263 position. This difference may be responsible for the enhanced transglucosylation activity of ClAgl29A.

We succeeded in the enzymatic synthesis of an α-L-glucosidic linkage which was not previously reported. Oligosaccharides composed of D-series monosaccharides support human health by stimulating the growth and activity of beneficial bacteria in the intestines.[23,24] Although it is not known whether α-L-glucosyl oligosaccharides have such a function, this study provides clues to reveal their hidden functionality.

Substrate Specificity in GH29. We demonstrated that ClAgl29A and ClAgl29B have elaborate α-L-glucoside recognition machinery that differs from that of α-L-fucosidases. The presence of ClAgl29A and ClAgl29B implies that α-L-glucoside-containing compounds exist in nature. Furthermore, α-L-glucosidases may be widespread. A BLAST search using ClAgl29B as the query sequence revealed that the top 45 amino acid sequences of the hits had substrate recognition mechanisms of α-L-glucosidases (Table S1). Thus, these sequences which were derived from the *Bacteroidetes*, *Verrucomicrobia*, *Kiritimatiellaeota*, *Planctomycetes*, *Terrabacteria*, and *Armatimonadetes* bacterial phyla, and Asgard archaea, can be regarded as α-L-glucosidases. A phylogenetic tree of representative proteins was constructed based on 15% co-membership thresholds (RP15) of PF01120 (Pfam accession number), 45 probable α-L-glucosidases, and the sequences annotated as characterized proteins in CAZy using the unweighted pair group method with arithmetic averages (UPGMA) (Figure S2). The tree indicates that a single cluster for α-L-glucosidase is embedded in GH29-A and diverged later than the cluster of GH29-B, a subfamily of α-L-fucosidases with different aglycone substrate specificities.

The discovery of α-L-glucosidase in the present study was based on the diversity of the conserved His residue that stabilizes OH-4 of α-L-fucoside in GH29 α-L-fucosidase. ClAgl29A and ClAgl29B, containing Asp instead of His, were identified as α-glucosidases in this study; however, we found substitutions for other amino acids in GH29. The genome of *Coraliomargarita akajimensis* encodes 11 GH29 enzymes. Five of these enzymes had the His residue substituted for Val (Caka_0522 and Caka_0523), Cys (Caka_0506 and Caka_0536), and Phe (Caka_0074). Using facilely purified recombinant enzymes, the activities toward PNP α-L-Fuc, PNP α-L-Glc, and PNP α-L-Gal were investigated. The activity was quite low for all of these substrates (Table 2).

https://doi.org/10.1021/acsomega.2c06991
ACS Omega 2022, 7, 47411−47423

GH29-B enzymes have strict substrate specificity on the aglycone side and hardly hydrolyze PNP α-L-Fuc.[12] However, phylogenetic analysis showed that the five enzymes from *C. akajimensis* were classified as GH29-A with relaxed specificity (Figure S3). Caka_0506 prefers PNP α-L-Fuc, Caka_0522 prefers PNP α-L-Glc and PNP α-L-Gal, while Caka_0536 accepts all substrates (PNP α-L-Glc, PNP α-L-Fuc, and PNP α-L-Gal). However, detailed evaluation of their substrate specificity was not performed due to the relatively slow reaction rates. The specificity of GH29 enzymes may be more divergent than known to date and the substrate specificity of these enzymes requires further investigation.

This study identified the α-L-glucosidases ClAgl29A and ClAgl29B in GH29 by analyzing an expanded sequence space enabled by genome analysis. The active-site architectures exclude α-L-fucoside but allow α-L-glucoside, suggesting that ClAgl29A and ClAgl29B may have evolved as α-L-glucosidases. On the other hand, further analysis of substrate specificity for α-L-glycosides, including α-L-xylosides and α-L-quinovosides (6-deoxy L-glucosides), will be necessary to demonstrate that these enzymes are true α-L-glucosidases. It is unclear whether these enzymes evolved under selection pressure to utilize L-glucoside, or if they were neutrally created and retained by chance. The discovery of α-L-glucosidase raises the possibility that α-L-glucoside exists in nature.

EXPERIMENTAL PROCEDURES

Chemicals. L-Galactose was prepared from D-galacturonic acid (Fujifilm Wako Pure Chemical Corporation, Osaka, Japan) according to a previous method except that Amberlite IR120 (H^+) was used instead of sulfuric acid to adjust the pH during the reduction of L-galactono-1,4-lactone.[25] The PNP group addition reaction was performed as previously described.[26] ^{1}H NMR (500 MHz) spectra were recorded using a Bruker AVANCE I spectrometer (Bruker Corporation, Billerica, MA). PNP α-L-Glc: ^{1}H NMR (500 MHz, D_2O) δ (ppm): 8.26 (d, J = 9.3 Hz, 2H, PNP), 7.30 (d, J = 9.3 Hz, 2H, PNP), 5.82 (d, J = 3.7 Hz, 1H, H1), 3.94 (t, J = 9.5 Hz, 1H, H3), 3.78 (dd, J = 3.7 and 9.8 Hz, 1H, H2), 3.74 (m, 2H, H6), 3.68 (dd, J = 3.5 and 6.3 Hz, 1H, H5), 3.53 (t, J = 9.5 Hz, 1H, H4). PNP α-L-Gal: ^{1}H NMR (500 MHz, D_2O) δ (ppm): 8.26 (d, J = 9.3 Hz, 2H, PNP), 7.30 (d, J = 9.3 Hz, 2H, PNP), 5.86 (d, J = 3.8 Hz, 1H, H1), 4.11 (dd, J = 3.3 and 10.3 Hz, 1H, H3), 4.06 (d, J = 3.3 Hz, 1H, H4), 4.04 (dd, J = 3.8 and 10.3 Hz 1H, H2), 3.98 (dd, J = 4.8 and 7.5 Hz, 1H, H5), 3.72 (dd, J = 7.5 and 11.9 Hz, 1H, H6a), 3.68 (dd, J = 4.8 and 11.9 Hz, 1H, H6b). L-glucose was purchased from TCI (Tokyo, Japan). PNP α-L-Fuc was obtained from Sigma-Aldrich (St. Louis, MA).

Sequence Analysis. Full-length amino acid sequences and RP15 belonging to the alpha_L_fucos family (PF01120) were obtained from the Pfam database (https://pfam.xfam.org/family/PF01120). The 8,266 obtained sequences were aligned using MAFFT version 7 (https://mafft.cbrc.jp/alignment/software/).[27] Refinement of the multiple sequence alignment involved excluding sequences with 220 amino acids or less. FASTA and BLAST searches were performed using GenomeNet (https://www.genome.jp/tools/fasta/) and NCBI (https://blast.ncbi.nlm.nih.gov/Blast.cgi) databases, respectively. Signal peptide prediction was performed using SignalP-5.0[28] (https://services.healthtech.dtu.dk/service.php?SignalP-5.0). Phylogenetic trees were constructed using the UPGMA and neighbor-joining method using the Jones−Taylor−Thornton model on the MAFFT version 7 website. iTOL v6 (https://itol.embl.de/)[29] was used to visualize the tree.

Cloning. Genes encoding ClAgl29A (EKB48091.1) and ClAgl29B (EKB48090.1), whose locus_tags were B879_03288 and B879_03287, respectively, were synthesized at Eurofins Genomics K.K. (Tokyo, Japan). The genes were amplified by PCR and cloned into the Novagen pET28a vector (Merck; Darmstadt, Germany) using an iVEC3 in vivo *E. coli* cloning system.[30] The gene encoding TmaFuc (TM0306; nucleotide 326116 to 327465 of AE000512.1) in *T. maritima* NBRC 100826^T genomic DNA was amplified by PCR and cloned into pET28a using *Nde*I and *Hin*dIII restriction sites. Genes encoding Caka_0074, Caka_0506, Caka_0522, Caka_0523, and Caka_0536 were cloned based on the *C. akajimensis* DSM 45221 complete genome (CP001998.1): 79209..80672, 649949..651400, 670716..672635, 672662..674473, and 688656..690170, respectively. The regions were amplified by PCR using genomic DNA of *C. akajimensis* JCM 23193^T as a template which was provided by RIKEN BRC through the National BioResource Project of MEXT, Japan. The PCR products were cloned into pET28a using a T5 exonuclease DNA assembly method.[31] PrimeSTAR MAX DNA Polymerase (Takara Bio Inc., Shiga, Japan) was used for all PCR amplifications, and reaction conditions were followed according to the attached protocol. The primers used for the PCR are listed in Table S2.

Production and Purification of Recombinant Enzymes. ClAgl29A and ClAgl29B: *E. coli* Rosetta (DE3) (Novagen-Merck) transformants with pET28a-derived plasmids were inoculated in ZYP-5052 medium[32] supplemented with 100 μg·mL^{-1} kanamycin and 30 μg·mL^{-1} chloramphenicol and cultured at 30 °C for 24 h. Cells collected by centrifugation were resuspended in 10 mM sodium phosphate buffer (pH 7.5) containing 300 mM sodium chloride (buffer A), and sonicated. Cell debris was removed by centrifugation, and the supernatant was loaded onto a Ni-chelating Sepharose Fast Flow column (Cytiva; Tokyo, Japan) preequilibrated with buffer A containing 10 mM imidazole. After washing the column with buffer A containing 30 mM imidazole, the absorbed proteins were eluted using 300 mM imidazole in buffer A. The active fractions were dialyzed against 10 mM sodium phosphate buffer (pH 7.5) and concentrated using Amicon Ultra-15 centrifugal filter units (30,000 NMWL, Millipore, Merck).

TmaFuc: The procedure followed a previously published method;[5] however, 50 mM sodium phosphate buffer (pH 8.0) containing 300 mM sodium chloride was used as the initial buffer for Ni^{2+} affinity column chromatography, and the purified enzyme obtained by Ni^{2+} affinity chromatography was dialyzed against 50 mM sodium phosphate buffer (pH 8.0) and concentrated using a Vivaspin 20 10,000 MWCO ultrafiltration unit (Sartorius AG; Göttingen, Germany).

Caka_0074, Caka_0506, Caka_0522, Caka_0523, and Caka_0536: *E. coli* BL21 (DE3) for Caka_0506 and Caka_0523 and Rosetta-gami 2 (DE3) (Novagen-Merck) for Caka_0074 Caka_0522, and Caka_0536 were transformed with the appropriate pET28a-derived plasmids. Each transformant was inoculated in LB medium supplemented with 100 μg·mL^{-1} kanamycin. Rosetta-gami 2 (DE3) transformants also contained 30 μg·mL^{-1} chloramphenicol in the medium. Expression of recombinant proteins was induced by incubation at 18 °C (Caka_0074, Caka_0522, and Caka_0536) or 30 °C

https://doi.org/10.1021/acsomega.2c06991
ACS Omega 2022, 7, 47411−47423

Table 3. Summary of Data Collection and Refinement Statistics[a]

	ClAgl29B/L-glucose	ClAgl29B	ClAgl29A
PDB ID	7XSH	7XSG	7XSF
	data collection		
beamline	PF AR-NE3A	PF AR-NE3A	SLS X06SA
wavelength	1.00000	1.00000	1.00000
space group	$P2_12_12_1$	$P2_12_12_1$	$P12_11$
unit cell			
a, *b*, *c* (Å)	72.39 121.58 166.37	73.13 123.41 166.60	74.58 95.52 83.55 β = 97.33°
resolution range (Å)	46.60−1.71 (1.81−1.71)	45.42−1.61 (1.71−1.61)	47.81−2.01 (2.13−2.01)
total no. of reflections	1,058,673 (157,759)	1,292,170 (201,560)	539,809 (86368)
no. of unique reflections	159,024 (25,291)	194,275 (30357)	76,908 (12,196)
multiplicity	6.7 (6.2)	6.7 (6.6)	7.0 (7.1)
completeness (%)	99.8 (99.1)	99.5 (97.0)	98.9 (97.3)
mean $I/\sigma(I)$	17.44 (2.37)	22.17 (4.51)	14.80 (2.45)
R_{mean}	10.1 (94.2)	6.9 (45.9)	9.0 (96.2)
R_{sym}	9.3 (86.3)	6.3 (42.3)	8.3 (89.1)
CC1/2	0.999 (0.810)	0.999 (0.942)	0.999 (0.90)
	refinement		
R_{work}	17.9	17.2	20.5
R_{free}	20.2	18.8	24.3
number of atoms			
macromolecules	8938	8985	8634
ligands	76	40	15
solvent		1008	203
RMSD values from ideal			
bonds	0.009	0.009	0.007
angles	1.39	1.43	1.25
Ramachandran (%)			
favored	96.88	97.23	96.56
allowed	3.12	2.77	3.44
outliers	0.00	0.00	0.00
rotamer outliers (%)	1.13	0.82	0.76
clashscore	1.42	1.80	2.19

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

(Caka_0506 and Caka_0523) without the addition of IPTG. The recombinant proteins were purified using the Amicon Pro purification system (30,000 NMWL, Millipore−Merck). Sodium phosphate buffer (20 mM, pH 7.5) containing 500 mM sodium chloride was used as the basic buffer during purification and was subsequently replaced with 20 mM sodium phosphate buffer, pH 7.5.

Protein concentrations of ClAgl29A and ClAgl29B 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 extinction coefficient values calculated from the amino acid analysis were 1.38×10^5 $M^{-1}cm^{-1}$ for ClAgl29A and 1.34×10^5 $M^{-1}cm^{-1}$ for ClAgl29B. 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.

Enzyme Assay. α-L-Glucosidase activity was determined by measuring the increase in PNP during its hydrolysis. The standard reaction was performed in McIlvaine buffer (pH 5.5) containing 2 mM PNP α-L-Glc and an appropriately diluted enzyme at 35 °C for 10 min. The enzyme was diluted with McIlvaine buffer (pH 5.5) containing 0.1 mg/mL bovine serum albumin (BSA). The reaction was terminated by the addition of two volumes of 1 M sodium carbonate. The amount of PNP released was measured based on the absorption at 400 nm using a 1 cm cuvette with a molar extinction coefficient of 5,560 M^{-1} cm^{-1}. One unit of enzyme was defined as the amount of enzyme that produced 1 μmol of PNP per minute under the above conditions. The hydrolysis rates for 2 mM PNP α-L-Gal and 2 mM PNP α-L-Fuc were measured under the same conditions as described above.

The optimal pH for hydrolysis of 2 mM PNP α-L-Glc was measured at various pH values using McIlvaine buffer (pH 3.0−8.0) with 67 nM ClAgl29A and 55 nM ClAgl29B. The pH stability was assessed by incubating ClAgl29A (1.34 μM) and ClAgl29B (1.37 μM) in twofold-diluted Britton−Robinson buffer (pH 2.5−11.5) and 50 mM glycine−NaOH (pH 11.5−12.0) containing 0.1 mg/mL BSA at 4 °C for 24 h or at 35 °C for 3 h, followed by measurement of the residual activity. The optimal reaction temperature was determined by measuring the hydrolysis rate at various temperatures using 67 nM ClAgl29A and 16 nM (ClAgl29B). The thermal stability of ClAgl29A (680 nM) and ClAgl29B (540 nM) was measured by diluting each enzyme with an equal volume of McIlvaine buffer (pH 5.5) containing 0.5 mg/mL BSA and incubation at each temperature (30−70 °C) for 15 min, followed by measurement of their residual activities.

https://doi.org/10.1021/acsomega.2c06991
ACS Omega 2022, 7, 47411−47423

The initial velocities of PNP α-L-Glc (0.5−6 mM for ClAgl29A, 0.25−4 mM for ClAgl29B) and those of PNP α-L-Fuc (0.04−0.6 mM for ClAgl29A, 0.05−0.5 mM for ClAgl29B) were measured under standard conditions. The enzyme concentrations were 13 nM ClAgl29A for PNP α-L-Glc, 890 nM ClAgl29A for PNP α-L-Fuc, 8.2 nM ClAgl29B for PNP α-L-Glc, and 550 nM ClAgl29B for PNP α-L-Fuc. Kinetic parameters 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).

Crystallization, Data Collection, Structure Determination, and Refinement. The enzymes used for crystallization were purified via anion exchange column chromatography using DEAE Sepharose Fast Flow (Cytiva) and Ni-affinity column chromatography. ClAgl29B was further purified by gel filtration column chromatography using a Sephacryl S-200 HR column (Cytiva). Chromatography was performed using 10 mM sodium phosphate buffer (pH 7.5) as the mobile phase. Anion exchange column chromatography involved the elution of proteins by linearly increasing the ionic strength with sodium chloride. The active fractions were dialyzed against 10 mM sodium phosphate buffer (pH 7.5) and concentrated to 10 mg/mL using an Amicon Ultra-15 centrifugal filter unit (30,000 NMWL).

All crystallization procedures were performed at 20 °C. Initial crystallization trials were performed using the sitting-drop vapor diffusion method in 96-well plates using PEGRx1, PEGRx2, SaltRx1, and SaltRx2 screen kits (Hampton Research, Aliso Viejo, CA). Microcrystals appeared under several conditions of PEGRx1 and PEGRx2. The crystallization conditions were optimized by varying the concentration of poly(ethylene glycol) (PEG) using the hanging-drop vapor diffusion method. ClAgl29A crystals obtained in a drop composed of 2 μL of ClAgl29A (10 mg·mL^{-1}), 2 μL of reservoir solution [0.1 M sodium citrate tribasic dihydrate (pH 5.5), 16% (w/v) PEG 3,350], and 1 μL of 1 mM L-fucose. Crystals were soaked in a cryoprotectant [0.1 M sodium citrate tribasic dihydrate (pH 5.5), 18% (w/v) PEG 3,350, 10% (v/v) glycerol, 1 mM L-fucose]. ClAgl29B crystals were obtained in a drop composed of 2 μL of ClAgl29B (5 mg·mL^{-1}), 2 μL of reservoir solution [0.1 M citric acid (pH 3.5), 5% (v/v) 2-propanol, 5% (w/v) PEG 20,000], and 1 μL of 1 mM L-glucose. The crystals were cryoprotected with 0.1 M citric acid (pH 3.5), 5% (v/v) 2-propanol, 6% (w/v) PEG 20,000, 20% (v/v) glycerol, and 1 mM L-Glc for the ligand-free from, and 0.1 M citric acid (pH 3.5), 5% (v/v) 2-propanol, 6% (w/v) PEG 20,000, 20% (v/v) glycerol, and 30 mM L-Glc for the L-glucose-bound form. All crystals were flash-frozen in liquid nitrogen. The diffraction datasets were collected at the X06SA beamline of the Swiss Light Source, Switzerland, for ClAgl29A, and at the AR-NE3A beamline of the Photon Factory, KEK, Japan, for ClAgl29B using an EIGER 16M detector (Dectris; Dättwil AG, Switzerland) and Pilatus 2m-F pixel detector (Dectris), respectively, at a wavelength of 1.00000 Å. The dataset was 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.[33]

The structure of ClAgl29A was determined by the molecular replacement method with Phenix. automr software[34,35] using the protein coordinates of TmaFuc (PDB, 2XSD) as a search model. The refinement was converged by several cycles of manual model corrections with Coot software[36,37] and further refined using Refmac5[38] and phenix.refine[39] programs. The data processing and refinement statistics are presented in Table 3.

Transglucosylation. The reaction mixture consisted of 30 mM PNP α-L-Glc, diluted enzymes (1.34 μM for ClAgl29A and 0.274 μM for ClAgl29B), and 0.1 mg/mL BSA in McIlvaine buffer (pH 5.5) and was incubated at 35 °C. An aliquot of the reaction solution was collected at 3, 6, 9, and 12 min, and the reaction was stopped by treatment at 100 °C for 2 min. Two volumes of 1.0 M sodium carbonate were added to the collected sample, and the absorbance value of the solution was measured at 400 nm to determine the PNP concentration. The L-glucose concentration was determined by high-performance anion exchange chromatography−pulsed amperometric detection (HPAEC-PAD) using a Dionex ICS-3000 system (Dionex-Thermo Fisher Scientific; Waltham, MA) equipped with a CarboPac PA1 analytical column (4 mm × 250 mm; Dionex/Thermo Fisher Scientific). The column was preequilibrated with the eluent before chromatographic analysis. Separation conditions included a flow rate of 0.8 mL.min^{-1}, sample injection of 10 μL (containing sorbitol internal standard) under isocratic elution with 160 mM sodium hydroxide, which was prepared from super special grade 50% sodium hydroxide solution (Fujifilm Wako Pure Chemical Corporation). The concentration was calculated using a calibration curve prepared from the chromatographic peak areas of sorbitol and known L-glucose standards. Chromatograms were evaluated using Chromeleon software (Dionex/Thermo Fisher Scientific).

PNP glycosides were quantified using 1.34 μM ClAgl29A and 0.274 μM ClAgl29B. The sample was desalted using Amberlite MB-4 resin (Organo Co., Tokyo, Japan) and separated by reversed-phase HPLC (RP-HPLC) using a Jasco system (Tokyo, Japan) equipped with a Cosmosil 5C18-PAQ analytical column (4 mm × 150 mm; Nacalai Tesque, Kyoto, Japan). The separation was performed at 40 °C, the mobile phase consisted of methanol/water = 20/80 (v/v), and the flow rate was 1.0 mL.min^{-1}. Each PNP glycoside was detected by measuring the absorbance at 313 nm, and the concentration was calculated using a calibration curve prepared from the chromatographic peak areas of PNP α-D-glucoside standard.

The molecular masses of the products were confirmed by electrospray ionization (ESI)-MS using an Exactive spectrometer (Thermo Fisher Scientific, Inc.). ^{1}H NMR (500 MHz) and ^{13}C (126 MHz), COSY, HSQC, HSQC-TOCSY, H2BC, and HMBC NMR spectra were recorded with a Bruker AVANCE I system. *p*-Nitrophenyl trisaccharide was only detected by MS because of the low yield; MS: *m*/*z* calcd: 648.17 [M + Na]$^{+}$; found: 648.18.

p-Nitrophenyl (α-L-glucopyranosyl)-(1 → 3)-α-L-glucopyranoside [α-L-Glcp-(1→3)-α-L-Glcp-O-PNP] (Figure S4). ^{1}H NMR (500 MHz, D_2O) δ (ppm): 8.25 (d, *J* = 9.0 Hz, 2H, PNP), 7.31 (d, *J* = 9.0 Hz, 2H, PNP), 5.85 (s, 1H, H-1), 5.45 (d, *J* = 3.4 Hz, 1H, H-1′), 4.12 (t, *J* = 9.2, 1H, H-3), 4.08 (m, 1H, H-5′), 3.91 (d, *J* = 11.3 Hz, 1H, H-2), 3.90 (m, 1H, H-6a′), 3.83 (m, H, H-6b′), 3.80 (m, 1H, H-3′), 3.79 (m, 1H, H-4), 3.76 (brd, 2H, H-6), 3.71 (m, 1H, H-5), 3.63 (dd, *J* = 3.4 and 10.1, 1H, H-2′), 3.50 (t, *J* = 9.5, 1H, H-4′); ^{13}C NMR (126 MHz, D_2O) δ (ppm): 162.1 (PNP), 143.2 (PNP), 126.9 (PNP), 117.6 (PNP), 100.1 (C-1′), 97.6 (C-1), 80.4 (C-3), 73.7 (C-3′), 73.5 (C-5), 72.7 (C-5′), 72.6 (C-2′), 70.5 (C-4), 70.4 (C-2) 70.3 (C-4′), 61.3 (C-6′), 60.9 (C-6); MS: *m*/*z* calcd: 486.12 [M + Na]$^{+}$; found: 486.12.

https://doi.org/10.1021/acsomega.2c06991
ACS Omega 2022, 7, 47411−47423

p-Nitrophenyl (α-L-glucopyranosyl)-(1→2)-α-L-glucopyranoside [α-L-Glcp-(1→2)-α-L-Glcp-O-PNP] (Figure S5). ^{1}H NMR (500 MHz, D_2O) δ (ppm): 8.26 (d, J = 8.5 Hz, 2H, PNP), 7.35 (d, J = 8.5 Hz, 2H, PNP), 6.05 (s, 1H, H-1), 5.12 (d, J = 3.2 Hz, 1H, H-1′), 4.09 (t, J = 9.5, 1H, H-3), 3.99 (d, J = 9.8, 1H, H-5′), 3.92 (d, J = 12.4 Hz, 1H, H-2), 3.91 (m, 1H, H-6a′), 3.91 (m, H, H-6a), 3.83 (m, 1H, H-3′), 3.83 (m, 1H, H-6b′), 3.80 (m, 1H, H-6b), 3.75 (d, J = 13.0, 1H, H-5), 3.62 (t, J = 9.4, 1H, H4), 3.55 (dd, J = 3.2 and 9.6, 1H, H-2′), 3.48 (t, J = 9.6, 1H, H-4′); ^{13}C NMR (126 MHz, D_2O) δ (ppm): 162.3 (PNP), 143.3 (PNP), 126.9 (PNP), 117.8 (PNP), 96.9 (C-1′), 94.9 (C-1), 75.6 (C-2), 73.6 (C-5′), 73.5 (C-3′), 72.1 (C-3), 71.9 (C-2′), 70.2 (C-4′), 69.9 (C-4), 61.2 (C-6′), 61.1 (C-6); MS: m/z calcd: 486.12 $[M + Na]^+$; found: 486.12.

p-Nitrophenyl (α-L-glucopyranosyl)-(1→6)-α-L-glucopyranoside [α-L-Glcp-(1→6)-α-L-Glcp-O-PNP] (Figure S6). ^{1}H NMR (500 MHz, D_2O) δ (ppm): 8.31 (d, J = 8.4 Hz, 2H, PNP), 7.34 (d, J = 8.4 Hz, 2H, PNP), 5.87 (d, J = 2.3 Hz, 1H, H-1), 4.90 (s, 1H, H-1′), 3.98 (t, J = 9.4, 1H, H-3), 3.94 (m, 1H, H-6a), 3.91 (d, J = 11.3 Hz, 1H, H-2), 3.91 (m, 1H, H-5), 3.91 (m, H, H-6a′), 3.77 (dd, J = 4.6 and 12.0, 1H, H-6b′), 3.73 (m, 1H, H-6b), 3.69 (m, 1H, H-5′), 3.63 (dd, J = 9.1 and 9.4, 1H, H-4), 3.63 (m, 1H, H-3′), 3.55 (d, J = 11.0, 1H, H-2′), 3.43 (t, J = 9.3, 1H, H-4′); ^{13}C NMR (126 MHz, D_2O) δ (ppm): 162.1 (PNP), 143.3 (PNP), 126.9 (PNP), 117.7 (PNP), 98.5 (C-1′), 97.4 (C-1), 74.0 (C-3), 73.9 (C-3′), 72.6 (C-5), 72.2 (C-2′), 71.7 (C-2), 70.3 (C-4′), 70.2 (C-4), 66.3 (C-6), 61.3 (C-6′); MS: m/z calcd: 502.10 $[M + Na]^+$; found: 502.09.

■ ASSOCIATED CONTENT

Data Availability Statement

The atomic coordinates were deposited in the Protein Data Bank with accession codes 7XSF, 7XSG, and 7XSH. The authors declare that all data are contained within the manuscript.

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsomega.2c06991.

Michaelis−Menten analysis (Figure S1); UPGMA phylogenetic tree of RP15 of Alpha_L_fuc (PF01120) (Figure S2); neighbor-joining phylogenetic tree of structure-known α-L-fucosidases, five GH29 sequences from *C. akajimensis*, ClAgl29A, and ClAgl29B (Figure S3); NMR spectra of α-L-Glc*p*-(1 → 3)-α-L-Glc*p*-O-PNP (Figure S4), α-L-Glc*p*-(1→2)-α-L-Glc*p*-O-PNP (Figure S5); and α-L-Glc*p*-(1 → 6)-α-L-Glc*p*-O-PNP (Figure S6); list of putative α-L-glucosidases (Table S1); and oligonucleotide sequences used in this study (Table S2) (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Takayoshi Tagami − *Research Faculty of Agriculture, Hokkaido University, Sapporo, Hokkaido 060-8589, Japan*; Email: tagami@abs.agr.hokudai.ac.jp

Masayuki Okuyama − *Research Faculty of Agriculture, Hokkaido University, Sapporo, Hokkaido 060-8589, Japan*; orcid.org/0000-0001-7013-6032; Email: okuyama@abs.agr.hokudai.ac.jp

Authors

Rikako Shishiuchi − *Research Faculty of Agriculture, Hokkaido University, Sapporo, Hokkaido 060-8589, Japan*

Hyejin Kang − *Research Faculty of Agriculture, Hokkaido University, Sapporo, Hokkaido 060-8589, Japan*

Yoshitaka Ueda − *Research Faculty of Agriculture, Hokkaido University, Sapporo, Hokkaido 060-8589, Japan*

Weeranuch Lang − *Research Faculty of Agriculture, Hokkaido University, Sapporo, Hokkaido 060-8589, Japan*

Atsuo Kimura − *Research Faculty of Agriculture, Hokkaido University, Sapporo, Hokkaido 060-8589, Japan*

Complete contact information is available at: https://pubs.acs.org/10.1021/acsomega.2c06991

Author Contributions

‡R.S. and H.K. contributed equally to this work.

Funding

This work was partly supported by the Japanese Society for the Promotion of Science KAKENHI (grant no. 18K19159).

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors thank the staff of the Instrumental Analysis Division of the Creative Research Institution at Hokkaido University for conducting amino acid analysis and mass spectrometry analysis and the staff of the GC-MS & NMR Laboratory at the Research Faculty of Agriculture at Hokkaido University for NMR analysis. They acknowledge the Paul Scherrer Institute (Villigen, Switzerland) and Photon Factory (Tsukuba, Japan) for providing synchrotron radiation beamtime at beamline X06SA of the SLS (proposal numbers: 20191094 and 20191134) and AR-NE3A, respectively, and thank all beamline staff for assistance. They also thank Editage (www. editage. com) for English language editing.

■ ABBREVIATIONS

ESI, electrospray ionization; HPAEC-PAD, high-performance anion exchange chromatography−pulsed amperometric detection; UPGMA, unweighted pair group method with arithmetic averages

■ REFERENCES

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

(2) Henrissat, B.; Davies, G. Structural and sequence-based classification of glycoside hydrolases. *Curr. Opin. Struct. Biol.* **1997**, *7*, 637−644.

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

(4) Kempton, J. B.; Withers, S. G. Mechanism of *Agrobacterium* β-glucosidase: Kinetic studies. *Biochemistry* **1992**, *31*, 9961−9969.

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

(6) Summers, E. L.; Moon, C. D.; Atua, R.; Arcus, V. L. The structure of a glycoside hydrolase 29 family member from a rumen bacterium reveals unique, dual carbohydrate-binding domains. *Acta Crystallogr., Sect. F: Struct. Biol. Commun.* **2016**, *72*, 750−761.

(7) Schuster-Böckler, B.; Schultz, J.; Rahmann, S. HMM Logos for visualization of protein families. *BMC Bioinf.* **2004**, *5*, No. 7.
(8) Okuyama, M.; Saburi, W.; Mori, H.; Kimura, A. α-Glucosidases and α-1,4-glucan lyases: structures, functions, and physiological actions. *Cell. Mol. Life Sci.* **2016**, *73*, 2727−2751.
(9) Ernst, H. A.; Lo Leggio, L.; Willemoës, M.; Leonard, G.; Blum, P.; Larsen, S. S. Structure of the *Sulfolobus solfataricus* α-glucosidase: implications for domain conservation and substrate recognition in GH31. *J. Mol. Biol.* **2006**, *358*, 1106−1124.
(10) Okuyama, M.; Kaneko, A.; Mori, H.; Chiba, S.; Kimura, A. Structural elements to convert *Escherichia coli* α-xylosidase (YicI) into α-glucosidase. *FEBS Lett.* **2006**, *580*, 2707−2711.
(11) Janeček, Š.; Svensson, B.; MacGregor, E. A. A remote but significant sequence homology between glycoside hydrolase clan GH-H and family GH31. *FEBS Lett.* **2007**, *581*, 1261−1268.
(12) Sakurama, H.; Tsutsumi, E.; Ashida, H.; Katayama, T.; Yamamoto, K.; Kumagai, H. Differences in the substrate specificities and active-site structures of two α-L-fucosidases (glycoside hydrolase family 29) from *Bacteroides thetaiotaomicron*. *Biosci., Biotechnol., Biochem.* **2012**, *76*, 1022−1024.
(13) Fersht, A. *Structure and Mechanism in Protein Science: a Guide to Enzyme Catalysis and Protein Folding*; W. H. Freeman and Company: New York, 1998; pp 349−376.
(14) Kumar, P. A.; Srinivas, T. N. R.; Madhu, S.; Sravan, R.; Singh, S.; Naqvi, S. W. A.; Mayilraj, S.; Shivaji, S. *Cecembia lonarensis* gen. nov., sp. nov., a haloalkalitolerant bacterium of the family Cyclobacteriaceae, isolated from a haloalkaline lake and emended descriptions of the genera *Indibacter*, *Nitritalea* and *Belliella*. *Int. J. Syst. Evol. Microbiol.* **2012**, *62*, 2252−2258.
(15) Krissinel, E.; Henrick, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **2007**, *372*, 774−797.
(16) Kovalová, T.; Kovaľ, T.; Benešová, E.; Vodičková, P.; Spiwok, V.; Lipovová, P.; Dohnálek, J. 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* **2019**, *29*, 59−73.
(17) Coyle, T.; Wu, L.; Debowski, A. W.; Davies, G. J.; Stubbs, K. A. Synthetic and crystallographic insight into exploiting sp^2 hybridization in the development of α-L-fucosidase inhibitors. *ChemBioChem* **2019**, *20*, 1365−1368.
(18) Cobucci-Ponzano, B.; Trincone, A.; Giordano, A.; Rossi, M.; Moracci, M. Identification of the catalytic nucleophile of the family 29 α-L-fucosidase from *Sulfolobus solfataricus* via chemical rescue of an inactive mutant. *Biochemistry* **2003**, *42*, 9525−9531.
(19) van Bueren, A. L.; Ardèvol, A.; Fayers-Kerr, J.; Luo, B.; Zhang, Y.; Sollogoub, M.; Blériot, Y.; Rovira, C.; Davies, G. J. Analysis of the reaction coordinate of α-L-fucosidases: a combined structural and quantum mechanical approach. *J. Am. Chem. Soc.* **2010**, *132*, 1804−1806.
(20) Robb, C. S.; Hobbs, J. K.; Pluvinage, B.; Reintjes, G.; Klassen, L.; Monteith, S.; Giljan, G.; Amundsen, C.; Vickers, C.; Hettle, A. G.; Hills, R.; Nitin; Xing, X.; Montina, T.; Zandberg, W. F.; Abbott, D. W.; Boraston, A. B. Metabolism of a hybrid algal galactan by members of the human gut microbiome. *Nat. Chem. Biol.* **2022**, *18*, 501−510.
(21) Osanjo, G.; Dion, M.; Drone, J.; Solleux, C.; Tran, V.; Rabiller, C.; Tellier, C. Directed evolution of the α-L-fucosidase from *Thermotoga maritima* into an α-L-transfucosidase. *Biochemistry* **2007**, *46*, 1022−1033.
(22) Cobucci-Ponzano, B.; Conte, F.; Bedini, E.; Corsaro, M. M.; Parrilli, M.; Sulzenbacher, G.; Lipski, A.; Piaz, F. D.; Lepore, L.; Rossi, M.; Moracci, M. β-Glycosyl azides as substrates for α-glycosynthases: preparation of efficient α-L-fucosynthases. *Chem. Biol.* **2009**, *16*, 1097−1108.
(23) Slavin, J. Fiber and prebiotics: mechanisms and health benefits. *Nutrients* **2013**, *5*, 1417−1435.
(24) Lordan, C.; Thapa, D.; Ross, R. P.; Cotter, P. D. Potential for enriching next-generation health-promoting gut bacteria through prebiotics and other dietary components. *Gut Microbes* **2020**, *11*, 1−20.
(25) Teuber, M.; Bevill, R. D.; Osborn, M. J. Aldoheptoses in the *O*-antigenic lipopolysaccharide of *Salmonella typhimurium* and other gram-negative bacteria. I. Chemical synthesis of L-glycero-D-mannoheptose and β-L-glycero-D-mannoheptopyranosyl 1-phosphate. *Biochemistry* **1968**, *7*, 3303−3308.
(26) Okuyama, M.; Yoshida, T.; Hondoh, H.; Mori, H.; Yao, M.; Kimura, A. Catalytic role of the calcium ion in GH97 inverting glycoside hydrolase. *FEBS Lett.* **2014**, *588*, 3213−3217.
(27) Kuraku, S.; Zmasek, C. M.; Nishimura, O.; Katoh, K. aLeaves facilitates on-demand exploration of metazoan gene family trees on MAFFT sequence alignment server with enhanced interactivity. *Nucleic Acids Res.* **2013**, *41*, W22−W28.
(28) Armenteros, J. J. A.; Tsirigos, K. D.; Sønderby, C. K.; Petersen, T. N.; Winther, O.; Brunak, S.; Von Heijne, G.; Nielsen, H. SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat. Biotechnol.* **2019**, *37*, 420−423.
(29) Letunic, I.; Bork, P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* **2021**, *49*, W293−W296.
(30) Nozaki, S.; Niki, H. Exonuclease III (XthA) enforces *in vivo* DNA cloning of *Escherichia coli* to create cohesive ends. *J Bacteriol.* **2019**, *201*, No. e00660-18.
(31) Xia, Y.; Li, K.; Li, J.; Wang, T.; Gu, L.; Xun, L. T5 exonuclease-dependent assembly offers a low-cost method for efficient cloning and site-directed mutagenesis. *Nucleic Acids Res.* **2019**, *47*, No. e15.
(32) Studier, F. W. Protein production by auto-induction in high density shaking cultures. *Protein Expression Purif.* **2005**, *41*, 207−234.
(33) Kabsch, W. XDS. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 125−132.
(34) Adams, P. D.; Afonine, P. V.; Bunkóczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L.-W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 213−221.
(35) McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. Phaser crystallographic software. *J. Appl. Crystallogr.* **2007**, *40*, 658−674.
(36) Emsley, P.; Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *60*, 2126−2132.
(37) Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. Features and development of Coot. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 486−501.
(38) Murshudov, G. N.; Skubák, P.; Lebedev, A. A.; Pannu, N. S.; Steiner, R. A.; Nicholls, R. A.; Winn, M. D.; Long, F.; Vagin, A. A. *REFMAC5* for the refinement of macromolecular crystal structures. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2011**, *67*, 355−367.
(39) Afonine, P. V.; Grosse-Kunstleve, R. W.; Echols, N.; Headd, J. J.; Moriarty, N. W.; Mustyakimov, M.; Terwilliger, T. C.; Urzhumtsev, A.; Zwart, P. H.; Adams, P. D. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2012**, *68*, 352−367.
(40) Laskowski, R. A.; Swindells, M. B. LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *J. Chem. Inf. Model.* **2011**, *51*, 2778−2786.

https://doi.org/10.1021/acsomega.2c06991
ACS Omega 2022, 7, 47411−47423