Identification of the Catalytic Residue of *Thermococcus litoralis* 4-α-Glucanotransferase through Mechanism-Based Labeling†

Hiromi Imamura,‡ Shinya Fushinobu,‡ Beong-Sam Jeon,‡ Takayoshi Wakagi,*,‡ and Hiroshi Matsuzawa§

Department of Biotechnology, The University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-8657, Japan, and Department of Bioscience and Biotechnology, Aomori University, Kohbata, Aomori 030-0943, Japan

Received May 18, 2001; Revised Manuscript Received August 6, 2001

ABSTRACT: *Thermococcus litoralis* 4-α-glucanotransferase (TLGT) belongs to family 57 of glycoside hydrolases and catalyzes the disproportionation and cycloamylose synthesis reactions. Family 57 glycoside hydrolases have not been well investigated, and even the catalytic mechanism involving the active site residues has not been studied. Using 3-ketobutylidene-*β*-2-chloro-4-nitrophenyl maltopentaoside (3KBG5CNP) as a donor and glucose as an acceptor, we showed that the disproportionation reaction of TLGT involves a ping-pong bi-bi mechanism. On the basis of this reaction mechanism, the glycosyl-enzyme intermediate, in which a donor substrate was covalently bound to the catalytic nucleophile, was trapped by treating the enzyme with 3KBG5CNP in the absence of an acceptor and was detected by matrix-assisted laser desorption ionization time-of-flight mass spectrometry after peptic digestion. Postsource decay analysis suggested that either Glu-123 or Glu-129 was the catalytic nucleophile of TLGT. Glu-123 was completely conserved between family 57 enzymes, and the catalytic activity of the E123Q mutant enzyme was greatly decreased. On the other hand, Glu-129 was a variable residue, and the catalytic activity of the E129Q mutant enzyme was not decreased. These results indicate that Glu-123 is the catalytic nucleophile of TLGT. Sequence alignment of TLGT and family 38 enzymes (class II α-mannosidases) revealed that Glu-123 of TLGT corresponds to the nucleophilic aspartic acid residue of family 38 glycoside hydrolases, suggesting that family 57 and 38 glycoside hydrolases may have had a common ancestor.

Enzymatic hydrolysis of glycosidic linkages can be classified into two major mechanisms according to the anomeric configuration of the product, retaining and inverting, and in both cases the critical residues are two acidic amino acids (*1*, *2*). The retaining glycoside hydrolases are proposed to operate via a double displacement mechanism in which one carboxyl group acts as a nucleophile, attacking the anomeric center of the substrate sugar to form a glycosyl-enzyme intermediate (*3*). The other acts as an acid/base catalyst, protonating the glycosidic oxygen in the first step and deprotonating the acceptor molecule in the second step.

Thermococcus litoralis 4-α-glucanotransferase (TLGT)[1] (*4*) is a hyperthermostable transglycosylation enzyme that acts on the α-1,4-glucosidic linkage of a donor substrate and transfers the glucosyl or maltooligosyl moiety to the acceptor, a new α-1,4-glucosidic linkage being formed. The enzyme has a calculated molecular mass of 78 kDa and is a monomer, as determined on SDS−PAGE and gel filtration. It is proposed that TLGT is involved in maltose metabolism in vivo, producing glucose and a series of maltodextrins through intermolecular transglycosylation (disproportionation reaction) of maltose entering the process via a high-affinity binding protein-dependent ABC transporter (*5*). Glucose and maltodextrins are converted to glucose 6-phosphate by a novel ADP-dependent glucokinase, and phosphorylase and phosphoglucomutase, respectively, and then the resultant glucose 6-phosphate is subsequently catabolized via an Embden−Myerhof-type glycolytic pathway (*6*). In addition to intermolecular transglycosylation, TLGT also catalyzes intramolecular transglycosylation in vitro to produce cycloamyloses (CAs) with sixteen to several hundred glucose units[2] from linear amylose (*4*). CAs are cyclic α-1,4-glucans containing various numbers, from six to several hundreds, of glucose units. The smallest CAs containing 6, 7, and 8 glucose units are also known as α-, *β*-, and *γ*-cyclodextrins, respectively. Large CAs were recently found by Takaha et al. as products of the potato D-enzyme with linear amylose as the substrate (*7*). Small CAs with 6−8 glucose units (i.e., cyclodextrins) are well-known doughnut-shaped molecules and are able to accommodate guest molecules in their central cavities, which leads to the formation of inclusion complexes. Although small CAs are conformationally restricted, larger CAs have a more flexible single-helical conformation in an aqueous solution (*8*). It is likely that during complex

† This research was supported in part by Grants-in-Aid for Scientific Research (B) (10460035 and 12460047) to H.M. from the Japan Society for the Promotion of Science.

* To whom correspondence should be addressed. E-mail: atwakag@mail.ecc.u-tokyo.ac.jp. Telephone: 81-3-5841-5152. Fax: 81-3-5841-5337.

‡ The University of Tokyo.

§ Aomori University.

[1] Abbreviations: TLGT, *Thermococcus litoralis* 4-α-glucanotransferase; CA(s), cycloamylose(s); CGTase, cyclodextrin glucanotransferase; 3KBG5CNP, 3-ketobutylidene-*β*-2-chloro-4-nitrophenyl maltopentaoside; CNP, 2-chloro-4-nitrophenol; PSD, postsource decay; MALDI-TOF-MS, matrix-assisted laser desorption ionization time-of-flight mass spectrometry.

[2] B.-S. Jeon, T. Wakagi, and H. Matsuzawa, unpublished data.

10.1021/bi011017c CCC: $20.00
Published on Web 09/19/2001

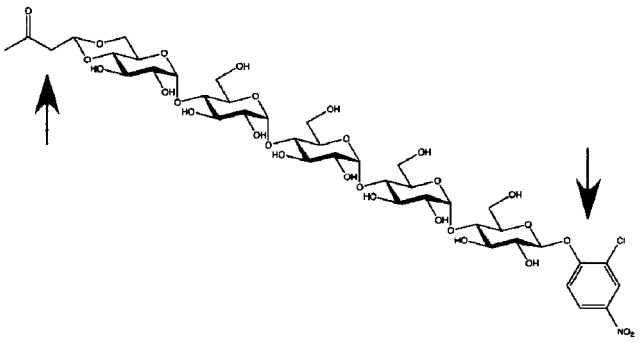


FIGURE 1: Structure of 3KBG5CNP. Positions different from those in maltopentaose are indicated by arrows.

formation large CAs lose their flexibility and fold into a compact structure (*9*) like the crystal structure of cycloamylose with 26 glucose units (*10*). Because of this structural difference from cyclodextrins, larger CAs show the advantageous features of the formation of inclusion complexes and higher solubility in water and thus are expected to be valuable for future industrial use (*11*).

Glycoside hydrolases have been classified into a group of glycoside hydrolase families on the basis of their amino acid sequences (*12*−*14*), there being more than 80 families at this time (*15*). The anomer-retaining glycoside hydrolases that act on the α-1,4- or α-1,6-glucosidic linkage of starch or amylose are categorized into three families, families 13, 57, and 77. Family 13 (also known as the α-amylase family) is widespread in a variety of organisms. The three-dimensional structures of many family 13 enzymes have been determined (*16*−*20*), and amino acid residues involved in catalysis including a nucleophile and an acid/base catalyst have also been well studied (for a review, see ref *21*). Family 77 is closely related to family 13 in both amino acid sequence and three-dimensional structure (*22*), and thus they are grouped into a clan (*15*). On the other hand, family 57, to which TLGT belongs, has been found only in anaerobic thermophiles, all of which are archaea with one exception, *Dictyoglomus thermophilum* (*23*). Family 57 glycoside hydrolases have not been well investigated, and their three-dimensional structures have not been determined yet. Moreover, the catalytic mechanism involving the active site residues has not been studied.

In this study, we identified the catalytic nucleophile of TLGT by means of mechanism-based labeling. Labeling of a retaining glycoside hydrolase nucleophile is usually performed with artificial sugars that have active groups such as fluoride (*24*−*26*) or epoxide (*27*−*29*) to accelerate the first step of the double replacement reaction and have different sugar conformations from the native substrate to decrease the reaction rate of the second step. Here, we employed a new mechanism-based labeling method involving 3-ketobutylidene-β-2-chloro-4-nitrophenyl maltopentaoside (3KBG5CNP), which is modified at the C4 and C6 hydroxyl groups at its nonreducing end and has 2-chloro-4-nitrophenol (CNP) as an aglycon (Figure 1). 3KBG5CNP was originally developed as a substrate for α-amylase (*30*). However, it can also be used as a donor substrate for cyclodextrin glucanotransferase (CGTase) (*31*). After proteolysis of the labeled enzyme, identification of the catalytic nucleophile was achieved by matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF-MS).

MATERIALS AND METHODS

Materials. All reagents were purchased from Wako Pure Chemical Industries, Ltd. (Osaka, Japan), unless otherwise noted. 3KBG5CNP was a gift from Toyobo (Osaka, Japan). Pepsin from porcine mucosa and 2,5-dihydroxybenzoic acid were from Sigma. Butyl-Toyopearl 650M and ceramic hydroxyapatite were from Tosoh (Tokyo, Japan) and Bio-Rad, respectively. Oligonucleotides were from Hokkaido System Science (Sapporo, Japan).

Gene Expression and Purification of TLGT. The wild-type and mutant TLGT genes were coexpressed in *Escherichia coli* cells with GroELS, $tRNA_{AGA}$, and $tRNA_{AGG}$ as described previously (*32*). The cells were grown for 3 h at 37 °C, and the T7 promoter was subsequently induced by the addition of 0.1 mM IPTG. The cells were allowed to grow for an additional 6 h at 30 °C, collected by centrifugation, and then suspended in 20 mM sodium phosphate buffer, pH 6.0 (buffer A). The suspension was sonicated and centrifuged (27000*g* for 60 min). The supernatant was incubated for 30 min at 80 °C and then centrifuged (27000*g* for 30 min) to remove denatured *E. coli* proteins. The resultant supernatant was mixed with an equal volume of buffer A containing 1.6 M ammonium sulfate and then applied to a butyl-Toyopearl 650 M column equilibrated with buffer A containing 0.8 M ammonium sulfate. The enzyme was eluted with buffer A containing 0.4 M ammonium sulfate. The active fraction was dialyzed against buffer A and then applied to a ceramic hydroxyapatite column equilibrated with buffer A. The enzyme was eluted with a linear gradient of sodium phosphate, from 20 to 400 mM. The active fraction was concentrated with a Centriplus-30 (Millipore) and then dialyzed against buffer A.

Enzyme Assay. The disproportionation reaction of TLGT was measured basically according to the method of Nakamura et al. (*31*). The enzyme was incubated with 3KBG5CNP and glucose in buffer A at 60 °C. After appropriate time intervals, 475 μL aliquots of the reaction mixture were mixed with 75 μL of 0.2 M HCl to stop the reaction and then left at room temperature for 10 min. Then, 100 μL of 0.5 M disodium hydrogen phosphate and 100 μL of buffer A containing 0.75 unit of α-glucosidase (from *Saccharomyces* sp.) and 0.5 unit of β-glucosidase (from almond) were added, followed by incubation for 1 h at 37 °C. Subsequently, 50 μL of 1 M sodium carbonate was added to increase the pH of the reaction mixture to above 8.5. The degradation of 3KBG5CNP was calculated from the increase in absorbance at 400 nm caused by CNP ($\epsilon = 16.8\ mM^{-1}$). In the disproportionation reaction, TLGT degraded 3KBG5CNP into β-glycosyl- or β-maltooligosyl-CNP, which was followed by α-glucosidase and β-glucosidase treatment to release CNP. Activity toward maltotriose was measured as described previously (*32*). One unit of activity was defined as the amount of enzyme that liberated 1 μmol of glucose from maltotriose per minute at 80 °C.

Labeling and Proteolysis of TLGT. Labeling of TLGT was performed by incubating it (50 mM) with 3KBG5CNP (5 mM) for 1 min at 60 °C in 20 μL of buffer A. The enzyme was immediately denatured by adding 30 μL of 50 mM sodium phosphate buffer, pH 2.0, followed by cooling to room temperature. Proteolysis was achieved by adding 20 μL of pepsin (10 μM in 50 mM sodium phosphate buffer,

pH 2.0), followed by incubation at 37 °C for 3 h.

MALDI-TOF-MS. Analysis of peptide samples was carried out using a Voyager DE-STR mass spectrometer (Applied Biosystems). The instrument was operated in the delayed extraction mode with an accelerating voltage of 20 kV. Ionization was achieved with a nitrogen laser (337 nm, 3 ns pulse width). A 10 μL aliquot of the digested sample was desalted with a Zip Tip C_{18} (Millipore) and then eluted with 10 μL of 50% acetonitrile. One microliter of the peptide solution was mixed with 4 μL of 2,5-dihydroxybenzoic acid (10 mg/mL in 10% ethanol). One microliter of the mixture was spotted onto sample plate and allowed to dry under ambient conditions before insertion into the mass spectrometer. Then the sample was first examined in the linear mode. The instrument was calibrated using angiotensin I (Sigma) as an external standard, and the mass range of 500−4300 Da was scanned. In the postsource decay (PSD) mode, different regions of 8−10 spectral segments were acquired by decreasing the mirror ratio stepwise, and these segments were stitched together to create a complete PSD spectrum. GRAMS/386 software was used for data processing.

Site-Directed Mutagenesis. The E123Q mutant was prepared with a USE mutagenesis kit (Amersham Pharmacia Biotech). Oligonucleotide 5′-GGAACTATATCCGCATATCCCGCAAGAGG-3′ was used as a selection primer, which eliminated the unique *Eco*RV site in the plasmid. Oligonucleotide 5′-GTGGCTCACACAACGCGTATGGCAGC-3′, which replaced the codon for Glu-123 (GAG) with CAA and introduced a *Mlu*I site for rapid screening of the mutation, was used as a mutagenic primer. The E129Q mutant was prepared with a QuikChange site-directed mutagenesis kit (Stratagene). Oligonucleotide 5′-GGTATGGCAGCCGCAGCTGGTAAAGTCG-3′ and a complementary oligonucleotide, which replaced the codon for Glu-129 (GAA) with CAG and introduced a *Pvu*II site, were used. Mutations were reconfirmed by sequencing with an ABI PRISM 310 DNA sequencer (Applied Biosystems).

RESULTS

In retaining glycoside hydrolases, a catalytic nucleophile binds covalently to a substrate sugar to form a glycosyl-enzyme intermediate in the course of the reaction (*3*). We expected that the catalytic nucleophile could be identified if the glycosyl-enzyme intermediate was trapped. Nakamura et al. (*31*) have proved that the disproportionation reaction of CGTase, which belongs to family 13 but catalyzes a similar reaction to TLGT, involves a ping-pong bi-bi mechanism with 3KBG5CNP as a donor and glucose as an acceptor. It seemed very likely that TLGT also used the same reaction mechanism for the disproportionation reaction, although these enzymes are members of different families. If the reaction of TLGT actually involves a ping-pong bi-bi mechanism and the enzyme does not have hydrolytic activity, a donor substrate itself should react with the enzyme, and the glycosyl-enzyme intermediate should accumulate in the absence of an acceptor substrate (Figure 2).

Reaction of TLGT with 3KBG5CNP. Using 3KBG5CNP and glucose, we investigated whether the reaction mechanism of TLGT is the same as that of CGTase. Because the reducing end of 3KBG5CNP is blocked, it only acts as a donor substrate for the transglycosylation reaction. In

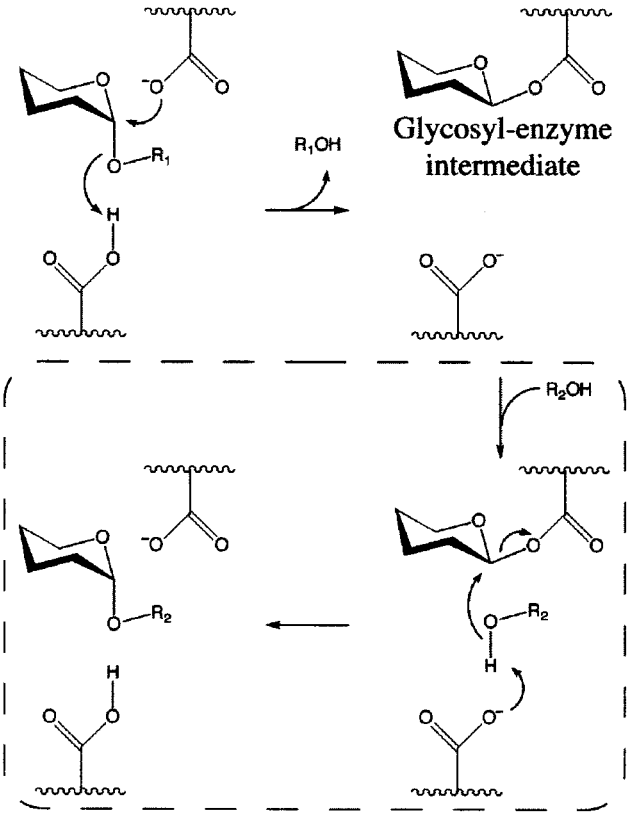


FIGURE 2: Trapping strategy for the glycosyl-enzyme intermediate. In the absence of an acceptor substrate (represented as R_2OH), the second step of the reaction (in the dotted square) does not proceed, and the covalent glycosyl-enzyme intermediate is expected to accumulate.

contrast, glucose only acts as an acceptor, because it does not have a glucosidic bond. TLGT was purified to homogeneity by heat treatment followed by two steps of column chromatography. At first, TLGT was incubated with 3KBG5CNP at 60 °C in the presence or absence of glucose as an acceptor. Although the optimum temperature of TLGT is 90 °C, the reaction temperature with 3KBG5CNP and glucose was set at 60 °C, because 3KBG5CNP was degraded above 70 °C (data not shown). In the presence of glucose, 3KBG5CNP was degraded, and the production of CNP increased in a linear manner with the reaction time (Figure 3A). The initial velocity of the reaction also increased as the concentration of glucose was increased (Figure 3A). This means that glucose is actually able to act as an acceptor. In contrast, 3KBG5CNP was hardly degraded in the absence of glucose, and only a little CNP was produced (Figure 3A). This indicates that the hydrolytic activity of TLGT is low enough to ignore under these conditions. The initial reaction rates with various substrate concentrations for the disproportionation reaction were determined to perform steady-state kinetic analysis. Lineweaver−Burk plots showed parallel lines with various glucose concentrations (Figure 3B). Likewise, Lineweaver−Burk plots showed parallel lines with various 3KBG5CNP concentrations (Figure 3C). These results indicate that the disproportionation reaction of TLGT involves a ping-pong bi-bi mechanism. Because TLGT hydrolyzes 3KBG5CNP very slowly, the glycosyl-enzyme intermediate was expected to accumulate when TLGT reacted with 3KBG5CNP in the absence of an acceptor, such as glucose, as shown in Figure 2.

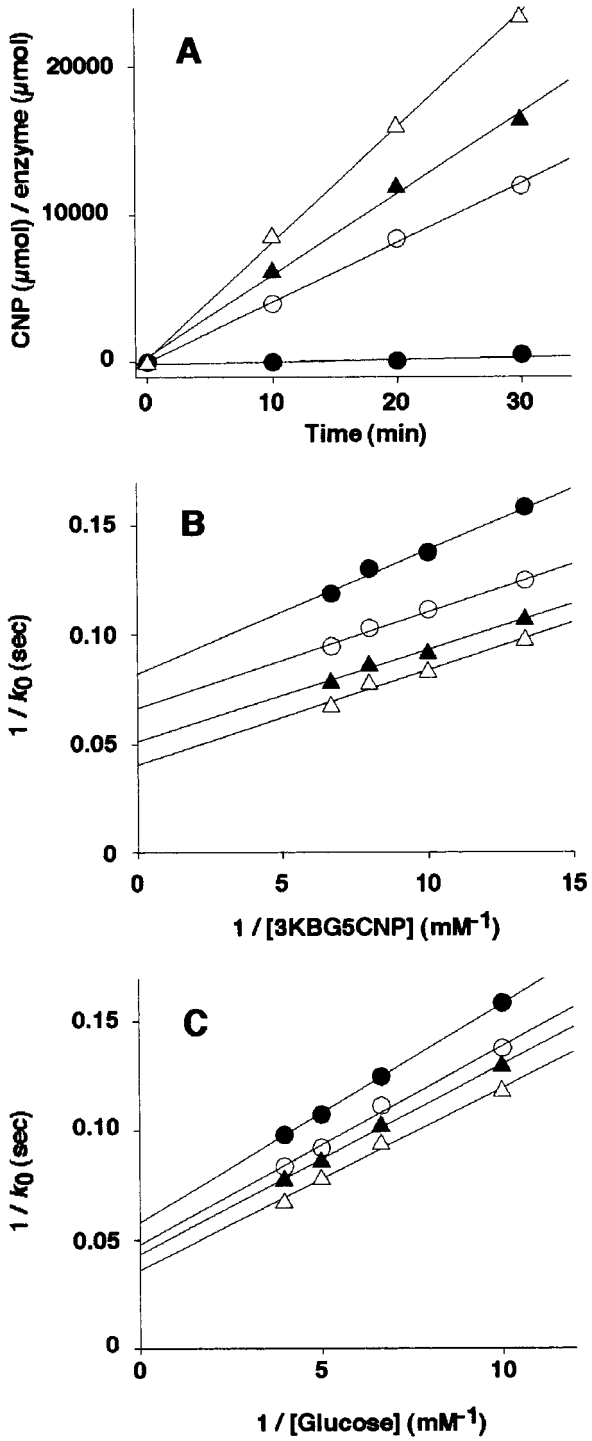


FIGURE 3: Disproportionation reaction of 4-α-glucanotransferase involves a ping-pong bi-bi mechanism. (A) Plot of released CNP (μmol) per micromole of enzyme versus reaction time with the following glucose concentrations: 0 (●), 0.25 (○), 0.5 (▲), and 1.0 mM (△). 3KBG5CNP was fixed at 1.0 mM. (B) Lineweaver−Burk plots with various concentrations of 3KBG5CNP with fixed concentrations of glucose. Values are the averages of three independent experiments. The glucose concentrations were 0.10 (●), 0.15 (○), 0.20 (▲), and 0.25 mM (△). (C) Lineweaver−Burk plots with various concentrations of glucose with fixed concentrations of 3KBG5CNP. Values are the averages of three independent experiments. The 3KBG5CNP concentrations were 0.075 (●), 0.10 (○), 0.125 (▲), and 0.15 mM (△).

Detecting the Labeled Peptide. The enzyme was incubated with and without 3KBG5CNP, followed by digestion with pepsin. After peptic digestion, no obvious bands were detected on SDS−PAGE (data not shown). The digested enzyme was desalted and then analyzed directly by MALDI-TOF-MS in the linear mode. Peptic digestion of the labeled enzyme gave a complex mixture of peptides, and many peaks were observed in the mass spectrum (Figure 4). The peaks were reproducibly observed on several measurements (data not shown), although pepsin exhibits a broad specificity. In the case of 3KBG5CNP labeling, a *m*/*z* 1824.7 peak appeared (Figure 4B), while this signal did not appear in the absence of 3KBG5CNP (Figure 4A). No other distinguishable signals could be detected in either spectrum. We thought that this signal was derived from a labeled peptide and then attempted to determine the amino acid sequence of the labeled peptide by PSD.

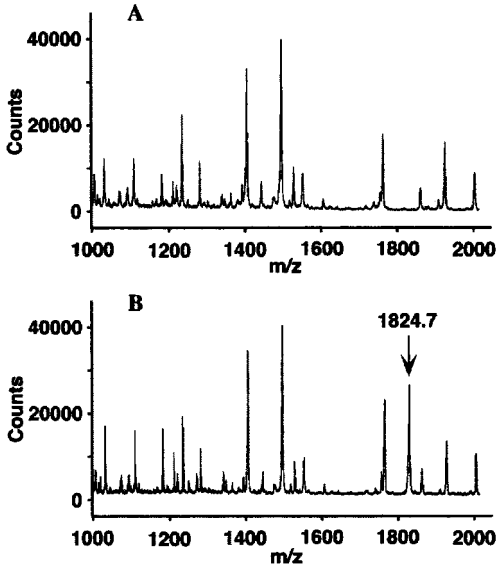
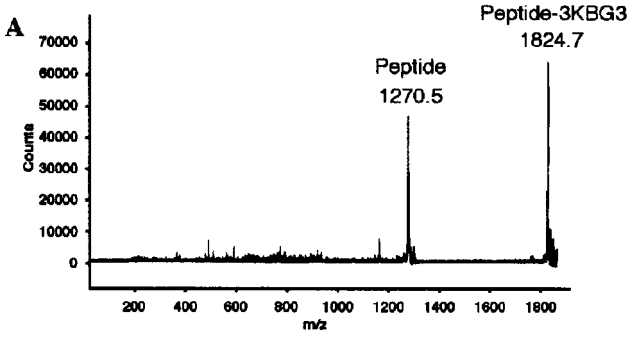
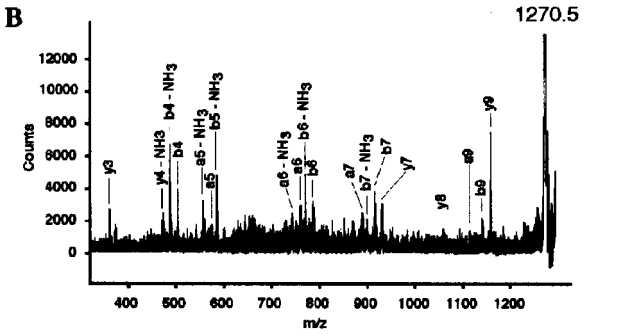


FIGURE 4: Mass spectra of peptic digests with 4-α-glucanotransferase. Samples with no labeling (A) and 3KBG5CNP labeling (B).

Identification of the Nucleophile Residue. To identify the labeled position, the *m*/*z* 1824.7 signal was analyzed by PSD. When the timed ion selector was fixed at 1824.7, only the *m*/*z* 1824.7 signal was isolated, since other signals, except those due to fragmentation ions, disappeared in the reflector mode (data not shown), and then the isolated signal was analyzed in the PSD mode. The highest mass peak next to the parent ion was *m*/*z* 1270.5, and there were no obvious peaks between them (Figure 5A). The decrease of 554.2 Da seemed to correspond to a 3-ketobutylidenemaltotriosyl group (molecular mass 555.5 Da) generated through cleavage of the linkage between the carboxyl oxygen of the nucleophile and the C1 atom of the substrate. The facile cleavage yielding a *m*/*z* 1270.5 peak is probably due to a hydrogen transfer from the 3-ketobutylidenemaltotriosyl group to the carboxylate of the peptide, which may cause the one mass unit difference between the molecular mass of the 3-ketobutylidenemaltotriosyl group (555.5) and the mass decrease in the spectrum (554.2). Therefore, the *m*/*z* 1270.5 peak was the $[M + H]^+$ ion of a peptide with no label, and the *m*/*z* 1824.7 peak was the $[M + H]^+$ ion of a peptide labeled by the 3-ketobutylidenemaltotriosyl group. In the absence of 3KBG5CNP, the *m*/*z* 1270.5 peak did not appear in the linear mode spectrum (Figure 5A), probably because the nonlabeled peptide was further degraded by pepsin and the labeled peptide was not degraded due to the masking effect of the





C

a 572 758 886 1112

b 501 599 786 914 1140

Leu[121]-Thr-Glu-Arg-Val-Trp-Gln-Pro-Glu-Leu

y 1158 1057 928 486 358

D

		Labeled fragment	
TLGT	117	KGVWLTERVWQPELVK	132
PFAM	117	RGVWLTERVWQPELVK	132
DTAM	118	KGMWLAERVWEPHLVK	133
PFAP	312	RGGWAAESALNDKTLE	327
THPU	313	KGGWAAESALNDKTLE	328

FIGURE 5: MALDI−PSD analysis of the labeled peptide. (A) Whole PSD spectrum. (B) Observed fragmentation ions in the PSD spectrum. (C) Amino acid sequence of the labeled peptide deduced on PSD analysis and a sequence search. Observed fragmentation ions are also presented. (D) Amino acid sequences of family 57 enzymes around the labeled region. Identical residues are in bold. Abbreviations: TLGT, *T. litoralis* 4-α-glucanotransferase; PFAM, *Pyrococcus furiosus* α-amylase; DTAM, *D. thermophilum* amylase A; PFAP, *P. furiosus* amylopullulanase; THPU, *Thermococcus hydrothermalis* pullulanase.

cross-linked substrate. Many ions derived from fragmented peptides were observed below the ion at 1270.5 (Figure 5A,B), but immonium ions were not observed (data not shown). Although four a-ions, five b-ions, and five y-ions were identified (Figure 5B), complete sequencing of this peptide was impossible on the basis of only the PSD spectrum due to the lack of some fragmentation ions. The combination of fragment analysis of the PSD spectrum and a search for the amino acid sequence of TLGT revealed that the labeled peptide corresponds to the peptide ^{121}LTERVWQPEL^{130} (Figure 5C). The average mass of the discharged peptide is 1270.45 and will give the $[M + H]^+$ about 1271.4. This value is consistent with the signal of $[M + H]^+$ = 1270.5 in the PSD spectrum, although the reason for the one mass difference is unclear. In the sequence, there are two candidates for the catalytic nucleophile, Glu-123 and Glu-129, because in retaining glycoside hydrolases the nucleophile residue is generally Glu or Asp.

Table 1: Specific Activity of the Wild Type and Mutants

	specific activity (units/mg)
wild type	17.7
E123Q	0.0012
E129Q	20.3

Activity of TLGT Mutants. To confirm which glutamic acid is really the catalytic nucleophile, we constructed the E123Q and E129Q mutants of TLGT, whose Glu-123 and Glu-129 were replaced with Gln, respectively. These mutants were expressed in *E. coli* and purified to homogeneity by heat treatment followed by two steps of column chromatography. The activity of the E123Q mutant toward maltotriose decreased about 15 000-fold (Table 1). On the other hand, the activity of the E129Q mutant was almost the same as that of the wild-type enzyme (Table 1). The E123Q mutant seemed to have the same secondary structure as the wild-type enzyme, because there were no obvious differences in circular dichroism (CD) spectra between them (data not shown).

DISCUSSION

Kinetic analysis of transglycosylation enzymes such as CGTase or TLGT has been very difficult, because it is hard to distinguish the donor substrate, acceptor substrate, and products when native substrates such as maltodextrin or amylose are used. It has been reported that 3KBG5CNP is a useful substrate for analyzing the reaction mechanism of CGTase, which belongs to family 13 of glycoside hydrolases (*31*). We thought that this method was also applicable to analysis of the reaction mechanism of TLGT, because they catalyze similar reactions. This method could be used for TLGT by increasing the reaction temperature from 37 to 60 °C, because TLGT is more thermophilic than CGTase. At this temperature, TLGT exhibits activity, but only 40% of the maximum activity (*4*), and 3KBG5CNP does not degrade. Judging from the results of kinetic analysis shown in Figure 3, the transglycosylation reaction of TLGT is based on a ping-pong bi-bi mechanism. Although TLGT and CGTase belong to different families, we have proved that they have the same reaction mechanism.

For retaining glycoside hydrolases, it is thought that in the course of the reaction the enzyme and substrate form a glycosyl-enzyme intermediate (*3*), in which the substrate sugar is bound covalently to the catalytic nucleophile, and then the substrate is transferred to an acceptor molecule, which is water in the hydrolysis reaction and a sugar in transglycosylation (*1*, *2*). Because TLGT involves a ping-pong bi-bi mechanism, TLGT and a donor substrate form a glycosyl-enzyme intermediate even without an acceptor substrate (Figure 2). It is difficult to trap the glycosyl-enzyme intermediate if TLGT has hydrolytic activity, because the glycosyl-enzyme intermediate would be quickly degraded by the attacking of water. However, it seemed possible to trap the intermediate, because TLGT has only weak hydro-

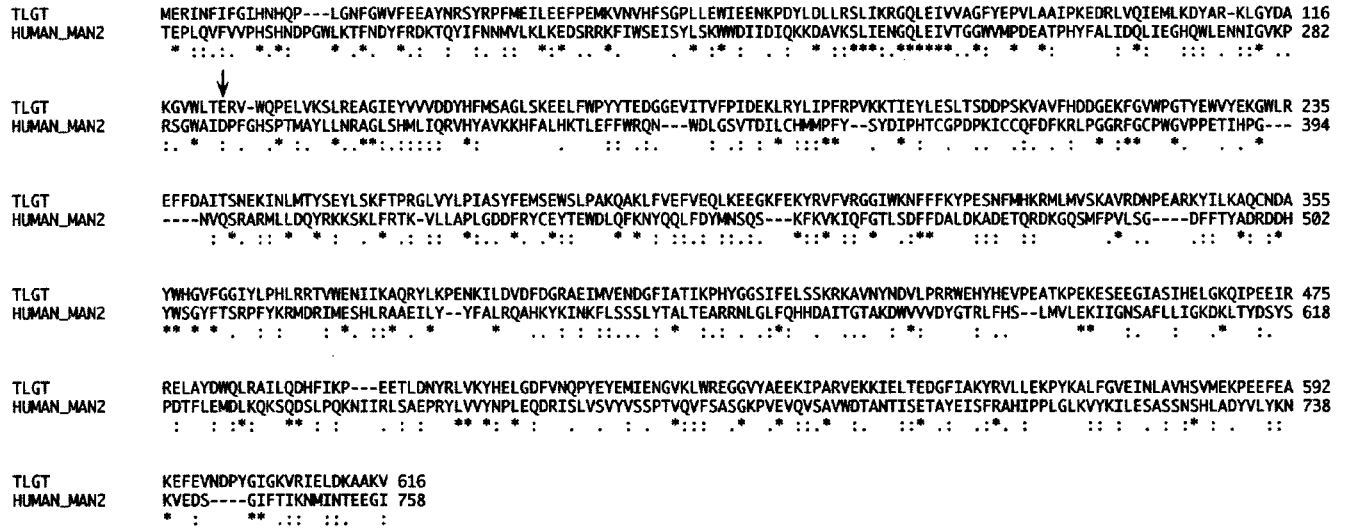


FIGURE 6: Nucleophile residues of families 57 and 38 are at the same position. Amino acid sequence alignment of *T. litoralis* 4-α-glucanotransferase (TLGT) and human Golgi α-mannosidase II (HUMAN_MAN2) is shown. Alignment was performed with Clastal W (*42*). The position of the nucleophile residue is indicated by an arrow. Asterisks indicate identical residues. Colons indicate conserved substitutions. Dots indicate semiconserved substitutions.

lytic activity (Figure 3A). By incubating TLGT with 3KBG5CNP but without an acceptor substrate, we could trap the glycosyl-enzyme intermediate, which was successfully detected by MALDI-TOF-MS after peptic digestion. We attempted to identify the glycosylation position by sequencing this signal by the PSD technique. Peptide sequencing by PSD has several advantages compared with classical Edman degradation (*33*). The most important advantage is the ability to sequence a peptide within a mixture because it can be mass selected by the instrument (timed ion selector), so further purification by, for example, HPLC is not necessary. We could determine the sequence of the labeled peptide by PSD together with a search for the amino acid sequence of TLGT. The intermediate detected was proved to be a peptide of 10 amino acids (^{121}LTERVWQPEL^{130}) cross-linked with a 3-ketobutylidenemaltotriosyl group. This intermediate forms when TLGT attacks the second α-1,4-glucosidic bond from the reducing end. In 3KBG5CNP there are three other α-1,4-glucosidic bonds, which TLGT can potentially attack (Figure 1). However, no other intermediates with other sugar units, such as a 3-ketobutylideneglucosyl group, could be trapped (Figure 4). One possibility is the attacking site bias of TLGT. However, a product bias was not observed when maltopentaose was used as substrate (data not shown). Another explanation is the structural crash of the 3-ketobutylidene and nitrochlorophenol moieties of 3KBG5CNP, whose natures are quite different from native substrates, with the active site residues of TLGT. This may force 3KBG5CNP into the position at which catalytic residues can attack the second α-1,4-glucosidic linkage from the reducing end. Unfortunately, the ester bond between the sugar and the peptide was cleaved more easily than the bonds in the peptide chain on PSD analysis, and therefore the residue bound to the substrate could not be identified directly. However, there were only two potential nucleophile residues, Glu-123 and Glu-129, because the catalytic nucleophile of a glycoside hydrolase is a Glu or an Asp in general. Additional mutation experiments suggest that Glu-123 is involved in the catalysis, whereas Glu-129 is not required for the enzymatic activity. Our results clearly indicate that Glu-123 acts as the catalytic nucleophile of TLGT. In general, two catalytic residues are completely conserved within each family of glycoside hydrolases. Within the labeled peptide fragment, Glu-123 is the only residue conserved among family 57 enzymes (Figure 5D). Recently, we determined the crystal structure of TLGT, and in this structure Glu-123 is located in the cleft of the molecule and is very close to the C1 atom of a substrate analogue.[3] This observation also supports our conclusion.

Sequence Comparison of Family 57 and Family 38 Glycoside Hydrolases. On using PSI-BLAST (*34*) to search the GenBank and SwissProt databases, very low homology was found between TLGT and family 38 glycoside hydrolases; they showed only 22−31% similarity and 9−18% identity. Family 38 enzymes are retaining α-mannosidases and are also categorized as class II α-mannosidases. Family 38 α-mannosidases, which are mainly found in eukarya, play important roles in the maturation and degradation of glycoproteins in mammals (*35*, *36*) and are medically highly important. In particular, Golgi α-mannosidase II, a key enzyme in the N-glycosylation pathway, is a target in the development of anticancer drugs (*37*). Also, a deficiency in mammalian lysosomal α-mannosidase, which is involved in glycoprotein degradation, causes α-mannosidosis, an autosomal recessive disorder (*38*, *39*). Interestingly, sequence alignment of human Golgi α-mannosidase II and TLGT revealed that an aspartic acid residue of family 38 glycoside hydrolases corresponds to the nucleophile glutamic acid residue of family 57 (indicated by an arrow in Figure 6). This aspartic acid residue was recently proved to be the catalytic nucleophile of family 38 glycoside hydrolases by labeling using 5-fluoro-*β*-gulosyl fluoride (*24*, *40*). This finding suggests that family 57 and 38 glycoside hydrolases may have had a common ancestor, although they show only low similarity. So far, the evolutionary relationship between the two families has not been referred to, nor have they been grouped into a clan. Very recently, the crystal structure of

[3] H. Imamura, S. Fushinobu, T. Kumasaka, M. Yamamoto, B.-S. Jeon, T. Wakagi, and H. Matsuzawa, unpublished results.

Golgi α-mannosidase II from *Drosophila melanogaster*, which belongs to family 38, was reported (*41*). The structures of the catalytic domains of TLGT and α-mannosidase II are different in the overall fold, but both folds are similar in that they have a β-sheet core surrounded by α-helices.[3] Golgi α-mannosidase II has an active site zinc that is involved in substrate specificity and catalysis, while there is no ion in the active site of TLGT.[3] When TLGT was incubated with 2-α-, 3-α-, or 4-α-mannobiose, α-mannosidase activity could not be detected (data not shown). From the structural differences it seems very difficult to convert TLGT into an α-mannosidase-type enzyme or α-mannosidase into a 4-α-glucanotransferase only through the substitution of some active site residues. Naturally, there is also the possibility that the two evolutionarily unrelated proteins converged to obtain the sugar hydrolytic function. However, the fact that family 57 and family 38 glycoside hydrolases are distributed mainly in archaea and eukarya, respectively, also suggests that the two families have evolved from a common ancestor, which had glycolytic activity, and have acquired their respective functions in the course of their evolution.

ACKNOWLEDGMENT

We thank Dr. K. Haga (University of Tsukuba) for useful advice regarding 3KBG5CNP.

REFERENCES

1. Davies, G., and Henrissat, B. (1995) *Structure (London) 3*, 853−859.
2. McCarter, J. D., and Withers, S. G. (1994) *Curr. Opin. Struct. Biol. 4*, 885−892.
3. Sinnott, M. L. (1990) *Chem. Rev. 90*, 1171−1202.
4. Jeon, B. S., Taguchi, H., Sakai, H., Ohshima, T., Wakagi, T., and Matsuzawa, H. (1997) *Eur. J. Biochem 248*, 171−178.
5. Xavier, K. B., Peist, R., Kossmann, M., Boos, W., and Santos, H. (1999) *J. Bacteriol. 181*, 3358−3367.
6. Selig, M., Xavier, K. B., Santos, H., and Schonheit, P. (1997) *Arch. Microbiol. 167*, 217−232.
7. Takaha, T., Yanase, M., Takata, H., Okada, S., and Smith, S. M. (1996) *J. Biol. Chem. 271*, 2902−2908.
8. Kitamura, S. A., Isuda, H., Shimada, J., Takada, T., Takaha, T., Okada, S., Mimura, M., and Kajiwara, K. (1997) *Carbohydr. Res. 304*, 303−314.
9. Kitamura, S., Nakatani, K., Takaha, T., and Okada, S. (1999) *Macromol. Rapid Commun. 20*, 612−615.
10. Gessler, K., Uson, I., Takaha, T., Krauss, N., Smith, S. M., Okada, S., Sheldrick, G. M., and Saenger, W. (1999) *Proc. Nat. Acad. Sci. U.S.A. 96*, 4246−4251.
11. Takaha, T., and Smith, S. M. (1999) *Biotechnol. Genet. Eng. Rev. 16*, 257−280.
12. Henrissat, B. (1991) *Biochem. J. 280*, 309−316.
13. Henrissat, B., and Bairoch, A. (1993) *Biochem. J. 293*, 781−788.
14. Henrissat, B., and Bairoch, A. (1996) *Biochem. J. 316*, 695−696.
15. Coutinho, P. M., and Henrissat, B. (1999) Carbohydrate-active enzymes server, http://afmb.cnrs-mrs.fr/~pedro/CAZY/db.html.
16. Boel, E., Brady, L., Brzozowski, A. M., Derewenda, Z., Dodson, G. G., Jensen, V. J., Petersen, S. B., Swift, H., Thim, L., and Woldike, H. F. (1990) *Biochemistry 29*, 6244−6249.
17. Buisson, G., Duee, E., Haser, R., and Payan, F. (1987) *EMBO J. 6*, 3909−3916.
18. Katsuya, Y., Mezaki, Y., Kubota, M., and Matsuura, Y. (1998) *J. Mol. Biol. 281*, 885−897.
19. Klein, C., and Schulz, G. E. (1991) *J. Mol. Biol. 217*, 737−750.
20. Matsuura, Y., Kusunoki, M., Harada, W., and Kakudo, M. (1984) *J. Biochem. 95*, 697−702.
21. Svensson, B. (1994) *Plant Mol. Biol. 25*, 141−157.
22. Przylas, I., Tomoo, K., Terada, Y., Takaha, T., Fujii, K., Saenger, W., and Strater, N. (2000) *J. Mol. Biol. 296*, 873−886.
23. Fukusumi, S., Kamizono, A., Horinouchi, S., and Beppu, T. (1988) *Eur. J. Biochem. 174*, 15−22.
24. Howard, S., He, S., and Withers, S. G. (1998) *J. Biol. Chem. 273*, 2067−2072.
25. Miao, S., Ziser, L., Aebersold, R., and Withers, S. G. (1994) *Biochemistry 33*, 7027−7032.
26. Mosi, R., He, S. M., Uitdehaag, J., Dijkstra, B. W., and Withers, S. G. (1997) *Biochemistry 36*, 9927−9934.
27. Iwanami, S., Matsui, H., Kimura, A., Ito, H., Mori, H., Honma, M., and Chiba, S. (1995) *Biosci., Biotechnol., Biochem. 59*, 459−463.
28. Kimura, A., Takata, M., Fukushi, Y., Mori, H., Matsui, H., and Chiba, S. (1997) *Biosci., Biotechnol., Biochem. 61*, 1091−1098.
29. Surarit, R., Matsui, H., Chiba, S., Svasti, J., and Srisomsap, C. (1996) *Biosci., Biotechnol., Biochem. 60*, 1265−1268.
30. Teshima, S., Hayashi, Y., Emi, S., and Ishimaru, K. (1991) *Clin. Chim. Acta 199*, 23−32.
31. Nakamura, A., Haga, K., and Yamane, K. (1994) *FEBS Lett. 337*, 66−70.
32. Imamura, H., Jeon, B. S., Wakagi, T., and Matsuzawa, H. (1999) *FEBS Lett. 457*, 393−396.
33. Chaurand, P., Luetzenkirchen, F., and Spengler, B. (1999) *J. Am. Soc. Mass Spectrom. 10*, 91−103.
34. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) *Nucleic Acids Res. 25*, 3389−3402.
35. Beccari, T., Stinchi, S., and Orlacchio, A. (1999) *Biosci. Rep. 19*, 157−162.
36. Chui, D., Oh, E. M., Liao, Y. F., Panneerselvam, K., Lal, A., Marek, K. W., Freeze, H. H., Moremen, K. W., Fukuda, M. N., and Marth, J. D. (1997) *Cell 90*, 157−167.
37. Goss, P. E., Baker, M. A., Carver, J. P., and Dennis, J. W. (1995) *Clin. Cancer Res. 1*, 935−944.
38. Bennet, J. K., Dembure, P. P., and Elsas, L. J. (1995) *Am. J. Med. Genet. 55*, 21−26.
39. Berg, T., Riise, H. M., Hansen, G. M., Malm, D., Tranebjaerg, L., Tollersrud, O. K., and Nilssen, O. (1999) *Am. J. Hum. Genet. 64*, 77−88.
40. Numao, S., He, S., Evjen, G., Howard, S., Tollersrud, O. K., and Withers, S. G. (2000) *FEBS Lett. 484*, 175−178.
41. van den Elsen, J. M., Kuntz, D. A., and Rose, D. R. (2001) *EMBO J. 20*, 3008−3017.
42. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) *Nucleic Acids Res. 22*, 4673−4680.

BI011017C