The loop structure of *Actinomycete* glycoside hydrolase family 5 mannanases governs substrate recognition

Yuya Kumagai[1,2], Keitaro Yamashita[3], Takayoshi Tagami[2], Misugi Uraji[1], Kun Wan[1], Masayuki Okuyama[2], Min Yao[3,4], Atsuo Kimura[2], and Tadashi Hatanaka[1]

[1]Okayama Prefectural Technology Center for Agriculture, Forestry and Fisheries, Research Institute for Biological Sciences (RIBS), Okayama 716-1241, Japan

[2]Research Faculty of Agriculture, Hokkaido University, Sapporo 060-8589, Japan

[3]Graduate School of Life Science and [4]Faculty of Advanced Life Science, Hokkaido University, Sapporo 060-0810, Japan

To whom correspondence may be addressed:

Tadashi Hatanaka

Research Institute for Biological Sciences (RIBS), Okayama, 7549–1 Kibichuo–cho, Kaga–gun, Okayama 716–1241, Japan

Tel: +81 866 56 9452, Fax: +81 866 56 9454; E-mail: hatanaka@bio–ribs.com

Running title: Substrate Recognition by *Actinomycete* mannanases

Abbreviations: ABEE, ethyl 4-aminobenzoate; CBM, carbohydrate binding module; GGM3, 6^{I},6^{II}-α-D-galactosyl mannotriose; GGM4, 6^{II},6^{III}-α-D-galactosyl mannotetraose; GGM5, 6^{III},6^{IV}-α-D-galactosyl mannopentaose; GH, glycoside hydrolase family; HPAEC-PAD, high-performance anion-exchange chromatography with pulsed amperometric detection; LBG, locust bean gum; M1 to M6, mannose to mannohexaose; StMan, *Streptomyces thermolilacinus* mannanase; StMandC, catalytic domain of StMan; STMan3dC, chimeric enzyme consisting of 36–246 of StMan and 227–330 of TfMan (*Thermobifida fusca* mannanase); STMan4dC, chimeric enzyme

consisting of 36–260 of StMan and 250–330 of TfMan; STMan5dC, chimeric enzyme consisting of 36–308 of StMan and 288–330 of TfMan; S(S-L7➔T-L7)dC, loop7 (276–283, GPPDQWGD) of StMandC changed to loop7 (256–262, HDHSDGN) of TfMan; S(S-L8➔T-L8)dC, loop8 (308–312, TDPV) of StMandC changed to loop8 (288–293, GGGVEY) of TfMan; S(S-L7/L8➔T-L7/L8)dC, loop7 and loop8 of StMandC changed to TfMan; TfMan, *Thermobifida fusca* mannanase; TfMandC, catalytic domain of TfMan; TSMan3dC, chimeric enzyme consisting of 29–226 of TfMan and 247–349 of StMan; T(T-L7➔S-L7)dC, loop7 of TfMandC changed to loop7 of StMan; T(T-L8➔S-L8)dC, loop8 of TfMandC changed to loop8 of StMan; T(T-L7/L8➔S-L7/L8)dC, loop7 and loop8 of TfMandC changed to loop7 and loop8 of StMan.

Keywords: mannanase; actinomycete; galactosylmannooligosaccharide; chimeric enzyme; glycoside hydrolase family 5

Abstract

Endo-β-1,4-mannanases from *Streptomyces thermolilacinus* (StMan) and *Thermobifida fusca* (TfMan) showed different substrate specificities. StMan hydrolyzed galactosylmannooligosaccharide (GGM5; 6^{III},6^{IV}-α-D-galactosyl mannopentaose) to GGM3 and M2 while TfMan hydrolyzed GGM5 to GGM4 and M1. To determine the region involved in the substrate specificity, we constructed chimeric enzymes of StMan and TfMan and evaluated their substrate specificities. Moreover, the crystal structure of the catalytic domain of StMan (StMandC) and the complex structure of the inactive mutant StE273AdC with M6 were solved at 1.60 and 1.50 Å resolution, respectively. Structural comparisons of StMandC and TfMandC lead to the identification of a subsite around −1 in StMandC which could accommodate a galactose branch. These findings demonstrate that the two loops (loop7 and loop8) are responsible for substrate recognition in GH5 actinomycete mannanases. In particular, Trp281 in loop7 of StMan, which is located in a narrow and deep cleft, plays an important role in its affinity toward linear substrates. Asp310 in loop8 of StMan specifically bound to the galactosyl unit in the −1 subsite.

Introduction

Plant biomass is an abundant carbon source and important sustainable biomaterial. The plant cell wall comprises cellulose, hemicelluloses, and lignin. Microbial enzymatic degradation is a key requirement for the involvement of the carbon cycle in the utilization of hard biomass [1-3]. Cellulose, a β-1,4-linked glucose polymer, can be converted to bioenergy [4]. Hemicellulose, a polymer composed of various types of sugars, also has the potential to be a sustainable biomaterial for bioenergy and bioactive compounds [5-7]. Mannan is one of the major hemicellulose components that exist as glucomannan or galactomannan in softwood or bean seeds [8]. Mannan consists of β-1,4-linked mannose polymers which are decorated with α-1,6-linked galactose branches. The amount of branched chains is dependent on the species and is also related to the physical properties of mannan [9-11].

For the deconstruction of mannan, cooperative degradation via a series of glycoside hydrolases [e.g., β-1,4-mannanase (**EC 3.2.1.78**), β-1,4-mannosidase (**EC 3.2.1.25**), α-galactosidase (**EC 3.2.1.22**)] and accessory enzymes (e.g., carbohydrate esterases) are necessary [12-14]. Mannanases, a major enzyme group used for mannan degradation, has been classified into three groups, glycoside hydrolase family (GH) 5, 26, and 113 (http://www.cazy.org/). Among these groups, GH5 mannanase is considered to primarily function in hemicellulose deconstruction because most GH5 mannanases are extracellular hydrolases and have carbohydrate binding modules (CBMs) [15, 16]. CBMs usually bind to insoluble or soluble saccharides with various conformations via a linker domain to increase the catalytic efficiency of the enzymes [17]. Recently, accessory enzymes with CBMs have been speculated to be involved in the cooperative deconstruction of the plant cell wall [18, 19]. The molecular architecture of most GH26 mannanases consists of only a catalytic domain which cannot hydrolyze mannan in the plant cell wall, while some GH26 mannanases with CBMs do exhibit hydrolysis activity toward mannan [20], suggesting that target saccharides are differentiated by mannanases depending on whether they have CBMs or not [21]. Therefore, the analysis of the

molecular architecture of these enzymes is important for improving our understanding of the deconstruction of hemicellulose. The degradation of linear saccharides by catalytic domains has been extensively studied [22-25]. Some GH5 mannanases exhibit high activity toward short substrates [26-28]. Specificity toward branched mannan has been studied in *Streptomyces* sp. SirexAA-E mannanase [29]. However, it is still unclear how the catalytic domains of these enzymes recognize and degrade linear or branched mannan structures.

We have studied the relationship between the structure and function of GH5 actinomycete mannanases [30-33]. The end-products of locust bean gum (LBG: galactosyl mannan) degraded by a catalytic domain of mannanase from *Streptomyces thermolilacinus* (StMandC) were mainly M2 and M3, while that from *Thermobifida fusca* (TfMandC) were mainly M1, and M2 [33]. Mannanase activity toward mannan usually decreases as the degree of modification of branched chain galactoses increases. Understanding the mechanism of mannanase catalysis of galactosyl-mannan should provide further insight into mannan degradation, and improve the use of an optimal combination of mannanase and other hydrolases or accessory enzymes for mannan degradation.

In this study, by using linear mannooligosaccharides and branched galactosyl mannooligosaccharide, we investigated the substrate specificities of StMandC and TfMandC. Moreover, we investigated the role of two loops (loop7 and loop8) in their substrate specificity toward linear mannooligosaccharides and branched galactosyl mannooligosaccharides (GGM5) by using StMan, TfMan, their chimeric enzymes, and mutants. To confirm the kinetic results, we solved the crystal structure of StMandC and a complex structure of the inactive mutant StE273AdC with M6. Our data demonstrated the relationship between the structure and substrate specificity of GH5 actinomycete mannanases.

Results

Identification of the region involved in substrate specificity

When StMandC hydrolyzed GGM5, two peaks were detected by HPAEC-PAD analysis. One peak was determined to be mannobiose (M2) and the remaining degradation product was identified as GGM3 (Fig. 1A and B). However, when TfMandC hydrolyzed GGM5, mannose (M1) was detected (Fig. 1B). Mannose residues at the reducing and non-reducing termini were present in GGM5. Using a labelling agent (ABEE), we modified the reducing terminus of the sugar to evaluate the mannanase hydrolysis pattern of GGM5. The hydrolysis product of GGM5-ABEE by TfMandC was M1-ABEE, indicating that TfMandC hydrolyzed the reducing terminus of GGM5 (Fig. 1C).

Four chimeric mannanases (STMan3dC consisting of residues 36–246 of StMan and 227–330 of TfMan, STMan4dC consisting of 36–260 of StMan and 250–330 of TfMan, STMan5dC consisting of 36–308 of StMan and 288–330 of TfMan, and TSMan3dC consisting of 29–226 of TfMan and 247–349 of StMan) were constructed to evaluate the GGM5 hydrolysis pattern (Fig. 2A). STMan3dC and STMan4dC hydrolyzed GGM5 to form GGM4 and M1 and TSMan3dC produced GGM3 and M2 (Fig. 3A). The C-terminal region (amino acid residues 258–349 in StMan) has been speculated to determine the substrate specificity of StMan. STMan5dC was significantly less reactive toward GGM5 compared with the parent enzymes and the three other chimeric mutants (for example, GGM5 activity of StMandC and STMan5dC were 10.1 and 0.012 U/mg, respectively). STMan5dC exhibited the GGM5-cleavage reactions of two wild-type enzymes in terms of its GGM5 hydrolysis products (M1, M2, GGM3, and GGM4), indicating that the GGM5 hydrolysis pattern could be determined by two regions: the amino acid sequences 258–302 and 303–349 of StMan (Figs. 2A and 3A). The alignment of these regions of StMandC and TfMandC showed that both regions contain two single-loop-structures with low sequence identity [loop7: 276–283 (GPPDQWGD) for StMan and 256–262 (HDHSDGN) for TfMan; loop8: 309–312 (TDPV) for StMan and 288–293 (GGGVEY) for TfMan] (Fig. 2B). We then constructed

six substitution mutants, of which loop7 and loop8 of StMandC were replaced with those of TfMandC and vice versa (S(S-L7➔T-L7)dC, S(S-L8➔T-L8)dC, S(S-L7/L8➔T-L7/L8)dC, T(T-L7➔S-L7)dC, T(T-L8➔S-L8)dC, and T(T-L7/L8➔S-L7/L8)dC; see Fig. 2C for construction of mutants), and evaluated the GGM5 hydrolysis patterns with these mutants. S(S-L7➔T-L7)dC, S(S-L7/L8➔T-L7/L8)dC, T(T-L7➔S-L7)dC could not hydrolyze GGM5. S(S-L8➔T-L8)dC and T(T-L8➔S-L8)dC hydrolyzed GGM5 with significantly lower activity to produce M1, M2, GGM3, and GGM4 (Fig. 3B). The GGM5 hydrolysis products of T(T-L7/L8➔S-L7/L8)dC having loop7 and loop8 of StMan were identical to those of StMan (Fig. 3B). Based on these results, we concluded the substrate specificity of mannanases is determined by the two loop-structures.

Kinetic parameters of StMandC and TfMandC

Both of StMandC and TfMandC displayed no activity on M2 and drastically low activity on M3, making it impossible to determine the kinetic parameters for these two substrates. Therefore, the kinetic parameters of mannanases toward mannooligosaccharides (M4, M5, and M6) and GGM5 are listed in Table 1. Among these substrates, both enzymes showed the highest activity toward M6. The k_{cat}/K_m values toward M5 and M4 of StMandC were approximately 10- and 100-fold lower than that of M6. The k_{cat}/K_m values toward M4 of TfMandC were approximately 100-fold lower than that of M5 and M6. The k_{cat}/K_m values of StMandC toward M4, M5, and M6 was 1.6-, 13-, and 2.4-fold lower than those of TfMandC, respectively. In particular, TfMandC had a 13-fold higher k_{cat} value toward M5 than that of StMandC. However, both enzymes were the least active toward GGM5. The values of k_{cat}/K_m toward GGM5 were very low compared with linear mannooligosaccharides. The k_{cat}/K_m and k_{cat} of StMandC toward GGM5 were 13- and 10-fold higher than those of TfMandC, respectively.

Crystal structures of StMandC and M6-complexed StE273AdC

The crystal structures of StMandC and its inactive mutant, StE273AdC, with substrate were solved. The overall structure of StMandC revealed that the protein was composed of a $(\beta/\alpha)_8$-barrel fold similar to that of other GH5 mannanases belonging to Clan GH-A (Fig. 4A). StMandC possessed calcium ion-binding site and one residue (Glu286) has been found as calcium ion-binding residue [31]. From crystal structure of StMandC, the side chain of Glu286 and Asp283, and main chain of Gly276 and Pro284 were involved in calcium ion-binding in loop7. The structure of StE273AdC in complex with substrate was successfully determined by soaking with M6 (Fig. 4A). The electron density map of mannose residues was found at subsites −4 to +3 (Fig. 4B), indicating various binding patterns of M6 to StE273AdC, since seven subsites (i.e., seven mannose units) were found by soaking with M6. The subsite −1 was occupied by two moieties of mannose, an internal moiety in the boat-form conformation $B_{2,5}$ and a reducing-terminal moiety in the skew-boat conformation $^{1}S_{5}$, whose equatorial O1 was interacted with Tyr246 by a 2.72 Å hydrogen bond (Fig. 4C). These conformations represented one of the structures of the mannose chain [34]. The plus subsites in StMandC consisted of a hydrophobic cleft, Trp215 at subsite +1, Trp281 at subsite +2, and Trp219 at subsite +3. Gln217 in loop5 bound to the OH-C(3) of mannose at subsite +2 with a 2.84 Å hydrogen bond. Binding of OH–C(2) mannose was thought to play a key role in determining the substrate specificity between mannanases and cellulases [22]. Trp219 and Asn182 bound to the OH–C(2) of mannose at subsite +2 with a 2.99 and 3.33 Å hydrogen bond, respectively. These residues created a narrow and deep cleft at the plus subsite side. The crystal structure of StE273AdC complexed with substrate revealed that the cleft was suitable for the incorporation of mannose residues containing a twisted glycoside linkage at subsites −1 and +1. The superimposed model of StMandC and StE273AdC showed that loop7 of StE273AdC shifted toward the substrate (Fig. 4A and B). Trp281 in loop7 moved into proximity to subsite +2 mannose, which would increase the interaction with the substrate (Fig. 4D). Thus, we speculated that Trp281 in loop7 plays an important role in substrate recognition.

The superimposed model of StE273AdC with M6 and *T. fusca* KW3 mannanase with M3 (PDB code: **3MAN**) showed that the mannose main chain was almost coincident at subsites −3 to −2, and the substrate-binding residues were well conserved, with the exception of Val263 (3MAN) bound to mannose at subsite −1, which did not correspond to the StMandC residue (Fig. 5A) [22]. Thr309 in loop8 of StMan, equivalent to Gly260 in 3MAN, was most likely involved in the binding of OH-C(6) mannose at subsite −3 (Fig. 5A). The plus subsites were compared with *T. fusca* KW3 mannanase (PDB code: **1BQC**) because the structure of 3MAN was lacking loop7 (Fig. 5B). Amino acid residues in loop4 and loop5 comprising the plus subsites were conserved (Fig. 5B). However, those of loop7 and loop8 comprising the opposite plus subsites were different: 1BQC had a wide cleft because of a short loop7 compared with that of StMandC, while Trp281 in loop7 of StMan constituted a narrow and deep cleft with loo4 and loop5 (Fig. 6). Additionally, loop7 and loop8 in StMandC generated a space around subsite −1 (Fig. 6B). It was supposed that this space accommodated the galactose branch, resulting in the different GGM5 hydrolysis patterns. The crystal structure of StMandC revealed that Asp310 in loop8 binds to the galactose branch around the −1 subsite.

Amino acid residues related to the substrate specificity of mannanases

From the structure of StMandC, we speculated that Trp281 in loop7 and Asp310 in loop8 were a mannose-main-chain binding residue and galactose-branch-chain binding residue, respectively. We constructed alanine mutants of Trp281 (StW281AdC) and Asp310 (StD310AdC). The kinetic parameters for StW281AdC decreased for the substrates. Their k_{cat}/K_m toward linear mannooligosaccharides (M4–M6) was 7- to 2-fold lower than those of StMandC (Table 1). The k_{cat}/K_m of StD310AdC toward GGM5 decreased approximately 5-fold compared with that of StMandC (Table 1).

Discussion

Understanding enzymatic properties is essential for determining an enzyme's role

in nature. Herein, we found that the substrate specificity of two types of GH5 actinomycete mannanases were related to their loop structures. Namely, a combination of loop7 and loop8 of StMandC is essential for accommodation of the galactose branch at the –1 subsite, however the residues critical for determining the substrate specificity were unclear.

The value of k_{cat} of StMandC toward GGM5 was the main factor determining the value of k_{cat}/K_m. The value of k_{cat} was 10^2–10^4-fold lower compared with those for the linear substrates. GGM5 is composed of a mannopentaose main chain with two galactose branches, suggesting that the galactose branches affected the k_{cat}. The crystal structure of GH5 endoglucanase and β-mannosidase revealed that the sugar at subsite −1 was distorted, and the glycoside linkage between subsites −1 and +1 was twisted [28, 35]. The complex structure of GH26 mannanase with a mannooligosaccharide possessing a galactose branch at the −1 subsite indicated that the galactose residue affected the mannose chain at the +1 subsite [36]. The complex structure of StE273AdC also revealed that the glycoside linkage between subsites −1 and +1 was twisted. The distortion of the sugar at the −1 resulted in a decrease in the distance between the OH–C(6) mannose at subsite −1 and the OH–C(3) mannose at subsite +1 and a decrease in the catalytic efficiency toward GGM5. Asp310 in loop8, which could not bind to the mannose main chain, is located around the −1 subsite. The decrease in k_{cat}/K_m toward GGM5 was primarily related to the decrease in k_{cat}. This indicated that Asp310 may assist in the distortion of the sugar at the −1 subsite by binding the galactose residue. Trp281 in loop7 of StMandC created a narrow cleft at subsite +2, thereby allowing a mannose residue to bind with a twisted glycoside linkage between subsites −1 and +1. The parameter most affected by the mutation was the k_{cat} toward M4, indicating that Trp281 was an important residue for the catalytic activity toward the short substrate M4 (Table 1). It was thought that Trp281 was also necessary for distortion of the sugar at subsite −1. These results suggest that Trp281 in loop7 is involved in the catalytic activity toward short linear substrates, while Asp310 in loop8 is specifically involved in the binding of branched

substrates. The hydrolysis patterns of GGM5 resulting from these mutants were similar to that of StMan (i.e., GGM5 was hydrolyzed to GGM3 and M2) (data not shown). Thus, other residues were also involved in determining the hydrolysis pattern of GGM5.

Streptomyces sp. SirexAA-E mannanase (SACTE_2347), which showed 79% identity with StMandC and possessed loops L1 and L2, corresponding to loop7 and loop8 of StMandC, respectively, hydrolyzed GGM5 to GGM3 and M2 [29]. The crystal structure of SACTE_2347 showed the space which could accommodate the galactose branched around the −1 subsite. Trp281 and Asp310 of StMan were equivalent to Tyr281 and Asp310 of SACTE_2347, respectively, implying that their possible functions are short substrate-binding (Tyr281 of SACTE_2347) and distortion of the sugar at the −1 subsite (Asp310 of SACTE_2347). Thr309 in loop8, which was conserved in both enzymes, might be the key residue to determine substrate specificity toward branched substrates by disturbing the accommodation of galactose branches around the −3 subsite.

The value of k_{cat}/K_m for TfMandC toward GGM5 decreased 10^5-fold compared with that for M5. For StMandC, the difference in k_{cat}/K_m between GGM5 and M5 was 10^3-fold. From analysis of the hydrolysis products of GGM5, it was determined that TfMandC could accommodate the galactose branches in the vicinity of subsites −3 and −2, while StMandC accommodated galactose branches around subsites −2 and −1. When using M5 as a substrate, both TfMandC and StMandC remove M1 and M2 from M5 to a similar amount (Fig. 7). However, the two enzymes had different reactivities toward GGM5 (Fig. 1). Therefore, the decrease in k_{cat}/K_m for TfMandC and StMandC toward GGM5 was related to the steric hindrance created by the accommodation of the galactose branches. Around subsite −2, loop2 and loop3 could accommodate the galactose branch (Fig. 5). The putative galactose binding residues between StMandC and TfMandC appear to be conserved: Arg108 in loop2, and Thr139 and Asp144 in loop3 of StMandC correspond to Arg58 in loop2, Thr89 and Gln94 in loop3 of TfMandC. Around the −3 subsite, the complex structure of StMandC with M6 showed a 2.7 Å hydrogen bond between Thr309 and the OH–C(6) of mannose at subsite −3. Therefore, Thr309 should

inhibit the accommodation of the galactose branch at the −3 subsite, leading to the substrate specificity of mannanases. At the −1 subsite, Val263 in loop8 of TfMandC would fill the space [22] (Fig. 6C). Loop8 was necessary for the substrate specificity of TfMan-type enzymes because the accommodation of a galactose branch at the −3 subsite would be attributed to loop8. We assumed that the substrate specificity of the loop8 substitution mutant S(S-L8➔T-L8)dC changed to a TfMandC-type, i.e., galactose branches were accommodated at subsites −3 and −2. However, S(S-L8➔T-L8)dC displayed less hydrolytic reaction: GGM5 to GGM4 and M1. Additionally, the loop8 substitution mutant T(T-L8➔S-L8)dC also could not hydrolyze GGM5 to GGM3 and M2, suggesting that loop7 contained important residues for substrate specificity. Therefore, we speculated that the difference in the accommodation of mannose residues at the plus subsites would be directly related to the enzyme's substrate specificity.

The synergistic degradation of hemicelluloses has been proposed to be carried out by GH5 and GH26 mannanases [21]. The genome of *Actinomycete* sp. contains both GH5 and GH26 mannanase genes. Most GH5 mannanases in *Actinomycetes* have loop-structures similar to those of StMan or TfMan (Fig. 8). Interestingly, several species such as *Streptomyces coelicolor* A3(2) and *Streptomyces scabiei* 87.22 possess more than two kinds of GH5 mannanase genes similar to those of StMan and TfMan (Fig. 8). SACTE_2347 was classified as a StMan-type enzyme, supporting the close relationship between the sequence alignment and the function of enzymes. The end-products from LBG by these enzymes were mainly M2 [29, 33]. β-1,4-Mannosidases from GH1 and GH2 catalyzed further hydrolysis of M2, and some GH5 mannanases could hydrolyze M2 [26, 27]. GH5 β-1,4-mannosidase from *Actinomycete* sp. has not been found, however it is possible to hydrolyze mannan to mannose by symbiotic degradation among GH5 mannanases. A combination of GH5 mannanases would have the potential to act synergistically to deconstruct lignocellulosic materials.

In conclusion, this study highlights the different enzyme properties of GH5 mannanases from *Actinomycete*s sp. Both loop7 and loop8 are key regions which

determine substrate specificity. StMan-loops can contribute to the hydrolysis of a mannose chain decorated with a galactose branch around the −1 subsite, while TfMan-loops show high activity toward linear mannan. The alignment shows that the many characteristics of mannanases are conserved in *Actinomycetes* sp.

Materials and methods

Construction, expression, and purification of Actinomycete mannanases

To evaluate the region involved in mannanase substrate specificity, we used expression plasmids encoding the catalytic domains of StMan (StMandC) and TfMan (TfMandC) and chimeric enzymes combining StMandC and TfMandC (STMan3dC and TSMan3dC) [30]. The other chimeric enzymes (STMan4dC and STMan5dC) were prepared using the Repeat-length-Independent Broad-Spectrum (RIBS) DNA shuffling method described in the following section. The recombinant proteins were expressed in *Escherichia coli* BL21-Gold (DE3) cells (Agilent Technologies, Palo Alto, CA, USA) harboring the pET28a construct (mannanase). The recombinant proteins were purified from cells as previously described [31] and concentrated using a 10,000-MW cutoff Amicon® ultra membrane (Millipore, Billerica, MA, USA) for use in subsequent experiments. The purities of the recombinant proteins were confirmed by using SDS-PAGE [37]. The protein concentrations were determined by the Bradford method [38] using BSA as the standard.

Preparation of chimeric enzymes using the RIBS shuffling system

Chimeric enzymes were prepared using the RIBS *in vivo* DNA shuffling system, which is an improved method of chimera genesis based on highly frequent deletion formation using the *E. coli ssb*-3 strain [39]. The parental mannanase genes, gentamicin resistance gene (Gm^r) and *E. coli rpsL*$^+$ gene [streptomycin-sensitive (Sm^s)], were tandemly cloned into the *Nde*I-*Hin*dIII sites of pET28a. The rank order of these genes was *stman*-Gm^r-*E.coli rpsL*$^+$-*tfman*, and then *E. coli* MK1019 [*ssb*-3 *rpsL* (Sm^r)]

harboring pET28a (StMan/Gm^r-*rpsL*/TfMan) was obtained. Thirty transformants were cultured overnight in Luria Broth (LB) medium containing 50 μg/mL chloramphenicol and each culture was spread and then cultivated on LB plates containing 50 μg/mL chloramphenicol and 50 μg/mL streptomycin. Plasmids containing chimeric mannanase (*c-man*) genes were isolated from 80 colonies. A total of 10 clones with different *c-man* genes were obtained. The catalytic domain of the enzymes was produced by PCR using a previously reported method [39]. Each chimeric construct was named based on recombination positions from the N-terminus of StMan. For example, chimera1-Man (*c1-man*) was designated as ST-C85. Among the *c-man* genes, the catalytic domain mutants of ST-C258dC and ST-C302dC, which were designated STMan4dC and STMan5dC, respectively were used in this study.

Preparation of mutant enzymes

To evaluate the effects of loop-structure on substrate specificity, we generated loop7 and loop8 substitution mutants (S(S-L7➔T-L7)dC, S(S-L8➔T-L8)dC, S(S-L7/L8➔T-L7/L8)dC, T(T-L7➔S-L7)dC, T(T-L8➔S-L8)dC, and T(T-L7/L8➔S-L7/L8)dC) by PCR, restriction enzyme digestion, and ligation. PCR was performed in the following two conditions: 1) PCR was performed to substitute the amino acid residues in the loop with a set of primers (S(S-L7➔T-L7)dC, StL7Fw1 and StCDRe; S(S-L8➔T-L8)dC, StFw and StL8Re1, and StL8Fw1 and StCDRe); and 2) PCR was performed to insert the restriction enzyme site for the digestion and ligation of mannanase gene fragments into pET28a using a set of primers (S(S-L7➔T-L7)dC, StFw and StL7Re1; StL7Fw2 and StCDRe; and S(S-L8➔T-L8)dC, StFw and StL8Re2; StL8Fw2 and StCDRe) (Table S1). The primer sets used to construct the loop-substitution mutants for TfMandC using a previously reported procedure are listed in Table S1 [30]. The mutant genes (StW281AdC, StD310AdC, and inactive mutant StE273AdC) were generated by PCR using site-directed mutagenesis (Prime Star GXL DNA polymerase kit; Takara Bio Inc.) with the following sets of primers (StW281A-S:

5′-CGACCAGGCGGGCGACCCGGACGAGGAC-3' and StW281A-AS: 5'-GGTCGCCCGCCTGGTCGGGCGGCCCGCC-3'; and StD310A-S: 5'-AACACCGCCCCCGTCCTCGACCTGGCG-3' and StD310A-AS: 5'-GACGGGGGCGGTGTTCCCGCTCCACGA-3'; StE273A-S: 5'-ATCGGGGCGTTCGGCGGGCCGCCCGAC-3' and StE273A-AS: 5'-GCCGAACGCCCCGATGACCAGCGGCAG-3') (the underline shows the position changed for alanine) using pET28a-StMandC as a template.

Mannanase activity assay

Mannanase activity was determined at 40°C for 10 min in a reaction mixture (0.1 mL) containing an appropriate amount of enzyme, 1% (w/v) LBG, 1 mM $CaCl_2$, and 50 mM piperazine-*N*,*N′*-bis-(2-ethanesulfonic acid) PIPES (pH 7.0) [30, 33]. The amount of reducing sugars produced by the reaction was determined using the 3,5-dinitrosalicylic acid method [40]. One unit of mannanase activity was defined as the amount of enzyme that liberates reducing sugars equivalent to 1.0 μmol mannose per min. All activity assays were performed in triplicate.

Evaluation of GGM5 hydrolysis patterns

GGM5 (Megazyme International Ireland Ltd., Bray, Ireland) was derivatized with ethyl 4-aminobenzoate (ABEE) [41]. GGM5-ABEE was hydrolyzed by StMandC and TfMandC, and degradation products were analyzed using an HPLC equipped with a C18 column (Cadenza CD-C18, 150 x 4.6 mm, Imtakt Corp., Kyoto, Japan). The samples were eluted using the following gradient with a flow rate of 1.0 mL/min: 20% acetonitrile for 0–2 min, a linear gradient to 50% acetonitrile for 2–15 min, and 50% acetonitrile for 15–20 min. The eluted sugars were detected by UV absorption at 305 nm.

Evaluation of the kinetic parameters of mannanases

The kinetic parameters of StMandC, TfMandC, StW281AdC, and StD310AdC toward mannooligosaccharides (Megazyme International Ireland Ltd., Bray, Ireland) were determined using a stopped-assay method involving high-performance anion-exchange chromatography with pulsed amperometric detection (HPAEC-PAD) (Dionex, Sunnyvale, CA, USA). A Carbopac PA1 column (4 × 250 mm) (Dionex) with an isocratic flow of 100 mM NaOH was used for separation. Assay conditions were 0.02–8 mM M4, 0.04–6 mM M5, 0.04–2 mM M6, and 1–50 mM GGM5 with an appropriate amount of enzymes. Aliquots were withdrawn at four time points during the 5–60 min incubation time, and the reactions were terminated by boiling. The formed hydrolysis products of GGM5 were determined by HPAEC-PAD. Kinetics for M4–M6 were determined by following the decrease of substrate after hydrolysis because several products were formed. The k_{cat} and K_m values were determined by relationship between substrate concentrations and initial hydrolytic velocities using Origin Software (Lightstone Corp., Tokyo, Japan).

Crystallization and data collection of StMandC and StE273AdC with M6

Crystallization was performed using the hanging-drop vapor diffusion method at 20°C. StMandC and StE273AdC were concentrated to 10 mg/ml in 0.1 M Tris-HCl, pH 7.0. The crystallization condition used was 1.1 M sodium malonate (pH 7.0), 0.1 M HEPES (pH 7.0), and 0.5% (v/v) Jeffamine® ED-2001 (pH 7.0). Glycerol was used at a concentration of 25% as the cryoprotectant. The structure of StE273AdC in complex with substrate was determined by soaking with M6. X-ray diffraction data of the StMandC crystal was collected on beamline BL44XU at SPring-8 (Hyogo, Japan) at a wavelength of 1.0000 Å using MX255HE CCD detector (Rayonix, USA). The X-ray diffraction data of the StE273AdC crystal was collected on beamline NE3A at the Photon Factory Advanced Ring, KEK (Tsukuba, Japan) at a wavelength of 1.0000 Å using the Q270 CCD detector (ADSC, USA). Both diffraction datasets were collected from single crystals under a stream of nitrogen at 100 K. The diffraction data sets were indexed,

integrated, and scaled with *XDS* [42]. The StMandC crystal belonged to the $P2_12_12_1$ space group with the cell dimensions a = 65.86 Å, b = 100.85 Å, and c = 105.24 Å and diffracted to 1.60 Å resolution, and the StE273AdC crystal belonged to the $P2_12_12_1$ space group with the cell dimensions a = 65.71 Å, b = 100.71 Å, and c = 104.74 Å and diffracted to 1.50 Å resolution. The StMandC structure was determined by the molecular replacement method with *phenix.automr* [43, 44] using **1BQC** (Protein Data Bank code of *Thermobifida fusca* KW3 mannanase) as a search model. Two molecules are present in the asymmetric unit. The resultant model was automatically rebuilt by *ARP*/*wARP* [45] using the calculated phases. The refinement was converged by several cycles of manual model corrections with *Coot* [46] and refinement using *phenix.refine* [47]. The StE273AdC structure was determined by rigid body refinement of the StMandC structure, followed by several cycles of manual model corrections with *Coot* [46] and refinement using *phenix.refine* [47]. Ramachandran plot analysis was performed using *MolProbity* [48]. Coordinates and structure factors have been deposited in the Protein Data Bank under codes **3WSU** and **4Y7E**. Data processing and refinement statistics are given in Table 2. The graphical representations were prepared using *PyMOL* (DeLano Scientific; The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC.) and *MolSoft* [49].

References

1. Pauly M & Keegstra K (2008) Cell-wall carbohydrates and their modification as a resource for biofuels. *Plant J* **54**, 559–568.

2. Demirbas MF (2009) Biorefineries for biofuel upgrading: a critical review. *Appl Energy* **86**, S151–S161.

3. Peralta-Yahya PP, Zhang F, del Cardayre SB & Keasling JD (2012) Microbial engineering for the production of advanced biofuels. *Nature* **488**, 320–328.

4. Sticklen MB (2008) Plant genetic engineering for biofuel production: towards affordable cellulosic ethanol. *Nat Rev Genet* **9**, 433–443.

5. Olsson L & Hahn-Hagerdal B (1996) Fermentation of lignocellulosic hydrolysates for ethanol production. *Enzyme Microb Technol* **18**, 312–331.

6. Himmel ME, Ding SY, Johnson DK, Adney WS, Nimlos MR, Brady JW & Foust TD (2007) Biomass recalcitrance: engineering plants and enzymes for biofuels production. *Science* **315**, 804–807.

7. Scheller HV & Ulvskov P (2010) Hemicelluloses. *Annu Rev Plant Biol* **61**, 263–289.

8. Moreira LR & Filho EX (2008) An overview of mannan structure and mannan-degrading enzyme systems. *Appl Microbiol Biotechnol* **79**, 165–178.

9. Bulpin PV, Gidley MJ, Jeffcoat R & Underwood DR (1990) Development of a biotechnological process for the modification of galactomannan polymers with plant alpha-galactosidase. *Carbohydr Polym* **12**, 155–168.

10. Daas PJ, Schols HA & de Jongh HH (2000) On the galactosyl distribution of commercial galactomannans. *Carbohydr Res* **329**, 609–619.

11. Rinaudo M (2004) Role of substitutents on the properties of some polysaccharides. *Biomacromolecules* **5**, 1155–1165.

12. Clarke JH, Davidson K, Rixon JE, Halstead JR, Fransen MP, Gilbert HJ & Hazlewood GP (2000) A comparison of enzyme-aided bleaching of softwood paper pulp using combinations of xylanase, mannanase and α-galactosidase. *Appl Microbiol Biotechnol* **53**, 661–667.

13. Puchart V & Biely P (2005) Glycosylation of internal sugar residues of oligosaccharides catalyzed by α-galactoside from *Aspergillus fumigatus*. *Biochim Biophys Acta* **1726**, 206–216.

14. Dhawan S & Kaur J (2007) Microbial mannanases: an overview of production and applications. *Crit Rev Biotechnol* **27**, 197–216.

15. Boraston AB, Bolam DN, Gilbert HJ & Davies GJ (2004) Carbohydrate-binding modules: fine-tuning polysaccharide recognition. *Biochem J* **382**, 769–781.

16. Gilbert HJ, Knox JP & Boraston AB (2013) Advances in understanding the molecular basis of plant cell wall polysaccharide recognition by carbohydrate-binding

modules. *Curr Opin Struct Biol* **23**, 669-677.

17. da Silva VM, Colussi F, Neto MO, Braz ASK, Squina FM, Oliveira CL & Garcia W (2014) Modular hyperthermostable bacterial endo-beta-1,4-mannanase: molecular shape, flexibility and temperature-dependent conformational changes. *PLoS One* **9(3)**, e92996.

18. Montanier C, Money VA, Pires VM, Flint JE, Pinheiro BA, Goyal A, Prates JA, Izumi A, Stålbrand H, Morland C, Cartmell A, Kolenova K, Topalas E, Dodson EJ, Bolam DN, Davies GJ, Fontes CM & Gilbert HJ (2009) The active site of a carbohydrate esterase displays divergent catalytic and noncatalytic binding functions. *PLoS Biol* **7**, e71.

19. Zhang X, Rogowski A, Zhao L, Hahn MG, Avci U, Knox JP & Gilbert HJ (2014) Understanding how the complex molecular architecture of mannan degrading hydrolases contributes to plant cell wall degradation. *J Biol Chem* **289**, 2002–2012.

20. Couturier M, Roussel A, Rosengre A, Leone P, Stålbrand H & Berrin JG (2013) Structural and biochemical analyses of glycoside hydrolase families 5 and 26 β-(1,4)-mannanases from *Podospora anserina* reveal differences upon manno-oligosaccharide catalysis. *J Biol Chem* **288**, 14624–14635.

21. Hogg D, Pell G, Dupree P, Goubet F, Martin-Orue SM, Armand S & Gilbert HJ (2003) The modular architecture of *Cellvibrio japonicus* mannanases in glycoside hydrolase families 5 and 26 points to differences in their role in mannan degradation. *Biochem J* **371**, 1027–1043.

22. Hilge M, Gloor SM, Rypniewski W, Sauer O, Heightman TD, Zimmermann W, Winterhalter K & Piontek K (1998) High–resolution native and complex structures of thermostable beta-mannanase from *Thermomonospora fusca* – substrate specificity in glycosyl hydrolase family 5. *Structure* **6**, 1433–1444.

23. Larsson AM, Anderson L, Xu B, Muñoz IG, Usón I, Janson JC, Stålbrand H & Ståhlberg J (2006) Three-dimensional crystal structure and enzymatic characterization of β-mannanase Man5A from blue mussel *Mytilus edulis*. *J Mol Biol* **357**, 1500–1510.

24. Tailford LE, Ducros VM, Flint JE, Roberts SM, Morland C, Zechel DL, Smith N, Bjørnvad ME, Borchert TV, Wilson KS, Davies GJ & Gilbert HJ (2009) Understanding how diverse β-mannanases recognize heterogeneous substrates. *Biochemistry* **48**, 7009–7018.

25. Santos CR, Paiva JH, Meza AN, Cota J, Alvarez TM, Ruller R, Prade RA, Squina FM & Murakami MT (2012) Molecular insights into substrate specificity and thermal stability of a bacterial GH5-CBM27 endo-1,4-β-D mannanase. *J Struct Biol* **177**, 469–476.

26. Dias FM, Vincent F, Pell G, Prates JA, Centeno MS, Tailford LE, Ferreira LM, Fontes CM, Davies GJ & Gilbert HJ (2004) Insights into the molecular determinants of substrate specificity in glycoside hydrolase family 5 revealed by the crystal structure and kinetics of *Cellvibrio mixtus* mannosidase 5A. *J Biol Chem* **279**, 25517–25526.

27. Goncalves AM, Silva CS, Madeira TI, Coelho R, de Sanctis D, San Romão MV & Bento I (2012) Endo-β-D-1,4-mannanase from *Chrysonilia sitophila* displays a novel loop arrangement for substrate selectivity. *Acta Crystallogr* D *Biol Crystallogr* **68**, 1468–1478.

28. Zhou, P, Liu Y, Yan Q, Chen Z, Qin Z & Jiang Z (2014) Structural insights into the substrate specificity and transglycosylation activity of a fungal glycoside hydrolase family 5 beta-mannosidase. *Acta Crystallogr* D *Biol Crystallogr* **70**, 2970–2982.

29. Takasuka TE, Acheson JF, Bianchetti CM, Prom BM, Bergeman LF, Book AJ, Currie CR & Fox BG (2014) Biochemical properties and atomic resolution structure of a proteolytically processed β-mannanase from cellulolytic *Streptomyces* sp. SirexAA-E. *PLoS One* **9(4)**, e94166.

30. Kumagai Y, Usuki H, Yamamoto Y, Yamasato A, Arima J, Mukaihara T & Hatanaka T (2011) Characterization of calcium ion sensitive region for beta-mannanase from *Streptomyces thermolilacinus*. *Biochim Biophys Acta* **1814**, 1127–1133.

31. Kumagai Y, Kawakami K, Mukaihara T, Kimura M & Hatanaka T (2012) The structural analysis and the role of calcium binding site for thermal stability in mannanase.

Biochimie **94**, 2783–2790.

32. Kumagai Y, Kawakami K, Uraji M & Hatanaka T (2013) Binding of bivalent ions to actinomycete mannanase is accompanied by conformational change and is a key factor in its thermal stability. *Biochim Biophys Acta* **1834**, 301–307.

33. Kumagai Y, Kawakami K, Uraji M & Hatanaka T (2013) Effect of the binding of bivalent ion to the calcium-binding site responsible for the thermal stability of actinomycete mannanase: potential use in production of functional mannooligosaccharides. *J Mol Catal* B *Enzym* **94**, 63–68.

34. Ardevol A, Biarnes X, Planas A & Rovira C (2010) The conformational free-energy landscape of beta-D-mannopyranose: Evidence for a $^{1}S_{5} \rightarrow B_{2,5} \rightarrow {}^{O}S_{2}$ catalytic itinerary in beta-mannosidases. *J Am Chem Soc* **132**, 16058–16065.

35. Kim HW & Ishikawa K (2011) Functional analysis of hyperthermophilic endocellulase from *Pyrococcus horikoshii* by crystallographic snapshots. *Biochem J* **437**, 223–230.

36. Cartmell A, Topakas E, Ducros VM, Suits MD, Davies GJ & Gilbert HJ (2008) The *Cellvibrio japonicus* mannanase CjMan26C displays a unique exo-mode of action that is conferred by subtle changes to the distal region of the active site. *J Biol Chem* **283**, 34403–34413.

37. Porzio MA & Pearson AM (1977) Improved resolution of myofibrillar proteins with sodium dodecyl sulfate-polyacrylamide gel electrophoresis. *Biochim Biophys Acta* **490**, 27–34.

38. Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein–dye binding. *Anal Biochem* **72**, 248–254.

39. Mori K, Mukaihara T, Uesugi Y, Iwabuchi M & Hatanaka T (2005) Repeat-Length-Independent Broad-Spectrum Shuffling, a novel method of generating a random chimera library *in vivo*. *Appl Environ Microbiol* **71**, 754–760.

40. Miller GL (1959) Use of dinitrosalicylic acid reagent for determination of reducing

sugar. *Anal Chem* **31**, 426–428.

41. Wang WT, LeDonne NC, Ackerman B & Sweeley CC (1984) Structural characterization of oligosaccharides by high-performance liquid chromatography, fast-atom bombardment-mass spectrometry, and exoglycosidase digestion. *Anal Biochem* **141**, 366–381.

42. Kabsch W (2010) XDS. *Acta Crystallogr* D *Biol Crystallogr* **66**, 125–132.

43. Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC & Zwart PH (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr* D *Biol Crystallogr* **66**, 213–221.

44. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC & Read RJ (2007) Phaser crystallographic software. *J Appl Crystallogr* **40**, 658–674.

45. Langer G, Cohen SX, Lamzin VS & Perrakis A (2008) Automated macromolecular model building for X-ray crystallography using ARP/ wARP version 7. *Nat Protoc* **3**, 1171–1179.

46. Emsley P, Lohkamp B, Scott WG & Cowtan K (2010) Features and development of Coot. *Acta Crystallogr* D *Biol Crystallogr* **66**, 486–501.

47. Afonine PV, Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M, Terwilliger TC, Urzhumtsev A, Zwart PH & Adams PD (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr* D *Biol Crystallogr* **68**, 352–367.

48. Chen VB, Arendall III WB, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS & Richardson DC (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr* D *Biol Crystallogr* **66**, 12–21.

49. Abagyan RA, Totrov MM & Kuznetsov DA (1994) ICM: A new method for protein modeling and design: applications to docking and structure prediction from the distorted

native conformation. *J Comp Chem* **15**, 488–506.

Supporting information

Additional supporting information may be found in the online version of this article at the publisher's web site:

Table S1. Primers sequences for loop exchange mutants

Table 1 Kinetic parameters of mannanases

Enzyme	Substrate	k_{cat} s^{-1}	Relative k_{cat}	K_m mM	Relative K_m	k_{cat}/K_m $s^{-1}\cdot mM^{-1}$	Relative k_{cat}/K_m
StMandC	GGM5	$7.0\times10^{-3} \pm 4.0\times10^{-4}$	1.0	7.61±1.44	1.0	9.2×10^{-4}	1.0
	M4	$1.3 \pm 4.7\times10^{-2}$	1.0	3.13±0.25	1.0	4.2×10^{-1}	1.0
	M5	$3.7 \pm 5.7\times10^{-1}$	1.0	1.15±0.36	1.0	3.2	1.0
	M6	2.8 ± 6.2	1.0	1.01±0.46	1.0	28	1.0
TfMandC	GGM5	$8.2\times10^{-4} \pm 1.0\times10^{-4}$	0.1	12.0±4.0	1.6	6.8×10^{-5}	0.07
	M4	$7.4\times10^{-1} \pm 4.8\times10^{-2}$	0.6	1.09±0.21	0.3	6.8×10^{-1}	1.6
	M5	48.1 ± 3.3	13.2	1.02±0.17	0.9	47	12.7
	M6	60 ± 11	2.1	0.88±0.36	0.9	68	2.4
StW281AdC	GGM5	$4.8\times10^{-3} \pm 3.0\times10^{-4}$	0.7	9.08±1.97	1.2	5.3×10^{-4}	0.6
	M4	$2.9\times10^{-1} \pm 3.7\times10^{-2}$	0.2	4.50±1.10	1.4	6.3×10^{-2}	0.2
	M5	$4.2 \pm 9.0\times10^{-1}$	1.1	2.58±1.00	2.2	1.6	0.4
	M6	20.1 ± 5.2	0.7	1.62±0.37	1.6	12	0.4
StD310AdC	GGM5	$2.3\times10^{-3} \pm 2.0\times10^{-4}$	0.3	10.3±2.9	1.4	2.3×10^{-4}	0.2
	M4	$1.5 \pm 1.7\times10^{-1}$	1.1	2.97±0.79	0.7	5.0×10^{-1}	1.2
	M5	$4.7 \pm 4.7\times10^{-1}$	1.3	1.19±0.32	1.0	3.9	1.1
	M6	30.2 ± 3.2	1.1	1.34±0.27	1.3	22	0.8

The relative k_{cat}, K_m, and k_{cat}/K_m values of StMandC were 1.0 toward each substrate.

Table 2 Data collection and refinement statistics

	StMandC	StE273AdC
Space group	$P2_12_12_1$	$P2_12_12_1$
Unit cell parameters (*a*, *b*, *c*; Å)	65.86, 100.85, 105.24	65.71, 100.71, 104.74
Resolution range (Å)	50–1.6 (1.7–1.6)	50–1.50 (1.59–1.50)
No. of unique reflections	90737 (14792)	109438 (15735)
R_{meas}	0.202 (0.913)	0.106 (0.587)
Completeness (%)	97.5 (99.3)	98.0 (88.2)
$\langle I/\sigma(I)\rangle$	9.96 (2.25)	13.11 (2.53)
Multiplicity	5.1 (4.9)	4.7 (3.5)
Refinement		
R_{work}	0.1601	0.1458
R_{free}	0.1938	0.1727
No. of protein atoms	4610	4644
No. of sugar atoms	0	191
No. of glycerol atoms	36	42
No. of water molecules	798	795
No. of ions	4 (Na^+)	5 (Ca^{2+})
Averaged B-factors (Å^2)		
Protein	9.6	11.2
Sugar	-	26.4
Glycerol	27.7	24.2
Water	25.4	26.8
Ion	16.6	18.3
r.m.s.d.[a] values from ideal		
Bond lengths (Å)	0.008	0.011
Bond angles (°)	1.188	1.398
Ramachandran plot analysis		
Favored region (%)	96.82	97.36
Allowed region (%)	3.01	2.31
Outlier region (%)	0.17	0.33

[a] r.m.s.d., root mean square deviation.

Figure legends

Figure 1. GGM5 hydrolysis by StMandC and TfMandC. (A) Schematic representation of GGM5 hydrolysis by mannanases. The parenthesis and numbers show the minus subsites in StMandC and TfMandC. The triangles show the cleavage site for StMandC (closed) and TfMandC (open). (B) HPAEC-PAD analysis of the hydrolysis products. (a) GGM5 and StMandC; (b) GGM5 and TfMandC; and (c) GGM5. (C) HPLC analysis of GGM5-ABEE hydrolysis products. (a) GGM5-ABEE and StMandC; (b) GGM5-ABEE and TfMandC; and (c) GGM5-ABEE.

Figure 2. Evaluation of the region involved in GGM5 hydrolysis activity. (A) Schematic representation of the structures of the parent and chimeric enzymes. Amino acid numbers and the length of StMan are shown above and inside the bar, respectively. GGM5 hydrolysis products are shown on the right side of the bar. (B) Alignment of the amino acid sequences responsible for GGM5 hydrolysis specificity of StMan and TfMan. The numbers represent each amino acid residue. (C) Schematic representation of the structures of the loop-substitution mutants. As shown in Fig. 2A, the hydrolysis products from GGM5 are shown on the right side of each bar. The “N.D.” represents no hydrolysis products detected.

Figure 3. HPAEC-PAD analysis of GGM5 hydrolysis products by mutated mannanases. GGM5 was hydrolyzed by chimeric enzymes (A) and loop-substitution mutants of StMandC and TfMandC (B).

Figure 4. Overall structure of StMandC and StE273AdC with M6. (A) The superimposed structure of StMandC and StE273AdC with M6. The peptide colors (from purple to red) correspond to the N-terminus to the C-terminus. L1-L8 in the figure shows the loop structures from loop1 to loop8, respectively. Mannose residues are shown as yellow and green sticks. (B) A superimposed structure of StMandC and StE273AdC. Amino acid

residues for mutagenesis in this study are shown. White and blue sticks indicate StMandC (free form) and StE273AdC (complex with substrate), respectively. (C) Mannose structure at subsites −2 to +1. Mannobiose at subsites −2 and −1 is shown by a green stick. Mannobiose at subsites −1 and +1 which distorted the mannose structure at subsite −1 is shown as a yellow stick. (D) Plus subsite binding residues in StMandC and StE273AdC. White and blue sticks indicate StMandC and StE273AdC, respectively. The numbers in the figure indicate subsites. σ_A-weighted F_o-F_c omit map is contoured at 3σ (C and D).

Figure 5. Structural comparison between StE273AdC and *T. fusca* KW3 mannanase (PDB code: **1BQC** and **3MAN**). The minus and plus subsites were compared (A) between StE273AdC and 3MAN, and (B) between StE273AdC and 1BQC, respectively. Green and dark blue sticks indicate StE273AdC and TfMandC. Amino acid residues of StE273AdC, which are related to substrate-binding, are shown in the figure, except for Val263 of 3MAN. Mannose residues from StE273AdC and 3MAN are shown as yellow and gray sticks. The numbers in the panels indicate the subsites.

Figure 6. Surface view and superimposed model of StE273AdC with *T. fusca* KW3 mannanase (PDB code: **1BQC**). Surface view of the superimposed model of StMandC (green) and 1BQC (orange) (A), StMandC (B), and 1BQC (C) from the upper (left panel) and side clefts from the reducing terminus (right panel). Mannose residues from StE273AdC are shown as a stick structure. The numbers in the figure indicate the subsites.

Figure 7. Hydrolysis products of M5 by StMandC and TfMandC. Hydrolysis products of M5 by StMandC and TfMandC were analyzed by HPAEC–PAD. Hydrolysis products were separated by an isocratic flow of 200 mM NaOH.

Figure 8. Primary structure alignment of bacterial mannanases. Alignment focused on loop7 and loop8 of bacterial mannanase: (A) similar to StMan and (B) similar to TfMan. The amino acid sequences (A) of StMan from Streptomyces thermolilacinus (BAK26781); WP_023591108, Streptomyces violaceusniger mannan endo-1,4-β-mannosidase; YP_004802777, Streptomyces sp. SirexAA-E cellulose-binding family protein; CAJ88324, Streptomyces ambofaciens ATCC 23877 putative secreted β-mannosidase; WP_018555858, Streptomyces sp. ATexAB-D23 beta-mannosidase; NP_733506, Streptomyces coelicolor A3(2) β-mannosidase; WP_005475300, Streptomyces bottropensis glycosylhydrolase; YP_003493383, Streptomyces scabiei 87.22 putative secreted glycosyl hydrolase; YP_007859663, Streptomyces sp. PAMC26508 putative secreted β-mannosidase; YP_004924956, Streptomyces flavogriseus ATCC 33331 glycoside hydrolase 5; WP_004003745, Streptomyces viridochromogenes Tue57 putative Secreted β-mannosidase; WP_020123328, Streptomyces canus β-mannosidase; WP_020140407, Streptomyces sp. 351MFTsu5.1 β-mannosidase. The amino acid sequence (B) of TfMan from Thermobifida fusca (AAZ54938); StManII from Streptomyces thermoluteus (BAM62868); SlMan from Streptomyces lividans 1326 (AAA26710); ADK91085, Streptomyces sp. S27 β-1,4-mannanase; WP_005474921, Streptomyces bottropensis glycosylhydrolase; CAA20610, Streptomyces coelicolor A3(2) β-mannosidase; CBG75158, Streptomyces scabiei 87.22 putative secreted glycosyl hydrolase; CCA60191, Streptomyces venezuelae ATCC 10712 Endo-1,4-β-xylanase A precursor; CCA60180, Streptomyces venezuelae ATCC 10712 Endo-1,4-β-xylanase A precursor; YP_008735110, Actinoplanes friuliensis DSM 7358 putative glycosyl hydrolase; YP_007953378, Actinoplanes sp. N902-109 secreted β-mannosidase; WP_020640659, Amycolatopsis balhimycina β-mannosidase; YP_003637895, Cellulomonas flavigena DSM 20109 glycoside hydrolase family protein; YP_004081647, Micromonospora sp. L5 β-mannanase-like protein ; WP_018788190, Micromonospora sp. CNB394 mannan endo-1,4-β-mannosidase; YP_004404351, Verrucosispora maris AB-18-032 glycoside

hydrolase family protein. The conserved regions and gaps are indicated by highlight and (-), respectively. The number of the alignment is indicated by the amino acid residues of each protein. “*1” shows the species possessed more than two kinds of GH5 mannanase genes: NP_733506 and WP_005474921 from S. coelicolor A3(2); YP_003493383 and CBG75158 from S. scabiei 87.22; WP_005475300 and WP_005474921 from S. bottropensis. “*2” shows the species possessed GH26 mannanase gene in addition to GH5 mannanase gene.

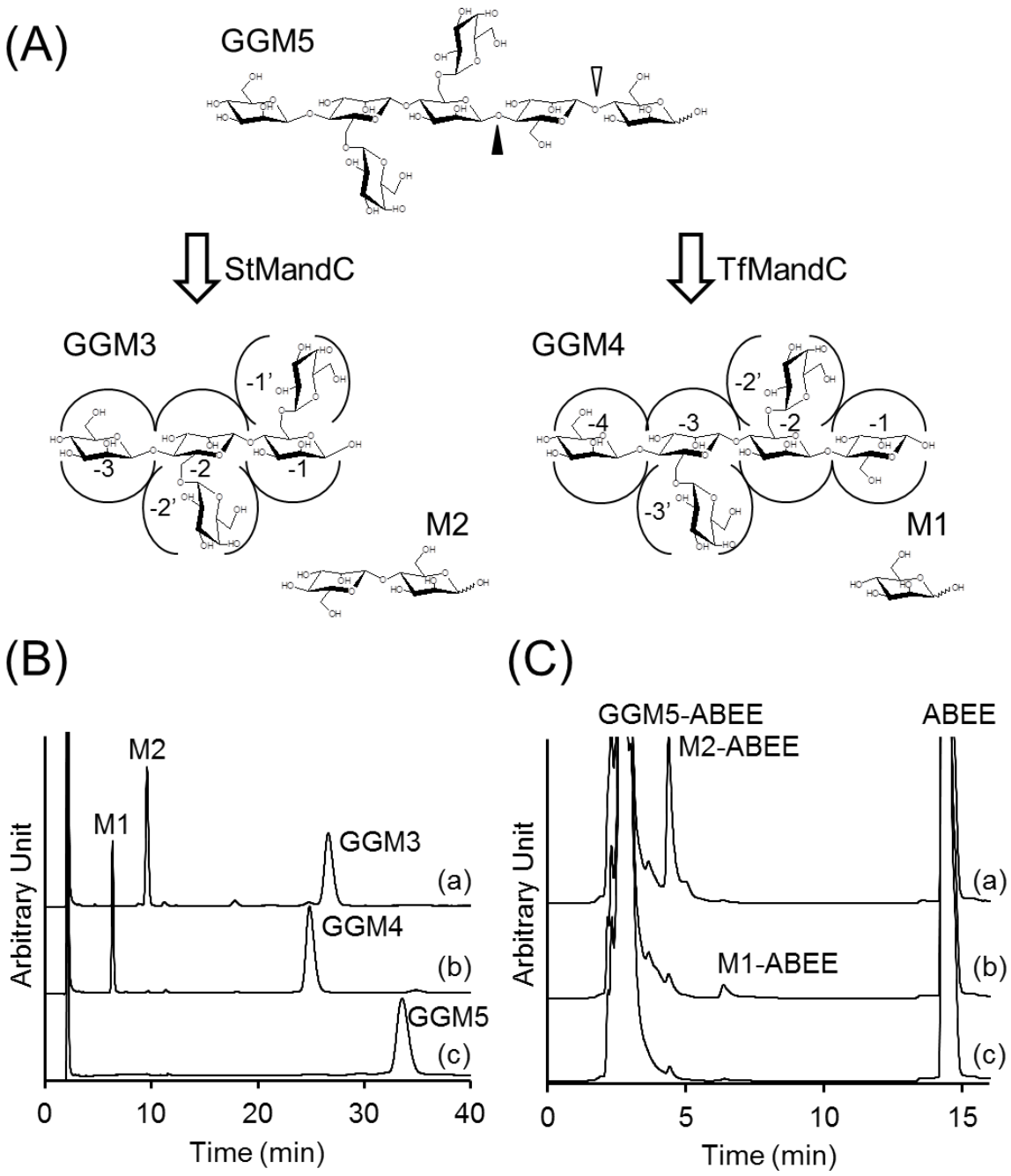


Figure 1 Kumagai et al

1

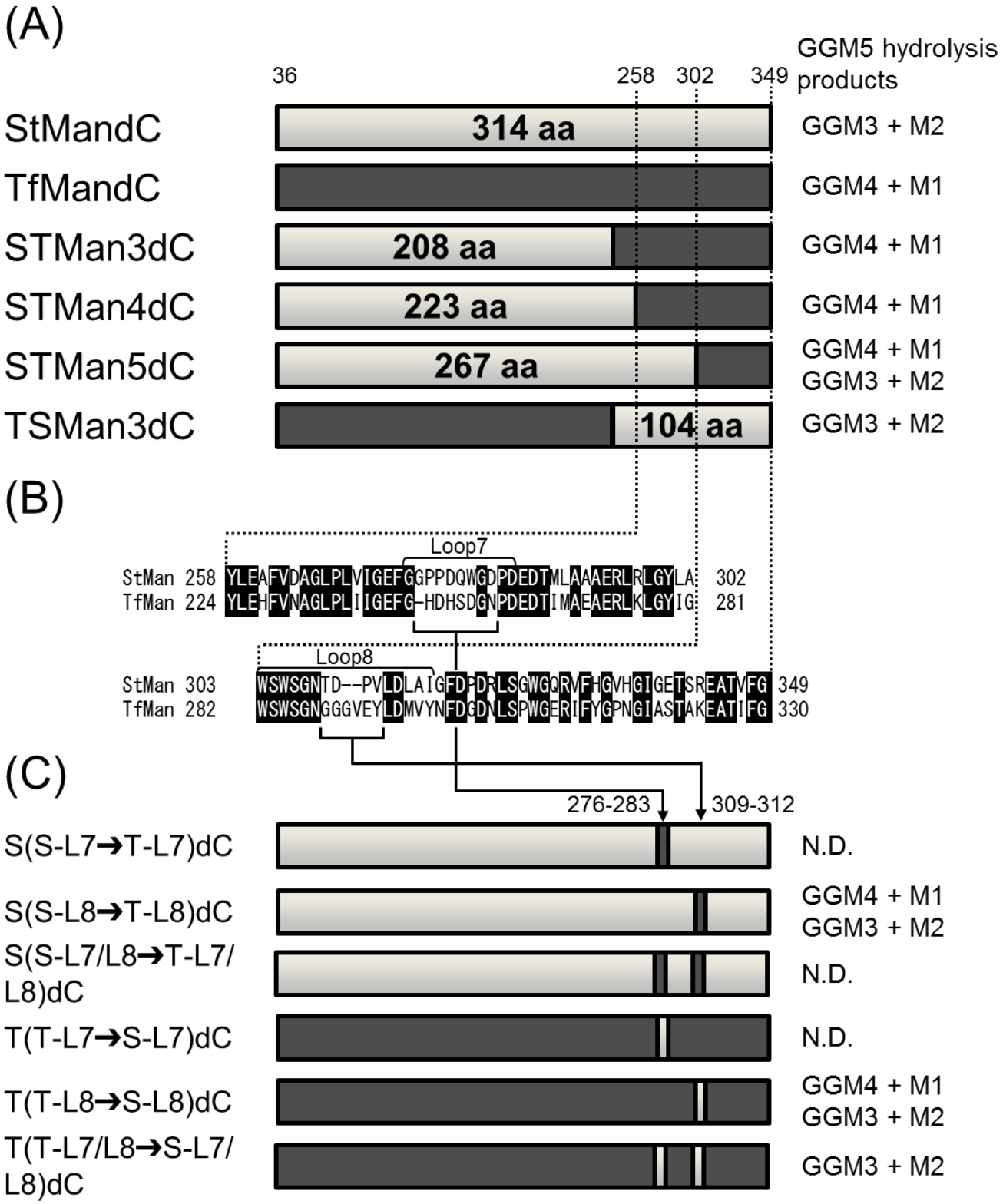


Figure 2 Kumagai et al

1

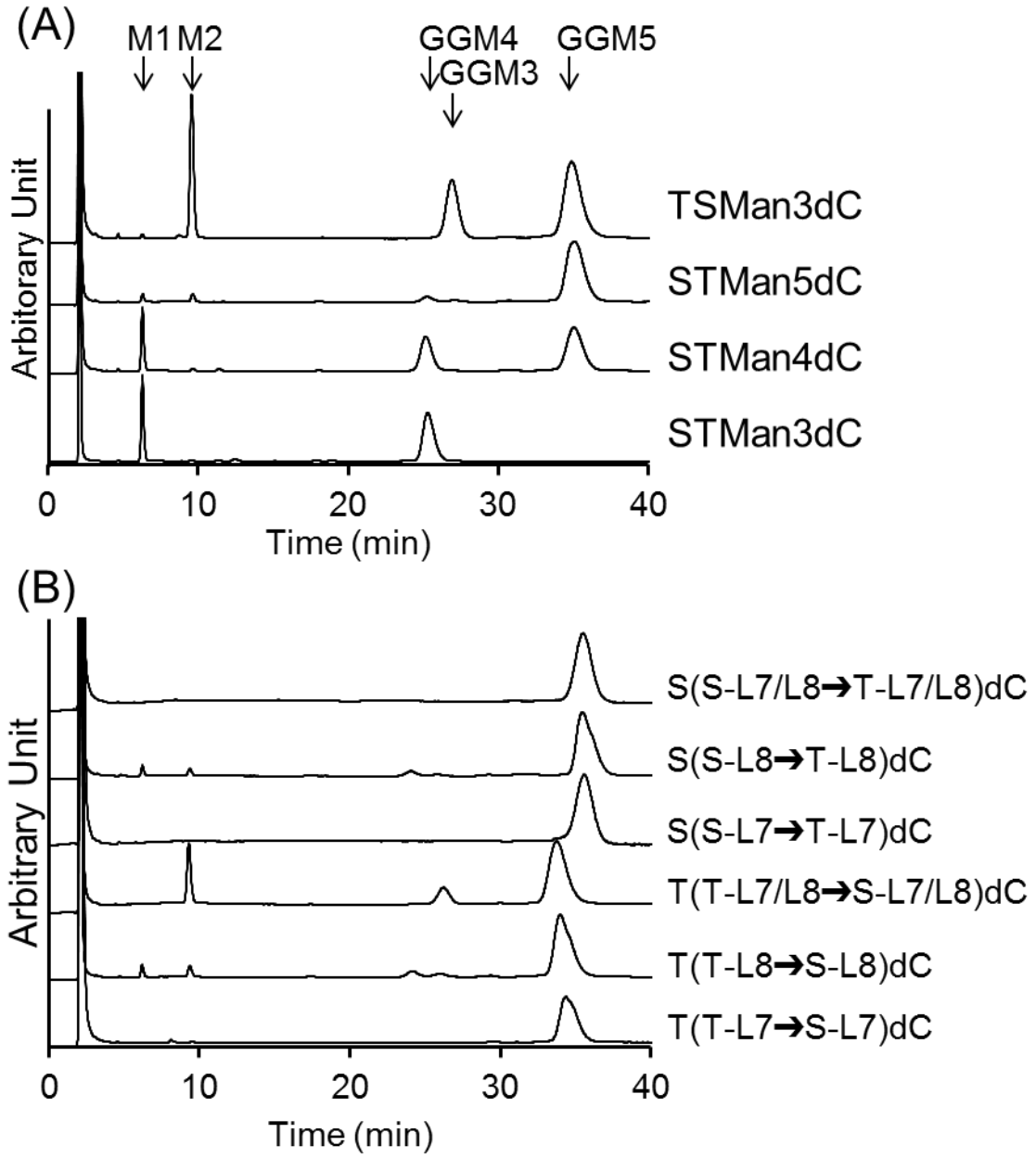


Figure 3 Kumagai et al

1

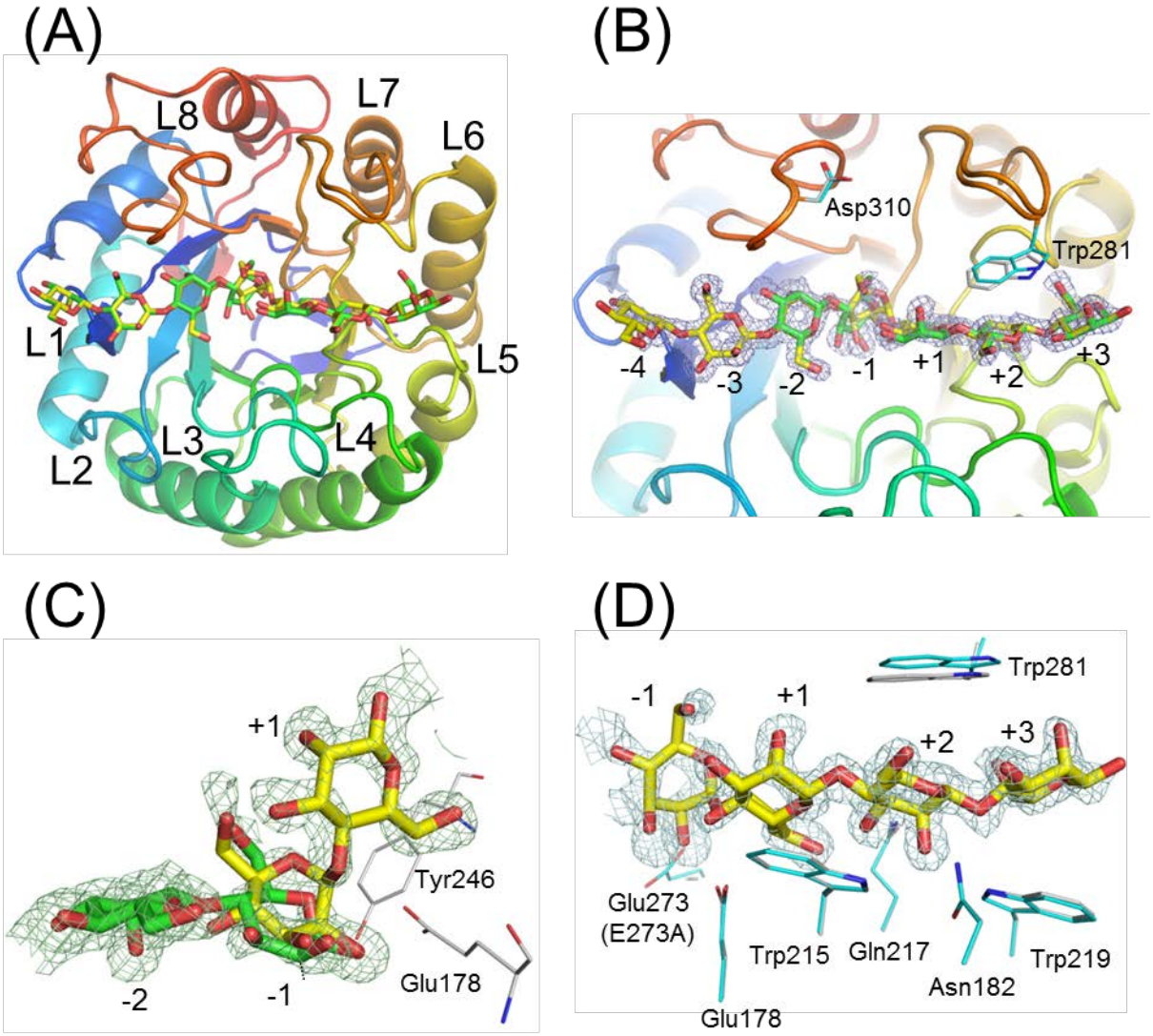
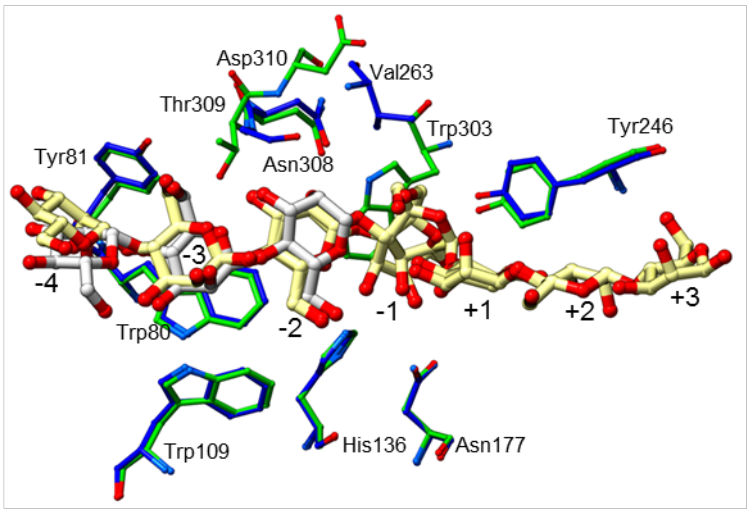


Figure 4 Kumagai et al

1

(A) StE273AdC + 3MAN



(B) StE273AdC + 1BQC

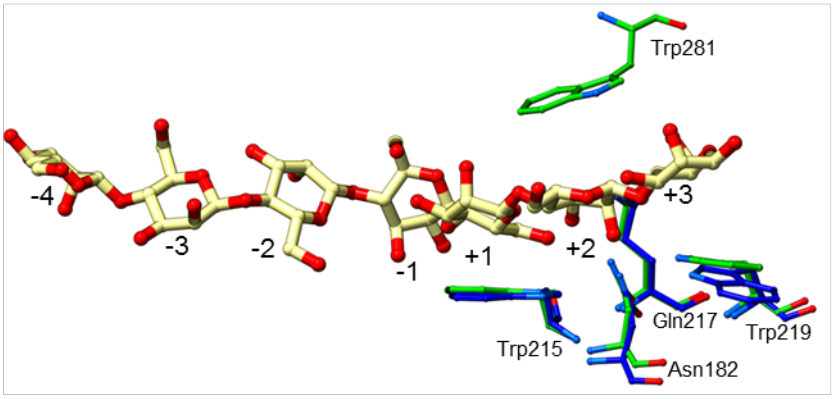


Figure 5 Kumagai et al

1

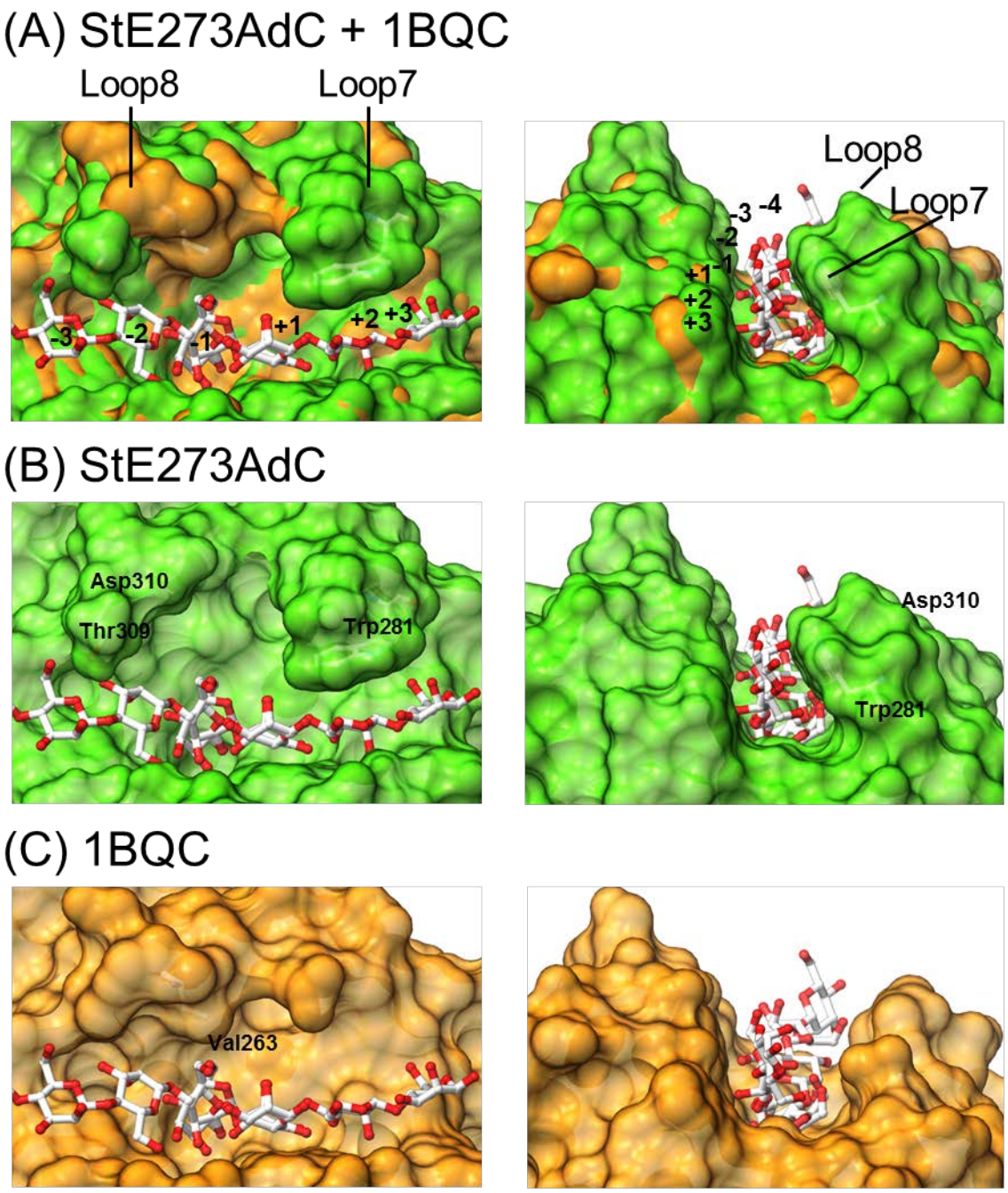


Figure 6 Kumagai et al

1

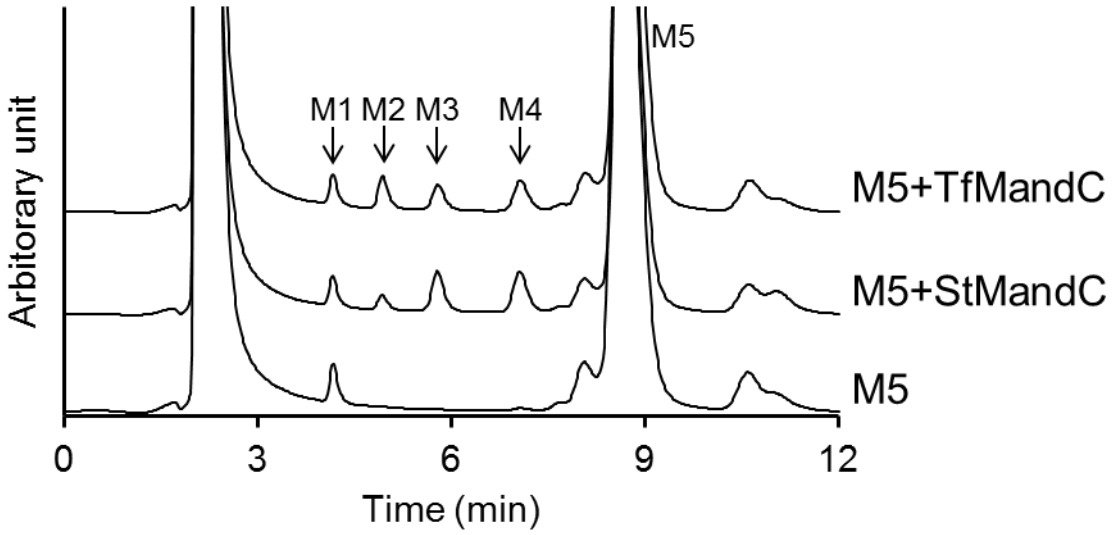


Figure 7 Kumagai et al

1

```
(A)                                Loop-7                          Loop-8
StMan              271 IGEFGGPPDCWGDPDEDTMLAAAERLRLGYLAWSWSGN--TDPVLDLAIG 318
WP_023591108       271 IGEFGGPPDCWGDPDEDTMLAAAERLRLGYLAWSWSGN--TDPVLDLAIG 318
YP_004802777       271 IGEFGGPADCYGDPDEDTMMATAEELGLGYLAWSWSGN--TDPVLDLVLD 318
CAJ88324           267 IGEFGGPADCWGDPDEDTMMAAAERLDLGYLAWSWSGN--TDPVLDLAID 314
WP_018555858       261 IGEFGGPADCYGDPDEDTMMADAEQLGLGWIAWSWSGN--TDPVLDLAID 308
NP_733506 *1       267 IGEFGGPADCYGDPDEDTMMATAEQLRLGYLAWSWSGN--TDPVLDLALD 314
WP_005475300 *1    267 IGEFGGPADCWGDPDEDTMMAAAEQLDLGYLAWSWSGN--TDPVLDLSIG 314
YP_003493383 *1,2  267 IGEFGGPADCWGDPDEDTMMAAAERLDLGYLAWSWSGN--TDPVLDLSIG 314
YP_007859663 *2    268 IGEFGGPADCWGDPDEDTMMATAEQLDLGYLAWSWSGN--TDPILDLAID 315
YP_004924956 *2    249 IGEFGGPADCWGDPDEDTMMATAEQLDLGYLAWSWSGN--TDPILDLAID 296
WP_004003745 *2    246 IGEFGGPPDCWGDPDEDTMMAAAQQLKLGYLAWSWSGN--TDPILDLAID 293
WP_020123328 *2    240 IGEFGGPADCWGDPDEDTMMATAEQLHLGYLAWSWSGN--TDTILDLVLD 287
WP_020140407 *2    241 IGEFGGPADCWGDPDEDTMMATAQRLGLGYLAWSWSGN--TDPSLDLVLG 288
(B)
TfMan              251 IGEFG-HDHSDGNPDEDTIMAEAERLKLGYIGWSWSGNGGGVEYLDMVYN 299
StManII            246 VGEFG-HDHSDGNPDEDAILSVTRQLGIGYLGWSWSGNGGGVEYLDMVEN 304
SlMan              257 VGEFG-DQHSDGNPDEDAIMATAQSLGVGYLGWSWSGNGGGVEYLDMVNG 305
ADK91085           259 VGEFG-HNHGDGDPDENAIMATAQSLRVGYLGWSWSGNGGGVEYLDMVNG 307
WP_005474921*1     261 VGEFG-HDHSDGNPDEDAILATAQRLGLGYLGWSWSGNGGGVEYLDMVTG 309
CAA20610*1         258 VGEFG-DQHSDGNPDEDAIMATAQSLGVGYLGWSWSGNGGGVEYLDMVNG 306
CBG75158*1, 2      333 IGEFG-HEHSDGNPDEDAILAAAQRLGLGYLGWSWSGNGGGVEYLDLVTG 381
CCA60191*2         258 VGEFG-DNHSDGNPDENAIMATTQSLRVGYLGWSWSGNGSGVEYLDMVTG 306
CCA60180*2         253 VGEFG-YDHSDGNPDEDAIMATARRLDLGYMGWSWSGNGGGVEYLDLATG 301
YP_008735110*2     251 VGEFG-HNHSDGNPDEDTILATAQAQGIGYLGWSWSGNGGGVEYLDLVTN 299
YP_007953378*2     244 VGEFG-FDHSDGNPDEDTIMATAQRLGIGYLGWSWSGNGGGVEYLDMVTG 292
WP_020640659*2     462 VGEFG-NMHTDGNPDEDTIMAQAQARGLGYLGWSWSGNSSDVAYLDMTNN 510
YP_003637895*2     256 IGEFG-IDHSDGDPDEATIMREATERGIGYYGWSWSGNSGGVEYLDMVTG 305
YP_004081647*2     249 VGEFG-FNHSDGNPDEDAIMAYAQANGIGYLGWSWSGNGGGVEYLDMTTA 297
WP_018788190*2     249 VGEFG-FNHSDGNPDEDAIMAYAQANGIGYLGWSWSGNGGGVEYLDMTTG 297
YP_004404351*2     249 VGEFG-HYHSDGDPDEDAILSYTQANGIGWLGWSWSGNGGGVEYLDMATN 297
```

Figure 8 Kumagai et al

Table S1. Primers sequences for loop exchange mutants

Primer	Restriction site	Sequence[a]
StFw	*Nde*I	5'-CATATGCGGACCGCCCCGCCCG-3'
StL7Re1	*Eco*RI	5'-GAATTCCCCGATGACCAGCGGCAGC-3'
StL7Fw1		5'-CACGACCACTCCGACGGCAACCCGGACGAGGACACGATGC-3'
StL7Fw2	*Eco*RI	5'-CCCGAATTCCACGACCACTCCGACGGCAACC-3'
StL8Re1		5'-AGGTACTCGACCCCGCCGCCGTTGCCGCTCCACGACCAGG-3'
StL8Re2	*Bgl*II	5'-TTCAGATCTAGGTACTCGACCCCGCCGCCGTTGC-3'
StL8Fw1		5'-ACGGCGGCGGGGTCGAGTACCTCGACCTGGCGATCGGGTT-3'
StL8Fw2	*Bgl*II	5'-AATAGATCTGGCGATCGGGTTCGACCCCGAC-3'
StCDRe	*Hin*dIII	5'-AAGCTTTCAGGTGTCGCCGGGGGTTTCCCC-3'
TfFw	*Nde*I	5'-CATATGGCCACCGGGCTCCACGTCAAG-3'
TfL7Re1	*Eco*RI	5'-GAATTCGCCGATGATGAGCGGCAGG-3’
TfL7Fw1		5'-CGCCCGACCAGTGGGGCGACCCCGACGAGGACACGATCAT-3'
TfL7Fw2	*Eco*RI	5'-AGGGAATTCGGCGGGCCGCCCGACCAGTGGGGCG-3'
TfL8Re1		5'-ATGTCGAGGACGGGGTCGGTGTTGCCGCTCCACGACCAGC-3'
TfL8Re2	*Bgl*II	5'-CAGATCTAGGACGGGGTCGGTGTTCCCG-3'
TfL8Fw1		5'-GCGGCAACACCGACCCCGTCCTCGACATGGTGTACAACTT-3'
TfL8Fw2	*Bgl*II	5'-AATAGATCTGTGTACAACTTCGACGGCGAC-3'
TfCDRe	*Hin*dIII	5'-AAGCTTTCACGGGCCCGGCTGGGAGCC-3'

[a]Underline shows the restriction site

1