Available online at www.sciencedirect.com
SCIENCE DIRECT
Biochemical and Biophysical Research Communications 344 (2006) 253–262
BBRC
www.elsevier.com/locate/ybbrc

# Substrate recognition by unsaturated glucuronyl hydrolase from *Bacillus* sp. GL1 ☆

Takafumi Itoh [a], Wataru Hashimoto [b], Bunzo Mikami [a,1], Kousaku Murata [b,*,1]

[a] *Division of Agronomy and Horticultural Science, Graduate School of Agriculture, Kyoto University, Gokasho, Uji, Kyoto 611-0011, Japan*
[b] *Division of Food Science and Biotechnology, Graduate School of Agriculture, Kyoto University, Gokasho, Uji, Kyoto 611-0011, Japan*

Received 25 February 2006
Available online 30 March 2006

**Abstract**

Bacterial unsaturated glucuronyl hydrolases (UGLs) together with polysaccharide lyases are responsible for the complete depolymerization of mammalian extracellular matrix glycosaminoglycans. UGL acts on various oligosaccharides containing unsaturated glucuronic acid (ΔGlcA) at the nonreducing terminus and releases ΔGlcA through hydrolysis. In this study, we demonstrate the substrate recognition mechanism of the UGL of *Bacillus* sp. GL1 by determining the X-ray crystallographic structure of its substrate-enzyme complexes. The tetrasaccharide-enzyme complex demonstrated that at least four subsites are present in the active pocket. Although several amino acid residues are crucial for substrate binding, the enzyme strongly recognizes ΔGlcA at subsite −1 through the formation of hydrogen bonds and stacking interactions, and prefers *N*-acetyl-D-galactosamine and glucose rather than *N*-acetyl-D-glucosamine as a residue accommodated in subsite +1, due to the steric hindrance.


*Keywords:* Biofilm; Chondroitin; Gellan; Glycoside hydrolase family GH-88; Hyaluronan; Substrate recognition; Unsaturated glucuronyl hydrolase; X-ray crystallography

Glycosaminoglycans such as chondroitin, hyaluronan, and heparin are negatively charged polysaccharides with a repeating disaccharide unit that consists of a uronic acid (glucuronic acid or iduronic acid) and an amino sugar (glucosamine or galactosamine) [1–3]. Chondroitin consists of D-glucuronic acid (GlcA) and *N*-acetyl-D-galactosamine (GalNAc) along with a sulfate group that is attached at position 4 and/or 6. Hyaluronan is composed of GlcA and *N*-acetyl-D-glucosamine (GlcNAc). These glycosaminoglycans are major components of the mammalian extracellular matrix and are responsible for cell-to-cell association; they are widely present in human tissues such as the eye, brain, liver, skin, and blood [4].

The UGL of *Bacillus* sp. GL1 is a member of the glycoside hydrolase family GH-88 in the CAZy database (Henrissat, B., Coutinho, P., and Deleury, E., 1999; http://afmb.cnrs-mrs.fr/~cazy/CAZY/index.html). The enzyme acts on oligosaccharides having ΔGlcA at the nonreducing terminus, e.g., ΔGlcA-GalNAc produced from chondroitin by chondroitin lyase [5,6], ΔGlcA-GlcNAc produced from hyaluronan by hyaluronate lyase [7,8], ΔGlcA-mannose (Man)-glucose (Glc) produced from a bacterial polysaccharide xanthan by xanthan lyase [9–11], and ΔGlcA-Glc-rhamnose (Rha)-Glc produced from a bacterial gellan by gellan lyase (Figs. 1A–D) [12–14]. The enzyme produces ΔGlcA (4-deoxy-L-threo-5-hexosulose-uronate) from these substrates. The resultant product ΔGlcA is converted into the chain form because the ring form of ΔGlcA is unstable due to the keto-enol equilibrium (Fig. 1D) [15,16].

GH-88 family proteins are found in bacteria and fungi. UGL and its gene were first identified in *Bacillus* sp. GL1, and along with xanthan and gellan lyases, they were

☆ Protein Data Bank accession numbers. The coordinates of UGL are in the RCSB Protein Data Bank (D88N/ΔGlcA-Glc-Rha-Glc, 2FV0; D88N/ΔGlcA-GlcNAc, 2FV1; and hex_wild type/apo structure, 2FUZ).

* Corresponding author. Fax: +81 774 38 3766.
*E-mail address:* kmurata@kais.kyoto-u.ac.jp (K. Murata).

[1] These authors have contributed equally to this work.

0006-291X/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.bbrc.2006.03.141

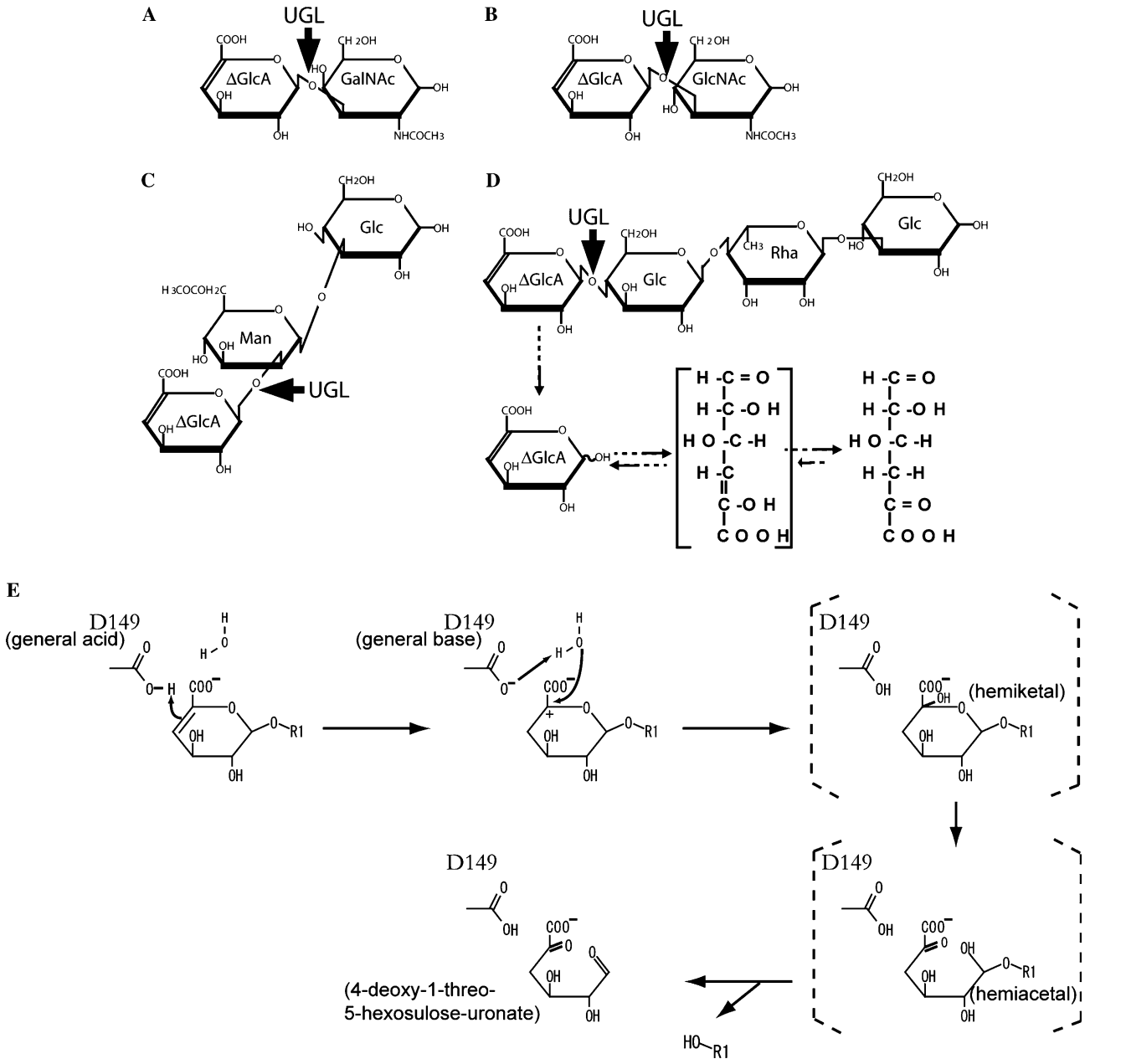


Fig. 1. (A) Unsaturated chondroitin disaccharide (ΔGlcA-GalNAc), (B) unsaturated hyaluronate disaccharide (ΔGlcA-GlcNAc), (C) unsaturated xanthan trisaccharide (ΔGlcA-Man-Glc), and (D) unsaturated gellan tetrasaccharide (ΔGlcA-Glc-Rha-Glc). Arrows indicate the cleavage site for UGL. Dashed-line arrows denote the degradation pathway for polysaccharides. (E) The proposed catalytic mechanism for UGL (Itoh et al., submitted for publication).

considered to be involved in the degradation of bacterial biofilms (*Xanthomonas* xanthan and *Sphingomonas* gellan) (Figs. 1C and D) [17,18]. On the other hand, other hypothetical proteins of the GH-88 family are found in pathogenic bacteria such as bacteroides, enterococci, mycoplasma, streptococci, and vibrio. These pathogenic bacteria invade host cells and degrade extracellular matrix glycosaminoglycans by the action of chondroitin and hyaluronate lyases [16,19]. These bacteria completely degrade glycosaminoglycans by using polysaccharide lyases and GH-88 family proteins (Hashimoto et al. manuscript in preparation).

The structural analysis of UGL is indispensable for elucidating the molecular mechanisms underlying the recognition and degradation of glycosaminoglycans and/or biofilms, and it provides us with valuable information on bacterial infectious process. We have determined the three-dimensional structure of UGL (hex_wild type/glycine, PDB Accession No. 1VD5) by the multiple isomorphous replacement method using hexagonal crystals and have demonstrated that the enzyme has an $\alpha_6/\alpha_6$-barrel topology because of which it is categorized into the $\alpha_6/\alpha_6$ toroidal fold family. One side of the $\alpha_6/\alpha_6$-barrel structure comprises long loops and contributes to the formation of a

deep pocket, which is an active site. However, this structure (hex_wild type/glycine) revealed that one glycine and two dithiothreitol (DTT) molecules, which are derived from the crystallization solution, are bound into the active site [20]. As a result, determination of the structure of the enzyme-substrate complex becomes difficult.

Thus, we attempted to crystallize the enzyme under other conditions in the absence of glycine and DTT and obtained its orthorhombic crystals. Structural analysis of the orthorhombic crystals (wild type and its inactive mutant D88N) revealed that the enzyme is in a ligand-free form (ortho_wild type/apo, PDB Accession No. 2D5J; D88N/apo, PDB Accession No. 2AHF). This crystal enabled us to analyze the mechanisms underlying enzyme catalysis and substrate recognition. More recently, we analyzed the enzyme's catalytic mechanism by using the crystal structure of UGL complexed with chondroitin disaccharide (D88N/GlcA-GalNAc, PDB Accession No. 2AHG) (Itoh et al., submitted for publication). The substrate recognition by UGL, however, remains to be clarified. This paper elucidates the crystal structures of the inactive mutant UGL in complexes with two different substrates (unsaturated hyaluronan disaccharide and gellan tetrasaccharide) and clarifies its substrate recognition mechanism.

## Materials and methods

*Enzyme assay.* UGL reactions were conducted at 30 °C as follows [20]. The reaction mixture (0.5 ml) consisted of 50 mM sodium phosphate (pH 6.5), 10–500 μM of the substrate (ΔGlcA-Glc-Rha-Glc), and 10 nM of the enzyme. Enzyme activity was measured by monitoring the decrease in absorbance at 235 nm; this corresponded to the loss of the C=C double bond of the substrate because the released ΔGlcA is converted to α-keto acid through the loss of the double bond (Fig. 1D) [15,16]. Enzyme concentration was determined by UV spectrophotometry using the theoretical molar extinction coefficient $\varepsilon_{280} = 99{,}570\ M^{-1}\,cm^{-1}$. The unsaturated gellan tetrasaccharide (ΔGlcA-Glc-Rha-Glc) used as the UGL substrate was prepared, as described previously [15]. The parameters $k_{cat}$ and $K_m$ were determined by nonlinear fitting to the Michaelis–Menten equation. The enzyme reaction was also performed in the presence of various concentrations (0–100 mM) of an inhibitor (glycine). The type of inhibition was determined by the Dixon plot ($[I]_0$ vs. $[E]_0/V_0$) at two substrate concentrations ($[S]_0 = 56$ and 115 μM).

*Crystallization and X-ray diffraction.* UGL and its mutant D88N from *Bacillus* sp. GL1 were overexpressed in *Escherichia coli*, purified, and crystallized by the vapor diffusion method, as described previously [20–22]. Orthorhombic crystals of D88N were placed into the substrate solution containing 400 mM ΔGlcA-GlcNAc (Seikagaku Corporation, Tokyo, Japan) or 400 mM ΔGlcA-Glc-Rha-Glc in a mixture of 25% (w/v) polyethylene glycol 10,000 and 0.15 M Tris–HCl (pH 7.5); the crystals were incubated for 10 min at 20 °C. Subsequently, they were placed in a cold nitrogen gas stream at −173 °C. X-ray diffraction images of the crystals were collected using a MarCCD detector (Mar USA, Illinois, USA) at the BL-41XU station of SPring-8 (Hyogo, Japan) or a Rigaku Jupiter 210 CCD detector (Tokyo, Japan) at the BL-38B1 station of SPring-8 with synchrotron radiation at a wavelength of 1.00 Å. The images were processed using the DENZO and SCALEPACK softwares [23] to a resolution of 1.73–1.90 Å (Table 1).

*Structure determination and refinement.* The structures were determined by the molecular replacement method and refined using CNS program ver. 1.1 [24]; the previously reported UGL structure (PDB Accession No. 1VD5) was used as the reference model. Several rounds of positional and *B*-factor refinement were performed. Following this, manual model building was attempted using the TURBO-FRODO program (AFMB-CNRS, Marseille, France) to improve the model by increasing the resolution of the data from 1.73 to 1.90 Å with the CNS program (Table 1). The stereo quality of the model was assessed using the PROCHECK [25]

Table 1
Synchrotron radiation data collection and refinement statistics

| | D88N/ΔGlcA-Glc-Rha-Glc | D88N/ΔGlcA-GlcNAc | Hex_wild type/apo |
|---|---|---|---|
| Space group | $P2_12_12_1$ | $P2_12_12_1$ | $P6_522$ |
| Unit cell parameters (Å) | $a = 87.7$, $b = 93.5$, $c = 95.4$ | $a = 87.7$, $b = 94.9$, $c = 95.4$ | $a = b = 102.8$, $c = 223.2$ |
| *Data collection* | | | |
| Resolution limit (last shell)[a] (Å) | 50.0–1.90 (1.97–1.90) | 50.0–1.73 (1.79–1.73) | 50.0–1.80 (1.86–1.80) |
| Measured reflections | 284,976 | 429,344 | 646,399 |
| Unique reflections (last shell) | 62,121 (6030) | 79,850 (6526) | 63,925 (6265) |
| Redundancy (last shell) | 4.6 (4.0) | 5.4 (4.2) | 10.1 (6.2) |
| Completeness ($\|I\| > \sigma\|I\|$) (last shell) (%) | 99.2 (97.6) | 95.3 (78.9) | 98.2 (98.4) |
| $R_{merge}$ (%)[b] | 5.7 (22.4) | 7.1 (33.3) | 8.7 (30.8) |
| *Refinement* | | | |
| Final model | 377 × 2 residues, 692 water, ΔGlcA-Glc-Rha-Glc | 377 × 2 residues, 789 water, 2 ΔGlcA-GlcNAc | 377 residues, 396 water, 9 MPD |
| Resolution limit (last shell) (Å) | 15.0–1.90 (1.97–1.90) | 15.0–1.73 (1.79–1.73) | 15.0–1.80 (1.86–1.80) |
| Used reflections (last shell) | 60,367 (5147) | 79,669 (6567) | 64,105 (6291) |
| Completeness ($\|F\| > \sigma\|F\|$) (last shell) (%) | 99.0 (83.3) | 95.4 (79.6) | 98.4 (98.5) |
| Average *B*-factor (Å$^2$) | 22.6 | 19.9 | 29.1 |
| *R*-factor (last shell) (%)[c] | 19.8 (24.4) | 17.9 (26.5) | 16.7 (24.6) |
| $R_{free}$ (last shell) (%)[d] | 25.4 (31.2) | 21.1 (29.3) | 18.8 (28.8) |
| *Root-mean-square deviations* | | | |
| Bond (Å) | 0.005 | 0.005 | 0.005 |
| Angle (deg) | 1.20 | 1.20 | 1.18 |

[a] Data in highest resolution shells are given in parentheses.
[b] $R_{merge} = \sum|I_i - \langle I\rangle|\sum I_i \times 100$, where $I_i$ is the intensity of individual reflection and $\langle I\rangle$ is the mean intensity of all reflections.
[c] $R_{factor} = \sum|F_o - F_c|\sum|F_o \times 100$, where $F_o$ is the observed structure factor and $F_c$ is the calculated structure factor.
[d] $R_{free}$ was calculated from 10% of reflections that were randomly selected, as defined by the CNS.

and WHATCHECK [26] programs. Structural alignment was carried out by superimposing with the aid of a fitting program in TURBO-FRODO. Ribbon plots and surface models were prepared using the MOLSCRIPT [27], RASTER3D [28], and GRASP [29] programs.

## Results

### *Inhibition by glycine*

The hex_wild type/glycine structure, which was the first UGL structure that was determined (PDB Accession No. 1VD5), contains one glycine molecule derived from the crystallization solution, as described previously [20]. Based on the enzyme-substrate complex structures that are described later, the location of this glycine was found to correspond to that of the carboxyl group of ΔGlcA at the nonreducing terminus of the substrate (Fig. 2A). Next, we examined the effect of glycine on UGL activity. Enzyme activity was inhibited by glycine in a dose-dependent manner (Fig. 2B). The results obtained with the double reciprocal plot revealed that glycine competes with the substrate to bind at subsite −1, which is located at the bottom of the active site, and has a $K_i$ value of 6.5 mM (Fig. 2B, inset graph).

### *Structure determination*

UGL is a monomeric enzyme with a molecular mass of approximately 43 kDa (377 amino acid residues) [15]. Other UGL crystals with another space group ($P2_12_12_1$) have been reported recently (Itoh et al., submitted for publication). However, the substrate recognition mechanism of UGL is still unclear because of lack of information on structures that are complexed with various substrates. In this study, we analyzed two orthorhombic crystal complex structures (D88N/ΔGlcA-Glc-Rha-Glc and D88N/ΔGlcA-GlcNAc) by using different substrates (Figs. 3A and B). Results of data collection and refined models are summarized in Table 1.

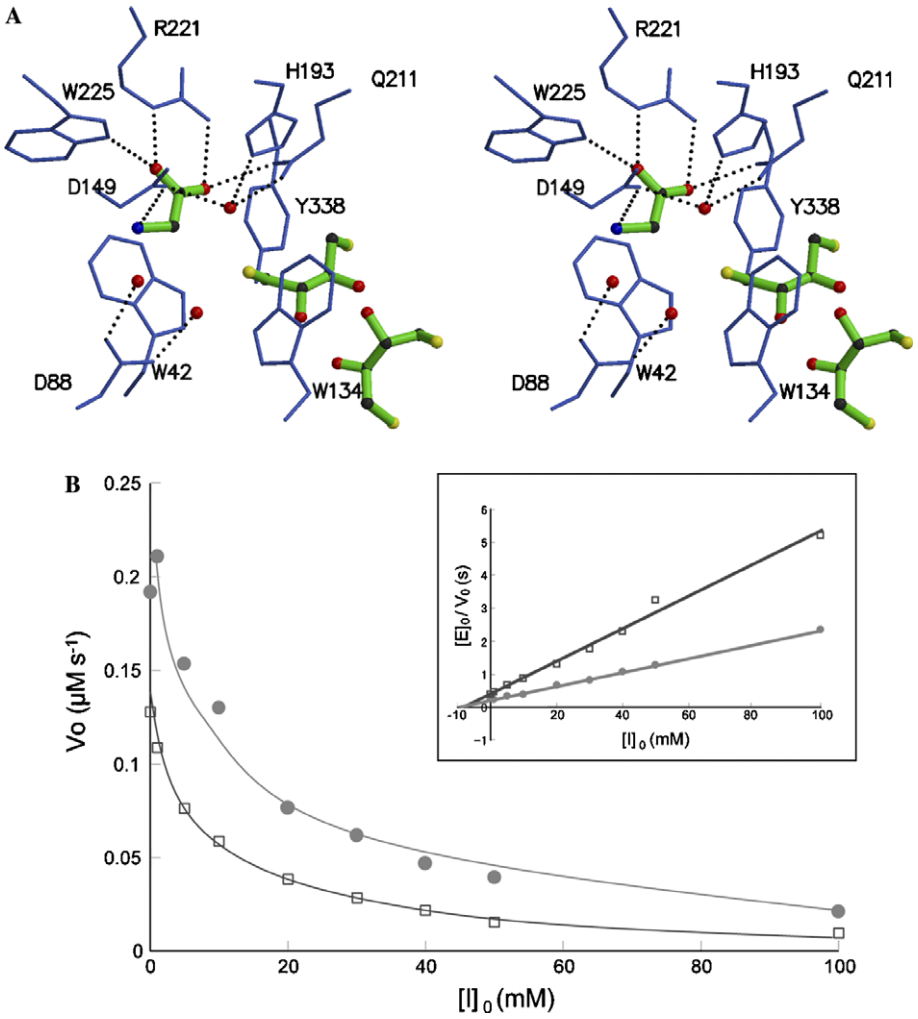


Fig. 2. (A) The active site of hex_wild type/glycine (PDB Accession No. 1VD5) [20]. The amino acid residues are indicated by blue bond models. Glycine and DTT molecules are also represented by green bond models. Water molecules are represented using red ball models. Hydrogen bonds are shown as dashed lines. (B) Inhibition of UGL activity by glycine. The inset graph shows double reciprocal plots of the inhibition kinetics of UGL by glycine. The circles represent experiments carried out with a substrate concentration $[S]_0$ of 115 μM, while squares represent those with $[S]_0$ of 56 μM.

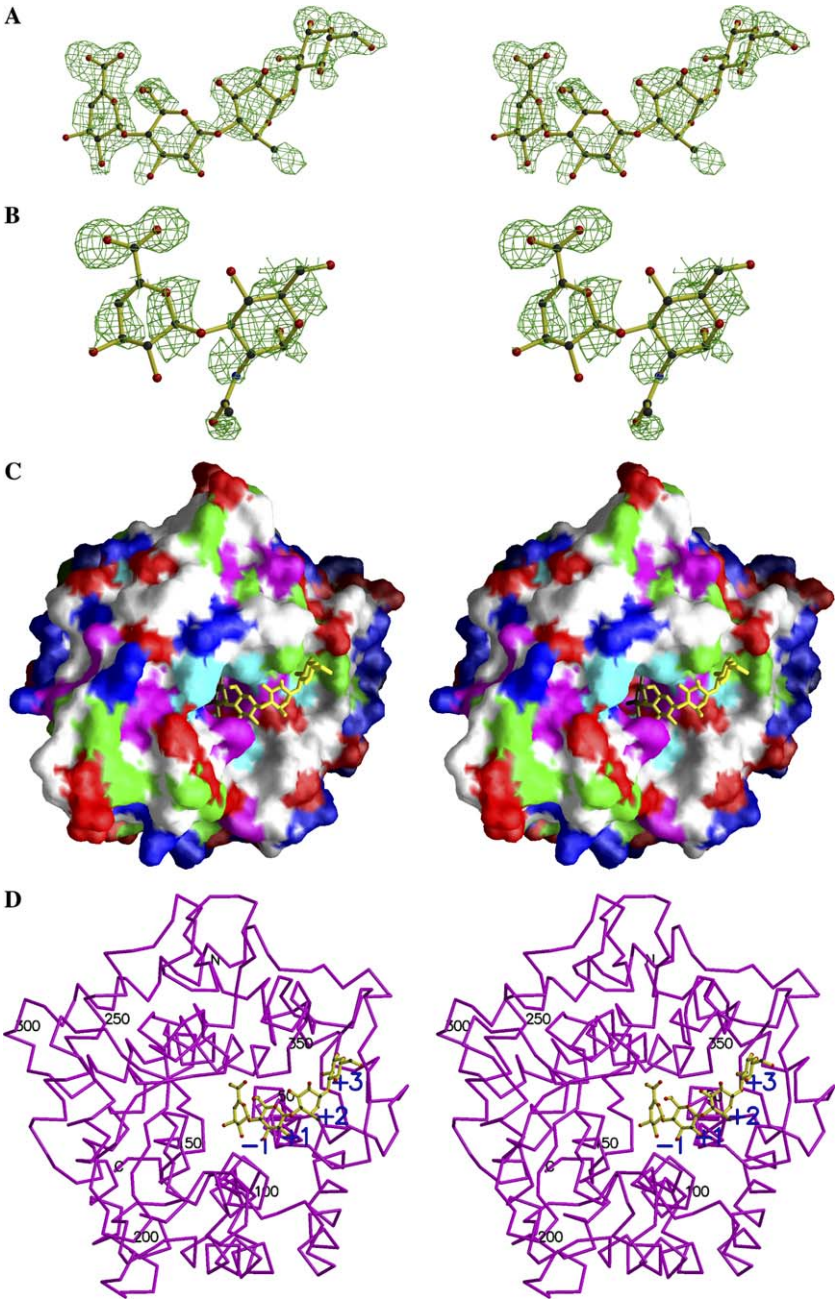


Fig. 3. The structure of UGL complexed with a substrate. Electron density of the ΔGlcA-Glc-Rha-Glc (A) or ΔGlcA-GlcNAc (B) substrates (yellow bond model) in the $2F_o - F_c$ map (green) is shown countered at the 1.0 (A) or 0.7 (B) sigma level. (C) ΔGlcA-Glc-Rha-Glc bound in the pocket of the enzyme (yellow bond model). Amino acid residues are indicated in magenta (Trp, Phe, and Tyr), green (Ala, Val, Leu, Ile, Pro, and Met), cyan (His), red (Asp and Glu), and blue (Lys and Arg). (D) Overall structure of UGL (Cα tracing) bound with the tetrasaccharide substrate (yellow bond model). The numeric values indicate the amino acid residues or subsites.

The hexagonal and orthorhombic crystals are similar in their overall structures, with the exception of their active sites. The details of their overall structures have been described previously [20]. The enzyme molecule has the $\alpha_6/\alpha_6$-barrel structure that is composed of 12 long α-helices, three antiparallel β-sheets consisting of 2–3 strands, and some loops. The active site of the enzyme is situated at the bottom of the large pocket that is surrounded by aromatic amino acid residues; the pocket is formed by the long loops and the inner helices of the barrel (Figs. 3C and D).

### *ΔGlcA-Glc-Rha-Glc bound at the active site*

The unsaturated gellan tetrasaccharide (ΔGlcA-Glc-Rha-Glc) was bound along the wall leading to the active site and fitted into a concavity of the wall (Figs. 3C, D, and 4A). UGL strongly recognizes ΔGlcA at subsite −1 rather than the saccharides at subsites +1 to +3. Contrary to most glycosidase-substrate complexes, no direct interaction was observed between UGL and the glycosidic bond to be cleaved. ΔGlcA was stabilized by several tight hydrogen

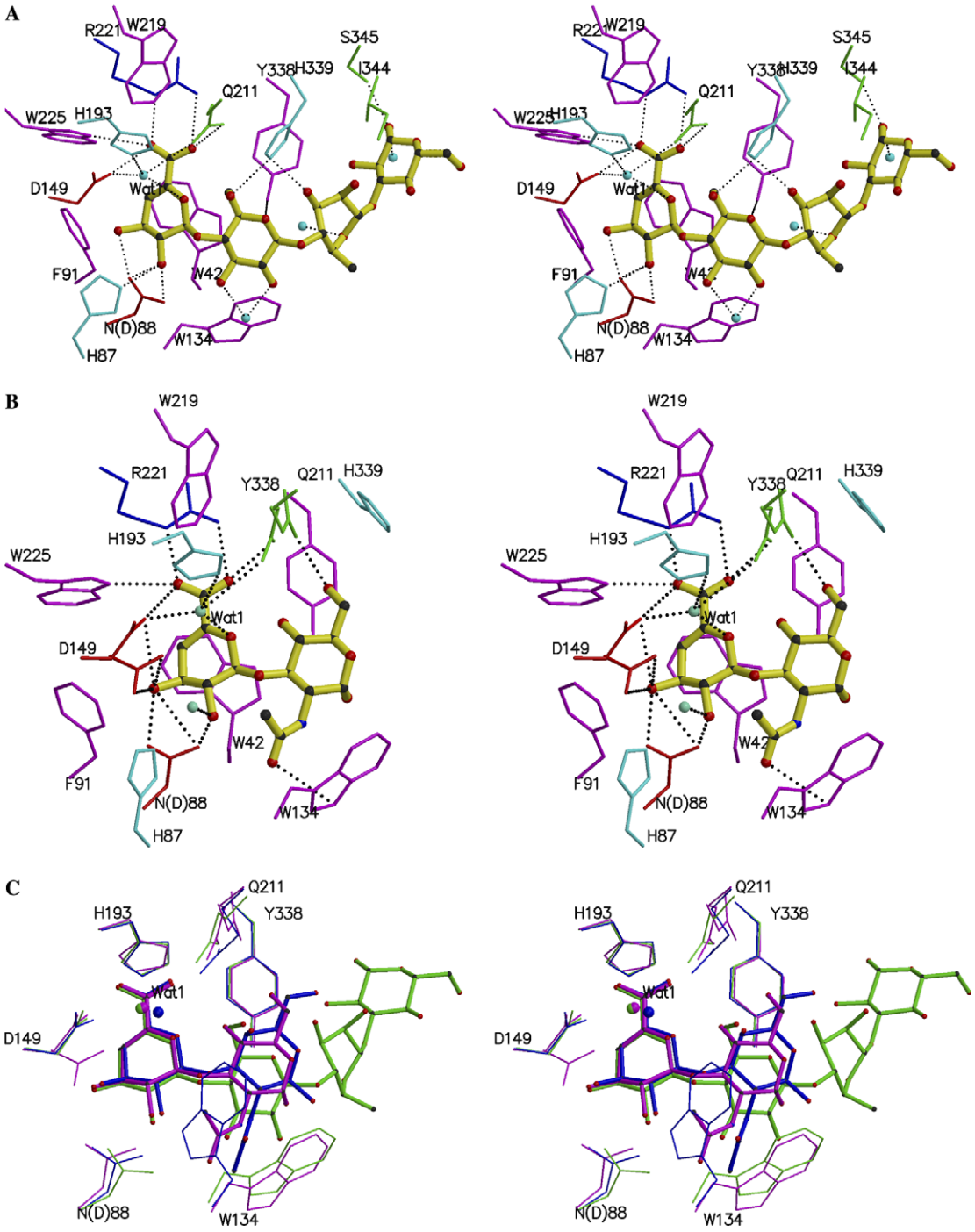


Fig. 4. Structure of UGL bound to the substrate at the active site (A) D88N/ΔGlcA-Glc-Rha-Glc and (B) D88N/ΔGlcA-GlcNAc. Residues interacting with the substrate are represented by the bond models. Side chains are indicated in red (Asp and Asn), blue (Arg), purple (Trp, Tyr, and Phe), cyan (His), and green (Val, Ser, Ile, and Gln). The substrate is denoted by a yellow bond model. Water molecules are represented as light blue balls. Hydrogen bonds are shown as dashed lines. (C) Structural comparison of D88N/ΔGlcA-Glc-Rha-Glc (green), D88N/ΔGlcA-GlcNAc (magenta), and D88N/ΔGlcA-GalNAc (blue) in the active site.

bonds (<3.5 Å) from residues $His^{87}$, $Asn^{88}$, $Asp^{149}$, $Gln^{211}$, $Arg^{221}$, and $Trp^{225}$ and van der Waals contacts (<4.5 Å) from residues $Trp^{42}$, $His^{87}$, $Asn^{88}$, $Phe^{91}$, $Asp^{149}$, $Trp^{219}$, and $Arg^{221}$ (Fig. 4A and Table 2). In particular, the stacking interaction with $Trp^{42}$ was observed to play an essential role in ΔGlcA binding. These residues are strictly conserved in GH-88 family proteins. The hydrophobic interactions are important for subsite +1 to +3 saccharides (Figs. 3C, 4A and Table 2). Glc at subsite +1 interacts with $Trp^{42}$, $Gln^{211}$, $Tyr^{338}$, and $His^{339}$ residues. The $His^{339}$ residue in UGL is replaced with serine in the UGL homologues [15]. Rha at subsite +2 interacts with $Tyr^{338}$, $His^{339}$, and $Ile^{344}$ residues. Glc at subsite +3 interacts with $Ile^{344}$ and $Ser^{345}$ residues. Variations are observed in residues responsible

Table 2
Interactions between UGL and substrate

| Protein residues | Subsites (saccharide) |
|---|---|
| Hydrogen bonds (<3.5 Å) | |
| $His^{87}$, Asn(Asp)$^{88}$, $Asp^{149}$, $Gln^{211}$, $Arg^{221}$, $Trp^{225}$ | −1 (ΔGlcA) |
| $Tyr^{338}$, $His^{339}$ | +1 (Glc) |
| $Trp^{134}$, $Gln^{211}$ | +1 (GlcNAc) |
| $His^{339}$ | +2 (Rha) |
| $Ser^{345}$ | +3 (Glc) |
| Hydrophobic interactions (<4.5 Å) | |
| $Trp^{42}$, $His^{87}$, Asn(Asp)$^{88}$, $Phe^{91}$, $Asp^{149}$, $Trp^{219}$, $Arg^{221}$ | −1 (ΔGlcA) |
| $Trp^{42}$, $Gln^{211}$, $Tyr^{338}$, $His^{339}$ | +1 (Glc) |
| $Trp^{134}$, $Tyr^{338}$, $His^{339}$ | +1 (GlcNAc) |
| $Tyr^{338}$, $His^{339}$, $Ile^{344}$ | +2 (Rha) |
| $Ile^{344}$ | +3 (Glc) |

for interactions of saccharides at subsites +1 to +3 among UGL homologues ($Ile^{344}$ in UGL is His or Lys in other GH-88 proteins; $Ser^{345}$ is Ala or Gly) [15]. This indicates that the strict interaction between UGL and ΔGlcA at subsite −1 is more important for enzyme catalysis.

*ΔGlcA-GlcNAc bound at the active site*

The unsaturated hyaluronan disaccharide (ΔGlcA-GlcNAc) was bound at the active site in D88N/ΔGlcA-GlcNAc, although its electron density and average *B*-factor (46.3 $Å^2$) were thin and large, respectively (Fig. 3B). The side chain of $Asp^{149}$ was modeled with two alternate conformations. These features differed from those for ΔGlcA-GalNAc in D88N/ΔGlcA-GalNAc (average *B*-factor, 30.8 $Å^2$) (Itoh et al., submitted for publication). However, the molecular weight of ΔGlcA-GlcNAc was the same as that of ΔGlcA-GalNAc since both are isomeric at the C4 hydroxyl group. The substrate concentration of the complex solution is much higher than its $K_m$ value (approximately 2000-fold). Thus, ΔGlcA-GlcNAc is believed to be disordered at the active site (Figs. 3B and 4B).

The interactions between ΔGlcA and UGL are common to the structures that complexed with either unsaturated chondroitin disaccharide, hyaluronan disaccharide, or gellan tetrasaccharide. The absence of a direct interaction between the enzyme and the glycosidic bond that is to be cleaved is also characteristic of UGL. Hydrophobic interactions are also important for GlcNAc (Fig. 4B and Table 2). While some hydrogen bonds are involved in binding GlcNAc, the two aromatic side chains of $Tyr^{338}$ and $Trp^{134}$ are observed to sandwich the pyranose ring of GlcNAc and are important for accommodating GlcNAc at subsite +1, as in the case of unsaturated chondroitin or gellan oligosaccharide.

*Wild type crystal structure without glycine*

The crystal structure of the wild type enzyme without glycine (hex_wild type/apo) determined in this study is similar to those of other UGLs; these structures differ only with respect to the position of the side chains of $Asp^{149}$ and $Trp^{134}$ at the active site, the position of which changes depending on ligand binding (Fig. 5). The side chain of $Asp^{149}$ forms a hydrogen bond with that of $Asp^{88}$ in the hex_wild type/apo, while in the complex (e.g., D88N/ΔGlcA-Glc-Rha-Glc), it forms a hydrogen bond with a water molecule ($Wat^1$) and is directed to the C4 atom of the substrate by approximately 70° rotation (Fig. 5B). Although two MPD molecules in hex_wild type/apo and one Glc molecule in D88N/ΔGlcA-Glc-Rha-Glc show similar hydrophobic interactions with $Trp^{134}$ and $Tyr^{338}$, the direction of the side chain of $Trp^{134}$ is different across the structures (Figs. 4A and C). Glc is restrained at the α-glycoside bond, and this results in the two pyranose rings of ΔGlcA and Glc almost being parallel. The MPD molecule is smaller than saccharide. Thus, the two MPD molecules could bind to two sites in the space between $Trp^{134}$ and $Tyr^{338}$, accompanying the conformational change of the side chain of $Trp^{134}$. In other words, the side chain of $Trp^{134}$ can suitably reorient by rotating approximately 120°, depending on the type of ligand attached (Fig. 5C).

## Discussion

In this study, we analyzed the crystal structures of UGL that was complexed with various substrates in order to clarify the substrate recognition mechanism of the enzyme. ΔGlcA is specifically recognized at subsite −1 through the formation of several hydrogen bonds and stacking interactions in which highly conserved residues of the GH-88 family participate (Fig. 4 and Table 2). In particular, the carboxyl group of ΔGlcA interacts with the enzyme through the formation of hydrogen bonds, and it is stabilized by the positive end of the inner α-helix dipole of the $\alpha_6/\alpha_6$-barrel. This observation has been supported by an inhibition experiment performed using glycine. The glycine molecule, which also contains a carboxyl group, could bind to the active site of UGL and competitively inhibits the enzyme (Fig. 2).

The unsaturated gellan tetrasaccharide (ΔGlcA-Glc-Rha-Glc) was bound from the bottom of the pocket to the wall that was completely covered with aromatic or hydrophobic amino acid residues (Fig. 3C). Although several amino acid residues at subsite +3 are not conserved in GH-88 family proteins and some hydrogen bonds are involved in binding the substrate, hydrophobic interactions are important for binding the saccharides at subsites +1 to +3 (Figs. 3C, 4 and Table 2). Subsite +1 is mainly formed in the space between the side chains of $Tyr^{338}$ and $Trp^{134}$. The side chain of $Tyr^{338}$ forms hydrogen bonds with $Trp^{42}$ and $Asp^{41}$. Thus, $Tyr^{338}$ is stabilized and does not change its position. However, the side chain of $Trp^{134}$ has high flexibility and can move suitably by rotation in order to bind substrates or ligands (Figs. 4 and 5). This flexibility is believed to provide a wider range of substrate specificity at subsite +1 (e.g., GalNAc, GlcNAc, Man, or

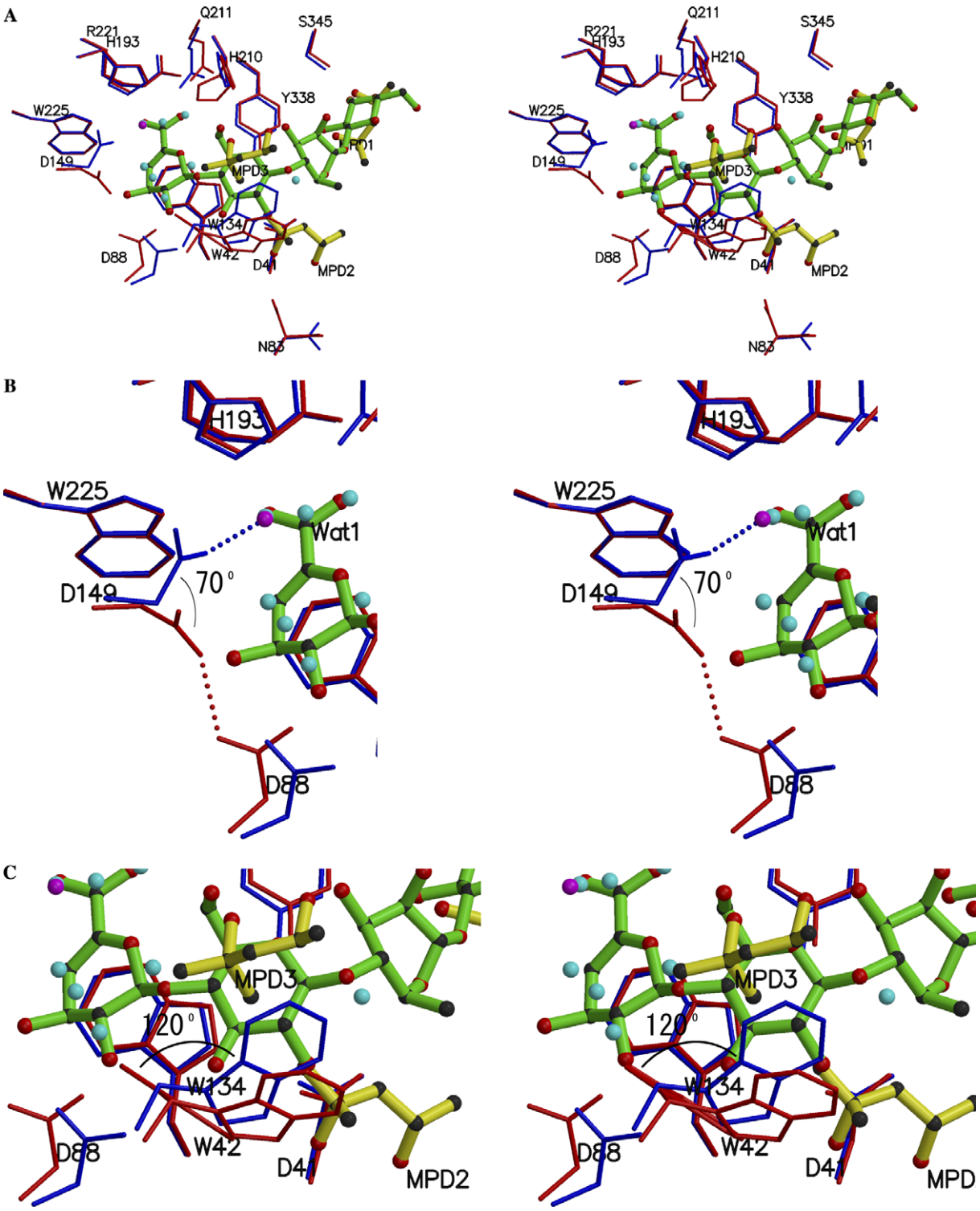


Fig. 5. (A) Structural comparisons of hex_wild type/apo (red) and D88N/ΔGlcA-Glc-Rha-Glc (blue). MPD molecules of hex_wild type/apo are denoted by yellow bond models, while ΔGlcA-Glc-Rha-Glc is denoted by the green bond model. (B) The surroundings of $Asp^{149}$ side chain. The side chain rotates by 70° to facilitate substrate binding. (C) The surroundings of the $Trp^{134}$ side chain. The side chain rotates by 120° to facilitate the substrate binding.

Glc) (Figs. 1A–D). The concavities at subsites +2 and +3 are mainly formed by the aromatic ($Tyr^{338}$ and $His^{339}$) and hydrophobic ($Ile^{344}$) side chains, which contribute to hydrophobic interactions with the substrate. The two pyranose rings fit into this hydrophobic surface at subsites +2 and +3. In comparison with the strong recognition for ΔGlcA through formation of hydrogen bond networks, the interactions between the enzyme and the saccharides at subsites +2 and +3 are very weak. The amino acid residues that participate in the interactions at these subsites

are not highly conserved in the GH-88 family. It is believed that subsites +2 and +3 are not essential for catalysis.

On the other hand, the catalytic mechanism of UGL has been clarified in which the enzyme triggers the hydration of the vinyl ether group in ΔGlcA (Fig. 1E) (Itoh et al., submitted for publication). $Asp^{149}$ acts as a general acid catalyst and a general base catalyst to protonate the C4 atom of ΔGlcA and deprotonate a water molecule, respectively. The deprotonated water molecule attacks the C5 atom of ΔGlcA to yield an unstable hemiketal, which spontaneously converts into an aldehyde (4-deoxy-L-threo-5-hexosulose-uronate) and a saccharide through the formation of a hemiacetal and the cleavage of the glycosidic bond.

In order to donate a proton to the C4 atom of ΔGlcA, which acts as a general acid catalyst, $Asp^{149}$ must be protonated. In the substrate-free structure (hex_wild type/apo), we observed an unusual contact between $Asp^{149}$ and $Asp^{88}$ through the formation of a hydrogen bond (Fig. 5B). This hydrogen bond would provide the protonation of $Asp^{149}$. It is speculated that $Asp^{88}$ may act as a proton donor for $Asp^{149}$ and modulate the $pK_a$ of $Asp^{149}$. This aspartic acid dyad is also observed in cellobiohydrolase Cel6A from *Trichoderma reesei* [30,31]. $Asp^{88}$ plays a crucial role not only in controlling the $pK_a$ of $Asp^{149}$ but also in connecting protons of the hydroxyl groups of ΔGlcA O2 and O3 through the formation of hydrogen bonds. Otherwise, the O3 hydroxyl groups would prevent $Asp^{149}$ protonation and also direct it away from C4 and toward O3.

The average *B*-factor of ΔGlcA-GlcNAc was much higher than that of ΔGlcA-GalNAc although ΔGlcA-GlcNAc and ΔGlcA-GalNAc are isomeric at the C4 hydroxyl group. The side chain of $Asp^{149}$ was modeled with two alternate conformations. The two alternate positions in D88N/ΔGlcA-GlcNAc were similarly located at the positions in the substrate binding or the free structures. The $k_{cat}$ value ($0.022\ \mu M^{-1}\ s^{-1}$) of ΔGlcA-GlcNAc is much smaller than that of ΔGlcA-GalNAc ($0.075\ \mu M^{-1}\ s^{-1}$), although their $K_m$ values are comparable [20]. Thus, it was considered that the binding arrangement of ΔGlcA-GalNAc was more favorable for catalysis than that of ΔGlcA-GlcNAc. GlcNAc causes steric hindrance to the active site in comparison with GalNAc. The *N*-acetyl group of GlcNAc also moves and affects the position of $Trp^{134}$ (Fig. 4C). As a result, it changes the relationships among $Asp^{88}$, $Asp^{149}$, and the O2 and O3 hydroxyl groups of ΔGlcA. This movement probably prevents the proper orientation of the $Asp^{149}$ side chain for catalysis and reduces the $k_{cat}$ value.

In conclusion, these enzyme complex structures clarify the active site architecture of the four subsites. UGL strongly recognizes ΔGlcA at subsite −1 through the formation of hydrogen bonds and stacking interactions, and prefers GalNAc and Glc rather than GlcNAc as a residue accommodated in subsite +1, due to the steric hindrance.

## Acknowledgments

We thank Drs. S. Kawamoto, N. Shimizu, K. Hasegawa, and H. Sakai, the Japan Synchrotron Radiation Research Institute (JASRI), for their assistance in data collection. Diffraction data for the crystals were collected at the BL-41XU and BL-38B1 stations of SPring-8 (Hyogo, Japan) with the approval of JASRI. Computation time was provided by the Supercomputer Laboratory, Institute for Chemical Research, Kyoto University. This work was supported in part by the Program for Promotion of Basic Research Activities for Innovative Biosciences (PROBRAIN) of Japan (K.M. and B.M.), a grant from the National Project on Protein Structural and Functional Analyses (B.M.), and Grants-in-Aid (K.M. and W.H.) from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

## References

[1] J.E. Scott, Extracellular matrix, supramolecular organisation and shape, J. Anat. 187 (1955) 259–269.
[2] S. Ernst, R. Langer, C.L. Cooney, R. Sasisekharan, Enzymatic degradation of glycosaminoglycans, Crit. Rev. Biochem. Mol. Biol. 30 (1995) 387–444.
[3] R.V. Iozzo, Matrix proteoglycans: from molecular design to cellular function, Annu. Rev. Biochem. 67 (1998) 609–652.
[4] U. Lindahl, M. Hook, Glycosaminoglycans and their binding to biological macromolecules, Annu. Rev. Biochem. 47 (1978) 385–417.
[5] Y.M. Michelacci, C.P. Dietrich, A comparative study between a chondroitinase B and a chondroitinase AC from *Flavobacterium heparinum*: isolation of a chondroitinase AC-susceptible dodecasaccharide from chondroitin sulphate B, Biochem. J. 151 (1975) 121–129.
[6] Y.M. Michelacci, C.P. Dietrich, Chondroitinase C from *Flavobacterium heparinum*, J. Biol. Chem. 251 (1976) 1154–1158.
[7] I.W. Sutherland, Polysaccharide lyases, FEMS Microbiol. Rev. 16 (1995) 323–347.
[8] M.J. Jedrzejas, L.V. Mello, B.L. de Groot, S. Li, Mechanism of hyaluronan degradation by *Streptococcus pneumoniae* hyaluronate lyase. Structures of complexes with the substrate, J. Biol. Chem. 277 (2002) 28287–28297.
[9] Y. Maruyama, W. Hashimoto, B. Mikami, K. Murata, Crystal structure of *Bacillus* sp. GL1 xanthan lyase complexed with a substrate: insights into the enzyme reaction mechanism, J. Mol. Biol. 350 (2005) 974–986.
[10] W. Hashimoto, H. Nankai, B. Mikami, K. Murata, Crystal structure of *Bacillus* sp. GL1 xanthan lyase, which acts on the side chains of xanthan, J. Biol. Chem. 278 (2003) 7663–7673.
[11] W. Hashimoto, H. Miki, N. Tsuchiya, H. Nankai, K. Murata, Polysaccharide lyase: molecular cloning, sequencing, and overexpression of the xanthan lyase gene of *Bacillus* sp. strain GL1, Appl. Environ. Microbiol. 67 (2001) 713–720.
[12] W. Hashimoto, T. Inose, H. Nakajima, N. Sato, S. Kimura, K. Murata, Purification and characterization of microbial gellan lyase, Appl. Environ. Microbiol. 62 (1996) 1475–1477.
[13] W. Hashimoto, N. Sato, S. Kimura, K. Murata, Polysaccharide lyase: molecular cloning of gellan lyase gene and formation of the lyase from a huge precursor protein in *Bacillus* sp. GL1, Arch. Biochem. Biophys. 354 (1998) 31–39.
[14] O. Miyake, E. Kobayashi, H. Nankai, W. Hashimoto, B. Mikami, K. Murata, Posttranslational processing of polysaccharide lyase: maturation route for gellan lyase in *Bacillus* sp. GL1, Arch. Biochem. Biophys. 422 (2004) 211–220.

[15] W. Hashimoto, E. Kobayashi, H. Nankai, N. Sato, T. Miya, S. Kawai, K. Murata, Unsaturated glucuronyl hydrolase of*Bacillus* sp. GL1: novel enzyme prerequisite for metabolism of unsaturated oligosaccharides produced by polysaccharide lyases, Arch. Biochem. Biophys. 368 (1999) 367–374.
[16] A. Linker, K. Meyer, B. Weissmann, Enzymatic formation of monosaccharides from hyaluronate, J. Biol. Chem. 213 (1955) 237–248.
[17] W. Hashimoto, K. Maesaka, N. Sato, S. Kimura, K. Yamamoto, H. Kumagai, K. Murata, Microbial system for polysaccharide depolymerization: enzymatic route for gellan depolymerization by *Bacillus* sp. GL1, Arch. Biochem. Biophys. 339 (1997) 17–23.
[18] H. Nankai, W. Hashimoto, H. Miki, S. Kawai, K. Murata, Microbial system for polysaccharide depolymerization: enzymatic route for xanthan depolymerization by *Bacillus* sp. strain GL1, Appl. Environ. Microbiol. 65 (1999) 2520–2526.
[19] G.J. Boulnois, Pneumococcal proteins and the pathogenesis of disease caused by *Streptococcus pneumoniae*, J. Gen. Microbiol. 138 (1992) 249–259.
[20] T. Itoh, S. Akao, W. Hashimoto, B. Mikami, K. Murata, Crystal structure of unsaturated glucuronyl hydrolase, responsible for the degradation of glycosaminoglycan, from *Bacillus* sp. GL1 at 1.8 Å resolution, J. Biol. Chem. 279 (2004) 31804–31812.
[21] S. Mori, S. Akao, H. Nankai, W. Hashimoto, B. Mikami, K. Murata, A novel member of glycoside hydrolase family 88: overexpression, purification, and characterization of unsaturated β-glucuronyl hydrolase of *Bacillus* sp. GL1, Protein Expr. Purif. 29 (2003) 77–84.
[22] S. Mori, S. Akao, O. Miyake, H. Nankai, W. Hashimoto, B. Mikami, K. Murata, Crystallization and preliminary X-ray analysis of a novel unsaturated glucuronyl hydrolase from *Bacillus* sp. GL1, Acta Crystallogr. D Biol. Crystallogr. 59 (2003) 946–949.
[23] Z. Otwinowski, W. Minor, Processing of X-ray diffraction data collected in oscillation mode, Methods Enzymol. 276 (1997) 307–326.
[24] A.T. Brünger, P.D. Adams, G.M. Clore, W.L. DeLano, P. Gros, R.W. Grosse-Kunstleve, J.S. Jiang, J. Kuszewski, M. Nilges, N.S. Pannu, R.J. Read, L.M. Rice, T. Simonson, G.L. Warren, Crystallography and NMR system: a new software suite for macromolecular structure determination, Acta Crystallogr. D Biol. Crystallogr. 54 (1998) 905–921.
[25] R.A. Laskowski, M.W. MacArthur, D.S. Moss, J.M. Thornton, PROCHECK: a program to check the stereochemical quality of protein structures, J. Appl. Crystallogr. 26 (1993) 283–291.
[26] R.W. Hooft, G. Vriend, C. Sander, E.E. Abola, Errors in protein structures, Nature 381 (1996) 272.
[27] P.J. Kraulis, MOLSCRIPT: a program to produce both detailed and schematic plots of protein structure, J. Appl. Crystallog. 24 (1991) 946–950.
[28] E.A. Merrit, M.E.P. Murphy, RASTER3D Version 2.0, a program for photorealistic molecular graphics, Acta Crystallogr. D Biol. Crystallogr. 50 (1994) 869–873.
[29] A. Nicholls, K. Sharp, B. Honig, Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons, Proteins Struct. Funct. Genet. 11 (1991) 281–296.
[30] A. Koivula, L. Ruohonen, G. Wohlfahrt, T. Reinikainen, T.T. Teeri, K. Piens, M. Claeyssens, M. Weber, A. Vasella, D. Becker, M.L. Sinnott, J.Y. Zou, G. Kleywegt, M. Szardenings, J. Stahlberg, T.A. Jones, The active site of cellobiohydrolase Cel6A from *Trichoderma reesei*: the roles of aspartic acids D221 and D175, J. Am. Chem. Soc. 124 (2002) 10015–10024.
[31] G. Wohlfahrt, T. Pellikka, H. Boer, T.T. Teeri, A. Koivula, Probing pH-dependent functional elements in proteins: modification of carboxylic acid pairs in Trichoderma reesei cellobiohydrolase Cel6A, Biochemistry 42 (2003) 10095–10103.