Crystal Structure of Unsaturated Glucuronyl Hydrolase, Responsible for the Degradation of Glycosaminoglycan, from Bacillus sp. GL1 at 1.8 Å Resolution*

Takafumi Itoh‡, Sae Akao§, Wataru Hashimoto§, Bunzo Mikami‡¶, and Kousaku Murata§‡

From the §Division of Agronomy and Horticultural Science, Graduate School of Agriculture, Kyoto University, Gokasho, Uji, Kyoto 611-0011, Japan and the ¶Division of Food Science and Biotechnology, Graduate School of Agriculture, Kyoto University, Gokasho, Uji, Kyoto 611-0011, Japan

Unsaturated glucuronyl hydrolase (UGL) is a novel glycosaminoglycan hydrolase that releases unsaturated D-glucuronic acid from oligosaccharides produced by polysaccharide lyases. The x-ray crystallographic structure of UGL from Bacillus sp. GL1 was first determined by multiple isomorphous replacement (mir) and refined at 1.8 Å resolution with a final R-factor of 16.8% for 25 to 1.8 Å resolution data. The refined UGL structure consists of 377 amino acid residues and 478 water molecules, four glycine molecules, two dithiothreitol (DTT) molecules, and one 2-methyl-2,4-pentanediol (MPD) molecule. UGL includes an α/αβ-barrel, whose structure is found in the six-hairpin enzyme superfamily of an α/αβ-toroidal fold. One side of the UGL α/αβ-barrel structure consists of long loops containing three short β-sheets and contributes to the formation of a deep pocket. One glycine molecule and two DTT molecules surrounded by highly conserved amino acid residues in UGLs were found in the pocket, suggesting that catalytic and substrate-binding sites are located in this pocket. The overall UGL structure, with the exception of some loops, very much resembled that of the Bacillus subtilis hypothetical protein Yter, whose function is unknown and which exhibits little amino acid sequence identity with UGL. In the active pocket, residues possibly involved in substrate recognition and catalysis by UGL are conserved in UGLs and Yter. The most likely candidate catalytic residues for glycosyl hydrolysis are Asp⁸⁸⁸ and Asp¹⁴⁰. This was supported by site-directed mutagenesis studies in Asp⁸⁸⁸ and Asp¹⁴⁰.

Polysaccharides exist ubiquitously in nature as components of the extracellular matrix on the cell surface of many different organisms, ranging from bacteria to mammals (1). These polysaccharides are important to a variety of biological and functional activities, and are divided into three groups: i.e. storage, e.g. starch; structural, e.g. cellulose; and functional, e.g. glycosaminoglycan.

Glycosaminoglycans such as hyaluronan, chondroitin, and heparin are linear, negatively charged polysaccharides with a repeating disaccharide unit consisting of an uronic acid residue (glucuronic or iduronic acid) and an amino sugar residue (glucosamine or galactosamine) (2). Hyaluronan consists of triple strustural GlcA and N-acetyl-D-glucosamine (GlcNAc) (Fig. 1a) and plays an important role in cell-to-cell association in mammals and as a capsule in streptococcal bacteria (3). This polysaccharide is widely present in such human tissues as the eye, brain, liver, skin, and blood (4). Chondroitin, also a member of the glycosaminoglycan family, consists of GlcA and N-acetyl-D-galactosamine (GalNAc) with a sulfate group(s) at position 4 or 6 or both (5) (Fig. 1b). Mammalian chondroitin covalently bound with proteins plays an important role in cellular architecture and permeability (5). These glycosaminoglycans in the extracellular matrix may be a target for pathogens that invade host cells, and many pathogens have been reported to show specific interaction with these polysaccharides (6).

Certain streptococci such as Streptococcus pyogenes and Streptococcus pneumoniae cause severe infectious disease, e.g. pneumonia, bacteremia, sinusitis, and meningitis, and produce polysaccharide lyases, which function as virulence factors in the degradation of hyaluronan and chondroitin (7, 8). Lyases for hyaluronan and chondroitin recognize GlcA residues in polysaccharides and produce unsaturated disaccharides with a GlcA residue having a C=C double bond at the nonreducing terminus through the β-elimination reaction. Unsaturated glucuronyl hydrolase (UGL) catalyzes the hydrolysis of unsaturated disaccharides to an amino sugar and an unsaturated GlcA (ΔGlcA), which is nonenzymatically converted immediately to α-keto acid (Fig. 1) (9). The enzyme is thus thought to be another virulent factor responsible for the complete degradation of glycosaminoglycans. Although a gene coding for UGL was first cloned from Bacillus sp. GL1 that degrades bacterial biofilms such as xanthan, which is produced by pathogenic xanthomonads, and gellan, which is produced by pathogenic sphingomonads (Fig. 1, c and d) (10, 11), highly homologous genes are distributed in pathogenic streptococci that produce polysaccharide lyases (12). We thus defined UGL as a novel member of hydrolase for the degradation of glycosaminoglycans and bacterial biofilms, and the enzyme be-

* This work was supported in part by the Program for Promotion of Basic Research Activities for Innovative Biosciences (PROBRAIN) of Japan (to K. M. and B. M.), a grant for the National Project on Protein Structural and Functional Analyses (to B. M.), and grants-in-aid (to K. M. and W. H.) from the Ministry of Education, Culture, Sports, Science, and Technology of Japan. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates and structure factors (code 1VD5) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

† These authors contributed equally to this work.

‡ To whom correspondence should be addressed. Tel.: 81-774-38-3763; Fax: 81-774-38-3764; E-mail: mikami@kais.kyoto-u.ac.jp.

The abbreviations used are: GlcA, D-glucuronic acid; UGL, unsaturated glucuronyl hydrolase; DTT, dithiothreitol; MPD, 2-methyl-2,4-pentanediol; GlcNAc, N-acetyl-D-glucosamine; ΔGlcA, unsaturated D-glucuronic acid; GalNAc, N-acetyl-D-galactosamine; Glc, D-glucose; Rha, L-rhamnose; Man, D-mannose; mir, multiple isomorphous replacement; r.m.s., root mean square; AGE, N-acetyl-D-glucosamine 2-epimerase.
longs to a new glycoside hydrolase family, GH-88, in the CAZY data base. Inhibitors of polysaccharide lyases and UGL are expected to become potent pharmaceuticals for treating streptococci infectious disease.

Structural analysis of polysaccharide lyases and UGL is indispensable for clarifying molecular mechanisms underlying catalysis and recognition of substrates, and sequential reaction mechanisms involved in polysaccharide depolymerization by bacteria. The crystal structures of polysaccharide lyases such as those for pectate (13–15), alginate (16), hyaluronate (17), chondroitin (18, 19), and xanthan (20) have been determined, and the structural and functional relationship of these lyases have been studied. No three-dimensional structure has been clarified for any UGL, however. This article deals with the crystal structure of UGL, a novel hydrolase for the degradation of glycosaminoglycans and bacterial biofilms, from Bacillus sp. GL1 determined by x-ray crystallography at 1.8 Å resolution and the identification of the active site. The structure provides useful information on the catalytic mechanism and for the molecular design of drugs for the treatment of streptococci, xanthomonad, and sphingomonad infectious diseases.

MATERIALS AND METHODS

Crystallization and X-ray Diffraction—UGL of Bacillus sp. GL1 was overexpressed in Escherichia coli, purified, and crystallized by sitting-drop vapor diffusion as described elsewhere (21). UGL crystals were soaked in a heavy atom derivative solution containing 1 mM of K2PtCl4, 1 mM of Hg(CH3COO)2, or 0.5 mM of NaAuCl4 for 15–50 min at 20 °C. These heavy atom solutions were prepared in 52% (v/v) 2-methyl-2,4-pentanediol (MPD), 0.12M of sodium chloride, 0.1 M of glycine, and 0.1 M of Tris-HCl buffer (pH 7.6). The crystals we used were removed from a droplet on a mounted nylon loop (Hampton Research, Laguna Niguel, CA), and placed in a cold nitrogen gas stream at 100 K. X-ray diffraction images of the UGL crystal (Native 1) were collected using a Quantum 4R CCD area detector (ADSC) with synchrotron radiation at a wave-

---

2 B. Henrissat, P. Coutinho, and E. Deleury, afmb.cnrs-mrs.fr/~cazy/CAZY/index.html.
length of 0.72 Å at the BL-38B1 station of SPring-8. Images were processed with DENZO and SCALEPACK software (22) to a resolution of 1.8 Å (Table I). Diffraction images of another crystal (Native 1) and derivative crystals for phasing were collected with a Bruker Hi-Star multiwire area detector using CuKα radiation. Images were processed with DENZO and SCALEPACK software (22) to a resolution of 25.0 Å–1.80 Å (96.7% completeness). The R-free value calculated for randomly separated 10% data was 18.9%. The stereoquality of the model was assessed using the PROCHECK (26) and WHATCHECK (27) programs. Structural similarity was searched for in the RCSB Protein Data Bank (28) using the DALI program (29). Coordinates of hypothetical protein Yter (INC5) were taken from the RCSB Protein Data Bank. UGL and Yter models were superimposed by a fitting program in TURBO-FRODO. Ribbon plots were prepared using the MOLSCRIPT (30), RASTER3D (31), and GRASP (32) programs.

**Mutagenesis and CD Spectra Measurement**—Asp88 or Asp149 in UGL was replaced with an asparagine residue by the use of the QuikChange site-directed mutagenesis kit (Stratagene), and mutation was confirmed by DNA sequencing. The plasmid pET3a-UGL, which is the expression vector for wild-type UGL (12), was used as a template. Primers were as follows: D88N (Asp88→Asn88), 5'-AGAATCTCGATC-ATCACACACATGCGTCTCTCCTACCTC-3' and 5'-GAATATAGAAGAC-CCAGTTGTTGATGATCGAGATTCT-3'; and D149N (Asp149→Asn149), 5'-GGACCGCATCATCACAATGCCTGCTGATG-3' and 5'-CAGATTCCAGCGCGTATGATGATCGAGATTCT-3' (mutations are indicated by bold letters). The E. coli BL21 (DE3) was transformed with plasmids having a mutation, t. e. pET3a-UGL(D88N) and pET3a-UGL(D149N). Mutant enzymes were expressed and purified for wild-type UGL (12) Structural conformations of purified wild-type and mutant enzymes were evaluated by far-UV CD spectroscopy using a Jasco J720 spectropolarimeter at 260 nm with a demountable accessory attachment.

**Enzyme Assay**—UGL reactions for the wild-type and mutants were conducted at 30 °C as follows: The reaction mixture consisted of 50 mM sodium phosphate buffer (pH 6.5), 20–500 μM of substrate, and enzymes in a 500-μL reaction volume. Enzyme activity was measured by monitoring the decrease in absorbance at 235 nm, corresponding to the loss of the C=C double bond of the substrate because the pyranose ring of the released AGlcA readily opens so that it is nonenzymatically.

### Table I

| Crystal system | Space group | Unit cell parameters (Å) | Molecules/asym. unit |
|---------------|-------------|--------------------------|----------------------|
| Hexagonal     | a = b = 103.01, c = 223.04 |

| Data collection | Resolution limit (last shell) (Å) | Measured reflections (last shell) | Unique reflections (last shell) | Redundancy (last shell) | Completeness (I > σI) (last shell) (%) |
|----------------|----------------------------------|-----------------------------------|---------------------------------|------------------------|--------------------------------------|
|                | 25.0–1.80 (1.86–1.80)            |                                   |                                 | 63,892 (6175)          | 4.7 (3.9)                           |

| Refinement | Final model | Resolution limit (last shell) (Å) | Used reflections (last shell) | Completeness (F > 2σF) (last shell) (%) |
|------------|-------------|----------------------------------|-------------------------------|----------------------------------------|
|            | 377 residues, 478 water, 4 glycine, 2 DTT, and 1 MPD | 25.0–1.80 (1.91–1.80) | 63,516 (9624) | 97.6 (96.0) |

| Phasing statistics | Table II |
|--------------------|----------|
| Compound           | Native 2 | K₃PtCl₄ | Hg(CH₃COO)₂ | NaAuCl₄ |
| Concentration (mM) | 1.0      | 1.0    | 0.5       |
| Soaking time (min) | 50       | 30     | 15        |
| Resolution limit (Å) | 2.7 | 3.5 | 3.5 |
| Phasing power       | 1.08     | 1.23   | 1.06 |
| Rmerge ( Å² )        | 0.612    | 0.652  | 0.701 |
| Rwork ( Å² )         | 0.088    | 0.211  | 0.104 |
| Number of sites      | 2        | 2      | 2        |
| Binding sites        | Met¹, Met¹²¹ | Cys¹⁵⁰, Cys¹⁷⁸ | Cys¹⁵⁰, His¹⁵⁰ |
| Mean figure of merit | 0.426    |        |          |

| Additional heavy atom sites were modeled into four glycine molecules, two dithiothreitol (DTT) molecules, and one MPD molecule from the crystallization medium, and density was excellent for the whole molecule. The final R-factor was 16.8% for 63,316 data points with F > 2σ(F) in a resolution of 25.0–1.8 Å (96.7% completeness).

The R-free value calculated for randomly separated 10% data was 18.9%. The stereoquality of the model was assessed using the PROCHECK (26) and WHATCHECK (27) programs. Structural similarity was searched for in the RCSB Protein Data Bank (28) using the DALI program (29). Coordinates of hypothetical protein Yter (INC5) were taken from the RCSB Protein Data Bank. UGL and Yter models were superimposed by a fitting program in TURBO-FRODO. Ribbon plots were prepared using the MOLSCRIPT (30), RASTER3D (31), and GRASP (32) programs.

**Mutagenesis and CD Spectra Measurement**—Asp88 or Asp149 in UGL was replaced with an asparagine residue by the use of the QuikChange site-directed mutagenesis kit (Stratagene), and mutation was confirmed by DNA sequencing. The plasmid pET3a-UGL, which is the expression vector for wild-type UGL (12), was used as a template. Primers were as follows: D88N (Asp88→Asn88), 5'-AGAATCTCGATC-ATCACACACATGCGTCTCTCCTACCTC-3' and 5'-GAATATAGAAGAC-CCAGTTGTTGATGATCGAGATTCT-3'; and D149N (Asp149→Asn149), 5'-GGACCGCATCATCACAATGCCTGCTGATG-3' and 5'-CAGATTCCAGCGCGTATGATGATCGAGATTCT-3' (mutations are indicated by bold letters). The E. coli BL21 (DE3) was transformed with plasmids having a mutation, t. e. pET3a-UGL(D88N) and pET3a-UGL(D149N). Mutant enzymes were expressed and purified for wild-type UGL (12) Structural conformations of purified wild-type and mutant enzymes were evaluated by far-UV CD spectroscopy using a Jasco J720 spectropolarimeter at 190–260 nm with a demountable quartz cell having a 0.1-mm path length.

**Enzyme Assay**—UGL reactions for the wild-type and mutants were conducted at 30 °C as follows: The reaction mixture consisted of 50 mM sodium phosphate buffer (pH 6.5), 20–500 μM of substrate, and enzymes in a 500-μL reaction volume. Enzyme activity was measured by monitoring the decrease in absorbance at 235 nm, corresponding to the loss of the C=C double bond of the substrate because the pyranose ring of the released AGlcA readily opens so that it is nonenzymatically.
converted to α-keto acid through the loss of the double bond (Fig. 1d) (7, 9). Enzyme concentration was determined by UV spectrophotometry using theoretical molar extinction coefficient ε₂₈₀ 99,570 (M⁻¹ cm⁻¹). Enzyme purity was assessed by SDS-PAGE followed by Coomassie Brilliant Blue staining.

Enzyme purity was assessed by SDS-PAGE followed by Coomassie Brilliant Blue staining. The gellan lyase product (GlcA-Glc-Rha-Glc) for the UGL substrate was prepared as described elsewhere (10). Hyaluronate lyase products (GlcA-GlcNAc) and chondroitin lyase products (GlcA-GalNAc) were obtained from Seikagaku Corporation (Tokyo, Japan).

RESULTS AND DISCUSSION

Crystallization and Structure Determination—UGL of Bacillus sp. GL1 is a monomeric enzyme with a molecular mass of about 43 kDa (377 amino acid residues) (9). A UGL crystal (0.1 × 0.1 × 0.5 mm) was obtained by sitting-drop vapor diffusion as described elsewhere (21). The space group was determined to be P6₃22 with unit cell dimensions of a = b = 103.01 and c = 223.04 Å, and the solvent content was 69% assuming one molecule per asymmetric unit. Results of native data collection using synchrotron radiation at the BL-38B1 station of SPring-8 are summarized in Table I. The phase of the structure was solved by MIR. Table II shows phasing statistics at a resolution of 3.5 Å. The protein model was built after solvent flattening with the PHASES program (23) and refined by simulated annealing and the restrained least-squares method using CNS (25) (Table I).

Quality of Refined Model—The refined model of UGL consists of 377 amino acid residues, and 478 water molecules, four glycine molecules, two DTT molecules, and one MPD molecule. The entire polypeptide chain sequence was well traced, and the electron densities of the main chain and side chain were generally very well defined in the 2Fᵦ - Fᵦ map, except for N-terminal amino acid residue Met¹ and C-terminal amino acid residue Arg³⁷⁷, whose electron density was too low for them to be identified completely. Other ligand molecules were also well fitted. The final overall R-factor for the refined model was 16.8%, with 63,316 unique reflections at a resolution of 25.0–1.8 Å. The final free R-factor was 18.9%. Final root mean square (r.m.s.) deviations from standard geometry were 0.005 Å for bond lengths and 1.22° for bond angles. Based on theoretical curves in the plot calculated according to Luzzati (33), the absolute positional error was estimated to be close to 0.17 Å at a resolution of 5.0–1.8 Å. Most (88.3%) non-glycine residues lie within most favored regions, and other residues (11.3%) within additionally allowed regions of the Ramachandran plot as defined in PROCHECK (26). Ser³⁴⁵ (ϕ = 38°, ψ = 65°), however, falls in generously allowed regions, exhibiting well

Fig. 2. Overall structure of UGL bound with molecules. a and b, structure is prepared by ribbon stereodiagrams using MolScript (30) and Raster3D (31). Colors denote secondary structure elements (pink, α-helices; cyan, β-strands; yellow, loops and coils), glycine (blue), DTT (red), and MPD (green) molecules (a, front; b, side). c, structure is represented as a white molecular surface model. Aromatic residues are pink. Bound molecules are presented by bond models and colored yellow. The figure was drawn using the GRASP (32) program.
and form the wall of the deep funnel-shaped pocket roughly 15–20 Å in diameter at the lip and about 15 Å deep (Fig. 2c). This pocket is widely surrounded by aromatic residues.

**Structural Comparison**—UGL consists of an α/α-toroidal fold. This basic fold is common in glycosyl hydrolases, polysaccharide lyases, and terpenoid cyclases/protein prenyltransferases in the SCOP data base (scop.mrc-lmb.cam.ac.uk/scop/data/scop.b.b.bcl.html) (34). UGL has the α/α-barrel found in the six-hairpin enzyme superfamily of the SCOP data base, which includes glucosamylases (35–37), cellulase catalytic domains (38–40), N-acetyl-d-glucosamine 2-epimerase (AGE) (41), the maltose phosphorylase central domain (42), and hypothetical protein Yter. Based on the structural similarity in the RCSB Protein Data Bank (28) observed with the DALI (29) program, three proteins, *i.e.*, hypothetical protein Yter from *B. subtilis*, glucosamylase from *Thermoaerobacterium thermosaccharolyticum* (36), and AGE from the porcine kidney (41), in the SCOP data base superfamily exhibited the highest degree of similarity. These proteins exhibited Z-scores of 32.7, 21.1, and 19.4. The r.m.s. distance was 2.80 Å for the superimpositioning of 323 Cα atoms of UGL on those of Yter, 3.50 Å for that of 287 Cα atoms of UGL on those of glucosamylase, and 3.30 Å for that of 278 Cα atoms of UGL on those of AGE, although they exhibit no amino acid sequence similarity and catalyze different types of reactions. UGL is an exo-hydrolase acting on unsaturated oligosaccharides produced by polysaccharide lyases, while Yter is a hypothetical protein of unknown function. The crystal structure of Yter determined by the Midwest Center for Structural Genomics has not, to the best of our knowledge, been published, although its coordinates are available in the RCSB Protein Data Bank. Glucosamylase is a polysaccharide exo-hydrolase found in some prokaryotic and many eukaryotic microorganisms. AGE has the α/α-barrel structure we determined, and catalyzes the epimeric reaction for N-acetyl-d-glucosamine and N-acetyl-d-mannosamine (41). Fig. 4 shows the superimpositioning of UGL on Yter; their structures show the best fit overall. The 12 helices of the α/α-barrel are very similar for UGL and Yter in position, direction, and angle, but significant differences exist between UGL and Yter in the loop structure. The L-H1:H2 loop of UGL is longer (23 residues) than that of Yter (13 residues). The L-H7:H8 loop of UGL protrudes from the wall to the outside, while that of Yter is directed into the pocket. L-H11:H12 of UGL makes the pocket open wide, while that of Yter makes the pocket closed.

**Ligands in the Active Site**—The UGL structure contains four glycine molecules (Gly401, 402, 403, 404), two DTT molecules (DTT405, 406), and one MPD molecule (MPD601) (Fig. 2). These molecules were present in the crystallization medium (21). Glycine molecules are bound on the surface mainly through ionic interactions. The positive-charged side chains of Arg221, Arg334, and Arg363 face the carboxyl groups of three Gly401, Gly402, and Gly404 molecules, and the negative-charged side chains of Asp149 and Asp56 face amino groups of two Gly403 and Gly405 molecules. DTT and MPD molecules are bound in the cavity predominantly through a hydrophobic interaction. Aromatic side chains of Tyr328 and Trp334 contact the DTT405 and DTT602 molecules, and Tyr25 the MPD601 molecule.

Notably, one glycine molecule (Gly401) and two DTT molecules (DTT405, 406) are bound in the deep pocket, which was confirmed to be an active site in α/α-barrel enzymes, and is formed by residues belonging to inner helices and long loops (35–47) (Figs. 2c and 5). The two O atoms of the carboxyl group of the Gly401 molecule, bound to the bottom surface of the pocket, are hydrogen-bonded to the Arg221 Nε2 atom (2.9 Å), Gln211 Nδ2 atom (3.0 Å), and Trp225 Nε1 atom (3.0 Å), and the N...
atom of its amino group to the Asp^{149} O^\textit{51} atom (2.9 Å). For DTT molecules, in addition to hydrophobic interaction, the O2 hydroxyl group of DTT^{501} is hydrogen-bonded to the His^{210} N^\textit{2} atom (3.0 Å) (Table III). The inherent substrates, i.e. products of lyases (Fig. 1), of UGL include carboxyl and hydroxyl groups, and pyranose rings of substrates appear to exhibit a stacked hydrophobic interaction with aromatic residues, as is often seen in complexes of polysaccharide lyases with substrates (20, 48–51). Glycine has a carboxyl group, and DTT contains two hydroxyl groups and a hydrophobic moiety. These structural characteristics of glycine and DTT molecules resemble those of substrates. The distance between Gly^{401} and DTT^{502} is about 14 Å, which corresponds to the distance of a trisaccharide. Since UGL can act on unsaturated di-, tri-, and tetrasaccharides (Fig. 1), these molecules (Gly^{401}, DTT^{501}, and DTT^{502}) are thought to interact with UGL in the active site, and to be

### Table III

| Molecule | Atom | Paired atom in UGL | Element | Distance (Å) |
|----------|------|--------------------|---------|-------------|
| Gly^{401} | O1  | Arg^{221} | N^{52}  | H8         | 2.9 |
|          | O1  | Gln^{211} | N^{52}  | H8         | 3.0 |
|          | O2  | Trp^{225} | N^{3}   | L-H7:H8    | 3.0 |
| DTT^{501} | N  | Asp^{149} | O^{1}   | H6         | 3.0 |
|          | O2  | His^{210} | N^{52}  | L-H7:H8    | 3.0 |

### van der Waals contact (C-C distance < 4.5 Å)

| Molecule | Atom | Paired atom in UGL | Element |
|----------|------|--------------------|---------|
| Gly^{401} | C   | Trp^{134} | C^{60}, C^{2}, C^{52}, C^{53}, C^{52}, C^{62} | L-H5:H6 |
|          | C   | Trp^{138} | C^{60}, C^{2}, C^{52}, C^{53}, C^{52}, C^{62} | L-H5:H6 |
|          | C   | Arg^{221} | C^{2}   | L-H7:H8    |
|          | C   | Asp^{149} | C^{1}   | L-H11:H12  |
| DTT^{502} | C   | His^{230} | C^{52}  | L-H7:H8    |
|          | C   | His^{230} | C^{52}  | L-H7:H8    |
|          | C4  | His^{230} | C^{52}  | L-H7:H8    |
|          | C   | Gly^{211} | C^{1}, C^{6} | L-H7:H8  |
|          | C   | Trp^{225} | C^{52}  | L-H7:H8    |
|          | C   | His^{230} | C^{52}  | L-H7:H8    |

FIG. 4. Structural comparison of overall structures of UGL (pink) and hypothetical protein Yter (blue). Superimposed results are shown schematically in C traces. Coordinates of Yter (1NC5) were taken from the RCSB Protein Data Bank (28). The figure was drawn using the GRASP (32) program.

FIG. 5. Stereoview of the active site in the UGL pocket. Amino acid residues face the putative active site, glycine, and DTT molecules. Side chains are red (Asp), blue (Arg), purple (Trp, Tyr, Phe), cyan (His), and green (Gln). Glycine and DTT molecules are yellow. Hydrogen bonds are shown as dotted lines. The figure was prepared using MOLSCRIPT (30) and RASTER3D (31).
located at the three subsites for saccharide substrates. At the active site, side chains of aromatic residues (Trp 42, Trp 134, Trp 219, Trp 225, Phe 91, and Tyr 338), positively charged residues (Arg 221, His 86, His 87, and His 193), negatively charged residues (Asp 88 and Asp 149), and a polar residue (Gln 211), which are completely conserved in UGL and its homologs (Fig. 6), were observed. Other residues, whose side chains also face the solvent at the active site, are His 210 of Bacillus sp. GL1 (Arg of other homologs), His 339 (Ser), Glu 141 (Ser), and Asn 142 (Asp, His). These differences in amino acid residues may imply substrate specificity.

**Structural Basis for Catalysis**—UGL releases nonreducing terminal GlcA from unsaturated oligosaccharides by splitting the bond between the anomeric carbon of GlcA and glycosidic oxygen (C1-O; GlcA-1,3, GlcA-1,2, or GlcA-1,4). The reaction mechanism underlying unsaturated glucuronyl hydrolysis and the anomeric configuration of the released GlcA have yet to be clarified. The pyranose ring of the released GlcA, however, readily

![Fig. 6. Amino acid sequence alignment of UGL and other homologous bacterial proteins made using the ClustalW program (clustalw.genome.ad.jp).](#)
has an ion-pair interaction with the side chain of Arg 221, as described above. Arg 221 is therefore a residue important for catalysis of UGL. This inversion is generally accepted for almost all glycoside hydrolases that catalyze through inversion (average 10.0%), as discussed above (Fig. 4). Asp88 in UGL corresponds to Asp88 in Yter, and in the same way, Asp149 to Asp143, His193 to His189, Arg221 to Arg213, Trp134 to Trp141, Trp219 to Trp211, and Trp225 to Trp217. Trp217 in UGL is replaced by Tyr41 in Yter. The side chains of residues in Yter that correspond to those of His86 and His87 in UGL are in opposite directions. Instead of these residues, the space corresponding to Phe91 in UGL is occupied by Tyr85 in Yter. The putative active site of Yter is smaller than that of UGL. This high structural similarity of putative active sites suggests that Yter binds to a similar UGL substrate (or ligand) or exhibits similar activity.

Intriguingly, the active site of UGL resembles that of hypothetical protein Yter (Fig. 7) more than that of any other structurally similar six-hairpin enzyme. The amino acid sequence identity is very low (less than 10%) between UGL and Yter, however, and long loops are arranged differently, as discussed above (Fig. 4). Asp88 in UGL corresponds to Asp88 in Yter, and in the same way, Asp149 to Asp143, His193 to His189, Arg221 to Arg213, Trp134 to Trp141, Trp219 to Trp211, and Trp225 to Trp217. Trp217 in UGL is replaced by Tyr41 in Yter. The space corresponding to Phe91 in UGL is occupied by Tyr85 in Yter. The putative active site of Yter is smaller than that of UGL. This high structural similarity of putative active sites suggests that Yter binds to a similar UGL substrate (or ligand) or exhibits similar activity.

**Protein Data Bank Accession Number**—The coordinates of open accompanying the loss of the double bond in C4=C5, and the saccharide is nonenzymatically converted to ω-keto acid (Fig. 1d) (7, 9). This makes it difficult to analyze the anemic configuration in the UGL reaction. To verify implications about the active site of UGL above, and to determine the anemic configuration of the intermediate monosaccharide, we tried, but failed, to prepare enzyme crystals bound with sugars, substrates, or products by soaking or cocrytallization. The presence of ligand molecules (one glycine and two DTT) in the active pocket interfered with the binding of sugars. We therefore proposed a catalytic mechanism, based on the UGL structure, conserved amino acid residues, and knowledge of glycoside hydrolysis, detailed below. Asp88, Asp149, His86, His87, His193, and Arg221, the ionizable residues nearest to the glucose molecule corresponding to the GlcA binding site, are completely conserved in UGL and its homologs (Figs. 5 and 6), and may play an important role in glycosyl hydrolysis. The carboxyl group of Gly101 is believed to be located on that of GlcA and has an ion-pair interaction with the side chain of Arg221, as described above. Arg221 is therefore a residue important for GlcA recognition. Two basic mechanisms, inversion and retention, have been proposed for glycoside hydrolases and are classified based on net retention or inversion of the anemic configuration of the reaction product (52). Both reaction mechanisms involve two catalytic ionizable groups of carboxyl and carboxylate groups. Asp88 and Asp149 are thus the most likely catalytic residues for UGL. One is thought to act as the carboxylate group and nucleophile/base for anomeric carbon and the other as the carboxyl group and proton donor/acid for glycosidic oxygen in the first step in both reactions. Asp88 and Asp149 are 7.3 Å apart, a distance suitable for glycoside hydrolases that catalyze through inversion (average 10.0 Å) (53). This inversion is generally accepted for almost all glycoside hydrolases with an α/α̅-barrel, such as glucoamylases (35–37), endoglucanase CelD (38), CelA (39), and endo/exocellulase E4 (40). In addition to glycoside hydrolases, another α/α̅-barrel enzyme, *i.e.* *Lactobacillus* maltose phosphorylase, is also an inversion enzyme (42).

To investigate the two candidates, Asp88 and Asp149, for catalytic residues, two mutants (D88N and D149N) having a substitution of Asn for Asp were prepared and assayed (Table IV). CD spectra for wild-type and mutant enzymes showed almost the same profiles (data not shown), indicating that both mutants have no significant conformational change compared with the wild-type enzyme. Specific activity ($k_{cat}/K_m$) of D88N (2.9 × 10⁻⁶ μM⁻¹ s⁻¹) and D149N (9.8 × 10⁻⁷ μM⁻¹ s⁻¹) was, however, significantly lower than that of the wild-type enzyme (8.1 × 10⁻² μM⁻¹ s⁻¹) for gellan tetrasaccharide (ΔGlcA-Glc-Rha-Glc). Both mutants are almost inactive. The $k_{cat}/K_m$ values of D88N (0.00057 s⁻¹) and D149N (0.00059 s⁻¹) were reduced by ~1000–10,000-fold compared with that of the wild-type enzyme (7.3 s⁻¹), while few fluctuations were observed in $K_m$ values among the three enzymes (wild-type 90 μM; D88N, 200 μM; D149N, 60 μM). These kinetics for D88N and D149N suggest that Asp88 and Asp149 are essential for catalysis and support our hypothesis postulated by x-ray crystallography that one acts as the nucleophile/base and the other as the proton donor/acid.

**Fig. 7. Structural comparison of the active site arrangement of UGL and hypothetical protein Yter.** Main chains are pink for UGL and blue for Yter. Side chains are red for UGL and cyan for Yter. The figure was prepared using MOLESCRIPT (30) and RASTER3D (31).
UGL are in the RCSB Protein Data Bank under accession number 1VD5.

Acknowledgments—Diffraction data for the UGL native crystal was collected at the BL-38B1 station of SPring-8 (Japan). We thank Dr. K. Hasegawa of the Japan Synchrotron Radiation Research Institute (JASRI) for his kind help in data collection. Computation time was provided by the Supercomputer Laboratory, Institute for Chemical Research, Kyoto University.

REFERENCES

1. Scott, J. E. (1995) J. Anat. 187, 259–269
2. Ernst, S., Langer, R., Cooney, C. L., and Sasisekharan, R. (1995) Crit. Rev. Biochem. Mol. Biol. 30, 387–444
3. Laurent, T. C., and Fraser, J. R. (1992) FASEB J. 6, 2397–2404
4. Lindahl, U., and Hook, M. (1978) Annu. Rev. Biochem. 47, 385–417
5. Inoue, R. V. (1998) Annu. Rev. Biochem. 67, 699–652
6. Sawitzky, D. (1989) Med. Microbiol. Immunol. (Berl.) 164, 153–161
7. Linker, A., Meyer, K., and Weissmann, B. (1995) J. Biol. Chem. 270, 237–248
8. Boulnois, G. J. (1992) J. Gen. Microbiol. 138, 249–259
9. Hashimoto, W., Kohayashi, E., Nankai, H., Sato, N., Miya, T., Kawai, S., and Murata, K. (1999) Arch. Biochem. Biophys. 368, 367–374
10. Hashimoto, W., Maesaka, K., Sato, N., Kimura, S., Yamamoto, K., Kumagai, H., and Murata, K. (1997) Arch. Biochem. Biophys. 339, 17–23
11. Nankai, H., Hashimoto, W., Miya, T., Kawai, S., and Murata, K. (1999) Appl. Environ. Microbiol. 65, 2520–2526
12. Mori, S., Akao, S., Nankai, H., Hashimoto, W., Miya, T., and Murata, K. (2003) Protein Expr. Purif. 29, 77–84
13. Yoder, M. D., Keen, N. T., and Jurnak, F. (1993) Science 260, 1503–1507
14. Akita, M., Suzuki, A., Kohayashi, T., Itu, S., and Yamane, T. (2001) Acta Crystallogr. Sect. D. Biol. Crystallogr. 57, 1786–1792
15. Charnock, S. J., Brown, I. E., Tuckenberg, J. P., Black, G. W., and Davies, G. J. (2000) Proc. Natl. Acad. Sci. U. S. A. 97, 12067–12072
16. Yoon, H.-J., Mikami, B., Hashimoto, W., and Murata, K. (1999) J. Mol. Biol. 290, 505–514
17. Li, S., Kelly, S. J., Laman, E., Ferraroni, M., and Jedrzejas, M. J. (2000) EMBO J. 19, 1228–1240
18. Jansson, P. E., Lindberg, B., and Sandford, P. A. (1983) Carbohydr. Res. 124, 133–139
19. Féthière, J., Eggimann, B., and Cygler, M. (1999) J. Mol. Biol. 288, 635–647
20. Hashimoto, W., Nankai, H., Miya, T., and Murata, K. (2003) J. Biol. Chem. 278, 7663–7673
21. Mori, S., Akao, S., Miyaake, O., Nankai, H., Hashimoto, W., Miya, T., and Murata, K. (2003) Acta Crystallogr. Sect. D. Biol. Crystallogr. 59, 946–949
22. Owczynowski, Z., and Muzzo, W. (1997) Methods Enzymol. 276, 307–326
23. Furey, W., and Swaminathan, S. (1997) Methods Enzymol. 277, 590–620
24. Wang, B. C. (1985) Methods Enzymol. 115, 90–112
25. Brugger, A. T., Adams P. D., Clare, G. M., DeLano, W. L., Gros, P., Grusee, Kanteleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read,
26. R. J., Rice, L. M., Simonsen, T., and Warren, G. L. (1998) Acta Crystallogr. Sect. D. Biol. Crystallogr. 54, 905–921
27. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) J. Appl. Crystallogr. 26, 283–291
28. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) Nucleic Acid Res. 28, 235–242
29. Holm, L., and Sander, C. (1995) J. Mol. Biol. 233, 123–138
30. Kraulis, P. J. (1991) J. Appl. Crystallogr. 24, 946–950
31. Merrit, E. A., and Murphy, M. E. P. (1994) Acta Crystallogr. Sect. D. Biol. Crystallogr. 50, 869–873
32. Nicholls, A., Sharp, K., and Honig, B. (1991) Proteins Struct. Funct. Genet. 11, 281–286
33. Luzzati, V. (1952) Acta Crystallogr. 5, 802–810
34. Murzin, A. G., Brenner, S. E., Hubbard, T. J., and Chothia, C. (1995) J. Mol. Biol. 247, 536–540
35. Alshin, A. E., Golubev, A., Firsov, L. M., and Honzatko, R. B. (1992) J. Biol. Chem. 267, 19291–19298
36. Alshin, A. E., Feng, P.-H., Honzatko, R. B., and Reilly, P. J. (2003) J. Mol. Biol. 327, 61–73
37. Ševcık, J., Solovevová, A., Hostinová, E., Gašperik, J., Wilson, K. S., and Dauter, Z. (1998) Acta Crystallogr. Sect. D. Biol. Crystallogr. 54, 854–866
38. Juy, M., Amit, A. G., Alzari, P. M., Paljak, R. J., Claeyssens, M., Biquin, P., and Aubert, J. (1992) Nature 357, 89–91
39. Alzari, P. M., Souchon, H., and Domiguez, R. (1996) Structure 4, 265–275
40. Saken, J., Irwin, D., Wilson, D. B., and Karplus, P. A. (1997) Nat. Struct. Biol. 4, 810–818
41. Rob, T., Mikami, B., Maru, I., Ohta, Y., Hashimoto, W., and Murata, K. (2000) J. Mol. Biol. 303, 733–744
42. Egly, M. P., Uppenberg, J., Haack, L., and Van, Tilbeurgh, H. (2001) Structure 9, 689–697
43. Alshin, A. E., Stoffer, B., Firsov, L. M., Svensson, B., and Honzatko, R. B. (1996) Biochemistry 35, 8319–8328
44. Guezin, D. M. A., Lascombe, M.-B., Costabel, M., Souchon, H., Laman, V., Beguin, P., and Alzari, P. M. (2002) J. Mol. Biol. 316, 1061–1069
45. Parsieglia, G., Belaich, A., Belaich, J. P., and Haser, R. (2002) Biochemistry 41, 11134–11142
46. Van, Petegem, F., Collins, T., Meuwis, M. A., Gerday, C., Feller, G., and Van Beeumen, J. (2003) J. Biol. Chem. 278, 7531–7539
47. Parsieglia, G., Reverbel-Leroy, C., Tartsid, C., Belaich, J. P., Driguez, H., and Haser, R. (2002) J. Biol. Chem. 277, 36678–36688
48. Huang, W., Matte, A., Li, Y., Kim, Y. S., Linhardt, R. J., Su, H., and Cygler, M. (1999) J. Mol. Biol. 289, 1257–1269
49. Mello, L. V., de Groot, B. L., Laman, V., and Cygler, M. (2001) Biochemistry 40, 2359–2372
50. Pennura, K., and Jedrzejas, M. J. (2000) J. Mol. Biol. 299, 885–895
51. Davies, G., and Henzszat, B. (1995) Structure 3, 853–859
52. McCarter, J. D., and Withers, S. G. (1994) Curr. Opin. Struct. Biol. 4, 885–892
