Hydration of vinyl ether groups by unsaturated glycoside hydrolases and their role in bacterial pathogenesis

Wataru Hashimoto,¹* Takafumi Itoh,¹ Yukie Maruyama,¹ Bunzo Mikami,² Kousaku Murata¹

¹Division of Food Science and Biotechnology, Graduate School of Agriculture, Kyoto University, Uji, Kyoto, Japan. ²Division of Applied Life Sciences, Graduate School of Agriculture, Kyoto University, Uji, Kyoto, Japan

Received 15 September 2007 · Accepted 30 November 2007

Summary. Many pathogenic microorganisms invade mammalian and/or plant cells by producing polysaccharide-degrading enzymes (lyases and hydrolases). Mammalian glycosaminoglycans and plant pectins that form part of the cell surface matrix are typical targets for these microbial enzymes. Unsaturated glycoside hydrolase catalyzes the hydrolytic release of an unsaturated uronic acid from oligosaccharides, which are produced through the reaction of matrix-degrading polysaccharide lyase. This enzymatic ability suggests that unsaturated glycoside hydrolases function as virulence factors in microbial infection. This review focuses on the molecular identification, bacterial distribution, and structure/function relationships of these enzymes. In contrast to general glycoside hydrolases, in which the catalytic mechanism involves the retention or inversion of an anomeric configuration, unsaturated glycoside hydrolases uniquely trigger the hydrolysis of vinyl ether groups in unsaturated saccharides but not of their glycosidic bonds. [Int Microbiol 2007; 10(4):233-243]

Key words: bacterial lyases · bacterial hydrolases · glycoside hydrolase family · glycosaminoglycan · pathogenic/saprophytic bacteria

Introduction

The adhesion of bacteria to eukaryotic cells may be a primary mechanism of bacterial infection, in addition to playing a role in secondary effects of the infectious process. Polysaccharides, including glycosaminoglycans and pectins, that form part of the cell surface matrix are typical targets of microbial pathogens that invade host cells, and many specific interactions between pathogens and these polysaccharides have been described [16]. In this context, bacterial polysaccharide lyases function as virulence factors.

Several different enzymes were found to act on the unsaturated saccharides produced through the reactions of polysaccharide lyases. These enzymes included other types of lyases as well as hydrolases. For example, Erwinia chrysanthemi oligogalacturonate lyase (Ogl) [48] and Sphingomonas sp. A1 oligoalginate lyase (A1-IV) [12] degrade pectin- and alginate-unsaturated oligosaccharides produced through reactions of pectin and alginate lyases. Ogl and A1-IV release an unsaturated monosaccharide and unsaturated saccharide between the unsaturated glucuronyl residue (ΔGlcA) at the nonreducing termi-
nus and the saccharide linked to the residue, with subsequent release of ΔGlcA and saturated saccharide from substrate oligosaccharides. The enzyme was thus designated an unsaturated glucuronyl hydrolase (UGL) [11]. Although the enzyme showed no homology with other proteins in DNA and protein databases, as determined following the cloning and sequencing of the enzyme gene in Bacillus [11], recent complete genome sequence analyses indicate that homologous genes are encoded in pathogenic bacteria, such as streptococci and vibrio [37]. In addition to UGL, an unsaturated galacturonyl hydrolase (UGH) has recently been identified in Bacillus subtilis [20]. In unsaturated saccharides, the enzyme specifically hydrolyzes the glycosidic bond between the unsaturated galacturonyl residue (ΔGlcA) at the nonreducing terminus and the saccharide linked to the residue.

UGL and UGH belong to two huge novel families GH-88 and GH-105, respectively, which (as of Sept. 2007) comprise 76 (GH-88) and 61 (GH-105) bacterial and fungal members. Recent structural analyses showed that UGL and UGH share a unique catalytic mechanism, one that differs from the retention or inversion mechanism of general glycoside hydrolases [17,21]. Here, we detail the enzyme characteristics, biological roles, and structure/function relationships of the GH-88 and GH-105 families of unsaturated glycoside hydrolases, including, respectively, UGL and UGH.

**Classification of carbohydrate-related enzymes**

Carbohydrate-related enzymes, such as glycoside hydrolases, polysaccharide lyases, glycosyl transferases, and carbohydrate esterases, are categorized into more than 200 families based on amino acid sequences in the Carbohydrate-Active enZymes (CAZy) database [http://afmb.cnrs-mrs.fr] [3]. Among these enzymes, glycoside hydrolases account for the greatest number of families, numbering over 100 (families GH-1 to GH-110, as of Sept. 2007). Tertiary structures of hydrolases in 69 families have been determined and their structure/function relationships well-analyzed. Almost all glycoside hydrolases have two catalytic residues as general acid and base/nucleophile catalysts. Most of the catalytic residues consist of acidic amino acids, Asp and/or Glu. Catalytic mechanisms of glycoside hydrolases are generally divided into two types based on the anomeric configuration of the substrate in the reaction products, i.e., retention or inversion. These glycoside hydrolase mechanisms have been extensively studied and reviewed [4,45].

Polysaccharide lyases are classified into 18 families (PL-1 to PL-18, as of Sept. 2007). These enzymes commonly recognize uronic acid residues in polysaccharides, catalyze a β-elimination reaction, and produce unsaturated saccharides with C=Å double bonds at C4 and C5 sites in nonreducing terminal uronic acid residues. The characteristics of lyases indicate that these enzymes share common structural features determining uronate recognition and the β-elimination reaction. Crystal structures of polysaccharide lyases belonging to 11 families have been determined, and structure/function relationships of enzymes, such as lyases for pectate, pectin, alginate, chondroitin, hyaluronan, xanthan, and rhamnagalacturonan (RG), are being analyzed [13,26,32–34,53,54]. In the catalytic reaction carried out by polysaccharide lyases, positively charged residues bind to and/or neutralize the carboxyl group in uronic acid residues. A general base catalyst extracts the proton from C5 of uronic acid residues and a general acid catalyst donates the proton to the glycosidic bond to be cleaved [46].

**Enzyme characteristics**

**Unsaturated glucuronyl hydrolase (UGL).** Substrates for UGL are unsaturated saccharides with a ΔGlcA residue at the nonreducing terminus, released from polysaccharides through polysaccharide lyase reactions. Typical substrates include mammalian extracellular matrices, glycosaminoglycans such as chondroitin, heparin, and hyaluronan, and bacterial biofilms such as xanthan and gellan. Unsaturated saccharides show an absorbance at 235 nm due to the double bond in the ΔGlcA residue at the nonreducing terminus. UGL activity is measured by monitoring the decrease of absorbance at 235 nm, because the resulting ΔGlcA is nonenzymatically converted to an α-keto acid, followed by the loss of the double bond at C4 and C5 in this residue [11,52]. UGL activity was first detected in F. heparinum, which degrades chondroitin and heparin [27]. The bacterial enzyme acts on unsaturated chondroitin disaccharide produced from chondroitin through the reaction of chondroitin lyase and releases ΔGlcA from the disaccharide [7]. Enzyme activity was also observed in Bacillus sp. strain GL1, in which xanthan and gellan are degraded [11]. The bacillus enzyme acts on xanthan trisaccharide, produced from xanthan through reactions of xanthan lyase, β-D-glucanase, and β-D-glucosidase, and gellan tetrasaccharide, produced from gellan through the reaction of gellan lyase [11]. The unsaturated disaccharides derived from chondroitin and hyaluronan become substrates for the bacillus enzyme.

UGLs have been purified from F. heparinum and Bacillus sp. strain GL1 and enzymatically characterized [7,11,27,37,52]. Except for N-terminal modification and amino acid
composition, the flavobacterial and bacillus enzymes are similar in molecular mass (42 kDa), optimum pH (pH 6.0–6.5) and temperature (37–45°C), the absence of a metal requirement, and substrate specificity (sulfate-free unsaturated disaccharides from chondroitin and hyaluronan, and 6S-monomethylated chondroitin disaccharide). Neither enzyme acts on the substrate p-nitrophenyl glucuronide, as do general glucuronidases. The N-terminus of the flavobacterial enzyme was modified uncertainly, while the N-terminal amino acid sequence of the bacillus enzyme was readily determined. Both enzymes specifically release ΔGlcA from unsaturated saccharides but not GlcA from polysaccharide. Residual saccharide other than ΔGlcA in the resulting reaction products is saturated and has no C=C double bond. These results demonstrate that UGL is a novel exotype glycoside hydrolase rather than a general glucuronidase.

In 1999, the gene encoding the enzyme was cloned in Bacillus sp. strain GL1. It consists of 1134 bp coding for a 42,859-kDa protein (377 amino acids) that has no known homology with other proteins in DNA and protein databases [11]. In 2002, the flavobacterial gene (1209 bp) for UGL was also isolated; it encodes a 45,621-kDa protein (402 amino acids) [40]. In the flavobacterial enzyme, the N-terminal 20 amino acid residues probably function as a signal peptide. Although the enzyme properties of bacillus and flavobacterial UGLs are similar, the amino acid sequence identity between the two is not high (30% identity). Recent complete genome sequence analyses indicated that a large number of microorganisms, ranging from bacteria to fungi (over 70 species), have a UGL-homologous gene in their genome. In the CAZy database, UGL and its homologues form a new family, GH-88. Microbial producers of UGL include pathogenic bacteria such as clostridia, streptococci, and vibrios. Genes homologous to that encoding the bacillus enzyme have been found mainly in streptococci: Streptococcus agalactiae, Streptococcus pneumoniae, Streptococcus pyogenes, and Streptococcus suis. The genetic features of UGL in the streptococcal genome are detailed below.

**Unsaturated galacturonyl hydrolase (UGH).** In 2006, a novel enzyme (YteR) was identified in the saprophytic bacterium Bacillus subtilis [20]. YteR is responsible for the metabolism of a plant cell wall polysaccharide, RG, which consists of L-rhamnose (Rha) and GalA. Although YteR was initially regarded as a function-unknown protein, after completion of genomic analysis, determination of the enzyme’s tertiary structure showed that it was similar to that of UGL [16]. YteR acts on unsaturated RG disaccharides (ΔGlcA-Rha) produced from RG through the reaction of RG lyases, and releases ΔGlcA and Rha through the hydrolytic reaction. Similar to the UGL reaction, the resulting ΔGalA is nonenzymatically converted to α-keto acid. Based on this reaction, YteR has been designated an unsaturated galacturonyl hydrolase (UGH).

Since GalA is the C4 epimer of GlcA, ΔGalA is identical to ΔGlcA due to the lack of the hydroxyl group at the C4 site, which in turn is a result of the reaction of polysaccharide lyases. Thus, UGL and UGH release the same molecule nonenzymatically converted to α-keto acid. Based on this reaction, YteR has been designated an unsaturated galacturonyl hydrolase (UGH).

Since GalA is the C4 epimer of GlcA, ΔGalA is identical to ΔGlcA due to the lack of the hydroxyl group at the C4 site, which in turn is a result of the reaction of polysaccharide lyases. Thus, UGL and UGH release the same molecule nonenzymatically converted to α-keto acid. Based on this reaction, YteR has been designated an unsaturated galacturonyl hydrolase (UGH).

Since GalA is the C4 epimer of GlcA, ΔGalA is identical to ΔGlcA due to the lack of the hydroxyl group at the C4 site, which in turn is a result of the reaction of polysaccharide lyases. Thus, UGL and UGH release the same molecule nonenzymatically converted to α-keto acid. Based on this reaction, YteR has been designated an unsaturated galacturonyl hydrolase (UGH).

Since GalA is the C4 epimer of GlcA, ΔGalA is identical to ΔGlcA due to the lack of the hydroxyl group at the C4 site, which in turn is a result of the reaction of polysaccharide lyases. Thus, UGL and UGH release the same molecule nonenzymatically converted to α-keto acid. Based on this reaction, YteR has been designated an unsaturated galacturonyl hydrolase (UGH).

Since GalA is the C4 epimer of GlcA, ΔGalA is identical to ΔGlcA due to the lack of the hydroxyl group at the C4 site, which in turn is a result of the reaction of polysaccharide lyases. Thus, UGL and UGH release the same molecule nonenzymatically converted to α-keto acid. Based on this reaction, YteR has been designated an unsaturated galacturonyl hydrolase (UGH).

Since GalA is the C4 epimer of GlcA, ΔGalA is identical to ΔGlcA due to the lack of the hydroxyl group at the C4 site, which in turn is a result of the reaction of polysaccharide lyases. Thus, UGL and UGH release the same molecule nonenzymatically converted to α-keto acid. Based on this reaction, YteR has been designated an unsaturated galacturonyl hydrolase (UGH).

Since GalA is the C4 epimer of GlcA, ΔGalA is identical to ΔGlcA due to the lack of the hydroxyl group at the C4 site, which in turn is a result of the reaction of polysaccharide lyases. Thus, UGL and UGH release the same molecule nonenzymatically converted to α-keto acid. Based on this reaction, YteR has been designated an unsaturated galacturonyl hydrolase (UGH).

Since GalA is the C4 epimer of GlcA, ΔGalA is identical to ΔGlcA due to the lack of the hydroxyl group at the C4 site, which in turn is a result of the reaction of polysaccharide lyases. Thus, UGL and UGH release the same molecule nonenzymatically converted to α-keto acid. Based on this reaction, YteR has been designated an unsaturated galacturonyl hydrolase (UGH).

Since GalA is the C4 epimer of GlcA, ΔGalA is identical to ΔGlcA due to the lack of the hydroxyl group at the C4 site, which in turn is a result of the reaction of polysaccharide lyases. Thus, UGL and UGH release the same molecule nonenzymatically converted to α-keto acid. Based on this reaction, YteR has been designated an unsaturated galacturonyl hydrolase (UGH).

Since GalA is the C4 epimer of GlcA, ΔGalA is identical to ΔGlcA due to the lack of the hydroxyl group at the C4 site, which in turn is a result of the reaction of polysaccharide lyases. Thus, UGL and UGH release the same molecule nonenzymatically converted to α-keto acid. Based on this reaction, YteR has been designated an unsaturated galacturonyl hydrolase (UGH).

Since GalA is the C4 epimer of GlcA, ΔGalA is identical to ΔGlcA due to the lack of the hydroxyl group at the C4 site, which in turn is a result of the reaction of polysaccharide lyases. Thus, UGL and UGH release the same molecule nonenzymatically converted to α-keto acid. Based on this reaction, YteR has been designated an unsaturated galacturonyl hydrolase (UGH).

Since GalA is the C4 epimer of GlcA, ΔGalA is identical to ΔGlcA due to the lack of the hydroxyl group at the C4 site, which in turn is a result of the reaction of polysaccharide lyases. Thus, UGL and UGH release the same molecule nonenzymatically converted to α-keto acid. Based on this reaction, YteR has been designated an unsaturated galacturonyl hydrolase (UGH).

Since GalA is the C4 epimer of GlcA, ΔGalA is identical to ΔGlcA due to the lack of the hydroxyl group at the C4 site, which in turn is a result of the reaction of polysaccharide lyases. Thus, UGL and UGH release the same molecule nonenzymatically converted to α-keto acid. Based on this reaction, YteR has been designated an unsaturated galacturonyl hydrolase (UGH).

Since GalA is the C4 epimer of GlcA, ΔGalA is identical to ΔGlcA due to the lack of the hydroxyl group at the C4 site, which in turn is a result of the reaction of polysaccharide lyases. Thus, UGL and UGH release the same molecule nonenzymatically converted to α-keto acid. Based on this reaction, YteR has been designated an unsaturated galacturonyl hydrolase (UGH).

Since GalA is the C4 epimer of GlcA, ΔGalA is identical to ΔGlcA due to the lack of the hydroxyl group at the C4 site, which in turn is a result of the reaction of polysaccharide lyases. Thus, UGL and UGH release the same molecule nonenzymatically converted to α-keto acid. Based on this reaction, YteR has been designated an unsaturated galacturonyl hydrolase (UGH).

Since GalA is the C4 epimer of GlcA, ΔGalA is identical to ΔGlcA due to the lack of the hydroxyl group at the C4 site, which in turn is a result of the reaction of polysaccharide lyases. Thus, UGL and UGH release the same molecule nonenzymatically converted to α-keto acid. Based on this reaction, YteR has been designated an unsaturated galacturonyl hydrolase (UGH).

Since GalA is the C4 epimer of GlcA, ΔGalA is identical to ΔGlcA due to the lack of the hydroxyl group at the C4 site, which in turn is a result of the reaction of polysaccharide lyases. Thus, UGL and UGH release the same molecule nonenzymatically converted to α-keto acid. Based on this reaction, YteR has been designated an unsaturated galacturonyl hydroses UGL and UGH are responsible for the metabolism of mammalian extracellular matrices, plant cell walls, and bacterial biofilms. The biological role of each enzyme in its typical bacterial producer, i.e., Bacillus sp. strain GL1, F. heparinum and Streptococcus for UGL, and B. subtilis for UGH, are the focus of this section.

**Bacillus sp. strain GL1.** This bacterium degrades bacterial biofilms made up of xanthan and gellan as sole carbon source [8,9]. Xanthan is an exopolysaccharide produced by the plant-pathogenic bacterium Xanthomonas campestris. The polymer has a cellulosic backbone with linear trisaccha-
ride side chains consisting of a pyruvylated mannose (Pyr-Man)-GlcA-acetylated Man sequence attached at the C3 position of alternate glucosyl (Glc) residues [22]. The xanthan depolymerization system in Bacillus sp. strain GL1 was determined through genetic, enzymatic, and glycochemical analyses [41] (Fig. 1A). Extracellular xanthan lyases (97 and 81 kDa) first attack xanthan to release PyrMan from the side chains [9]. A β-D-glucanase (350 kDa) hydrolyzes glycosidic bonds in the main chain and degrades the polymer into the tetrasaccharide (ΔGlcA-Man-(Glc)Glc) [41]. The latter product is degraded into its constituent monosaccharides through subsequent reactions of intracellular β-D-glucosidase (51 kDa), UGL (42 kDa), and α-D-mannosidase (330 kDa) [41].

Gellan is produced by the Gram-negative bacterium Sphingomonas paucimobilis and consists of a linear repeating tetrasaccharide (Glc-GlcA-Glc-Rha) with O-acetyl and glyceryl moieties on Glc adjacent to GlcA as side chains [23]. Gellan degradation in Bacillus sp. strain GL1 is shown in Fig. 1B. Gellan is first depolymerized to a tetrasaccharide-repeating unit (ΔGlcA-Glc-Rha-Glc) by extracellular gellan lyases (260 and 130 kDa) [10]. The resulting tetrasaccharide is then degraded into its constituent monosaccharides through subsequent reactions of intracellular UGL (42 kDa), β-D-glucosidase (51 kDa), and α-L-rhamnosidase (100 kDa) [11].

UGL is inductively produced in bacillus cells in the presence of xanthan or gellan, while the enzyme is not expressed in bacterial cells grown on glucose or pectin [11]. The UGL substrate specificity and expression profile indicate that the enzyme is essential for complete depolymerization of bacterial biofilms (xanthan and gellan) by Bacillus sp. strain GL1.

**Glycosaminoglycans.** F. heparinum and some streptococci degrade mammalian glycosaminoglycans from extracellular matrices. Glycosaminoglycans such as chondroitin, heparin, and hyaluronan are linear and negatively charged polysaccharides with a repeating disaccharide unit consisting of a uronic acid residue (GlcA or iduronic acid, IdoA) and an aminosugar residue (glucosamine or galactosamine) [6]. Chondroitin consists of GlcA and N-acetyl-d-galactosamine (GalNAc), with a sulfate group at position 4 or 6 or both [15]. Chondroitin sulfates A and C have GalNAc with a sulfate group at positions 4 and 6. Heparin and heparan sulfate are composed of D-glucosamine (GlcN) and GlcA or IdoA [44]. Hyaluronan is composed of GlcA and N-acetyl-D-glucosamine (GlcNAc) [5]. Except for hyaluronan, mammalian glycosaminoglycans covalently bind to proteins; the resulting proteoglycans play an important role in cellular architecture and permeability [15]. Hyaluronan is crucial for cell-to-cell association in mammals and for the capsule in streptococcal bacteria [25]. These glycosaminoglycans in extracellular matrices become a target for pathogens to invade host cells and serve as sites of specific interaction for many pathogens [47].

**Flavobacterium heparinum.** Gram-negative F. heparinum, isolated from a soil sample, degrades chondroitin, heparin, and hyaluronan by the actions of several enzymes [27,30,39,49,51]. In the presence of chondroitin and hyaluronan, the bacterium inducibly produces chondroitin AC lyase (77 kDa), chondroitin B lyase (54 kDa), sulphatase (50 kDa), and UGL (42 kDa) in the periplasm [7]. Chondroitin AC lyase acts on chondroitin, chondroitin sulfate A, chondroitin sulfate C, and hyaluronan, and releases unsaturated disaccharides with ΔGlcA at the nonreducing terminus. Dermatan sulfate is depolymerized into unsaturated disaccharide through the reaction of chondroitin B lyase. The resulting unsaturated disaccharides derived from chondroitin, chondroitin sulfate C, and hyaluronan are subsequently degraded into constituent monosaccharides through the reaction of UGL, while the enzyme is inert on unsaturated disaccharide 4-sulfate from chondroitin sulfate A and unsaturated disaccharide from dermatan sulfate [7]. In heparin degradation, F. heparinum also produces three heparin lyases: I (43 kDa), II (85 kDa), and III (73 kDa) [30]. Unsaturated disaccharides resulting from the reactions of lyases are further degraded by UGL, making it a key enzyme in glycosaminoglycan metabolism in F. heparinum.

**Streptococcus spp.** Infections with certain streptococcal species, such as S. agalactiae, S. pneumoniae, and S. pyogenes, cause severe diseases, e.g., pneumonia, bacteremia, sinusitis, and meningitis [29,36]. Septicemia, meningitis, and pneumonia are induced in pigs through infection by S. suis [31]. All of these bacterial species produce polysaccharide lyases, which function virulence factors that degrade glycosaminoglycans [24,28]. As described above, a UGL-homologous gene is present in the genome of these streptococci. Interestingly, this gene is located in a gene cluster for the putative phosphotransferase system (PTS), which imports aminosugar, a component of glycosaminoglycans. The cluster also includes the polysaccharide lyase gene (Fig. 1C), strongly suggesting that UGL functions as another streptococcal virulence factor, one that is responsible for the complete degradation of glycosaminoglycans. Unlike the bacillus and flavobacterial enzymes, streptococcal UGL acts on unsaturated chondroitin disaccharide 4-sulfate [unpublished results], i.e., the streptococcal enzyme is capable of degrading highly sulfated mammalian glycosaminoglycans. This property enables the bacteria to readily invade mammalian cells through the degradation of extracellular sulfated glycosaminoglycans.
**Bacillus subtilis.** This bacterium is a typical representative of saprophytic microbes [38]. For such saprophytic and/or plant pathogenic bacteria, plant cell wall degradation is a major event in plant cell invasion. The plant cell wall consists of pectin, cellulose, and hemicellulose [2]. Pectin is further divided into three components, a linear main chain (polygalacturonic) and two branched chains, RG-I and -II [50]. RG-I consists of alternating Rha and GalA as a backbone with additions of arabinans and galactans as side chains [35]. RG-II contains polygalacturonic as a backbone and a complex of approximately 30 monosaccharides as a side chain [43]. In the presence of RG-I, *B. subtilis* induces the expression of three gene clusters, *yes*, *yte*, and *ybc*-ybd [42]. Gene clusters *yes* and *yte* include genes for YesR and YteR. The degradation pathway of RG-I in *B. subtilis* is shown in Fig. 1D. An extracellular RG lyase, YesW (68 kDa), catalyzes the initial cleavage of the glycosidic bond in the RG-I main chain, with the resulting oligosaccharides converted to the unsaturated RG disaccharide through the reaction of extracellular exotype YesX (68 kDa). The unsaturated disaccharide is finally degraded into constituent monosaccharides through the reaction of intracellular YteR (43 kDa) or YesR (39 kDa). Generally, the bacterium cannot assimilate polysaccharides until monosaccharides have been released, indicating that the UGHs YteR and YesR are essential to the metabolism of plant cell wall RG-I in *B. subtilis*.

**Structure/function relationships.** Unsaturated glycoside hydrolases UGL and UGH are unique in that they act on unsaturated saccharides but not on saturated saccharides, and release unsaturated uronic acid. To clarify the structure/function relationships of these enzymes, the three-dimensional structures of *Bacillus* sp. strain GL1 and *S. agalactiae* UGLs and *B. subtilis* UGH were determined by X-ray crystallography [16,20].

**Structures of UGL and UGH.** The overall structure of bacillus UGL is shown in Fig. 2A. The enzyme has approximate dimensions of $45 \times 45 \times 40$ Å and shows a single globular domain consisting of 12 α-helices (H1-H12), three antiparallel β-sheets composed of two or three strands (SA, SB, and SC), two short 3_10-helices, and loops connecting each element. The 12 α-helices constitute an α/α_β-barrel structure with a deep cleft, considered to be the active site. The
\(\alpha/\alpha\)-barrel structure is made up of six outer \(\alpha\)-helices (H1, H3, H5, H7, H9, and H11) oriented in roughly the same direction, and six inner \(\alpha\)-helices (H2, H4, H6, H8, H10, and H12) running in the opposite direction. These 12 helices are connected by short and long loops in a nearest-neighbor, up-down pattern. Short loops extend along one side of the barrel. On the other side, long loops constitute the wall of a deep pocket-like cleft, which has a diameter of 15–20 Å at the lip and is 15 Å deep.

In structural comparisons to bacillus UGL, three proteins, i.e., the hypothetical protein YteR from \textit{B. subtilis} [55] (Fig. 2C), glucoamylase from \textit{Thermoanaerobacterium thermosaccharolyticum} [1], and N-acetyl-d-glucosamine 2-epimerase (AGE) [19] from porcine kidney, showed the highest similarity. These proteins also adopt an \(\alpha/\alpha\)-barrel structure as a basic scaffold, although they show no sequence similarity at the amino acid level. The crystal structure of YteR was determined as a function-unknown protein by the Midwest Center for Structural Genomics [55]. The high structural similarity between UGL and YteR enabled us to clarify that YteR exhibits UGL-like activity. The detailed intrinsic function of YteR was described in the previous section.

**Catalytic mechanism.** The structure of an enzyme that is complexed with substrates contributes to the identification of its active site and clarification of its mechanisms for catalytic reaction and substrate recognition. Thus, the crystal structure of bacillus UGL complexed with a substrate, unsaturated chondroitin disaccharide (\(\Delta\text{GlcA-GalNAc}\)), was determined at 1.9 Å resolution [17] (Fig. 2B). No significant conformational change occurs in the overall structure between substrate-free and bound enzymes. To study the complexed structure, enzyme D88N, which has a mutation at position 88 (Asp\(\rightarrow\)Asn), is used because the wild-type enzyme readily degrades the substrate and thus cannot be isolated in the substrate-bound form. The disaccharide is bound to the deep pocket at the center of the UGL molecule, which indicates that the active site is in the pocket. A molecular surface model of the complex structure shows that the active site is surrounded by aromatic and positively charged residues (Fig. 2D).

\(\Delta\text{GlcA}\) is bound to residues Trp42, Asn(Asp)88, Phe91, Trp134, Asp149, Gln211, Trp219, Arg221, Trp225, and Tyr338, while GalNAc interacts with residues Trp42, His86, Trp134, His210, Tyr338, and His339 (Fig. 3A). Almost all residues responsible for binding to the disaccharide are well

**Fig. 2.** Overall structure of unsaturated glycoside hydrolases. (A) Ligand-free UGL from \textit{Bacillus} sp. strain GL1. (B) Substrate (\(\Delta\text{GlcA-GalNAc}\))-bound UGL (D88N) from \textit{Bacillus} sp. strain GL1. (C) Substrate (\(\Delta\text{GalA-Rha}\))-bound YteR (D143N) from \textit{B. subtilis}. Colors denote elements with a secondary structure (pink, \(\alpha\)-helices; cyan, \(\beta\)-strands; and yellow, turns and coils). A green or blue ball-and-stick model indicates the substrate (red, oxygen atom; yellow, carbon atom; and blue, nitrogen atom). (D) Molecular surface of the substrate-bound UGL (D88N) from \textit{Bacillus} sp. strain GL1. Colors denote amino acid residues (red, Asp or Glu; blue, Lys or Arg; cyan, His; and magenta, Trp, Tyr, or Phe).
conserved in family GH-88. The complex structure is unique in that no atom forms hydrogen bonds to the glycosidic oxygen (O3 atom of GalNAc). The arrangement of amino acid residues and water molecules at the active site is detailed as follows (Fig. 3A): a water molecule (Wat) is situated at an “α-face position” of ΔGlcA and binds through hydrogen bonds to Asp149 (3.1 Å), His193 (2.7 Å), Gln211 (2.6 Å), and O5 atom (3.2 Å) of ΔGlcA, but not to glycosidic oxygen (4.2 Å); Wat is close to C5 carbon (3.4 Å) of ΔGlcA, but not to the C1 carbon atom (3.9 Å); Asp149 is close to the C4 atom (2.9 Å) of ΔGlcA, but not to glycosidic oxygen (5.8 Å); Asn88 at the mutation site (wild-type, Asp88) is not close to the C1 carbon (4.3 Å) of ΔGlcA and glycosidic oxygen (4.9 Å); no water molecule is present at the “β-face position” of ΔGlcA. These characteristics are common to crystal structures of UGL complexed with unsaturated hyaluronan disaccharide (ΔGlcA-GlcNAc) and unsaturated gellan tetrasaccharide (ΔGlcA-Glc-Rha-Glc) [18] (Fig. 3B, C). This arrangement of amino acid residues and water molecules is unusual in glycoside hydrolases [4,45].

The novel catalytic mechanism of UGL, deduced from its structural characteristics, consists of the following two reaction steps: (i) the water addition reaction (hydration) of the vinyl ether group (C4=C5–O5) in ΔGlcA and (ii) subsequent hydrolysis of the glycosidic bond between ΔGlcA and GalNAc. In the hydration reaction, Asp149 donates a proton to the double bond (C4 atom) of ΔGlcA as a general acid catalyst and then deprotonates Wat as a general base catalyst. The resulting deprotonated water molecule attacks the C5 atom of ΔGlcA. The product, a hemiketal, is unstable and readily converted to α-keto acid (hemiacetal) in an enzyme-independent manner. This hydration is partially caused by the electron-rich double bond in C4=C5, which acts as an “electron sink.” In the hydrolysis reaction, due to aldehyde-hemiacetal equilibrium, the hemiacetal is

Fig. 3. Interaction between unsaturated glycoside hydrolases and substrates (stereodiagram). (A) Bacillus sp. strain GL1 UGL (D88N) complexed with unsaturated chondroitin disaccharide (ΔGlcA-GalNAc). (B) Bacillus sp. strain GL1 UGL (D88N) complexed with unsaturated hyaluronan disaccharide (ΔGlcA-GlcNAc). (C) Bacillus sp. strain GL1 UGL (D88N) complexed with unsaturated gellan tetrasaccharide (ΔGlcA-Glc-Rha-Glc). (D) B. subtilis YteR (D143N) complexed with unsaturated RG disaccharide (ΔGalA-Rha). The bound substrate as well as the surrounding amino acid residues and water molecules (Wat, gray ball) interacting with the substrate or amino acid residues are shown. The substrate is represented by a yellow ball-and-stick model (red, oxygen atom; and cyan, nitrogen atom). Direct hydrogen bonds (≤3.2 Å), shown as dotted lines, are formed between the substrate and amino acid residues. (E) Superpositioning of active site structures of the bacillus UGL and YteR (magenta, UGL and blue, YteR).
finally converted to an aldehyde (4-deoxy-L-threo-5-hexosulose-uronate), leaving behind a saccharide through cleavage of the glycosidic bond. Asp88 probably stabilizes the carbocation at C5 of ΔGlcA during the catalytic reaction. GC/MS analysis of the reaction product in the presence of H$_2$O directly showed hydration of the vinyl ether group in ΔGlcA but not of the glycosidic bond [17]. The reaction product has $^{18}$O on the C5 atom but not on the C1 atom. Site-directed mutagenesis and kinetic studies support this catalytic mechanism and indicate that the proposed role of each amino acid residue in the active site is plausible [16–18].

The X-ray crystallographic and GC/MS studies provide evidence of a novel catalytic mechanism of UGL, i.e., the enzyme triggers hydration of the vinyl ether group in ΔGlcA but not of the glycosidic bond to be cleaved. Thus, UGL is clearly unlike general glycoside hydrolases, in which the catalytic mechanism consists of retention or inversion of an anomeric configuration.

The tertiary structure of the YteR mutant (D143N, Asp143 to Asn) complexed with the substrate (ΔGalA-Rha) was also determined by X-ray crystallography [21] (Fig. 2C). Similar to UGL, the enzyme accommodates the substrate in the deep pocket at the center of the molecule. ΔGalA interacts with residues His40, Tyr41, Asp88, His132, Trp141, Asn(Asp)143, Met147, Arg213, Trp217, Gly332, Thr333, and His352, while Rha is bound to residues His40, Tyr41, His42, and Gly332 (Fig. 3D). Ionizable residues His132, Asn(Asp)143, His189, and Arg213 interact with the substrate and are conserved in YteR and its homologs. The arrangement of amino acid residues at the active site is similar for YteR and UGL. Asp143 in YteR corresponds to Asp149 in UGL, likewise for His189 to His193 and Arg213 to Arg221. The interaction between YteR and ΔGalA-Rha is as follows (Fig. 3D): Wat hydrogen bonds to Asn(Asp)143 (2.9 Å) but not to glycosidic oxygen; Wat is also close to the C5 carbon atom of ΔGalA (3.5 Å); Asn(Asp)143 is close to the C4 atom of ΔGalA (3.4 Å). These residues of YteR can be superimposed on those of UGL (Fig. 3E), indicating that UGH also catalyzes the first hydration reaction to the vinyl ether group in ΔGalA. The catalytic mechanism of YteR is as follows: Asp143 donates a proton to the double bond (C4 atom) of ΔGalA as a general acid catalyst. Asp143 then deprotonates Wat as a general base catalyst and the deprotonated Wat molecule attacks the C5 atom of ΔGalA. The product (hemiketal) is unstable and readily converted to α-keto acid (hemiacetal). Due to aldehyde-hemiacetal equilibrium, the resulting hemiacetal is finally converted to aldehyde (4-deoxy-L-threo-5-hexosulose-uronate) and saturated saccharide (Rha) through cleavage of the glycosidic bond.

**Substrate recognition.** In the complex structure of the UGL mutant (D88N) and ΔGlcA-GalNAc (Fig. 3A), there are fewer interactions between D88N and GalNAc than between D88N and ΔGlcA, indicating that the enzyme interacts strongly with ΔGlcA rather than with GalNAc through hydrogen bonds and van der Waals contacts. UGL acts on different unsaturated saccharides derived from chondroitin, hyaluronan, xanthan, and gellan, but these substrates without...
exception have ΔGlcA residues at their nonreducing termini. The structural feature of UGL binding to its substrates was also examined by determination of the crystal structures of UGL complexed with unsaturated hyaluronan disaccharide and unsaturated gellan tetrasaccharide [18] (Fig. 3B,C). Analysis of the structures of UGL complexed with different substrates showed that ΔGlcA is specifically recognized through the formation of several hydrogen bonds and stacking interactions by amino acid residues highly conserved in the GH-88 family of enzymes (Fig. 3A–C). Aromatic residues in particular are responsible for the stacking interactions (Fig. 3). The carboxyl group of ΔGlcA interacts with the enzyme through the formation of hydrogen bonds, and interaction is stabilized by the positive end of the inner α-helix dipole of the α/α6-barrel (Fig. 2B,C). This structural feature of UGL, crucial to strict recognition of the unsaturated uronic acid residue, is also found in YteR.

Unlike the bacillus and flavobacterial UGLs, streptococcal UGL acts on unsaturated chondroitin disaccharide containing a sulfate group at C4 of GalNAc. The crystal structure of S. agalactiae UGL was determined at 1.8 Å resolution by X-ray crystallography to clarify the difference in substrate specificity among UGLs [unpublished results]. The in silico model, based on this crystal structure and the coordinate of unsaturated chondroitin disaccharide 4-sulfate, indicated that Arg236 of streptococcal UGL specifically recognizes the sulfate group (Fig. 4). The bacillus and flavobacterial UGLs have a histidine residue in place of Arg236. The arginine residue is well-conserved in UGLs from S. agalactiae, S. pneumoniae, and S. pyogenes (Fig. 4). The binding ability of Arg236 to the sulfate group probably contributes to the broad substrate specificity of streptococcal UGLs.

**Fig. 5.** The α/α-barrel structure shared by carbohydrate-related enzymes. Magenta, UGL; cyan, YteR; yellow, AGE; and green, A1-III.

---

**Concluding remarks**

Most genes coding for unsaturated glycoside hydrolases UGL and UGH are encoded in the genome of microorganisms producing polysaccharide lyases, since hydrolases act on unsaturated saccharides released from polysaccharides through the reaction of polysaccharide lyases. Many microbial polysaccharide lyases attack extracellular matrices such as glycosaminoglycan or plant cell wall pectin. Mammalian pathogenic and/or symbiont bacteria, including bacteroides, clostridia, enterococci, lactobacilli, streptococci, and vibrios, predominate among UGL producers in the enzyme family GH-88. In enzyme family GH-105, plant pathogenic, endophytic, and/or plant symbiont bacteria, e.g., agrobacteria, Erwinia, enterobacteria, and rhizobia, are typical producers of UGH. Thus, UGL and UGH, together with polysaccharide lyases, play important roles in allowing bacteria to associate with mammals or plants and function as virulence factors for host cells. Inhibitors of polysaccharide lyases and unsaturated glycoside hydrolases may become potent pharmaceuticals for treating microbial infectious diseases, because most of these enzymes are produced by microorganisms but not by plants or mammals. Clarification of the structure/function relationships of polysaccharide lyases and unsaturated glycoside hydrolases is expected to facilitate the design of novel drugs targeting such diseases. Since the catalytic mechanism of UGL and UGH in particular differs substantially from that of general glycoside hydrolases, inhibitors of these enzymes should be effective without exerting side effects on mammalian and plant host cells.

The α/α-barrel structure is categorized as a α/α toroid fold in the SCOP database [http://scop.mrc-lmb.cam.ac.uk/scop/]. The α/α toroid-fold proteins are further divided into five superfamilies, among which the unsaturated glycoside hydrolases UGL and UGH belong to a superfamily of six-hairpin glycoside hydrolases. This superfamily includes carbohydrate-related enzymes such as glucoamylases (Protein Data Bank (PDB) entry, 1GAI), cellulases (PDB entry, 1KWF), maltose phosphorylase (PDB entry, 1H54), AGE (PDB entry, 1FP3), and alginate lyase A1-III (PDB entry, 1QAZ). We found that the basic scaffold, i.e., the α/α-barrel, of UGL, YteR, AGE, and A1-III, are well-superimposed (Fig. 5).
Structural similarity provides valuable information with which to clarify the intrinsic function of hypothetical proteins. Analogous to the identification of YteR as a UGH, *Salmonella typhimurium* hypothetical protein YihS, containing an α/α-barrel structure, has recently been identified as a novel carbohydrate-related enzyme that catalyzes isomerization and epimerization [manuscript in preparation]. Regarding the structural similarity observed in α/α-barrel proteins, we hypothesize that these carbohydrate-related enzymes developed from a common ancestral protein, although they catalyze different reactions, i.e., hydrolyase, phosphorylase, epimerase, or lyase, and no similarities among their primary structures have been found. Further study of the molecular evolution of enzymes is expected to reveal common structural features.

**Acknowledgements.** This work was supported in part by Grants-in-Aid (KM, BM, and WH) and the Targeted Proteins Research Program from the Ministry of Education, Culture, Sports, Science, and Technology of Japan.

**References**


