Title	Crystal structure of a bacterial unsaturated glucuronyl hydrolase with specificity for heparin.
Author(s)	Nakamichi, Yusuke; Mikami, Bunzo; Murata, Kousaku; Hashimoto, Wataru
Citation	The Journal of biological chemistry (2014), 289(8): 4787-4797
Issue Date	2014-02-21
URL	http://hdl.handle.net/2433/203065
Right	This research was originally published in [The Journal of Biological Chemistry, 289, 4787-4797, doi: 10.1074/jbc.M113.522573]. © the American Society for Biochemistry and Molecular Biology
Type	Journal Article
Textversion	author

Crystal structure of a bacterial unsaturated glucuronyl hydrolase with specificity for heparin*

Yusuke Nakamichi[1], Bunzo Mikami[2], Kousaku Murata[1,3], and Wataru Hashimoto[1]

[1]From the Laboratory of Basic and Applied Molecular Biotechnology, Division of Food Science and Biotechnology, Graduate School of Agriculture, Kyoto University, Uji, Kyoto 611-0011, Japan
[2]Laboratory of Applied Structural Biology, Division of Applied Life Sciences, Graduate School of Agriculture, Kyoto University, Uji, Kyoto 611-0011, Japan
[3]Present address
Faculty of Science and Engineering, Setsunan University, Neyagawa, Osaka 572-8508, Japan

*Running title: *Complete bacterial degradation of heparin*

To whom correspondence should be addressed: Wataru Hashimoto, Laboratory of Basic and Applied Molecular Biotechnology, Division of Food Science and Biotechnology, Graduate School of Agriculture, Kyoto University, Uji, Kyoto 611-0011, Japan, Tel.: +81-774-38-3756; Fax: +81-774-38-3767; E-mail: whasimot@kais.kyoto-u.ac.jp

Keywords: glycosaminoglycan; heparin; heparan sulfate; chondroitin; X-ray crystallography; enzyme structure; glycoside hydrolases; *Streptococcus*

Background: Bacterial unsaturated glucuronyl hydrolase (UGL) is essential for complete degradation of host glycosaminoglycans.
Results: Crystal structure of *Pedobacter heparinus* UGL, Phep_2830 specific for heparin degradation, was determined.
Conclusion: The pocket-like structure and lid loop of Phep_2830 are involved in heparin disaccharide recognition.
Significance: This work contributes to understanding of bacterial degradation of host extracellular matrix components.

ABSTRACT

Extracellular matrix molecules such as glycosaminoglycans (GAGs) are typical targets for some pathogenic bacteria, which allow adherence to host cells. Bacterial polysaccharide lyases depolymerize GAGs in β-elimination reactions, and the resulting unsaturated disaccharides are subsequently degraded to constituent monosaccharides by unsaturated glucuronyl hydrolases (UGLs). UGL substrates are classified as 1,3- and 1,4-types based on the glycoside bonds. Unsaturated chondroitin and heparin disaccharides are typical members of 1,3- and 1,4-types, respectively. Here we show the reaction modes of bacterial UGLs with unsaturated heparin disaccharides by X-ray crystallography, docking simulation, and site-directed mutagenesis. Although streptococcal and bacillus UGLs were active on unsaturated heparin disaccharides, those preferred 1,3- rather than 1,4-type substrates. The genome of GAG-degrading *Pedobacter heparinus* encodes 13 UGLs. Of these, Phep_2830 is known to be specific for unsaturated heparin disaccharides. The crystal structure of Phep_2830 was determined at 1.35 Å resolution. In comparison with structures of streptococcal and bacillus UGLs, a pocket-like structure and lid loop at subsite +1 are characteristic of Phep_2830. Docking simulations of Phep_2830 with unsaturated heparin disaccharides demonstrated that the direction of substrate pyranose rings differs from that in unsaturated chondroitin disaccharides. Acetyl groups of unsaturated heparin disaccharides are well accommodated in the pocket at subsite +1, and aromatic residues of the lid loop are required for stacking interactions with substrates. Thus, site-directed mutations of the pocket and lid loop led to significantly reduced enzyme

activity, suggesting that the pocket-like structure and lid loop are involved in the recognition of 1,4-type substrates by UGLs.

Glycosaminoglycan (GAG) is a heteropolysaccharide comprising uronic acid and amino sugar residues (1). Based on the sugar composition and mode of glycosidic bonds (i.e., 1,3- and 1,4-types), GAGs are categorized as hyaluronan, chondroitin sulfate, dermatan sulfate, heparan sulfate, and heparin (Fig. 1A, B). The polysaccharide chondroitin sulfate has repeating disaccharide units comprising β-D-glucuronic acid (GlcUA) and *N*-acetyl-D-galactosamine (GalNAc), which are linked via 1,3-glycosidic bonds. In contrast, heparin and heparan sulfate have only 1,4-type glycosidic bonds, which comprise repeating disaccharide units with uronic acid residues [α-L-iduronic acid (IdoUA) or β-GlcUA] and amino sugar residues [D-glucosamine (GlcN) or *N*-acetyl-D-glucosamine (GlcNAc)]. Constituent sugar residues of chondroitin sulfate and heparin have varying numbers of sulfate groups (2). With the exception of hyaluronan, these GAGs are present as core protein bound proteoglycans and play multiple roles in the architecture of extracellular matrices and cell growth and differentiation (3).

Adhesion of microbes to eukaryotic cells may be a primary mechanism for residence of normal flora and pathogenic infections. GAGs are typical microbial targets for interactions with host cells, and some specific interactions between microbes and these polysaccharides have been described (4,5). Such microbes include GAG-degrading bacteria. GAGs are depolymerized by a variety of bacterial polysaccharide lyases. Based on primary structures, these enzymes are categorized into several families (PL-1–22) in the Carbohydrate-Active enZymes (CAZy) database (6-10). GAG-degrading enzymes include hyaluronate and chondroitin ABC lyases (family PL-8), and heparinase I (PL-13), -II (PL-21), and -III (PL-12), which have differing sequence and substrate specificities (9,10). Polysaccharide lyases depolymerize GAGs through β-elimination reactions and generally produce unsaturated disaccharides containing unsaturated uronic acid with C–C double bonds at nonreducing termini (Fig. 1B) (11). Carbon atoms C-3, C-4, C-5, and C-6 of unsaturated GlcUA (ΔGlcUA) and IdoUA (ΔIdoUA) are located in a single plane because of the formation of double bonds between C-4 and C-5 (12). Because GlcUA is a C-5 epimer of IdoUA, ΔGlcUA and ΔIdoUA are identical in structure. Hence, both ΔGlcUA and ΔIdoUA are referred to as ΔGlcUA in this paper.

Unsaturated GAG disaccharides are degraded to monosaccharides by unsaturated glucuronyl hydrolase (UGL), which is classified as a member of the glycoside hydrolase family 88 in the CAZy database (Fig. 1B) (7,13). In contrast to other hydrolases, UGL recognizes a double bond in ΔGlcUA and triggers the hydration of C-5 (14,15). Because ΔGlcUA is a component of all unsaturated GAG oligosaccharides, UGLs are essential for the complete degradation of GAGs. To date, all bacterial UGLs are classified as members of the GH-88 family. However, different substrate specificities accommodate structural diversities of unsaturated GAG oligosaccharides with different sugar residues, glycosidic bonds, and degrees of sulfation. For example, streptococcal UGLs and *Bacillus* sp. GL1 enzymes (BacillusUGL) prefer sulfated and unsulfated unsaturated disaccharides with 1,3-glycosidic bonds, respectively (16,17). In contrast, Phep_2830, the UGL of *Pedobacter heparinus* (formerly known as *Flavobacterium heparinum*) degrades only unsaturated heparin disaccharides with 1,4-glycosidic bonds (18).

Recent studies have focused on the physiological functions and structures of UGLs and have revealed peculiar mechanisms of UGL catalysis using artificial substrates (15) and that UGL gene disruption leads to reduced upper respiratory tract colonization by *Streptococcus pneumoniae* (19). Specific inhibitors of UGL are therefore expected to provide anti-bacterial drugs with no side effects. In addition, we demonstrated inducible mRNA expression of the streptococcal enzyme in the presence of GAG (16) and identified structural determinants of the preference of streptococcal UGL for sulfated substrates with 1,3-glycosidic bonds (12). However, the enzyme recognition mechanism

for 1,4-glycosidic bond-type substrates from heparin and heparan sulfate remains unknown. Degradation of heparins and heparan sulfates with 1,4-glycosidic bonds is also considered important for bacterial adherence and invasion to host cells because heparin, heparan sulfate, hyaluronan, and chondroitin sulfate are constituents of mammalian extracellular matrices (20). In this study, we examined the crystal structure of a *P. heparinus* UGL that has specific activity for unsaturated disaccharides with 1,4-glycosidic bonds and demonstrated the binding modes of these substrates with bacterial UGLs by X-ray crystallography, docking simulations, site-directed mutagenesis, and measurements of mutant enzyme kinetics.

Experimental procedures

Materials- Unsaturated chondroitin disaccharides were purchased from Seikagaku Biobusiness (Tokyo, Japan). Unsaturated heparin disaccharides were purchased from Iduron (Manchester, UK) and Sigma-Aldrich (St. Louis, MO, USA). Restriction endonucleases and DNA-modifying enzymes were purchased from Toyobo (Osaka, Japan). All other analytical grade chemicals were obtained from commercial sources.

Overexpression- Overexpression systems for *P. heparinus* UGLs (Phep_2238, Phep_2649, and Phep_2830) were constructed in *Escherichia coli* cells as follows. Three *P. heparinus* UGL genes were amplified from the genome of *P. heparinus* by polymerase chain reaction (PCR). PCR was performed using *P. heparinus* genomic DNA as a template, synthetic oligonucleotide primers, and KOD-FX-neo polymerase (Toyobo). Forward and reverse primers for the Phep_2238 gene were 5′-GGCATATGAAAAGAAACCTGATTTTAAAAA-3′ and 5′-CCCTCGAGTTATTTCTGTAGTTTTTTATAACGC-3′; those for the Phep_2649 gene were 5′-GGCATATGAACATTATCAAGTCAAGTCC-3′ and 5′-CCCTCGAGTCATAAACCTCTTTCTTTTA-3′, and those for the Phep_2830 gene were 5′-GGCATATGAAATCACTACTCAGTGCGTTTG-3′ and 5′-CCCTCGAGTTAAGACTGATTAATTGTTTTC-3′. Restriction sites for *Nde*I and *Xho*I are shown as underlined 5′ regions. The PCR conditions were as follows: 94°C for 2 min followed by 30 cycles of 98°C for 10 s, 55°C for 30 s, and 68°C for 30 s. PCR products were ligated with *Hin*cII-digested pUC119 (Takara, Shiga, Japan) using Ligation High Ver. 2 (Toyobo) in accordance with the manufacturer's protocol, and the resulting plasmids were digested with *Nde*I and *Xho*I to isolate UGL gene fragments. Phep_2238 and Phep_2649 DNA fragments were ligated with *Nde*I- and *Xho*I-digested pET21b vectors (Novagen, Darmstadt, Germany), and the Phep_2830 DNA fragment was ligated into *Nde*I- and *Xho*I-digested pCold IV vector (Takara Bio, Shiga, Japan) using Ligation High Ver. 2. The resulting plasmids were designated pET21b-Phep_2238, pET21b-Phep_2649, and pCold IV-Phep_2830, respectively. To overexpress Phep_2238, Phep_2649, and Phep_2830, *E. coli* BL21(DE3) host cells (Novagen) were transformed with pET21b-Phep_2238 and pET21b-Phep_2649, and *E. coli* Rosetta-gami B host cells (Novagen) were transformed with pCold IV-Phep_2830.

Microorganisms and culture conditions- To express Phep_2238 or Phep_2649, *E. coli* BL21(DE3) cells harboring pET21b-Phep_2238 or pET21b-Phep_2649 plasmids were cultured at 37°C in Luria broth (Sigma-Aldrich) supplemented with sodium ampicillin (100 μg ml^{-1}). When turbidity at 600 nm reached 0.3–0.7, isopropyl-β-D-thiogalactopyranoside was added to the culture to a final concentration of 0.1 mM, and the cells were further cultured at 16°C for 44 h. To overexpress Phep_2830, *E. coli* Rosetta-gami B cells harboring pCold IV-Phep_2830 were cultured at 37°C in Luria broth supplemented with sodium ampicillin (100 μg ml^{-1}) and chloramphenicol (33 μg ml^{-1}). When turbidity at 600 nm reached 1–1.2, the cells were cooled to 15°C with ice water, isopropyl-β-D-thiogalactopyranoside was added to the culture to a final concentration of 0.4 mM, and the cells were further cultured at 15°C for 44 h.

Purification- Purification of *S. agalactiae* UGL (SagUGL), *S. pneumoniae* UGL

(SpnUGL), *S. pyogenes* UGL (SpyUGL), SagUGL mutant S368G (12)], and BacillusUGL were expressed and purified as described previously (12,16,21). *E. coli* cells harboring pET21b-Phep_2238, pET21b-Phep_2649, or pCold IV-Phep_2830 plasmids were collected by centrifugation at 6700 ×*g* for 10 min at 4°C. Cells harboring pET21b-Phep_2238 or pET21b-Phep_2649 were resuspended in 20 mM potassium phosphate buffer (KPB, pH 6.0), and the cells harboring pCold IV-Phep_2830 were resuspended in a solution containing 20 mM tris (hydroxymethyl) aminomethane–hydrochloric acid (Tris–HCl; pH 7.5), 1 mM dithiothreitol, and 1 mM ethylenediaminetetraacetic acid. Cells were ultrasonically disrupted (Insonator model 201M; Kubota, Osaka, Japan) at 9 kHz for 10 min at 0°C, and the supernatant obtained by centrifugation at 28,000 ×*g* for 20 min at 4°C was used as a cell extract. Phep_2238 and Phep_2649 were purified by cation exchange chromatography (TOYOPEARL CM-650, Tosoh, Tokyo, Japan) followed by gel filtration chromatography (HiLoad 16/60 Superdex 75 pg, GE Healthcare, Little Chalfont, UK). Phep_2830 was purified by anion exchange chromatography (HiLoad 16/10 Q Sepharose, GE Healthcare) followed by gel filtration chromatography (HiLoad 16/60 Superdex 75 pg). Purity was assessed using sodium dodecyl sulfate–polyacrylamide gel electrophoresis (22).

Enzyme assay- Reactions of SagUGL, SpnUGL, SpyUGL, and BacillusUGL were conducted at 30°C in 500 µl solutions of 20 mM Tris–HCl (pH 7.5), 0.2 mM substrate, and enzyme. Reactions of Phep_2238, Phep_2649, and Phep_2830 were conducted at 30°C in 500 µl solutions containing 100 mM KPB (pH 6.0), 0.2 mM substrate, and enzyme. Concentrations and pH of each buffer were adopted in accordance with previous reports (12,18). Enzyme activity was measured by monitoring decreases in absorbance at 235 nm, which corresponded to the loss of substrate C–C double bonds.

Unsaturated chondroitin disaccharides were used as substrates, and their molar absorption coefficients at 235 nm [ε_{235} (M^{-1} cm^{-1})] were as follows: ΔGlcUA-1,3-GalNAc (CΔ0S), ε_{235} = 4800; ΔGlcUA-1,3-GalNAc with a sulfate group at the C-4 position of GalNAc (CΔ4S), ε_{235} = 4800; and ΔGlcUA-1,3-GalNAc with a sulfate group at the C-6 position of GalNAc (CΔ6S), ε_{235} = 4800 (12). Unsaturated heparin disaccharides were also used as substrates, and their molar absorption coefficients at 235 nm were as follows: ΔGlcUA-1,4-GlcNAc (HΔNAc0S), ε_{235} = 4524; ΔGlcUA-1,4-GlcN with a sulfate group at the *N* position of GlcN (HΔNS), ε_{235} = 6600; ΔGlcUA-1,4-GlcN with a sulfate group at the C-6 position of GlcN (HΔ6S), ε_{235} = 4826; ΔGlcUA-1,4-GlcNAc with a sulfate group at the C-6 position of GlcNAc (HΔNAc6S), ε_{235} = 4300; ΔGlcUA-1,4-GlcN with sulfate groups at *N* and C-6 positions of GlcN (HΔNS6S), ε_{235} = 6075; ΔGlcUA-1,4-GlcN with sulfate groups at the C-2 position of ΔGlcUA and the *N* position of GlcN (HΔ2'SNS), ε_{235} = 4433 (18).

Kinetic parameters of SagUGLs (wild-type, S365A, S365G, S368A, S368G, and K370A), BacillusUGL, Phep_2238, Phep_2649, and Phep_2830s (wild-type and F164A) with various unsaturated GAG disaccharides were determined using assays of UGL activity at 30°C, in which decreases in substrate absorbance at 235 nm were monitored. Reaction mixtures of SagUGLs and BacillusUGL contained substrate, 20 mM Tris–HCl (pH 7.5), and enzyme. Reaction mixtures for Phep_2238, Phep_2649, and Phep_2830s contained substrate, 100 mM KPB (pH 6.0), and enzyme. Substrate concentrations ranges were fixed at 0.05–1.0 mM because the absorbance at 235 nm of the substrate at over 1.0 mM exceeds measurement limitations on the spectrometer. Michaelis–Menten constants (K_m) and turnover numbers (k_{cat}) were calculated by fitting the data using the *KaleidaGraph* software (Synergy Software, Reading, PA, USA).

Docking simulations- To perform docking analyses of UGLs and substrates, our coordinates of SagUGL were used from the Protein Data Bank (PDB, http://www.rcsb.org). The structure of ligand-free SagUGL (PDB ID: 2ZZR) was selected as a receptor model. Docking analyses were performed using the *AutoDock 4.2* program (23). The coordinates of CΔ6S were obtained from coordinates of

SagUGL complexed with CΔ6S (PDB ID: 3ANK), and coordinates of unsaturated heparin disaccharides were obtained using *ACD/ChemSketch Freeware*, version 5.12 (Advanced Chemistry Development, Inc., Toronto, ON, Canada) and the *OpenBabel* program (24). Asp-175 and Lys-370 of SagUGL were treated as flexible residues because Asp-175 of SagUGL was critical for catalysis and Lys-370 showed movement during interactions with CΔ6S (12,16). The number of genetic algorithm runs was set to 20. Figures for protein structures and docking forms were prepared using the *Pymol* program (25).

Site-directed mutagenesis- Three residues (Ser-365, Ser-368, and Lys-370) of SagUGL were substituted with Ala or Gly, Ala, and Ala, and the resulting mutants were designated S365A, S368A, S368A, and K370A, respectively. Asp-182 of Phep_2238, and Arg-57, Phe-164, Asp-174, and Gly-362 of Phep_2830 were also substituted with Asn, Ala, Ala, Asn, and Tyr, and the resulting mutants were designated D182N, R57A, F164A, D174N, and G362Y, respectively. With the exception of D174N and G362Y, UGL mutants were constructed using a Quik Change site-directed mutagenesis kit (Agilent Technologies, Santa Clara, CA, USA). D174N and G362Y were constructed using a KOD-Plus Mutagenesis Kit (Toyobo) in accordance with the manufacturer's protocol. The plasmids pET21b-SagUGL and pCold IV-Phep_2830 were used as PCR templates for SagUGL and Phep_2830, respectively, and synthetic oligonucleotides were used as sense and antisense primers. Sense and antisense primers sequences were as follows (mutation sites are underlined): S365A, 5′-CTATTGCACGGTGTGTATGCGTGGCATTCAGGTAAAGGAG-3′ and 5′-CTCCTTTACCTGAATGCCACGCATACACACCGTGCAATAG-3′; S365G, 5′-CTATTGCACGGTGTGTATGGGTGGCATTCAGGTAAAGGAG-3′ and 5′-CTCCTTTACCTGAATGCCACCCATACACACCGTGCAATAG-3′; S368A, 5′-CACGGTGTGTATTCGTGGCATGCAGGTAAAGGAGTAGATG-3′ and 5′-CATCTACTCCTTTACCTGCATGCCACGAATACACACCGTG-3′; K370A, 5′-GTATTCGTGGCATTCAGGTGCAGGAGTAGATGAAGG-3′ and 5′-CCTTCATCTACTCCTGCACCTGAATGCCACGAATAC-3′; D182N, 5’-GAATTTAAGGTGATCATTAACAATATGATGAACCTGGAAC-3’ and 5’-GTTCCAGGTTCATCATATTGTTAATGATCACCTTAAATTC-3’; R57A, 5’-CCAGGAATGAACCCAGCGTCTGTCAATCCGGACGGGACGG-3’ and 5’-CCGTCCCGTCCGGATTGACAGACGCTGGGTTCATTCCTGG-3’; F164A, 5’-GGCCATCCGCTCATGGGATGCCGGACACTGGCAATTTCCG-3’ and 5’-CGGAAATTGCCAGTGTCCGGCATCCCATGAGCGGATGGCC-3’; D174N, 5’-AACAACCTGATGAACCTGGAGTATTTATACTGGGCAGGAA-3’ and 5’-TATAATTACCGGAAATTGCCAGTGTCCGAAATCCCATGAG-3’; and G362Y, 5’-GCATTGCTGTACAATTCGGAAATCGATACACCTTTGAATT-3’ and 5’-ATATACACTATGTTTCAATATAAAAACTGGTTCTCGCCG-3’. PCR was performed using KOD-Plus-Neo polymerase (Toyobo). Mutations were confirmed using the dideoxy-chain termination method with the automated DNA sequencer model 3730xl (Applied Biosystems, Foster City, CA, USA) (26). *E. coli* HMS174 (DE3) host cells (Novagen) were transformed with SagUGL mutant plasmids. Expression and purification of SagUGL mutants were conducted using the procedure for wild-type SagUGL as described previously (12). Phep_2830 mutant plasmids were digested with *Nde*I and *Xho*I, and the resulting DNA fragments were ligated with *Nde*I and *Xho*I-digested pCold II vector (Takara Bio) to express Phep_2830 mutant proteins with N-terminal histidine tags for purification. N-terminal histidine tagged Phep_2830 reportedly exhibits comparable enzyme activity to that of the histidine tag-free enzyme (18). The Phep_2830 mutants R57A, F164A, and G362Y were purified to almost homogeneity by affinity chromatography (TALON, Clontech, Mountain View, CA, USA).

X-ray crystallography- To determine the three-dimensional structure of Phep_2830, the purified enzyme was crystallized using sitting drop

vapor diffusion. Solutions containing 1 µl proteins (10 mg ml^{-1}) in 20 mM Tris–HCl (pH 7.5), 1 mM dithiothreitol, and 1 mM ethylenediaminetetraacetic acid were mixed with equal volumes of reservoir solution containing 25% polyethylene glycol 6000, 0.1 M 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES; pH 7.5), and 0.1 M lithium chloride. Protein solutions were then incubated at 20°C, and single crystals were grown for about 2 months. The crystal was flash-cooled under a cold nitrogen stream. Diffraction data were collected at λ = 1 Å using an ADSC Q315 detector at the BL38B1 station of SPring-8, Harima, Japan. The data were processed using the *HKL2000* program (27). Molecular replacements for structure determinations were conducted using the *Molrep* program supplied in the *CCP4* program package with coordinates of SagUGL (PDB ID: 2ZZR) as an initial model (28,29). Structure refinement was conducted with the *Phenix.refine* program supplied in the *PHENIX* program package (30,31). Randomly selected 5% reflections were excluded from refinement and were used to calculate R_{free}. After each cycle of refinement, the model was manually adjusted using the *Coot* program (32). The final model quality was checked using the *PROCHECK* program (33). Superpositioning of protein models and calculation of their root mean square deviations (r.m.s.d.) were conducted using the *LSQKAB* program supplied with the *CCP4* program package (34).

Results and Discussion

Enzyme activity of streptococcal UGLs toward unsaturated heparin disaccharides- Three streptococcal UGLs (SagUGL, SpnUGL, and SpyUGL) were purified to homogeneity, and their enzyme activities (U mg^{-1}) toward a variety of unsaturated heparin disaccharides were measured (Table 1). All three streptococcal UGLs degraded HΔNS with greater efficiency than the other unsaturated heparin disaccharides. Kinetic parameters of SagUGL toward HΔNS were determined as follows: K_m, 1.9 mM; k_{cat}, 3.9 s^{-1}. SagUGL is known to exhibit the highest enzymatic activity toward CΔ6S with 1,3-glycosidic bond (12) as follows: K_m, 0.10 mM; k_{cat}, 10 s^{-1}, indicating that the enzyme activity of SagUGL toward unsaturated heparin disaccharides was lower than that toward unsaturated chondroitin disaccharides. Indeed, k_{cat} K_m^{-1} with HΔNS was 50-fold lower than with CΔ6S (12). The enzyme activities of BacillusUGL with each unsaturated heparin disaccharide were also measured. BacillusUGL exhibited the highest enzyme activity toward HΔNAc0S distinct from streptococcal UGLs (Table 1). BacillusUGL also demonstrated specificity for unsulfated unsaturated chondroitin disaccharides (17). These results and our previous reports (16,17) demonstrate that streptococcal UGLs exhibit high enzyme activity toward disaccharides with specific sulfate groups, whereas BacillusUGL preferentially degrades unsulfated substrates.

Recognition of unsaturated heparin disaccharides by streptococcal UGL- Structure determinations of SagUGL in complex with HΔNS were difficult, possibly due to its lower affinity for this disaccharide. Thus, to elucidate the binding modes of unsaturated heparin disaccharide with SagUGL, the structure of SagUGL in complex with HΔNS was predicted using the docking simulation program *AutoDock*. To evaluate the suitability of the *AutoDock* program, docking simulations were performed with our coordinates of ligand-free SagUGL (PDB ID: 2ZZR) and CΔ6S. The simulated complex structure with the lowest binding energy was almost identical to the crystal structure of SagUGL/CΔ6S except that Lys-370 forms hydrogen bond with a sulfate group (PDB ID: 3ANK; Fig. 2A, B), indicating that the *AutoDock* docking simulation accurately determined the substrate-bound UGL structure. Thus, the binding mode of HΔNS to SagUGL was calculated using an *AutoDock* program, indicating the accommodation of HΔNS in the active site of SagUGL (Fig. 2C). The direction of the pyranose ring of GlcN toward ΔGlcUA of HΔNS opposed that of GalNAc in CΔ6S (Fig. 1C). Due to these differences in direction, the amino group of GlcN in HΔNS is located at a similar position to the C-6 of GalNAc in CΔ6S, and the position of a sulfate group in HΔNS corresponds to that of the sulfate group in CΔ6S. The directions of HΔNS pyranose rings indicated by docking simulations were similar to

those of heparin in solution, which were determined by X-ray scattering (35). This docking simulation suggests that Ser-365, Ser-368, and Lys-370 residues of SagUGL form hydrogen bonds with the sulfate group of HΔNS. To confirm the functions of Ser-365, Ser-368, and Lys-370 of SagUGL, the SagUGL mutants S365A, S365G, S368A, S368G, and K370A were purified and assayed. Kinetic parameters of each mutant were determined and compared with those of wild-type SagUGL (Table 2). K_m values of S365G, S365A, S368G, and S368A for HΔNS were higher than that of the wild-type enzyme, whereas that of K370A was lower. The k_{cat} values of S365G, S368G, and in particular K370A for HΔNS were also decreased (k_{cat} of K370A was approximately 50-fold lower than that of WT). This mutant analysis shows that Ser-365 and Ser-368 of SagUGL are involved in binding to HΔNS, whereas Lys-370 enhances the reaction efficiency of SagUGL rather than the binding affinity for HΔNS. In contrast, although the binding modes of the other unsaturated heparin disaccharides to SagUGL were calculated, no structures of substrate-bound SagUGL were obtained (data not shown). These docking simulations and enzyme assays of SagUGL (Table 1) suggest that, with the exception of HΔNS, its affinity toward unsaturated heparin disaccharides is also remarkably low.

Docking simulations of HΔNS with SagUGL, site-directed mutagenesis, and kinetic analyses demonstrate that Ser-365 and Ser-368 residues of SagUGL contribute to recognition of the sulfate group of HΔNS and that the Lys-370 residue is involved in degradation of the substrate. These amino acid residues constitute the motif SXXSXK, which plays an important role in the recognition of the sulfate group of CΔ6S (12). Streptococcal UGLs acted on unsaturated heparin disaccharides, although their enzyme activities toward these substrates were low. In addition, putative heparan sulfate lyase is encoded in the vicinity of the streptococcal UGL gene, suggesting that streptococcal UGLs contribute to the degradation of heparan sulfate.

Substrate specificity of P. heparinus *UGLs*- The structural determinants of SagUGL recognition of HΔNS sulfate groups were calculated using docking simulations. However, the intrinsic UGL-binding mechanism of substrates containing 1,4-glycosidic bonds remains to be clarified. Therefore, the mechanisms behind UGL recognition of unsaturated heparin disaccharides were analyzed using high affinity UGL interactions with unsaturated heparin disaccharides. One (Phep_2830) of the UGLs of *P. heparinus*, which assimilates heparin as a carbon source, is known to specifically degrade substrates containing 1,4-glycosidic bonds. Thirteen UGL genes of *P. heparinus* were assigned by complete genome sequences (36) and the three UGLs Phep_2238, Phep_2649, and Phep_2830 were overexpressed in *E. coli* cells and were purified to homogeneity and assayed for activity (Table 1).

Phep_2238 degraded HΔNAc0S and HΔNS more efficiently than other unsaturated heparin disaccharides. Moreover, Phep_2238 exhibited comparably lower enzyme activity toward substrates with 1,3-glycosidic bonds, such as CΔ0S, CΔ4S, and CΔ6S, than those with 1,4-glycosidic bonds. The kinetic parameters K_m and k_{cat} of Phep_2238 toward HΔNS were 0.073 mM and 2.3 s^{-1}, respectively. Phep_2649 preferred CΔ6S but also degraded CΔ0S, HΔNAc0S, and HΔNS, although its specific activity was lower than that of the streptococcal UGLs, BacillusUGL, Phep_2238, and Phep_2830. K_m and k_{cat} values of Phep_2649 toward HΔNS were 0.052 mM and 0.17 s^{-1}, respectively. Although a UGL that degrades unsaturated chondroitin disaccharide (CΔ6S) was previously isolated as a 1,3-glycuronidase from *P. heparinus* (37), this enzyme differs from Phep_2238 and Phep_2649 in amino acid composition, isoelectric point, and molecular weight. In contrast, Phep_2830 exhibited enzyme activity toward only unsaturated heparin disaccharides, as described previously (18). K_m and k_{cat} values of Phep_2830 toward HΔNAc6S were 0.029 mM and 16 s^{-1}, respectively.

Crystal structure of Phep_2830- Because Phep_2238 and Phep_2830 showed high enzyme activity toward various unsaturated heparin disaccharides, these enzymes were

crystallized to clarify the mechanisms by which UGL recognizes substrates with 1,4-glycosidic bonds based on tertiary structures. Phep_2830 was successfully crystallized, and X-ray diffraction data were collected. The crystal structure of Phep_2830 was determined at a resolution of 1.35 Å using molecular replacement with the SagUGL structure (PDB ID: 2ZZR) as an initial model. Data collection and model refinement statistics are summarized in Table 3. The final model contains one monomer enzyme from Gly-28 to Thr-398, a molecule of HEPES, and 422 water molecules. The N-terminal region from Met-1 to Asn-27 could not be assigned because the electron density map was too thin. Similar to SagUGL and BacillusUGL, the overall structure of Phep_2830 has a α_6/α_6-barrel architecture, and Phep_2830 contains 12 α helices and 5 β strands (Fig. 3A). The r.m.s.d. for all of the 336 C^{α} atoms between Phep_2830 and SagUGL was 1.7 Å, indicating that both have a common basic scaffold structure. In contrast, approximately 10 C-terminal amino acid residues of Phep_2830 protrude forward from the outside of the protein. Previous studies show that Asp-149 of BacillusUGL and Asp-175 of SagUGL act as critical catalysts (14,16). Thus, to investigate the catalytic mechanisms of *P. heparinus* UGLs, corresponding residues of Phep_2238 (Asp-182) and Phep_2830 (Asp-174) were substituted with Asn. Both mutant enzymes were inactive, indicating that the catalytic mechanisms of Phep_2238 and Phep_2830 are similar to those of SagUGL and BacillusUGL.

Active sites were structurally compared by superimposing coordinates of Phep_2830 on those of SagUGL (PDB ID: 2ZZR) and BacillusUGL/CΔ0S (PDB ID: 2AHG; Fig. 3B–D). Subsites were defined such that −n represents the nonreducing terminus, +n represents the reducing terminus, and cleavage occurs between these sites (38). The amino acid residues and their positions at subsite −1, which is the binding site of ΔGlcUA, were common to all three UGLs (Fig. 3B). However, the structures at subsite +1, which is the binding site of an amino sugar, differed significantly (Fig. 3C, D). In particular, Ser-365, Ser-368, and Lys-370 of SagUGL comprise the motif SXXSXK, which contributes to the recognition of a sulfate group (12). These three residues are not conserved in Phep_2830 and correspond to the Ala-363, Tyr-366, and Ser-368 residues of Phep_2830, respectively. These observations indicate that Phep_2830 recognition of substrate sulfate groups differs from that of SagUGL. Regarding hydrophobic amino acid residues forming a stacking interaction with an amino sugar at the subsite +1, the Tyr-338 residue of BacillusUGL (or the Tyr-364 residue of SagUGL), which is located around C-6 position of GalNAc in BacillusUGL/CΔ0S is substituted with the Gly-362 residue in Phep_2830, while the Trp-134 residue of BacillusUGL (or the Trp-161 residue of SagUGL) is conserved in Phep_2830. Due to this difference, this large space in Phep_2830 occurs around Gly-362, and is distinct from that of SagUGL and BacillusUGL. As described above, the Phe-164 residue of Phep_2830 that corresponds to Glu-163 of SagUGL or Pro-136 of BacillusUGL also lies at the binding site of an amino sugar (Fig. 3C, D). The loop comprising amino acid residues 163–169 of Phep_2830, which is designated loop A and corresponds to residues 162–170 of SagUGL and 135–144 of BacillusUGL, covers the active site. However, these loops of both SagUGL and BacillusUGL are distal from those at the active site. Accordingly, the Phe-164 residue of Phep_2830 is proximal to the active site (Fig. 3D). These Phep_2830 specific amino acid residues may be involved in stacking interactions with substrates, especially with amino sugars. Indeed, whereas Arg-57, Arg-66, and Glu-369 residues of Phep_2830 are arranged at subsite +1, these are not conserved in the active sites of either SagUGL or BacillusUGL.

Binding mode of unsaturated heparin disaccharides to Phep_2830- To clarify the mechanism by which Phep_2830 recognizes unsaturated heparin disaccharides, we attempted to prepare complexes of Phep_2830 with these substrates but failed. Thus, the binding modes of unsaturated heparin disaccharides (HΔNAc0S, HΔNS, HΔ6S, and HΔNAc6S) to Phep_2830 were estimated in a similar way to those of SagUGL using the *AutoDock* program. In these simulations, structures of Phep_2830-bound

HΔNAc0S, HΔNS, and HΔNAc6S were successfully calculated and that of the Phep_2830–HΔNAc0S complex is shown in Fig. 2D. Two common features were observed in the three docking structures that exhibited low binding energies. Similar to the docking simulation of SagUGL with HΔNS, the direction of the pyranose ring of GlcNAc (or GlcN sulfated at the *N* position) toward ΔGlcUA was opposed to that of GalNAc in complex with SagUGL/CΔ6S and BacillusUGL/CΔ0S (Fig. 1C). In addition, the acetyl groups of HΔNAc0S and HΔNAc6S (or the sulfate group of HΔNS) were predicted to be accommodated in the pocket-like structure comprising Arg-57, Arg-66, Trp-73, Gly-362, Ala-363, Tyr-366, Ser-368, and Glu-369 residues, which was designated the acetyl/sulfate group-binding pocket (Fig. 2D, E). No similar pocket-like structures were observed in SagUGL and BacillusUGL, which prefer substrates with 1,3-glycosidic bonds (Fig. 3C). The Trp-162 and Phe-164 residues in loop A of Phep_2830 (Fig. 3D) are located close to the C-6 position of an amino sugar of unsaturated heparin disaccharides. As described above, Trp-161, Trp-134, and Trp-162 residues of SagUGL, BacillusUGL, and Phep_2830, respectively, are situated at almost identical positions, whereas the residues of SagUGL and BacillusUGL that correspond to Phe-164 of Phep_2830 are distal from subsite +1.

To confirm that this pocket-like structure and loop A contribute to recognition of unsaturated heparin disaccharides, Arg-57, Phe-164, and Gly-362 residues of Phep_2830 were substituted with Ala, Ala, and Tyr, respectively. The resulting mutants R57A, F164A, and G362Y were expressed in *E. coli* with N-terminal His-tags and were purified using affinity columns. Although their expression was confirmed by SDS-PAGE (Fig. 4), two mutants of pocket-like structure (R57A and G362Y) exhibited no detectable enzyme activity with unsaturated heparin disaccharides as well as unsaturated chondroitin disaccharides. The K_m value (1.4 mM) of F164A toward HΔNAc6S was about 50 fold higher than that (0.029 mM) of the wild-type enzyme, whereas there was no significant difference in k_{cat} between wild-type enzyme (16 s^{-1}) and F164A (22 s^{-1}). These site-directed mutagenesis experiments indicated that the pocket-like structure and loop A play a significant role in recognizing substrates. In contrast, the absence of the acetyl group may have led to low specific activity of Phep_2830 toward HΔ6S and hampered docking simulations of the complex of Phep_2830 with HΔ6S. Hence, substrates containing 1,4-glycosidic bonds are readily inserted into the acetyl/sulfate group-binding pocket of Phep_2830 through stacking interactions of loop A residues, Phe-164 and Trp-162, with an amino sugar. Simultaneously, Trp-162 and Phe-164 residues may inhibit binding of Phep_2830 with substrates containing 1,3-glycosidic bonds.

Comparisons of the active site structures and docking simulations suggest that the acetyl/sulfate group-binding pocket comprising Arg-57, Arg-66, Trp-73, Gly-362, Ala-363, Ser-368, and Glu-369, and the loop A including Trp-162 and Phe-164, are important for binding of Phep_2830 to unsaturated heparin disaccharides. Among these amino acid residues, Arg-57, Trp-73, Gly-362, Glu-369, and Trp-162 are conserved in Phep_2238 and Phep_2649 (Fig. 5). Moreover, Arg-66, Ala-363, Tyr-366, Ser-368, and Phe-164 of Phep_2830 correspond with Val-73, His-373, Gly-376, Ser-378, and Ser-171 of Phep_2238 and Arg-59, Ser-362, His-365, Asn-367, and Val-158 of Phep_2649, respectively. These differences in primary structure may lead to changes in pocket shapes but do not occupy the space. However, Ser-171 of Phep_2238 and Val-158 of Phep_2649, which correspond to Phe-164 of Phep_2830, do not form stacking interactions because they are smaller than Phe. Hence, these structural features may enable recognition of substrates with 1,3-glycosidic bonds by Phep_2238 and Phep_2649. Although the pocket-like structure of Phep_2649 was arranged with high probability, the activity of the enzyme for unsaturated heparin disaccharides was lower than that of Phep_2238 and Phep_2830. In addition, the relative length of amino acid residues of Phep_2649 that correspond to the loop A may contribute to low enzyme activity.

Bacteroides species, human intestinal bacteria, are known to produce family PL-12, 13,

and 21 heparinases depolymerizing heparin to unsaturated disaccharides through β-elimination reaction (39-43), while complete degradation of heparin to monosaccharides remains to be clarified. Four UGL-homologous proteins are encoded in the genomes of *Bacteroides stercoris* (BACSTE_02470, BACSTE_03202, BACSTE_03210, and BACSTE_03707) and *Bacteroides thetaiotaomicron* (BT_0146, BT_2913, BT_3348, and BT_4658). Arg-57 and Gly-362, key residues comprising the pocket-like structure of Phep_2830, are conserved in these bacteroides UGLs (Supplemental Fig. S1), suggesting that, similar to Phep_2238 and Phep_2830, bacteroides enzymes preferentially degrade unsaturated heparin disaccharides to constituent monosaccharides. Consequently, structural determinants of Phep_2830 for 1,4-specificity found in this study promotes a better understanding of the complete degradation of heparin by *Bacteroides* species.

In conclusion, tertiary and active site structures of UGLs that are specific for unsaturated heparin disaccharides containing 1,4-glycosidic bonds were determined for the first time. Comparisons of the active site structures, docking simulations, and site-directed mutagenesis experiments demonstrate the significance of the acetyl/sulfate group-binding pocket and the lid loop at subsite +1 in Phep_2830 for recognition of substrates with 1,4-glycosidic bonds.

Acknowledgements

We thank Drs. S. Baba and N. Mizuno from the Japan Synchrotron Radiation Research Institute (JASRI) for their kind help in data collection. Diffraction data for crystals were collected at the BL-38B1 station of SPring-8 (Hyogo, Japan), with the approval (2012B1265 and 2013A1106) of JASRI.

References

1. Ernst, S., Langer, R., Cooney, C. L., and Sasisekharan, R. (1995) Enzymatic degradation of glycosaminoglycans. *Crit. Rev. Biochem. Mol. Biol.* **30**, 387-444
2. Esko, J. D., and Selleck, S. B. (2002) Order out of chaos: Assembly of ligand binding sites in heparan sulfate. *Annu. Rev. Biochem.* **71**, 435-471
3. Gandhi, N. S., and Mancera, R. L. (2008) The structure of glycosaminoglycans and their interactions with proteins. *Chem. Biol. Drug Des.* **72**, 455-482
4. Sava, I. G., Zhang, F. M., Toma, I., Theilacker, C., Li, B. Z., Baumert, T. F., Holst, O., Linhardt, R. J., and Huebner, J. (2009) Novel interactions of glycosaminoglycans and bacterial glycolipids mediate binding of enterococci to human cells. *J. Biol. Chem.* **284**, 18194-18201
5. Baron, M. J., Bolduc, G. R., Goldberg, M. B., Auperin, T. C., and Madoff, L. C. (2004) Alpha C protein of group B *Streptococcus* binds host cell surface glycosaminoglycan and enters cells by an actin-dependent mechanism. *J. Biol. Chem.* **279**, 24714-24723
6. Shim, K. W., and Kim, D. H. (2008) Cloning and expression of chondroitinase AC from *Bacteroides stercoris* HJ-15. *Protein Expr. Purif.* **58**, 222-228
7. Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V., and Henrissat, B. (2009) The Carbohydrate-Active EnZymes database (CAZy): an expert resource for glycogenomics. *Nucleic Acids Res.* **37**, D233-D238
8. Li, S. L., Kelly, S. J., Lamani, E., Ferraroni, M., and Jedrzejas, M. J. (2000) Structural basis of hyaluronan degradation by *Streptococcus pneumoniae* hyaluronate lyase. *EMBO J.* **19**, 1228-1240
9. Sasisekharan, R., Bulmer, M., Moremen, K. W., Cooney, C. L., and Langer, R. (1993) Cloning and expression of heparinase-I gene from *Flavobacterium heparinum*. *Proc. Natl. Acad. Sci. USA* **90**, 3660-3664
10. Su, H. S., Blain, F., Musil, R. A., Zimmermann, J. J. F., Gu, K. F., and Bennett, D. C. (1996) Isolation and expression in *Escherichia coli* of hepB and hepC, genes coding for the glycosaminoglycan-degrading enzymes heparinase II and heparinase III, respectively, from *Flavobacterium heparinum*. *Appl. Environ. Microbiol.* **62**, 2723-2734
11. Linhardt, R. J., Avci, F. Y., Toida, T., Kim, Y. S., and Cygler, M. (2006) CS lyases: structure, activity, and applications in analysis and the treatment of diseases. *Adv. Pharmacol.* **53**, 187-215
12. Nakamichi, Y., Maruyama, Y., Mikami, B., Hashimoto, W., and Murata, K. (2011) Structural determinants in streptococcal unsaturated glucuronyl hydrolase for recognition of glycosaminoglycan sulfate groups. *J. Biol. Chem.* **286**, 6262-6271
13. Mori, S., Akao, S., Nankai, H., Hashimoto, W., Mikami, B., and Murata, K. (2003) 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**, 77-84
14. Itoh, T., Hashimoto, W., Mikami, B., and Murata, K. (2006) Crystal structure of unsaturated glucuronyl hydrolase complexed with substrate. *J. Biol. Chem.* **281**, 29807-29816
15. Jongkees, S. A. K., and Withers, S. G. (2011) Glycoside cleavage by a new mechanism in unsaturated glucuronyl hydrolases. *J. Am. Chem. Soc.* **133**, 19334-19337
16. Maruyama, Y., Nakamichi, Y., Itoh, T., Mikami, B., Hashimoto, W., and Murata, K. (2009) Substrate specificity of streptococcal unsaturated glucuronyl hydrolases for sulfated glycosaminoglycan. *J. Biol. Chem.* **284**, 18059-18069
17. Hashimoto, W., Kobayashi, E., Nankai, H., Sato, N., Miya, T., Kawai, S., and Murata, K. (1999) Unsaturated glucuronyl hydrolase of *Bacillus* sp. GL1: Novel enzyme prerequisite for metabolism of unsaturated oligosaccharides produced by polysaccharide lyases. *Arch. Biochem. Biophys.* **368**, 367-374
18. Myette, J. R., Shriver, Z., Kiziltepe, T., McLean, M. W., Venkataraman, G., and Sasisekharan, R. (2002) Molecular cloning of the heparin/heparan sulfate Δ4,5 unsaturated glycuronidase from *Flavobacterium heparinum*, its recombinant expression in *Escherichia coli*, and biochemical determination of its unique substrate specificity. *Biochemistry* **41**, 7424-7434

19. Marion, C., Stewart, J. M., Tazi, M. F., Burnaugh, A. M., Linke, C. M., Woodiga, S. A., and King, S. J. (2012) *Streptococcus pneumoniae* can utilize multiple sources of hyaluronic acid for growth. *Infect. Immun.* **80**, 1390-1398
20. Esko, J. D., and Lindahl, U. (2001) Molecular diversity of heparan sulfate. *J. Clin. Invest.* **108**, 169-173
21. Itoh, T., Akao, S., Hashimoto, W., Mikami, B., and Murata, K. (2004) 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**, 31804-31812
22. Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685
23. Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., and Olson, A. J. (2009) AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. *J. Comput. Chem.* **30**, 2785-2791
24. O'Boyle, N. M., Banck, M., James, C. A., Morley, C., Vandermeersch, T., and Hutchison, G. R. (2011) Open Babel: an open chemical toolbox. *J. Cheminform.* **3**, 33
25. DeLano, W. L. (2004) *The PyMOL Molecular Graphics System.*, DeLano Scientific LLC, San Carlos, CA
26. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**, 5463-5467
27. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307-326
28. Vagin, A. A., and Isupov, M. N. (2001) Spherically averaged phased translation function and its application to the search for molecules and fragments in electron-density maps. *Acta Crystallogr. D. Biol. Crystallogr.* **57**, 1451-1456
29. Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. W., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A., and Wilson, K. S. (2011) Overview of the *CCP4* suite and current developments. *Acta Crystallogr. D. Biol. Crystallogr.* **67**, 235-242
30. Afonine, P. V., Grosse-Kunstleve, R. W., and Adams, P. D. (2005) The Phenix refinement framework. *CCP4 Newsletter* **42**, 8
31. Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) *PHENIX*: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 213-221
32. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of *Coot*. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 486-501
33. Laskowski, R. A., Macarthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *PROCHECK* - a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* **26**, 283-291
34. Kabsch, W. (1976) Solution for best rotation to relate two sets of vectors. *Acta Crystallogr. Sect. A* **32**, 922-923
35. Khan, S., Gor, J., Mulloy, B., and Perkins, S. J. (2010) Semi-rigid solution structures of heparin by constrained X-ray scattering modelling: new insight into heparin-protein complexes. *J. Mol. Biol.* **395**, 504-521
36. Han, C., Spring, S., Lapidus, A., Del Rio, T. G., Tice, H., Copeland, A., Cheng, J. F., Lucas, S., Chen, F., Nolan, M., Bruce, D., Goodwin, L., Pitluck, S., Ivanova, N., Mavromatis, K., Mikhailova, N., Pati, A., Chen, A., Palaniappan, K., Land, M., Hauser, L., Chang, Y. J., Jeffries, C. C., Saunders, E., Chertkov, O., Brettin, T., Goker, M., Rohde, M., Bristow, J., Eisen, J. A., Markowitz, V., Hugenholtz, P., Kyrpides, N. C., Klenk, H. P., and Detter, J. C. (2009) Complete genome sequence of *Pedobacter heparinus* type strain (HIM 762-3(T)). *Stand. Genomic Sci.* **1**, 54-62

37. Gu, K. N., Linhardt, R. J., Laliberte, M., Gu, K. F., and Zimmermann, J. (1995) Purification, characterization and specificity of chondroitin lyases and glycuronidase from *Flavobacterium heparinum*. *Biochem. J.* **312**, 569-577
38. Davies, G. J., Wilson, K. S., and Henrissat, B. (1997) Nomenclature for sugar-binding subsites in glycosyl hydrolases. *Biochem. J.* **321**, 557-559
39. Kim, W. S., Kim, B. T., Kim, D. H., and Kim, Y. S. (2004) Purification and characterization of heparin lyase I from *Bacteroides stercoris* HJ-15. *Journal of Biochemistry and Molecular Biology* **37**, 684-690
40. Hyun, Y. J., Lee, J. H., and Kim, D. H. (2010) Cloning, overexpression, and characterization of recombinant heparinase III from *Bacteroides stercoris* HJ-15. *Appl. Microbiol. Biotechnol.* **86**, 879-890
41. Hyun, Y. J., Lee, K. S., and Kim, D. H. (2010) Cloning, expression and characterization of acharan sulfate-degrading heparin lyase II from *Bacteroides stercoris* HJ-15. *J. Appl. Microbiol.* **108**, 226-235
42. Luo, Y. D., Huang, X. Q., and McKeehan, W. L. (2007) High yield, purity and activity of soluble recombinant *Bacteroides thetaiotaomicron* GST-heparinase I from *Escherichia coli*. *Arch. Biochem. Biophys.* **460**, 17-24
43. Dong, W., Lu, W. Q., McKeehan, W. L., Luo, Y. D., and Ye, S. (2012) Structural basis of heparan sulfate-specific degradation by heparinase III. *Protein & Cell* **3**, 950-961

Footnotes

*This work was supported in part by grants-in-aid from the Japan Society for the Promotion of Science (to K. M. and W. H.) and by the Targeted Proteins Research Program (to W. H.) from the Ministry of Education, Culture, Sports, Science, and Technology (MEXT) of Japan. This work was also supported in part by research fellowships from the Japan Society for the Promotion of Science for Young Scientists (to Y. N.).

The atomic coordinates and structure factors for Phep_2830 (code 3WIW) have been deposited in the Protein Data Bank (http://www.rcsb.org/).

The abbreviations used are: GAG, glycosaminoglycan; GlcUA, D-glucuronic acid; GalNAc, *N*-acetyl-D-galactosamine; IdoUA, L-iduronic acid; GlcN, D-glucosamine; GlcNAc, *N*-acetyl-D-glucosamine; CAZy, The Carbohydrate-Active enZyme; ΔGlcUA, 4,5-unsaturated GlcUA; ΔIdoUA, 4,5-unsaturated IdoUA; UGL, unsaturated glucuronyl hydrolase; BacillusUGL, UGL of *Bacillus* sp. GL1; SagUGL, UGL of *Streptococcus agalactiae*; SpnUGL, UGL of *S. pneumoniae*; SpyUGL, UGL of *S. pyogenes*; PCR, polymerase chain reaction; Tris–HCl, tris (hydroxymethyl) aminomethane-hydrochloric acid; ε_{235}, molar absorption coefficient at 235 nm; CΔ0S, ΔGlcUA-1,3-GalNAc; CΔ4S, ΔGlcUA-1,3-GalNAc with a sulfate group at the C-4 position of GalNAc; CΔ6S, ΔGlcUA-1,3-GalNAc with a sulfate group at the C-6 position of GalNAc; HΔNAc0S, ΔGlcUA-1,4-GlcNAc; HΔNS, ΔGlcUA-1,4-GlcN with a sulfate group at the *N* position of GlcN; HΔ6S, ΔGlcUA-1,4-GlcN with a sulfate group at the C-6 position of GlcN; HΔNAc6S, ΔGlcUA-1,4-GlcNAc with a sulfate group at the C-6 position of GlcNAc; HΔNS6S, ΔGlcUA-1,4-GlcN with sulfate groups at *N* and C-6 positions of GlcN; HΔ2'S6S, ΔGlcUA-1,4-GlcN with sulfate groups at the C-2 position of ΔGlcUA and the *N* position of GlcN; PDB, Protein Data Bank; HEPES, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid.

Figure Legends

Figure 1. Degradation of GAGs
(A) Examples of GAGs containing 1,3-glycosidic bonds (hyaluronan) and 1,4-glycosidic bonds (heparan sulfate); (B) schematic of chondroitin degradation by lyase and UGL; (C) directions of pyranose rings of an unsaturated chondroitin disaccharide and an unsaturated heparin disaccharide in complex with UGLs.

Figure 2. Binding mode of unsaturated disaccharides to bacterial UGLs (stereo diagram)
(A) Our crystal structure of SagUGL in complex with CΔ6S determined by X-ray crystallography (PDB ID: 3ANK); (B) an active site structure of SagUGL/CΔ6S; (C) an active site structure of SagUGL/HΔNS; (D) an active site structure of Phep_2830/HΔNAc0S; (E) surface of the pocket recognizing acetyl group; (B)–(E) structures estimated via docking simulations; numbers of carbon atoms are shown (B) and (C). Atoms C of SagUGL, C of Phep_2830, O, N, and S are in yellow, purple, red, blue, and green, respectively.

Figure 3. Crystal structure of Phep_2830 (stereo diagram)
(A) Overall structure of Phep_2830. Blue, pink, and green denote the secondary structure elements α-helices, β-sheets, and loops, respectively; (B) amino acid residues of the three UGLs Phep_2830, SagUGL, and BacillusUGL/CΔ0S at subsite −1; (C) amino acid residues of the three UGLs at subsite +1; (D) each loop comprises amino acid residues 163–169 of Phep_2830, 162–170 of SagUGL, and 135–144 of BacillusUGL. Sticks indicate side chains of Trp-162 and Phe-164 of Phep_2830, Trp-161 of SagUGL, and Trp-134 of BacillusUGL. (B)–(D) Phep_2830, SagUGL (PDB ID: 2ZZR), and BacillusUGL in complex with CΔ0S (PDB ID: 2AHG) are indicated in purple, cyan, and yellow, respectively. Atoms O and N of CΔ0S are colored red and blue, respectively.

Figure 4. SDS-PAGE profile
Lane M, molecular weight standards (from top): synthetic polypeptides with molecular weights of 250,000, 150,000, 100,000, 75,000, 50,000, 37,000, 25,000, and 20,000; lane 1, R57A; lane 2, F164A; lane 3, G362Y.

Figure 5. Multiple sequence alignments of UGLs
Primary structures of Phep_2238, Phep_2649, Phep_2830, SagUGL, and BacillusUGL were used for sequence alignments. Identical and similar amino acid residues among proteins are denoted by asterisks and dots, respectively. Amino acid residues of the acetyl/sulfate group-binding pocket are indicated in boxes. The residues corresponding to loop A are shaded. Arrows indicate positions of Trp-162 and Phe-164 of Phep_2830.

Table 1
Substrate specificity of bacterial UGLs

	Specific activity (U mg^{-1})						
	SagUGL	SpnUGL	SpyUGL	BacillusUGL	Phep_2238	Phep_2649	Phep_2830
HΔNAc0S	0.016	0.0055	0.0033	5.3	12	0.30	15
HΔNS	0.77	0.14	0.10	1.8	11	1.1	7.1
HΔ6S	0.0088	0.0018	0.00094	0.081	1.2	ND[a]	1.9
HΔNAc6S	0.035	0.0059	0.0066	0.016	7.9	ND	18
HΔNS6S	0.0080	0.0031	ND	0.0050	4.2	ND	4.5
HΔ2'SNS	0.0023	0.0024	ND	0.0095	ND	ND	ND
CΔ0S	0.72	0.99	0.63	10	2.6	0.70	ND
CΔ4S	0.0019	ND	0.0046	ND	0.31	ND	ND
CΔ6S	17	6.1	5.3	1.8	3.2	1.7	ND

[a]Not detected

Table 2
Kinetic parameters of SagUGL mutants for HΔNS

	K_m (mM)	k_{cat} (s^{-1})	$k_{cat}\,K_m^{-1}$ ($mM^{-1}\,s^{-1}$)
WT	1.9 ± 0.2	3.8 ± 0.4	2.0 ± 0.3
S365A	3.4 ± 1.3	1.9 ± 0.6	0.56 ± 0.14
S365G	5.2 ± 2.7	0.76 ± 0.35	0.15 ± 0.07
S368A	3.0 ± 1.0	6.0 ± 1.7	2.0 ± 0.4
S368G	3.4 ± 0.7	2.8 ± 0.5	0.83 ± 0.05
K370A	0.86 ± 0.29	0.075 ± 0.014	0.086 ± 0.013

Table 3
Data collection and refinement statistics

Space group	*C*2
Unit cell parameters (Å, deg)	a = 91.84, b = 50.96, c = 85.82 β = 115.73
Data collection	
Wavelength (Å)	1.0000
Resolution limit (last shell)[a] (Å)	50-1.35 (1.40-1.35)[a]
Measured reflections	368,700
Unique reflections	78,862 (7,783)
Redundancy	4.7 (3.9)
Completeness (%)	99.9 (100)
I/σ (I)	32.5 (3.72)
R_{merge} (%)	0.091 (0.37)
Wilson *B*-factor (Å^2)	13.5
Refinement	
Resolution limit (Å)	45.64-1.35 (1.38-1.35)
Used reflections	78,821 (5,304)
Completeness (%)	99.7 (97.4)
R-factor (%)	13.2 (14.9)
R_{free} (%)	16.0 (21.4)
Final model	
No. of molecules /asymmetric unit	1
No. of nonhydrogen atoms	
Protein	3,244
HEPES	15
H_2O	422
Average isotropic *B*-factor (Å^2)	
Protein	16.7
HEPES	30.1
H_2O	27.5
r.m.s.d. from ideal	
Bond length (Å)	0.009
Bond angle (deg)	1.226
Ramachandran plot (%)	
Favored region	97.8
Allowed region	2.2
Outlier region	0

[a]Data in highest resolution shells are given in parentheses.

Figure 1

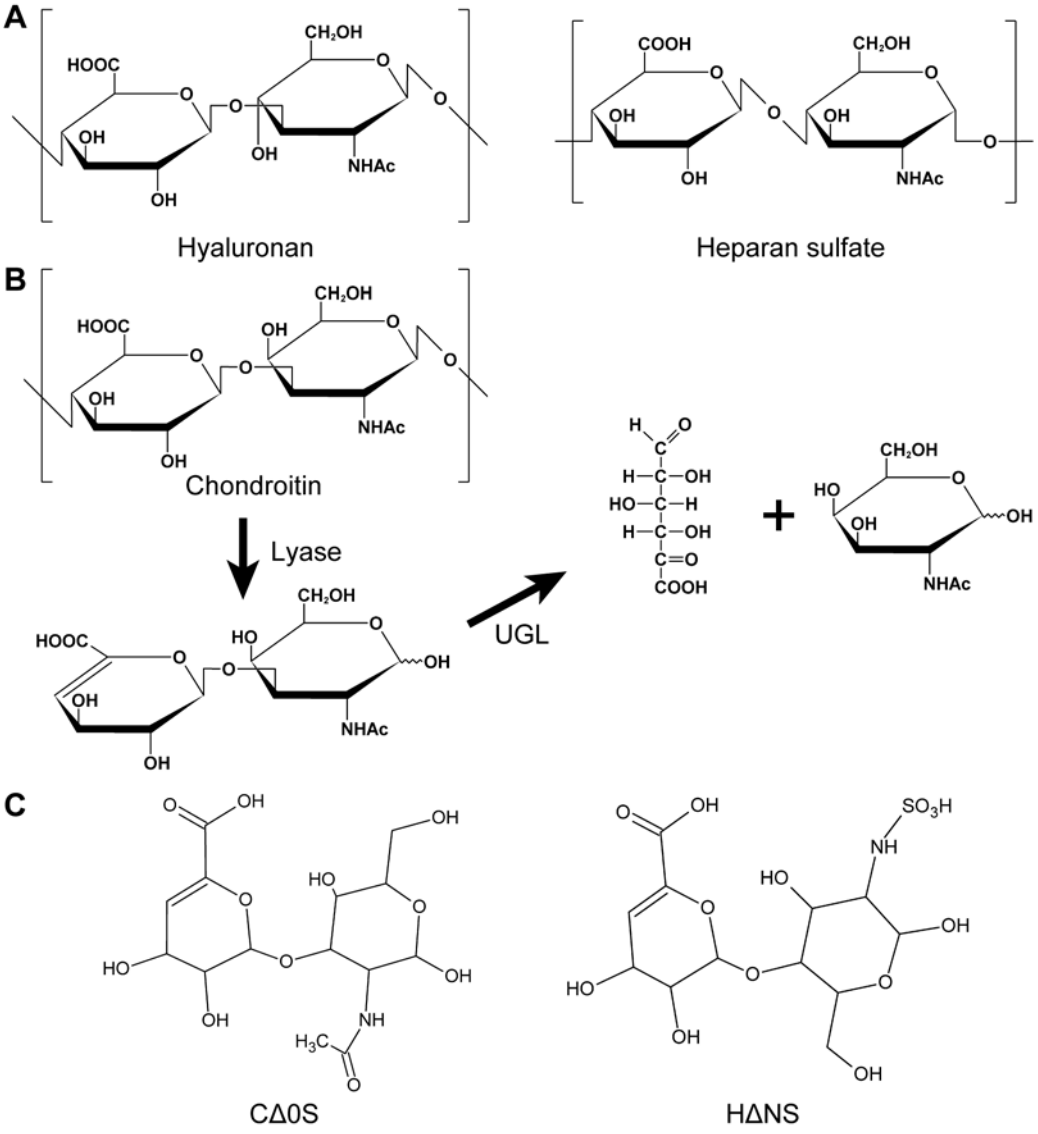

A
Hyaluronan
Heparan sulfate
B
Chondroitin
Lyase
UGL
C
CΔ0S
HΔNS

Figure 2

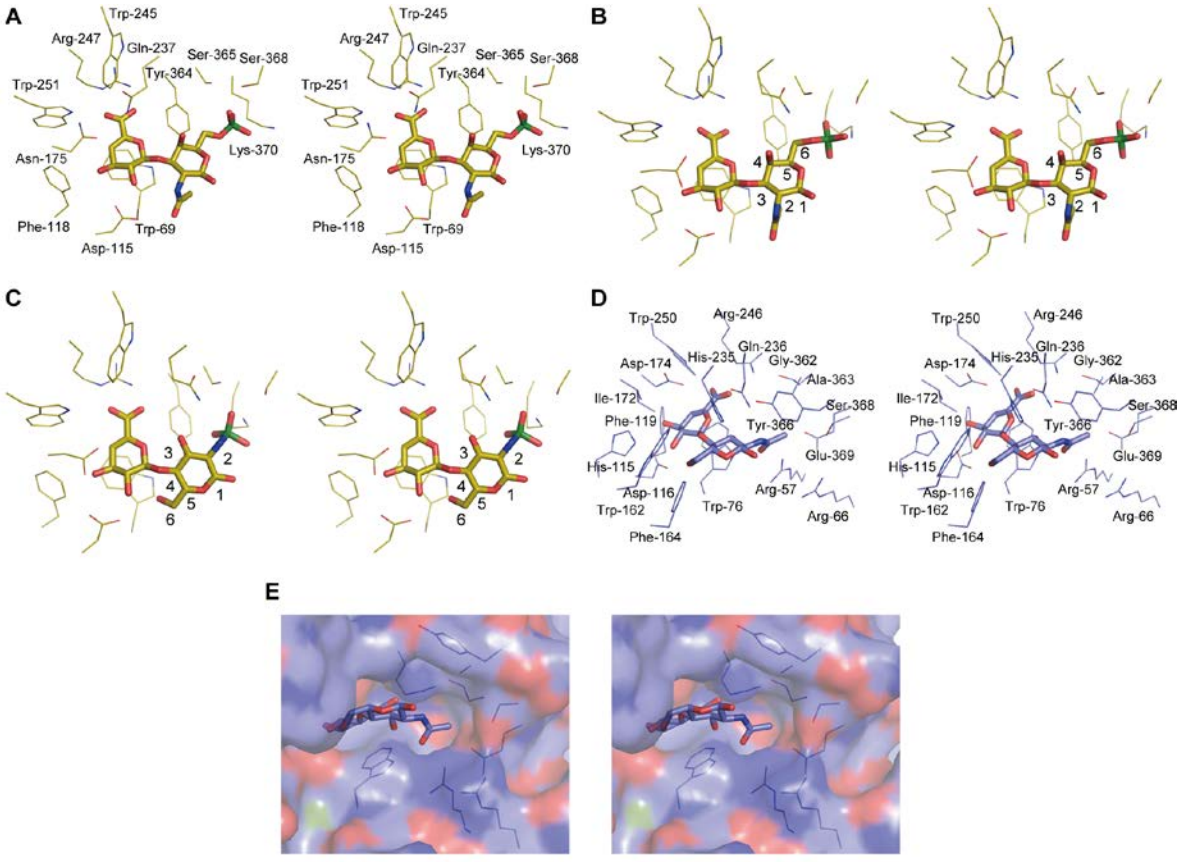

A
Trp-245
Arg-247
Gln-237
Ser-365
Ser-368
Tyr-364
Trp-251
Asn-175
Lys-370
Phe-118
Trp-69
Asp-115
B
C
D
Trp-250
Arg-246
Gln-236
Asp-174
His-235
Gly-362
Ala-363
Ile-172
Ser-368
Phe-119
Tyr-366
His-115
Glu-369
Asp-116
Arg-57
Trp-162
Trp-76
Arg-66
Phe-164
E

Figure 3

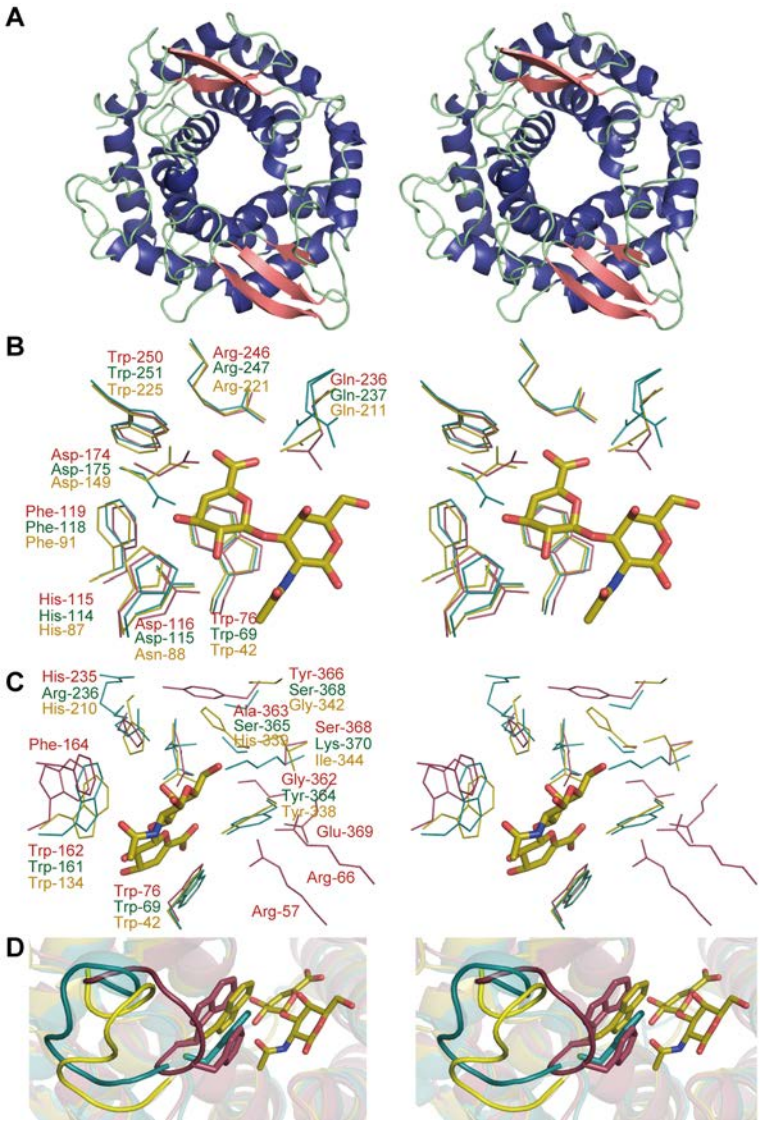

A
B
Trp-250
Trp-251
Trp-225
Arg-246
Arg-247
Arg-221
Gln-236
Gln-237
Gln-211
Asp-174
Asp-175
Asp-149
Phe-119
Phe-118
Phe-91
His-115
His-114
His-87
Asp-116
Asp-115
Asn-88
Trp-76
Trp-69
Trp-42
C
His-235
Arg-236
His-210
Tyr-366
Ser-368
Gly-342
Ala-363
Ser-365
His-339
Ser-368
Lys-370
Ile-344
Phe-164
Gly-362
Tyr-364
Tyr-338
Glu-369
Trp-162
Trp-161
Trp-134
Arg-66
Trp-76
Trp-69
Trp-42
Arg-57
D

Figure 4

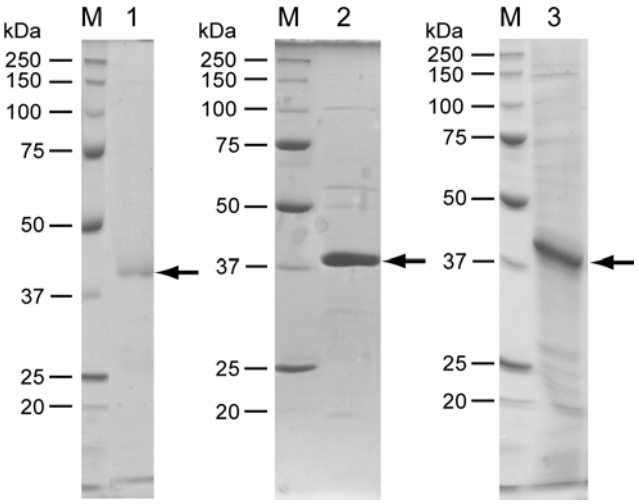
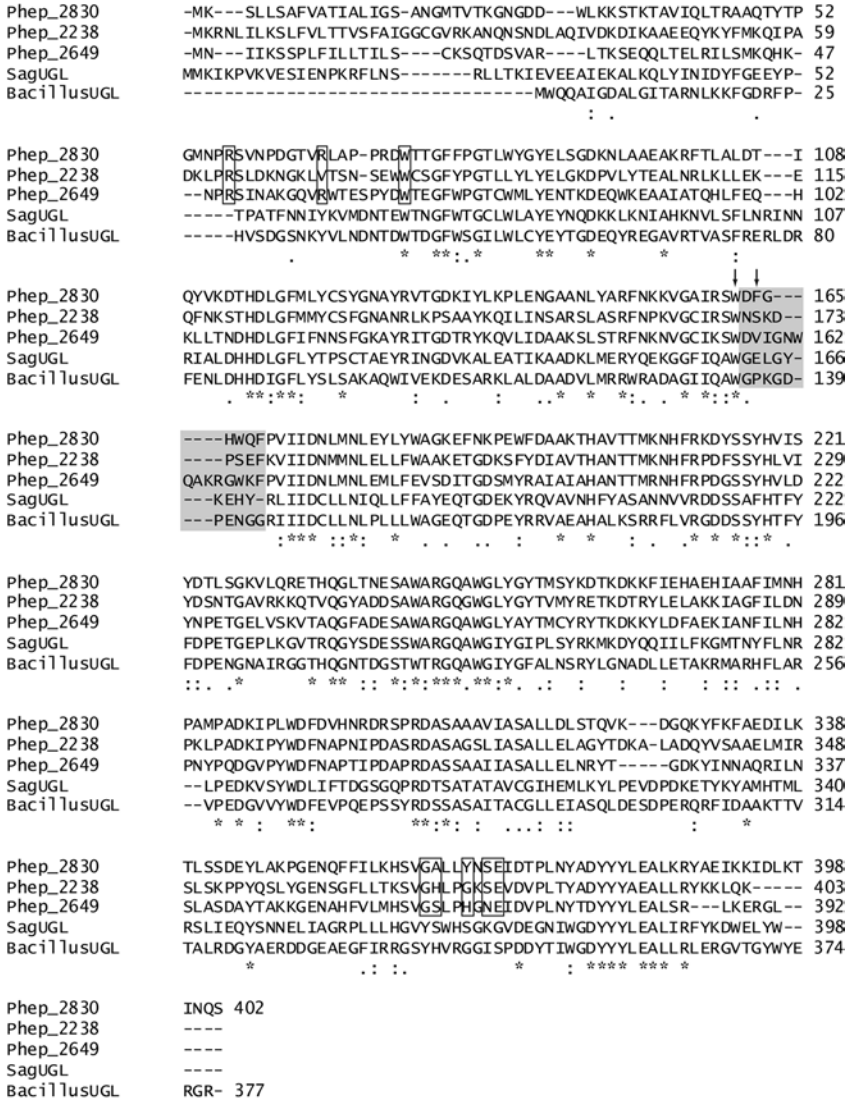

kDa
M
1
250
150
100
75
50
37
25
20
kDa
M
2
250
150
100
75
50
37
25
20
kDa
M
3
250
150
100
75
50
37
25
20

Figure 5


Phep_2830 -MK---SLLSAFVATIALIGS-ANGMTVTKGNGDD---WLKKSTKTAVIQLTRAAQTYTP 52
Phep_2238 -MKRNLILKSLFVLTTVSFAIGGCGVRKANQNSNDLAQIVDKDIKAAEEQYKYFMKQIPA 59
Phep_2649 -MN---IIKSSPLFILLTILS----CKSQTDSVAR----LTKSEQQLTELRILSMKQHK- 47
SagUGL MMKIKPVKVESIENPKRFLNS-------RLLTKIEVEEAIEKALKQLYINIDYFGEEYP- 52
BacillusUGL ----------MWQQAIGDALGITARNLKKFGDRFP- 25
Phep_2830 GMNPRSVNPDGTVRLAP-PRDWTTGFFPGTLWYGYELSGDKNLAAEAKRFTLALDT---I 108
Phep_2238 DKLPRSLDKNGKLVTSN-SEWWCSGFYPGTLLYLYELGKDPVLYTEALNRLKLLEK---E 115
Phep_2649 --NPRSINAKGQVRWTESPYDWTEGFWPGTCWMLYENTKDEQWKEAAIATQHLFEQ---H 102
SagUGL -----TPATFNNIYKVMDNTEWTNGFWTGCLWLAYEYNQDKKLKNIAHKNVLSFLNRINN 107
BacillusUGL -----HVSDGSNKYVLNDNTDWTDGFWSGILWLCYEYTGDEQYREGAVRTVASFRERLDR 80
Phep_2830 QYVKDTHDLGFMLYCSYGNAYRVTGDKIYLKPLENGAANLYARFNKKVGAIRSWDFG--- 165
Phep_2238 QFNKSTHDLGFMMYCSFGNANRLKPSAAYKQILINSARSLASRFNPKVGCIRSWNSKD-- 173
Phep_2649 KLLTNDHDLGFIFNNSFGKAYRITGDTRYKQVLIDAAKSLSTRFNKNVGCIKSWDVIGNW 162
SagUGL RIALDHHDLGFLYTPSCTAEYRINGDVKALEATIKAADKLMERYQEKGGFIQAWGELGY- 166
BacillusUGL FENLDHHDIGFLYSLSAKAQWIVEKDESARKLALDAADVLMRRWRADAGIIQAWGPKGD- 139
Phep_2830 ----HWQFPVIIDNLMNLEYLYWAGKEFNKPEWFDAAKTHAVTTMKNHFRKDYSSYHVIS 221
Phep_2238 ----PSEFKVIIDNMMNLELLFWAAKETGDKSFYDIAVTHANTTMKNHFRPDFSSYHLVI 229
Phep_2649 QAKRGWKFPVIIDNLMNLEMLFEVSDITGDSMYRAIAIAHANTTMRNHFRPDGSSYHVLD 222
SagUGL ---KEHY-RLIIDCLLNIQLLFFAYEQTGDEKYRQVAVNHFYASANNVVRDDSSAFHTFY 222
BacillusUGL ---PENGGRIIIDCLLNLPLLLWAGEQTGDPEYRRVAEAHALKSRRFLVRGDDSSYHTFY 196
Phep_2830 YDTLSGKVLQRETHQGLTNESAWARGQAWGLYGYTMSYKDTKDKKFIEHAEHIAAFIMNH 281
Phep_2238 YDSNTGAVRKKQTVQGYADDSAWARGQGWGLYGYTVMYRETKDTRYLELAKKIAGFILDN 289
Phep_2649 YNPETGELVSKVTAQGFADESAWARGQAWGLYAYTMCYRYTKDKKYLDFAEKIANFILNH 282
SagUGL FDPETGEPLKGVTRQGYSDESSWARGQAWGIYGIPLSYRKMKDYQQIILFKGMTNYFLNR 282
BacillusUGL FDPENGNAIRGGTHQGNTDGSTWTRGQAWGIYGFALNSRYLGNADLLETAKRMARHFLAR 256
Phep_2830 PAMPADKIPLWDFDVHNRDRSPRDASAAAVIASALLDLSTQVK---DGQKYFKFAEDILK 338
Phep_2238 PKLPADKIPYWDFNAPNIPDASRDASAGSLIASALLELAGYTDKA-LADQYVSAAELMIR 348
Phep_2649 PNYPQDGVPYWDFNAPTIPDAPRDASSAAIIASALLELNRYT-----GDKYINNAQRILN 337
SagUGL --LPEDKVSYWDLIFTDGSGQPRDTSATATAVCGIHEMLKYLPEVDPDKETYKYAMHTML 340
BacillusUGL --VPEDGVVYWDFEVPQEPSSYRDSSASAITACGLLEIASQLDESDPERQRFIDAAKTTV 314
Phep_2830 TLSSDEYLAKPGENQFFILKHSVGALLYNSEIDTPLNYADYYYLEALKRYAEIKKIDLKT 398
Phep_2238 SLSKPPYQSLYGENSGFLLTKSVGHLPGKSEVDVPLTYADYYYAEALLRYKKLQK----- 403
Phep_2649 SLASDAYTAKKGENAHFVLMHSVGSLPHGNEIDVPLNYTDYYYLEALSR---LKERGL-- 392
SagUGL RSLIEQYSNNELIAGRPLLLHGVYSWHSGKGVDEGNIWGDYYYLEALIRFYKDWELYW-- 398
BacillusUGL TALRDGYAERDDGEAEGFIRRGSYHVRGGISPDDYTIWGDYYYLEALLRLERGVTGYWYE 374
Phep_2830 INQS 402
Phep_2238 ----
Phep_2649 ----
SagUGL ----
BacillusUGL RGR- 377