Discovery of exolytic heparinases and their catalytic mechanism and potential application

Qingdong Zhang

Shandong University

Hai-Yan Cao

Shandong University

Lin Wei

Shandong University

Danrong Lu

Shandong University

Min Du

Shandong University

Min Yuan

Shandong University

Deling Shi

Shandong University

Xiangxue Chen

Tiandong Pharma

Peng Wang

Shandong University

Xiu-Lan Chen

Shandong University

Lianli Chi

Shandong University

Yu-Zhong Zhang

Shandong University

fuchuan Li (✉ fuchuanli@sdu.edu.cn)

Shandong University

Article

Keywords: Heparinases, exoHepases, DUF4962, heterogeneous heparin, heparan sulfate

Posted Date: August 14th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-54275/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Heparinases are critical tools for the studies of highly heterogeneous heparin (HP)/heparan sulfate (HS). However, the exolytic heparinases urgently needed for the sequencing of HP/HS chains are seemingly inaccessible. Herein, a type of exolytic heparinases (exoHepases) was identified from the genomes of different bacteria. These exoHepases share almost no homology with known Hepases and prefer to digest HP rather than HS chains by sequentially releasing unsaturated disaccharides from their reducing ends. The structural study of an exoHepase (exoHep) shows that an N-terminal conserved DUF4962 superfamily domain is essential to the exolytic activity of these exoHepases, which is involved in the formation of a unique L-shaped catalytic cavity controlling the sequential digestion of substrates through electrostatic interactions. Further, several HP octasaccharides were preliminarily sequenced by using exoHep. Overall, this study fills the research gap of exoHepases and provides urgently needed tools for the structural and functional studies of HP/HS chains.

Introduction

Heparin (HP)/heparan sulfate (HS), a kind of polyanionic and heterogeneous polysaccharides belonging to glycosaminoglycans^{1,2}, are ubiquitously presented on cell surfaces and in the extracellular matrix (ECM) as well as the intracellular environment (in mast cells)^{2,3}. The various structural characteristics of HP/HS permit these molecules to interact with various proteins, such as enzyme inhibitors³, chemokines⁴, growth factors⁵, morphogens⁶, and other signaling proteins, to participate in various physiological and pathological processes, such as coagulation⁷, cell adhesion⁸, inflammation⁴, cell migration³, differentiation⁹ and even pathogenic infections¹⁰. These vital biological roles have attracted great attention to the structural and functional studies and the clinical applications of HP/HS. Since being discovered in the 1920s, HP and its low-molecular preparations, as the most important class of anticoagulants, have been widely applied in clinical treatment¹¹.

The backbone of HP/HS polysaccharides is composed of repeating disaccharide units consisting of D-glucuronic acid (GlcA)/L-iduronic acid (IdoA) and N-acetyl-D-glucosamine (GlcNAc)¹². During the biosynthesis process of HP/HS, their common precursors composed of repeating disaccharide GlcA-GlcNAc units are initially synthesized and then further modified by several enzymes, including *N*-deacetylase-*N*-sulfotransferase and 3-*O*- and 6-*O*-sulfotransferase, which modify GlcNAc residues; C5-epimerase, which converts the GlcA to an IdoA residue; and 2-*O*-sulfotransferase, which catalyzes the sulfation of GlcA/IdoA residues¹³. These modifications can result in 32 disaccharide variations in HP/HS polysaccharides, which make HP/HS the most sophisticated polymers in nature¹³. Consequently, the high structural complexity has severely hindered the structural and functional studies of HP/HS.

Heparinases (Hepases) from bacteria are polysaccharide lyases that specifically catalyze the β -eliminative reaction of glycosidic bonds between GlcNAc and GlcA/IdoA residues in HP/HS chains to produce oligosaccharides containing a 4,5-unsaturated uronic acid residue at the nonreducing end and are indispensable tools for the structural and functional studies of HP/HS^{14,15}. To date, based on their substrate specificity, the identified Hepases can be divided into three types: Hepase I (EC4.2.2.7), which specifically degrades highly sulfated and IdoA-rich HP; Hepase III (EC 4.2.2.8), which prefers to degrade low sulfated and GlcA-rich HS; and Hepase II (EC 4.2.2.-), which digests both HP and HS^{16,17}. Structurally, Hepase I folds into a β -jelly roll-type structure and prefers to cleave the α -1,4-linkages connected to IdoA2S/GlcA2S residues^{18,19}. In contrast, Hepase III is composed of an N-terminal (α/α)₅ barrel domain and a C-terminal antiparallel β -sandwich domain and prefers to cleave the α -1,4-linkages connected to GlcA/IdoA residues^{18,20}. Hepase II adopts a topology similar to Hepase III and has no selectivity for certain GlcA/IdoA or IdoA2S/GlcA2S structures^{18,21,22}. All Hepases share similar catalytic mechanisms in which Tyr-His acts as a Brønsted base and acid^{19,22,23}.

Notably, all identified Hepases belong to the family of endolytic lyases, which randomly cleave the HP/HS chains on the inside, while exo-type Hepases, which can be very useful tools for sequencing of HP/HS chains, have not been found until now. Here, a new type of Hepase with exolytic activity has been identified for the first time. These exolytic Hepases belong to the PL15 family and prefer to degrade highly sulfated HP chains at their reducing end by sequentially releasing unsaturated disaccharides. Further structural studies have shown that the active center of the exolytic Hepase is an L-shaped semiopen tunnel with a positively charged entrance and a negatively charged exit, which precisely controls the sequential cleavage of highly negatively charged HP. This exolytic mechanism is significantly different from those of other identified exolyases in the PL15 family.

Results

Sequence Information of the New Hepases. The *exohep* gene (GenBank: EDV07780.1) is 2607 bp long, contains 46.11% GC, and encodes a protein consisting of 868 amino acid residues. The molecular mass and isoelectric point (pI) of the putative exoHep protein are 98.9 kDa and 6.54, respectively. SignalP 4.1 and LipoP 1.0 analyses indicate that the exoHep protein has a type I signal peptide containing 23 amino acid residues. Furthermore, Carbohydrate-Active Enzyme database and Simple Modular Architecture Research Tool analyses showed that the exoHep protein is composed of three domains: an N-terminal signal peptide, a DUF4962 superfamily domain and a C-terminal Hepase II/III superfamily domain. Phylogenetic analysis showed that exoHep clustered with other putative Hepases in the PL15 family, such as the Hepase II (BTexoHep) from *Bacteroides thetaiotaomicron* VPI-5482²⁴ and heparan-sulfate lyase (BCexoHep) from *Bacteroides cellulosilyticus* WH2 in the PL15_2 family²⁵. This subfamily of the PL15 family is separated from other Hepase clades, as shown in Fig. 1. Notably, the enzymatic characteristics of these putative Hepases in this subfamily, including BCexoHep and BTexoHep, remain to be studied.

To investigate the possibility of exoHep belonging to an unidentified Hepase family, several proteins sharing various similarities with exoHep in the PL15_2 family, such as the Hepase II/III-like protein (BFexoHep) (GenBank: EEX44367.1) from *Bacteroides fingoldii* DSM 17565, the hypothetical protein JCM10512_3474 (BRexoHep) (GenBank: GAE85076.1) from *Bacteroides reticulotermitis* JCM 10512s, the DUF4962 domain-containing protein (PAexoHep) (NCBI accession number: WP_026063245.1) from *Pedobacter arcticus*, the DUF4962 domain-containing protein (BXexoHep) (GenBank: WP_008020419.1) from *Bacteroides xyloisolvans* and the DUF4962 domain-containing protein (PHexoHep) (GenBank: WP_015808643.1) from *Pedobacter heparinus* were synthesized for the following studies. The sequence properties of these proteins are listed in Table 1. These PL15_2 proteins were mostly composed of an N-terminal signal peptide, a DUF4962 superfamily domain and a C-terminal Hepase II/III superfamily domain, except PAexoHep and BFexoHep, which do not have the N-terminal signal peptide and the C-terminal Hepase II/III superfamily domain. In the phylogenetic tree, these proteins all clustered with exoHep in the PL15 family (Supplementary Fig. S1).

Table 1
Comparison of sequence and enzymatic properties of various PL15_2 family Hepases.

Comparison of sequence and enzymatic properties of various FETC_Z family Hepases.											
Enzyme activity (U/mg)		EDTA effect (5 mM)	Effect of divalent metal cation (5 mM)		Optimal condition			Sequence character			
HP	HP-F _{all}		Inhibitor	Enhancer	Salt	pH	Temperature(°C)	Signal	Isoelectric point	Molecular mass (KDa)	
0.75 ± 0.06	^b 49.89 ± 3.18	19%	Co ²⁺ (0%)	Ca ²⁺ + ^a (142%)	-	6.0 (HAc-NaAc)	30	SPI	6.54	98.9	exoHep
			Hg ²⁺ (0%)	Ba ²⁺ (115%)							
			Pb ²⁺ (0%)								
			Ni ²⁺ (0%)								
			Cu ²⁺ (0%)								
			Zn ²⁺ (0%)								
1.13 ± 0.03	61.86 ± 2.85	99%	Co ²⁺ (33%)	Ca ²⁺ (135%)	NaCl (230%) (100 mM)	8.0 (Tris-HCl)	20	SPII	6.91	98.5	BCexoHep
			Hg ²⁺ (4%)	Ba ²⁺ (150%)							
			Pb ²⁺ (28%)	Mg ²⁺ (133%)							
			Ni ²⁺ (4%)								
			Cu ²⁺ (1%)								
			Zn ²⁺ (11%)								
-	26.75 ± 1.78	95%	Co ²⁺ (36%)	Ca ²⁺ (192%)	NaCl (788%) (250 mM)	8.0 (NaH ₂ PO ₄ ⁻ -Na ₂ HPO ₄)	20	SPII	7.37	100.3	BTexoHep
			Hg ²⁺ (22%)	Ba ²⁺ (198%)							
			Ni ²⁺ (1%)	Mg ²⁺ (187%)							
			Cu ²⁺ (10%)								
			Zn ²⁺ (10%)								
-	24.25 ± 1.24	78%	Ni ²⁺ (0%)	Ba ²⁺ (130%)	NaCl (140%) (5 mM)	8.0 (NaH ₂ PO ₄ ⁻ -Na ₂ HPO ₄)	10	SPI	8.83	98.3	PAexoHep
			Cu ²⁺ (0%)	Mg ²⁺ (142%)							
			Zn ²⁺ (11%)								

^a: The residual activities were shown as the percentage of that (100%) of WT-exoHep.

^b: The enzyme activities are the means ± S.D. for at least three experiments.

Enzyme activity (U/mg)		EDTA effect (5 mM)	Effect of divalent metal cation (5 mM)		Optimal condition			Sequence character			
-	12.13 ± 1.20	97%	Co ²⁺ (1%)	Ca ²⁺ (200%)	KCl (485%)	8.0	20	SPI	9.76	100.2	BFexoHep
			Hg ²⁺ (31%)	Ba ²⁺ (262%)	(50 mM)	(NaH ₂ PO ₄ ⁻ Na ₂ HPO ₄)					
			Ni ²⁺ (1%)	Mg ²⁺ (179%)							
			Cu ²⁺ (8%)								
			Zn ²⁺ (5%)								
a: The residual activities were shown as the percentage of that (100%) of WT-exoHep.											
b: The enzyme activities are the means ± S.D. for at least three experiments.											

Enzymatic Characteristics of the New Hepases. ExoHep was able to slowly degrade HP to mainly produce disaccharides with absorbance at 232 nm (Supplementary Fig. S2a), hardly degrade HS (Supplementary Fig. S2b) and did not degrade chondroitin sulfate (CS), hyaluronan (HA), dermatan sulfate (DS) or alginate (data not shown). Interestingly, compared to its action on the untreated HP polysaccharide, exoHep showed stronger activity to the resistant fractions prepared from the exhaustive digestion of HP polysaccharides with Hepase III (HP-F_{all}) (Table 1). These results indicate that exoHep is an HP lyase instead of an HS lyase.

To further characterize the proteins sharing various similarities with exoHep in the PL15_2 family, these proteins were heterologously expressed in *Escherichia coli* (*E. coli*) and used to estimate the enzymatic activity. The results showed that the proteins BCexoHep, BTexoHep, BFexoHep and PAexoHep had similar substrate specificity to exoHep (Supplementary Fig. S2a-b). In contrast, the proteins BRexoHep, BXexoHep and PHexoHep showed no activity against any tested polysaccharides. These results indicate that the PL15_2 family is an assembly of novel Hepases that prefer to degrade HP rather than HS.

The optimum temperatures of the recombinant enzymes exoHep, BCexoHep, BTexoHep, PAexoHep and BFexoHep are 30 °C, 20 °C, 20 °C, 10 °C and 20 °C, respectively (Table 1). Notably, the low optimum temperature of PAexoHep may be due to the cold living environment of *P. arcticus*²⁶. The optimal pH of the recombinant enzymes exoHep, BCexoHep, BTexoHep, PAexoHep and BFexoHep were determined to be 6.0 in 50 mM NaAc-HAc buffer, 8.0 in 50 mM Tris-HCl, 8.0 in 50 mM NaH₂PO₄-Na₂HPO₄, 8.0 in 50 mM NaH₂PO₄-Na₂HPO₄ and 8.0 in 50 mM NaH₂PO₄-Na₂HPO₄, respectively (Table 1). In addition, the activities of these enzymes were significantly stimulated by 5 mM divalent cations such as Ca²⁺ and Ba²⁺ as well as Mg²⁺ in most cases (Table 1), indicating that the divalent cations may involved in the catalytic mechanism of these types of Hepases. The activity of these novel Hepases is strongly inhibited by various heavy metal ions, such as Co²⁺, Hg²⁺, Ni²⁺, Pb²⁺, Cu²⁺, and Zn²⁺ (Table 1). Moreover, salts (NaCl or KCl) at certain concentrations are strong stimulators for the abovementioned enzymes except exoHep, and the activities of BCexoHep, BTexoHep, PAexoHep and BFexoHep can be optimally enhanced by 100 mM NaCl (230%), 250 mM NaCl (788%), 5 mM NaCl (140%) and 50 mM KCl (485%), respectively (Table 1).

Finally, the enzymatic activities of these novel Hepases were determined under their respective optimum reaction conditions by using HP, HS and HP-F_{all} as substrates. As shown in Table 1, the enzymatic activities of exoHep, BCexoHep, BTexoHep, PAexoHep and BFexoHep toward HP-F_{all} were 49.89 U/mg protein (pH 6.0, 50 mM NaAc-HAc containing 5 mM Ca²⁺, 30 °C), 61.86 U/mg protein (pH 8.0, 50 mM Tris-HCl containing 5 mM Ba²⁺ and 100 mM NaCl, 20 °C), 26.75 U/mg protein (pH 8.0, 50 mM NaH₂PO₄-Na₂HPO₄ containing 5 mM Ba²⁺ and 250 mM NaCl, 20 °C), 24.25 U/mg protein (pH 8.0, 50 mM NaH₂PO₄-Na₂HPO₄ containing 5 mM Mg²⁺ and 5 mM NaCl, 10 °C) and 12.13 U/mg protein (pH 8.0, 50 mM NaH₂PO₄-Na₂HPO₄ containing 5 mM Ba²⁺ and 50 mM KCl, 20 °C), respectively; the activities of exoHep and BCexoHep toward HP polysaccharides were 0.75 and 1.13 U/mg protein, respectively (Table 1). However, the enzymatic activities of BTexoHep, PAexoHep and BFexoHep toward HP polysaccharide were too low to be accurately measured. As expected, these PL15_2 enzymes showed extremely weak activities toward HS polysaccharides.

Digestion Pattern of Substrate by the PL15_2 Hepases. To determine the action pattern of the novel Hepases in the PL15_2 family, a saturated HP DP13 fraction without specific absorbance at 232 nm was prepared from the unsaturated Hepase III-resistant HP UDP14 fraction by treatment with O₃²⁷ and used as substrate for digestion by each enzyme in a time-course assay as described under “Methods”. The results showed that all the tested PL15_2 Hepases generated only disaccharides with specific absorbance at 232 nm, and no larger oligosaccharide products were notably detected during the whole process of degradation (Fig. 1a-e), which suggested that these enzymes should be exolytic Hepases and cleave HP chains from the reducing end to sequentially release unsaturated disaccharides.

Substrate Specificity of ExoHep. To investigate the substrate specificity of the PL15_2 Hepases, the capacity of the typical enzyme exoHep degrading HP tetrasaccharides with specific structures was estimated. Five structure-defined HP tetrasaccharide fractions P4-4 (ΔUA1-4GlcNAc6S1-4GlcA1-4GlcNS6S and ΔUA1-4GlcNAc6S1-4IdoA1-4GlcNS6S), P4-5 (ΔUA1-4GlcNS1-4GlcA1-4GlcNS6S and ΔUA1-4GlcNS1-4IdoA1-4GlcNS6S), P4-6 (ΔUA1-4GlcNS6S1-

4GlcA1-4GlcNS6S), P4-7 (Δ UA2S1-4GlcNS6S1-4GlcA1-4GlcNS6S) and P4-8 (Δ UA2S1-4GlcNS6S1-4IdoA2S1-4GlcNS6S) were prepared and structurally determined (Supplementary Figs. S3, S4, S5 and Table S1) and individually used as substrates for digestion assays by exoHep. When these tetrasaccharides (5 pmol) were treated with excess exoHep overnight, all could be almost completely degraded (Fig. 2a-e), indicating that exoHep could digest both GlcA- and IdoA-containing tetrasaccharides regardless of their sulfation patterns. However, the tetrasaccharides (5 pmol) were treated with limited enzyme (10 mU) for 12 h, and the conversion efficiencies of tetrasaccharides P4-4, P4-5, P4-6, P4-7 and P4-8 to disaccharides were 15.90%, 9.37%, 34.68%, 100% and 100%, respectively (Supplementary Table S2). A further study shows that the enzyme activity of exoHep toward P4-4, P4-5, P4-6, P4-7 and P4-8 are < 1, 3.98, 24.22, 49.12 and 79.49 U/mg proteins (Supplementary Table S3), respectively, suggesting that the activity of exoHep against these tetrasaccharides is significantly different and that this enzyme prefers to digest highly sulfated substrates. In addition, a time course experiment showed that when size-defined HP tetrasaccharides were treated with exoHep the signal of Anomeric ^1H corresponding to IdoA2S residues much more quickly decreased than those of unsulfated IdoA/GlcA residues, and further nonsulfated GlcA showed more susceptible to exoHep than IdoA (Supplementary Fig. S6 and Supplementary Table. S4). Taken together, these results suggest that highly sulfated domains in particular those containing IdoA2S residues in HP/HS chains can be the optimal substrates to exoHepases in the PL15_2 family.

Overall Structure of ExoHep. To investigate the structural characteristics of the PL15_2 proteins, the crystal structure of exoHep with an HP disaccharide product was solved at a resolution of 1.98 Å by preparing a crystal of selenomethionine (SeMet)-labeled exoHep in the presence of HP disaccharide products followed by single-wavelength anomalous dispersion (SAD) phasing. Furthermore, a crystal of an inactive exoHep-Y390A/H555A mutant with an unsaturated HP tetrasaccharide substrate was obtained, and the structure was solved at a resolution of 1.73 Å by molecular replacement with Phaser using the complex structure of exoHep with the product as a template. In the crystal structures, each asymmetric unit consists of one exoHep monomer, an unsaturated disaccharide (Δ UA2S1-4GlcNS6S) or tetrasaccharide (Δ UA1-4GlcNS6S1-4IdoA2S1-4GlcNS6S), and two Ca^{2+} ions that were confirmed by inductively coupled plasma-mass spectrometry (ICP-MS) analysis. Notably, the exoHep complex with the disaccharide product and the structure with the tetrasaccharide substrate do not show significant differences (the root-mean-square deviation (r.m.s.d.) between the two structures was 0.140 Å based on 786 atoms).

The structure of exoHep consists of three domains: an N-terminal small β -sheet domain (Ala²⁴-Asn¹⁵⁰) composed of a large loop and a five-stranded β -sheet (β 1- β 5), a central (α/α)₅ barrel domain (Pro¹⁵¹-Glu⁵¹⁵) containing 20 α -helices (α 1- α 20) to form an (α/α)₅ incomplete toroid structure, and a C-terminal β -sandwich domain (Leu⁵¹⁶-Pro⁸⁶⁴) containing 23 β -strands and 5 short α -helices that fold into a three-layered β -sheet sandwich structure (α 21- α 25 and β 6- β 28) (Fig. 3a). The two Ca^{2+} ions, named Ca1 and Ca2 are coordinated by nitrogen and oxygen atoms of several residues in near pentacoordination with distances of 3.1–3.5 Å (Fig. 3b). Ca1 is close to the uronic acid residue of the disaccharide at a distance of 5.6 Å compared with the 14 Å distance of Ca2.

A structure-based homology search for exoHep was performed using the DALI server²⁸. The results showed that the closest structure to exoHep is the exo-type alginate lyase Atu3025 (PDB code 3AFL)²⁹ with a Z-score of 31.7 and an r.m.s.d. of 3.1 Å on 766 aligned Ca atoms. Additionally, exoHep shows some degree of structural similarity to Hepase II (PDB code 2FUQ, Z-score: 29.3, and r.m.s.d.: 4.3 Å on 747 atoms) from *Pedobacter heprinus* DSM 2366²¹, and Hepase III (PDB code 4FNV, Z-score: 23.8, and r.m.s.d.: 4.0 Å on 659 atoms) from *P. heprinus* DSM 2366²⁰. However, exoHep has an additional small β -sheet domain (Ala²⁴-Asn¹⁵⁰) at the N-terminus compared with Hepase II and III (Fig. 3c).

The Unique L-Shaped Catalytic Cavity of ExoHep. Compared to the open catalytic cavities of Hepase I, II and III (Supplementary Fig. S7a-c), exoHep possesses a semiopen, narrow, elongated and L-shaped catalytic cavity composed of residues from the three domains of the N-terminal small β -sheet, central (α/α)₅ barrel and C-terminal β -sandwich (Supplementary Fig. S7d and Fig. 4a). The unique catalytic tunnel consists of two almost mutually perpendicular sections: the positively charged section generated by residues from α 6, α 8, α 10, α 12, α 14, and the loops connecting α 9/ α 10, α 18/ α 19, β 8/ β 9, and β 10/ α 21, which acts as the entrance for the binding and degradation of negatively charged HS/HP substrates, and the negatively charged section composed of residues from α 7, α 8, α 9 and the big loop of the N-terminal small β -sheet domain, especially residues Phe⁶² to Pro⁸⁰, which acts as the exit to facilitate the release of the disaccharide product (Supplementary Fig. S8).

Substrate-Binding and Product-Releasing Sites of the L-shaped Tunnel. According to the hydrogen-bonding networks, a series of residues, such as Asp⁷⁰, Ser²²⁶, Arg²³⁰, Arg²³³, Arg³³², Asn³³⁶, His³³⁷, Gln³⁴⁰, Tyr³⁹⁰, Ser⁵⁵⁴, His⁵⁵⁵, Arg⁵⁷⁹, Arg⁶⁴⁸, Lys⁷⁵⁵ and Arg⁷⁵⁷, participate in the binding of the disaccharide product (Δ UA2S1-4GlcNS6S) and the tetrasaccharide substrate (Δ UA1-4GlcNS6S1-4IdoA2S1-4GlcNS6S) (Fig. 4c-d). The reducing end sugar rings of the tetrasaccharide substrate at the + 1 and + 2 subsites are tightly hugged by the side chains of the residues in the positively charged section of the L-shaped tunnel via a hydrogen bond network (Fig. 4c). The residues His³³⁷, Tyr³⁹⁰ and His⁵⁵⁵ match well with three key residues His²⁰², Tyr²⁵⁷ and His⁴⁰⁶ in the active center of Hepase II from *P. heprinus* DSM 2366 (Supplementary Fig. S9), indicating that they may also play key roles in the catalytic mechanism of exoHep. This speculation was proven by mutation of three residues to Ala, which caused almost complete loss of enzymatic activity (Supplementary Fig. S10a-b). In contrast, the nonreducing end sugar rings at subsites – 1 and – 2 have fewer direct interactions with the residue sidechains because the “+” subsites are deeply buried in the tunnel, while the “-” subsites are exposed to the environment (Fig. 4d).

The negatively charged residues Pro⁶⁷, Asp⁷⁰, Pro⁷³, Asp²⁸¹, Glu²³⁴ and Asp³³⁵ from the N-terminal small β -sheet and central (α/α)₅ barrel domains compose the negatively charged section of the tunnel for quick delivery of the disaccharide product (Fig. 4b). Residue Asp³³⁵ is located at the corner of the tunnel and close to the + 2 subsite of the substrate (Fig. 4c), and mutation of this residue to Ala almost completely destroyed the enzymatic activity of exoHep (Supplementary Fig. S10a). The other residues compose the tunnel exit to approximately 8.3 Å in width (measurement of CG^{Pro67} to

OD2^{Asp281}) and 12.8 Å in length (measurement of O^{Asp70} to O^{Glu324}). ExoHep-D281A and exoHep-P67A mutants lost 70% and 90% enzymatic activity, respectively (Supplementary Fig. S5a). These results show that the negatively charged residues play very important roles in the release process of disaccharide products. Furthermore, an N-terminal small β -sheet domain deletion mutant (residues 1-150) caused complete loss of exoHep enzymatic activity (Supplementary Fig. S10a), which further confirmed the essential role of the additional small N-terminal domain for the activity of the exoHepases.

Preliminary Sequencing of HP Octasaccharides.

To investigating the potential application of exoHepases in the sequencing of HS/HP chains, the HP octasaccharide fraction prepared from HP-F_{all} were subfractionated by anion-exchange HPLC on a Propack PA1 column, and five main subfractions (P8-1, P8-2, P8-3, P8-4 and P8-5) were obtained as shown in Supplementary Fig S11. To further obtain the preliminary disaccharide sequences of these octasaccharide preparations, an enzymatic sequencing method was established based on the exolytic activity of exoHep (Fig. 6). Taking the sequencing of subfraction P8-4 as an example, an aliquot of P8-4 was partially digested by exoHep for isolating and preparing its nonreducing end hexasaccharide (UDP_{NE-P8-4}) and tetrasaccharide (UDP_{NE-P8-4}) fractions as described under “Methods”. Then, the disaccharide compositions of P8-4 (UDP_{NE-P8-4}), UDP_{NE-P8-4} and UDP_{NE-P8-4} were analyzed by complete digestion with Hepases I and II followed by 2-AB labeling and anion-exchange HPLC. As shown in Supplementary Fig S12d, the complete digestion of UDP_{NE-P8-4} with Hepases produced two unsaturated disaccharides Δ UA1-4GlcNS6S and Δ UA2S1-4GlcNS6S with a molar ratio of 1:3, and the molar ratios of these two disaccharides in its nonreducing end hexasaccharide UDP_{NE-P8-4} and tetrasaccharide UDP_{NE-P8-4} were 1:2 and 1:1, respectively (Supplementary Table. S5). By comparing the molar ratio change of disaccharides in UDP_{NE-P8-4} and UDP_{NE-P8-4}, the reducing end disaccharide (D) of P8-4 could be deduced as HexUA2S1-4GlcNS6S, and similarly the disaccharide located at C site in P8-4 could be determined as HexUA2S1-4GlcNS6S by comparing the disaccharide compositions of UDP_{NE-P8-4} and UDP_{NE-P8-4} too (Fig. 6 and Supplementary Table. S5). To determine the type of disaccharide at B site in P8-4, the UDP_{NE-P8-4} was treated with O₃ to remove the unsaturated uronic acid at the nonreducing end and further digested with Hepases I and II for disaccharide composition analysis as described in “Method”. As shown in Supplementary Fig S12d a single disaccharide peak corresponding to Δ UA2S1-4GlcNS6S was detected, indicating that the disaccharide at B site was HexUA2S1-4GlcNS6S too. Finally, based on the disaccharide composition of UDP_{NE-P8-4}, the nonreducing end disaccharide (A) of P8-4 must be Δ UA1-4GlcNS6S. Taken together, the preliminary sequence of P8-4 can be deduced as Δ UA1-4GlcNS6S1-4HexUA2S1-4GlcNS6S1-4HexUA2S1-4GlcNS6S1-4HexUA2S1-4GlcNS6S (Supplementary Table. S6). By using the same method, the preliminary sequences of P8-1, P8-2, P8-3 and P8-5 were determined as shown in Supplementary Fig S12 and Supplementary Table. S5 and Supplementary Table. S6. The molecular masses and compositions of these octasaccharides were confirmed by the HILIC-ESI-MS (Supplementary Fig S13 and Table. S7). Notably, the monosulfated disaccharide (Δ UA1-4GlcNAc(S)) located at the nonreducing end of P8-3 or P8-5 could not be deduced due to lack of its corresponding disaccharide standard. Considering the fact that Hepase III cannot cleave the 1–4 glycosidic bonds to 2-O-sulfated uronic acid residues, the unknown disaccharide should be sulfated on its GlcNAc residue, such as Δ UA1-4GlcNAc3S.

Discussion

In this study, the discovery of the novel Hepase exoHep with exolytic activity led to the identification of a new Hepase family, PL15_2, a subfamily of the PL15 family that has been known as the assembly of exolytic alginate lyases²⁹, which is different from Hepases I, II and III, belong to the PL13, PL21 and PL12 families¹⁶, respectively. In this new family, five exolytic Hepases, including exoHep, BCexoHep, BTexoHep, PAexoHep and BFexoHep, have been identified and studied in detail for the first time, and compared with the previously identified Hepases (I, II and III), all PL15_2 enzymes have an additional conserved DUF4962 superfamily domain. Structural and functional studies showed that this DUF4962 superfamily domain forms a small N-terminal β sheet domain involved in the formation of the active cleft in the 3-D structure of exoHep, and deletion of this domain will cause a complete loss of exoHep activity, indicating that the conserved DUF4962 superfamily domain is essential to the exolytic activity of PL15_2 Hepases.

Although the PL15_2 Hepases have a preference to degrade HP rather than HS, their activities against HP polysaccharides are still very low, which can be attributed to the complex structures of HP polysaccharides. The inevitable existence of some resistant structures, such as non- or low-sulfated HS domains³⁰ and linkage regions (GlcA-Gal-Gal-Xyl)^{3,31} at the reducing ends of HP polysaccharides will prevent the exolysis of HP chains by these enzymes. This may be the key reason why this type of Hepase is not easy to find, as researchers usually use HP/HS polysaccharides as substrates to screen potential Hepases. Consistent with this hypothesis, these enzymes showed relatively higher activity against the Hepase III-resistant HP fraction HP-F_{all}. In fact, by using structure-defined HP tetrasaccharides as substrates we found that the enzymatic activity of exoHep increased with the increase of sulfation degree of the substrates, and further study showed that IdoA2S residues in HP tetrasaccharides were much more sensitive than nonsulfated uronic acids, which can explain why this enzyme exhibited the highest activity toward the hexasulfated tetrasaccharide Δ UA2S1-4GlcNS6S1-4IdoA2S1-4GlcNS6S. Interestingly, we also find that nonsulfated GlcA residues are more susceptible to exoHep than IdoA residues in tetrasaccharides, suggesting that exoHepases have different preference to Glc- and IdoA-containing saccharides. However, the detailed effects of uronic acid residues on the activity of exoHepases remain to be further investigated by using various structure-defined HP/HS oligosaccharides.

The active centers of exoHep and Hepase II partially overlap, indicating that their catalytic mechanism might have some features in common. As shown in this study, the residues His³³⁷, Tyr³⁹⁰ and His⁵⁵⁵ of exoHep corresponding to the three key residues His²⁰², Tyr²⁵⁷ and His⁴⁰⁶ of Hepase II²² play key roles in the catalytic mechanism of exoHep, indicating that similar to other lyases, including Hepases I, II and III, exolytic Hepases catalyze the degradation of HP via a catalytic mechanism in which Tyr-His acts as the Brønsted base and acid^{19,22,23}. However, unlike in the case of the His³³⁷ to Ala

mutation that completely destroyed the ability of exoHep to cleave the tetrasaccharide Δ UA2S1-4GlcNS6S1-4IdoA2S1-4GlcNS6S (Supplementary Fig. S10b), a previous study showed that the mutation of His²⁰² to Ala in Hepase II did not completely destroy the enzymatic activity toward this tetrasaccharide²², indicating that there are some significant differences between the catalytic mechanisms of these two enzymes.

Compared with the open active centers of endolytic Hepase I, II and III^{19, 20, 21}, exoHep has a semiopen active center, indicates that the substrate of HP chains cannot randomly bind to the catalytic cleft of exoHepases. Furthermore, the active center of exoHep is an L-shaped tunnel, which is similar to that of Hepase I from *B. thetaiotaomicron* VPI-5482¹⁹. However, the L-shaped tunnel of Hepase I is open, longer, and positively charged, which allows larger HP chains to randomly bind to the tunnel and thus results in mainly endolytic activity. In contrast, the semiopen L-shaped tunnel of exoHep is composed of two sections: the positively charged entrance and negatively charged exit, which can strictly control the HP chains that go through the tunnel from the positively charged entrance and be sequentially cleaved to produce an unsaturated disaccharide product that is released from the negatively charged exit.

Based on this study, the exolytic mechanism of PL15.2 Hepases is proposed: 1) first, the reducing end of the free HP chain in the environment binds to the positively charged entrance of the L-shaped tunnel via charge attraction; 2) through electrostatic interactions, the bound HP chain moves into the tunnel, and once the reducing terminal disaccharide reaches the +1 and +2 subsites, the disaccharides will be cleaved via a β -elimination mechanism; 3) with the aid of the negatively charged residue Asp³³⁵, the resulting disaccharide product is pushed into the negatively charged exit and rapidly released with assistance from Pro⁶⁷ and Asp²⁸¹ (Fig. 5). Notably, this exolytic mechanism is different from those of other enzymes in the PL15 family, such as the exo-type alginate lyase Atu3025²⁹. The active center of Atu3025 is a pocket-like structure that releases the monosaccharide product mainly by the opening of a short α -helix, which is simple in structure and suitable for the production and release of small products. In contrast, the L-shaped tunnel of exoHep is more complex and suitable for the production and release of larger disaccharide products.

Currently, sequencing of HP/HS chains is still a big challenge though various techniques have been developed over the past three decades^{32, 33, 34}. Compared with methods involved in mass spectroscopy and NMR, which need sophisticated instruments and professional experience, enzymatic sequencing of HP/HS saccharides is easier to be mastered and applied in common laboratories and is particularly suitable for the analysis of trace samples. However, due to the lack of exoHepases, existing enzymatic sequencing methods require a combined use of various animal-derived exoenzymes to specifically hydrolyze different sulfate groups and glycosidic bonds, which causes the sequencing process very cumbersome^{32, 35}. In this study, by utilizing the exolytic feature of exoHepases that sequentially cleave HP chains from the reducing ends, a novel enzymatic sequencing method (Fig. 6) was established and successfully determined these preliminary sequences of five HP octasaccharides. Different from the existing enzymatic methods using specific lysosomal exoenzymes to analyze sulfate groups and monosaccharide residues one by one, this method using exoHepase can sequence HP oligosaccharides by analyzing disaccharide as a basic unit, which greatly simplifies the analysis process. Technically, this method can be used to sequence longer HP/HS oligosaccharides. Additionally, bacteria-derived exoHepases are easier to be recombinantly expressed and are more stable and higher activity comparing to animal-derived exoenzymes. We believe that these exoHepases identified in this study can be very powerful tools in the sequencing of HP/HS chains by combining with analytical techniques, such as nitrous acid degradation³² and Ion Mobility Mass Spectrometry³⁴.

In conclusion, the first identification of an exolytic Hepase family can provide not only novel enzymatic tools for structural and functional studies, particularly the sequencing of HP/HS chains, but also a clue for the discovery of additional novel Hepases. Moreover, revelation of the exolytic mechanism of these enzymes will certainly enrich our knowledge regarding the catalytic mechanism of exolytic lyases and be helpful for the engineering of related enzymes.

Methods

Materials. The strain *Bacteroides intestinalis* DSM 17393 was purchased from German Collection of Microorganisms and Cell Cultures (Braunschweig, Germany). The genes encode PL15 family proteins were downloaded from National Center for Biotechnology Information (NCBI) database and synthesized by GENEWIZ, Inc (Suzhou, China). The Fast Mutagenesis Kit V2 was purchased from Vazyme Biotech Co., Ltd (Nanjing, China). Standard unsaturated disaccharides were purchased from Iduron (Manchester, United Kingdom). 2-Aminobenzamide (2-AB), cyanoborohydride (NaBH_3CN), hyaluronan (HA), chondroitin sulfate A (CSA) from whale cartilage, chondroitin sulfate C (CSC) from shark cartilage, dermatan sulfate (DS) from porcine skin and alginate were purchased from Sigma-Aldrich, HP and HS polysaccharides from porcine intestinal mucosa were provided by Tiandong Pharma (Dongying, China). All other chemicals and reagents used in this paper were of the highest quality available. Hepase I (EC 4.2.2.7), II (EC 4.2.2.-), III (EC 4.2.2.8), 2-O-sulfatase and glycuronidase were cloned from *Pedobacter heprinus* DSM 2366^{36, 37}. The anti-Hepase III fraction of HP (HP-F_{all}) was prepared by exhaustively degraded of HP polysaccharides with Hepase III and then precipitated with 5 volume anhydrous alcohol followed by centrifugation to collect the deposit. Size-defined HP oligosaccharide fractions were prepared by the digestion of HP polysaccharides with Hepase II or Hepase III followed by gel filtration chromatography on a Superdex Peptide 10/300 GL column³⁸.

Sequence Analysis of the Gene and Protein of ExoHep. The gene and protein sequence of exoHep were downloaded from the NCBI database (GenBank: EDV07780.1). Then, the analysis of the similarity between exoHep and the known Hepases sequences was performed using the BLASTp algorithm online. Secretion signal peptides and their types were analyzed using the SignalP 4.1 server and LipoP 1.0 server online, respectively. The molecular mass of the protein was calculated using the peptide mass tool on the ExPASy server of the Swiss Institute of Bioinformatics. Sequence alignment and phylogenetic analysis were performed using MEGA version 7.0. Protein modules and domains were identified using the Simple Modular Architecture Research Tool, Pfam database, and Carbohydrate-Active Enzyme database.

Heterologous Expression and Purification of ExoHep and Its Derivative with Selenomethionine. To express exoHep in *E. coli* strains, the gene *exohep* was constructed onto plasmid pET-30a (+) (Novagen). Then, the plasmid pET-30a-*exohep* was transformed into *E. coli* BL21 (DE3) cells. *E. coli* cells harboring an expression vector (pET30a-*exohep*) were initially cultured in LB broth and then induced to start the expression of wild-type exoHep (WT-exoHep) proteins with 5 mM IPTG (isopropyl 1-thio- β -D-galactopyranosid) when the cell density reached an A_{600} of 0.6–1.0. An overexpression system for the exoHep derivative with selenomethionine (SeMet-exoHep) was also constructed in *E. coli* BL21 (DE3) cells as previously reported³⁹. Cells grown overnight in LB medium were harvested by centrifugation at $3,700 \times g$ for 5 min and washed twice with M9 medium. Then, the cells were transferred into M9 medium containing 1 mM $MgSO_4$, 0.1 mM $CaCl_2$ and 0.4% (w/v) glucose and cultured at 37 °C. When the A_{600} reached 0.6, the culture was cooled to 22 °C, and then 100 mg/L lysine, phenylalanine, and threonine and 50 mg/L isoleucine, leucine, valine and L-SeMet were added. Fifteen minutes later, the culture was incubated at 16 °C overnight under the induction of 5 mM IPTG.

The cells were collected by centrifugation at $8,000 \times g$ for 5 min, washed with ice-cold buffer A (50 mM Tris-HCl, 150 mM NaCl, pH 8.0) and resuspended in the same buffer, followed by ultrasonic disruption of cells (50 repetitions, 5 s) in an ice-cold environment. After centrifugation at $15,000 \times g$ for 30 min, the supernatant containing soluble proteins was collected for further steps. The supernatant was loaded onto a column packed with nickel-Sepharose 6 Fast Flow and washed with a gradient concentration of imidazole ranging from 50 to 250 mM. Then, the 250 mM imidazole eluate was diluted and loaded onto a Source 15Q column and washed with a gradient concentration of NaCl ranging from 0 to 1 M. Finally, the fractions containing the target protein were concentrated and further fractionated by gel filtration on a Superdex G-200 column. The purity of the exoHep protein was analyzed using SDS-PAGE. SDS-PAGE was performed using 13.2% polyacrylamide gels.

Expression and Preliminary Characterization of PL15_2 Genes. A series of unidentified homogenous genes of exoHepases were synthesized by GENEWIZ, Inc. (Suzhou, China) and characterized as described for exoHep above. These genes, including Hepase II (pET 30a-*BTexoHep*) from strain *Bacteroides thetaiotaomicron* VPI-5482, heparan-sulfate lyase (pET 30a-*BCexoHep*) from *Bacteroides cellulosilyticus* WH2, Hepase II/III-like protein (BFexoHep) (GenBank: EEX44367.1) from strain *Bacteroides fingoldii* DSM 17565, the hypothetical protein JCM10512_3474 (BRexoHep) (GenBank: GAE85076.1) from strain *Bacteroides reticulotermitis* JCM 10512, the DUF4962 domain-containing protein (PAexoHep) (NCBI accession number: WP_026063245.1) from *Pedobacter arcticus*, the DUF4962 domain-containing protein (BXexoHep) (GenBank: WP_008020419.1) from *Bacteroides xylanisolvens* and the DUF4962 domain-containing protein (PHexoHep) (GenBank: WP_015808643.1) from *Pedobacter heparinus*, share a sequence similarity of 66.26%, 60.57%, 49.27%, 47.79% and 31.13% with exoHep, respectively. The expression plasmids were transferred into *E. coli* BL21 (DE3) cells and overexpressed. The activities of the expression products were analyzed as described for exoHep above.

Biochemical Characterization of the Recombinant PL15_2 Family of Proteins. The optimal conditions for each PL15_2 protein were determined as described previously⁴⁰. Briefly, to determine the optimal pH for each protein, HP-F_{all} (1 mg/ml) was digested with the indicated protein (0.5 μ g) in buffers with a gradient of pH values, including a final concentration of 50 mM NaAc-HAc buffer (pH 5.0–6.0), 50 mM NaH_2PO_4 - Na_2HPO_4 buffer (pH 6.0–8.0), and 50 mM Tris-HCl buffer (pH 7.0–10.0) in a total volume of 100 μ l, at 30 °C for 30 min. After the optimum pH was determined, the effects of temperature on each protein activity were tested in each optimal buffer at temperatures ranging from 0 to 70 °C for 30 min. The effects of metal ions/chelating reagents (5 mM) on the HP-F_{all} degrading activities of each protein were investigated at the optimum pH and temperature determined above. The effects of salt concentrations on the activities of each protein against HP-F_{all} were investigated at the optimum pH and temperature of each corresponding protein.

Activity Assay of PL15_2 Enzymes. The activities of PL15_2 enzymes were measured using HP, HS and HP-F_{all} as substrates. Briefly, each enzyme (1 μ g) was added to 1 mg/ml HP, HS or HP-F_{all} in its optimal buffer (50 mM) and an additional optimal concentration of ions or salts in a total volume of 1 ml. The reaction mixture was incubated at the corresponding optimal temperature. At various time intervals 0, 0.5, 1, 2 and 5 min, aliquots of 100 μ l were withdrawn in duplicate, boiled for 10 min, and then cooled in ice-cold water for 10 min. After being centrifuged at $15,000 \times g$ for 10 min, the supernatant was collected, and diluted three times. All reactions were performed in triplicate, and the activity of enzyme was estimated by measuring the absorbance of the diluted products at 232 nm. One unit of enzyme was defined as the amount of enzyme that produced 1 μ mol of unsaturated carbon bonds per/min.

Degradation Pattern of Substrates by PL15_2 Enzymes. Because the activity of the PL15_2 family of proteins toward HP and HS polysaccharides was relatively low, HP UDP14, derived from the final products of HP degraded by Hepase III, was used to investigate the action pattern of the PL15_2 family of proteins. First, 200 μ g of HP UDP14 was treated with O_3 to remove unsaturated uronic acid at the nonreducing ends as previously reported, which resulted in saturated HP DP13 without specific absorbance at 232 nm. After desalting by gel filtration, saturated HP DP13 (30 μ g) was digested with various PL15_2 family proteins (50 mU) and traced under their optimal conditions. Aliquots (3 μ g) of the reactants were removed at different time points and analyzed through gel filtration chromatography on a Superdex Peptide 10/300 GL column eluted with 0.20 M NH_4HCO_3 at a flow rate of 0.4 ml/min and monitored at 232 nm using a UV detector. Online monitoring and data analysis (e.g., molar ration determination) were performed using the software LC solution version 1.25.

Substrate Specificity of ExoHep. To determine the substrate specificity of exoHep, 5 pmol of various structure-defined HP tetrasaccharides (P4-4 (mixture of Δ UA1-4GlcNAc6S1-4GlcA1-4GlcNS6S and Δ UA1-4GlcNAc6S1-4IdoA1-4GlcNS6S), P4-5 (mixture of Δ UA1-4GlcNS1-4GlcA1-4GlcNS6S and Δ UA1-4GlcNS1-4IdoA1-4GlcNS6S), P4-6 (Δ UA1-4GlcNS6S1-4GlcA1-4GlcNS6S), P4-7 (Δ UA2S1-4GlcNS6S1-4GlcA1-4GlcNS6S) and P4-8 (Δ UA2S1-4GlcNS6S1-4IdoA2S1-4GlcNS6S)) (Supplementary Figs. S3, S4, S5 and Table S1) were degraded by 100 mU of exoHep protein overnight. Then, the products were 2-

AB labeled and analyzed by anion-exchange HPLC on a Pack Polyamine II column eluted using a NaH_2PO_4 gradient (0–1 M) and monitored using a fluorescence detector with excitation and emission wavelengths of 330 and 420 nm⁴¹, respectively.

To determine the preference of exoHep to different uronic acid residues, size-defined tetrasaccharides (3 mg) prepared from the partial digest of HP treated with Hepase II was treated with 0.9 U exoHep for 0 min, 1 min, 5 min, 3 min or 12 h and then boiled to stop the reaction. Resulted oligosaccharides in each reactant were purified and pooled through gel filtration chromatography on a Superdex Peptide 10/300 GL column eluted with 0.20 M NH_4HCO_3 , and freeze-dried repeatedly to remove NH_4HCO_3 . Each oligosaccharide sample was dissolved in 0.5 ml of 99.9% D₂O and lyophilized for twice, and was finally dissolved in 0.5 mL D₂O in a 5 mm NMR tube for the following analysis. The type of the internal uronic acids in each tetrasaccharide preparation was directly determined by ¹H NMR spectroscopy. ¹H NMR spectroscopy was performed on a JNM-ECP600 (JEOL, Japan) instrument set at 600 MHz.

Mutation of ExoHep. To investigate the key residues involved in the active center of exoHep, the residues in and around the active center were individually mutated to Ala, using the Fast Mutagenesis Kit V2 from Vazyme Biotech Co., Ltd (Nanjing, China). The primers are listed in Supplementary Table S8. The residual activities of the mutants were measured by using HP-F_{all} as substrate, and compared with that (100%) of the WT-exoHep. Furthermore, the activity of the mutant exoHep-H337A was tested by using the structure-defined tetrasaccharide subfraction P4-8 as substrate (5 pmol), and the resultants were 2-AB labeled and analyzed by anion-exchange HPLC on a Pack Polyamine II column as described above.

Crystallization and Data Collection. The protein exoHep (12 mg/ml) was premixed with HP disaccharides (10 mg/ml) and then crystallized at 18 °C by the sitting drop method in buffer containing 16% (w/v) polyethylene glycol (PEG) 1500 and 100 mM MMT (DL-malic acid, MES, and Tris) (pH 6.0). The crystal of the SeMet-exoHep protein was prepared under the same conditions used for exoHep except for at pH 5.3. The inactive mutant exoHep-Y390A/H555A (8 mg/ml) mixed with HP UDP4 at a molar ratio of 1:10 was crystallized at 18 °C by the hanging drop method in buffer containing 100 mM MMT (DL-malic acid, MES, and Tris) (pH 6.5) and 23% PEG 1500. X-ray diffraction data were collected on a BL17U1 beam line at the Shanghai Synchrotron Radiation Facility⁴². The initial diffraction datasets were processed by HKL2000. Data collection statistics are shown in Supplementary Table S9.

Structure Determination and Refinement. The selenium single-wavelength anomalous dispersion (Se-SAD) phasing method was used to determine the phase of exoHep. Experimental phases were solved using the Phenix program⁴³ AutoSol. Initial model building was performed using the Phenix program AutoBuild. Refinement of the exoHep structure was performed alternately by the Phenix program Refine and Coot⁴⁴. The ion type was confirmed by inductively coupled plasma-mass spectrometry (ICP-MS) analysis. The structure of the exoHep-Y390A/H555A complex with a tetrasaccharide was determined by molecular replacement with Phaser using the complex structure of exoHep with a disaccharide as the search model. Data collection and refinement statistics are summarized in Supplementary Table S9. All figures of the structures were generated using the program PyMOL.

Preparation and Sequencing of HP Octasaccharides. Size-defined HP octasaccharide fraction prepared from HP-F_{all} was further subfractionated by anion-exchange HPLC on a Propack PA1 column eluted using a NaCl gradient (0.5–1.5 M) in 70 min by monitoring at 232 nm. The main peaks (P8-1, P8-2, P8-3, P8-4 and P8-5) were collected and desalted through gel filtration chromatography on a G10 column.

To determine the sequences of the octasaccharides P8-1, P8-2, P8-3, P8-4 and P8-5, an enzymatic sequencing method was established as shown in Fig. 6. Firstly, 500 pmol of each octasaccharides were partially digested by 50, 50, 50, 9 and 1.5 mU exoHep proteins for 5 min, respectively. The UDP6 and UDP4 products were collected and freeze-dried repeatedly to remove NH_4HCO_3 to get the purified oligosaccharides. Secondly, the UDP8, UDP6 and UDP4 of the octasaccharides were individually digested by Hepases I and II, 2-AB labeled and then analyzed by anion-exchange HPLC on a Polyamine II column eluted using a NaH_2PO_4 gradient (0–550 mM) in 60 min and monitored using a fluorescence detector with excitation and emission wavelengths of 330 and 420 nm, respectively. Thirdly, the nonreducing end UDP4 from the octasaccharides were individually treated with O₃, then digested by Hepases I and II, 2-AB labeled and analyzed as described in step 2.

The molecular mass and composition of these octasaccharides were further analyzed and confirmed by HILIC-ESI-MS on a LTQ-Orbitrap XL FT MS (Thermo Fisher Scientific, San-Jose, CA) as previously reported⁴⁵. LC separation of octasaccharides was performed on a Thermo Ultimate 3000 system (Thermo Fisher Scientific, San-Jose, CA) using a Luna HILIC column (2.0 mm × 150 mm, 200 Å, Phenomenex, Torrance, CA). Mobile phase A was 5 mmol/L ammonium acetate, while mobile B was 5 mmol/L ammonium acetate in 98% acetonitrile. Mobile B was started at 90% and lasted for 5 min, then decreased to 65% in 20 min and kept for 5 min at a flow rate of 150 µL/min. The MS parameters included a spray voltage of -4.2 kV, a capillary voltage of -40 V, a tube lens voltage of -50 V, a capillary temperature of 275 °C, a sheath flow rate of 20, an aux gas flow rate of 5. All FT mass spectra were acquired at a resolution 60,000 with m/z 300–1500.

Declarations

Competing interests

The authors declare no competing interests.

Data deposition

The atomic coordinates and structure factors of the structures in this study have been deposited in the Protein Data Bank, www.pdb.org (PDB ID codes 6LJA and 6LJL).

Author contributions:

F. Li, Y. Z. Zhang and Q. Zhang designed research; Q. Zhang, L. Wei, D. Lu, M. Du, M. Yuan and D. Shi performed the research; Q. Zhang, H.Y. Cao, P. Wang and X. L. Chen, L. Wei, L. Chi analyzed data; X. Chen prepared some key substrates; Q. Zhang and F. Li wrote the paper.

Acknowledgments

We thank the staffs from BL18U1&BL19U1 beamlines of National Facility for Protein Sciences Shanghai (NFPS) and Shanghai Synchrotron Radiation Facility, for assistance during data collection. This work was supported by the National Science Foundation of China (31971201, 31570071, U1706207, 31630012, and 91851205), the Science and Technology Development Project of Shandong Province (No.2018GSF121002), the Project of Taishan Industry Leading Talent of Shandong Province (tscy20160311), the National Key R&D Program of China (2018YFC1406700, 2018YFC0310704), the Program of Shandong for Taishan Scholars (tspd20181203), AoShan Talents Cultivation Program Supported by Qingdao National Laboratory for Marine Science and Technology (2017ASTCP-OS14), and the Major Scientific and Technological Innovation Project (MSTIP) of Shandong Province (2019JZZY010817).

References

1. Linhardt, R. J. 2003 Claude S. Hudson Award address in carbohydrate chemistry. Heparin: structure and activity. *J. Med. Chem.* **46**, 2551–2564 (2003).
2. Laremore, T. N., Zhang, F., Dordick, J. S., Liu, J. & Linhardt, R. J. Recent progress and applications in glycosaminoglycan and heparin research. *Curr. Opin. Chem. Biol.* **13**, 633–660 (2009).
3. Capila, I. & Linhardt, R. J. Heparin $\pm$ Protein Interactions. *Angew. Chem. Int. Ed.* **41**, 390–412 (2002).
4. Sepuru, K. M., Nagarajan, B., Desai, U. R. & Rajarathnam, K. Structural basis, stoichiometry, and thermodynamics of binding of the chemokines KC and MIP2 to the glycosaminoglycan heparin. *J. Biol. Chem.* **293**, 17817–17828 (2018).
5. García-Jiménez, M. J. et al. Interactions between a heparin trisaccharide library and FGF-1 analyzed by NMR methods. *Int. J. Mol. Sci.* **18**, 1293 (2017).
6. Lee, J. et al. Structural determinants of heparin-transforming growth factor- β 1 interactions and their effects on signaling. *Glycobiology* **25**, 1491–1504 (2015).
7. Piran, S. & Schulman, S. Treatment of bleeding complications in patients on anticoagulant therapy. *Blood* **133**, 425–435 (2019).
8. Guillem-Marti, J. et al. RGD mutation of the heparin binding II fragment of fibronectin for guiding mesenchymal stem cell behavior on titanium surfaces. *ACS Appl. Mater. Interfaces.* **11**, 3666–3678 (2019).
9. Lin, Y. et al. Heparin promotes cardiac differentiation of human pluripotent stem cells in chemically defined albumin-free medium, enabling consistent manufacture of cardiomyocytes. *Stem. Cells. Transl. Med.* **6**, 527–538 (2017).
10. Kizlik-Masson, C. et al. Cleavage of anti-PF4/heparin IgG by a bacterial protease and potential benefit in heparin-induced thrombocytopenia. *Blood* **133**, 2427–2435 (2019).
11. Hemker, H. C. A century of heparin: past, present and future. *J. Thromb. Haemost.* **14**, 2329–2338 (2016).
12. Gama, C. I. & Hsieh-Wilson, L. C. Chemical approaches to deciphering the glycosaminoglycan code. *Curr. Opin. Chem. Biol.* **9**, 609–619 (2005).
13. Sasisekharan, R. & Venkataraman, G. Heparin and heparan sulfate: biosynthesis, structure and function. *Curr. Opin. Chem. Biol.* **4**, 626–631 (2000).
14. Linhardt, R. J., Galliher, P. M. & Cooney, C. L. Polysaccharide lyases. *Appl. Biochem. Biotechnol.* **12**, 135–176 (1987).
15. Linhardt, R. J., Wang, H. & Ampofo, S. A. "New methodologies in heparin structure analysis and the generation of LMW heparins" in heparin and related polysaccharides. (Springer press, Boston, 1992).
16. Ulaganathan, T. et al., Conformational flexibility of PL12 family heparinases: structure and substrates pecificity of heparinase III from *Bacteroides thetaiotaomicron* (BT4657). *Glycobiology* **27**, 176–187 (2017).
17. Desai, U. R., Wang, H. M. & Linhardt, R. J. Substrate specificity of the heparin lyases from *Flavobacterium heparinum*. *Arch. Biochem. Biophys.* **306**, 461–468 (1993).
18. Desai, U. R., Wang, H. M. & Linhardt, R. J. Specificity studies on the heparin lyases from *Flavobacterium heparinum*. *Biochemistry* **32**, 8140–8145 (1993).
19. Han, Y. H. et al. Structural snapshots of heparin depolymerization by heparin lyase I. *J. Biol. Chem.* **284**, 34019–34027 (2009).
20. Hashimoto, W., Maruyama, Y., Nakamichi, Y., Mikami, B. & Murata, K. Crystal structure of *Pedobacter heparinus* heparin lyase Hep III with the active site in a deep cleft. *Biochemistry* **53**, 777–786 (2014).

21. Shaya, D. et al. Crystal structure of heparinase II from *Pedobacter heparinus* and its complex with a disaccharideproduct. *J. Biol. Chem.* **281**, 15525–15535 (2006).
22. Shaya, D. et al. Catalyticmechanism of heparinase II investigated by site-directed mutagenesis and the crystalstructure with its substrate. *J. Biol. Chem.* **285**, 20051–20061 (2010).
23. Dong, W., Lu, W., McKeethan, W. L., Luo, Y. & Ye, S. Structural basis of heparan sulfate-specific degradation by heparinase III. *Protein. Cell.* **3**, 950–961 (2012).
24. Xu, J. et al., A genomic view of the human-*Bacteroides thetaiotaomicron* symbiosis. *Science.* **299**, 2074–2086 (2003).
25. Helbert, W. et al. Discovery of novel carbohydrate-active enzymes through the rational exploration of the protein sequences space. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 6063–6068 (2019).
26. Zhou, Z. et al. *Pedobacter arcticus* sp. nov., a facultative psychrophile isolated from Arctic soil and emended descriptions of the genus *Pedobacter*, *Pedobacter heparinus*, *Pedobacter daechungensis*, *Pedobacter terricola*, *Pedobacter glucosidilyticus* and *Pedobacter lentus*. *Int. J. Syst. Evol. Microbiol.* **62**, 1963–1969 (2012).
27. Masuko, S. et al. Ozonolysis of the double bond of the unsaturated uronate residue in low-molecular-weight heparin and K5 heparosan. *Carbohydr. Res.* **346**, 1962–1966 (2011).
28. Holm, L. & Rosenström, P. Dali server: conservation mapping in 3D. *Nucleic. Acids Res.* **38**, 545–549 (2010).
29. Ochiai, A., Yamasaki, M., Mikami, B., Hashimoto, W. & Murata, K. Crystal structure of exotype alginate lyase Atu3025 from *Agrobacterium tumefaciens*. *J. Biol. Chem.* **285**, 24519–24528 (2010).
30. Sasisekharan, R. & Venkataraman, G. Heparin and heparan sulfate: biosynthesis, structure and function. *Curr. Opin. Chem. Biol.* **4**, 626–631 (2000).
31. Sugahara, K. & Kitagawa, H. Recent advances in the study of the biosynthesis and functions of sulfated glycosaminoglycans. *Curr. Opin. Struct. Biol.* **10**, 518–527 (2000).
32. Turnbull, J. E., Hopwood, J. J. & Gallagher, J. T. A strategy for rapid sequencing of heparan sulfate and heparin saccharides. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 2698–2703 (1999).
33. Gill, V. L., Wang, Q., Shi, X. & Zaia, J. Mass spectrometric method for determining the uronic acid epimerization in heparan sulfate disaccharides generated using nitrous acid. *Anal. Chem.* **84**, 7539–7546 (2012).
34. Miller, R. L. et al. Shotgun ion mobility mass spectrometry sequencing of heparan sulfate saccharides. *Nat. Commun.* **11**, 1481 (2020).
35. Stringer, S. E., Kandola, B. S., Pye, D. A. & Gallagher, J. T. Heparin sequencing. *Glycobiology* **13**, 97–107 (2003).
36. Myette, J. R. et al. The heparin/heparan sulfate 2-O-sulfatase from *Flavobacterium heparinum*. Molecular cloning, recombinant expression, and biochemical characterization. *J. Biol. Chem.* **278**, 12157–12166 (2003).
37. Myette, J. R. et al. Molecular cloning of the heparin/heparan sulfate delta 4,5 unsaturated glycuronidase from *Flavobacterium heparinum*, its recombinant expression in *Escherichia coli*, and biochemical determination of its unique substrate specificity. *Biochemistry* **41**, 7424–7434 (2002).
38. Li, F. et al. Structure of pleiotrophin- and hepatocyte growth factor-binding sulfated hexasaccharide determined by biochemical and computational approaches. *J. Biol. Chem.* **285**, 27673–27685 (2010).
39. Xu, F. et al. Novel molecular Insights into the catalytic mechanism of marine bacterial alginate lyase AlyGC from polysaccharide lyase family 6. *J. Biol. Chem.* **292**, 4457–4468 (2017).
40. Lu, D. et al. Biochemical characteristics and synergistic effect of two novel alginate lyases from *Photobacterium* sp. FC615. *Biotechnol. Biofuels.* **12**, 1–17 (2019).
41. Han, W. et al., A novel eliminase from a marine bacterium that degrades hyaluronan and chondroitin sulfate. *J. Biol. Chem.* **289**, 27886–27898 (2014).
42. Wang, Z. et al., Automatic crystal centering procedure at the SSRF macromolecular crystallography beamline. *J. Synchrotron. Radiat.* **23**, 1323–1332 (2016).
43. Adams, P. D. et al., PHENIX: building new software for automated crystallographic structure determination. *Acta. Crystallogr. D. Biol. Crystallogr.* **58**, 1948–1954 (2002).
44. Emsley, P. & K. Cowtan, Coot: model-building tools for molecular graphics. *Acta. Crystallogr. D. Biol. Crystallogr.* **60**, 2126–2132 (2004).
45. Li, L., Zhang, F., Zaia, J. and Linhardt, R. J. Top-down approach for the direct characterization of low molecular weight heparins using LC-FT-MS. *Anal. Chem.* **84**, 8822–8829 (2012).

Figures

Figure 1

Time course experiments of degradation of saturated HP DP13 by PL15_2 family enzymes. Saturated HP DP13 (30 μg) was treated with 50 mU of exoHep (a), BCexoHep (b), BTexoHep (c), PAexoHep (d) or BFexoHep (e). Aliquots (3 μg) were taken at different time points for gel filtration analysis as described under “Methods”. Saturated HP DP13 was prepared via the treatment of the Hepase III-resistant HP UDP14 fraction with O3 as described under “Methods”. Di: HP disaccharide fractions.

Figure 2

Substrate specificity of exoHep. Structure-defined HP tetrasaccharides P4-4 (a), P4-5 (b), P4-6 (c), P4-7 (d), and P4-8 (e) were individually digested with exoHep, and the digests were labeled with 2-AB followed by analysis using anion-exchange HPLC on a Pack Polyamine II column eluted using a NaH_2PO_4 gradient (0-1 M) and monitored using a fluorescence detector with excitation and emission wavelengths of 330 and 420 nm, respectively. 6S: $\Delta\text{UA}(1-4)\text{GlcNAc}6\text{S}$; NS: $\Delta\text{UA}(1-4)\text{GlcNS}$; NS6S: $\Delta\text{UA}(1-4)\text{GlcNS}6\text{S}$; 2SNS6S: $\Delta\text{UA}2\text{S}(1-4)\text{GlcNS}6\text{S}$.

Figure 3

Overall structure of exoHep. (a) Overall structure (stereo diagram) of exoHep. α -Helices and β -sheets are colored red and yellow, respectively. The three domains are labeled with elliptical coils. (b) Ca^{2+} binding sites. The pentacoordination of $\text{Ca}1$ and $\text{Ca}2$ ions in the exoHep structure. A cut off distance of 3.5 Å was applied to choose neighboring residues. The distances between the residues and $\text{Ca}1$ atom are as follows: NE2Asp340, 3.2 Å; NE1Trp384, 3.3 Å; ND2Asn386, 3.1 Å; ND2Asn394, 3.3 Å; and water molecule, 3.3 Å. The distances between the residues and $\text{Ca}2$ atom are as follows: NGly440, 3.1 Å; NH1Arg546, 3.4 Å; NLeu557, 3.3 Å; ND2Asn559, 3.5 Å; and water molecule, 3.2 Å. (c) Structural overlap of exoHep with Hepase II and Hepase III from *P. heprinus* DSM 2366. The structures are shown as cartoons, and Hepase II (PDB code 2FUT), Hepase III (PDB code 4MMH) and exoHep (PDB code 6LJL) are colored pink, yellow and green, respectively. The additional N-terminal β -sheet domain of exoHep is labeled with an elliptical coil.

Figure 4

The L-shaped catalytic tunnel and relative sites of exoHep. Sectional view of the L-shaped tunnel with a tetrasaccharide substrate. (a) The positively charged entrance (blue) and negatively charged exit (red) of the tunnel are indicated with white dashed lines, and the tetrasaccharide substrate is denoted by a yellow arrow. (b) Key residues in the negatively charged exit. The residues are shown in surface representation. Binding of an HP disaccharide product to the positively charged entrance of exoHep. (c) The disaccharide ($\Delta\text{UA}2\text{S}1\text{-}4\text{GlcNS}6\text{S}$) is shown in thick lines and its carbon atoms are colored green, and the side chains of surrounding residues are shown in thick lines and their carbon atoms are colored fuchsia. (d) Binding of an HP tetrasaccharide substrate to the positively charged entrance of exoHep-Y390A/H555A. The tetrasaccharide ($\Delta\text{UA}1\text{-}4\text{GlcNS}6\text{S}1\text{-}4\text{IdoA}2\text{S}1\text{-}4\text{GlcNS}6\text{S}$) and surrounding residue side chains are shown as yellow sticks and as fuchsia sticks, respectively. Hydrogen bonds are shown as dashed lines.

Figure 5

Schematic diagram of the proposed exolytic mode of exoHep.

Figure 6

Strategy for the exo-sequencing of unsaturated HP octasaccharides.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementdataforDiscoveryofexolyticheparinasesandtheir catalytic mechanism and potential application.docx](#)
- [FigS1.jpg](#)
- [FigS2.jpg](#)
- [FigS3.jpg](#)
- [FigS4.jpg](#)
- [FigS5.jpg](#)
- [FigS6.jpg](#)
- [FigS7.jpg](#)
- [FigS8.jpg](#)
- [FigS9.jpg](#)
- [FigS10.jpg](#)
- [FigS11.jpg](#)
- [FigS12.jpg](#)
- [FigS13.jpg](#)
- [TableS1.png](#)
- [TableS2.png](#)
- [TableS3.png](#)
- [TableS4.png](#)
- [TableS5.png](#)
- [TableS6.png](#)

- [TableS7.png](#)
- [TableS8.png](#)
- [TableS9.png](#)