




# Discovery of exolytic heparinases and their catalytic mechanism and potential application

Qingdong Zhang<sup>1,6</sup>, Hai-Yan Cao<sup>2,3,4,6</sup>, Lin Wei<sup>1</sup>, Danrong Lu<sup>1</sup>, Min Du<sup>1</sup>, Min Yuan<sup>1</sup>, Deling Shi<sup>1</sup>, Xiangxue Chen<sup>5</sup>, Peng Wang<sup>3,4</sup>, Xiu-Lan Chen<sup>2,4</sup>, Lianli Chi<sup>1</sup>, Yu-Zhong Zhang<sup>3,4</sup>  & Fuchuan Li  <sup>1</sup> 

Heparinases (Hepases) are critical tools for the studies of highly heterogeneous heparin (HP)/heparan sulfate (HS). However, exolytic heparinases urgently needed for the sequencing of HP/HS chains remain undiscovered. Herein, a type of exolytic heparinases (exoHepases) is 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 (BlexoHep) shows that an N-terminal conserved DUF4962 superfamily domain is essential to the enzyme activities 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 have been preliminarily sequenced by using BlexoHep. Overall, this study fills the research gap of exoHepases and provides urgently needed tools for the structural and functional studies of HP/HS chains.

<sup>1</sup>National Glycoengineering Research Center and Shandong Key Laboratory of Carbohydrate Chemistry and Glycobiology, Shandong University, Qingdao, China. <sup>2</sup>State Key Laboratory of Microbial Technology, Shandong University, Qingdao, China. <sup>3</sup>College of Marine Life Sciences, and Frontiers Science Center for Deep Ocean Multispheres and Earth System, Ocean University of China, Qingdao, China. <sup>4</sup>Laboratory for Marine Biology and Biotechnology, Pilot National Laboratory for Marine Science and Technology (Qingdao), Qingdao, China. <sup>5</sup>Tiandong Pharma, Dongying, China. <sup>6</sup>These authors contributed equally: Qingdong Zhang, Hai-Yan Cao. ✉email: [zhangyz@sdu.edu.cn](mailto:zhangyz@sdu.edu.cn); [fuchuanli@sdu.edu.cn](mailto:fuchuanli@sdu.edu.cn)

**H**eparin (HP)/heparan sulfate (HS), the polyanionic and heterogeneous polysaccharides belonging to glycosaminoglycan family<sup>1,2</sup>, are ubiquitously presented on cell surfaces and in the extracellular matrix (ECM) as well as the intracellular environment (in mast cells)<sup>2,3</sup>. The various structural characteristics of HP/HS permit these molecules to interact with various proteins, such as enzyme inhibitors<sup>3</sup>, chemokines<sup>4</sup>, growth factors<sup>5</sup>, morphogens<sup>6</sup>, and other signaling proteins, to participate in various physiological and pathological processes, such as coagulation<sup>7</sup>, cell adhesion<sup>8</sup>, inflammation<sup>4</sup>, cell migration<sup>3</sup>, differentiation<sup>9</sup>, and even pathogenic infections<sup>10</sup>. 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 1916, HP and its low-molecular preparations, as the most important class of anticoagulants, have been widely applied in clinical treatment<sup>11</sup>.

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)<sup>12</sup>. 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 residues<sup>13</sup>. These modifications can result in numerous disaccharide variations in HP/HS polysaccharides, which make HP/HS the most sophisticated polymers in nature<sup>13</sup>. 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  $\beta$ -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<sup>14,15</sup>. To date, based on their substrate specificity, the identified Hepases can be divided into three types: Heparase I (EC4.2.2.7), which specifically degrades highly sulfated and IdoA-rich HP; Heparase III (EC 4.2.2.8), which prefers to degrade low sulfated and GlcA-rich HS; and Heparase II (EC 4.2.2.-), which digests both HP and HS<sup>16,17</sup>. Structurally, Heparase I folds into a  $\beta$ -jelly roll-type structure and prefers to cleave the  $\alpha$ -1,4-linkages connected to IdoA2S/GlcA2S residues<sup>18,19</sup>. In contrast, Heparase III is composed of an N-terminal ( $\alpha/\alpha$ )<sub>5</sub> barrel domain and a C-terminal antiparallel  $\beta$ -sandwich domain and prefers to cleave the  $\alpha$ -1,4-linkages connected to GlcA/IdoA residues<sup>18,20</sup>. Heparase II adopts a topology similar to Heparase III and has no selectivity for certain GlcA/IdoA or IdoA2S/GlcA2S structures<sup>18,21,22</sup>. All Hepases share similar catalytic mechanisms in which His–Tyr acts as a Brønsted base and acid<sup>19,22,23</sup>. 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, we discover a type of Heparase with exolytic activity. 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 Heparase 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 Hepases.** The *biexohep* 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 BlExoHep protein are 98.9 kDa and 6.54, respectively. SignalP 4.1 and LipoP 1.0 analyses indicate that the BlExoHep 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 BlExoHep protein is composed of three domains: an N-terminal signal peptide, a DUF4962 superfamily domain and a C-terminal Heparase II/III superfamily domain. Phylogenetic analysis showed that BlExoHep clustered with other putative Hepases in the PL15 family, such as the Heparase II (BTExoHep) from *Bacteroides thetaiotaomicron* VPI-5482<sup>24</sup> and heparan-sulfate lyase (BCexoHep) from *Bacteroides cellulosilyticus* WH2 in the PL15\_2 family<sup>25</sup>. This subfamily of the PL15 family is separated from other Heparase clades, as shown in Supplementary 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 BlExoHep belonging to an unidentified Heparase family, several proteins sharing various similarities with BlExoHep in the PL15\_2 family, such as the Heparase 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 xylinisolvans* 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 Heparase II/III superfamily domain, except PAexoHep and BFexoHep, which do not have the N-terminal signal peptide and the C-terminal Heparase II/III superfamily domain. In the phylogenetic tree, these proteins all clustered with BlExoHep in the PL15 family (Supplementary Fig. 1).

**Enzymatic characteristics of the Hepases.** BlExoHep was able to slowly degrade HP to mainly produce disaccharides with absorbance at 232 nm (Supplementary Fig. 2a), hardly degrade HS (Supplementary Fig. 2b). Interestingly, compared to its action on the untreated HP polysaccharides, BlExoHep showed stronger activity to the substrates prepared from the exhaustive digestion of HP polysaccharides with Heparase III (HP-F<sub>III</sub>) (Table 1). These results indicate that BlExoHep is an HP lyase instead of an HS lyase.

To further characterize the proteins sharing various similarities with BlExoHep 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 BlExoHep (Supplementary Fig. 2a, 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 Hepases that prefer to degrade HP rather than HS.

The optimum temperatures of the recombinant enzymes BlExoHep, 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

**Table 1 Comparison of sequence and enzymatic properties of various PL15\_2 family Hepases.**

		<b>BlexoHep</b>	<b>BCexoHep</b>	<b>BTexoHep</b>	<b>PAexoHep</b>	<b>BFexoHep</b>
Sequence character	Molecular mass (KDa)	98.9	98.5	100.3	98.3	100.2
	Isoelectric point	6.54	6.91	7.37	8.83	9.76
	Signal	SPI	SPII	SPII	SPI	SPI
Optimal condition	Temperature(°C)	30	20	20	10	20
	pH	6.0	8.0	8.0	8.0	8.0
	Salt	-	NaCl (230%) (100 mM)	NaCl (788%) (250 mM)	NaCl (140%) (5 mM)	KCl (485%) (50 mM)
Effect of divalent metal cation (5 mM)	Enhancer	Ca <sup>2+</sup> (142%)	Ca <sup>2+</sup> (135%)	Ca <sup>2+</sup> (192%)	Ba <sup>2+</sup> (130%)	Ca <sup>2+</sup> (200%)
		Ba <sup>2+</sup> (115%)	Ba <sup>2+</sup> (150%)	Ba <sup>2+</sup> (198%)	Mg <sup>2+</sup> (142%)	Ba <sup>2+</sup> (262%)
			Mg <sup>2+</sup> (133%)	Mg <sup>2+</sup> (187%)		Mg <sup>2+</sup> (179%)
	Inhibitor	Co <sup>2+</sup> (0%)	Co <sup>2+</sup> (33%)	Co <sup>2+</sup> (36%)	Ni <sup>2+</sup> (0%)	Co <sup>2+</sup> (1%)
		Hg <sup>2+</sup> (0%)	Hg <sup>2+</sup> (4%)	Hg <sup>2+</sup> (22%)	Cu <sup>2+</sup> (0%)	Hg <sup>2+</sup> (31%)
		Pb <sup>2+</sup> (0%)	Pb <sup>2+</sup> (28%)	Ni <sup>2+</sup> (1%)	Zn <sup>2+</sup> (11%)	Ni <sup>2+</sup> (1%)
		Ni <sup>2+</sup> (0%)	Ni <sup>2+</sup> (4%)	Cu <sup>2+</sup> (10%)		Cu <sup>2+</sup> (8%)
	Cu <sup>2+</sup> (0%)	Cu <sup>2+</sup> (1%)	Zn <sup>2+</sup> (10%)		Zn <sup>2+</sup> (5%)	
	Zn <sup>2+</sup> (0%)	Zn <sup>2+</sup> (11%)				
EDTA effect (5 mM)		19%	99%	95%	78%	97%
EGTA effect (5 mM)		62%	69%	59%	35%	84%
Enzyme activity (U/mg)	HP-F <sub>αIII</sub>	<sup>b</sup> 49.89 ± 3.18	61.86 ± 2.85	26.75 ± 1.77	24.25 ± 1.24	12.15 ± 1.48
	HP	0.75 ± 0.06	1.14 ± 0.03	-	-	-

<sup>a</sup>The residual activities were shown as the percentage of that (100%) of WT-BlexoHep.

<sup>b</sup>The enzyme activities are the means ± S.D. for at least three experiments. Source data are provided as a Source data file.

to the cold living environment of *P. arcticus*<sup>26</sup>. The optimal pH of the recombinant enzymes BlexoHep, 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 NaH<sub>2</sub>PO<sub>4</sub>-Na<sub>2</sub>HPO<sub>4</sub>, 8.0 in 50 mM NaH<sub>2</sub>PO<sub>4</sub>-Na<sub>2</sub>HPO<sub>4</sub> and 8.0 in 50 mM NaH<sub>2</sub>PO<sub>4</sub>-Na<sub>2</sub>HPO<sub>4</sub>, respectively (Table 1). In addition, the activities of these enzymes were significantly stimulated by 5 mM divalent cations such as Ca<sup>2+</sup> and Ba<sup>2+</sup> as well as Mg<sup>2+</sup> in most cases (Table 1). Moreover, the activities of the PL15\_2 enzymes were strongly inhibited by divalent metal ion chelating agents, in particular EGTA (Table 1), indicating that the divalent cations may involve in the catalytic mechanism of these types of Hepases. The activities of these Hepases is strongly inhibited by various heavy metal ions, such as Co<sup>2+</sup>, Hg<sup>2+</sup>, Ni<sup>2+</sup>, Pb<sup>2+</sup>, Cu<sup>2+</sup>, and Zn<sup>2+</sup> (Table 1). Moreover, salts (NaCl or KCl) at certain concentrations are strong stimulators for the above-mentioned enzymes except BlexoHep, 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). This promotion of enzyme activity by salts may be own to the presence of salt ions enhances the binding of substrates to enzyme by removing the water coat from HP chains<sup>27</sup>, meanwhile, salts also can affect the stability, solubility, and surface charge distribution of the enzymes to affect the activity.

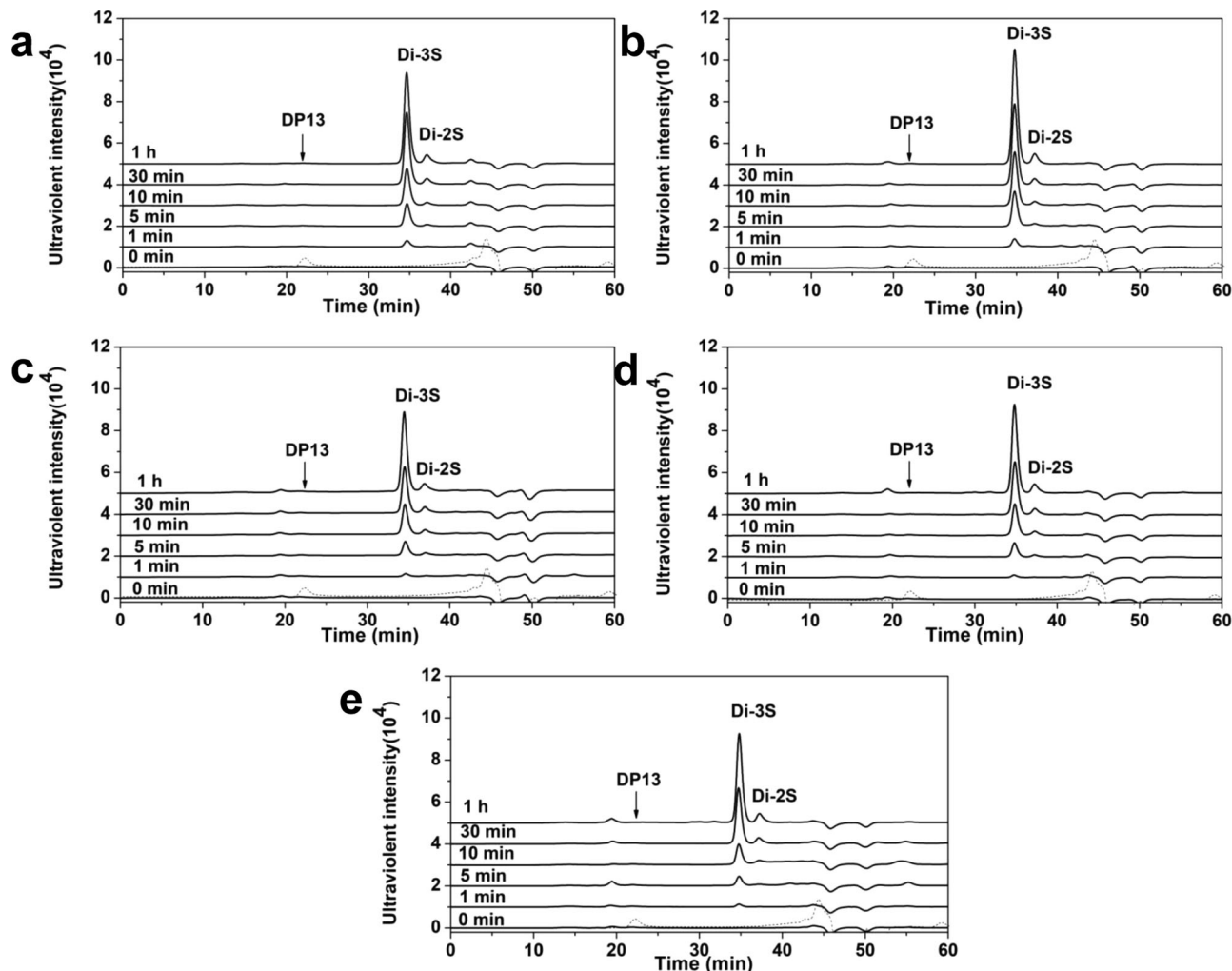
Finally, the enzymatic activities of these Hepases were determined under their respective optimum reaction conditions by using HP, HS, and HP-F<sub>αIII</sub> as substrates. As shown in Table 1, the enzymatic activities of BlexoHep, BCexoHep, BTexoHep, PAexoHep, and BFexoHep toward HP-F<sub>αIII</sub> were 49.89 U/mg protein (pH 6.0, 50 mM NaAc-HAc containing 5 mM Ca<sup>2+</sup>, 30 °C), 61.86 U/mg protein (pH 8.0, 50 mM Tris-HCl containing 5 mM Ba<sup>2+</sup> and 100 mM NaCl, 20 °C), 26.75 U/mg protein (pH 8.0, 50 mM NaH<sub>2</sub>PO<sub>4</sub>-Na<sub>2</sub>HPO<sub>4</sub> containing 5 mM Ba<sup>2+</sup> and 250 mM NaCl, 20 °C), 24.25 U/mg protein (pH 8.0, 50 mM NaH<sub>2</sub>PO<sub>4</sub>-Na<sub>2</sub>HPO<sub>4</sub> containing 5 mM Mg<sup>2+</sup> and 5 mM NaCl, 10 °C) and 12.15 U/mg protein (pH 8.0, 50 mM

NaH<sub>2</sub>PO<sub>4</sub>-Na<sub>2</sub>HPO<sub>4</sub> containing 5 mM Ba<sup>2+</sup> and 50 mM KCl, 20 °C), respectively; the activities of BlexoHep and BCexoHep toward HP polysaccharides were 0.75 and 1.14 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 Hepases in the PL15\_2 family, the HP polysaccharides were used as substrate for digestion by each enzyme in a time-course assay as described under “Methods”. As results shown in Supplementary Fig. 4a–e, all the tested PL15\_2 Hepases generated only disaccharides with specific absorbance at 232 nm, and no obvious amount of the larger oligosaccharide products were detected throughout the degradation process (Supplementary Fig. 4a–e), indicating that these enzymes maybe exolytic Hepases. However, the enzyme activities of these Hepases against HP polysaccharides are too weak to determine the action pattern of these enzymes exactly.

To confirm the possible exolytic character of the PL15\_2 Hepases, 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 O<sub>3</sub><sup>28</sup> and used as substrate for digestion by each enzyme in a time-course assay as described under “Methods”. The results showed that all enzymes could quickly degrade HP DP13, and just like the case against HP polysaccharides the digestion of HP DP13 produced only disaccharides with specific absorbance at 232 nm during the whole process of degradation (Fig. 1a–e), confirming that these Hepases should be exolytic enzymes and cleave HP chains from the reducing end to sequentially release unsaturated disaccharides.

To further confirm the substrate-degrading direction of the PL15\_2 Hepases, the HP DP13 was labeled with 2-AB and then treated with these enzymes, respectively. As the results shown in

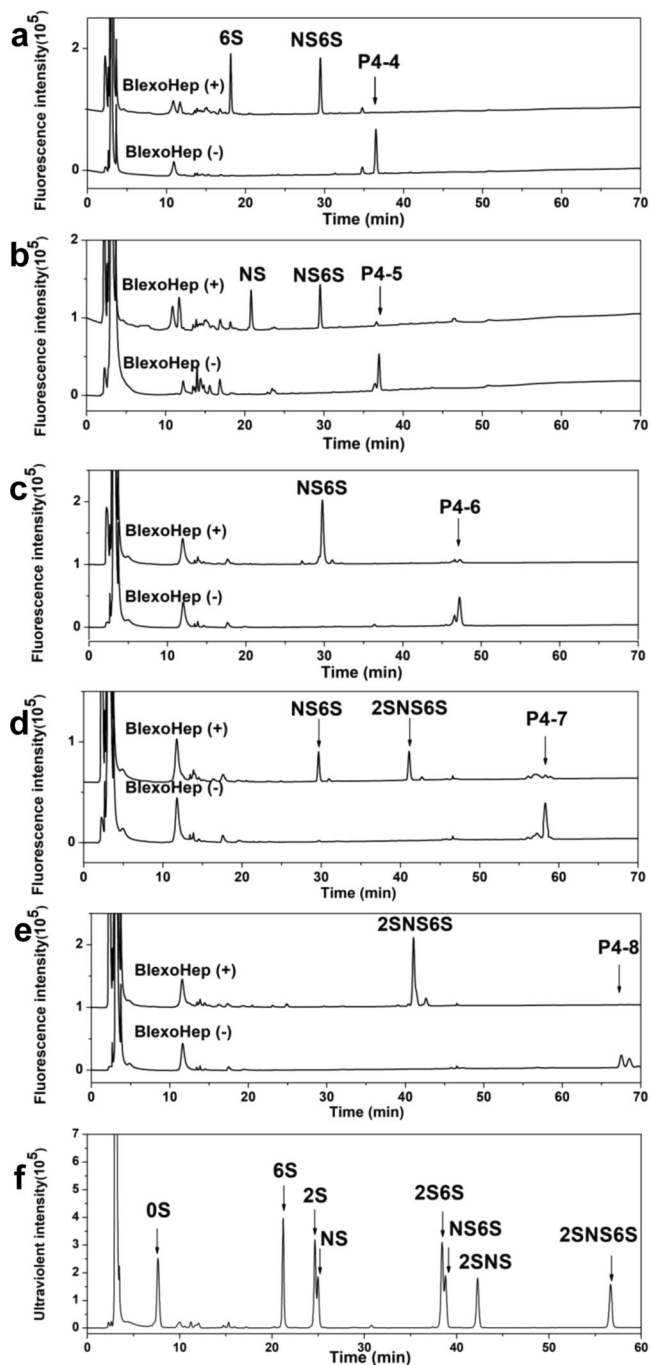


**Fig. 1** Time course experiments of degradation of saturated HP DP13 by PL15\_2 family enzymes. Saturated HP DP13 (30  $\mu$ g) was treated with 50 mU of BlexoHep (a), BCexoHep (b), BTexoHep (c), PAexoHep (d) or BFexoHep (e). Aliquots (3  $\mu$ g) were taken at different time points for SEC analysis as described under “Methods”. Saturated HP DP13 was prepared via the treatment of the Hepase III-resistant HP UDP14 fraction with  $O_3$  as described under “Methods”. The elution profile of HP DP13 was detected under 210 nm and shown as dash lines in the figures. Di-3S, the trisulfated HP disaccharide; Di-2S, the disulfated HP disaccharide. The elution positions of the disaccharides were determined by comparison with those of HP standard oligosaccharides (Supplementary Fig. 3).

Supplementary Fig. 5, the 2-AB-labeling at the reducing ends of HP DP13 completely inhibited the action of the PL15\_2 Hepases compared with the quick degradation of unlabeled HP DP13 by these enzymes (Fig. 1a–e), indicating that the introduction of 2-AB group at the reducing end of HP chain hindered the action of the PL15\_2 Hepases and further confirmed that these enzymes are exolytic Hepases that act from the reducing ends of substrates.

**Substrate specificity of BlexoHep.** To investigate the substrate specificity of the PL15\_2 Hepases, the capacity of the typical enzyme BlexoHep degrading HP tetrasaccharides with specific structures was estimated. Five structure-defined HP tetrasaccharide fractions P4-4 ( $\Delta$ UA1-4GlcNAc6S1-4GlcA1-4GlcNS6S and  $\Delta$ UA1-4GlcNAc6S1-4IdoA1-4GlcNS6S), P4-5 ( $\Delta$ UA1-4GlcNS1-4GlcA1-4GlcNS6S and  $\Delta$ UA1-4GlcNS1-4IdoA1-4GlcNS6S), P4-6 ( $\Delta$ UA1-4GlcNS6S1-4GlcA1-4GlcNS6S), P4-7 ( $\Delta$ UA2S1-4GlcNS6S1-4GlcA1-4GlcNS6S), and P4-8 ( $\Delta$ UA2S1-4GlcNS6S1-4IdoA2S1-4GlcNS6S) were prepared and structurally determined (Supplementary Figs. 6–8 and Supplementary Table 1) and individually used as substrates for digestion assays by BlexoHep. When these tetrasaccharides (5 pmol) were

treated with excess BlexoHep overnight, all could be almost completely degraded (Fig. 2a–e), indicating that BlexoHep 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 2). A further study shows that the enzyme activity of BlexoHep 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 3), respectively, suggesting that the activity of BlexoHep 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 BlexoHep 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 BlexoHep than IdoA (Supplementary Fig. 9 and Supplementary Table 4). 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.



**Fig. 2** Substrate specificity of BlexoHep. Structure-defined HP tetrasaccharides P4-4 (a), P4-5 (b), P4-6 (c), P4-7 (d), and P4-8 (e) were individually digested with BlexoHep, and the digests were labeled with 2-AB followed by analysis using anion-exchange HPLC on a Pack Polyamine II column eluted using a  $\text{NaH}_2\text{PO}_4$  gradient (0–1 M) and monitored using a fluorescence detector with excitation and emission wavelengths of 330 and 420 nm, respectively. The types of disaccharides were determined by comparing their elution positions with those of authentic 2-AB-derivatized unsaturated HP/H5 disaccharides indicated by arrows in (f). OS,  $\Delta\text{UA}(1-4)\text{GlcNAc}$ ; 6S,  $\Delta\text{UA}(1-4)\text{GlcNAc6S}$ ; 2S,  $\Delta\text{UA}2\text{S}(1-4)\text{GlcNAc}$ ; NS,  $\Delta\text{UA}(1-4)\text{GlcNS}$ ; 2S6S,  $\Delta\text{UA}2\text{S}(1-4)\text{GlcNAc6S}$ ; NS6S,  $\Delta\text{UA}(1-4)\text{GlcNS6S}$ ; 2SNS,  $\Delta\text{UA}2\text{S}(1-4)\text{GlcNS}$ ; 2SNS6S,  $\Delta\text{UA}2\text{S}(1-4)\text{GlcNS6S}$ .

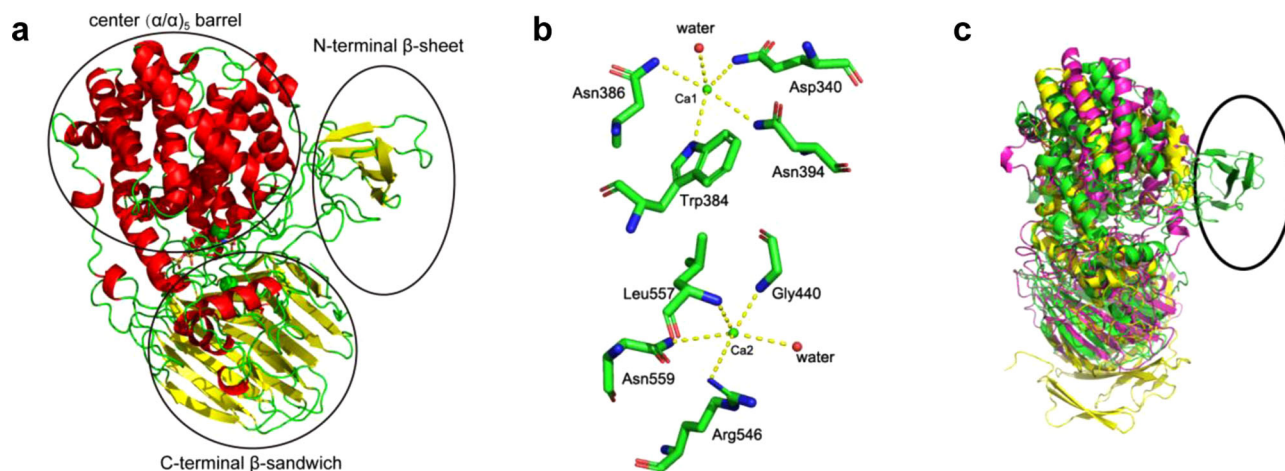
As we know, Hepase I, II, and III cannot effectively work on the 3-*O*-sulfated HP substrates. To test whether the exoHepases can degrade this kind of substrates, a synthetic HP pentasaccharide, fondaparinux (Arixtra), was used as substrate to investigate the possibility. Results show that all these exoHepases can degrade the fondaparinux to generate only the unsaturated trisulfated disaccharide  $\Delta\text{UA}2\text{S}(1-4)\text{GlcNS6S}(\text{OCH}_3)$  without the trisulfated HP disaccharide  $\Delta\text{UA}(1-4)\text{GlcNS3S6S}$ , which is same as the case of fondaparinux treated with Hepases I<sup>29</sup> (Supplementary Fig. 10). These results indicate that the exoHepases could not degrade the 3-*O*-sulfated substrates, too.

**Overall structure of BlexoHep.** To investigate the structural characteristics of the PL15\_2 proteins, the crystal structure of BlexoHep with an HP disaccharide product was solved at a resolution of 1.98 Å by preparing a crystal of selenomethionine (SeMet)-labeled BlexoHep in the presence of HP disaccharide products followed by single-wavelength anomalous dispersion (SAD) phasing. Furthermore, a crystal of an inactive BlexoHep-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 BlexoHep with the product as a template. In the crystal structures, each asymmetric unit consists of one BlexoHep monomer, an unsaturated disaccharide ( $\Delta\text{UA}2\text{S}1-4\text{GlcNS6S}$ ) or tetrasaccharide ( $\Delta\text{UA}1-4\text{GlcNS6S}1-4\text{IdoA}2\text{S}1-4\text{GlcNS6S}$ ), and two  $\text{Ca}^{2+}$  ions that were confirmed by inductively coupled plasma-mass spectrometry (ICP-MS) analysis. Notably, the BlexoHep 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).

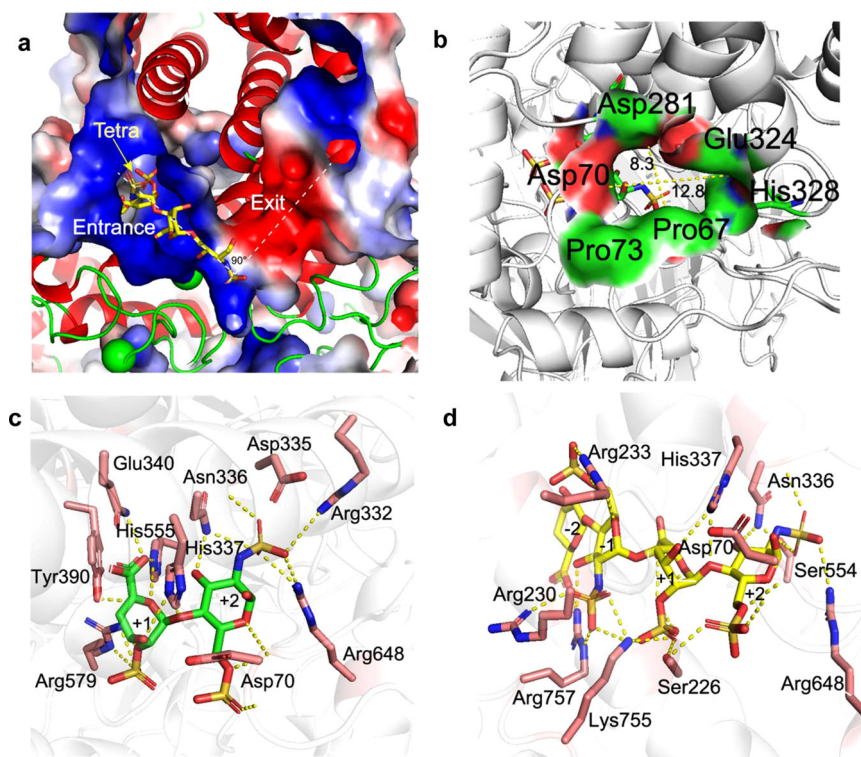
The structure of BlexoHep consists of three domains: an N-terminal small  $\beta$ -sheet domain (Ala<sup>24</sup>–Asn<sup>150</sup>) composed of a large loop and a five-stranded  $\beta$ -sheet ( $\beta 1$ – $\beta 5$ ), a central ( $\alpha/\alpha$ )<sub>5</sub> barrel domain (Pro<sup>151</sup>–Glu<sup>515</sup>) containing 20  $\alpha$ -helices ( $\alpha 1$ – $\alpha 20$ ) to form an ( $\alpha/\alpha$ )<sub>5</sub> incomplete toroid structure, and a C-terminal  $\beta$ -sandwich domain (Leu<sup>516</sup>–Pro<sup>864</sup>) containing 23  $\beta$ -strands and 5 short  $\alpha$ -helices that fold into a three-layered  $\beta$ -sheet sandwich structure ( $\alpha 21$ – $\alpha 25$  and  $\beta 6$ – $\beta 28$ ) (Fig. 3a). The two  $\text{Ca}^{2+}$  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 BlexoHep was performed using the DALI server<sup>30</sup>. The results showed that the closest structure to BlexoHep is the exo-type alginate lyase Atu3025 (PDB code 3AFL)<sup>31</sup> with a Z-score of 31.7 and an r.m.s.d. of 3.1 Å on 766 aligned Ca atoms. Additionally, BlexoHep 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 2366<sup>21</sup>, 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 2366<sup>20</sup>. However, BlexoHep has an additional small  $\beta$ -sheet domain (Ala<sup>24</sup>–Asn<sup>150</sup>) at the N-terminus compared with Hepase II and III (Fig. 3c).

**The unique L-shaped catalytic cavity of BlexoHep.** Compared to the open catalytic cavities of Hepase I, II, and III (Supplementary Fig. 11a–c), BlexoHep possesses a semiopen, narrow, elongated, and L-shaped catalytic cavity composed of residues from the three domains of the N-terminal small  $\beta$ -sheet, central ( $\alpha/\alpha$ )<sub>5</sub> barrel and C-terminal  $\beta$ -sandwich (Supplementary Fig. 11d



**Fig. 3 Overall structure of BlexoHep.** **a** Overall structure (stereo diagram) of BlexoHep.  $\alpha$ -Helices and  $\beta$ -sheets are colored red and yellow, respectively. The three domains are labeled with elliptical coils. **b**  $\text{Ca}^{2+}$  binding sites. The pentacoordination of  $\text{Ca}1$  and  $\text{Ca}2$  ions in the BlexoHep 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:  $\text{NE}2^{\text{Asp}340}$ , 3.2 Å;  $\text{NE}1^{\text{Trp}384}$ , 3.3 Å;  $\text{ND}2^{\text{Asn}386}$ , 3.1 Å;  $\text{ND}2^{\text{Asn}394}$ , 3.3 Å; and water molecule, 3.3 Å. The distances between the residues and  $\text{Ca}2$  atom are as follows:  $\text{N}^{\text{Gly}440}$ , 3.1 Å;  $\text{NH}1^{\text{Arg}546}$ , 3.4 Å;  $\text{N}^{\text{Leu}557}$ , 3.3 Å;  $\text{ND}2^{\text{Asn}559}$ , 3.5 Å; and water molecule, 3.2 Å. **c** Structural overlap of BlexoHep 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 BlexoHep (PDB code 6LJL) are colored pink, yellow, and green, respectively. The additional N-terminal  $\beta$ -sheet domain of BlexoHep is labeled with an elliptical coil.



**Fig. 4 The L-shaped catalytic tunnel and relative sites of BlexoHep.** 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 BlexoHep. **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 BlexoHep-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.

and Fig. 4a). The unique catalytic tunnel consists of two almost mutually perpendicular sections: the positively charged section generated by residues from  $\alpha 6$ ,  $\alpha 8$ ,  $\alpha 10$ ,  $\alpha 12$ ,  $\alpha 14$ , and the loops connecting  $\alpha 9/\alpha 10$ ,  $\alpha 18/\alpha 19$ ,  $\beta 8/\beta 9$ , and  $\beta 10/\alpha 21$ , 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  $\alpha 7$ ,  $\alpha 8$ ,  $\alpha 9$ , and the big loop of the N-terminal small  $\beta$ -sheet domain, especially residues Phe<sup>62</sup> to Pro<sup>80</sup>,

which acts as the exit to facilitate the release of the disaccharide products (Supplementary Fig. 12).

**Substrate-binding and product-releasing sites of the L-shaped tunnel.** According to the hydrogen-bonding networks, a series of residues, such as Asp<sup>70</sup>, Ser<sup>226</sup>, Arg<sup>230</sup>, Arg<sup>233</sup>, Arg<sup>332</sup>, Asn<sup>336</sup>, His<sup>337</sup>, Gln<sup>340</sup>, Tyr<sup>390</sup>, Ser<sup>554</sup>, His<sup>555</sup>, Arg<sup>579</sup>, Arg<sup>648</sup>, Lys<sup>755</sup>, and Arg<sup>757</sup>, participate in the binding of the disaccharide product ( $\Delta$ UA2S1-4GlcNS6S) and the tetrasaccharide substrate ( $\Delta$ 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 His<sup>337</sup>, Tyr<sup>390</sup> and His<sup>555</sup> match well with three key residues His<sup>202</sup>, Tyr<sup>257</sup> and His<sup>406</sup> in the active center of Hepase II from *P. heprinus* DSM 2366 (Supplementary Fig. 13), indicating that they may also play key roles in the catalytic mechanism of BlExoHep. This speculation was proven by mutation of three residues to Ala, which caused almost complete loss of enzymatic activity (Supplementary Fig. 14a, 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).

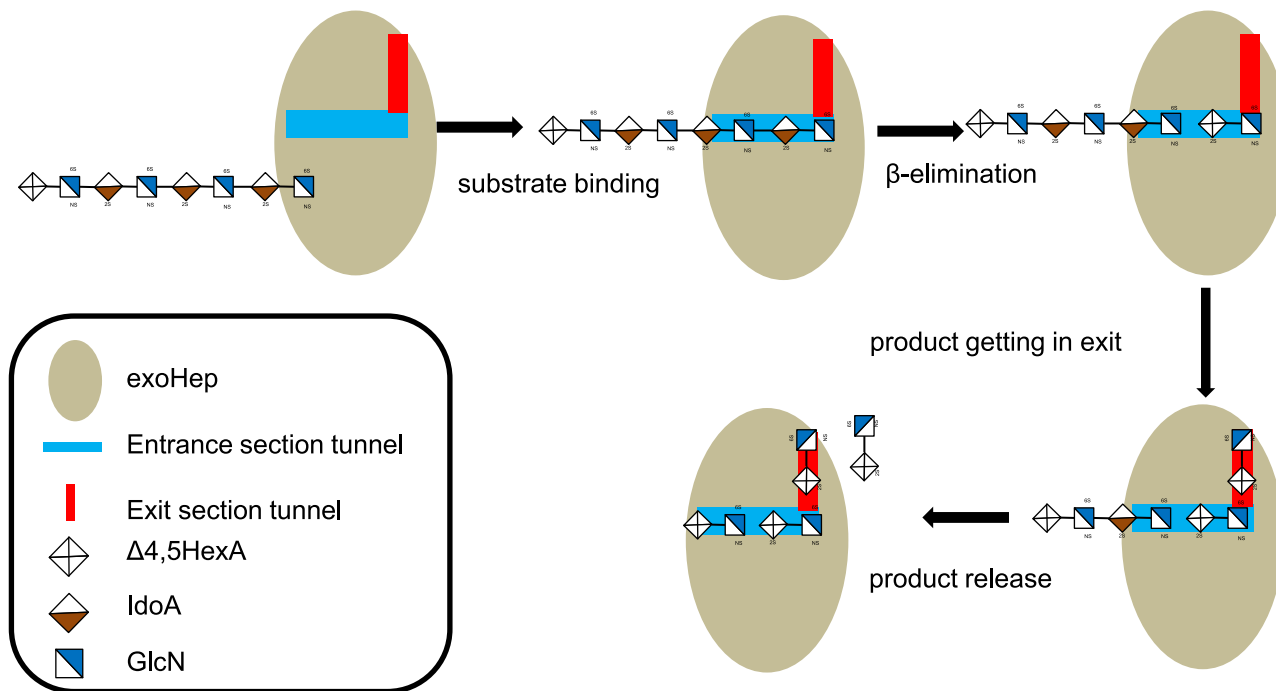
The negatively charged residues Pro<sup>67</sup>, Asp<sup>70</sup>, Pro<sup>73</sup>, Asp<sup>281</sup>, Glu<sup>234</sup>, and Asp<sup>335</sup> from the N-terminal small  $\beta$ -sheet and central ( $\alpha/\alpha$ )<sub>5</sub> barrel domains compose the negatively charged section of the tunnel for quick delivery of the disaccharide product (Fig. 4b). Residue Asp<sup>335</sup> 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 BlExoHep (Supplementary Fig. 14a). The other residues compose the tunnel exit to approximately 8.3 Å in width (measurement of CG<sup>Pro67</sup> to OD<sup>2Asp281</sup>) and 12.8 Å in length (measurement of O<sup>Asp70</sup> to O<sup>Glu324</sup>). The mutation of negatively charged Asp residues to uncharged Ala or Asn or positively charged His in the exit tunnel caused a dramatic decrease of enzyme activity, such as the enzyme activities of mutants BlExoHep-P67A, BlExoHep-D70N, BlExoHep-D281A, BlExoHep-D281N and BlExoHep-D281H are reduced by 90%, 98%, 70%, 68%, and 85%, respectively, compared with that of the WT-BlExoHep (Supplementary Fig. 14a). These results show that the negatively charged residues play very important roles in the release process of disaccharide products.

**Action mode switch of BlExoHep.** To determine the key factors leading to the exolytic character of BlExoHep, the N-terminal small  $\beta$ -sheet domain (residues 1–150) of BlExoHep was truncated firstly, which however caused a complete loss of enzyme activity (Supplementary Fig. 14a). This finding can confirm the essential role of the additional small N-terminal domain for the activity of the exoHepases but cannot reveal the role of this domain for the exolytic property of this enzyme. Then, we tried to investigate the effect of electrical property of the exit tunnel on the action mode of BlExoHep by the mutation of the acidic residues in tunnel exit. Consistent with the observation above, the mutants BlExoHep-D70H-D281H-D335H, BlExoHep-D70H-D281N-D335H, and BlExoHep-D70N-D281N-D335N show very weak enzyme activities toward HP polysaccharides as well as the HP-F<sub>all</sub> substrates (Supplementary Fig. 14a). Interestingly, in a time-course assay, we found that the mutant BlExoHep-D70H-D281N-D335H could degrade HP polysaccharides to produce many larger size oligosaccharides during the degradation process (Supplementary Fig. 14c), which was significantly different from the action mode of the WT-BlExoHep (Supplementary Fig. 4a)

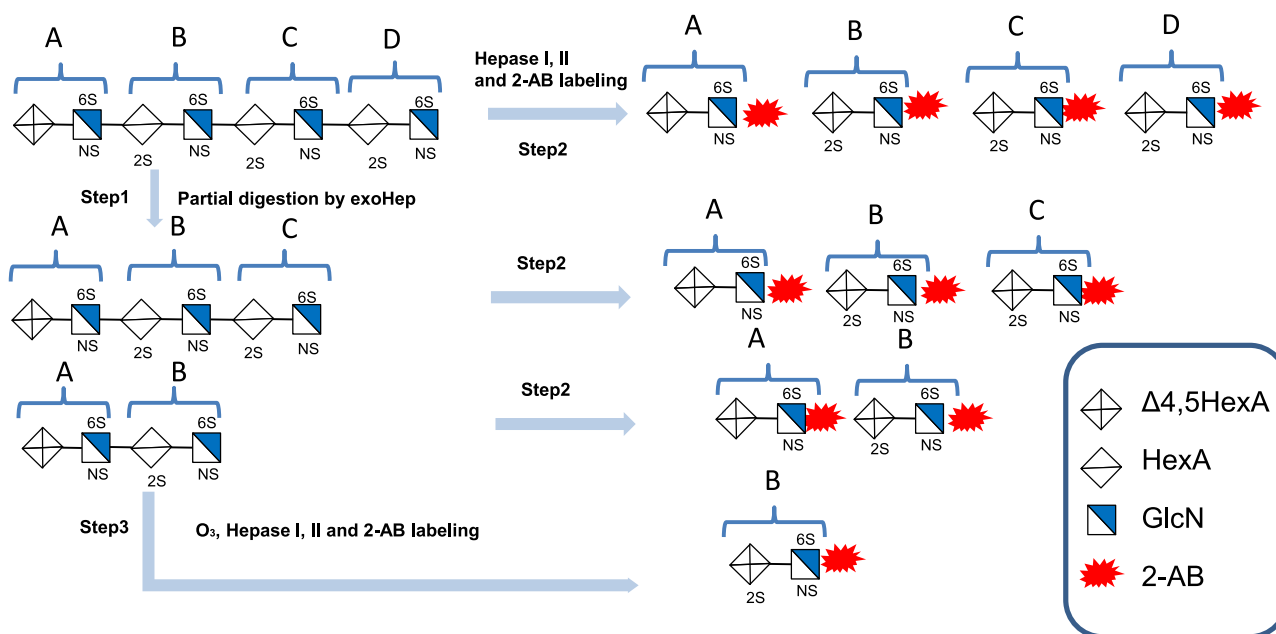
and thus showed that this mutant preliminarily possesses certain endolytic activity. These results indicate that the exolytic character of the exoHepases maybe attributed to the distribution of the negative charge in the exit tunnel but the exact reason remains to be investigated future. And the proposed exolytic mode of BlExoHep was shown in Fig. 5.

**Roles of Ca<sup>2+</sup> ions.** According to the initial electron density map, there are two Ca<sup>2+</sup> ions Ca1 and Ca2 in the structure of BlExoHep as described above. The Ca1 is surrounded by the nitrogen and oxygen atoms of residues Asp<sup>340</sup>, Trp<sup>384</sup>, Asn<sup>386</sup>, Asn<sup>394</sup> and a water molecule in near pentacoordination with distances of 3.1–3.3 Å (Fig. 3b). In contrast, the Ca2 is coordinated by the nitrogen and oxygen atoms of residues Gly<sup>440</sup>, Arg<sup>546</sup>, Leu<sup>557</sup>, Asn<sup>559</sup> and a water molecule with distances of less than 3.5 Å (Fig. 3b). To investigate the roles of these two Ca<sup>2+</sup> ions in the catalysis of BlExoHep, the residues surrounded Ca1 and Ca2 were individually mutated to Ala and the enzyme activity of each mutant were measured. As results shown in Supplementary Fig. 14a, the activity of BlExoHep was destroyed to varying degree by the mutation. Especially, the mutation of Asp<sup>340</sup> surrounded Ca1 caused the enzyme to completely lose the ability to degrade HP (Supplementary Fig. 14a). These results indicate that both Ca1 and Ca2 play important roles in the catalysis of BlExoHep. In comparison, Ca1, which is much closer to substrate than Ca2, may directly participate the catalytic process of BlExoHep, while Ca2 may play an important role in the structural stabilization of enzyme. Obviously, both biochemical and structural evidence prove that BlExoHep is a Ca<sup>2+</sup>-dependent enzyme.

**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<sub>all</sub> 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. 15. To further obtain the preliminary disaccharide sequences of these octasaccharide preparations, an enzymatic sequencing method was established based on the exolytic activity of BlExoHep (Fig. 6). Taking the sequencing of subfraction P8-4 as an example, an aliquot of P8-4 was partially digested by BlExoHep for isolating and preparing its nonreducing end hexasaccharide (UDP<sub>NE-P8-4</sub><sup>6</sup>) and tetrasaccharide (UDP<sub>NE-P8-4</sub><sup>4</sup>) fractions as described under “Methods”. Then, the disaccharide compositions of P8-4 (UDP<sub>NE-P8-4</sub><sup>8</sup>), UDP<sub>NE-P8-4</sub><sup>6</sup> and UDP<sub>NE-P8-4</sub><sup>4</sup> were analyzed by complete digestion with Hepases I and II followed by 2-AB labeling and anion-exchange HPLC. As shown in Supplementary Fig. 16d, the complete digestion of UDP<sub>NE-P8-4</sub><sup>8</sup> with Hepases produced two unsaturated disaccharides  $\Delta$ UA1-4GlcNS6S and  $\Delta$ UA2S1-4GlcNS6S with a molar ratio of 1:3, and the molar ratios of these two disaccharides in its nonreducing end hexasaccharide UDP<sub>NE-P8-4</sub><sup>6</sup> and tetrasaccharide UDP<sub>NE-P8-4</sub><sup>4</sup> were 1:2 and 1:1, respectively (Supplementary Table 5). By comparing the molar ratio change of disaccharides in UDP<sub>NE-P8-4</sub><sup>8</sup> and UDP<sub>NE-P8-4</sub><sup>6</sup>, 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<sub>NE-P8-4</sub><sup>6</sup> and UDP<sub>NE-P8-4</sub><sup>4</sup> too (Fig. 6 and Supplementary Table 5). To determine the type of disaccharide at B site in P8-4, the UDP<sub>NE-P8-4</sub><sup>4</sup> was treated with O<sub>3</sub> 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. 16d a single disaccharide peak corresponding to  $\Delta$ UA2S1-



**Fig. 5 Schematic diagram of the proposed exolytic mode of BlexoHep.** The exolytic mode of BlexoHep is proposed as: 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; then the bound HP chain moves into the tunnel, and the reducing end disaccharides are cleaved via a  $\beta$ -elimination mechanism; finally, the resulting disaccharide product is pushed into the negatively charged exit and rapidly released.



**Fig. 6 Strategy for the exo-sequencing of unsaturated HP octasaccharides.** The strategy is divided into three steps: first, the octasaccharide was partially digested by BlexoHep for isolating and preparing its nonreducing end hexasaccharide (UDP<sub>6NE</sub>) and tetrasaccharide (UDP<sub>4NE</sub>) fractions. Then, the disaccharide compositions of UDP<sub>8NE</sub>, UDP<sub>6NE</sub> and UDP<sub>4NE</sub> were analyzed by complete digestion with Hepsases I and II followed by 2-AB labeling and anion-exchange HPLC. By comparing the molar ratio change of disaccharides in UDP<sub>8NE</sub> and UDP<sub>6NE</sub>, the reducing end disaccharide (D) could be deduced, and similarly the disaccharide located at C site could be determined by comparing the disaccharide compositions of UDP<sub>6NE</sub> and UDP<sub>4NE</sub> too. Finally, to determine the type of disaccharide at B site, the UDP<sub>4NE</sub> was treated with O<sub>3</sub> to remove the unsaturated uronic acid at the nonreducing end and further digested with Hepsases I and II for disaccharide composition analysis. Combining all the disaccharides' information, the A site could be deduced.

4GlcNS6S was detected, indicating that the disaccharide at B site was HexUA2S1-4GlcNS6S too. Finally, based on the disaccharide composition of UDP<sub>4NE-P8-4</sub>, the nonreducing end disaccharide (A) of P8-4 must be  $\Delta$ UA1-4GlcNS6S. Taken together, the

preliminary sequence of P8-4 can be deduced as  $\Delta$ UA1-4GlcNS6S1-4HexUA2S1-4GlcNS6S1-4HexUA2S1-4GlcNS6S1-4HexUA2S1-4GlcNS6S (Supplementary Table 6). 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. 16 and Supplementary Table 5 and Supplementary Table 6.

To verify the enzymatic sequencing, the compositions and sequences of these octasaccharides were subjected to ESI-MS (Supplementary Fig. 17 and Supplementary Table 7) and ESI-MS/MS (Supplementary Figs. 18–22). Notably, the monosulfated disaccharide located at the nonreducing end of P8-3 or P8-5 could not be deduced by the enzymatic method because of lack of its corresponding disaccharide standard, and the exact structure of the disaccharide remains for further studies. In addition, we should note that the digestion of UDP<sub>4</sub><sup>NE-P8-3</sup> produces two disaccharides ΔUA(1-4)GlcNAc(S) and ΔUA2S(1-4)GlcNS6S but the proportion of ΔUA2S(1-4)GlcNS6S is too low to hold the line with that of ΔUA(1-4)GlcNAc(S) as shown in the Supplementary Fig. 16c. The reason why the detected ratio of these two disaccharides in the nonreducing end UDP<sub>4</sub> of P8-3 is so different is not clear, which may be due to the different labeling efficiency of the two disaccharides or the purity of P8-3.

## Discussion

In this study, the discovery of the Hepase BlexoHep with exolytic activity led to the identification of a Hepase family, PL15\_2, a subfamily of the PL15 family that has been known as the assembly of exolytic alginate lyases<sup>31</sup>, which is different from Hepases I, II, and III, belong to the PL13, PL21, and PL12 families<sup>16</sup>, respectively. In this family, five exolytic Hepases, including BlexoHep, BCexoHep, BTexoHep, PAexoHep, and BFeexoHep, have been identified and studied in detail, 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 BlexoHep, and deletion of this domain will cause a complete loss of BlexoHep activity, indicating that the conserved DUF4962 superfamily domain is essential to the activities 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 domains<sup>13</sup> and linkage regions (GlcA-Gal-Gal-Xyl)<sup>3,32</sup> 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<sub>III</sub>. In fact, by using structure-defined HP tetrasaccharides and pentasaccharides as substrates, we found that the enzymatic activity of BlexoHep increased with the increase of sulfation degree of the substrates, except for the 3-O-sulfated substrates which resistant to the activities of the exoHepases, 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 ΔUA2S1-4GlcNS6S1-4IdoA2S1-4GlcNS6S. Interestingly, we also find that nonsulfated GlcA residues are more susceptible to BlexoHep 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 BlexoHep and Hepase II partially overlap, indicating that their catalytic mechanism might have some

features in common. As shown in this study, the residues His<sup>337</sup>, Tyr<sup>390</sup>, and His<sup>555</sup> of BlexoHep corresponding to the three key residues His<sup>202</sup>, Tyr<sup>257</sup>, and His<sup>406</sup> of Hepase II<sup>22</sup> play key roles in the catalytic mechanism of BlexoHep, 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 His-Tyr acts as the Brønsted base and acid<sup>19,22,23</sup>. However, unlike in the case of the His<sup>337</sup> to Ala mutation that completely destroyed the ability of BlexoHep to cleave the tetrasaccharide ΔUA2S1-4GlcNS6S1-4IdoA2S1-4GlcNS6S (Supplementary Fig. 14b), a previous study showed that the mutation of His<sup>202</sup> to Ala in Hepase II did not completely destroy the enzymatic activity toward this tetrasaccharide<sup>22</sup>, 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<sup>19–21</sup>, BlexoHep 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 BlexoHep is an L-shaped tunnel, which is similar to that of Hepase I from *B. thetaiotaomicron* VPI-5482<sup>19</sup>. However, the L-shaped tunnel of Hepase I is open, longer, and whole 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 BlexoHep 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. When mutate the acidic residues in the negatively charged exit to the basic amino acids, the mutant protein BlexoHep-D70H-D281N-D335H exhibited the endolytic activity. These results indicate that the exolytic character of the exoHepases might affected by the negative charge region in the semiopen L-shaped tunnel and the key reason contributes to the exolytic characters of the exoHepases are still remains for the further studies.

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 Asp<sup>335</sup>, the resulting disaccharide product is pushed into the negatively charged exit and rapidly released with assistance from Pro<sup>67</sup> and Asp<sup>281</sup> (Fig. 5). Notably, this exolytic mechanism is different from those of other enzymes in the PL15 family, such as the exo-type alginate lyase Atu3025<sup>31</sup>. 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 BlexoHep 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<sup>33–35</sup>. 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<sup>33,36</sup>. In this study, by utilizing the exolytic feature of exoHepases that sequentially cleave HP chains from the reducing ends, an enzymatic sequencing method (Fig. 6) was established and successfully determined the 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 degradation<sup>33</sup> and Ion Mobility Mass Spectrometry<sup>35</sup>.

In conclusion, the identification of an exolytic Hepase family can provide not only enzymatic tools for structural and functional studies, particularly the sequencing of HP/HS chains, but also a clue for the discovery of additional 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<sub>3</sub>CN), 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<sup>37,38</sup>. The anti-Hepase III fraction of HP (HP-F<sub>III</sub>) 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 size exclusion chromatography (SEC) on a Superdex Peptide 10/300 GL column<sup>39</sup>.

**Sequences analysis of the gene and protein of BlexoHep.** The gene and protein sequence of BlexoHep were downloaded from the NCBI database (GenBank: EDV07780.1). Then, the analysis of the similarity between BlexoHep 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 BlexoHep and its derivative with selenomethionine.** To express BlexoHep in *E. coli* strains, the gene *bioxhep* was constructed onto plasmid pET-30a (+) (Novagen). Then, the plasmid pET-30a-*bioxhep* was transformed into *E. coli* BL21 (DE3) cells. *E. coli* cells harboring an expression vector (pET30a-*bioxhep*) were initially cultured in LB broth and then induced to start the expression of wild-type BlexoHep (WT-BlexoHep) proteins with 5 mM IPTG (isopropyl 1-thio-β-D-galactopyranosid) when the cell density reached an A<sub>600</sub> of 0.6–1.0. An overexpression system for the BlexoHep derivative with selenomethionine (SeMet-BlexoHep) was also constructed in *E. coli* BL21 (DE3) cells as previously reported<sup>40</sup>. Cells grown overnight in LB medium were harvested by centrifugation at 3700 × g for 5 min and washed twice with M9 medium. Then, the cells were transferred into M9 medium containing 1 mM MgSO<sub>4</sub>, 0.1 mM CaCl<sub>2</sub>, and 0.4% (w/v) glucose and cultured at 37 °C. When the

A<sub>600</sub> 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 8000 × 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 × 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 BlexoHep 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 BlexoHep above. These genes, including Hepase II (BTexoHep) from strain *Bacteroides thetaiotaomicron* VPI-5482, heparan-sulfate lyase (BCexoHep) from *Bacteroides cellulosilyticus* WH2, Hepase II/III-like protein (BExoHep) (GenBank: EEX44367.1) from strain *Bacteroides finegoldii* 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 BlexoHep, respectively. The expression plasmids were transferred into *E. coli* BL21 (DE3) cells and over-expressed. The activities of the expression products were analyzed as described for BlexoHep above.

**Biochemical characterization of the recombinant PL15\_2 family of proteins.** The optimal conditions for each PL15\_2 protein were determined as described previously<sup>41</sup>. Briefly, to determine the optimal pH for each protein, HP-F<sub>III</sub> (1 mg/ml) was digested with the indicated protein (0.5 μg) 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<sub>2</sub>PO<sub>4</sub>-Na<sub>2</sub>HPO<sub>4</sub> buffer (pH 6.0–8.0), and 50 mM Tris-HCl buffer (pH 7.0–10.0) in a total volume of 100 μl, 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<sub>III</sub> 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<sub>III</sub> 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<sub>III</sub> as substrates. Briefly, each enzyme (1 μg) was added to 1 mg/ml HP, HS or HP-F<sub>III</sub> 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 μl were withdrawn in duplicate, boiled for 10 min, and then cooled in ice-cold water for 10 min. After being centrifuged at 15,000 × 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. The activity was calculated according to the change of absorbance per minute applying a molar extinction coefficient of 3800 M<sup>-1</sup> cm<sup>-1</sup> for products [1 U = 1 μmol of (ΔA)-containing product formed per min].

**Degradation pattern of substrates by PL15\_2 enzymes.** To determine the degradation patterns of the PL15\_2 enzymes, the digests of HP (1 mg/ml) by the enzymes (10 unit/ml) were traced at each optimal condition. Aliquots (20 μg) of the reactants were removed at different time points and analyzed through SEC on a Superdex Peptide 10/300 GL column eluted with 0.20 M NH<sub>4</sub>HCO<sub>3</sub> 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.

To further confirm the digest patterns and directions of the enzymes, 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 μg of HP UDP14 was treated with O<sub>3</sub> to remove unsaturated uronic acid at the nonreducing ends<sup>28</sup>, which resulted in saturated HP DPI3 without specific absorbance at 232

nm. After desalting by gel filtration, saturated HP DP13 (30  $\mu$ g) was digested with various PL15\_2 family proteins (50 mU) and traced under their optimal conditions. Aliquots (3  $\mu$ g) of the products were removed at different time points and analyzed through SEC as described above.

To further confirm the digest directions of the enzymes, the 1  $\mu$ g of the HP DP13 was 2-AB labeled in advance and then treated with 5 mU of each PL15\_2 enzymes for 4 h. Then the products were analyzed by SEC using a fluorescence detector with excitation and emission wavelengths of 330 and 420 nm, respectively.

**Substrate specificity of BlexoHep.** To determine the substrate specificity of BlexoHep, 5 pmol of various structure-defined HP tetrasaccharides (P4-4 (mixture of  $\Delta$ UA1-4GlcNAc6S1-4GlcA1-4GlcNS6S and  $\Delta$ UA1-4GlcNAc6S1-4IdoA1-4GlcNS6S), P4-5 (mixture of  $\Delta$ UA1-4GlcNS1-4GlcA1-4GlcNS6S and  $\Delta$ UA1-4GlcNS1-4IdoA1-4GlcNS6S), P4-6 ( $\Delta$ UA1-4GlcNS6S1-4GlcA1-4GlcNS6S), P4-7 ( $\Delta$ UA2S1-4GlcNS6S1-4GlcA1-4GlcNS6S), and P4-8 ( $\Delta$ UA2S1-4GlcNS6S1-4IdoA2S1-4GlcNS6S)) (Supplementary Figs. 6, 7, 8 and Supplementary Table 1) were degraded by 100 mU of BlexoHep protein overnight. Then, the products were 2-AB labeled and analyzed by anion-exchange HPLC on a Pack Polyamine II column eluted using a  $\text{NaH}_2\text{PO}_4$  gradient (0–1 M) and monitored using a fluorescence detector with excitation and emission wavelengths of 330 and 420 nm<sup>42</sup>, respectively.

To determine the preference of BlexoHep 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 BlexoHep for 0 min, 1 min, 5 min, 3 min or 12 h and then boiled to stop the reactions. Resulted oligosaccharides in each reactant were purified and pooled through SEC on a Superdex Peptide 10/300 GL column eluted with 0.20 M  $\text{NH}_4\text{HCO}_3$ , and freeze-dried repeatedly to remove  $\text{NH}_4\text{HCO}_3$ . Each oligosaccharide sample was dissolved in 0.5 ml of 99.9%  $\text{D}_2\text{O}$  and lyophilized for twice, and was finally dissolved in 0.5 ml  $\text{D}_2\text{O}$  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  $^1\text{H}$  NMR spectroscopy.  $^1\text{H}$  NMR spectroscopy was performed on a JNM-ECP600 (JEOL, Japan) instrument set at 600 MHz and analyzed on MestReNova (9.0.1).

To determine if the exoHepases can degrade the 3-O-sulfated HP substrates, a synthetic HP pentasaccharide, fondaparinux (Arixtra), GlcNS6S1-4GlcA1-4GlcNS3S6S1-4IdoA1-4GlcNS6S, was used as substrate to test the preference of BlexoHep and other PL15\_2 enzymes. Briefly, 5  $\mu$ g Arixtra was treated with each enzyme and then the product was analyzed by anion-exchange HPLC on a Pack Polyamine II column eluted using a linear gradient of  $\text{NaH}_2\text{PO}_4$  from 16 mM to 450 mM and monitored at 232 nm using a UV detector.

**Mutation of BlexoHep.** To investigate the key residues involved in the active center of BlexoHep, the residues in and around the active center were individually mutated to Ala or His and Asn, using the Fast Mutagenesis Kit V2 from Vazyme Biotech Co., Ltd (Nanjing, China). The primers are listed in Supplementary Table 8. The residual activities of the mutants were measured by using HP-F<sub>act</sub> as substrate, and compared with that (100%) of the WT-BlexoHep. Furthermore, the activity of the mutant BlexoHep-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.

To investigate the action mode of the mutant BlexoHep-D70H-D281N-D335H, the degradation process of HP polysaccharide by this mutant was traced in a time-course assay as described for the WT-BlexoHep above.

**Crystallization and data collection.** The protein BlexoHep (12 mg/ml) was pre-mixed 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-BlexoHep protein was prepared under the same conditions used for BlexoHep except for at pH 5.3. The inactive mutant BlexoHep-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 Facility<sup>43</sup>. The initial diffraction datasets were processed by HKL2000. Data collection statistics are shown in Supplementary Table 9.

**Structure determination and refinement.** The selenium single-wavelength anomalous dispersion (Se-SAD) phasing method was used to determine the phase of BlexoHep. Experimental phases were solved using the Phenix program<sup>44</sup> AutoSol. Initial model building was performed using the Phenix program Auto-Build. Refinement of the BlexoHep structure was performed alternately by the Phenix program Refine and Coot<sup>45</sup>. The ion type was confirmed by inductively coupled plasma-mass spectrometry (ICP-MS) analysis. The structure of the BlexoHep-Y390A/H555A complex with a tetrasaccharide was determined by molecular replacement with Phaser using the complex structure of BlexoHep with a disaccharide as the search model. Data collection and refinement statistics are

summarized in Supplementary Table 9. 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<sub>act</sub> was further subfractionated by anion-exchange HPLC on a Propac 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 SEC 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 BlexoHep proteins for 5 min, respectively. The UDP6 and UDP4 products were collected and freeze-dried repeatedly to remove  $\text{NH}_4\text{HCO}_3$  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  $\text{NaH}_2\text{PO}_4$  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  $\text{O}_3$ , 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 reported<sup>46</sup>. 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  $\times$  150 mm, 200 Å, Phenomenex, Torrance, CA). Mobile phase A was 5 mM ammonium acetate, while mobile B was 5 mM 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  $\mu$ l/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 gas 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.

To further analyze and verify the sequences of the octasaccharides, each desalted fraction dissolved in 50% methanol containing 3 mM NaOH was directly introduced to the ESI-MS interface of a Thermo LTQ-Orbitrap XL MS (Thermo Fisher Scientific, San Jose, CA) as previously reported<sup>47</sup>. The MS parameters included a spray voltage of  $-3.5$  kV, a capillary voltage of  $-40$  V, a tube lens voltage of  $-50$  V, a capillary temperature of 275 °C, a sheath gas flow rate of 10, an aux gas flow rate of 2. The MS/MS parameters were set as following: Iso width ( $m/z$ ), 3.0; normalized collision energy, 55.0–60.0.

**Statistics and reproducibility.** Each experiment was done at least three times by triplicates. Statistical analyses were performed using Excel and Origin 8.0. Error bars represent means of triplicates  $\pm$  SD. For comparison of the statistical differences between two groups, Student's *t*-test (two-sided) was carried out for statistical analysis.

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Data availability

All data supporting the findings of this study are available within the paper (and its Supplementary Information files). All relevant data generated during this study or analyzed in this published article (and its Supplementary Information files) are available from the corresponding author on reasonable request. The atomic coordinates and structure factors of the structures in this study have been deposited in the Protein Data Bank, [www.pdb.org](http://www.pdb.org) (PDB ID codes 6LJA and 6LJL). The sequences of BlexoHep (GenBank: EDV07780.1), BTexoHep (GenBank: AA079757.1), BCexoHep (GenBank: ALJ58962.1), PAexoHep (GenBank: WP\_026063245.1) and BFexoHep (GenBank: EEX44367.1) have already existed in NCBI database. The raw MS data of P8-1, P8-2, P8-3, P8-4, and P8-5 have been uploaded onto figshare and can be accessed by the link <https://doi.org/10.6084/m9.figshare.13413584>. Source data are provided with this paper.

Received: 5 August 2020; Accepted: 25 January 2021;  
Published online: 24 February 2021

## References

- Linhardt, R. J. 2003 Claude S. Hudson Award address in carbohydrate chemistry. Heparin: structure and activity. *J. Med. Chem.* **46**, 2551–2564 (2003).
- 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–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- $\beta$ 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-Martí, 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, 1992).
16. Ulaganathan, T. et al. Conformational flexibility of PL12 family heparinases: structure and substrate specificity 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 disaccharide product. *J. Biol. Chem.* **281**, 15525–15535 (2006).
22. Shaya, D. et al. Catalytic mechanism of heparinase II investigated by site-directed mutagenesis and the crystal structure with its substrate. *J. Biol. Chem.* **285**, 20051–20061 (2010).
23. Dong, W., Lu, W., McKeehan, 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. USA* **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. Hof, D. J., Versteeg, E. M. M., van de Lest, C. H. A., Daamen, W. F. & van Kuppevelt, T. H. A versatile salt-based method to immobilize glycosaminoglycans and create growth factor gradients. *Glycoconj. J.* **36**, 227–236 (2019).
28. 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).
29. Daud, A. N. et al. Synthetic heparin pentasaccharide depolymerization by heparinase I: molecular and biological implications. *Clin. Appl. Thromb. Hemost.* **7**, 58–64 (2001).
30. Holm, L. & Rosenström, P. Dali server: conservation mapping in 3D. *Nucleic Acids Res.* **38**, 545–549 (2010).
31. 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).
32. 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).
33. Turnbull, J. E., Hopwood, J. J. & Gallagher, J. T. A strategy for rapid sequencing of heparan sulfate and heparin saccharides. *Proc. Natl Acad. Sci. USA* **96**, 2698–2703 (1999).
34. 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).
35. Miller, R. L. et al. Shotgun ion mobility mass spectrometry sequencing of heparan sulfate saccharides. *Nat. Commun.* **11**, 1481 (2020).
36. Stringer, S. E., Kandola, B. S., Pye, D. A. & Gallagher, J. T. Heparin sequencing. *Glycobiology* **13**, 97–107 (2003).
37. 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).
38. 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).
39. 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).
40. 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).
41. Lu, D. et al. Biochemical characteristics and synergistic effect of two novel alginate lyases from *Photobacterium* sp. FC615. *Biotechnol. Biofuels.* **12**, 1–17 (2019).
42. Han, W. et al. A novel eliminase from a marine bacterium that degrades hyaluronan and chondroitin sulfate. *J. Biol. Chem.* **289**, 27886–27898 (2014).
43. Wang, Z. et al. Automatic crystal centering procedure at the SSRF macromolecular crystallography beamline. *J. Synchrotron Radiat.* **23**, 1323–1332 (2016).
44. Adams, P. D. et al. PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. D. Biol. Crystallogr.* **58**, 1948–1954 (2002).
45. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D. Biol. Crystallogr.* **60**, 2126–2132 (2004).
46. Li, L., Zhang, F., Zaia, J. & 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).
47. Wang, Z. et al. Sequencing the oligosaccharide pool in the low molecular weight heparin dalteparin with offline HPLC and ESI-MS/MS. *Carbohydr. Polym.* **183**, 81–90 (2018).

## Acknowledgements

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. We thank Xiaoju Li and Haiyan Sui from Life Science General Research Technology Platform of SKLMT (State Key Laboratory of Microbial Technology, Shandong University) for the assistance in XRD experiments. 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).

## 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.

## Competing interests

The authors declare no competing interests.

**Additional information**

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-021-21441-8>.

**Correspondence** and requests for materials should be addressed to Y.-Z.Z. or F.L.

**Peer review information** *Nature Communications* thanks Rebecca Miller and the other anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021