Discovery of exolytic heparinases and their catalytic mechanism and potential application

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

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.

\textsuperscript{1}National Glycoengineering Research Center and Shandong Key Laboratory of Carbohydrate Chemistry and Glycobiology, Shandong University, Qingdao, China. \textsuperscript{2}State Key Laboratory of Microbial Technology, Shandong University, Qingdao, China. \textsuperscript{3}College of Marine Life Sciences, and Frontiers Science Center for Deep Ocean Multispheres and Earth System, Ocean University of China, Qingdao, China. \textsuperscript{4}Laboratory for Marine Biology and Biotechnology, Pilot National Laboratory for Marine Science and Technology (Qingdao), Qingdao, China. \textsuperscript{5}Tiandong Pharma, Dongying, China. \textsuperscript{6}These authors contributed equally: Qingdong Zhang, Hai-Yan Cao. \textsuperscript{✉}email: zhangyz@sdu.edu.cn; fuchuanli@sdu.edu.cn
Heparin (HP)/heparan sulfate (HS), the polyanionic and heterogeneous polysaccharides belonging to glycosaminoglycan family1,2, are ubiquitously presented on cell surfaces and in the extracellular matrix (ECM) as well as the intracellular environment (in mast cells)3-5. The various structural characteristics of HP/HS permit these molecules to interact with various proteins, such as enzyme inhibitors3, chemokines5, growth factors5, morphogens8, and other signaling proteins, to participate in various physiological and pathological processes, such as coagulation2, cell adhesion2, 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 1916, 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-idurionate acid (IdoA) and N-acetyl-D-glucosamine (GlcNAc)12. During the biosynthesis 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 numerous 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/HS14,15. To date, based on their substrate specificity, the identified Hepases can be divided into three types: Hepase I (EC 4.2.2.7), 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 4.2.2.8), which digests both HP and HS16,17. Structurally, Hepase I folds into a β-jelly roll-type structure and prefers to cleave the α-1,4-linkages connected to IdoA2S/GlcA2S residues18,19. In contrast, Hepase III is composed of an N-terminal (α/α) barrel domain and a C-terminal antiparallel β-sandwich domain and prefers to cleave the α-1,4-linkages connected to GlcA/IdoA residues18,20. Hepase II adopts a topology similar to Hepase III and has no selectivity for certain GlcA/IdoA or IdoA2S/GlcA2S structures18,21,22. All Hepases share similar catalytic mechanisms in which His-Tyr acts as a Brønsted base and acid19,22,23. 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, we discover a type of Hepase 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 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 exoylases in the PL15 family.
to the cold living environment of *P. arcticus* [26]. 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 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). 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 are strongly inhibited by various heavy metal ions, such as Co²⁺, Cu²⁺, Ni²⁺, Pb²⁺, Zn²⁺, and Mn²⁺ (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 also be attributed to the presence of salt ions enhancing the binding of substrates to enzyme by removing the water coat from HP chains [27], 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-Fₐₕᵢᵢ as substrates. As shown in Table 1, the enzymatic activities of BlexoHep, BCexoHep, BTexoHep, PAexoHep, and BFexoHep toward HP-Fₐₕᵢᵢ 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 12.15 U/mg protein (pH 8.0, 50 mM NaH₂PO₄-Na₂HPO₄ containing 5 mM Ba²⁺ 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.

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

| Sequence character | Molecular mass (KDa) | Isoelectric point | Optimal condition | Effect of divalent metal cation (5 mM) | EDTA effect (5 mM) | EGTA effect (5 mM) | Enzyme activity (U/mg) HP-Fₐₕᵢᵢ | HP | αIII |
|--------------------|---------------------|------------------|------------------|--------------------------------------|-------------------|-------------------|-------------------------------|-----|-----|
| BlexoHep           | 98.9                | 6.54             | 30(Tris-HCl)     | Ca²⁺(142%) CaBa²⁺(150%) Mg²⁺(133%) | 19%               | 62%               | 0.75 ± 0.06                   | 49.89 ± 3.18 | 49.89 ± 3.18 |
| BCexoHep           | 98.5                | 6.91             | 30(Tris-HCl)     | Ca²⁺(135%) CaBa²⁺(192%) Mg²⁺(187%) | 99%               | 69%               | 6.18 ± 2.85                   | 61.86 ± 2.85 | 61.86 ± 2.85 |
| BTexoHep           | 100.3               | 7.37             | 80(Tris-HCl)     | Ca²⁺(133%) Mg²⁺(36%)               | 95%               | 59%               | 26.75 ± 1.77                  | 26.75 ± 1.77 | 26.75 ± 1.77 |
| PAexoHep           | 98.3                | 8.83             | 8.0(Tris-HCl)    | Mg²⁺(36%)                         | 95%               | 35%               | 24.25 ± 1.24                  | 24.25 ± 1.24 | 24.25 ± 1.24 |
| BFexoHep           | 100.2               | 9.76             | 8.0(Tris-HCl)    | Mg²⁺(36%)                         | 97%               | 84%               | 12.15 ± 1.48                  | 12.15 ± 1.48 | 12.15 ± 1.48 |

The residual activities were shown as the percentage of that (100%) of WT-BlexoHep.

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

### 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 exo-enzymes. 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 exo-lytic character of the PL15_2 Hepases, a saturated HP DP13 fraction without specific absorbance at 232 nm was prepared from the unsaturated Hepase PL15_2 enzymes, the HP DP13 was labeled with 2-AB and then treated with these enzymes, respectively. As the results shown in
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 BIexoHep. To investigate the substrate specificity of the PL15_2 Hepases, the capacity of the typical enzyme BIexoHep 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 (ΔUA2S1-4GlcNS6S1-4GlcA1-4GlcNS6S), and P4-8 (ΔUA2S1-4GlcNS6S1-4IdoA2S1-4GlcNS6S) were prepared and structurally determined (Supplementary Figs. 6–8 and Supplementary Table 1) and individually used as substrates for digestion assays by BIexoHep. When these tetrasaccharides (5 pmol) were treated with excess BIexoHep overnight, all could be almost completely degraded (Fig. 2a–e), indicating that BIexoHep 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 BIexoHep 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 BIexoHep 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 BIexoHep 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 BIexoHep 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.
As we known, 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 \(\text{UA2S(1-4)GlcNS6S(OCH}_3\text{)}\) without the trisulfated HP disaccharide \(\Delta\text{UA1(1-4)GlcNS3S6S}\), which is same as the case of fondaparinux treated with Hepases II\(^{29}\) (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 a 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 \((\text{UA2S1-4GlcNS6S})\) or tetrasaccharide \((\text{UA1-4GlcNS6S1-4idoA2S1-4GlcNS6S})\), and two 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 β-sheet domain (Ala\(^{24-Asn150}\)) composed of a large loop and a five-stranded β-sheet (β1-β5), a central (α/α)\(^{5}\) barrel domain (Pro\(^{151}\)…Glu\(^{317}\)) containing 20 α-helices (α1–α20) to form an (α/α)\(^{5}\) incomplete toroid structure, and a C-terminal β-sandwich domain (Leu\(^{316}\)…Pro\(^{466}\)) containing 23 β-strands and 5 short α-helices that fold into a three-layered β-sheets framework structure (α21–α25 and β6–β28) (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 BlexoHep was performed using the DALI server\(^{30}\). The results showed that the closest structure to BlexoHep is the exo-type alginate lyase Atu3025 (PDB code 3AFL)\(^{31}\) with a Z-score of 31.7 and an r.m.s. distance of 3.1 Å on 766 aligned C\(_\alpha\) 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\(^{32}\) 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\(^{30}\). However, BlexoHep has an additional small β-sheet domain (Ala\(^{24-Asn150}\)) 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 β-sheet, central (α/α)\(^{5}\) barrel and C-terminal β-sandwich (Supplementary Fig. 11d...
and Fig. 4a). The unique catalytic tunnel consists of two almost mutually perpendicular sections: the positively charged section generated by residues from α6, α8, α10, α12, α14, and the loops connecting α9/α10, α18/α19, β8/β9, and β10/α21, 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 α7, α8, α9, and the big loop of the N-terminal small β-sheet domain, especially residues Phe62 to Pro80, and Fig. 4a). The unique catalytic tunnel consists of two almost mutually perpendicular sections: the positively charged section generated by residues from α6, α8, α10, α12, α14, and the loops connecting α9/α10, α18/α19, β8/β9, and β10/α21, which acts as

The unique catalytic tunnel consists of two almost mutually perpendicular sections: the positively charged section generated by residues from α6, α8, α10, α12, α14, and the loops connecting α9/α10, α18/α19, β8/β9, and β10/α21, 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 α7, α8, α9, 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 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 Asp70, Ser236, Arg230, Arg233, Arg332, Asn336, His337, Gln340, Tyr390, Ser554, His555, Arg579, Arg636, Lys575, and Arg237, participate in the binding of the disaccharide product (ΔUA2S1-4GlcNS6S) and the tetrasaccharide substrate (ΔUA1-4GlcNS6S-14Oda2S1-4GlcNS6S) (Fig. 4c, d). The reducing end sugar rings of the tetrasaccharide substrate at the +1 and +2 subsites are tightly huggled 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 His355 match well with three key residues His202, Tyr257 and His406 in the active center of Hepase II from *P. heinus* 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 these 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 Pro67, Asp70, Pro73, Asp281, Glu236, and Asp335 from the N-terminal small β-sheet and central (a/a), 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 BlexoHep (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 Ca2+–dependent enzyme.

**Action mode switch of BlexoHep.** To determine the key factors leading to the exolytic character of BlexoHep, the N-terminal small β-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-Fall 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 Ca2+ ions.** According to the initial electron density map, there are two Ca2+ ions Ca1 and Ca2 in the structure of BlexoHep as described above. The Ca1 is surrounded by the nitrogen and oxygen atoms of residues Asp340, Thr384, Asp386, Asn394 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 Gly440, Arg456, Leu557, Asn559 and a water molecule with distances of less than 3.5 Å (Fig. 3b). To investigate the roles of these two Ca2+ 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 Asp340 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 Ca2+–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-Fall 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 (UDP6NE-P8-4) and tetrasaccharide (UDP4NE-P8-4) fractions as described under “Methods”. Then, the disaccharide compositions of P8-4 (UDP8NE-P8-4), UDP6NE-P8-4 and UDP4NE-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. 16d, the complete digestion of UDP8NE-P8-4 with Hepases produced two unsaturated disaccharides ΔUA1-4GlcNS6S and ΔUA2S1-4GlcNS6S with a molar ratio of 1:3, and the molar ratios of these two disaccharides in its nonreducing end hexasaccharide UDP6NE-P8-4 and tetrasaccharide UDP4NE-P8-4 were 1:2 and 1:1, respectively (Supplementary Table 5). By comparing the molar ratio change of disaccharides in UDP8NE-P8-4 and UDP6NE-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 UDP6NE-P8-4 and UDP4NE-P8-4 too (Fig. 6 and Supplementary Table 5). To determine the type of disaccharide at B site in P8-4, the UDP4NE-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. 16d a single disaccharide peak corresponding to ΔUA2S1-
4GlcNS6S was detected, indicating that the disaccharide at B site was HexUA2S1-4GlcNS6S too. Finally, based on the disaccharide composition of UDP4NE-P8-4, the nonreducing end disaccharide (A) of P8-4 must be ΔUA1-4GlcNS6S. Taken together, the preliminary sequence of P8-4 can be deduced as ΔUA1-4GlcNS6S1-4HexUA2S1-4GlcNS6S1-4HexUA2S1-4GlcNS6S (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 UDP4 of P8-3 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 UDP4 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 lyases31, which is different from Hepases I, II, and III, belong to the PL13, PL21, and PL12 families20, respectively. In this family, five exolytic Hepases, including BlexoHep, BCexoHep, BTexoHep, PAexoHep, and BFexoHep, 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 domains13 and linkage regions (GlcA-Gal-Gal-Xyl)32 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-F14II. 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 His337, Tyr300, and His555 of BlexoHep corresponding to the three key residues His202, Tyr257, and His406 of Hepase II 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 acid19,22,23. However, unlike in the case of the His337 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 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 III19–21, 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-548219. 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 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 Atu302531. 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 decades33–35. 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\(^3,3^8\). 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 weapons 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\(^3^3\) and Ion Mobility Mass Spectrometry\(^3^5\).

In conclusion, the identification of an exolytic Hepase family candidate and the enzymatic characterization of the enzyme 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.

Method

Materials

The strain Bacteroides intestinalis DSM 17393 was purchased from German Collection of Microorganisms and Cell Cultures (Braunschweig, Germany). The genes encode PL15.1 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 Idurom (Manchester, United Kingdom). 2-Aminobenzamide (2-AB), cyanoborohydride (NaBH\(_3\)CN), HP and HS polysaccharides from porcine esophagus, was established and successfully determined the preliminary sequences of HP/HS chains by combining with analytical techniques, such as nitrous acid degradation\(^3^3\) and Ion Mobility Mass Spectrometry\(^3^5\).

Sequences analysis of the gene and protein of BlexoHp

The gene and protein sequence of BlexoHp were downloaded from the NCBI database (GenBank: EDV077801.1). Then, the analysis of the similarity between BlexoHp and the known Hepas sequences was performed using the BLASTp algorithm online. Secretion signal peptides and their types were analyzed using the SignalP 4.1 server (Center for Biological Sequence Analysis, Genomics, and Informatics, Denmark) and the Secretion signal peptide 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 BlexoHp and Its derivative with selenomethionine.

To express BlexoHp in E. coli strains, the gene blexoHp was constructed onto plasmid PET-30a (+) (Novagen). Then, the plasmid PET-30a-blexoHp was transformed into E. coli BL21 (DE3) cells. E. coli cells harboring an expression vector (pET30a-blexoHp) were initially cultured in LB broth and then induced to start the expression of wild-type BlexoHp (WT-BlexoHp) proteins with 5 mM IPTG (isopropyl-\(\alpha\)-\(D\)-galactopyranosid) when the cell density reached an \(A\_600\) of 0.6–1.0. An overexpression system for the BIexoHep derivative with selenomethionine (SerMet- BlexoHp) was also constructed in E. coli BL21 (DE3) cells as previously reported^41. Cells grown overnight in LB medium were harvested by centrifugation at 3700 × \(g\) 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\(^{-1}\) cm\(^{-1}\) for products [1 \(\mu\)M = 1 μmol of (Δ\(A\))/absorbent 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 μl) of the reaction mixture were added to 1 mg/ml HP, HS, or HP-FIII degrading activities of each enzyme were investigated at the optimum \(\pH\) and temperature above. The effects of salt concentrations on the activities of each protein against HP-FIII were investigated at the optimum \(\pH\) and temperature above. The effects of salt concentrations on the activities of each protein against HP-FIII were investigated at the optimum \(\pH\) and temperature above. The optimal conditions for each PL15.2 protein were determined as described previously^41. Briefly, to determine the optimal \(\pH\) for each protein, HP-FIII (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\(_2\)PO\(_4\)-NaPO\(_4\) buffer (pH 6.0–8.0), and 50 mM Tris-\(\text{HOCl}\) 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-FIII 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-FIII 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-FIII as substrates. Briefly, each enzyme (1 μg) was added to 1 mg/ml HP, HS or HP-FIII in its optimal buffer containing 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\(^{-1}\) cm\(^{-1}\) for products [1 \(\mu\)M = 1 μmol of (Δ\(A\))/absorbent product formed per min].

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 BlexoHp above. These genes, including Hepase II (BTeXoHp) from strain Bacteroides thetaiotaomicron VPI-5482, heparan-sulfate lyase (BCexoHp) from Bacteroides cellulosolvens W12, Hepase III (BlexoHp) (GenBank: EEX44367.1) from strain Bacteroides fragilis DSM 17565, the hypothetical protein (CIM0152-3474) (BTeXoHp) (GenBank: GAE85076.1) from strain Bacteroides reticularitermitis JCM 10512, the DUF9462 domain-containing protein (PAexoHp) (NCBI accession number: WP_026063245.1) from Pedobacter arcticus, the DUF9462 domain-containing protein (BnexoHp) (GenBank: WP_000982393.1) from Bacteroides xylanisolvens and the DUF9462 domain-containing protein (PHexoHp) (GenBank: WP_015808643.1) from Pedobacter heparinus, share a sequence similarity of 66.26%, 60.57%, 49.27%, 47.79%, and 31.13% with BlexoHp, 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 BlexoHp above.

Biochemical characterization of the recombinant PL15.2 family of proteins.

The optimal conditions for each PL15.2 protein were determined as described previously^41. Briefly, to determine the optimum \(\pH\) for each protein, HP-FIII (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\(_2\)PO\(_4\)-NaPO\(_4\) buffer (pH 6.0–8.0), and 50 mM Tris-\(\text{HOCl}\) 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-FIII 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-FIII were investigated at the optimum \(\pH\) and temperature of each corresponding protein.
nm. After desalting by gel filtration, saturated HP DP13 (30 μg) was digested with various PL15_2 family proteins (50 μL) and traced under their optimal conditions. Aliquot samples 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 μg of the HP DP13 was 2-AB labeled in advance and then treated with 5 μL 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 ΔUA1-GlcNAcS61-GlcA1-4GlcNS6S, ΔUA1-GlcNAcS61-4GlcA1-4GlcNS6S, ΔUA1-GlcNS6S1-4GlcA1-4GlcNS6S) were individually digested by Hepases I and II, 2-AB labeled and then analyzed by anion-exchange HPLC on a Polyamide II column eluted using a NaH2PO4 gradient (0-1 M) and monitored using a fluorescence detector with excitation and emission wavelengths of 330 and 420 nm, 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 Hexas I-II was digested with 0.9 U BlexoHep for 0 min, 1 min, 5 min, 30 min and 12 h and then treated with 5 μL of each PL15_2 enzymes. Resulting oligosaccharides in each reaction were purified and pooled through SEC on a Superdex Peptide 10/300 GL column eluted with 0.20 M NH4HCO3, and freeze-dried repeatedly to remove NH4HCO3. Eluted with 0.20 M NH4HCO3, and freeze-dried repeatedly to remove NH4HCO3. Each oligosaccharide sample was dissolved in 0.5 mL of 99.9% D2O and lyophilized finally dissolved in 0.5 mL D2O in a 5 mm NMR tube for the 1H NMR spectroscopy. 1H 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), GlcNS6S1-4GlcA1-4GlcNS6S, was used as substrate to test the preference of BlexoHep and other PL15_2 enzymes. Brieﬂy, 5 μg Arixtra was treated with each enzyme and then the product was analyzed by anion-exchange HPLC on a Polyamide II column eluted using a linear gradient of NaH2PO4 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 identically 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-F11 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 in advance and then treated with 5 mU of each PL15_2 enzymes for 4 h. 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 octasaccharide was partially digested by 50, 50, 50, 9, and 1.5 mM BlexoHep proteins for 5 min, respectively. The UDPG and UDP products were collected and freeze-dried repeatedly to remove NH4HCO3 to get the purified oligosaccharides. Secondly, the UDPG, UDP and UDPG of the octasaccharides were individually digested by Hexas I and II, 2-AB labeled and then analyzed by anion-exchange HPLC on a Polyamine II column eluted using a NaH2PO4 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 UDPG from the octasaccharides were individually treated with O2, then digested by Hexas 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 HiILIC-ESI-MS on a LTQ-Orbitrap XL FT MS (Thermo Fisher Scientific, San Jose, CA) as previously reported. LC separation of octasaccharides was performed on a Thermo Ultimate 3000 system (Thermo Fisher Scientific, San Jose, CA) using a Luna HiILIC column (2.0 mm × 150 mm, 200 Å, Phenomenex, Torrance, CA). Mobile phase a was 5 mM ammonium acetate, while mobile phase b was 5 mM ammonium acetate in 98% acetonitrile. Mobile phase b was increased from 5% 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 −42 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. 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) 2.0, 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 ± SD. For comparison of the statistical differences between two groups, Student’s t-test (two-sided) was carried out for statistical analysis.

### 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 6J1A and 6L1J). The sequences of BlexoHep (GenBank: EDV07780.1), BTexoHep (GenBank: AA079757.1), BcexoHep (GenBank: AL893682.1), PaexoHep (GenBank: WP_026063245.1) and BfexoHep (GenBank: EEX44367.1) have already existed in NCBI database. The raw MS data of Ph8-3, Ph8-4, Ph8-5 have been uploaded onto figshare and can be accessed by the link [https://doi.org/10.6084/m9.figshare.13413584](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
1. Linhardt, R. J. 2003 Claude S. Hudson Award address in carbohydrate chemistry. Heparin: structure and activity. *J. Med. Chem.* 46, 2531–2564 (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).
