Glycobiology, 2017, vol. 27, no. 2, 176–187
doi: 10.1093/glycob/cww096
Advance Access Publication Date: 12 October 2016
Original Article

OXFORD

Structural Biology

Conformational flexibility of PL12 family heparinases: structure and substrate specificity of heparinase III from *Bacteroides thetaiotaomicron* (BT4657)

ThirumalaiSelvi Ulaganathan[2], Rong Shi[3], Deqiang Yao[4], Ruo-Xu Gu[5], Marie-Line Garron[6], Maia Cherney[2], D Peter Tieleman[5], Eric Sterner[7], Guoyun Li[7], Lingyun Li[7], Robert J Linhardt[7], and Miroslaw Cygler[2,1]

[2]Department of Biochemistry, University of Saskatchewan, Saskatoon, S7N 5E5 Saskatchewan, Canada, [3]Département de Biochimie, de Microbiologie et de Bio-informatique, PROTEO, and Institut de Biologie Intégrative et des Systèmes (IBIS), Université Laval, Pavillon Charles-Eugène-Marchand, Québec City, QC G1V 0A6, Canada, [4]National Center for Protein Science, Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, China and Shanghai Science Research Center, Chinese Academy of Sciences, Shanghai 201204, China, [5]Department of Biological Sciences and Centre for Molecular Simulation, University of Calgary, 2500 University Dr NW, Calgary, AB T2N 1N4, Canada H4P 2R2, Canada, [6]the Architecture et Fonction des Macromolécules Biologiques, UMR7257 CNRS, Aix-Marseille University, F-13288 Marseille, France, the INRA, USC1408 Architecture et Fonction des Macromolécules Biologiques, F-13288 Marseille, France, and [7]Department of Chemistry and Chemical Biology, Center for Biotechnology and Interdisciplinary Studies, Rensselaer Polytechnic Institute, Troy, NY 12180, USA

[1]To whom correspondence should be addressed: Tel: +1 (306) 966-4361, Fax: +1 (306) 966-4390;
e-mail: miroslaw.cygler@usask.ca

Received 1 July 2016; Revised 5 September 2016; Accepted 6 September 2016

Abstract

Glycosaminoglycans (GAGs) are linear polysaccharides comprised of disaccharide repeat units, a hexuronic acid, glucuronic acid or iduronic acid, linked to a hexosamine, *N*-acetylglucosamine (GlcNAc) or *N*-acetylgalactosamine. GAGs undergo further modification such as epimerization and sulfation. These polysaccharides are abundant in the extracellular matrix and connective tissues. GAGs function in stabilization of the fibrillar extracellular matrix, control of hydration, regulation of tissue, organism development by controlling cell cycle, cell behavior and differentiation. Niche adapted bacteria express enzymes called polysaccharide lyases (PL), which degrade GAGs for their nutrient content. PL have been classified into 24 sequence-related families. Comparison of 3D structures of the prototypic members of these families allowed identification of distant evolutionary relationships between lyases that were unrecognized at the sequence level, and identified occurrences of convergent evolution. We have characterized structurally and enzymatically heparinase III from *Bacteroides thetaiotaomicron* (BtHepIII; gene BT4657), which is classified within the PL12 family. BtHepIII is a 72.5 kDa protein. We present the X-ray structures of two crystal forms of BtHepIII at resolution 1.8 and 2.4 Å. BtHepIII contains two domains, the N-terminal α-helical domain forming a toroid and the C-terminal β-sheet domain. Comparison with recently determined structures of two other heparinases from the same PL12 family allowed us to identify structural flexibility in the arrangement of the domains indicating open–close movement. Based on

comparison with other GAG lyases, we identified Tyr301 as the main catalytic residue and confirmed this by site-directed mutagenesis. We have characterized substrate preference of BtHepIII toward sulfate-poor heparan sulfate substrate.

Key words: active site, crystal structure, heparin lyase, protein dynamics, substrate specificity

Introduction

Glycosaminoglycans (GAGs) are polysaccharides present abundantly both in vertebrates and in many invertebrates, located primarily on the cell surface or in the extracellular matrix (Bülow and Hobert 2006). They are linear polymers built from disaccharide repeating units of hexuronic acid and hexosamine sugar. Based on the chemical structure and post polymerization modification, GAGs are classified into chondroitin sulfate, dermatan sulfate, heparan sulfate, heparin, keratan sulfate and hyaluronan. GAGs in animals undergo relatively rapid turnover and are degraded by GAG hydrolases (Owens and Wagner 1991; Yung et al. 2004). Bacteria, especially from soil, have enzymes that break down GAGs by a lytic elimination mechanism, to serve as carbon source (Ernst et al. 1995). These bacterial enzymes are classified within the polysaccharide lyases (PL) families. While GAG hydrolases present in eukaryotes and prokaryotes break the glycosidic bond at the reducing end of the glucosamine, GAG lyases, found only in prokaryotes (Linhardt et al. 1986), break the glycosidic linkage at the nonreducing end of the uronic acid with a concomitant formation of the C4–C5 double bond (Gacesa 1987). PL have been classified into 24 families, with over 300 unclassified sequences (Cantarel et al. 2009). Of these, seven families, PL6, 8, 12, 13, 16, 21 and 23, contain enzymes degrading GAGs (Garron and Cygler 2010, 2014).

The breaking down of GAGs by PL has been extensively studied both biochemically and structurally. Four folds were found among their structures, with the most common being the $(\alpha/\alpha)_{5/6}$ barrel + antiparallel β-sheet found in PL8, 12, 21 and 23 (viruses) families. The other folds observed among GAG lyases are parallel β-helix, β-jelly roll and triple strand β-helix (PL16) (Garron and Cygler 2014).

Heparin and heparan sulfate are highly negatively charged and consist of repeating units of 1→4-linked uronic acid (L-iduronic, IdoA or D-glucuronic acid, GlcA) and glucosamine; heparin has a high proportion of IdoA (~90%) and is highly sulfated (Iozzo 1998). Heparin interacts with a variety of proteins such as growth factors and chemokines, and participates in various physiological and pathological processes (reviewed by Capila et al. 2002; Imberty et al. 2007). The lytic depolymerization of heparin/heparan sulfate is carried out by three families of enzymes with distinct specificities and no recognizable sequence similarities, namely heparinases I, II and III (Lohse and Linhardt 1992; Godavarti and Sasisekharan 1996). Heparin lyase I (heparinase I, HepI, family PL13) is specific for heparin, cleaving next to IdoA, heparin lyase III (heparinase III, HepIII, family PL12) cleaves heparan sulfate next to GlcA while heparin lyase II (heparinase II, HepII, family PL21) can depolymerize both these substrates (Desai et al. 1993a, 1993b). The representative structures for all three heparin lyase families are known (Shaya et al. 2006; Han et al. 2009; Dong et al. 2012). HepI (PDB code 3IKW, Han et al. 2009) folds into a β-jelly roll, its structure with oligosaccharide substrate has been determined and the catalytic mechanism established as utilizing a tyrosine (Tyr357) and histidine (His151) as a general acid and base, respectively. HepII (PDB code 2FUQ (Shaya et al. 2006)) and HepIII (PDB code 4FNV (Dong et al. 2012)) display the $(\alpha/\alpha)_5$ barrel + antiparallel β-sheet fold and also utilize tyrosine as the general base.

Two structures of heparinases III have been described in the literature, that of BT4662 from *Bacteroides thetaiotaomicron* (PDB code 4FNV) (Dong et al. 2012) and an HepC from *Pedobacter heparinus* (PDB code 4MMH, 4MMI) (Hashimoto et al. 2014). Unlike HepI and HepII, neither of them could be crystallized in complex with an oligosaccharide substrate and the catalytically important residues were deduced by comparison with HepII-substrate complex. These two HepIII proteins share 32% sequence identity and their individual N- and C-terminal domains align closely. However, these domains differ significantly in their relative orientation in these two proteins. In a continuing effort to capture the enzyme–substrate complex and to address the question of the conformational flexibility, we have undertaken the investigation of a third enzyme from the PL12 family, a second *B. thetaiotaomicron* enzyme specific for heparan sulfate, coded by the gene BT4657 (refer to henceforth as BtHepIII). The crystal structure showed that this enzyme displays yet another relative orientation of the N- and C-terminal domains, with a wide open substrate binding site. Despite extensive efforts, we were unable to capture substrate in the binding site of BtHepIII, which suggested that this as well as the other HepIII enzymes were captured in inactive conformations. The conformational differences displayed by these enzymes prompted us to investigate the role of conformational flexibility in PL12 enzymes by applying normal mode analysis. We discovered that the flexible segment is not at the connection between the α-helical and β-sheet domains but rather within the α-helical domain following the third α-helical hairpin. The active site residues are located on the same C-terminal rigid segment and are unaffected by this large-scale motion but the residue neutralizing the acidic group of GlcA is located on N-terminal segment and moves relative to the active site and assumes a proper place concomitantly with the substrate binding.

Results and Discussion

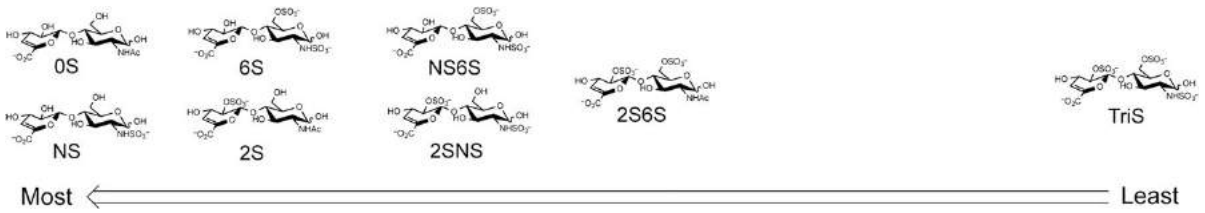
BtHepIII substrate specificity

The natural substrate for heparinase III isolated from *B. thetaiotaomicron* is likely the GAG heparan sulfate. But by probing the specificity of this enzyme using other natural and chemically modified heparin/heparan sulfate-type polysaccharides, information can be gleaned of the structural requirements for this enzyme and its activity can be compared to heparinases from other organisms.

The action of BtHepIII on various polysaccharide substrates was examined by disaccharide analysis using liquid chromatography–mass spectrometry (LC–MS) (Table I). With the exception of heparosan (the homo-copolymer →4) β-D-GlcA (1→4) α-D-GlcNAc (1→), these polysaccharide substrates are structurally heterogeneous, being consisted of multiple types of disaccharide repeating units. The relative amounts

Table I. Degradation of various GAGs to disaccharide units by BtHepIII as compared to a total degradation by complement of heparinases I, II and III. Preference for sulfation level and position extracted from this table is shown in Schema 1

Substrate	Enzymatic treatment	Disaccharide products formed (relative percentage)						Absolute quantity of disaccharide products (%)
		0 S (%)	NS (%)	6 S/2 S (%)	NS6S/NS2S (%)	2S6S (%)	*TriS* (%)	
Heparan sulfate	Heparinase III	47.4	33.2	15.2	3.94	0.23	0.03	89.5
	Heparinase I, II, III	47.5	27.3	15.5	3.27	0.79	5.60	100
Heparin	Heparinase III	57.3	14.3	20.1	2.58	5.73	0.03	29.6
	Heparinase I, II, III	9.26	6.04	11.0	7.35	3.17	63.1	100
Heparosan	Heparinase III	99.9	<0.01	<0.01	<0.01	0.10	<0.01	98.2
	Heparinase I, II, III	99.8	0.01	<0.01	<0.01	0.23	<0.01	100
N-sulfoheparosan	Heparinase III	19.0	80.3	0.19	0.24	0.24	<0.01	99.8
	Heparinase I, II, III	17.5	80.3	0.56	1.22	0.49	<0.01	100
Desulfated N-sulfated heparin	Heparinase III	60.3	14.5	19.30	4.83	1.09	<0.01	92.9
	Heparinase I, II, III	58.1	17.4	17.6	3.17	3.17	0.60	100
N-desulfated, N-acetylated heparin	Heparinase III	33.0	0.01	65.93	0.00	1.10	<0.01	31.4
	Heparinase I, II, III	7.86	0.01	33.7	0.02	58.4	0.01	100



Schema 1. Efficiency of release from polysaccharide on treatment with BtHepIII.

of specific disaccharides within each polysaccharide were determined by exhaustive treatment with heparinase I, II and III, followed by LC–MS (Table I). The disaccharide structures, shown in Schema 1, are ordered in their preference for BtHepIII release from polysaccharide substrate. Unsulfated disaccharides or monosulfated disaccharides with an *N*-sulfo are equally likely to be released, next come the disaccharides with a single *O*-sulfo group, followed by the disulfated disaccharides and finally the trisulfated disaccharide, which is highly resistant to BtHepIII release. The specificity of BtHepIII with regard to uronic acid epimers is somewhat more difficult to discern because BtHepIII is a lyase and the chirality at carbon-5 of the uronic acid at the cleavage site is lost, making it impossible to discern whether the uronic acid at the linkage that was cut was GlcA or IdoA. There are two ways to deduce the specificity of BtHepIII of uronic acid epimers. First, it is well known that polysaccharides with higher sulfation levels have higher IdoA:GlcA ratios (i.e., heparin with ~2.7 sulfates/disaccharide has an 8:2 ratio, while heparan sulfate with ~1 sulfate/disaccharide has a 1:9 ratio of IdoA:GlcA). Thus, it was unclear whether the presence of IdoA or the presence of high sulfation levels is more important in the resistance to BtHepIII cleavage. Most noteworthy is that 93% of the linkages to uronic acid in chemically modified heparin (desulfated N-sulfated heparin), containing an 8:2 ratio of IdoA:GlcA but with a low level of sulfation, displayed a similar BtHepIII sensitivity to a chemically modified heparosan (*N*-sulfoheparosan), containing only GlcA and a low level of sulfation. These results suggest that BtHepIII acts preferentially on polysaccharide substrates having low O-sulfation including substrates linkages containing an IdoA residue. The specificity of *B. thetaiotaomicron* BtHepIII acting on polysaccharide substrates closely resembles that of *Flavobacterium heparinum* (*P. heparinus*) reported (Desai et al. 1993a, 1993b).

Determination of apparent kinetic parameters of BtHepIII

The determination of the apparent kinetic parameters of BtHepIII followed a well-established heparinase characterization protocol that has been previously published (Lohse and Linhardt 1992). Importantly, though the BtHepIII enzyme used in this study was of high purity and homogeneity, the heparan sulfate substrate is polydisperse and highly heterogeneous. Heparan sulfate contains approximately 20 unique disaccharides that vary based on structural and chemical composition. Resultantly, catalytic efficiency of the heparinase III enzyme at each unique site may vary. In this regard, heparan sulfate is not a typical substrate, but remains the commercial and academic standard substrate for measuring heparin lyase activity.

The kinetic profile of the BtHepIII enzyme was well fit to both a general Michaelis–Menten nonlinear regression profile (Fig. 1A) and an Eadie–Hofstee linearization plot (Fig. 1B). The calculated K_M of the BtHepIII enzyme for heparan sulfate was 593 ± 64 μM. The calculated apparent V_{MAX} was 1.9 ± 0.2 (μmol min^{-1} $mg_{BtHepIII}^{-1}$). This compares to a K_M of 29.4 ± 3.16 μM and V_{MAX} of 141 ± 3.88 (μmol min^{-1} $mg_{BtHepIII}^{-1}$) reported for HepC from *F. heparinum* (*P. heparinus*) (Lohse and Linhardt 1992) and a K_M of 6.8 ± 0.33 μM and V_{MAX} of 126 (μmol min^{-1} $mg_{BtHepIII}^{-1}$) reported for BT4662 from *B. thetaiotaomicron* (Dong et al. 2012). These results suggest that BtHepIII from *B. thetaiotaomicron* has a lower catalytic efficiency than HepC or BT4662.

Overall structure of BtHepIII

BtHepIII is an α + β protein that can be divided into two domains (Fig. 2A–C). The N-terminal domain of ~370 residues is

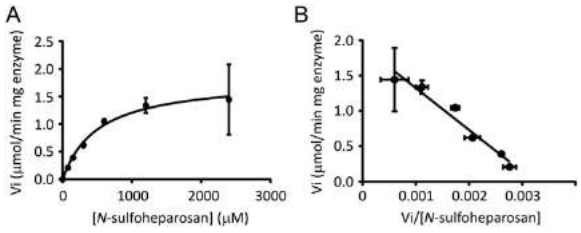


Fig. 1. Kinetic studies on HepIII acting on heparan sulfate. (**A**) Michaelis–Menten plot for heparan sulfate between the concentrations of 0 and 300 μM. (**B**) Eadie–Hofstee plot of the data presented in panel A. All experiments were done in triplicate.

predominantly α-helical, while the C-terminal domain is composed of β-sheets. The N-terminal domain contains 16 α-helices and is assembled around five helical hairpins hp1 (α6–α7), hp2 (α9–α10), hp3 (α11–α12), hp4 (α13–α14) and hp5 (α15–α16), arranged into a partial toroid form designated $(\alpha/\alpha)_5$. The first and the last hairpins are separated by ~20 Å leaving ample space for the substrate to enter. The inner surface of this partial toroid is lined with helices α6, α9, α11, α13 and α15, while helices α7, α10, α12, α14 and α16 form an outer wall of the toroid. The loops within the hairpins are rather short, while the loops connecting the hairpins are significantly longer. Particularly long insertions occur between hairpins hp1 and hp2 and between hairpins hp4 and hp5. The latter insertion forms a β-strand hairpin that stacks against the C-terminal domain. The N-terminal segment of ~110 residues folds along the side of the hairpin hp1 and on the top of the intra-hairpin loops of hp1–3; this segment starts with three α-helices (α1–α3) followed by an 18 residue long linker connecting to an α-helix (α4), β-hairpin and α-helix (α5) (Fig. 2B). The C-terminal domain is folded into three twisted, anti-parallel β-sheets; the first two are rather extensive with eight antiparallel β-strands each and two crossovers between the sheets (Fig. 2C). The fifth strand of the first sheet has a 28 amino acid long hairpin insertion that extends perpendicularly to the sheet. The third sheet contains only three β-strands. The N-terminal domain comprises residues 21–387 and the C-terminal domain residues 397–666. A nine-residues linker connects the two domains.

One metal ion was observed in the crystal structure of BtHepIII located between β-hairpin at the end of the N-terminal domain and the first β-sheet of the C-terminal domain. The metal ion forms octahedral coordination with Gln433, Asp451, His476 and three water molecules. All three waters are hydrogen bonded to the backbone CO or NH of the neighboring residues and in particular to the β-hairpin from the N-terminal domain. The octahedral coordination together with the ion-liganding atom distances varying between 2.16 and 2.29 Å in the higher resolution form 1 structure suggest Mg^{2+}, Mn^{2+} or Zn^{2+} as the likely ion (Zheng et al. 2008). Refinement with each of these in turn showed that only for Mg^{2+}, there is no negative density at the metal site and the B-factor is very similar to the liganding atoms.

The methylation of surface-exposed lysine residues was essential for the crystallization of HepIII. Inspection of the HepIII electron density indicated that many lysine side chains are partially disordered and the identification of methylated lysines from electron density map is inconclusive. Only very few lysines are involved in the crystal contacts and they differ between form 1 and form 2 crystals. We rationalize that rather than specific contacts formed by few methylated lysine side chains, the effect of their methylation was a global change of the electrostatic potential on the surface of the protein, allowing for a closer approach of molecules and thus promoting crystallization.

Active site

Bacteroides thetaiotaomicron BtHepIII has a similar two-domain architecture as *P. heparinus* HepII although they are classified into different PL families. The C-terminal β-sheet domains superimpose very well, in particular the first two large β-sheets with r.m.s.d. of 1.58 Å for 171 Ca atoms (out of 247). The metal ion in BtHepIII is in exactly the same position as the Zn^{2+} ion in HepII; moreover, it is covered in a similar fashion by a β-hairpin in both proteins. The superposition of the N-terminal α-helical domains shows good overlap of the first two α-helical hairpins (r.m.s.d. of 1.49 Å for 76 Ca), with the remaining three hairpins showing larger divergence (twist perpendicular to the toroidal plane). One can also superimpose the fourth and fifth hairpins quite well with divergence of the first three hairpins. This suggests that the difference in arrangement of the five hairpins of the N-terminal domains is related to conformational changes around the third hairpin. Three residues of HepII were identified as the active site residues, His202, Tyr257 and His406 (Shaya et al. 2006). HepII is capable of breaking the C4–O1 bond next to either GlcA or IdoA. For GlcA, it is Tyr257 and His406 that perform the catalysis, while His202 is essential for cleavage near the iduronic acid (Shaya et al. 2010). The superposition of BtHepIII and HepII based on their C-terminal domains superimposes Tyr257 and His406 of HepII on Tyr301 and His431 of BtHepIII (Fig. 2D). Therefore, these two residues, Tyr301 and His431, are the best candidates for active site in BtHepIII. On the other hand, the HepII segment containing His202 does not align well with BtHepIII and there is no histidine side chain in close proximity. The key feature necessary for proton abstraction from the C4 atom of GlcA sugar is neutralization of the charge on the carboxylic group at C5. In this superposition of BtHepIII and HepII active sites the HepII Glu205, which in a protonated state helps to neutralize the acidic group of GlcA (Shaya et al. 2010), overlaps with the Asn247 of BtHepIII, suggesting that this asparagine is involved in charge neutralization. Further support for this role of Asn247 comes from comparison with chondroitinase AC and hyaluronan lyase. In these proteins, charge neutralization is accomplished by the side chain of asparagine that forms double hydrogen bond with the acidic group of GlcA (Li et al. 2000, Lunin et al. 2004). Inspection of the BtHepIII shows the Asn247 residue, located at the first helix of the third α-hairpin, extendings toward Tyr301 and the substrate binding site (Fig. 2E).

Site-directed mutagenesis of active site and putative binding site residues

Two single mutants, Y301F and H431A, were made and assayed to test the importance of the two putative active site residues for catalysis. Depolymerization of GAG polysaccharide by a lyase results in the formation of an unsaturated disaccharide product, which can be monitored by measuring the absorbance at 230 nm. The activity of wild-type BtHepIII and the mutants was followed using heparan sulfate as a substrate by observing the appearance of an unsaturated product. This assay easily detected activity of the BtHepIII and showed that after methylation of lysines the enzyme could still depolymerize heparan sulfate, albeit with lower activity. However, mutating the putative active site residues rendered the enzyme inactive (Fig. 3), supporting their key role in catalysis. Mutations of equivalent residues, Tyr314 in BT4662 and Tyr294 and His424 in HepC, also led to loss of activity (Dong et al. 2012, Hashimoto et al. 2014). We next made the N247A mutation to test if Asn247 is indeed involved in the catalysis. In accordance with our hypothesis,

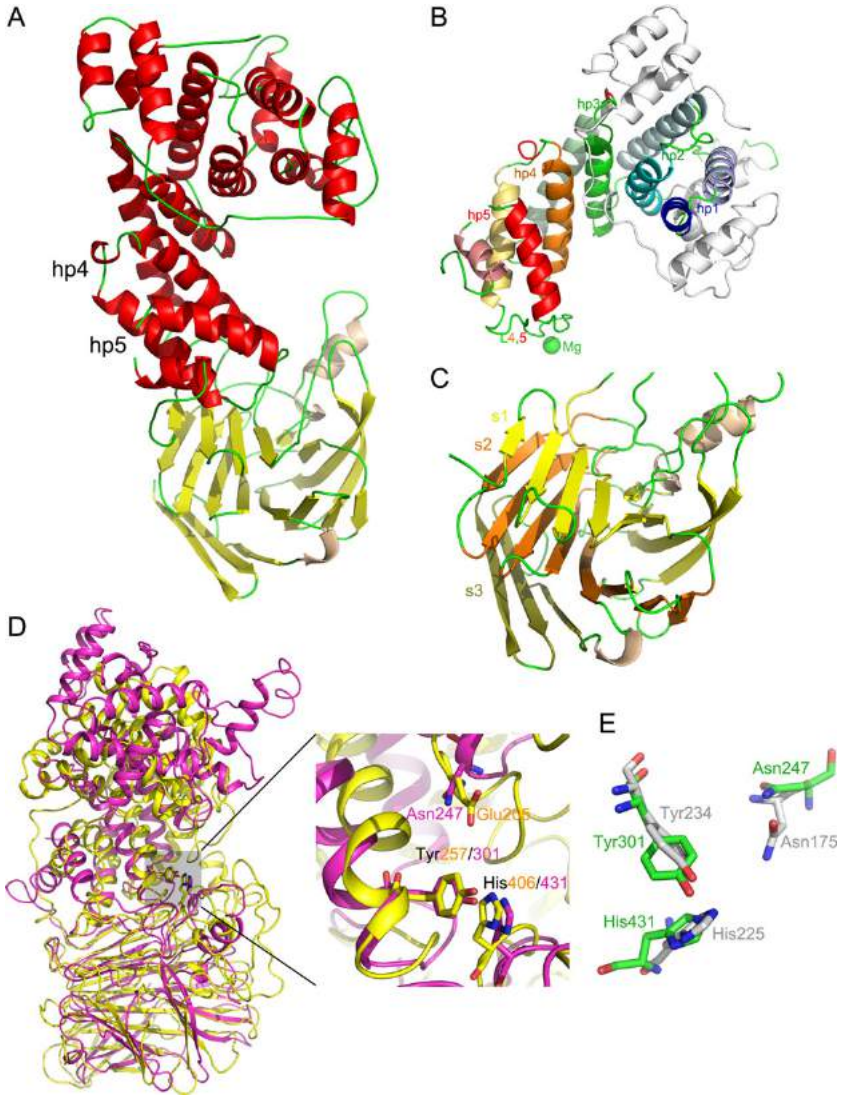


Fig. 2. (**A**) Cartoon representation of BtHepIII showing α helices (red) in the N-terminal domain and β sandwich (green) in the C-terminal domain; (**B**) the N-terminal α-helical domain forming an incomplete $(\alpha/\alpha)_5$ toroid fold. Helices that are not part of the toroid are colored white, the hairpins (hp) are marked in different colors, the inner helix is a darker shade and outer helix is a lighter shade. The loop connecting hp4 and hp5 (L4,5 green) and the metal ion (Mg, green) tying this loop to the β-sheet domain are marked; (**C**) the C-terminal β-sheet domain showing stacks of three β sheets colored yellow, orange and olive, respectively; (**D**) superposition of BtHep3 (magenta) and heparinase II (yellow, PDB code 2FUQ) based on the C-terminal domains only. The inset shows the close-up of the active site Tyr301, His431 and Asn247 are shown in stick representation. Tyr457, a potential catalytic residue, is also shown; (**E**) Superposition of the three active site residues of BtHepIII (Asn247, Tyr301 and His431 colored gray) and chondroitinase AC (Asn175, His225 and Tyr234 colored green, PDB code 1HMW). This figure is available in black and white in print and in color at *Glycobiology* online.

this mutant also showed no activity. The residue equivalent to Asn247 in BT4662 is Asn260 and in HepC it is Asn240 and their mutation to alanines also led to dramatically reduced activity (Dong et al. 2012, Hashimoto et al. 2014).

Inspection of the structure of BtHepIII shows the presence of a second tyrosine side chain, Tyr457, pointing toward the putative location of the GlcA residue and opposite to Tyr301. Mutation of this residue to a phenylalanine, Y457F, only moderately decreased the activity of the protein, suggesting that it is involved in substrate binding but not catalysis. This is corroborated by the lack of activity of Y301F mutant, in which Tyr457 is present but not able to rescue the activity.

Comparison with other PL12 enzymes

The structures of two other enzymes from family PL12 have been determined, namely *B. thetaiotaomicron* heparinase III (BT4662, PDB code 4FNV) and *P. heparinus* heparin sulfate lyase/heparinase III (HepC, PDB code 4MMH). Despite extensive trials, neither group was able to capture enzyme–substrate/product/analog complex. The comparison of these three structures as well as comparison with other lyases containing this fold shed some light on the reasons for the inability to capture the substrate in the binding site.

Comparison of BtHepIII, BT4662 and HepC shows that their N- and C-domains have very similar structures but their relative

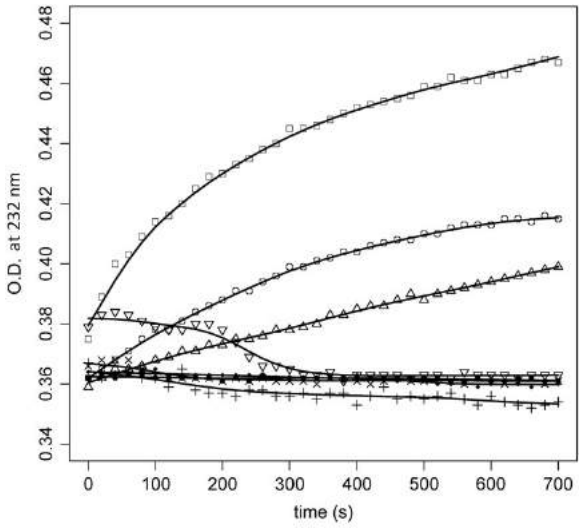


Fig. 3. Measurement of activity of the enzyme with heparan sulfate substrate. The reaction mixture contained 0.15 mg/mL heparan sulfate polymer, and 1 μg enzyme. The absorbance at 230 nm was measured as a function of time. The measurements were performed three times and each point is an average of three measurements. Symbols: □, native BtHepIII; ○, methylated BtHepIII; Δ, Y457F mutant; +, Y301F mutant; ×, N247A mutant; ▽,H431A mutant; ♦,heparan sulfate substrate; ●, buffer solution. The figure shows typical curves. The initial decrease of absorbance for H431A mutant is a likely artifact of the experiment.

Table II. Root-mean-squares deviations between the N- and C-domains of BtHepIII molecules from two crystal forms and two other heparinases from the PL12 family

Mol 1	No. of residues	Mol 2	No. of residues	R.m.s.d.	No. of equivalent Cα atoms
HepIII_form1	622	HepIII_form2	646	0.98	610
HepIII_form1_N	339	HepIII_form2_N	364	0.64	330
HepIII_form1_C	272	HepIII_form2_C	272	0.44	271
HepIII_form2_N	339	4FNV_N	375	1.04	210
HepIII_form2_C	272	4FNV_C	284	1.11	250
HepIII_form2_N	339	4MMH_N	350	1.19	290
HepIII_form2_C	272	4MMH_C	284	1.06	249

orientation differs between the proteins (Table II, Fig. 4A). Even between the two crystal forms of BtHepIII, a small but easily recognizable difference in orientation of N- vs. C-domain can be seen (Fig. 4B). Comparison of the root-mean-square deviations upon superposition of individual domains shows clearly that the C-terminal domains are structurally better conserved than the N-terminal domains. Moreover, the N-terminal domains superimpose well for the N-terminal helical cap and the first three α-hairpins, while they start to deviate for the fourth and fifth α-hairpins. Coincidently, this difference in the last two α-hairpins leads to a more closed or more open partial toroid. The BT4662 is the most closed form, while the HepC is the most open. The superposition of entire proteins based only on their C-terminal β-sheet domains provides a new insight into the origins of the flexibility (Fig. 4A). Indeed, this superposition shows that the end of the second helix in the third α-hairpin, the following two antiparallel β-strands and the fourth (with the exception of the intra-hairpin loop) and fifth α-hairpins are rigidly attached to the C-terminal β-sheet domain. This tying together of the end of the N-terminal domain and the C-terminal domain may be helped by the presence of a metal ion coordinated by the side chains from the β-hairpin in the N-terminal domain (after third α-hairpin) and the β-strands in the C-terminal domain. Of importance, the active site residues are within this rigid region, Tyr301 at the beginning of the first helix in the fourth α-hairpin of the N-terminal domain, and His431 in the loop between third and fourth β-strand of the first β-sheet of the C-terminal domain (Fig. 4C). The hinge region is within the N-terminal domain at the bottom part of second helix of the third α-hairpins, around residue Asn281 in BtHepIII.

Comparison with GAG lyases having $(\alpha/\alpha)_5 + \beta$ fold

Comparison of BtHepIII with HepII has been presented already above. The HepII (Y257A/H202A) mutant was crystallized with a tetrasaccharide substrate ΔUA-GlcNAc-GluA-GlcNAc (where ΔUA is 4-deoxy-αsL-*threo*-hex-4-enopyranosyluronic acid). As we discussed above, the superposition of HepII and BtHepIII based on their C-terminal domains overlaps the Tyr-His catalytic dyads. The location of the tetrasaccharide in this superposition provides a good approximation to the expected location of the substrate in BtHepIII (Fig. 4D). It is evident from this superposition that the contacts of the substrate with the observed 'open' conformer of BtHepIII are rather limited, and there are no contacts with the other end of the open toroid. The same lack of contact of the GAG chain with the upper end of the N-terminal domain occurs for the HepC that was also crystallized in an open conformation. The BT4662 has been captured in a closed-like conformation and here the modeled

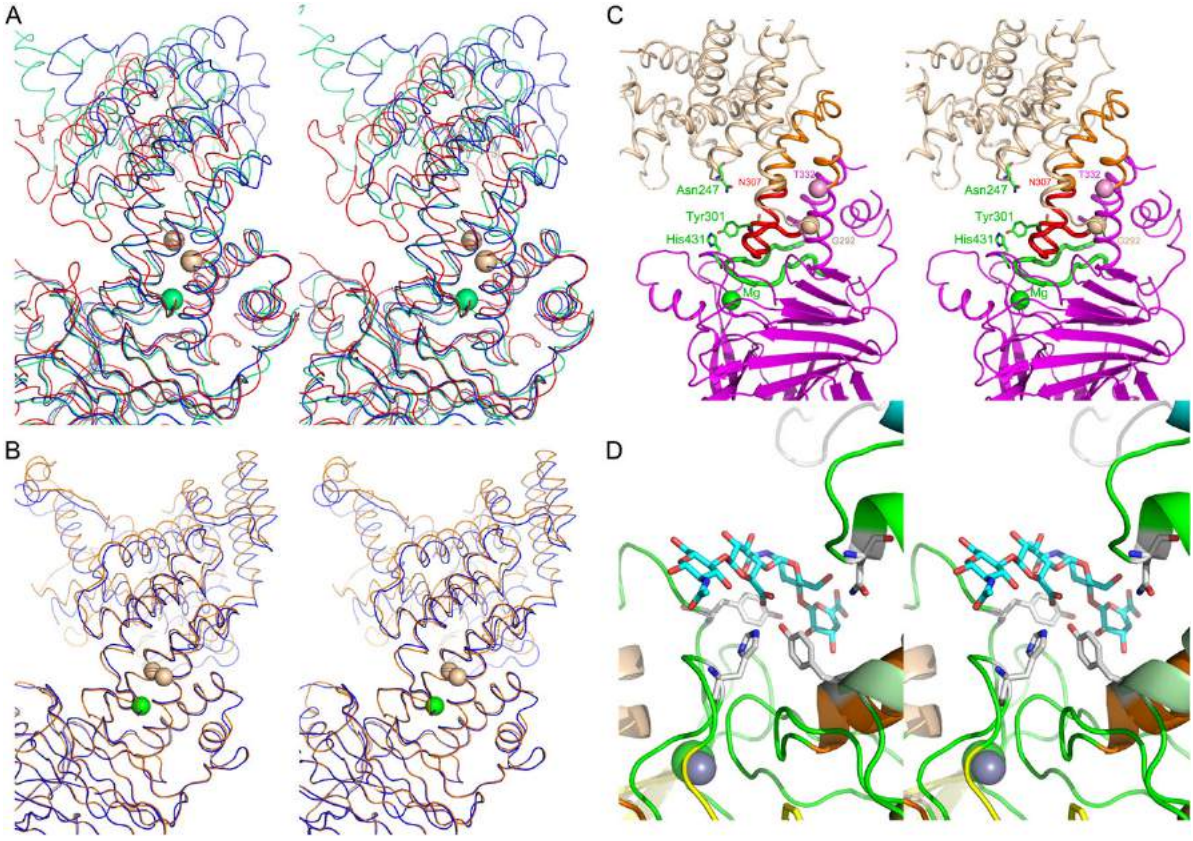


Fig. 4. Superposition of the structures of heparinase III enzymes based on their C-terminal domains. The N-terminal domains were found in different orientations in these enzymes. (**A**) Comparison of BtHepIII (green), BT4662 (red) and HepC (blue). In each case, the chain preceding the first helical hairpin was removed for clarity. The spheres indicate the position of the hinge regions. The main hinge is located within the loop between the hairpins h3 and h4 (green sphere). Additionally, the inner loop within hairpin h4 with the tips of the helices (between the wheat-colored spheres) follows the movement of the N-terminal part of the protein. (**B**) Two crystal forms of BtHepIII superimposed based on their C-terminal domains. A small difference in the orientation of the N-terminal domains is easily visible. The hinge points are the same as when comparing different heparinases III shown in (A). (**C**) Location of the active site residues (shown in stick mode, green) relative to the hinge points in BtHepIII. The hinge points are marked as spheres. The N-terminal segment 23–291 (colored wheat) and residues 308–331 (orange) move together, while the end of α-helix and the following loop (aa 292–307, red) containing Tyr301 move together with the β-sheet C-terminal domain (magenta). The loop between hp4 and hp5 (aa 347–361, red) is tied into the β-sheet domain by the metal ion (Mg). (**D**) Putative location of the tetrasaccharide substrate relative to the active site of BtHepIII. The tetrasaccharide was extracted from the structure of HepII-tetrasaccharide (PDB code 3E7J) upon superposition of the C-terminal domains. This figure is available in black and white in print and in color at *Glycobiology* online.

tetrasaccharide would make contacts with the inner α-helix of the first α-hairpin. However, the lack of success in crystallization of the complex of BT4662 with a saccharide (Dong et al. 2012) suggests that this conformation is also not compatible with substrate binding and catalysis.

There are other GAG lyases with a similar two-domain fold, namely hyaluronan lyase, chondroitinase AC and ABC and heparinase II. Their active sites also contain tyrosine and histidine, although in the protein sequence, the histidine is located in close proximity to the tyrosine and is part of the N-terminal domain. However, the active site in hyaluronan lyase and chondroitinase AC is a tetrad that includes additional arginine and glutamate connected together by hydrogen bonds. In chondroitinase AC, the active site is composed of Tyr234-His225-Arg288-Glu371, with the Glu371 coming from the C-terminal domain (Fethiere et al. 1999). Interestingly, His431 of BtHepIII is in a nearly identical location as the His225 (Fig. 2E), despite different linear order of the active site residues in primary sequences of these two proteins, suggesting path of the evolution of the active site within the enzymes containing this fold.

Normal mode analysis

The 3D structures of proteins from the PL12 family captured in the crystals display different conformations of their N-terminal α-helical domains and, based on our analysis described above, none of them represent the active conformation. This is different from enzyme of several other PL families with the same fold, namely hyaluronan lyase, chondroitinase AC and heparinase II, which crystallized in active conformations. However, this is reminiscent of chondroitinase ABC, which was captured in an open conformation. To understand better the conformational dynamic of PL12 family enzymes, we have resorted to molecular dynamics. Since we were interested in large-scale movements rather than local dynamics, we have opted for the normal mode analysis (Bahar et al. 2010). Starting from the crystal structures, we have performed normal mode analyses for all

three proteins. In each case, we found that two orthogonal normal modes describe sufficiently well the large-scale motions (Fig. 5A–C). In one mode, a bending motion brought the N- and C-terminal domains closer to each other, whereas in the other mode, a rotation motion adjusted the orientation of the two domains (see Movies M1–6 in supporting information). These motions occur within the N-terminal α-helical domain and the hinge region is located within the third α-hairpin. The active site residues Tyr301 and His431 in the C-terminal segment remain in the same position relative to each other. The putative C6-carboxylic group neutralizing residue, Asn247, is in the N-terminal segment and its position relative to the active site residues changes during the large-scale motion of the enzyme. The hinge points in trajectories of all three HepIII enzymes are very closely mirroring the hinge points deduced from comparing the three crystal structures. We were further interested if the movement along these two normal modes brings all three proteins to pass through a common conformation. This was indeed the case and the common conformation is shown in Fig. 5D. This might be an approximation of the active conformation. The position of the tetrasaccharide substrate, as transferred from heparinase II, in the observed conformation is shown in Fig 5E (top) and in the conformation common to all trajectories in Fig 5E (bottom).

The fact that the all structures captured in the crystal are in an inactive conformation suggests that the lowest energy state of these lyases does not represent the conformation required for activity and that the shift to active conformation is driven by the energy derived from substrate binding. It might be that this conformational change is stimulated more easily by a longer heparan sulfate chain rather

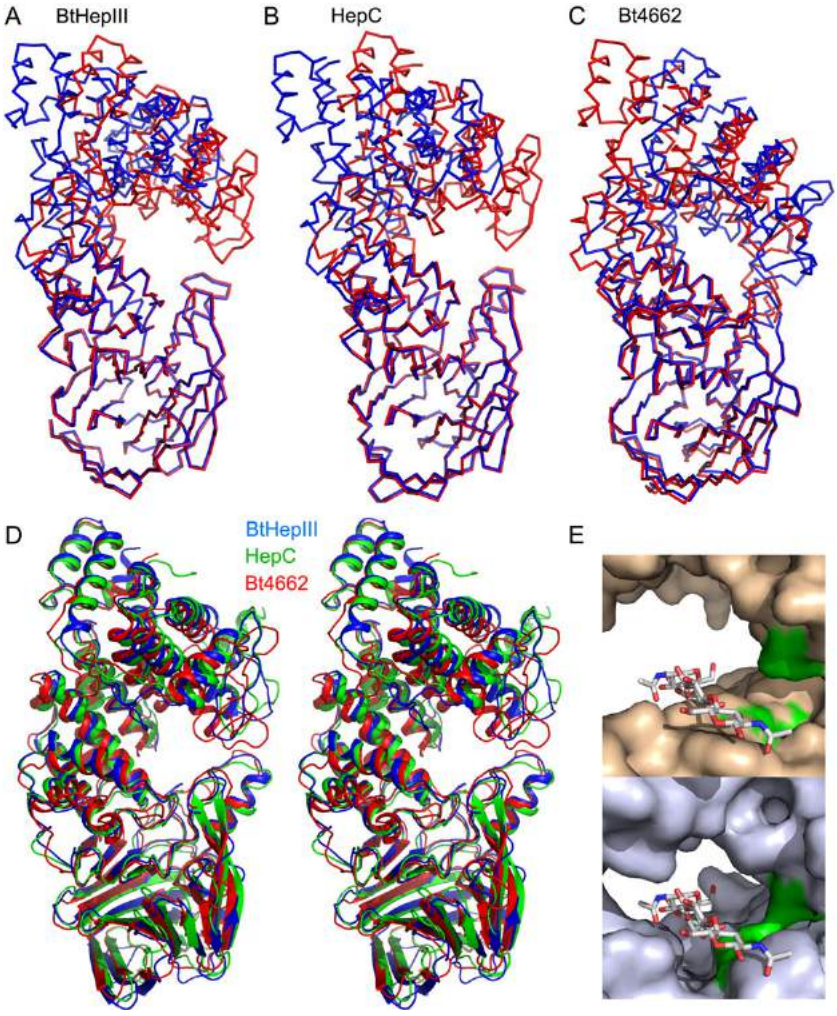


Fig. 5. Trajectory along the two normal modes of the normal mode analysis. All molecules are aligned using the C-terminal β-sheet domain only and shown in the same orientation. (**A**) Initial, open (blue, from the crystal structure) and the last, closed conformation (red) along the calculated trajectory for BtHepIII (see Supplementary –movie M1-2); (**B**) the same for HepC, first: blue, last: red (see Supplementary movie M3–4); (**C**) Bt4662, first: blue, last: red. Here the initial conformation is of the closed state and the final one is of the open state (see Supplementary movie M5–6); (**D**) stereo view of the conformations along the 51-frame trajectories of BtHepIII (blue), BT4662 (red) and HepC (green) that are closest to each other. They correspond to frame 25 for BtHepIII, frame 20 for Bt4662 and frame 51 for HepC and we reason that they approximate the active conformation of this family of heparinases; (**E**) the tetrasaccharide fitted into the observed conformation of BtHepIII (surface representation, top) and in the conformation common to all trajectories (bottom). The surface corresponding to the catalytically essential residues, Tyr301, His431 and Asn247, is colored green. This figure is available in black and white in print and in color at *Glycobiology* online.

than by a tetra- or hexa-saccharides that were used in the co-crystallization attempts.

Materials and methods

Molecular cloning

The gene BT4657 from *B. thetaiotaomicron* (BtHepIII) encodes a protein of 666 amino acids length with the signal peptide comprising the first 22 residues. The DNA segment corresponding to the protein region of 23–666 was PCR amplified from genomic DNA and cloned into the expression vector pRL652 containing N-terminal GST tag followed by the tobacco etch virus (TEV) cleavage site.

Protein expression and purification

The cloned BtHepIII was expressed in *Escherichia coli* BL21 (DE3) expression cell lines. An overnight inoculum of the *E. coli* expression strain harboring BtHepIII was subcultured in 1 l of TB medium supplemented with ampicillin and cells were allowed to grow at 37 °C until the absorbance at 600 nm reached ~1. The temperature was reduced to 20° and protein expression was induced by addition of IPTG to the final concentration of 1 mM. After an overnight growth, the cells were harvested by centrifugation at 6000 rpm for 20 min and re-suspended in the lysis buffer (15 mM Hepes pH 7.5, 150 mM NaCl, 5 mM dithiothreitol (DTT)). The cells were lysed using cell disruptor and the cell debris was separated by centrifugation at 20,000 rpm for 1 h. The soluble fraction was mixed with the GST resin and incubated for 2 h at 4 °C. The resin was loaded on a column and washed with phosphate buffer saline pH 7.4 (PBS-137 mm NaCl, 2.7 mM KCl, 4.3 mM Na_2HPO_4, 1.47 mM KH_2PO_4) containing additional 350 mM NaCl and 500 mM NaCl salt. The GST tag was cleaved at room temperature by TEV protease added at the mass ratio of 1:50 over the course of 8 h. The final solution was loaded on a Ni-NTA column to remove His-tagged TEV protease. The flow through contained highly pure BtHepIII.

Methylation of BtHepIII

Lysine methylation is one of the rescue strategies when native protein fails to crystallize. Purified recombinant BtHepIII was concentrated to 9 mg/mL. Methylating reagent, 20 µL of dimethylamine borane complex (Hampton Research, Aliso Viejo, CA) was added to 1 mL of protein followed by 40 µL of 10% formaldehyde, mixed well and incubated at 4°C for 2 h. The previous step was repeated and incubated again for 2 h. After 2 h, 10 µL of the methylating reagent was added and incubated overnight at 4°C. The reaction was stopped by adding 125 µL of 1 M glycine. The methylating reagents were removed by gel filtration. Methylated BtHepIII was loaded on the Superdex 200 column and the protein eluted as a single peak corresponding to a molecular weight of 72.5 kDa for a monomeric BtHepIII. The fractions were analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and those containing BtHepIII were pooled together.

BtHepIII specificity

Purified recombinant BtHepIII was tested on a number of potential substrates to examine their specificity. Substrates at the concentration of 1 mg/mL were digested overnight at room temperature with 10 µg/mL of enzyme in 15 mM HEPES at pH 7.5.

HILIC LC–FTMS analysis of digested disaccharides

Fully digested sample was analyzed by an LC–MS as previously described with minor modifications (Li et al. 2012). The column used in the current study was a Luna HILIC column, 2.0 × 50 mm (Phenomenex, Torrance, CA). Mobile phase A was 5 mM ammonium acetate prepared with HPLC grade water. Mobile B was 5 mM ammonium acetate prepared in 98% HPLC grade acetonitrile with 2% of HPLC grade water. The mobile phase used an 8 min gradient from 95% mobile phase B (5% mobile phase A) to 20% mobile phase B (80% mobile phase A). Sample (5 × 50 mm; Phenomenex). Mobile phase A was 5 mM ammonium acetate prepared with HPLC grade water. Mobile B was 5 mM ammonium acetate prepared in 98% HPLC grade acetonitrile with 2% of HPLC grade water. The mobile ph.2 kV, a capillary voltage of −40 V, a tube lens voltage of −50 V, a capillary temperature of 275 °C, a sheath flow rate of 30 and an auxiliary gas flow rate of 6. External calibration of mass spectra routinely produced a mass accuracy of better than 3 ppm. All FT mass spectra were acquired at a resolution 60,000 with 400–2000 Da mass range. Extracted disaccharides ion chromatographic peak area was integrated for relative quantification to compare the digestion efficiency for different substrates.

BtHepIII kinetics on heparan sulfate

The apparent Michaelis–Menten parameters were determined for BtHepIII acting on heparan sulfate as previously described (Lohse and Linhardt 1992). Briefly, porcine mucosal heparan sulfate (Celsus Laboratories, Cincinnati, OH) was dissolved in 50 mM sodium phosphate, 100 mM sodium chloride buffer (pH 7.6). Six concentrations of heparan sulfate were prepared, corresponding to final concentrations ranging from 75 to 2400 µM. A negative control (no heparan sulfate) was used to remove background absorbance. Each concentration of heparan sulfate was assayed in triplicate (three unique wells per concentration) on a quartz crystal 96-well plate. The quartz crystal plate was then incubated for 5 min at 35 °C before purified heparinase III was added to each well; the final BtHepIII concentration was 2 µg/mL. The plate was immediately placed in a temperature controlled SpectraMax plate reader (Molecular Devices, Sunnyvale, CA) pre-set to 35°C. Absorbance readings at 230 nm were taken every 20 s for 5 min. Enzyme velocities at each heparan sulfate concentration were determined from the kinetic data generated on the plate reader and analyzed using the nonlinear regression, Michaelis–Menten curve fitting function of GraphPad 6.0.

Crystallization and X-ray diffraction

Initial attempts to crystallize BtHepIII were unsuccessful but the trypsin-treated fragment of BtHepIII crystallized from 0.2 M Mg formate pH 5.9, 16% PEG 3350 and diffracted to 2.9 Å resolution at the synchrotron. The SeMet BtHepIII was purified, proteolyzed by trypsin and crystallized under the same conditions. These crystals diffracted to 3.5 Å.

In an attempt to obtain crystals of full-length BtHepIII, the surface of this protein was modified by methylation of surface exposed lysines. Methylated BtHepIII was concentrated to 15 mg/mL in a buffer containing 20 mM Hepes 7.5, 150 mM NaCl and 3 mM DTT. Crystallization trials with the use of Gryphon robot (ArtRobbins Instruments, Sunnyvale, CA) resulted in crystals of the full-length protein. Two different conditions led to crystals and both were optimized by microseeding in hanging drop vapor diffusion method in a 24-well plate format. The first crystal form was obtained from the condition

20% (W/V) PEG 3350, 100 mM *Bis-Tris* pH 5.5, 200 mM $MgCl_2$ and 500 mM of zwitterionic compound 3-(1-pyridino)-1-propane sulfonate (NDSB-201, Hampton Research). These crystals belong to the *C*2 space group and diffracted to 1.85 Å resolution at the synchrotron. The second crystal form was obtained from 25% (W/V) PEG 3350, 100 mM *Tris* pH 8.5, 200 mM $MgCl_2$ and 500 mM NDSB-201. They belong to space group $P2_1$ and diffracted to 2.4 Å resolution. For data collection, both sets of crystals were transferred to cryo solution containing 20% glycerol in mother liquor and were flash cooled with liquid nitrogen. Diffraction data were collected at the 08ID beamline at the Canadian Light Source synchrotron and were processed with HKL2000 (Otwinowski and Minor 1997). Data collection statistics are summarized in Table III.

Structure determination and refinement

The molecular weight of the protein recovered from the trypsin digested protein crystals was ~55 kDa as determined by SDS-PAGE. Mass spectrometry identified two likely fragments, one starting at Ile188 and the second at Leu203. Crystals of the truncated SeMet BtHepIII diffracted to 3.5 Å and the Se sub-structure was solved by Single Wavelength Anomalous diffraction data of the selenomethionine derivative crystal. The initial model was built using Phenix AutoBuild into an isomorphous native data set of 2.9 Å and completed manually. The residues visible in the electron density map are 219–666 suggesting that the crystallized fragment was 203–666. Several cycles of refinement with Refmac5 interlaced with manual rebuilding reduced the R_{work} to 0.231 and R_{free} to 0.297.

The structures of the methylated form of BtHepIII were solved by molecular replacement with the program Phaser (McCoy et al. 2007) using partial model obtained from the trypsin-treated BtHepIII. OASIS package (He et al. 2007) was used to improve the electron density map. Subsequent model building and refinement were performed using ARP/wARP (Langer et al. 2008), COOT (Emsley and Cowtan 2004) and PHENIX (Adams et al. 2011). The final R_{work}/R_{free} for the C2 form are 0.167/0.202 and for the $P2_1$ form are 0.215/0.256.

Site-directed mutagenesis for BtHepIII

Site-directed mutagenesis of several BtHepIII residues was performed using Quickchange site-directed mutagenesis kit (Agilent Technologies, Santa Clara, CA) using KOD polymerase following manufacturer instructions. The presence of designed mutations was confirmed by DNA sequencing. The mutants were expressed in the BL21 cell line.

Enzyme assay to test activity of various BtHepIII mutants

The activities of various BtHepIII mutants were tested by measuring the formation of unsaturated double bond at 230 nm. The assay solution composed of 0.15 mg/mL heparan sulfate in 20 mM Hepes 7.5, 150 mM NaCl and 2 mM DTT. Of note, 1 µg of the enzyme was added directly to 100 µL of assay solution to a final concentration of 138 nM. A control reaction without enzyme was used to remove background absorbance from substrate. The plate was immediately placed on a temperature controlled SpectraMax plate reader (Molecular Devices) set at 25 °C. Absorbance readings at 230 nm were taken every 20 s for 30 min.

Normal mode analysis

Normal mode analysis (NMA) is a useful tool to investigate large-scale conformational motions of proteins such as ligand-induced conformational changes (Bahar et al. 2010). In NMA, a Hessian matrix, in which the elements are mass weighted secondary derivatives of protein potential energy, is diagonalized. The resulted eigenvectors are normal modes, which describe the directions of protein motions/conformational changes, whereas the eigenvalues are the

Table III. X-ray data collection and refinement statistics for two crystal forms of BtHepIII

Crystal form	Form 1	Form 2	Truncated BtHepIII SeMet	Truncated BtHepIII
Space group	*C*2	$P2_1$	$P4_32_12$	$P4_32_12$
a, *b*, *c* (Å), *β*(°)	186.7, 45.7, 80.5, 94.05	66.5, 80.5, 83.3, 103.9	134.0, 134.0, 145.2	134.2, 134.2, 145.9
Wavelength (Å)	0.97931	0.9796	0.9786	0.97949
Resolution (Å)	20.5–1.85	40.5–2.4	50–3.50	39.8–2.92
Observed *hkl*	203,710	155,894	262,689	361,465
Unique *hkl*	58,055	33,448	17,203	27,894
Completeness (%)	98.7 (98.0)	99.5 (98.5)	99.8 (99.4)	94.3
Redundancy	3.6	4.7	15.3	12.9
R_{sym}	0.066 (0.590)	0.130 (1.27)	0.128 (0.496)	0.045 (0.128)
$I/(\sigma I)$	27.8 (2.2)	9.3 (1.6)	38.5 (6.5)	23.7 (2.3)
R_{work}	0.165	0.219		0.234
R_{free}	0.203	0.258		0.297
Wilson *B* (Å^2)	44.3	58.1	78.4	79.4
B-factor (Å^2, atoms)				
Protein	41.9 (5278)	67.8 (5146)		
Solvent	44.5 (536)	57.5 (122)		
Ramachandran plot				
Favored (%)	97.1	94.3		
Allowed (%)	2.7	5.1		
Disallowed (%)	0.2	0.6		
R.m.s.d.				
Bonds (Å)	0.010	0.003		
Angles (°)	0.98	0.6		
PDB code	5JMF	5JMD		

frequencies of the corresponding normal modes. The collective movements of the protein are approximately described by superposition of several low-frequency normal modes, whereas the other normal modes mostly represent local vibrations, which are not responsible for large-scale conformational changes.

To investigate the mechanism of the conformational changes of BtHepIII between the open state and the closed state, NMA was performed for the apo form of three homologous proteins: the crystal structure of the open state of BtHepIII, as well as the crystal structure of the closed state of *B. thetaiotaomicron* Bt4662 (PDB ID: 4FNV) (Dong et al. 2012) and the crystal structure of the open state of *P. heparinus* HepC (PDB ID 4MMH) (Hashimoto et al. 2014). Calculations were conducted by elNemo (Suhre and Sanejouand. 2004). The potential energy of the protein was described by an elastic network with a cutoff of 8 Å. All of the heavy atoms of the structures were used in NMA. Homology models of the closed state of BtHepIII, HepC and the open state of *B. thetaiotaomicron* Bt4662 were constructed and used to determine the contributions of each mode to the conformational changes. For NMA of each crystal structure, two low-frequency normal modes were selected and analyzed based on their contributions. Perturbation was applied on the crystal structures to get deformed conformations along these normal modes by g_nmtraj tool in GROMACS package (Christen et al. 2005). The magnitude of the deformation was determined based on the distance between the deformed structure and the homology model.

Supplementary data

Supplementary data for this article is available online at http://glycob.oxfordjournals.org/.

Authors' contributions

M.C., R.J.L. and D.T.P. designed the study, analyzed the data and wrote the paper. T.S.U., R.S. and D.Q. purified the protein and solved the structures of two crystal forms, M.L.G. solved the structure of truncated heparinase, M. Ch. crystallized the protein, R.X.G. performed normal mode analysis and interpreted the results, E.S., G.L., L.L. and T.S.U. performed the kinetic experiments and interpreted their results. All authors analyzed the results and approved the final version of the manuscript.

Funding

C.L.S. is supported by the Canada Foundation for Innovation, Natural Sciences and Engineering Research Council of Canada, the University of Saskatchewan, the Government of Saskatchewan, Western Economic Diversification Canada, the National Research Council Canada and the Canadian Institutes of Health Research. This research in M.C. laboratory was supported by the Natural Sciences and Engineering Research Council grant. D.P.T. is the Alberta Innovates Technology Futures Strategic Chair in (Bio)Molecular Simulation and an Alberta Innovates Health Solutions (AIHS) Scientist. R.X.G. is supported by postdoctoral fellowships from the CREATE Training Program in Bionanomachines, Canadian Institutes for Health Research (CIHR, funding reference number: MFE-140949), and AIHS. This work was supported by CIHR (MOP-62690 to D.P.T.). This research was supported in part by the US National Institutes of Health in the form of grants HL62244, and HL094463 to R.J.L.

Acknowledgements

We would like to acknowledge the help of Ms Yunge Li in crystallization of form 1 of BtHepIII, the Protein Characterization and Crystallization Facility, College of Medicine, University of Saskatchewan for access to the crystallization robot and the 08ID-1 beamline at the Canadian Light Source (CLS) for diffraction data collection.

Conflict of interest

None declared.

Abbreviations

DTT, dithiothreitol; GAG, glycosaminoglycan; GlcA, glucuronic acid; IdoA, iduronic acid; PL, polysaccharide lyases; HepII, heparinase II; BtHepIII, *Bacteroides thetaiotaomicron* heparinase III; HepC, *Pedobacter heparinus* heparinase III; NMA, normal mode analysis; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis.

References

teAdams PD, Afonine PV, Bunkoczi G, Chen VB, Echols N, Headd JJ, Hung LW, Jain S, Kapral GJ, Grosse Kunstleve RW, et al. 2011. The Phenix software for automated determination of macromolecular structures. *Methods*. 55:94–106.

Bahar I, Lezon TR, Yang LW, Eyal E. 2010. Global dynamics of proteins: Bridging between structure and function. *Annu Rev Biophys*. 39:23–42.

Bülow HE, Hobert O. 2006. The molecular diversity of glycosaminoglycans shapes animal development. *Annu. Rev. Cell Dev. Biol*. 22:375–407.

Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B. 2009. The carbohydrate-active enZymes database (CAZy): An expert resource for Glycogenomics. *Nucleic Acids Res*. 37:D233–D238.

Capila I, Wu Y, Rethwisch DW, Matte A, Cygler M, Linhardt RJ. 2002. Role of arginine 292 in the catalytic activity of chondroitin AC lyase from Flavobacterium heparinum. *Biochim.Biophys.Acta*. 1597:260–270.

Christen M, Hunenberger PH, Bakowies D, Baron R, Burgi R, Geerke DP, Heinz TN, Kastenholz MA, Krautler V, Oostenbrink C, et al. 2005. The GROMOS software for biomolecular simulation: GROMOS05. *J Comput Chem*. 26:1719–1751.

Desai UR, Wang HM, Linhardt RJ. 1993a. Specificity studies on the heparin lyases from Flavobacterium heparinum. *Biochemistry*. 32:8140–8145.

Desai UR, Wang HM, Linhardt RJ. 1993b. Substrate specificity of the heparin lyases from Flavobacterium heparinum. *Arch.Biochem.Biophys*. 306:461–468.

Dong D, Ako R, Wu B. 2012. Crystal structures of human sulfotransferases: Insights into the mechanisms of action and substrate selectivity. *Expert Opin Drug Metab Toxicol*. 8:635–646.

Dong W, Lu W, McKeehan W, Luo Y, Ye S. 2012. Structural basis of heparan sulfate-specific degradation by heparinase III. *Protein Cell*. 3:950–961.

Emsley P, Cowtan K. 2004. COOT: Model-building tools for molecular graphics. *Acta crystallogr. D*. 60:2126–2132.

Ernst S, Langer R, Cooney CL, Sasisekharan R. 1995. Enzymatic degradation of glycosaminoglycans. *Crit.Rev.Biochem.Mol.Biol*. 30:387–444.

Fethiere J, Eggimann B, Cygler M. 1999. Crystal structure of chondroitin AC lyase, a representative of a family of glycosaminoglycan degrading enzymes. *J. Mol. Biol*. 288:635–647.

Gacesa P. 1987. Alginate-modifying enzymes. A proposed unified mechanism of action for the lyases and epimerases. *FEBS Lett*. 212:199–202.

Garron ML, Cygler M. 2010. Structural and mechanistic classification of uronic acid-containing polysaccharide lyases. *Glycobiology*. 20:1547–1573.

Garron ML, Cygler M. 2014. Uronic polysaccharide degrading enzymes. *Curr. Opin. Struct. Biol*. 28C:87–95.

Godavarti R, Sasisekharan R. 1996. A comparative analysis of the primary sequences and characteristics of heparinases I, II, and III from Flavobacterium heparinum *Biochem. Biophys. Res. Commun*. 229:770–777.

Han YH, Garron ML, Kim HY, Kim WS, Zhang Z, Ryu KS, Shaya D, Xiao Z, Cheong C, Kim YS, et al. 2009. Structural snapshots of heparin depolymerization by heparin lyase I. *J. Biol. Chem*. 284:34019–34027.

Hashimoto W, Maruyama Y, Nakamichi Y, Mikami B, Murata K. 2014. Crystal structure of pedobacter heparinus heparin lyase HepIII with the active site in a deep cleft. *Biochemistry*. 53:777–786.

He Y, Yao DQ, Gu YX, Lin ZJ, Zheng CD, Fan HF. 2007. OASIS and molecular-replacement model completion. *Acta Crystallogr. D*. 63:793–799.

Imberty A, Lortat-Jacob H, Perez S. 2007. Structural view of glycosaminoglycan-protein interactions. *Carbohydr.Res*. 342:430–439.

Iozzo RV. 1998. Matrix proteoglycans: From molecular design to cellular function. *Annu.Rev.Biochem*. 67:609–652.

Langer G, Cohen SX, Lamzin VS, Perrakis A. 2008. Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nat. Protocols*. 3:1171–1179.

Li L, Zhang F, Zaia J, Linhardt RJ. 2012. Top-down approach for the direct characterization of low molecular weight heparins using LC-FT-MS. *Anal. Chem*. 84:8822–8829.

Li S, Kelly SJ, Lamani E, Ferraroni M, Jedrzejas MJ. 2000. Structural basis of hyaluronan degradation by Streptococcus pneumoniae hyaluronate lyase. *EMBO J*. 19:1228–1240.

Linhardt RJ, Galliher PM, Cooney CL. 1986. Polysaccharide lyases. *Appl. Biochem.Biotechnol*. 12:135–176.

Lohse DL, Linhardt RJ. 1992. Purification and characterization of heparin lyases from Flavobacterium heparinum. *J. Biol. Chem*. 267: 24347–24355.

Lunin VV, Li Y, Linhardt RJ, Miyazono H, Kyogashima M, Kaneko T, Bell AW, Cygler M. 2004. High resolution crystal structure of arthrobacter aurescens chondroitin ac lyase: Enzyme-substrate complex defines the catalytic mechanism. *J. Mol. Biol*. 337:367–386.

McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. 2007. Phaser crystallographic software. *J Appl Crystallogr*. 40: 658–674.

Otwinowski Z, Minor W. 1997. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol*. 276:307–326.

Owens RT, Wagner WD. 1991. Metabolism and turnover of cell surface-associated heparan sulfate proteoglycan and chondroitin sulfate proteoglycan in normal and cholesterol-enriched macrophages. *Arterioscler Thromb Vac Biol*. 11:1752–1758.

Shaya D, Tocilj A, Li Y, Myette J, Venkataraman G, Sasisekharan R, Cygler M. 2006. Crystal structure of heparinase II from Pedobacter heparinus and its complex with a disaccharide product. *J. Biol. Chem*. 281: 15525–15535.

Shaya D, Zhao W, Garron ML, Xiao Z, Cui Q, Zhang Z, Sulea T, Linhardt RJ, Cygler M. 2010. Catalytic mechanism of heparinase II investigated by site-directed mutagenesis and the crystal structure with its substrate. *J Biol Chem*. 285:20051–20061.

Suhre K, Sanejouand YH. 2004. ElNemo: A normal mode web server for protein movement analysis and the generation of templates for molecular replacement. *Nucleic Acids Res*. 32:W610–614.

Yung S, Hausser H, Thomas G, Schaefer L, Kresse H, Davies M. 2004. Catabolism of newly synthesized decorin in vitro by human peritoneal mesothelial cells. *Perit Dial Int*. 24:147–155.

Zheng H, Chruszcz M, Lasota P, Lebioda L, Minor W. 2008. Data mining of metal ion environments present in protein structures. *J Inorg Biochem*. 102:1765–1776.