HS is a sulfated glycosaminoglycan (GAG) polysaccharide produced by nearly all animal species. Structurally, HS is composed of a linear repeating [α-HexUA(β1→4)-GlcNX(α1→4)]n disaccharide motif, in which HexUA is either glucuronic acid (GlcUA) or its C5 epimer, iduronic acid (IdoUA), and GlcNX is N-acetylgalactosamine (GlcNAc) or N-sulfoglucosamine (GlcNS). These chains are further modified by O-sulfation at O2 of HexUA, O6 of GlcNX and (rarely) O3 of GlcNX. The deacetylation, sulfation and epimerization reactions involved in biosynthesis of HS are nontemplated and do not reach completion, thus resulting in substantial structural heterogeneity along chains. Mature HS is modular in composition, containing both 'NS' domains rich in GlcNS, IdoUA and O-sulfated sugars and separate 'NA' domains rich in GlcNAc and GlcUA but having lower O-sulfation (Supplementary Fig. 1). This heterogeneity of HS structure provides interaction sites for a large number of different binding partners and is central to the proper biological function of HS.

HS occurs in vivo in the form of heparan sulfate proteoglycans (HSPGs), which consist of one or more HS chains covalently linked to a core transmembrane or secreted protein. HSPGs are an important constituent of the ECM, which surrounds cells, and they perform important structural and signaling functions via HS-mediated interactions. The compositions of HS chains are adapted to their function and can differ between cells and tissues even when the core HSPG protein is the same. HS chains are also dynamically regulated in response to external stimuli: turnover in some cells occurs with a half-life as rapid as 2.5 h (refs. 5, 6). This turnover is underpinned by a network of enzymes that serve to efficiently synthesize and break down HS in a regulated fashion.

The principal enzyme involved in breakdown of HS is HPSE, a member of the carbohydrate-active enzyme (CAzy) GH79 family of carbohydrate-processing enzymes. HPSE catalyzes hydrolysis of internal GlcUA(β1→4)GlcNS linkages in HS, with net retention of anomeric configuration. HPSE-mediated breakdown of HS is not indiscriminate but instead is restricted to a small subset of GlcUAs, thus reflecting a requirement for specific N- and O-sulfation patterns on neighboring sugars. HPSE present in late endosomes and lysosomes performs an essential housekeeping role in catabolic processing of internalized HSPGs. In addition, HPSE can be trafficked to the cell surface or released into the ECM, where it effects breakdown of extracellular pools of HS.

HPSE-mediated breakdown of HS in the ECM has several effects on the behavior of nearby cells. Weakening of structural HS networks in the ECM and basement membranes directly facilitates cell motility and extravasation into surrounding tissues. Latent pools of growth factors stored by HS are released upon breakdown by HPSE and subsequently promote increased cell proliferation, motility and angiogenesis. HS fragments generated by HPSE activity can also activate downstream signaling cascades. Whereas controlled HPSE activity plays an important part in physiological processing of the ECM, aberrant HPSE expression is associated with inflammation and cancerous growth. The proliferative advantages conferred by HPSE lead to its upregulation in tumors in a variety of tissues, and HPSE overexpression correlates strongly with metastasis and worsened clinical prognoses.

Only one gene with heparanase-like catalytic activity has been identified in mammals to date, thus suggesting that loss of HPSE activity may not easily be compensated for by the cell. Accordingly, HPSE inhibition has attracted intense interest as an anticancer strategy, although the efficacy of small-molecule inhibitors has yet to rival those reported for oligosaccharide-like HS mimetics, such as the sulfated phosphomannopentaose derivative PI-88 (which is currently in phase III clinical trials as an adjuvant therapy for the treatment of...
viral hepatits–related hepatocellular carcinomas).

Despite intense biological and clinical interest, no three-dimensional (3D) structure has been reported for human HPSE. Structures of a GH79 exoglucuronidase from Acidobacterium capsulatum (AcaGH79) and, recently, of an endo–acting heparanase from Burkholderia pseudomallei (BpHPSE) have helped provide a structural overview for this enzyme family. To understand the mechanistic basis of HPSE function, we determined the crystal structures of human HPSE in apo and ligand-bound states. These data shed light on the structure and function of a key enzyme involved in human HS metabolism and may be of substantial utility in future efforts to rationally design inhibitors of HPSE.

RESULTS Tertiary structure of HPSE

HPSE is initially translated as a preproenzyme containing a signal sequence spanning Met1–Ala35. Cleavage of this signal sequence by signal peptidase leaves an inactive 65-kDa proHPSE, which must undergo further processing for activity. Proteolytic removal by cathepsin L of a linker spanning Ser110–Gln157 liberates an N-terminal 8-kDa subunit and a C-terminal 50-kDa subunit, which remain associated as a noncovalent heterodimer in mature active HPSE.

For our studies, we expressed HPSE through a baculovirus system, using a previously described dual-expression strategy. We placed cDNA encoding the 8-kDa and 50-kDa subunits into a single bacmid under the control of separate viral promoters. The two subunits cotranslationally fold into mature heterodimeric HPSE, thereby bypassing the 65-kDa proenzyme form and ensuring expression of only active enzyme (Fig. 1a).

The structure of apo-HPSE, refined to 1.75-Å resolution (Table 1), contains a single heterodimer in the asymmetric unit, comprising residues Gln36–Glu109 of the 8-kDa subunit and Lys159–Ile543 of the 50-kDa subunit (according to numbering based on the full proproenzyme). The domain architecture of HPSE comprises a (β/α)₈ domain flanked by a smaller β-sandwich domain. Both 8-kDa-and 50-kDa subunits are structurally involved in both domains: the 8-kDa subunit contributes one β-sheet to the β-sandwich and the first β-α-β fold of the (β/α)₈ domain, and the remaining folds are contributed by the 50-kDa subunit (Fig. 1b). Overall, the domain architecture of HPSE is superficially similar to that of previously characterized bacterial GH79s, with Ca r.m.s. differences of 2.35 Å over 392 residues (out of 457) and 2.59 Å over 387 residues for AcaGH79 and BpHPSE, respectively (Supplementary Fig. 2a).

HPSE contains six putative N-glycosylation sites, all residing on the 50-kDa subunit. N-linked GlcNAc residues (corresponding to N-glycan trees treated by endoglycosidase H during protein preparation) were present in the apo-HPSE structure at Asn162, Asn200, Asn217, Asn238 and Asn459. Additionally, we observed a core α1→6-linked fucose on the GlcNAc linked to Asn459. We found no noticeable density corresponding to GlcNAc at the N-glycosylation site Asn178, thus suggesting that this position may not be well N-glycosylated during baculoviral expression or that N-GlcNAc in this position is not compatible with crystal packing.

The GH79 family belongs to the larger GH-A clan, which is characterized by a (β/α)₈ domain containing the catalytic site. We clearly observed a cleft spanning ~10 Å in the (β/α)₈ domain of apo-HPSE, thus suggesting that the HS-binding site is contained within this part of the enzyme (Fig. 1c). This cleft contains residues Glu343 and Glu225, which have been previously identified as the catalytic nucleophile and acid-base of HPSE and are required for the retaining catalytic mechanism. In accordance with the negatively charged nature of its HS substrate, the HPSE binding cleft is lined by basic side chains contributed by Arg35, Lys158, Lys159, Lys161, Lys231, Arg272, Arg273 and Arg303.

The positioning of the 8-kDa-subunit C terminus and 50-kDa-subunit N terminus in our model indicates that the excised Ser110–Gln157 linker of proHPSE should lie very near or even within the active site cleft of the (β/α)₈ domain. This positioning would hinder incoming HS substrates and is consistent with a previously proposed steric-block mechanism for proHPSE inactivation by its own linker.

Structural basis of HPSE-substrate interactions

Interaction of HS with HPSE is influenced by substrate sulfation, and only HS sequences with particular sulfation patterns are hydrolyzed.
Studies using defined HS oligomers have suggested that HPSE preferentially cleaves a trisaccharide with sulfated GlcNX residues at the –2 and +1 positions\(^9\),\(^{10}\),\(^{34}\),\(^{35}\) (subsite nomenclature in ref. 36). However, a mechanistic understanding of how HPSE ‘reads’ the sulfation status of HS substrates and selects favorable cleavage sites has so far been lacking.

We mapped the substrate-binding and sulfate-interaction sites of HPSE by obtaining structures of the enzyme in complex with a set of HS analogs. We initially chose three ‘HepMers’, semisynthetic HS lacking.

| Data collection | Apo-HPSE (PDB 5E8M) | M04 S00a complex (PDB 5E97) | M04 S02a complex (PDB 5E98) | M09 S05a complex (PDB 5E9B) | Dp4 complex (PDB 5E9C) |
|-----------------|----------------------|-----------------------------|-----------------------------|-----------------------------|------------------------|
| Space group     | P2\(_1\)             | P2\(_1\)                    | P2\(_1\)                    | P2\(_1\)                    | P2\(_1\)               |
| Resolution (Å)  | 46.64–1.75           | 40.99–1.64                  | 38.42–1.63                  | 52.52–1.88                  | 52.64–1.73             |
| Completeness (%)| 99.5 (99.8)          | 100.0 (99.8)                | 99.8 (99.8)                 | 99.8 (99.7)                 | 99.0 (98.4)            |
| Redundancy      | 6.1 (5.6)            | 4.2 (4.1)                   | 4.0 (3.6)                   | 4.0 (3.8)                   | 4.1 (4.2)              |
| No. reflections | 49,329               | 57,972                      | 58,820                      | 38,830                      | 49,793                 |
| R\(_{merge}\) / R\(_{free}\)| 0.16 / 0.20          | 0.17 / 0.20                 | 0.17 / 0.20                 | 0.18 / 0.22                 | 0.17 / 0.21            |
| No. atoms       | 3,669                | 3,652                       | 3,657                       | 3,652                       | 3,660                  |
| Protein         | 249                  | 397                         | 223                         | 208                         | 228                    |
| Ligand/ion      | 63                   | 66                          | 51                          | 60                          | 72                     |
| Water           | 36.3                 | 28.4                        | 36.3                        | 42.1                        | 35.5                   |
| B factors       | 47.4                 | 41.5                        | 45.6                        | 59.2                        | 68.5                   |
| r.m.s. deviations | 1.74 (1.67)          | 0.014 (0.014)               | 0.014 (0.014)               | 0.015 (0.015)               | 0.015 (0.015)          |
| Bond angles (%) | 1.66 (1.75)          | 1.66                        | 1.66                        | 1.75                        | 1.78                   |

Values in parentheses are for highest-resolution shell.

Complexes of HPSE with the three HepMers showed clear electron density for the ligands within the active site cleft, thus revealing the basis of interaction at the –1 and –2 binding subsites (Fig. 2b–d). The electron density was progressively more disordered from the –3 sugar onward as the substrate exited the binding cleft, in accordance with HPSE recognizing a trisaccharide spanning the –2, –1 and +1 subsites. For M04 S00a and M04 S02a, the electron density for the reducing-end paranitrophenol (pNP) moiety was clearly visible at the +1-equivalent position, thus suggesting that these molecules are poor substrates for HPSE. In contrast, density at +1 for M09 S05a was almost absent.

Binding of GlcUA at the –1 subsite was identical in all HepMer complexes, in which the ring was in \(^{4}\)C\(_1\) conformation, and the anomeric carbon was close to the nucleophile Glu343. We observed direct hydrogen bonds to HPSE from GlcUA to the side chains of Asp62, Asn224 and Tyr391 as well as to the backbone NHs of Thr97, Gly349 and Gly350. A strong network of hydrogen bonds from the GlcUA C6 carboxylate to Tyr391, Gly349 and Gly350 was similar to that observed in the bacterial GH79s, thus suggesting the conservation of a key GH79 motif that has been tuned to recognize GlcUA. Our structures also showed that the presence of GlcUA(2S) or IdoUA(2S) at the –1 subsite cannot be tolerated by HPSE, owing to steric clashes between the bulky 2O-sulfate and Asn224.

GlcNX residues at the –2 subsite also adopt a \(^{3}\)C\(_2\) conformation in all complexes and illustrate HPSE’s ability to accommodate a variety of GlcNX sugars at this position. Surprisingly, only the N2 position of GlcNX appeared to make any direct hydrogen-bonding interactions to HPSE: for both GlcNAc and GlcNS, the amide NH formed a hydrogen bond to the side chain of Tyr391. For GlcNAc, the amide carbonyl formed a hydrogen bond to both Asn64 and a structural water molecule, whereas GlcNS made these same interactions as well as hydrogen bonds to a further structural water and the backbone NH of Gly389. For –2 GlcNS(6S), 6-sulfate was ideally placed to participate in electrostatic interactions with the side chain of Lys159. Hence, GlcNS(6S) at the –2 subsite should be favored over GlcNS, which is favored over GlcNAc, because of the formation of additional electrostatic and hydrogen-bonding interactions, respectively.

Lack of density at the +1 subsite for M09 S05a, in contrast to that for M04 S00a and M04 S02a, suggested that the observed structure reflected several related complexes differing at this position. Inspection of GlcNS(6S) at the –2 subsite also revealed that the electron density for 6-sulfate was considerably weaker than that for N-sulfate (Supplementary Fig. 2b), thus indicating that this
subsite was probably occupied by a mixture of GlcNS and GlcNS(6S). Although M09 S05a contains several GlcNS residues, binding of HPSE far toward the reducing end of M09 S05a is strongly disfavored in crystallo, owing to the presence of a clashing symmetry molecule at the ‘positive’ end of the binding cleft (Supplementary Fig. 2c). Hence, we concluded that the M09 S05a complex reflected partial HPSE binding only one disaccharide unit further from the reducing end. This would imply GlcNS(6S) occupancy of the +1 subsite or alternatively a product complex in which a +1 GlcNS(6S) had been hydrolyzed. We were thus able to determine kinetic parameters for HPSE hydrolysis of M09 S05a through a colorimetric reducing-end assay \( K_m = 7.70 \pm 1.42 \mu M, k_{cat} = 0.53 \pm 0.02 s^{-1} \) (mean ± s.e.m.; Supplementary Fig. 3) but not for M04 S00a or M04 S02a. Whereas \( K_m \) for M09 S05a hydrolysis was in the low-micromolar range, comparably to previously reported HPSE activity \( (1.64 \mu M \text{ for natural HS}^{37} \text{ and } 46 \mu M \text{ for the synthetic substrate fondaparinux}^{38} \), \( k_{cat} \) for M09 S05a was ~15% of that previously determined for fondaparinux \( (3.5 s^{-1}) \), possibly reflecting differences in substrate sulfation or oligosaccharide length.

Because +1 subsite interactions were not resolvable through HepMers, we sought to obtain a structure of HPSE in complex with a heparin-derived tetrasaccharide (hereafter dp4, generated by depolymerization of full-length heparin by heparin lyases; Fig. 2a). Like HS and heparin itself, dp4 is a mixture of structurally related disaccharide units, with the tetrasaccharide \( \Delta \text{HexUA(2S)-GlcNS(6S)-IdoUA(2S)-GlcNS(6S)} \) as the major component. The structure of HPSE with dp4 showed occupancy of the same binding site as for HepMers (Fig. 2e), consistently with the overall similarity of heparin to HS. The observed density was not, however, for the predominant dp4 tetrasaccharide but was better modeled by a molecule containing IdoUA at –1 (instead of IdoUA(2S)), thus confirming that 2O-sulfate at the –1 subsite is not tolerated. The electron density for dp4 in HPSE was correspondingly weaker than for HepMers across all subsites, a result reflecting the binding of a minor constituent.

The overall configuration of dp4 across the HPSE active site was similar to that of HepMers, but –1 IdoUA adopted a \( ^{2}S_{0} \) conformation instead of the \( ^{4}C_{1} \) conformation seen for GlcUA in HepMers. In free heparin, IdoUA residues exist in an easily traversable equilibrium between \( ^{1}C_{4} \) and \( ^{2}S_{0} \), with the O2 position held axial or equatorial, respectively\(^{39} \). Upon binding to HPSE, hydrogen-bonding to Asn224 and the presence of clashing Glu343 at the O2 axial position constrain –1 IdoUA to the \( ^{2}S_{0} \) conformation. IdoUA constrained in \( ^{2}S_{0} \) may be hindered from undergoing the conformational changes required for hydrolysis\(^{40} \), thus perhaps explaining why heparin-type ligands containing IdoUA residues act as competitive inhibitors of HS cleavage.

Some density at the +1 subsite, although weak, was resolved in the dp4 complex, thus suggesting the presence of +1 6O-sulfate, which we tentatively modeled as GlcNS(6S). In our model, GlcNS(6S) at the +1 subsite adopts an undistorted \( ^{4}C_{1} \) conformation, in which 6O-sulfate provides the main interactions to HPSE via hydrogen bonds to the backbone and side chain NH of Gln270, as well as electrostatic interactions with Arg272. According to its positioning in our model, further electrostatic interactions may also exist between +1 N-sulfate and Arg303, although we were unable to confirm this, given the poor density for the ligand at this position. Although dp4 contains IdoUA instead of GlcUA at the –1 subsite, both GlcUA \( ^{4}C_{1} \) and IdoUA \( ^{2}S_{0} \) conformations involve a similar all-equatorial arrangement of substituents around the sugar ring. Therefore, we propose that the position...
of GlcNS(6S) at the +1 subsite is also likely to be relevant for an HS substrate containing GlcUA instead of IdoUA.

**HPSE interaction induces distortion of the substrate chain**

The structure of free heparin has been extensively studied; it adopts a linear right-handed helix with a rotation of ~180° and translation of 0.82–0.87 nm per disaccharide unit11–42. Sulfates are presented in clusters on the surface of the sugar chain, and their distribution reflects whether IdOA residues are in the 1C4 or 2SO conformation. The helical heparin conformation is retained in complex with binding partners, as seen in crystal structures of heparin bound to fibroblast growth factor or antithrombin34–46, which also illustrate the role of sulfation in mediating electrostatic interactions with basic residues on heparin-interacting proteins. In contrast, the presence of a bulky negatively charged ‘coat’ surrounding the core glycan may hinder access to nucleophilic protein residues, thereby protecting the sugar residues from enzymatic attack.

Dp4 in complex with HPSE retains an approximately right-handed helical configuration within the active site cleft, similarly to free heparin. However, interactions from HPSE to –2 N-sulfate and +1 O-sulfate introduce a clear bend in the heparin chain across the –2, –1, +1 trisaccharide. In comparison to a previous model for an idealized heparin helix42, the distance between –2(N2) and –1(O2) for dp4 in HPSE increased from 4.8 Å to 7.2 Å, and the distance between –1(O2) and +1(C6) increased from 4.4 Å to 5.7 Å (Fig. 3a,b). This ‘bend’ separates the N2 and O6 sulfates adjacent to the –1 anomeric center and hence allows the catalytic residues of HPSE to access this position more easily. Although we did not observe a +1-subsite sugar in HepMer–HPSE complexes, we found a similar distortion for HepMers across the –2 and –1 subsites (lengthened distance between –2(N2) and –1(O2) of 7.1 Å for M09 S05a; Fig. 3c). These results suggest that the ‘bending’ observed for dp4 is also likely to apply to substrates with a –1 GlcUA instead of IdOA.

Our results indicate a dual role for HS sulfation in interactions with HPSE. Not only does sulfation serve as a molecular signal that directs the enzyme to cleave only certain glycan sites, but –2 N-sulfate and +1 O-sulfate moieties also act as mechanistic handles by which HPSE can prize open a substrate HS helix and more effectively access the anomeric center of the –1 sugar.

**DISCUSSION**

The 3D structure of the human GH79 endoglucuronidase HPSE and its interaction with HS substrates (Fig. 4a) provides a long-anticipated structural rationale that ties together the results of numerous biochemical studies on the HS sulfation patterns required for HPSE cleavage9–11. We confirmed that sulfation is key for HPSE interaction with HS and that the recognized cleavage site is a trisaccharide accommodated in the HPSE binding cleft. Structurally, –2 N-sulfate and +1 O-sulfate appear to be the main determinants for recognition, because these directly contact the enzyme through hydrogen-bonding networks. –2 O-sulfate and +1 N-sulfate may also further stabilize the HPSE-bound trisaccharide through electrostatic interactions to basic residues lining the active site cleft. Our observations are consistent with the major findings of previous biochemical studies.

Whereas the activity of human HPSE (hHPSE) is broadly similar to that of the recently reported bacterial enzyme BpHPSE28, differences exist between the substrates hydrolyzed by each enzyme. BpHPSE has been reported to preferentially cleave HS-containing GlcNAc residues, whereas GlcNS is highly favored by hHPSE. Furthermore, BpHPSE has also been found to degrade chondroitin sulfate GAGs, which are not known to be a substrate for hHPSE. Sequence alignment of several eukaryotic heparanases with BpHPSE and AcaGH79 revealed that residues at the –1 subsite (of hHPSE) are well conserved across all species (thus reflecting absolute specificity for GlcUA), residues at the –2 and +1 subsites show much poorer conservation in the bacterial enzymes, thus providing a rationale for observed differences in substrate specificity (Supplementary Fig. 4). Compared to hHPSE, BpHPSE may represent a more general catabolic enzyme that is used by the bacterium to break down GAGs encountered in its environment.

One point of particular interest is the evolutionary relationship between exo- and endo-acting GH79 enzymes, especially regarding the nature of the loop that forms part of the exo-acting substrate-binding pocket in AcaGH79 (Fig. 4b and Supplementary Fig. 5). In comparison to that in AcaGH79, this loop is substantially reduced in size in endo-acting BpHPSE, thus converting the binding pocket into a large cleft capable of accommodating long GAG chains. In contrast, the human enzyme appears to have evolved endo specificity though expansion of this loop into a proteolytically cleavable linker motif. Proteolytically activated proenzymes are abundant in higher eukaryotes and represent an efficient mechanism to induce enzymatic activity in response to stimuli. However, de novo evolution of proenzymes is extremely unlikely, because it would require the coevolution of both viable enzymatic and proteolyzable sequences. A more plausible scenario involves expansion of a small loop around the active site of an existing enzyme scaffold; the loop would eventually form a larger proteolytically cleavable steric block that could be removed for activation. This mechanism has already been hypothesized for serine proteases such
as the trypsinogen—trypsin and chymotrypsinogen—chymotrypsin
proenzyme-enzyme pairs. Structural observations on HPSE and
related enzymes have provided evidence that such evolutionary pro-
cesses may have also occurred for carbohydrate-processing enzymes.
We anticipate that further examples of such relationships will become
apparent in the near future, as efforts to solve the structures of higher
eukaryotic carbohydrate processing enzymes intensify.

In conclusion, our structures of HPSE and its ligand complexes
provide insight into the mechanisms of action of a key player in ECM
remodeling. The structures reported here should be of utility in future
efforts to design improved inhibitors of HPSE for use as therapeutics
and/or chemical biology tools. Indeed, our complex of HPSE with
the tetrasaccharide dp4 already suggests that introduction of sugars
at the –1 subsite, fixed in particular conformations, may be a viable
strategy for inhibitor design. More fundamentally, GAGs such as HS
have important roles in numerous biological processes, both normal
and pathological. An understanding of the enzymes involved in GAG
processing will be essential for determining how this important and
complex class of carbohydrates is regulated.

METHODS
Methods and any associated references are available in the online
version of the paper.

Accession codes. Coordinates and structure factors have been deposited
in the Protein Data Bank under accession codes PDB 5E8M (apo struc-
ture), PDB 5E97 (M04 S00a complex), PDB 5E98 (M04 S02a complex),
PDB 5E9B (M09 S05a complex) and PDB 5E9C (dp4 complex).

Note: Any Supplementary Information and Source Data files are available in the
online version of the paper.

ACKNOWLEDGMENTS
We thank Diamond Light Source for access to beamlines I02, I03 and
I04 (proposal mx-9948), which contributed to the results presented here.
DH10EMBacY cells used to generate recombinant bacmids were a gift from
I. Berger (University of Bristol). The full HPSE gene was obtained from a
baculoviral construct from D. Jarvis (University of Wyoming; US National
Institutes of Health grant RR005351). G.J.D. and L.W. acknowledge support
from the European Research Council through Advanced Grant Glycopoise (AdG
322942 (G.J.D.)). A.M.B. and C.M.V. thank the UK Medical Research Council for
funding (MR/K000179/1 (A.M.B.)).

Figure 4 Subsite interactions of human HPSE and comparison to other
GH79 enzymes. (a) Composite summary of HPSE-substrate interactions across
the –2, –1 and +1 subsites of the enzyme’s binding cleft. Interactions as mapped by complexes with HepMers and dp4. The catalytic
residues of HPSE have been annotated. Electrostatic interaction with Arg303 is postulated on the basis of the position of N-sulfate in the dp4 model. Nuc, nucleophile.
(b) Structural relationships between the active sites of exo- and endo-acting GH79 enzymes. An extended loop in the (βα)8 domain of AcaGH79 forms part of the
exo-acting substrate-binding pocket wall. This loop is considerably shortened in BpHPSE, thus creating an endo-acting binding cleft. In contrast, hHPSE
has expanded this loop into a large linker sequence, which is proteolytically removed to produce its endo-acting binding cleft.
AUTHOR CONTRIBUTIONS
L.W. and G.J.D. designed and interpreted the experiments. L.W. and C.M.V. cloned, expressed and purified proteins with help from A.M.B. in eukaryotic protein expression facilities. L.W. carried out kinetics experiments and protein crystallizations, and solved the structures of protein and ligand complexes. L.W. and G.J.D. wrote the manuscript.

COMPETING FINANCIAL INTERESTS
The authors declare no competing financial interests.

Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Lawrence, R. et al. Evolutionary differences in glycosaminoglycan fine structure detected by quantitative glycan reductive isotope labeling. J. Biol. Chem. 283, 33674–33684 (2008).
2. Tan, N., Elin, M. & Vlodavsky, I. Regulation, function and clinical significance of heparanase in cancer metastasis and angiogenesis. Int. J. Biochem. Cell Biol. 38, 2018–2039 (2006).
3. Xu, Y.J. et al. N-(4-(4-Hydroxybenzal)azol-2-yl)-arilamino-methyl phenyl-benzamide derivatives as small molecule heparanase inhibitors. Bioorg. Med. Chem. Lett. 16, 404–408 (2006).
4. Courtney, S.M. et al. 2,3-Di-O-hydroxy-1,3-dioxy-1-H-isodole-5-carboxylic acid derivatives: a novel class of small molecule heparanase inhibitors. Bioorg. Med. Chem. Lett. 14, 3269–3273 (2004).
5. Cassellini, G. et al. Antitumor efficacy of the heparanase inhibitor SST0001 alone and in combination with antiangiogenic agents in the treatment of human pediatric sarcoma models. Biochem. Pharmacol. 85, 1424–1432 (2013).
6. Hammond, E., Handley, P., Dedridge, K. & Bytheway, I. Mechanisms of heparanase inhibition by the heparan sulfate mimetic PG545 and three structural analogues. FEBS Open Bio 3, 346–351 (2013).
7. Lewis, K.D. et al. A phase II study of the heparanase inhibitor PI-88 in patients with advanced melanoma. Invest. New Drugs 26, 89–94 (2008).
8. Michikawa, M. et al. Structural and biochemical characterization of glycyl hydrolase family 79 (g-glucuronidase from Acidobacterium capsulatum. J. Biol. Chem. 287, 14069–14077 (2012).
9. Böhlmann, L. et al. Functional and structural characterization of a heparanase. Nat. Glycoconjug. J. 30, 1033–1043 (2013).
10. Fairbanks, M.B. et al. Processing of the human heparanase precursor and evidence that the active enzyme is a heterodimer. J. Biol. Chem. 274, 29587–29590 (1999).
11. McKenzie, E. et al. Biochemical characterization of the active heterodimer form of human heparanase (Hpa1) protein expressed in insect cells. Biochem. J. 371, 423–435 (2003).
12. Hulett, M.D. et al. Identification of active-site residues of the pro-metastatic endoglycosidase heparanase. Biochemistry 39, 15659–15667 (2000).
13. Davies, G. & Henriassat, B. Structures and mechanisms of glycosyl hydrolases. Structure 3, 853–869 (1995).
14. Nardella, C. et al. Mechanism of activation of human heparanase investigated by protein engineering. Biochemistry 43, 1862–1873 (2004).
15. Peterson, S. & Liu, J. Deciphering mode of action of heparanase using structurally defined oligosaccharides. J. Biol. Chem. 287, 34836–34843 (2012).
16. Peterson, S.B. & Liu, J. Unraveling the specificity of heparanase utilizing synthetic substrates. J. Biol. Chem. 285, 14504–14513 (2010).
17. Davies, G.J., Wilson, K.S. & Henriassat, B. Nomenclature for sugar-binding subsites in glycosyl hydrolases. Biochem. J. 321, 557–559 (1997).
18. Fairweather, J.K., Hammond, E., Johnstone, K.D. & Ferro, V. Synthesis and heparanase inhibitory activity of sulfated mannosialoglycosidases related to the antiangiogenic agent PI-88. Bioorg. Med. Chem. 16, 699–709 (2008).
19. Hammond, E., Li, C.P. & Ferro, V. Development of a colorimetric assay for heparanase activity suitable for kinetic analysis and inhibitor screening. Anal. Biochem. 396, 112–116 (2010).
20. Torri, G. et al. Mono- and bidimensional 500 MHz 1H-NMR spectra of a synthetic pentasaccharide corresponding to the binding sequence of heparin to antithrombin-III: evidence for conformational peculiarities of the sulfated iduronic residue. Biochim. Biophys. Acta. 128, 134–140 (1985).
21. Speciale, G., Thompson, A.J., Davies, G.J. & Wlodawer, A. The ternary complex of heparanase with heparin and the calcium ion. Curr. Opin. Struct. Biol. 52, 19–37 (1975).
22. Nardella, C. et al. Structure of a heparanase-hyaluronan complex reveals the basis for inhibitor specificity. Structure 3, 853–869 (1995).
23. Atkins, E.D. & Nieduszynski, I.A. Crystalline structure of heparin. Adv. Exp. Med. Biol. 52, 19–37 (1975).
24. Mulloy, B., Forster, M.J., Jones, C. & Davies, D.B. N.m.r. and molecular-modelling studies of the solution conformation of heparin. Biochem. J. 293, 849–858 (1993).
25. Mikhalkov, D., Linhardt, R.J. & Mayo, K.H. NMR solution conformation of heparin-derivadosaccharide. Biochem. J. 328, 51–61 (1997).
26. Faham, S., Hileman, R.E., Fromm, J.R., Linhardt, R.J. & Rees, D.C. Heparin structure and interactions with basic fibroblast growth factor. Science 271, 1116–1120 (1996).
27. Schlessinger, J. et al. Crystal structure of a ternary FGF-FGFR-heparin complex reveals a dual role for heparin in FGF binding and dimerization. Mol. Cell 6, 743–750 (2000).
28. Dementiev, A., Petitou, M., Herbert, J.M. & Gettins, P.G. The ternary complex of antithrombin-III-heparin-reveals the basis of inhibitor specificity. Nat. Struct. Mol. Biol. 11, 863–867 (2004).
29. Neurath, H. & Walsh, K.A. Role of proteolytic enzymes in biological regulation (a review). Proc. Natl. Acad. Sci. USA 73, 3825–3832 (1976).
ONLINE METHODS

Production of HPSE-expressing baculovirus. DNA was extracted from baculovirus encoding proHPSE N-terminally tagged with a honeybee melittin signal sequence, by use of a Qiagen miniprep kit, with 250 μL of virus stock as input. PCR with restriction site–tailed primers (Supplementary Table 1) was used to generate CDNAs for the melittin signal peptide, 8-kDa HPSE subunit, and 50-kDa HPSE subunit. N-mellitin-HPSE (8 kDa) was subcloned into the BamHI/PstI sites of pFastBac Dual, under the control of the PolH promoter. N-mellitin-HPSE (50 kDa) was subcloned into the XhoI/KpnI sites of the same vector, under the control of p10. For both fragments, melittin was ligated to the HPSE gene via an XmaI restriction site, which leaves an extra DFG tripeptide in the protein upon expression and signal peptidase cleavage.

Recombinant bacmid was produced with the Tn7 transposition method in DH10EMBacY (Geneva Biotech) and purified with the PureLink miniprep kit (Invitrogen) according to standard protocols. V1 baculovirus was produced by transfection of bacmid into low-passage adherent Sf21 cells (Invitrogen) with FuGENE HD transfection reagent (Promega), at a ratio of 2 μg DNA to 4.5 μL FuGENE. V1–V2 virus amplification was carried out with Sf21 cells in suspension culture, and the YFP marker present in EMBacY baculovirus was used to determine optimum amplification before harvesting (typically ~60% cells fluorescent). For expression, Trichoplusia ni cells (Invitrogen) were infected with V2 baculovirus at an MOI >1, and infection followed with the EMbAcY YFP marker to determine the optimum time point for harvesting (typically 72 h, with >80% cells fluorescent). All insect cells used had tested negative for mycoplasma contamination.

Expression and purification of HPSE. 3 L of conditioned medium was cleared of cells by centrifugation at 400g for 15 min at 4 °C. This was followed by further clearing of debris by centrifugation at 4,000g for 60 min at 4 °C. DTSS (1 mM) and AEBSF (0.1 mM) were added to cleared medium, which was then loaded onto a preequilibrated HiTrap Heparin HP 5-mL column (GE Healthcare). The heparin column was washed with 10 CV hepar buffer A (20 mM HEPES, pH 7.4, 100 mM NaCl, 1 mM DTT) and eluted with a linear gradient over 30 CV with hepar buffer B (20 mM HEPES, pH 7.4, 1.5 M NaCl, and 1 mM DTT). HPSE-containing fractions were pooled, diluted ten-fold into IEX buffer A (50 mM phosphate, pH 8.0, 100 mM NaCl, and 1 mM DTT), and loaded onto a preequilibrated HiTrap Sepharose SP HP 1-mL column at 4 °C. Protein was eluted from the SP column with a linear gradient over 30 CV with IEX buffer B (50 mM phosphate, pH 8.0, 1.5 M NaCl, and 1 mM DTT). HPSE-containing fractions were pooled, diluted to 2 mL with a 30-kDa Vivaspin concentrator (GE Healthcare) and treated with 5 μL Endo H (NEB) for 4 h at ambient temperature. Digested protein was purified by size-exclusion chromatography (SEC) with a Superdex 75 16/600 column (GE Healthcare) in SEC buffer (20 mM HEPES, pH 7.4, 200 mM NaCl, and 1 mM DTT). HPSE-containing fractions were concentrated to 10 mg/mL with a 30-kDa Vivaspin concentrator and were buffer-exchanged into hepar buffer A via at least three rounds of dilution/reconcentration.

Crystallization of HPSE. HPSE at 10 mg/mL was tested against a range of commercial crystallization screens. Large split crystals were found in several conditions of the index and PEG/ion screens, which were used for further optimization. Well-diffracting single crystals were grown using the sitting-drop vapor-diffusion method at 20 °C with 100 mM MES, pH 5.5, 100 mM MgCl2, 17% PEG 3350, and a protein/well solution ratio of 200:300 μL. Crystals typically appeared after 3 d. Apo-HPSE crystals were cryoprotected before flash freezing in liquid N2 for data collection, and the mother liquor solution was supplemented with 25% ethylene glycol.

Crystal-structure solution. The HPSE crystal structure was solved by single isomorphous replacement with anomalous scattering (SIRAS) with the apo crystal and an iodide derivative. The iodide derivative was prepared by a short soak (~10 s) of an apo crystal in cryoprotectant solution further supplemented with 0.5 M NaI. Data were collected at beamline I02 of the Diamond Light Source, processed with XDS, reduced with Aimless, and phased with the SHELX pipeline. Initial autobuilding was carried out with Buccaneer before subsequent rounds of manual model building and refinement with Coot and REFMACS (ref. 55), respectively. MolProbity and Privateer were used to assess model quality before deposition in the PDB. Crystal structure figures were generated with CCP4mg.

Substrate complexes. HPSE complexes were generated by soaking the apo crystal with M0H S00a, M04 S02a, and M09 S05a HepMers (Iduron) or dp4 heparin tetrasaccharide (Dextra Laboratories). In each case, a small piece of solid compound was directly dissolved in mother liquor containing the crystal. All soaks were carried out for 20 min, before direct flash freezing in liquid N2 for data collection without further cryoprotection. Ligand coordinates were built with JLigand.

HepMer-hydrolysis assays. 20-μL assay solutions comprising 50 nM HPSE added to a series of HepMer solutions in 40 mM NaOAc buffer, pH 5.0, were incubated at 37 °C for 20 min, before addition of NaOH and WST-1 dye (Santa Cruz Biotech) to a final volume of 40 μL and final concentrations of 0.1 M NaOH and 1 mM WST-1. Reactions were developed at 60 °C for 60 min, and the absorbance at 584 nm was measured in a 384-well microplate with a POLARStar Optima plate reader (BMG Labtech). A584 readings were quantified against a d-glucose standard set up on the same plate. For each compound, a no-enzyme control was used to account for nonenzymatic autohydrolysis of the pNP residue present on the HepMer. Data analysis and curve fitting were performed with SigmaPlot 12.5 (Systat Software).

84. Bieniossek, C., Imasaki, T., Takagi, Y. & Berger, I. MultiBac: expanding the research toolbox for multiprotein complexes. Trends Biochem. Sci. 37, 49–57 (2012).
85. Dauter, Z., Dauter, M. & Rajashankar, K.R. Novel approach to phasing proteins: derivatization by short cryo-soaking with halides. Acta Crystallogr. D Biol. Crystallogr. 56, 232–237 (2000).
86. Kabash, W. et al. Acta Crystallogr. D Biol. Crystallogr. 66, 125–132 (2010).
87. Evans, P.R. & Murshudov, G.N. How good are my data and what is the resolution? Acta Crystallogr. D Biol. Crystallogr. 69, 1204–1214 (2013).
88. Sheldrick, G.M. Experimental phasing with SHELXC/D/E: combining chain tracing with density modification. Acta Crystallogr. D Biol. Crystallogr. 66, 479–485 (2010).
89. Cowtan, K. The Buccaneer software for automated model building. 1. Tracing protein chains. Acta Crystallogr. D Biol. Crystallogr. 62, 1002–1011 (2006).
90. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. 60, 2126–2132 (2004).
91. Mushudov, G.N., Vagin, A.A. & Dodson, E.J. Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr. D Biol. Crystallogr. 53, 240–255 (1997).
92. Chen, V.B. et al. MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. D Biol. Crystallogr. 66, 12–21 (2010).
93. Agirre, J. et al. Privateer: software for the conformational validation of carbohydrate structures. Nat. Struct. Mol. Biol. 22, 833–834 (2015).
94. McNicholas, S., Potterton, E., Wilson, K.S. & Noble, M.E.M. Presenting your structures: the CCP4mg molecular-graphics software. Acta Crystallogr. D Biol. Crystallogr. 67, 386–394 (2011).
95. Lebedev, A.A. et al. JLigand: a graphical tool for the CCP4 template-restraint library. Acta Crystallogr. D Biol. Crystallogr. 68, 431–440 (2012).
Corrigendum: Structural characterization of human heparanase reveals insights into substrate recognition

Liang Wu, Cristina M Viola, Andrzej M Brzozowski & Gideon J Davies

Nat. Struct. Mol. Biol. 22, 1016–1022 (2015); published online 16 November 2015; corrected after print 8 December 2015

In the version of this article initially published, there were errors in two figures. The schematic in Figure 4a was mistakenly drawn with a β1→3 linkage between the −1 GlcUA and the +1 GlcNS(6S), rather than the correct β1→4 linkage. Supplementary Figure 1 incorrectly gave the names of the GlcNX monomers as N-acetyl-α-D-glucuronic acid and N-sulfo-α-D-glucuronic acid, rather than N-acetyl-α-D-glucosamine and N-sulfo-α-D-glucosamine. These errors have been corrected in the HTML and PDF versions of the article.