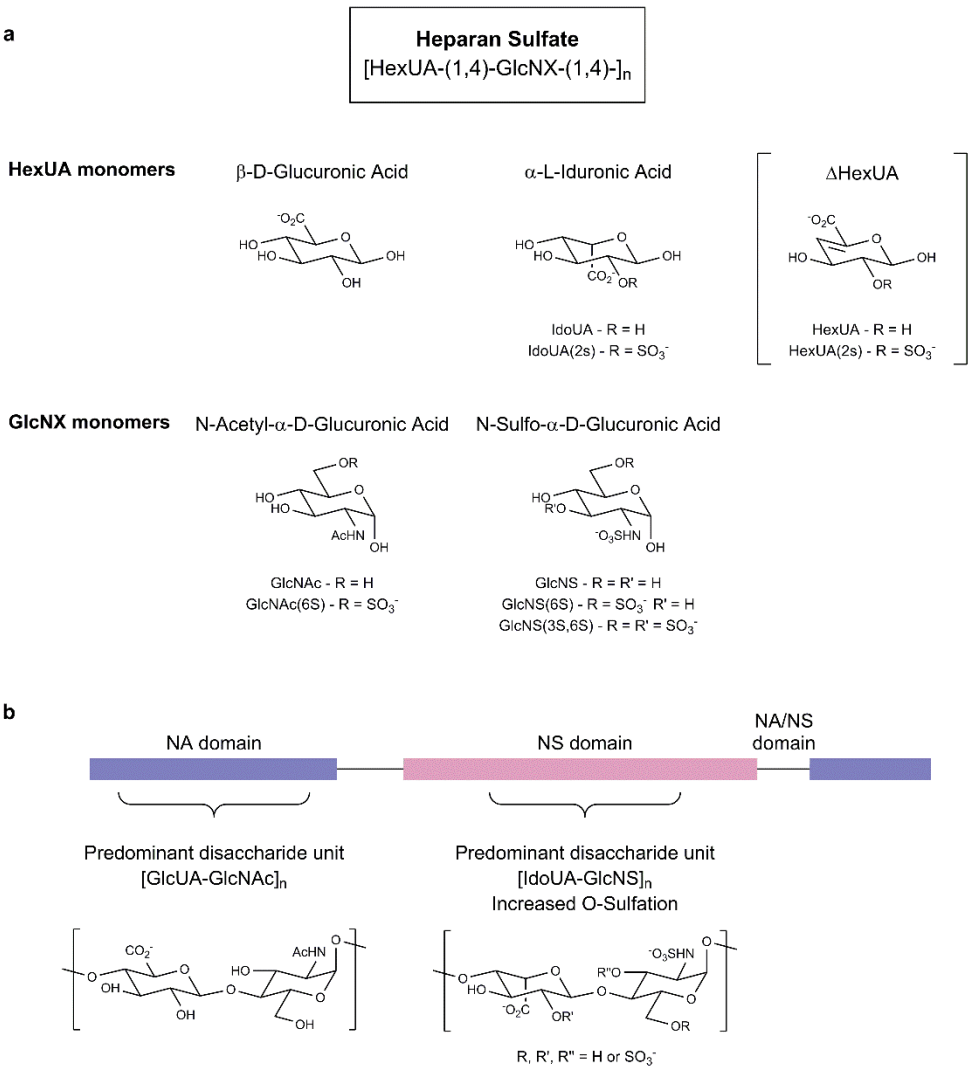
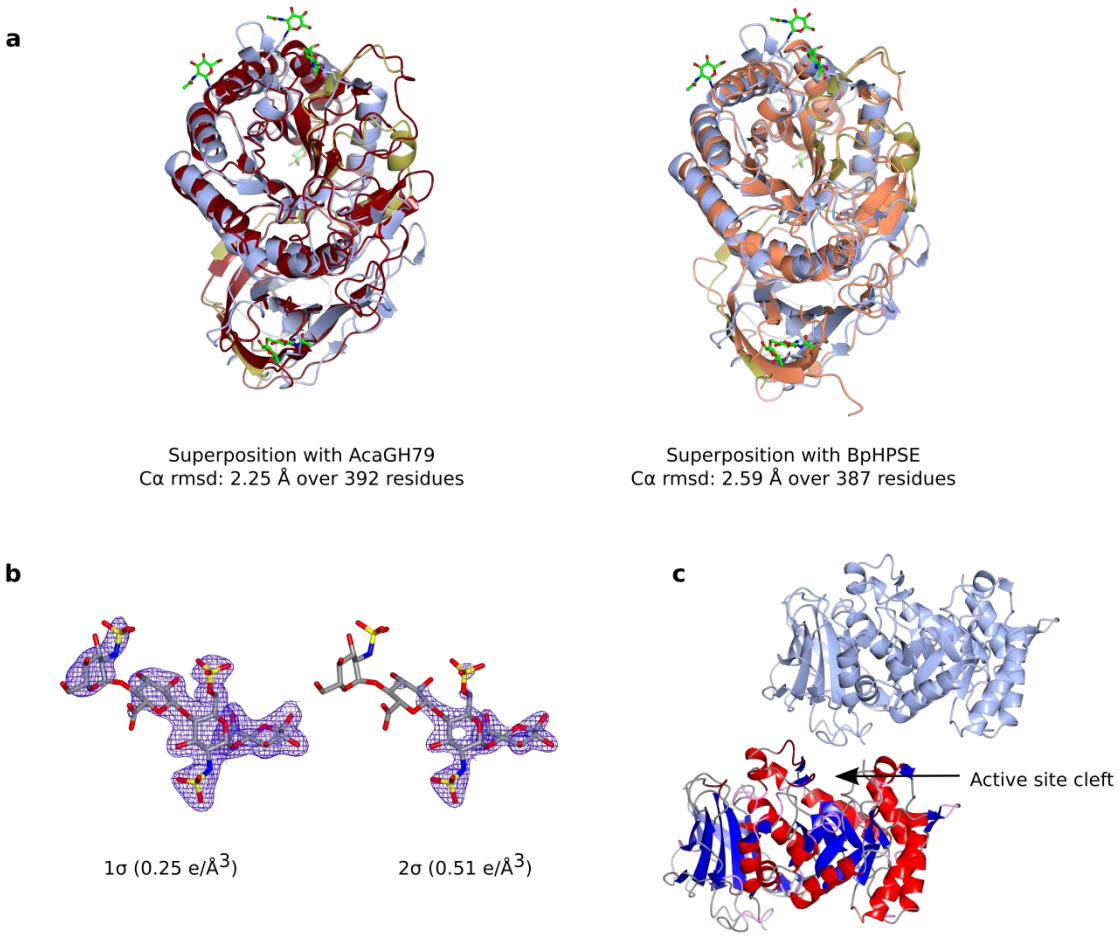
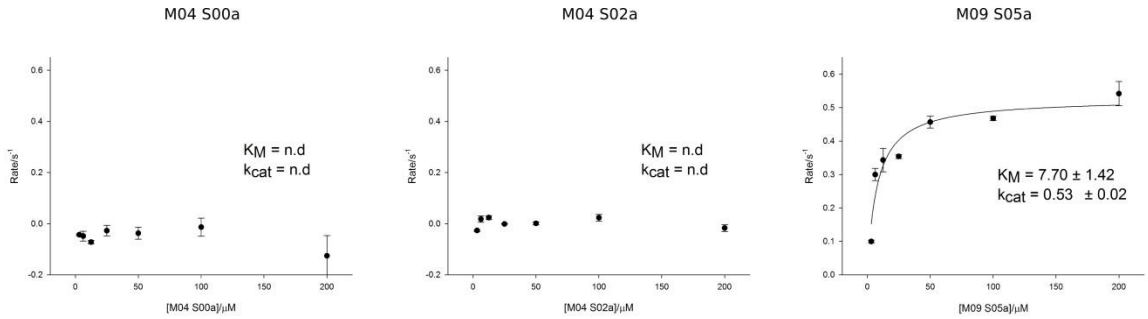


Supplementary Figure 1 Molecular and domain composition of HS. HS is a structurally heterogeneous linear polysaccharide composed of repeating HexUA-(1,4)-GlcNX-(1,4) disaccharide units. (**a**) Chemical structures of the constituent HS sugars. HexUA is either GlcUA or IdoUA, which differ only in stereochemistry of the 5 position. GlcNX is either GlcNAc or its N-deacetylated-N-sulfated congener GlcNS. IdoUA, GlcNAc and GlcNS are subject to varying degrees of O-sulfation. Also shown is the structure of ΔHexUA, an unsaturated uronic acid formed at the non-reducing end of heparins depolymerized by bacterial heparin lyases, such as the dp4 tetrasaccharide used in this study. (**b**) Macromolecular 'domain' like organization of mature HS. NA domains are rich in GlcUA and GlcNAc, and show low levels of O-sulfation. NS domains are enriched in IdoUA and GlcNS, with a higher degree of O-sulfation. NA and NS domains are bounded by mixed NA/NS domains, which show intermediate characteristics.



Supplementary Figure 2 Supplemental crystal structure illustrations (**a**) Ribbon representations of HPSE (blue and gold) superposed against the bacterial exoglucuronidase AcaGH79 (dark red; PDB accession code 3VNY), and bacterial endoglucuronidase BpHPSE (coral; PDB accession code 5BWI). The three GH79 proteins possess a high degree of similarity in their overall fold topology. (**b**) M09 S05a -2 subsite GlcNS(6S) with density contoured to 1σ (0.25 electrons/Å^3) and 2σ (0.51 electrons/Å^3). The relative weakness of 6O sulfate density compared to N-sulfate can be seen. Electron densities are REFMAC maximum-likelihood/$\sigma_{\text{Å}}$ weighted 2Fo−Fc syntheses. (**c**) View of HPSE (colored by secondary structure) along the active site cleft towards the 'positive' end, showing the presence of a symmetry molecule (ice blue) at the opening. This symmetry molecule prevents HPSE interactions *in crystallo* which involve substrates protruding too far out of the 'positive' end of the cleft.



Supplementary Figure 3 Michaelis-Menten kinetics for HPSE hydrolysis of the HepMers. Reaction was measured using the reducing end detection dye WST1. Baseline subtractions using a no enzyme control were carried out for all reactions, to control for non-enzymatic paranitrophenol autohydrolysis. Error bars are standard errors of the mean for technical replicates (n=3 for all points). N.d. stands for not determinable.

a

-1 Subsite

```
Homo_sapiens                FLSVTIDANLATDPRFLILLGSPKLRTLARGLSPAYLRFGGTKTDFLI   103
Pan_troglodytes             FLSVTIDANLATDPRFLIFLGSPKLRTLARGLSPAYLRFGGTKTDFLI   103
Mus_musculus                FLSITIDASLATDPRFLTFLGSPRLRALARGLSPAYLRFGGTKTDFLI   95
Danio_rerio                 FLSVAIDASLLTEEKFMNLLNSPKLRTLAKALTPAFLRFGGTKQDFLK   101
Burkholderia_pseudomallei   FAGLSIEKAALSYP-LLSGENGNMVGLFNRLG-AGVLRIGGNSSDASG   118
Acidobacterium_capsulatum   YTGLSYEQAQMANPNYFSGANTQLAGFLRTLGRQGVLRIGGNTSEYTF   86

Homo_sapiens                ISWELGNEPNSFLKK   232
Pan_troglodytes             ISWELGNEPNSFLKK   232
Mus_musculus                ISWELGNEPNSFWKK   224
Danio_rerio                 MSWELGNEPNSYEKK   230
Burkholderia_pseudomallei   AGFEIGNEPDLYAQH   194
Acidobacterium_capsulatum   LAFQLGNEPDLFYRN   180

Homo_sapiens                GKKVWLGETSSAYGGGAPLLSDTFAAGFMWLDKLGLSARMGIEVVMRQVFFGAGNYHLVD   395
Pan_troglodytes             GQKVWLGETSSAYGGGAPLLSDTFAAGFMWLDKLGLSARMGIEVVMRQVFFGAGNYHLVD   395
Mus_musculus                GKKVWLGETSSAYGGGAPLLSNTFAAGFMWLDKLGLSAQMGIEVVMRQVFFGAGNYHLVD   387
Danio_rerio                 GKKVWLGETSSAYGGGAVGLSDTFVAGFMWLDKLGLAAKLGLNLVIRQVLIGAGTYHLVD   394
Burkholderia_pseudomallei   GIGFRLAETNSYWGGGKPGVSDAHASALWVINFLFAVAQGGASG-VNLHTGGGASYSAIK   349
Acidobacterium_capsulatum   GLPFRLTETNSCYQGGKQGVSDTFAAALWAGDLMYQQAAAGSTG-INFHGGGYGWYTPVA   338
```

Gly349, Gly350 and Thr97 (Homo sapiens numbering) interact with ligands via backbone N-H

-2 Subsite

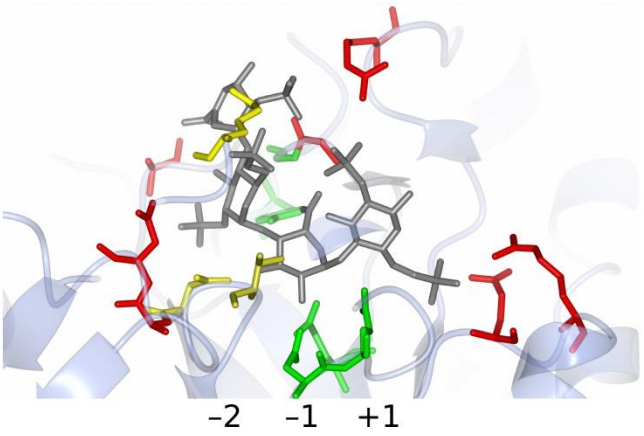
```
Homo_sapiens                VTIDANLATDP   69      Homo_sapiens                EHYQKKFKNS   163
Pan_troglodytes             VTIDANLATDP   69      Pan_troglodytes             EHYQKKFKNG   163
Mus_musculus                ITIDASLATDP   61      Mus_musculus                EQYQKEFKNS   155
Danio_rerio                 VAIDASLLTEE   67      Danio_rerio                 KELDGKYRNT   161
Burkholderia_pseudomallei   LSIEKAALSYP   86      Burkholderia_pseudomallei   ------ETSG   129
Acidobacterium_capsulatum   LSYEQAQMANP   52      Acidobacterium_capsulatum   ---HHAAARE   114

Homo_sapiens                FFGAGNYHLVD   395
Pan_troglodytes             FFGAGNYHLVD   395
Mus_musculus                FFGAGNYHLVD   387
Danio_rerio                 LIGAGTYHLVD   394
Burkholderia_pseudomallei   TGGGASYSAIK   349
Acidobacterium_capsulatum   GGGYGWYTPVA   338
```

Gly389 (Homo sapiens numbering) interacts with ligands via backbone N-H

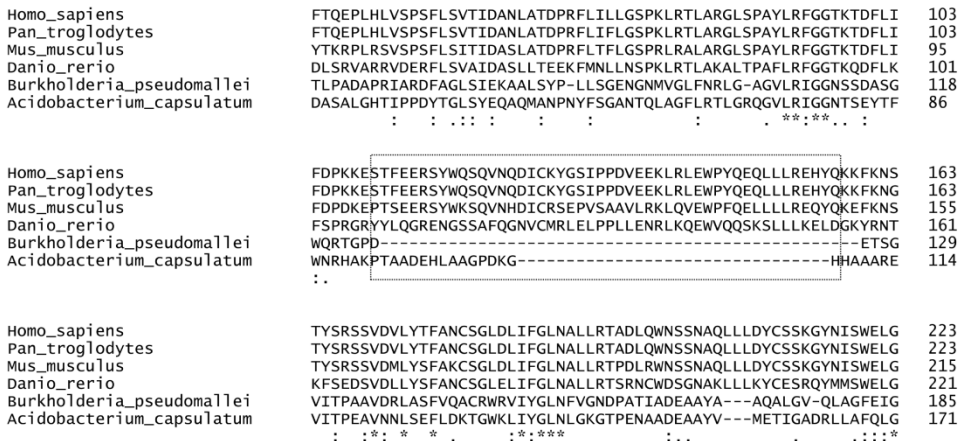
+1 Subsite

```
Homo_sapiens                YGPDVGQPRRKTAKMLKSFLKAGGEVIDSVTWHHYYLNGR----TAT   306
Pan_troglodytes             YGPDVGQPRRKTAKMLKSFLKAGGEVIDSVTWHHYYLNGR----TAT   306
Mus_musculus                YGPDIGQPRGKTVKLLRSFLKAGGEVIDSLTWHHYYLNGR----IAT   298
Danio_rerio                 YGPDVSQPKDHRKDLLTGFLETGGKVINACTWHHYYVNGR----DTS   305
Burkholderia_pseudomallei   TGPATAWNY---QRYTVPFASDAAGLVS-LLTQHHYRN----PDSAT   265
Acidobacterium_capsulatum   AGPDTAYNT----KWLVPFADKFKHDVK-FISSHYYAEGPPTDPSMT   254
```

b



Supplementary Figure 4 Relationship of HPSE active site residues to other GH79 enzymes. (**a**) Clustalω[1] alignment of 4 eukaryotic GH79 sequences, AcaGH79, and BpHPSE. Residues corresponding to those which interact with substrates in human HPSE (as outlined in **Figure 4**) are highlighted: green where identical with the human sequence, or orange where a non-conserved residue can interact in a similar fashion. Residues of the -1 subsite are hyperconserved, illustrating their requirement for interacting with GlcUA. In contrast, residues of the -2 and +1 subsites are only well conserved amongst the mammalian heparanases. (**b**) Active site view of the dp4-HPSE complex, with residues colored by conservation to bacterial enzymes AcaGH79 and BpHPSE (green – identical, yellow – partially conserved, red – not conserved). Dp4 ligand is shown in grey.



Supplementary Figure 5 Relationship of HSPE proenzyme linker sequence to other GH79 enzymes. Alignment of 4 eukaryotic GH79 sequences, AcaGH79, and BpHPSE, across the region corresponding to the proenzyme linker of human HPSE (dashed box). Eukaryotic GH79s show an extended sequence at this position, likely corresponding to a proteolytically cleavable linker as in human HPSE. The corresponding AcaGH79 sequence is shorter, and forms a loop which creates part of its *exo*-binding pocket. The corresponding BpHPSE sequence is effectively absent, and corresponds to a very short loop which reveals the *endo*- acting binding cleft of this enzyme.

Supplementary Table 1 Primers used for cloning HPSE constructs

8 kDa Subunit	
BamHI-Mellitin F primer	gatcggatccaccatgaaatttttggtg
Mellitin-XmaI R primer	gatcgcccgggtccgcataaatgtagctaatg
XmaI-HPSE(8 kDa) F primer	gatcgcccgggccaggacgtcgtggacctgg
HPSE(8 kDa)-PstI R primer	gatcctgcagtatcattccttcttgggatcgaaaattag
50 kDa Subunit	
XhoI-Mellitin F primer	gatcctcgagcaccatgaaatttttggtg
Mellitin-XmaI R primer	gatcgcccgggtccgcataaatgtagctaatg
XmaI-HPSE(50 kDa) F primer	gatcgcccgggcaaaaagttcaagaacagcacc
HPSE(50 kDa)-KpnI R primer	gatcggtacctatcagatgcaagcagcaac

References for supplementary materials

1. Sievers, F. et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology* **7**(2011).