Structural Analysis of Heparanase Inhibition

Cassidy A. Whitefield

January 2023

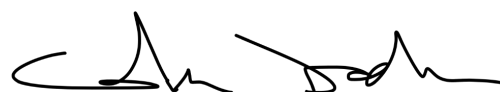*A thesis submitted for the degree of Doctor of Philosophy of The Australian National University*

Declaration

Unless otherwise stated, the work described in this thesis was carried out by the author between the dates of March 2019 and January 2023 under the supervision of Professor Colin Jackson at the Research School of Chemistry, Australian National University. This thesis contains no material which has been accepted for the award of any other degree or diploma in any university. To the best of the author's knowledge, it contains no material previously published or written by another person, except where due reference is made in the text.

However, this thesis contains does contain several co-authored papers published, submitted, or drafted for peer-review articles. The following outlines the status and contributions of Cassidy Whitefield to the research and authorship of each paper. Statements outlining the contributions of the authors are also provided at the beginning of each chapter.

Cassidy A. Whitefield

Professor Colin J. Jackson

(on behalf of all collaborating authors)

13 January 2023

13 January 2023

Chapter	Title	Authorship	Publication outlet and status	Overall contribution of CW
1	Computational design and experimental characterisation of a stable human heparanase variant	**Cassidy Whitefield***, Nansook Hong*, Joshua A. Mitchell, Colin J. Jackson	**RSC Chemical Biology** In press https://doi.org/10.1039/D1CB00239B	~50%
2	Complex inhibitory mechanism of glycomimetics with heparanase	**Cassidy Whitefield**, Yen Vo, Brett D. Schwartz, Caryn Hepburn, F. Hafna Ahmed, Hideki Onagi, Martin G. Banwell, Keats Nelms, Lara R. Malins, and Colin J. Jackson	**ACS Biochemistry** In press https://doi.org/10.1021/acs.biochem.3c00038	~85%
3	Allosteric small molecule inhibitors of human heparanase	**Cassidy Whitefield**, Junming He, Brett D. Schwartz, Ben Clifton, Nansook Hong, F. Hafna Ahmed, Elaaf Mohamed, Keats Nelms, Martin G. Banwell, Lara R. Malins, Colin J. Jackson	Manuscript Draft	~80%

*These authors contributed equally

Acknowledgements

To my family who supported me to follow my passion, I thank you. I also thank my friends that have supported me along the way; Mitchell Blyth, Kimberly Dunbar, Andie Delaney, Michael Thomas, Alex and Adam Barlow, Erin and Micah McDonald and the many others i haven't mentioned. Thanks to all my colleagues in the Jackson Lab (past and present) and the rest of the Research School of Chemistry.

I would also like to thank my PhD supervisory panel for helpful feedback. A special thanks to Mark Ellison for his support throughout my teaching endeavours. I would also like to thank all the students that I have had the pleasure of teaching and learning from during my time at the RSC.

I would like to express my sincere gratitude to my supervisor, Professor Colin Jackson, for his invaluable guidance and support throughout my PhD journey. I am very grateful for his optimism, patience, advice, encouragement and support during my time at ANU. I am thankful for the wonderful opportunities that Colin has given me to develop, both as a scientist and as a person, and I thank him for opening my eyes to the wonderful world of structural biology.

Abstract

Heparanase (HPSE) is the only known mammalian endo-β-glucuronidase that catalyzes the degradation of heparan sulfate, a major component of the extracellular matrix in animals. The overactivation of heparanase is associated with larger tumor size, increased metastasis, and poor prognosis, with all cancers to date showing elevated heparanase expression. Its enzymatic activity of cleaving heparan sulfate is well known, but research now shows that heparanase is a multifunctional protein with both enzymatic and non-enzymatic activities inside and outside the cell. Because of its importance in disease, heparanase has been the target of numerous therapeutic programs over the last two decades, with selected substrate mimetics currently being studied in clinical trials. However, despite tremendous efforts, no therapy with high clinical benefit and minimal side effects has been identified. The limited structural and kinetic knowledge of the various heparanase therapeutic classes has contributed to this issue. Access to easily produced protein has also contributed to this knowledge gap due to the complex processing path of heparanase in the body. In summary, this thesis focuses on the use of biophysical and structural analysis to develop and understand new heparanase therapeutics. Within this scope, our primary focus is on developing and understanding different classes of inhibitors, heparan sulfate mimetics, and small molecule inhibitors, each of which has its own therapeutic benefits and limitations.

Firstly, I demonstrate the computational redesign of heparanase to allow for high yield expression in *Escherichia coli*. This mutated form of heparanase exhibits essentially identical kinetics and inhibition. X-ray crystallography and molecular dynamics show that the heparanase variant has the same structure and dynamics as the wild-type protein. This variant is then used to understand the interactions of the therapeutics studied in subsequent chapters.

In the following chapter, I investigate pentosan polysulfate sodium (PPS), an FDA-approved inhibitor for interstitial cystitis and a known heparanase inhibitor. However, due to its heterogeneity, the mechanism of heparanase inhibition by pentosan has been limited in characterization. Three synthetic analogues of PPS showed that the mechanism of PPS inhibition is complex, with oligosaccaride length contributing to multiple inhibitor events. ITC and protein crystallography were used to determine that PPS binds at three binding domains on the surface of heparanase. Kinetic and structural studies support that this binding mechanism causes inhibitor-induced

aggregation and irreversible inhibition. This binding model could also be applied to other heparanase therapeutics currently in clinical trials.

Lastly, I study a new series of small molecule inhibitors. Small molecules can overcome many of the limitations of other therapeutics due to their increased bioavailability and specificity. A large structure-activity relationship for quinazoline-like compounds was completed for heparanase inhibition. Kinetics support that these inhibitors are non-competitive inhibitors, suggesting a potential binding site on heparanase. Despite unsuccessful crystallography studies, a broad screen of heparanase mutations was conducted to try to identify this binding site, but it was unsuccessful. This mutagenesis screen did demonstrate the importance of allosteric effects on heparanase. These studies highlight the importance of structural analysis of enzyme inhibition and the importance of designing and analyzing different therapeutics in order to design ideal inhibitors in the future.

List of Abbreviations

ΔG   Change in Gibbs free energy

ΔS   Change in entropy

ΔT   Change in enthalpy

k_{cat}/K_M Turnover number/substrate concentration at half maximum velocity

E. coli   Escherichia coli

K_M Substrate concentration at half maximum velocity

K_a Binding association constant

k_{cat} Turnover number

CD   Circular Dichroism

DDT   Dichlorodiphenyltrichloroethane

ECM   Extracellular matrix

ECM   Extracellular matrix

HBD   Heparan binding domain

HDX-MS Hydrogen Deuterium exchange Mass spectrometry

HEPES   4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

HPSE   Heparanase

HPSE P6 Heparanase pross 6 variant

HPSG   Heparan sulfate proteoglycan

HS   Heparan sulfate

HTS   High throughput screen

IC_{50} The half maximal inhibitory concentration

ITC   Isothermal titration calorimetry

MD   Molecular dynamics

PCA   Principle component analysis

PCR   Polymerase chain reaction

PDB   Protein Data Bank

PPS   Pentosan Polysulfate Sodium

PROSS Protein Repair One Stop Shop

RMSD   Root mean square deviation

RMSF   Root mean square fluctuation

SAR   Structure-Activity Relationship

SD   Standard deviation

SEM   Standard error from the mean

SS   Sum of squares

T_m Thermal stability

TCEP   tris(2-carboxyethyl)phosphine

WST-1 4-[3-(4-iodophenyl)-2-(4-nitrophenyl)-2H-5-tetrazolio]-1,3-benzene disulfonate

WT   Wild Type

Contents

For family

Chapter 1

Research Context

1.1 Heparan Sulfate Proteoglycans

Heparan sulfate proteoglycans (HSPGs) are a class of proteins commonly found in the extracellular matrix (ECM) of animals. These glycoproteins consist of a core protein covalently linked to one or more heparan sulfate (HS) chains. There are five main classes of HSPGs: syndecans, glypicans, perlecan, agrin and collagen type XVIII. HS chains are always present on the surface of these proteins, but other proteins such as CD44, betaglycan and testican can also have varying levels of HS expression (*1*). HSPGs can be found in three different locations: membrane HSPGs, which are attached to a transmembrane domain (syndecans) or to a glycosylphosphatidylinositol-anchored core linkage (glypicans); secreted extracellular matrix HSPGs, such as perlecan, agrin and collagen type XVIII, which are found in basement membranes (BMs) and the extracellular matrix (ECM); and secretory vesicle proteoglycan serglycin, which is the only intracellular proteoglycan (*2*). HSPGs are structurally diverse molecules due to the extensive chemical heterogeneity of the HS chains. This heterogeneity allows HS chains to bind to a wide range of bioactive molecules, including cytokines, chemokines, growth factors, cell adhesion proteins and proteases, providing a low-affinity storage depot for heparin/heparan binding proteins (*3*).

1.2 Heparan Sulfate

Heparan sulfate consists of repeating disaccharide units formed by a negatively charged uronic acid (D-glucuronic acid, GlcA, or L-iduronic acid, IdoA) and a glucosamine (N-acetylated, GlcNAc, or N-sulfated, GlcNSO$_3$). The intra-disaccharide glycosidic linkage is α (1→4) or β (1→4), whereas the inter-disaccharide linkages are always α (1→4) (*4*). It is worth noting that the intra-glycosiduronic linkage in HS possesses identical absolute configuration and regiochemistry (1-4). However, due to the complex and specialized naming conventions in carbohydrate chemistry, the terminology used for these linkages differs. Specifically, D-glucuronic acid is referred to as "beta," while L-iduronic acid is referred to as "alpha." Biosynthesis of HS is non-templated, which allows the composition of HS to vary substantially. This can include chain length variation between 40-300 sugar residues (20-150 nm) and the presence of regions of high or low sulfation (*5*).

HSPG core proteins are modified and folded in the endoplasmic reticulum to achieve their proper or mature conformation, then HS chains are assembled directly onto the core proteins in the Golgi apparatus via sequential polymerases and modifying enzymes (*6*). HS chains are initiated at a serine residue in the core protein and further modification of the HS chains includes sulfation at the N-, 3-O, or 6-O

position of N-acetylamine or the 2-O position of uronic acid (*7*). Importantly, the extent and arrangement of sulfation promotes the functional diversity of HS chains (*8*). This structural heterogeneity is crucial for HS function, enabling a single polysaccharide chain to interact with a variety of binding partners, including growth factors, chemokines, cytokines and enzymes with low-affinity binding, providing a storage depot. HS fragments are also capable of modulating the activity of other enzymes and growth factors (*9*). Therefore, the regulation of HS chains is important for biological function.
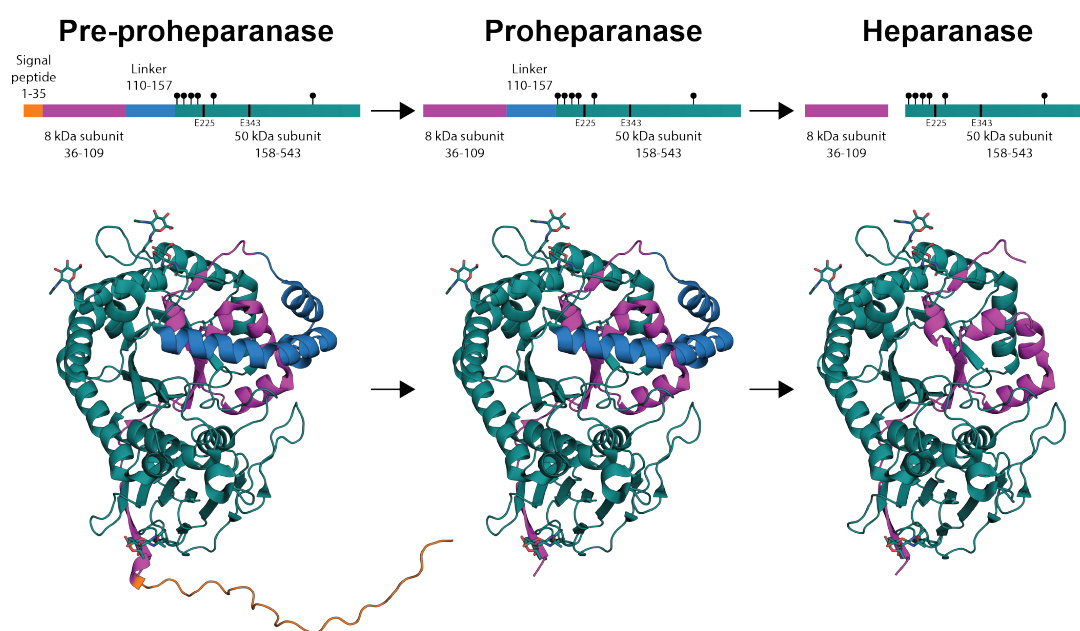
1.3   Heparanase

Heparanase is the only known endo-β (1,4)- D-glucuronidase capable of cleaving HS chains (*10*). It is a member of the Carbohydrate Active Enzymes (CAZy) Glycoside Hydrolase GH79 family of carbohydrate processing enzymes, characterised by the $(\beta/\alpha)_8$ domain containing the catalytic site of the enzyme. Heparanase splice variants will not be discussed in this thesis.

1.3.1   Structure and processing

Heparanase is synthesised as a 68 kDa precursor (pre-proheparanase), which consists of an N-terminal signal peptide (Met1-Ala35), a C-terminal hydrophobic peptide (Pro515-Ile534), 5 cysteine residues and 6 N-glycosylation sites (*11*). Pre-proheparanase is transported to the endoplasmic reticulum (ER) for glycosylation and cleavage of the signaling peptide to form 65 kDa inactive proheparanase 1.1. Proheparanase is then directed to the Golgi apparatus, where it is packaged into vesicles and secreted (*12*). Extracellular proheparanase interacts with cell surface HSPGs, low-density lipoprotein-receptor related protein and certain receptors. The inactive enzyme can be taken up by certain cells, where it undergoes pH-dependent proteolytic cleavage by the cysteine protease cathepsin L, resulting in the production of catalytically active heparanase in the form of a 58 kDa heterodimer (*13*).

Active heparanase has residues Gln36-Glu109 of the 8 kDa subunit and Lys159-Ile543 of the 50 kDa subunit (numbering based on the full preproenzyme) (*13*). The domain architecture of heparanase consists of a $(\beta/\alpha)_8$ domain flanked by a smaller beta-sandwich domain 1.1. Both the 8 kDa and 50 kDa subunits are structurally involved in both domains: the 8 kDa subunit contributes one β-strand to the β-sandwich and the first β-α-β fold of the $(\beta/\alpha)_8$ domain, with the remaining folds contributed by the 50 kDa subunit (*14*, *15*). Heparanase also contains three heparan

binding domains (HBDs) which were identified through sequence analysis (HBD-1: Lys158-Asp171; HBD-2: Gln270-Lys280; HBD-3: Lys411-Lys417, Lys427-Arg432) (*16*). These domains are involved in the interaction with the negatively charged HS chains and are crucial for activity. HBD-1 is located on the N-terminal of the 50 kDa chain and is fundamental in heparanase's ability to recognize HS. HBD-2 comprises an alpha-helix of the TIM barrel domain and has been shown to have a lower affinity than HBD-1, but it can also affect recognition. HBD-3 is on the β-sandwich domain of heparanase and has been suggested to play a role in protein stability or anti-coagulation activities, but not in HS binding (*16*, *17*)

Figure 1.1: A schematic representation of the heparanase domain structure and the processing steps to produce the active form. Heparanase is synthesised as the pre-proheparanase form (Structure generated from Alphafold, sugars modelled from PDB: 5LA4), which is processed into the proheparanase form upon the removal of the signal peptide (PDB: 5LA4). Proheparanase is then processed *via* cathapsin L to remove the linker peptide, producing the active heterodimeric enzyme (PDB: 5E8M) with the small 8 kDa subunit (residues 36-109) and a large 50 kDa subunit (residues 158-543). The location of the six N-linked glycosylation sites are indicated by solid circles and the catalytic residues are indicated by black lines.

Heparanase contains six N-glycosylation sites on the 50 kDa subunit, located at residues Asn162, Asn200, Asn217, Asn238, Asn459 and Asn178 (*18*). These glycosylation sites are important for the transport and secretion of heparanase, but have no involvement in catalytic activity(*19*). The substrate binding cleft of heparanase spans approximately 10 Å across the $(\beta/\alpha)_8$ domain of heparanase and contains the catalytic residues Glu343 (catalytic nucleophile) and Glu225 (acid-base) (*20*).The cleft is lined with basic side chains contributed by Arg35, Lys158, Lys159, Lys161, Lys231,

Arg272, Arg273 and Arg303, which helps to accommodate the negative charge of the enzyme's HS substrate (*18*).

The catalytic activity of active heparanase occurs within a pH range of 5 to 6 and has a half-life of approximately 30 hours (*21*). Upon activation, the enzyme is stored in lysosomes, which can then move to the cell surface for release into the extracellular matrix (ECM). Inflammatory cytokines can also stimulate the release of heparanase from endothelial cells and peripheral T lymphocytes (*22*).

1.3.2 Substrate Recognition

Heparanase exhibits a high degree of specificity in its interaction with HS structures, with only certain sites being susceptible to enzymatic attack. Substrate recognition is strongly influenced by the sulfation patterns of HS chains and the enzyme preferentially cleaves specific sulfation patterns. The recognition sequence is three oligosaccharides in length and favors cleavage between a glucuronic acid linked to a 6O-sulfated glucosamine that may be N-sulfated or N-acetylated (*23*). When acetylated 6O-sulfated glucosamine is present, heparanase will cleave every bond, but when sulfated 6O-sulfated glucosamine is present, heparanase will follow a gapped cleavage pattern, cleaving every second bond(*23*, *24*). 2O-sulfated glucuronic acid and 2O-sulfated iduronic acid cannot be part of the cleavage recognition site due to steric clashes with Asn224 in the active site and the bulky sulfate group. In general, iduronic acid residues cannot attain the necessary conformation to be in the active site (*18*, *25*).
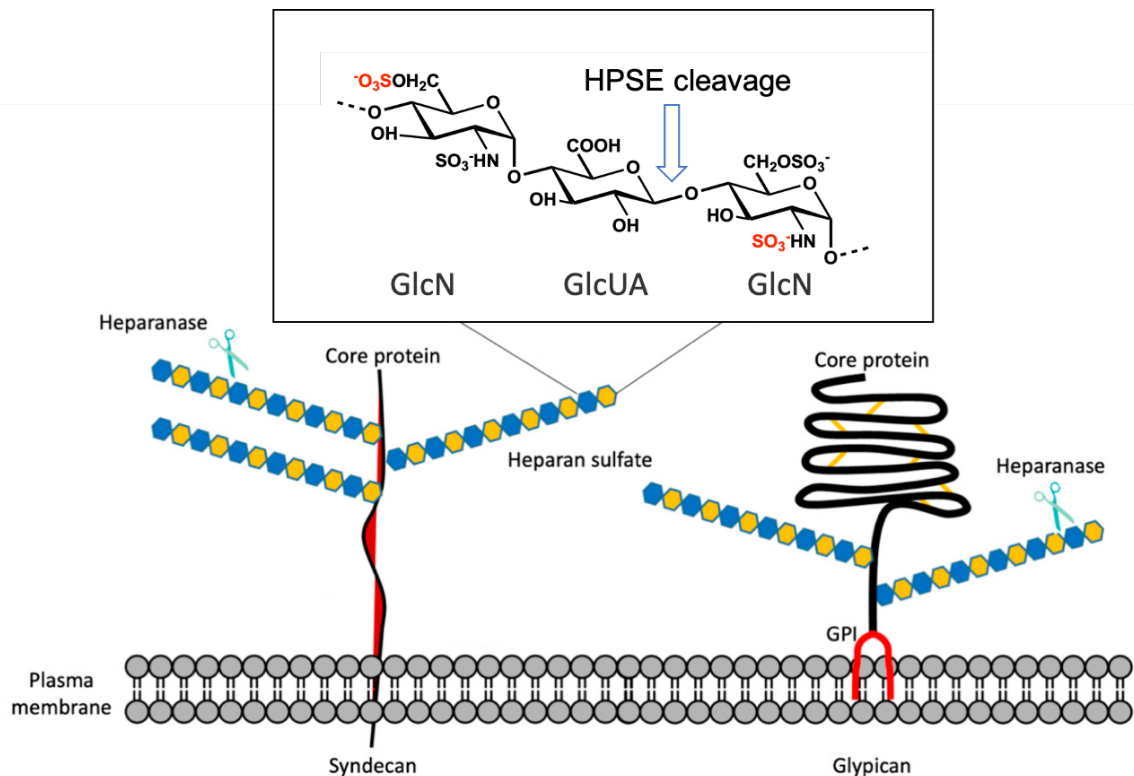
Figure 1.2: Schmatic depiction of HS chains on HSPGs of syndecan and glypican. The minimum trisaccharide sequence for heparanase recognition and cleavage is highlighted, with the arrow indicating the glucuronidic linkage cleaved by heparanase. The SO_3 on the reducing side GlcN(2-*N*-sulfate) and the GlcN(6-*O*-sulfate) on the non-reducing side of the cleavage site are considerably important for the substrate recognition. Figure adapted from (*26*)

1.4 Normal Physiological Role of Heparanase

Heparanase plays important roles in both normal physiological processes and in cancer development. Heparanase can exist in both an inactive proheparanase form and an enzymatically active heparanase form and both forms have many important functions in the body. In normal physiological conditions, heparanase is highly expressed in haematopoietic cells such as platelets (*27*), monocytes (*28*), macrophages (*29*, *30*), neutrophils (*31*), mast cells (*32*), dendritic cells (*28*), activated lymphocytes (*33*) and eosinophils (*34*), Lower levels of expression are also observed in the lymph node, bone marrow, spleen, liver and thymus (*35*, *36*).

Intracellular heparanase is predominantly located in lysosomes and late endosomes, where the acidic environment promotes its catalytic activity and allows it to perform a housekeeping role through the processing of internalized HSPGs (*37*). In lysosomes, heparanase modulates autophagy through autophagosomes and participates in normal lysosomal activity (*38*). Endosomes allow cellular communication by

transporting materials around the cell. Within these cells, heparanase trims HS or heparin from various HSPGs, such as syndecan and serglycan (*39*). Heparanase's enzymatic activity also results in the cleavage of syndecan-1 HS chains, allowing clustering of syndecan-1 and increasing the production of exosomes, which enhances cell-cell communication (*40*). Heparanase can also regulate gene expression through its enzymatic interactions with nuclear syndecan-1 and HS, leading to an increase in growth-related genes and directly binding to gene promoters (*41*).
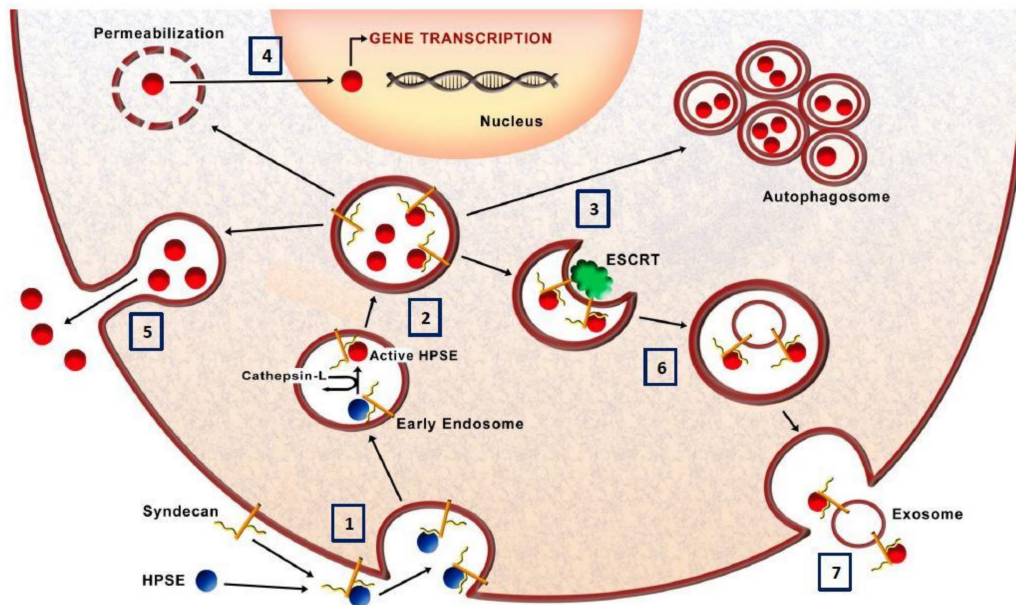
Figure 1.3: Schematic model of heparanase trafficking and biological activities."**1.** Inactive proheparanase can interact with HSPGs such as syndecan-1 in the extracellular space where the complex is endocytosed to form endosomes. **2.** These endosomes can fuse with lysosomes, resulting in acidification, where heparanase can be activated through cleavage by cathepsin-L. **3.** From here, heparanase can participate in the formation of autophagosomes, thereby controlling the basal levels of autophagy. **4.** Heparanase can also translocate into the nucleus, where it can modulate gene transcription. **5.** Active heparanase can be secreted into the extracellular space. **6.** In addition, heparanase modulates the formation and release of exosomes where **7.** heparanase can be released and anchored to syndecan on exosome surfaces. Figure obtained from (*42*)

Heparanase plays a key role in the immune system, where it has numerous functions due to the rapid growth and movement required for an immune response (*43*, *44*). It can regulate gene regulation, differentiation and migration/invasion of various immune cells and can enter the nucleus of activated T lymphocytes to regulate immune genes (*45*). Heparanase is also essential for the activation and function of macrophages (*29*) and has been shown to be important in dendritic and natural killer cells, where it enables increased migration, which is crucial for the immune response (*46*). Heparanase is seen in mast cells, where it regulates the

production of anti-coagulation compound, heparin. Heparanase is also vital for neutrophil granulocyte function and leukocyte recruitment to sites of inflammation (*44*, *47*).

Extracellular heparanase exists in its enzymatically active form, where it contributes to the degradation of HSPGs through the cleavage of HS (*48*). The breakdown of HSPGs affects the structure of basal membranes and the ECM, releasing a pool of HS-bound proteins into the environment. The remodeling of the ECM, along with the diffusion of cytokines, growth factors and lipoproteins, facilitates cell mobility, angiogenesis, inflammatory responses and coagulation (*9*, *49*, *50*).

1.5 Heparanase Involvement in Cancer

The increased activity of heparanase has been observed in a wide range of pathologies, including cancer, inflammatory diseases and viral infections. Much of the current knowledge about heparanase centers on its involvement in cancer and there are many comprehensive reviews on the topic (*9*, *51–53*). RNA and protein levels of heparanase are elevated in almost all cancers studied to date, including ovarian, pancreatic, myeloma, colon, bladder, brain, prostate, breast, liver and rhabdomyosarcoma cancers(*54–64*). Clinical studies have demonstrated that increased heparanase expression correlates with increased tumour size, tumour progression, metastasis and poor prognosis (*49*, *65*). Studies using knockdown approaches have also shown that decreasing heparanase expression can decrease tumour progression (*29*, *66*)

During cancer progression, heparanase is one of many enzymes that act on the extracellular matrix and basement membrane surrounding a primary tumor. This activity weakens the structures and so facilitate tumor cell invasion into the surrounding tissues and assist metastasis (*65*). The breaking down of HS chains in the ECM also releases a flurry of proteins allowing for metastasis and angiogenesis to occur (*2*).

Heparanase is also located in the nucleus, where elevated nuclear heparanase has been shown to contribute to disease in some cases. In glioma and breast carcinoma, nuclear heparanase has been observed to degrade nuclear HS (*41*). In mesenchymal tumours, heparanase co-localizes with syndecan-1 (*67*). While nuclear localization of heparanase is often seen as an indicator of poor prognosis in certain cancers, in some cases such as squamous cell carcinomas, it has been associated with a good outcome (*68*). Other studies suggest that heparanase can bind DNA and/or chromatin and modulate gene transcription (*69*, *70*)

In recent years, the connection between heparanase and chemotherapy resistance has become increasingly clear. It has been observed that during chemotherapy treatments, there is an elevated level of heparanase and heparanase has been found at an elevated level within exosomes (*71*). These exosomes interact with macrophages, leading to an increase in macrophage migration and the secretion of growth-promoting cytokines (*72*). Latent heparanase is located on the surface of these exosomes and is readily taken up by tumour cells, activating in the acidic environment. This increased ability to alter the tumour microenvironment through the degradation of the ECM can lead to disease progression and chemotherapy resistance (*38*, *71*).

1.5.1   Heparanase-2

Heparanase-2 is a homolog closely related to heparanase, sharing 40% sequence identity (*73*). Despite their structural similarities, heparanase-2 lacks the proteolytic cleavage of the linker region, which renders it incapable of cleaving HS (*73*). However, heparanase-2 exhibits a stronger affinity for HS compared to heparanase, even though it lacks enzymatic activity. This high-affinity binding enables heparanase-2 to competitively outcompete heparanase for HS binding sites, thus functioning as an endogenous inhibitor of heparanase (*74*). Although heparanase-2's role in cancer is less extensively studied, findings indicate that it exerts an opposing effect to heparanase by reducing tumor growth and offering improved prognoses (*75*, *76*). Vlodavsky *et al.*, have provided an excellent review on the distinct roles of heparanase-1 and -2 in cancer, offering further valuable insights into this topic (*77*).

1.6   Therapeutic Options

Since the discovery that heparanase plays a role in metastasis, there has been a search for heparanase inhibitors. This field has continued to grow since it was discovered that there is only one heparanase gene in the human genome. Several types of heparanase inhibitors are in development, including analogues of HS, synthetically produced small molecule compounds(*78–80*); nucleic acid-based drugs (*81*), vaccines (*82*), monoclonal antibodies (*83*), proteins (*34*) and natural products (*84*). To date, only four heparanase drugs have reached clinical trials: PI-88(*85*) SST0001 (*86*), M-402 (*87*) and PG545(*88*). These drugs are all polysaccharide mimetics of the native substrate.

1.6.1   Heparanase as a potential drug target

Heparanase plays multiple roles in cancer pathology and its increased expression has been found to be strongly correlated with larger tumor size and a poorer prognosis.

Research has also shown that heparanase knockouts can significantly reduce cancer progression with few side effects. Heparanase also has roles in a variety of other pathological conditions including inflammation, viral infections and cardiovascular disease. (*66*, *89*). Before developing drugs targeting heparanase, it is important to validate the target to increase the likelihood of success in drug development programs

Heparanase is an enzyme that has unique catalytic activity not found in any other protein in the human body. This makes it a promising target for drug development because its binding specificity and conserved active site make it easier to develop drugs that can bind to and inhibit it. In 2015, the heparanase crystal structure was determined, which has helped in the development of heparanase therapeutics (*18*).To ensure the success of heparanase inhibitor development, it is important to have reliable and easily accessible assays to test the activity of potential therapeutics on a larger scale. There are various heparanase assays that have been developed using both cellular and *in vitro* methods, which allow for the efficient testing of new inhibitors (*90*).

Heparanase expression is largely limited to specific tissues in the body, including the placenta, lymphoid organs, leukocytes, platelets, keratinocytes and endothelial cells.(*91*). This restricted expression means that drugs targeting heparanase may have fewer off-target effects on other tissues and systems in the body. Heparanase knockout mice do not have severe abnormalities, but do exhibit impaired wound repair and reduced immunity due to heparanase's role in cell self-renewal and angiogenesis (*92*). While heparanase is required at low levels for many physiological processes, drugs that target heparanase are not always 100% effective and the effects of heparanase knockouts may not be fully replicated with heparanase inhibitors for cancer treatment. Despite the potential side effects of anti-cancer therapeutics, heparanase has been considered a promising target due to its effects being generally tolerable. However, it has been challenging to design an effective therapeutic for heparanase.

1.6.2 Why heparanase is a challenging drug target

Heparanase is a multi-functional protein, with different regions participating in various biological activities. Both its enzymatic and non-enzymatic activities are thought to contribute to cancer pathology. The splice variant T5, for example, has been shown to contribute to the tumor environment even though it lacks HS binding ability (*93*). The enzymatic activity of heparanase, which involves cleaving HS,

enables cell invasion and metastasis, while its non-enzymatic activities, such as anti-coagulation and gene regulation, do not require enzymatic activity (*70*, *94*). While the HS binding cleft is involved in its catalytic activity, it is not yet clear which regions of heparanase are involved in its multiple non-enzymatic activities. Recent studies on heparanase's anti-coagulation activity have identified that the region around HBD-3 (Ala421-Thr438) is important, but it is unknown if this is the only region involved in heparanase's non-enzymatic roles (*17*).

The drugs discussed in this thesis are designed to inhibit heparanase's catalytic activity, but it is not clear whether they can effectively reduce disease because non-active variants of heparanase also contribute to the disease state. The most commonly studied type of heparanase inhibitor is the HS-mimetic, a sulfated oligosaccharide with a large structural variation. These inhibitors are also selected for development because they have the ability to inhibit other HS-interacting proteins involved in disease, making it challenging to understand how they interact with heparanase and how heparanase plays a role in cancer progression.

A another major challenge in this field is the difficulty of assaying the enzyme. Many assays have been developed over the years to help identify new inhibitors more effectively, but no protocol has emerged as the gold standard (*90*). This makes it difficult to compare results from different assays, which is why activities of compounds in this introduction are presented in different ways and may not be directly comparable (μg/mL vs μM). Most assays used today use colorimetric or fluorescence detection methods with substrates that have single cleavage sites. The method used in this research is from Hammond et al, 2010 (*95*) and is a simple assay using the synthetic substrate fondaparinux with one cleavage site. This assay is currently one of the most widely used for *in vitro* studies and allows for kinetic analysis, which is rarely undertaken for heparanase inhibitors. Despite its flexibility, this assay still has many limitations as it is only an endpoint experiment, meaning that accurate activities and rates cannot be determined.

Below, I discuss select inhibitors of heparanase from a large field of compounds that have been designed in the last few decades. For comprehensive reviews on inhibitors of heparanase, please refer to reviews on the topic (*96*, *97*).

1.6.3   Modified Heparins and Sulfated Polysaccharides

The search for heparanase inhibitors began with heparin, an oligosaccharide structurally related to HS that showed significant heparanase inhibition. Due to the structural similarities between HS and heparin, heparanase can cleave heparin as it acts as a

substrate, even though it also inhibits HS binding. However, heparin's anticoagulant activity and ability to displace HS-bound growth factors from their storage within HS networks made it difficult to use as a pharmacological inhibitor. As a result, research focused on developing analogues with reduced side effects, such as anticoagulation.

After large structure-activity relationship studies about the binding recognition and structural variability of heparanase were conducted, two modified heparins of particular interest were SST0001 and M402, both of which entered clinical trials. **SST0001**, also known as (Roneparstat), is a 100 % N-acetylated, 25 % glycol-split heparin-like derivative with an average molecular size of 20 kDa (*98*). Its N-acetylation and glycol-split structure significantly affect its coagulation effects by cleaving the C2-C3 bond of the GlcA oligosaccharide. It is one of the most potent heparanase inhibitors, with an IC_{50} of 3 nM (*99*). SST0001 has also been shown to interact with other HS-binding proteins, such as VEGF and HGF, which are involved in cancer diagnosis. Preclinical trials supported its potential as an antitumor, antiangiogenic, immunomodulatory and antimetastatic agent. Phase I trials showed limited side effects, but also had minimal efficacy, suggesting that it may not be effective as a standalone treatment (*100*).

M402, also known as Necuparanib, is a rationally designed low molecular weight heparin with an average molecular size of 6 kDa. It is obtained through nitrous acid controlled depolymerization of heparin, which creates an N-sulfated structure followed by glycol splitting. These changes reduced the anticoagulation activity of M402. In addition to inhibiting heparanase with an IC_{50} of 5 μg/mL, M402 also showed the ability to interfere with other HS-binding proteins (*87*). *In vivo* results showed positive inhibition and clinical trials were conducted to evaluate the use of M402 in combination with standard anticancer therapy to treat breast and pancreatic cancer. However, studies were halted in Phase II trials due to insufficient efficacy (*101*).

Many other modified heparins and sulfated polysaccharides have been studied and many of them show activity against heparanase. These include δ-carrageenan, fucoidan, pentosan polysulfate, dextran sulfate, laminarin sulfate and SCM-chitin III, a highly sulfated semisynthetic chitin derivative (*85*, *102*, *103*). However, the structural variation and anti-coagulation effects of sulfated polysaccharides continue to present challenges for drug development programs investigating these types of heparanase inhibitors .

1.6.4 Sulfated Oligosaccharides

Oligosaccharide mimetics are highly sulfated, synthetic or semisynthetic compounds that function as non-hydrolyzable analogues of heparan sulfate glycosaminoglycans (HS GAGs). These mimetics have shown potential as anti-cancer agents due to their ability to inhibit heparanase with high potency (*104*). Synthetic chemistry has been used to diversify existing oligosaccharides, allowing the size and functional groups of these molecules to be varied, optimizing them for various therapeutic uses (*88*).
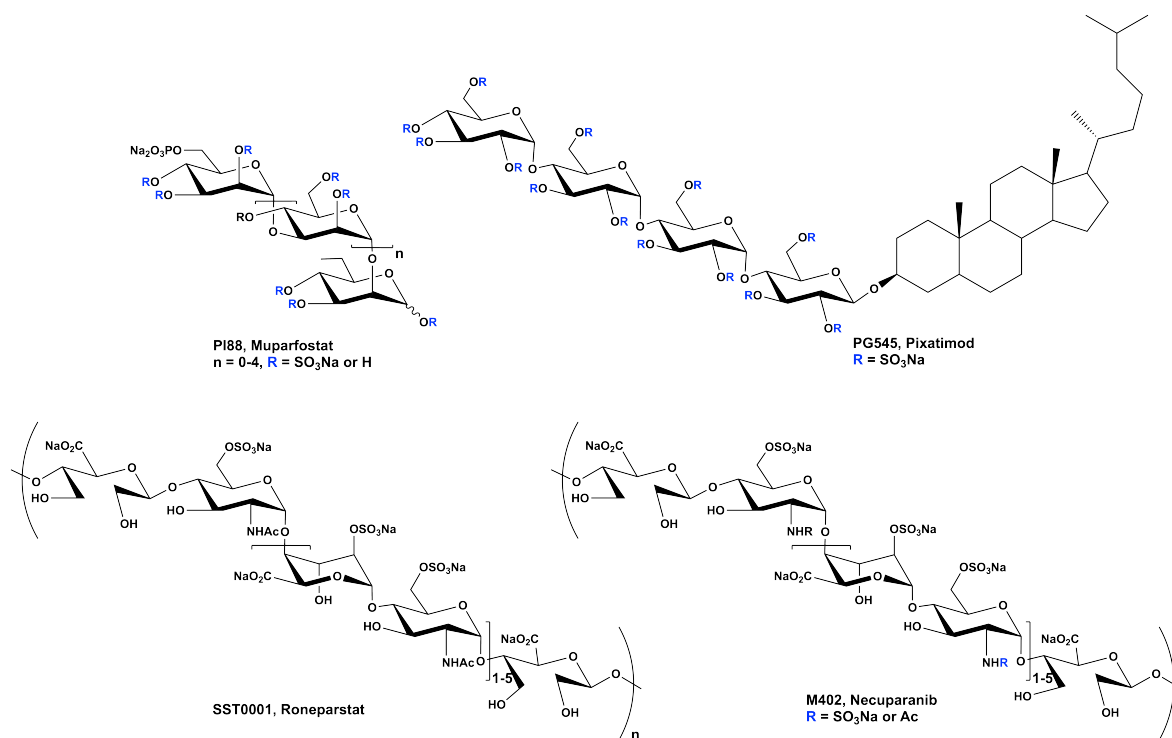
Figure 1.4: Structures of selected sulfated polysaccaride and sulfated oligosaccaride inhibitors of heparanase

PI-88 (Muparfostat) was the first heparanase inhibitor to reach Phase III clinical trials. It is a mixture of several highly sulfated mannose-oligosaccharides, ranging from two to six saccharides in length, derived from the yeast *Pichia (Hanensula) holstii* (*105*). PI-88 was designed to simultaneously inhibit not only heparanase, but also other HS-binding pro-angiogenic growth factors. It was also shown to block the enzymatic activity of endoglucosamine 6-sulfatases, which have pro-angiogenic activity. During preclinical trials, PI-88 was shown to act at multiple stages of carcinogenesis, inhibit cell proliferation, increase tumour cell apoptosis, impair angiogenesis and reduce the number of invasive carcinomas (*89*). However, it has shown adverse effects such as dehydration, fatigue, diarrhea, nausea, thrombocytopenia and high rates of febrile neutropenia, which led to the termination of the clinical trials (*106*).

PG545 (Pixatimod) is a second-generation PI-88 inhibitor and is a synthetic, fully sulfated HS mimetic that contains a cholesterol-conjugated maltotetraose sulfate. It has a longer half-life than PI-88 and benefits from being structurally homogeneous. PG545 has also been shown to have anti-angiogenic and anti-tumor activity, likely due to its ability to bind to multiple targets (*88*). This mimetic was tested in Phase Ia clinical trials and it was found that intravenous delivery was the only successful delivery method (limiting its use). However, it is currently undergoing Phase 1b trials to be used in combination with other anti-cancer drugs (*107*)

There are several limitations to the use of oligosaccharides as drugs, including their high molecular weight and heterogeneous nature due to multiple ring conformations and various lengths. These limitations can make it difficult to characterise the drug's structure, interpret biological assays due to molecular heterogeneity and develop delivery methods (*80*). To address many of the issues related to HS mimetics, further development of oligosaccharide inhibitors as single chemical entities may be necessary. Oligosaccarides also have poor physicochemical properties such as low logP values, resulting in poor bioavailability. Such as seen with PG545, intravenous delivery is required for these therapeutics, limiting their use and accessibility.

1.6.5   Small molecules

Oligosaccharides and polysaccharides have low lipophilicity due to the charge they carry, which results in poor absorption into the body (*108*). This means that these inhibitors need to be administered intravenously. Small molecules typically have superior pharmacokinetic properties compared to oligosaccharides due to their smaller size and more modifiable structure, which can improve uptake (*97*).
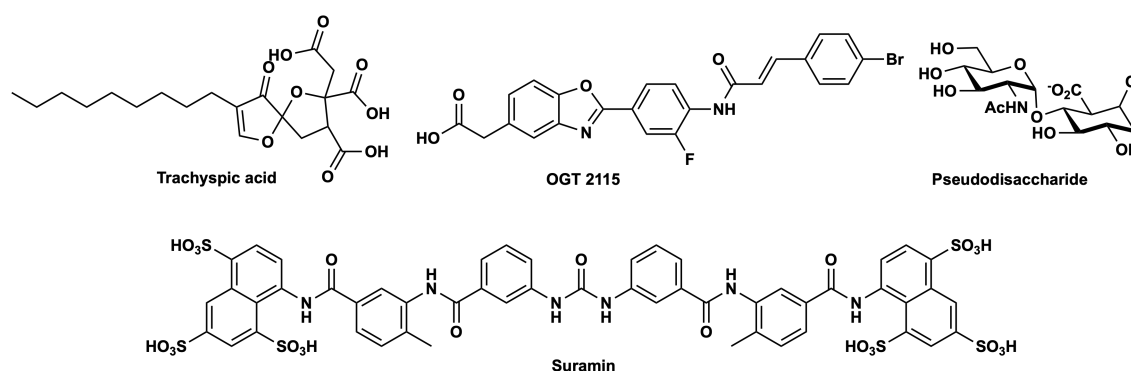
Figure 1.5: Structures of selected small molecule inhibitors of heparanase

Symmetric and asymmetric urea-based compounds are another common class of drugs used for heparanase inhibition. The first example of this class was **Suramin**,

polysulfonated naphthylurea used to treat African sleeping sickness and river blindness (*109*). Suramin is the only reported heparanase inhibitor that inhibits *via* a non-competitive binding mechanism and has an IC$_{50}$ of 46 μM (*110*). However, this compound failed to advance to clinical trials due to potential effects on multiple cellular pathways and the potential for serious side effects. Despite this, the suramin scaffold has continued to be a focus of drug development programs in an effort to increase its activity and decrease its side effects.

Another common class of heparanase small molecule inhibitors are the benzazoles, most commonly benzimidazol-2-yl and 2- or 5-substituted benzoxazolyl derivatives. **OGT-2115** is a benzoxazole acetic acid that was identified as a heparanase inhibitor *via* HTS and secondary screenings of leads, which showed positive results for heparanase binding *in vitro*. However, no *in vivo* data was ever released and the inhibitor never reached clinical trials (*79*). More inhibitors based on the OGT-2115 scaffold were subsequently designed, with some related derivatives exhibiting better inhibition (*78*). While many of these analogues exhibited reasonable IC$_{50}$ values, none of them have progressed to clinical trials due to poor physiochemical properties (*97*).

Recently, polyanionic saccharide-based covalent inhibitors have garnered interest as a potential solution to the limitations of both HS mimetics and standard small molecule inhibitors. These inhibitors selectively bind to the heparanase active site and the catalytic residue Glu343, leading to irreversible inhibition. In cellulo and in vivo cancer models have demonstrated that these inhibitors exhibit low micromolar inhibition at concentrations of 0.46 µM and reduce cancer aggression (*111*).

Many other computational and experimental screenings have resulted in the discovery of many small molecule inhibitors, including indoles, carbazoles, fluorenes, diphenylethers, rhodanines, triazolo-thiadiazoles, furanthiazoles, DMBO analogues and iminosugars, Most of these inhibitors have µM activity, but almost all of them lack kinetic and structural information about their binding mechanisms. Overall, there is a broad range of structural diversity among small molecule inhibitors for heparanase and a comprehensive chemical-based SAR has not been possible. Therefore, it has been suggested that there are multiple mechanisms of inhibition within the small molecule class of heparanase inhibitors, contributing to the difficulty of developing better therapeutics.

1.7 Techniques for Investigating Inhibitor Interactions

A detailed understanding of inhibitor binding mechanisms requires input from multiple techniques, including biophysical, structural and kinetic approaches. This thesis is centered on using X-ray protein crystallography, isothermal titration calorimetry and kinetics, in combination with other biophysical and computational approaches, to gain a strong understanding of inhibitor binding mechanisms.

1.7.1 Protein X-ray crystallography

X-ray crystallography is a powerful form of microscopy that allows us to visualize atoms and molecules. The goal of protein X-ray crystallography is to generate a three-dimensional model of a crystalline protein's structure from its X-ray diffraction pattern, as well as any target compounds that bind to the protein, in order to gain insights into the mechanisms of ligand binding. X-rays lie in the electromagnetic spectrum between ultraviolet light and gamma radiation, with a wavelength in the order of angstroms (A = 10^{-10} m) and can therefore be used to resolve individual atoms in a protein, which have an average distance of 1.5 Å. When an X-ray beam is scattered by a protein crystal, it produces a characteristic diffraction pattern. The scattering angles at which diffracted X-rays are observed are indicative of the geometry of the crystal lattice, while the intensities of the diffracted X-rays contain information about the spatial distribution of electron density within the crystal. For the theoretical background associated with protain X-ray crystallography, please refer to (*112*).

X-ray crystallography is a common technique used in drug discovery and it has become increasingly important as the idea of rational drug design has gained traction. The ability to design and optimize drug leads using the knowledge provided by an X-ray crystal structure of a macromolecular target in complex with an inhibitor has transformed the way drug design is approached. X-ray crystallography is also important for the development of structure-activity relationship (SAR) analysis, in which the structures of multiple complexes are determined to guide ligand optimization, as well as for high-throughput screening of crystals and new inhibitors (*113*).

X-ray crystallography experiments can reveal the intrinsic dynamics of proteins or regions of proteins that are highly flexible and associated with poor density. This can also be the case for ligands that are highly mobile in the crystal structure or that have low occupancy, only being present in a fraction of the unit cells of the crystal. B-factors can represent this, as they reflect the mean square displacement of non-hydrogen atoms in different regions of the protein (*114*). However, in some

cases, the dynamics of the protein or ligand may be constrained in the crystal lattice or a particular conformation may not be sampled. In these cases, other techniques such as molecular dynamics (MD) simulations or hydrogen-deuterium exchange mass spectrometry can be used to observe these dynamics.
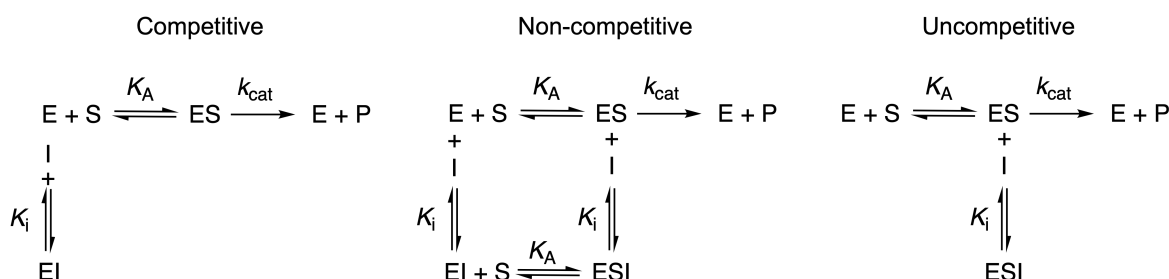
1.7.2 Inhibition Kinetics

The catalytic activity of enzymes can be inhibited by hundreds of natural compounds and thousands of synthetic chemical compounds. Inhibitors are defined as chemical compounds that reduce the rate of enzyme-catalyzed reactions when added to a reaction mixture. Inhibition can occur in a variety of ways, so there are several types of inhibition. There are two main groups of inhibitors: reversible inhibitors and irreversible inhibitors. Irreversible inhibitors are often chemically reactive compounds that enter into chemical reactions with the enzyme, forming covalent bonds and inactivating the enzyme. An enzyme that has undergone this reaction with an irreversible inhibitor cannot regain activity through dialysis or other inhibitor removal methods. The various kinetic models discussed here have been summarised from Leskovac, 2003. (*115*)

The more common class of inhibitor is the reversible inhibitor, which forms non-covalent bonds with the enzyme. This reaction reduces the amount of free enzyme available to participate in normal reactions in solution. Reversible inhibitors can be removed through dialysis, allowing the enzyme to regain its original catalytic activity. Reversible inhibitors can form different interactions with the free enzyme, resulting in different catalytic properties. Reversible inhibitors can be classified into three main types: competitive, non-competitive and uncompetitive. Competitive inhibition can occur in enzyme reactions with one or more substrates or products, while non-competitive and uncompetitive inhibition typically occur only in reactions with two or more substrates or products. These inhibitors can also be further divided into linear and non-linear kinetics based on their effect on the rate of the reaction.

In simple terms, **competitive inhibition** occurs when the inhibitor binds to the same site on the enzyme as the substrate, forming an inactive complex **(Figure 1.6, left)**. This means that the substrate and inhibitor compete for the same binding site and only one enzyme-inhibitor complex can be formed. A competitive inhibitor will increase the K_m (Michaelis-Menten constant) without affecting the V_{max} (maximum velocity), while IC_{50} (the concentration of inhibitor required to inhibit the enzyme by 50%) will increase indefinitely with an increased substrate concentration.

Non-competitive inhibition is when an inhibitor has no effect on substrate binding and vice versa. The inhibitor and substrate bind reversibly and independently at different sites on the enzyme **(Figure 1.6, middle)**. The substrate can bind to both the enzyme (E) and enzyme-inhibitor (EI) complex, while the inhibitor can bind to both the E and enzyme-substrate (EA) complex. However, the EAI complex is inactive. Therefore, a non-competitive inhibitor will decrease the V_{max} without affecting the K_m, while the IC_{50} remains independent of substrate concentration.

Uncompetitive inhibition occurs when an inhibitor binds reversibly to the enzyme-substrate (EA) complex, resulting in the formation of an inactive EAI complex **(Figure 1.6, right)**. The inhibitor is unable to bind to the free enzyme. This type of inhibition is characterised by a decrease in both the V_{max} and the K_m. Uncompetitive inhibitors also exhibit higher inhibition at higher substrate concentrations.

Figure 1.6: kinetic models for competitive, non-competitive and uncompetitive inhibition

The previously listed inhibition types all result in a dead-end EI complex or non-productive ESI complex, or both. Not all inhibition types fall within these three types of models and can result in complicated nonlinear types of inhibition. Here we will discuss hyperbolic and parabolic types of inhibition.

Hyperbolic Inhibition occurs when the enzyme can form an ESI complex that can still yield a product with equal or lesser efficiency than the ES complex **(Figure 1.7, left)**. In this case, the inhibition velocity functions will display a nonlinear dependence on inhibitor concentration. The EI complex may also have different affinities for the substrate compared to the free enzyme and the ES complex may have different affinities for the inhibitor. In this case, an infinitely high inhibitor concentration cannot drive the velocity to zero, resulting in incomplete or partial inhibition.

Parabolic inhibition can take different forms depending on the interaction of the inhibitor with the protein. In the case of parabolic competitive inhibition, which we will discuss in this thesis, an enzyme that can bind to one substrate can also bind two molecules of the same competitive inhibitor **(Figure 1.7, right)**. The binding of the inhibitor at either site should be sufficient to exclude the substrate. The binding

of the first inhibitor molecule increases the affinity of the subsequent molecule's interaction with the enzyme. This changes the dissociation constant of the active site, resulting in a parabolic relationship with respect to inhibitor concentration. Note that ligand-induced protein aggregation can be mistaken for this mechanism.
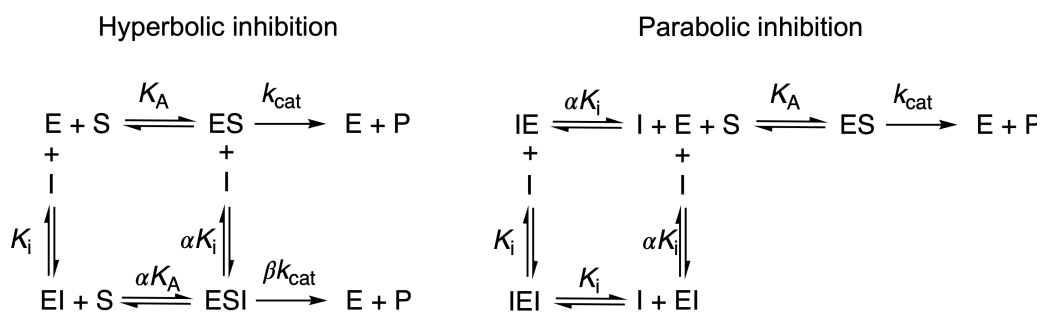

**Figure 1.7:** kinetic models hyperbolic and parabolic inhibition

1.7.3 Isothermal Titration Calorimetry

Isothermal titration calorimetry (ITC) is a label-free method that determines the thermodynamic parameters associated with protein-ligand interactions by directly measuring heat. From a single ITC experiment, it is possible to determine the enthalpic (ΔH) and entropic (ΔS) components of the Gibbs free energy of binding (ΔG), which together constitute the thermodynamic signature of a protein-ligand interaction. The association constant (K_a) and stoichiometry (n) of the interaction can also be determined. This method is considered the gold standard for understanding the thermodynamic signatures of macromolecular binding interactions because it is universal.

In a typical ITC experiment, the ligand is titrated into a solution of the protein of interest in a sample cell. The experiment is performed using a power compensation instrument, which applies constant power to keep the temperature of the sample cell and an identical water-filled reference cell at the same temperature. Each injection of ligand into the protein solution results in an associated peak in power applied to the sample cell **(Figure 1.8 A)**. The peak is then integrated to produce a binding isotherm that reflects the total amount of heat produced or absorbed during each injection of ligand as a function of the protein:ligand ratio **(Figure 1.8 B)** There are several binding models that can be used to analyze the binding mechanisms observed through ITC, depending on the number and affinity of binding sites. It is important to choose the appropriate binding model to obtain accurate binding parameters.

Proper experimental design is essential for obtaining accurate binding parameters from ITC experiments. The best conditions for the experiment are those in which

the heat change is measurable for each injection and the heat changes vary for subsequent injections. To achieve these conditions, it is important to consider the following variables: the concentration of the sample in the cell, the ligand:protein ratio, the number and volume of injections. A typical ITC experiment includes around 20 injections, but this number can be adjusted based on the expected Kd rates and affinity of the interaction. It is important to ensure that sufficient ligand is bound to the protein in order to calculate an accurate isotherm. For high affinity interactions, this may require a 2-3 molar excess of ligand, while lower affinity interactions may require a larger excess (*116, 117*).
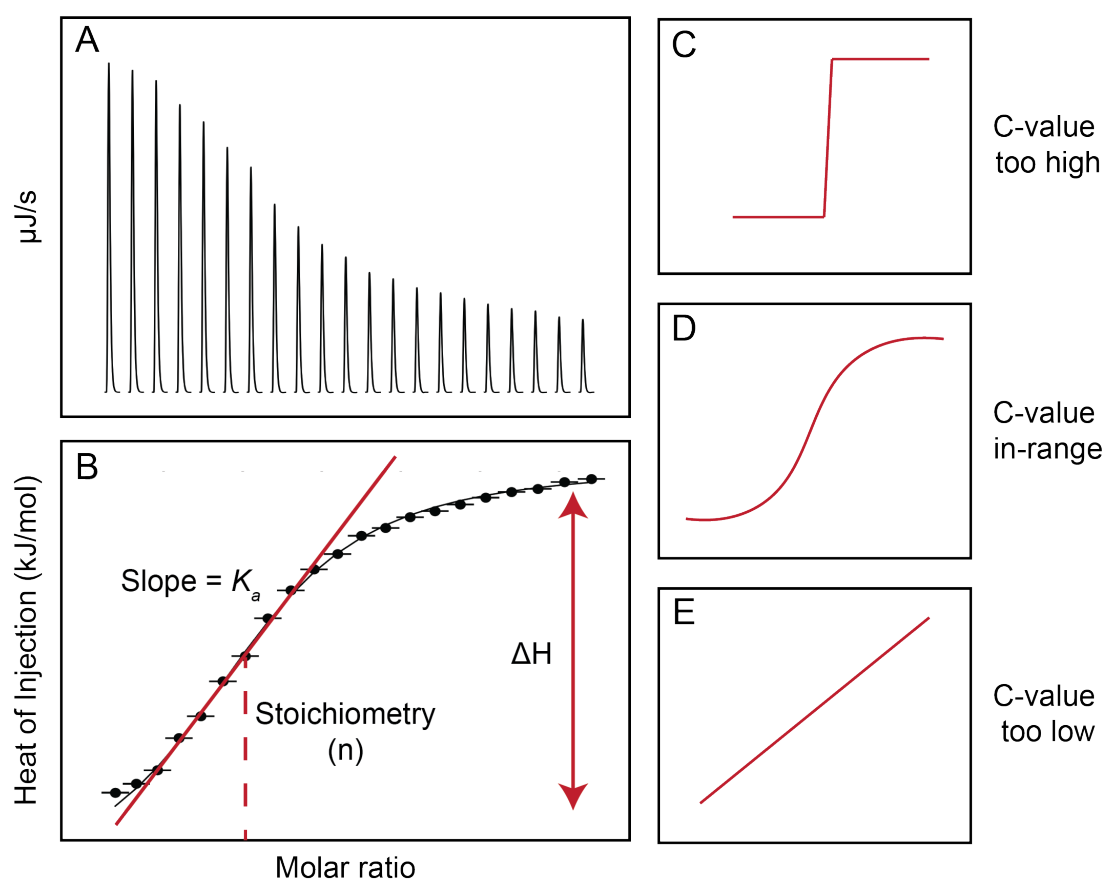
Figure 1.8: A typical isothermal titration calorimetry experiment. A) A power compensation ITC measure the power applied to the cell following an injection of the analyte into the reaction cell. **B)** Integration of the area under the peak for all injections produces a binding isotherm which an appropriate model can be fitted. From this plot, the thermodynamic parameters of binding can be determined, such as binding affinity, enthaply and stoichiometry. **C-E)** Illustration of the effect of the association constant value on the shape of the titration curve. When the C-value is large enough, a good estimate of the enthalpy can be determined.

To obtain the correct shape of the binding isotherm, it is important to optimize the concentration of protein in the cell ([P]). The shape of the isotherm is determined by

the value of the parameter c, which is calculated using the following equation:

$$\mathbf{c} = [\mathbf{P}] \ \times \mathbf{n} \ \times K_a \tag{1.1}$$

Where [P] is the protein concentration, n is the stoichiometry and K_a is the binding constant. A large c value corresponds to a tight binder, resulting in an isotherm that is rectangular in shape with minimal data points in the transition of binding, therefore making the estimation of K_a unreliable. On the other hand, a low c value corresponds to weak binding and can make the estimation of K_a and ΔH unreliable **(Figure 1.8 C-E)** (*118*). Typically, an experiment in which c is between 1 and 1000 can be used to determine binding constants (*118*). This information should be considered when designing ITC experiments.

ITC is a valuable technique for understanding the thermodynamics of protein-ligand interactions, but it has some limitations. One limitation is that it is not a high-throughput system and requires large amounts of sample to obtain useful information, which is especially important for low affinity systems. This can make it challenging to use ITC if the solubility or accessibility of compounds is limited. Additionally, ITC is not suitable for analyzing systems with multiple binding sites or those that involve conformational changes in the protein. Finally, ITC is limited on its throughput, as ITC experiments takes hours for a single replicate.

1.8 Thesis Motivation & Research Objectives

1.8.1 Research Objectives

The work presented in this thesis is structured around structural and biophysical techniques, combined with kinetics, to understand the interaction of different heparanase inhibitors. The goal of this research was to explore these interactions using an engineered heparanase model in order to see how different inhibitor types differ in their binding and inhibition of heparanase. To understand these interactions, a variety of methods were used, including inhibition assays, kinetics, affinity measurements, X-ray crystallography, HDX-MS and mutagenesis. The combination of these different techniques facilitated the discussion of binding mechanisms and inhibitions with heparanase. In each case, the mode of action of the inhibitors and their potential as therapeutics were discussed. The inhibitors tested here were also compared with other similar therapeutics in the literature, leading to a greater understanding of the heparanase therapeutic research field as a whole. This thesis also demonstrates

the importance of structural biology in therapeutic development and the need for synthetic chemistry, biology, engineering and medicine to work together to design better therapeutics in the future.

1.8.2 Research Outline

The following chapters highlight the importance of structural biology and detailed protein-ligand interaction analysis for the development of potential therapeutics. In the past, the lack of accessible heparanase protein has limited the ability to conduct such studies, resulting in a limited understanding of drugs that have been developed. These chapters aim to expand our knowledge of heparanase therapeutics in order to design new and better therapeutics in the future

In the following chapter (Chapter 2), I present a publication that describes the computational design of a soluble heparanase variant (HPSE P6). We showed that this variant is the first time heparanase could be expressed in high yields in *E. coli* and using kinetic studies, X-ray crystallography and molecular dynamics, we provided evidence that this variant can be used for *in vitro* methodologies. This work is a crucial step in heparanase-inhibitor interaction analysis because it provides a useful model that can be used with different biophysical techniques in the future.

In Chapter 3, I present a publication that investigates the mechanism of heparanase with HS mimetic inhibitors, focusing on the mimetic Pentosan polysulfate sodium (PPS). In this paper, we used three synthetic analogues of PPS to understand how the length of the mimetics affects heparanase inhibition. By using kinetics, ITC and X-ray crystallography, we showed that the binding of HS mimetics occurs at three binding sites on the surface of heparanase, rather than in the active site. Through the use of aggregation assays, CD and additional kinetic assays, we identified that this binding mechanism is contributed by aggregation and secondary structure changes, which increases the perceived potency of these inhibitors. We highlight that this mode of inhibition is common among other HS mimetics, such as those in clinical trials and discuss the potential connection between these interactions and the poor success of these drugs in clinical trials.

Chapter 4 provides a detailed analysis of a new class of small molecule inhibitors of heparanase. A large structure-activity relationship (SAR) study was conducted to understand how the structure of these compounds interacts with heparanase. The results showed that these compounds result in non-competitive inhibition of heparanase, as demonstrated by various assays and ITC experiments. Mutagenesis studies were also conducted to test the hypothesis that these compounds bind remote

from the active site, although the specific allosteric site could not be conclusively identified. The chapter highlights the lack of understanding surrounding many small molecule heparanase inhibitors and the need for more structural information to allow better therapeutics to be developed.

Finally, in Chapter 5, I summarize the work presented in this thesis, discuss the implications of this work on the heparanase therapeutics field and outline further research and challenges for the field moving forward. Overall, this research has made structural biology studies on heparanase more readily accessible through the use of HPSE P6, enabling the study of inhibitor interactions that were previously inaccessible.

Chapter 2

Computational design and experimental characterisation of a stable human heparanase variant

2.1 Declaration

The following article was peer reviewed and published in RSC chemical biology All experimental work except where otherwise stated, was carried out by the author under the supervision of Professor Colin Jackson The contribution of other authors are as follows: Dr. Nansook Hong conducted cloning and expression trials as well as analysed crystal structures. Dr Josh A. Mitchell helped the author with the running and analysis of molecular dynamics simulations and principal component analysis. Dr Nansook Hong, Professor Colin J. Jackson, and the author conceived the experiments. The manuscript was written by the author and the primary investigators. Note that the page numbering of the following article is from the published journal number. The online article can be found at the DOI listed below

'Computational design and experimental characterisation of a stable heparanase variant'

DOI: 10.1039/d1cb00239b

PAPER

Check for updates

Computational design and experimental characterisation of a stable human heparanase variant†

Cassidy Whitefield, [ID] ‡[a] Nansook Hong, [ID] ‡[a] Joshua A. Mitchell [ID][a] and Colin J. Jackson*[ab]

Heparanase is the only human enzyme known to hydrolyse heparin sulfate and is involved in many important physiological processes. However, it is also unregulated in many disease states, such as cancer, diabetes and Covid-19. It is thus an important drug target, yet the heterologous production of heparanase is challenging and only possible in mammalian or insect expression systems, which limits the ability of many laboratories to study it. Here we describe the computational redesign of heparanase to allow high yield expression in *Escherichia coli*. This mutated form of heparanase exhibits essentially identical kinetics, inhibition, structure and protein dynamics to the wild type protein, despite the presence of 26 mutations. This variant will facilitate wider study of this important enzyme and contributes to a growing body of literature that shows evolutionarily conserved and functionally neutral mutations can have significant effects on protein folding and expression.

Introduction

Heparan sulfate (HS) consists of 1–4 linked disaccharide units that are negatively charged and structurally heterogeneous due to variable sulfation, deacetylation and epimerization during biosynthesis.[1] HS is often covalently linked to proteins and peptides to form heparan sulfate proteoglycans (HSPGs).[2] HSPGs are themselves a major component of the extracellular matrix (ECM) and basement membranes, forming a protective barrier by interacting with other major components of the ECM such as collagen, fibronectin and laminin. Their structural diversity and negative charge attract various cationic proteins and water, forming porous hydrogels that are able to store bioactive molecules including growth factors,[3,4] chemokines[5] and enzymes.[6]

Heparanase (HPSE) is the only mammalian enzyme that is known to hydrolyse HS.[7–10] In adults, HPSE is normally expressed at low levels, found only in platelets, immune cells and the placenta,[11–13] but increased expression of HPSE has been observed in many disease states, including cancer and Covid-19.[14–16] When overexpressed, HPSE catalyses the hydrolysis of HS, resulting in weakening of the ECM barrier, which can promote inflammation,[17,18] cancer cell invasion, growth and migration,[19,20] as well as angiogenesis.[21,22] HPSE is also associated with tumour initiation by up-regulating pro-inflammatory mediators.[23,24] Moreover, animal studies have shown HPSE genetic knock-outs improve cancer prognosis and increased survival without significant side effects.[25,26]

Owing to its roles in many diseases, HPSE has been a drug target for many years. For instance, HPSE expression promotes resistance to chemotherapy, whereas targeting HPSE with inhibitors can overcome chemoresistance and tumour relapse.[27] Indeed, many groups have attempted to produce drug-like HPSE inhibitors over recent decades.[1,26,28,29] However, HPSE production currently relies on complex and expensive eukaryotic expression systems such as mammalian[7,8,30] and insect cells.[31,32] While some prokaryotic HPSE expression methods have been reported,[33,34] they have not been sufficiently robust for widespread adoption. HPSE has many features that are known to reduce soluble expression prokaryotic systems, such *as Escherichia coli*, including multiple disulfide bonds and large positive regions on the surface,[35–37] as well as *N*-glycosylation.[22,38] Moreover, HPSE is natively expressed as a pre-proheparanase which undergoes proteolytic cleavage of a signal peptide then a linker segment, resulting in an active heterodimer composed of 8 kDa and 50 kDa subunits (Fig. 1)[39] Thus, in prokaryotic expression systems the 8 kDa and 50 kDa subunits have to be expressed separately and assemble into a heterodimeric complex.[33,34]

[a] *Research School of Chemistry, Australian National University, Canberra, ACT, 2601, Australia. E-mail: colin.jackson@anu.edu.au*

[b] *Australian Research Council Centre of Excellence for Innovations in Peptide and Protein Science, Australian National University, Canberra, ACT 2601, Australia*

† Electronic supplementary information (ESI) available. See DOI: 10.1039/d1cb00239b

‡ Contributed equally to this work.
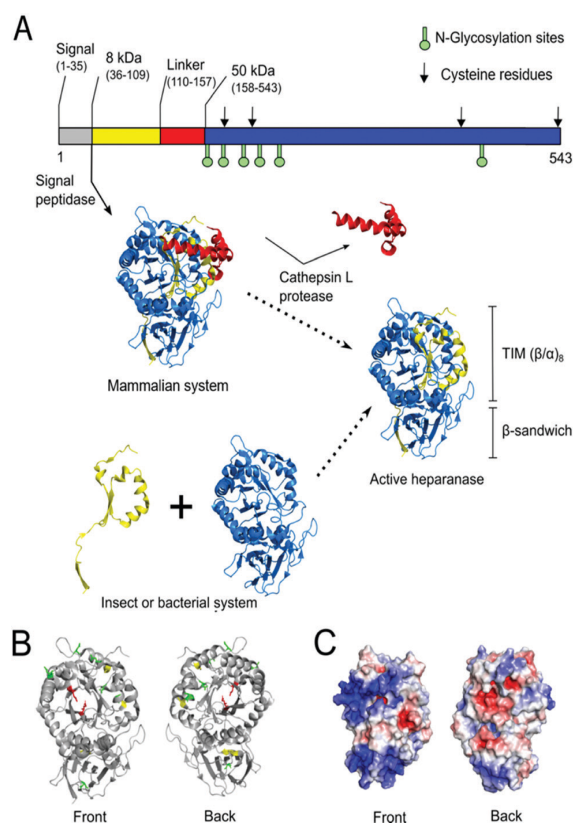
Fig. 1 Native maturation and folding of HPSE in mammalian cells compared with heterologous production in insect or bacterial systems. (A) In mammalian cells, pre-proheparanase (Met1–Ile543) undergoes successive cleavage events of the N-terminal signal peptide (Met1–Ala35) and linker (Ser110–Gln157, red cartoon) segments to produce mature HPSE. The resulting heterodimer assembly of two subunits (8 kDa subunit (Gln36–Glu109, yellow cartoon) and 50 kDa subunit (Lys158–Ile543, blue cartoon)) consists of a TIM barrel $(\beta/\alpha)_8$ and β-sandwich fold. The sequence of HPSE is shown on top as a bar representation in which glycosylation sites are shown as green sticks and cysteine residues are indicated as black arrows. In non-mammalian systems active protein must instead be produced *via* co-expression of the two subunits.[31] (B) Crystal structure of human HPSE expressed in an insect expression system (PDB ID: 5E8M) is shown as grey cartoons (bottom-left). Six *N*-glycosylation sites (Asn162, 178, 200, 217, 238, 459) are shown as green sticks and four cysteines (Cys179, 211, 437, 542) are shown as yellow sticks whereby two of them form a disulfide bond (Cys437-542) at the β-sandwich domain. Catalytic residues (Glu343, Glu225) at the TIM face are shown as red sticks. (C) Electrostatic potential surface was calculated using amino acid residues in the crystal structure by APBS (glycans were not included in the calculation). This shows two large positively-charged patches at the TIM domain and at the β-sandwich domain, which may promote aggregation in the nucleic acid rich micro environment.[37]

There are many experimental and computational methods that have been developed to improve enzyme function and stability, such as bioinformatics-based approaches like consensus design[40,41] or ancestral sequence reconstruction,[42,43] or forcefield-based approaches like Rosetta[44] or FoldX.[45] However, both approaches have limitations. The Protein Repair One Stop Shop (PROSS) algorithm combines forcefield-based Rosetta modelling and phylogenetic sequence information to create variants with improved stability.[46] Here, we describe the use of PROSS to

generate the first stable human HPSE variant to be expressed in *E. coli*. We demonstrate that it has significantly increased solubility, very similar catalytic activity and identical inhibition by competitive inhibitors, when compared to wild type human HPSE produced from mammalian cells. Our results are supported by an X-ray crystal structure and molecular dynamics simulations, which demonstrate that the introduced mutations stabilise HPSE with almost no effect on the three-dimensional structure or dynamics. This mutant HPSE should significantly reduce the costs and technical barriers to the development of HPSE inhibitors and its widespread study.

Results

Computational design of a soluble HPSE variant

We first tested bacterial expression of human HPSE (wild type) by cloning two subunits (8 kDa and 50 kDa, Tables S1 and S2, ESI†) into the dual expression vector (pETDuet-1). To optimize the chances of obtaining soluble, properly folded protein, we co-expressed the protein with chaperones (trigger factor and GroEL/GroES),[47–49] and used *E.coli* Shuffle T7 Express[50] cells, which allow disulfide bonds to form in the cytosol. Under these conditions, the 50 kDa subunit was totally insoluble, while the 8 kDa subunit was partially soluble (Fig S1, ESI†).

Given that the molecular structure of HPSE has recently been solved,[31,32] it is now an appropriate candidate to be engineered to allow expression in simple and inexpensive expression systems, such as *E. coli*. Recently the PROSS algorithm[46] has demonstrated its utility in designing stable variants of challenging proteins for soluble and functional expression in bacteria.[51–53] Unlike conventional consensus mutagenesis approaches, in which poorly conserved residues are mutated back to their consensus identity (from a multiple sequence alignment),[54] PROSS combines this approach with computational modelling with Rosetta,[55] generating a set variants, each containing multiple mutations that ideally act together to increase stability.[46] We therefore used PROSS to redesign the insoluble 50 kDa subunit based on the crystal structure of the insect cell expressed human HPSE (PDB ID: 5E9C). The substrate binding site and the heterodimer interface residues were restrained to maintain function and preserve the interaction with the 8 kDa subunit. Seven variants with accumulated mutations were generated (Fig. S2, ESI†), which were subsequently synthesized and sub-cloned into multiple cloning site 2 of pETDuet-1 vector. The 8 kDa subunit with a N-terminal poly histidine tag was sub-cloned into multiple cloning site 1. All variants were tested (Fig. S3, ESI†) and the most soluble design, containing 26 amino acid substitutions was identified and purified (HPSE P6), using Ni^{2+}-NTA, heparin and size exclusion chromatography. This resulted in pure, homogeneous, heterodimeric HPSE with a final yield of 4 mg from 1 L *E. coli* culture (Fig. 2). Notably, PROSS is not infallible; many of the designs did not produce soluble protein, which emphasises the need to test multiple different variants.

Paper
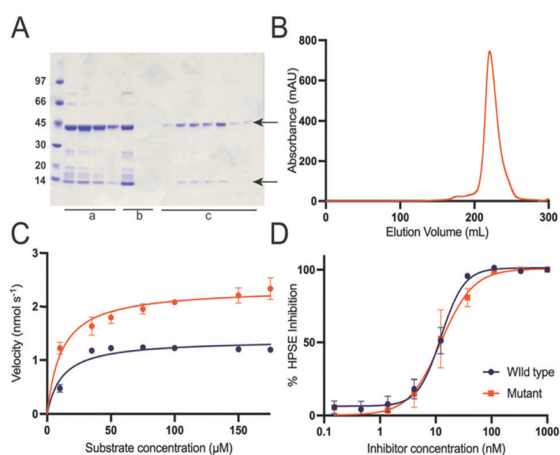
Fig. 2 Expression, purification and activity of the successful HPSE P6 (A) Ni-NTA elution fractions (lanes a), heparin column flow in and flow through fractions (lanes b) and size exclusion elution fractions (lanes c). LMW protein marker (GE healthcare) is on the first lane. The sizes corresponding to the two subunits of the HPSE sit at 44 kDa, and 8 kDa, noting that the size of the large subunit is smaller than the previously reported value of 50 kDa due to the lack of glycosylation. (B) Size exclusion chromatography (HiLoad 26/600 Superdex 200 pg, GE Healthcare) shows one major peak corresponding to monomeric HPSE. (C) Kinetics HPSE P6 and HPSE WT, where catalytic rate (k_{cat}) of HPSE P6 was 70% higher (k_{cat} 2.94 ± 0.13) compared to (k_{cat} 1.72 ± 0.07) for WT. The binding affinity (K_M 11.6 ± 2.8 μM and 11.83 ± 2.7 μM respectively) was the same. (D) Pentosan polysulfate was used to compare the design with the human HPSE expressed in mammalian cell. This was measured using colorimetric method with fondaparinux. Error bars represent standard error from a minimum of three measurements.

HPSE P6 exhibits wild-type like activity and response to inhibitors

Given the large number of mutations and loss of glycosylation sites, it was important to test whether these changes had any effect on the activity of the protein. The catalytic activity of purified HPSE P6 was tested by colorimetric assay using fondaparinux[56] (Arixtra), a synthetic analogue of HS. Although, the catalytic rate (k_{cat}) of HPSE P6 was slightly (less than 2-fold) higher (k_{cat} = 2.94 ± 0.13 s^{-1}) compared to HPSE WT (k_{cat} = 1.72 ± 0.07 s^{-1}), the binding affinity (K_M = 11.6 ± 2.8 μM and 11.83 ± 2.7 μM respectively) was the same, demonstrating that the introduced mutations have no effect on the interaction between enzyme and substrate (Fig. 2C, Table 1). In fact, the slight increase in k_{cat} is most likely due to the higher purity of the HPSE P6, compared to the commercially available HPSE WT. The loss of the six glycosylation sites no effect on the activity of the enzyme, suggesting these sites may be important for protein solubility in mammalian systems.

Having established the enzyme kinetic parameters are comparable to HPSE WT, we then tested whether the HPSE P6 variant would interact identically with heparan sulfate mimetic inhibitors; in this case pentosan polysulfate[57] (Fig. 2C and D). As with the enzyme kinetics, the inhibitory response to the model inhibitor pentosan polysulfate was near identical between HPSE WT and HPSE P6, with an IC$_{50}$ of 12.46 ± 1.26 nM and 12.43 ± 2.47 nM, respectively (Fig. 2D).

HPSE P6 is thermostable and structurally isomorphous to HPSE WT

The thermal stability HPSE P6 was measured using circular dichroism (CD) by observing the loss of helicity at 222 nm over 20–90 °C. This revealed that HPSE P6 is somewhat thermostable, undergoing a transition to an unfolded state with a T_m value of 63.6 ± 0.19 °C, (Fig. 3A). This Tm value is similar to other engineered variants of human proteins produced through the use of PROSS,[46,53] and significantly exceeds the normal temperature range that human proteins are exposed to (∼37 °C).

To understand how the 26 mutations in HPSE P6 result in enhanced protein folding and stability, we solved the crystal
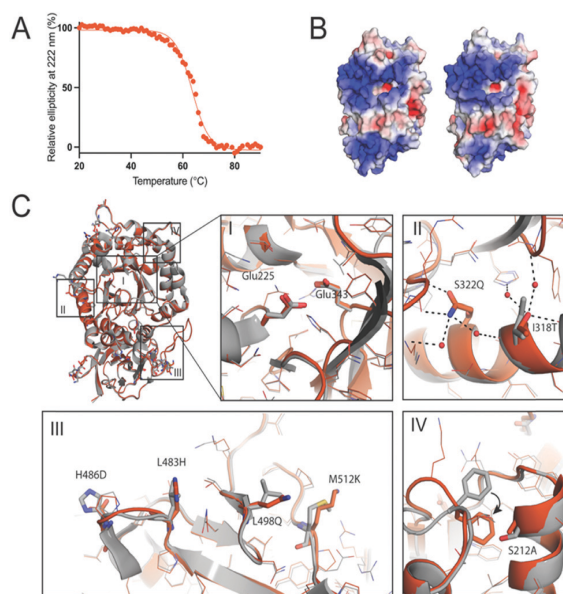
Fig. 3 Thermal stability and structural insight of the designed HPSE. (A) The ellipticity at 222 nm was measured using circular dichroism over 20–90 °C, resulting in the melting temperature of 63.6 °C, similar to the values of other engineered proteins by the PROSS algorithm.[46,51,53] (B) The front views of wild type (WT) and designed HPSE (HPSE P6) are calculated using APBS[58] and visualized using PyMol. (C) The superimposed structures of wild type (grey) and HPSE P6 (orange, mutated residues are shown as sticks) are shown as overall (top-left) and detailed views (I–IV). Overall (top-left) and active site view (I) show closely aligned Cα backbones (RMSD of 0.645 Å) and the side chain conformations in the active site. Overall, the mutations reduce the hydrophobicity and increase polarity (II and III), to introduce new hydrogen bonds (black dotted lines, II) and to increase the hydrophobic packing (IV). The Phe258 side chain folds into the hydrophobic packing area (shown as black arrow) with a nearby mutation Ser212Ala causes a loss of an interaction with Thr257. PDB ID: 5E8M (WT), 7RG8 (HPSE P6).

Table 1 Kinetic and Inhibition parameters for HPSE WT and P6 proteins

Parameter	HPSE WT	HPSE P6
K_M (μM)	11.8 ± 2.7	11.6 ± 2.8
k_{cat}	1.72 ± 0.07	2.94 ± 0.13
IC$_{50}$ (nM)	12.5 ± 1.3	12.4 ± 2.5

structure of HPSE P6 at 1.30 Å resolution (Table S3, ESI†). The protein crystallised in the $P2_12_12_1$ space group within 1 day, forming rod shaped crystals. This compares to WT HPSE crystallising in the $P2_1$ space groups, after 1–3 days.

Despite 26 mutations, the crystal structure of HPSE P6 shows almost identical overall backbone and active site structures to HPSE WT expressed in an insect system (with the exception of the absence of any glycosylation). The Cα RMSD was 0.645 Å, with an alignment score of 0.017 (Fig. 3C).

Many subtle changes were observed due to the 26 mutations. Firstly, surface polarity, which is known to positively contribute to folding and stability,[59] was increased by substitutions to surface leucine and alanine residues to more polar or smaller amino acids (*e.g.* Leu197Gly, Leu354Gly, Leu498Gln, Leu230Arg, Ala195Ser) (Fig. 3C). Secondly, additional stabilising interactions, including increased hydrogen bonding networks and hydrophobic packing were introduced, which stabilise the folded state. For example, new hydrogen bonding interactions were introduced by Leu483-His, Ile318Thr, Lys477Gln and Ser322Gln and new hydrophobic interactions were introduced by Ser530Ala, Ser292Ala and Arg307Leu in the partially solvent exposed areas. Interestingly, we observed indirect conformational change of Phe258 by Ser212-Ala (Fig. 3C.iv). Finally, the disulfide bond (Cys437–Cys542) was possibly stabilized by introduction of proline at the position 540 (Ala540Pro) on the loop (Fig. S4, ESI†).

In previous applications of PROSS, it was noted that large positively charged patches, which could promote aggregation in the nucleic acid rich micro environment,[37] were eliminated or reduced.[60,61] Here, in the case of HPSE, the electrostatic potential surface of HPSE WT shows two large positive patches around the active site and the β-sandwich domain (Fig. 3B). For HPSE P6, the theoretical isoelectric point was the same as the wild type (pI 9.4), and the electrostatic surface potential shows that while one of the large positive patches around the β-sandwich domain was slightly diminished by two lysine mutations (Lys427Asp and Lys477Gln), the electrostatic potential around the active site at the TIM face was maintained as the area was constrained during the design process (Fig. 3B).

Molecular dynamic simulations to account stabilization

It has previously been shown that the dynamics and function of similar proteins can be very different despite ground state structures appearing very similar in terms of Cα RMSD.[62] Crystallographic B-factors are commonly used to probe differences in the conformational flexibility of proteins within a crystal lattice, although this approach can be limited by the existence of crystallographic artifacts, whereby flexible regions on the protein surface could be stabilized by interactions with the lattice symmetry mates. Comparison between the B-factors of HPSE WT and HPSE P6 reveal the overall trend in terms of regions with high or low B-factors is conserved, although a decrease in the overall B factors of the P6 variant in the TIM (β/α)$_8$ domain fluctuations, mostly around the surface loops of the active site (Fig. S4A, ESI†). However, this analysis is confounded by the higher resolution, lower Wilson B-factor and different crystal packing of the HPSE P6 variant.
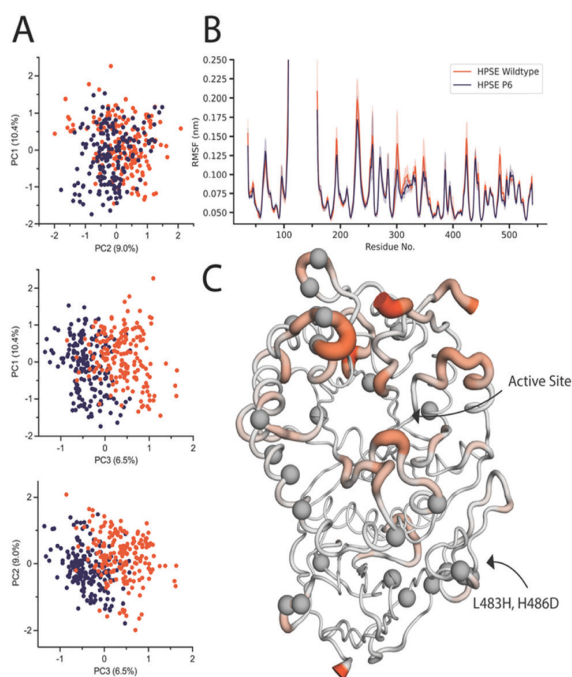

Fig. 4 Molecular dynamic simulations of the wild type the pross design (A) PCA of the two proteins comparing PC1, PC2 and PC3, showing PC3 to have slight differences between the two proteins. (B) RMSF plot of the 95% CI of the wild type, and the mutant average RMSF showing that the mutant stays within the 95% CI, suggesting similar fluctuations. (C) HPSE P6 with average RMSF overlaid. Large RMSF is represented in orange putty, mostly seen around the active site. Mutations are represented as grey spheres.

To complement the structural analysis, we also performed molecular dynamics simulations to examine the effects of these mutations on the conformational sampling and motions of the protein. To identify whether the dynamic range of HPSE P6 is the same as the HPSE WT, a total simulation time of 1 μs per protein was completed. Principal component analysis was conducted to visualize motions that represent the major fluctuations of the system. Principal components 1 and 2 of the HPSE WT and HPSE P6 (10.4% and 9.0%) overlap, demonstrating that the breathing motion of the active site is conserved (Fig. 4A). The third major component, which only contributes 6.5% of the total movement of the protein, shows slight differences, being comprised predominately of the movement of surface-exposed loops. No other principal component showed any difference between the two proteins (up to 20 components).

Root mean square fluctuations were also analysed to identify the displacement of amino acids throughout the course of the simulation (Fig. 4B). The average RMSF and their 95% confidence intervals, (where 95% of the residue's displacement occurs in that region) are overlaid for both proteins. This demonstrates that HPSE WT and P6 fluctuations overlap closely. There were very few differences overall, where the most consistent change is a decrease in magnitude of surface loops for HPSE P6 in comparison to WT simulations. Even though these residues have a very slightly decreased magnitude, the RMSF still has the same overall shape.

The only residues with RMSF values outside of the 95% CI are residues 488–495. This is a surface loop on the β-sandwich domain with two introduced mutations; Leu483His and His486Asp. These mutations allow an increase of hydrogen bonding causing a slight rigidification of this loop (Fig. 4C). Overall, the conservation of the protein dynamics despite 26 mutations is striking and unexpected. Indeed, the great majority of these mutations (identified as grey spheres in Fig. 4C) do not cause any significant difference on the dynamics of the protein. This analysis is also fully consistent with the functional data, which revealed almost to effect of the mutations on activity or inhibition.

Discussion

Despite the widespread interest in HPSE as an important enzyme in human physiology and a drug target, the difficulty related to obtaining large quantities of pure recombinant protein has limited the ability of many groups to study the protein. Bacterial expression systems, such as *E. coli*, are widely accessible, allow protein to be expressed in high yields and at low cost. While prokaryotic HPSE expression methods have been reported,[33,34] none have been repeated in the literature or have been widely adopted. Here, a stable version of human HPSE has been computationally designed, which allowed the mature heterodimeric enzyme to be expressed at reasonable yield in soluble form in *E.coli*. The subsequent characterization of this designed version (HPSE P6) showed the enzyme to behave essentially identically to HPSE WT in terms of its interactions with substrates (Table 1) and inhibitors (Fig. 1). Thus, this computationally designed HPSE P6 variant should be a useful surrogate for the wild-type enzyme in structural biology, inhibitor screening and kinetic analyses.

It is notable that despite 26 mutations, the enzyme is essentially structurally isomorphous to the wild-type, with no significant changes to the C-α backbone or side chain rotamer sampling, especially in the vicinity of the active site. Additionally, the dynamics of the enzyme were also largely identical to the wild-type enzyme, suggesting that there were no significant changes to the relative stabilities of different conformational substates. This reinforces the functional neutrality of many mutations and the power of bioinformatics inspired approaches such as consensus design and PROSS; these mutations were acquired through phylogenetic analysis *i.e.*, they are known to be tolerated in related enzymes. Indeed, while their individual effects might be small, the summation of the effects can become considerable. However, the route to HPSE P6 was not simple or trivial; P6 was the only design of the seven that we tested that was effective. Thus, while the combinatorial effects of the mutations can be powerful, the unpredictable effects of the mutations, and their epistatic interactions, make it imperative that a range of designs are trialled.

Our structural analysis of HPSE P6 shows that many of the mutations appear to have effects that can be rationalised in terms of our understanding of how proteins fold: increasing surface polarity, forming additional non-covalent interactions such as hydrogen bonds, increased packing within the hydrophobic core, *etc.* The lack of major structural changes, such as strong salt bridges or significant changes to internal cavities, which are characteristic of rational or computationally designed stabilising mutations, meant that the structural dynamics of the protein, and thus its catalytic function, was largely unperturbed. This study is thus an interesting example of protein stabilisation: on the one hand, 26 mutations could be considered to be a large number of mutations, but the counter argument is that 26 functionally neutral mutations that have almost no effect on the structure and dynamics is in fact a very conservative method for stabilizing a protein, in comparison to a smaller number of mutations that might have a larger effect on the structure, dynamics and function of the enzyme.

Experimental

Stability design

Chain A of the crystal structure of the ligand bound human HPSE (PDB ID: 5E9C) was submitted to the PROSS stability design algorithm[46] on the web server (http://pross.weizmann.ac.il), with constrained residues, which have contacts with the ligand (Dp4) and with the chain B. This generated 7 mutants.

Cloning

The linear 8 kDa (Gln36–Glu109) and 50 kDa (Lys158–Ile543) subunits of the human HPSE were *E. coli* codon optimized and synthesized by IDT (Australia). The seven PROSS designs were *E. coli* codon optimized and synthesized by Twist bioscience. The 8 kDa subunit was amplified and sub-cloned into the multiple cloning site 1 of the linearized pETDuet-1 vector (Novagen) through the BamHI and NotI restriction sites (Fast Digest,Thermo) by Gibson one-step isothermal assembly.[63] The resulting plasmid DNA was linearized using NdeI and XhoI restriction enzymes (Fast Digest, Thermo) and designs were inserted into the multiple cloning site 2 by Gibson assembly.[63] The ligated DNA was transformed to *E. coli* TOP10 cells and the plasmid DNA was extracted and sent to Garvan Institute (Australia) for Sanger sequencing to confirm the sequences.

Protein expression and purification

The wild type and the 7 designs were transformed in *E. coli* Shuffle T7 Express cells (NEB), together with different combinations of chaperones in a pACYC vector and spread on an Agar plate with ampicillin and chloramphenicol. 1% overnight seed culture from a single colony was inoculated into 1 L of LB medium supplemented with ampicillin (100 mg L^{-1}) and chloramphenicol (34 mg L^{-1}), then incubated at 37 °C for 5 hours. Overexpression was induced by adding IPTG to a final concentration of 0.05 mM and the culture was further incubated for 3 hours at 37 °C. The cell pellets were resuspended in buffer A (20 mM HEPES pH 8, 300 mM NaCl, 5 mM β-mercaptoethanol, 10% glycerol, 0.05% Tween, 20 mM Imidazole) with Turbonuclease

(Sigma) and lysed by sonication (Omni Sonic Ruptor 400 Ultrasonic homogenizer). The lysate was filtered (0.45 μm) and loaded onto Ni-NTA column (GE healthcare) and eluted with 100% buffer B (buffer A + 500 mM Imidazole). The peak eluent was diluted 5 times with buffer C (20 mM HEPES pH 7.4, 200 mM NaCl, 1 mM DTT, 10% glycerol, 0.05% Tween20) and loaded to heparin affinity column (GE healthcare) and eluted with 100% buffer D (buffer C + 1.5 M NaCl). The peak eluent was loaded onto a size exclusion column (HiLoad 26/600 Superdex 200 pg, GE healthcare) and eluted in a buffer E (20 mM HEPES pH 7.4, 200 mM NaCl, 10% glycerol, 1 mM DTT, 0.05% Tween20). The final concentration of the monomeric heparinase from the gel filtration was estimated by absorbance at 280 nm using NanoDrop One (Thermo) and the yield was more than 2 mg per litre of LB culture.

Colorimetric assay using fondaparinux

Assays were conducted using the colorimetric assay designed by Hammond *et al.*[56] Bovine serum albumin-coated 96 well microplates were used for all assays and were prepared by incubation of the plates with 1% BSA dissolved in phosphate-buffered saline (PBS) with 0.05% Tween-20 (PBST) at 37 °C for 75 minutes. The plates were then washed three times with PBST, dried and stored at 4 °C. Assay mixtures contain 40 mM sodium acetate buffer (pH 5.0), 0.8 nM HPSE in 0.01% Tween 20 sodium acetate buffer and 100 mM fondaparinux (GlaxoSmithKline) with or without increasing concentrations of inhibitor. Plates were incubated at 37 °C for 2–20 hours before the reaction was stopped with 100 μL of 1.69 mM 4-[3-(4-iodophenyl)-2-(4-nitrophenyl)-2*H*-5-tetrazolio]-1,3d-benzene disulfonate (WST-1) in 0.1 M NaOH. The plates were resealed and developed at 60 °C for 60 minutes, and the absorbance was measured at 584 nm. Kinetics were carried out with a standard curve constructed with D-galactose as the reducing sugar standard, prepared in the same buffer and volume over the range of 0–2 μM. All curve fitting to calculate IC_{50} values and Michaelis–Menten constants, was done using GraphPad Prism software (version 8.1).

Circular dichroism analysis

The size exclusion fraction was directly used to measure the CD using the Chirascan CD spectrometer (Applied Photophysics). The thermal stability of the protein (0.15 mg mL^{-1}) was measured in a range of temperatures 20–90 °C by monitoring the ellipticity at 222 nm using a cuvette with 1 mm path length. Data analysis was performed using GraphPad Prism, within which the mid-point of the melting curve was calculated using Boltzmann sigmoid equation.

Structure determination and refinement

Well diffracted single crystals were obtained by the hanging-drop vapor-diffusion method at 18 °C by combining the protein (6–8 mg mL^{-1}) and the well solution (1.9 M $(NH_4)_2SO_4$) with a ratio of 1.5 : 1.5 μL. Crystals appeared within a week and continued to grow for 1–2 months. The crystal was cryoprotected with additional 30% glycerol to the mother liquor before flash freezing in liquid nitrogen. Crystallographic data were collected

at 100 K at the Australian Synchrotron (MX2,[64] 0.9537 Å). The obtained diffraction data were indexed and integrated with XDS.[65] Resolution estimation and data truncation were performed using aimless program in CCP4[66] on the basis of the datasets overall half-dataset correlation, a $CC_{1/2}$ value of 0.3.[67] All structures were solved by molecular replacement using the Molrep program in CCP4[66] using the structure deposited under PDB accession code 5E9M as a starting model. The models were refined using phenix.refine,[68] and the model was subsequently optimized by iterative model building with the program COOT v0.8.[69] The alternative conformations were modelled based on mF_o–DF_c density and the occupancies and B-factors were determined using phenix.refine.[68] The structures were then evaluated using MolProbity[70] in Phenix. Details of the refinement statistics were produced by Phenix v1.17[71] and summarized in Table S3 (ESI†). The structures were visualized and analysed using PyMol v2.3[72] or Maestro,[73] whereby APBS[58] program in PyMol was used to calculated the electrostatic potential and protein alignment program in Maestro was used to calculate the Cα RMSD.

Molecular dynamics

Molecular dynamic simulations were performed using the GROMACS 2018.8 engine with parameters from the Charm22* force field.[74,75] All chain termini were capped with neutral acetyl or methylamide groups. Protonation states were assigned with the PDB2PQR server for pH 5.0.[58] Completed structures were solvated with a TIP3P water model[76] using a rhombic dodecahedron simulation box with a minimum distance of 12 Å between the protein and simulation box, followed by the addition of 200 mM NaCl to the aqueous phase and sufficient ions to neutralise the system charge. Simulation systems of WT and PROSS 6 were relaxed using the standard steepest descent minimization using at least 10 000 steps before being equilibrated for 1 ns in the isothermal–isobaric (*NPT*) ensemble to stabilize the system. Ten replicates of each system were simulated for 100 ns under *NPT*. Periodic boundary conditions were used, and long-range electrostatics were calculated using the particle-mesh Ewald method with a cutoff of 1.2 nm.[77] Non-bonded interactions were evaluated using a Verlet cut-off scheme. The temperature in all simulations was set to 300 K and controlled *via* the Bussi–Donadio–Parrinello stochastic velocity rescaling thermostat;[78] the initial velocities of all particles were pseudo-randomly generated. Pressure coupling was handled with the Berendsen barostat during equilibration and the Parrinello–Rahman barostat for production.[79,80] The LINCS (Linear Constraint Solver) algorithm was used to constrain bonds involving hydrogen in conjunction with an integration time step of 2 fs.[81] Constraints were applied to the starting configuration of the production run. Analyses of simulations were preformed using the tools provided in the GROMACS package. Data was collected from the last 90 ns of each production simulation, as RMSF had stabilised by this time.

Principal component analysis

Principal component analysis was performed using the MDTraj python library and the scikit-learn machine learning tool.[82,83]

Using the aligned and concatenated trajectory, a merged dataset was created, from which the WT and P6 systems were projected. Data was plotted in Graphpad prism.

Data availability

Coordinates and structure factors have been deposited in the Protein Data Bank under accession code PDB 7RG8.

Conclusions

This study describes the production of a new variant of HPSE, which is functionally identical to the wild-type protein in terms of activity, inhibition, structure and dynamics that is easily expressed in *E. coli* and crystallises within a day, yielding high resolution crystals. This protein should make the study of HPSE function and the development of inhibitors significantly easier and less expensive. It is notable that the large number of mutations in HPSE P6 were functionally neutral. This contributes to a growing understanding of the relationship between protein sequence and folding, where evolutionarily conserved and functionally neutral consensus-like mutations can be understood to significantly affect the efficiency of protein folding and expression and protein thermostability.

Author contributions

C. W., N. H. and C. J. J. conceived the study and analysed data. C. W designed variants. N. H. conducted cloning and expression trials. C. W. expressed, purified and crystallised HPSE P6. C. W and N. H. processed crystallography data. N. H. analysed crystal structures. C. W. and J. A. M. performed and analysed molecular dynamics simulations and principal component analysis. C. W., N. H. and C. J. J. wrote the manuscript. All authors provided intellectual input and edited and approved the final manuscript.

Conflicts of interest

CJ has received funding from Beta Therapeutics to work on heparanase inhibitors.

Acknowledgements

References

1 S. Rivara, F. M. Milazzo and G. Giannini, *Future Med. Chem.*, 2016, **8**, 647–680.

2 J. R. Bishop, M. Schuksz and J. D. Esko, *Nature*, 2007, **446**, 1030–1037.

3 D. Aviezer and A. Yayon, *Proc. Natl. Acad. Sci. U. S. A.*, 1994, **91**, 12173–12177.

4 T. F. Zioncheck, L. Richardson, J. Liu, L. Chang, K. L. King, G. L. Bennett, P. Fugedi, S. M. Chamow, R. H. Schwall and R. J. Stack, *J. Biol. Chem.*, 1995, **270**, 16871–16878.

5 A. Amara, O. Lorthioir, A. Valenzuela, A. Magerus, M. Thelen, M. Montes, J. L. Virelizier, M. Delepierre, F. Baleux, H. Lortat-Jacob and F. Arenzana-Seisdedos, *J. Biol. Chem.*, 1999, **274**, 23916–23925.

6 S. Eisenberg, E. Sehayek, T. Olivecrona and I. Vlodavsky, *J. Clin. Invest.*, 1992, **90**, 2013–2021.

7 I. Vlodavsky, Y. Friedmann, M. Elkin, H. Aingorn, R. Atzmon, R. Ishai-Michaeli, M. Bitan, O. Pappo, T. Peretz, I. Michal, L. Spector and I. Pecker, *Nat. Med.*, 1999, **5**, 793–802.

8 M. D. Hulett, C. Freeman, B. J. Hamdorf, R. T. Baker, M. J. Harris and C. R. Parish, *Nat. Med.*, 1999, **5**, 803.

9 P. H. Kussie, J. D. Hulmes, D. L. Ludwig, S. Patel, E. C. Navarro, A. P. Seddon, N. A. Giorgio and P. Bohlen, *Biochem. Biophys. Res. Commun.*, 1999, **261**, 183–187.

10 M. Toyoshima and M. Nakajima, *J. Biol. Chem.*, 1999, **274**, 24153–24160.

11 R. Goshen, A. A. Hochberg, G. Korner, E. Levy, R. Ishai-Michaeli, M. Elkin, N. De Groot and I. Vlodavsky, *Mol. Hum. Reprod.*, 1996, **2**, 679–684.

12 L. Gutter-Kapon, D. Alishekevitz, Y. Shaked, J. P. Li, A. Aronheim, N. Ilan and I. Vlodavsky, *Proc. Natl. Acad. Sci. U. S. A.*, 2016, **113**, E7808–E7817.

13 I. Vlodavsky, A. Eldor, A. Haimovitz-Friedman and Y. Matzner, *Invasion and metastasis*.

14 B. Buijsers, C. Yanginlar, A. de Nooijer, I. Grondman, M. L. Maciej-Hulme, I. Jonkman, N. A. F. Janssen, N. Rother, M. de Graaf, P. Pickkers, M. Kox, L. A. B. Joosten, T. Nijenhuis, M. G. Netea, L. Hilbrands, F. L. van de Veerdonk, R. Duivenvoorden, Q. de Mast and J. van der Vlag, *Front. Immunol.*, 2020, **11**, DOI: 10.3389/FIMMU.2020.575047.

15 I. Vlodavsky, Z. Fuks, M. Bar-Ner, Y. Ariav and V. Schirrmacher, *Cancer Res.*, 1983, **43**(6), 2704–2711.

16 M. Nakajima, T. Irimura, D. Di Ferrante, N. Di Ferrante and G. L. Nicolson, *Science*, 1983, **220**, 611–613.

17 M. Waterman, O. Ben-Izhak, R. Eliakim, G. Groisman, I. Vlodavsky and N. Ilan, *Mod. Pathol.*, 2007, **20**, 8–14.

18 I. Vlodavsky, R. V. Iozzo and R. D. Sanderson, *Matrix Biol.*, 2013, **32**, 220–222.

19 I. Vlodavsky, P. Beckhove, I. Lerner, C. Pisano, A. Meirovitz, N. Ilan and M. Elkin, *Cancer Microenviron.*, 2012, **5**, 115–132.

20 C. P. Baburajeev, C. D. Mohan, S. Rangappa, D. J. Mason, J. E. Fuchs, A. Bender, U. Barash, I. Vlodavsky, Basappa and K. S. Rangappa, *BMC Cancer*, 2017, **17**, 235.

21 E. A. McKenzie, *Br. J. Pharmacol.*, 2007, **151**, 1–14.

© 2022 The Author(s). Published by the Royal Society of Chemistry

RSC Chem. Biol., 2022, **3**, 341–349 | **347**

22 I. Vlodavsky, O. Goldshmidt, E. Zcharia, R. Atzmon, Z. Rangini-Guatta, M. Elkin, T. Peretz and Y. Friedmann, *Semin. Cancer Biol.*, 2002, **12**, 121–129.

23 F. Sanchez-Muñoz, A. Dominguez-Lopez and J. K. Yamamoto-Furusho, *World J. Gastroenterol. WJG*, 2008, **14**, 4280.

24 L. M. Coussens, B. Fingleton and L. M. Matrisian, *Science*, 2002, **295**, 2387–2392.

25 E. Zcharia, J. Jia, X. Zhang, L. Baraz, U. Lindahl, T. Peretz, I. Vlodavsky and J.-P. Li, *PLoS One*, 2009, **4**, e5181.

26 C. D. Mohan, S. Hari, H. D. Preetham, S. Rangappa, U. Barash, N. Ilan, S. C. Nayak, V. K. Gupta, Basappa, I. Vlodavsky and K. S. Rangappa, *iScience*, 2019, **15**, 360–390.

27 V. C. Ramani, F. Zhan, J. He, P. Barbieri, A. Noseda, G. Tricot and R. D. Sanderson, *Oncotarget*, 2016, **7**, 1598.

28 L. Jia and S. Ma, *Eur. J. Med. Chem.*, 2016, **121**, 209–220.

29 V. Masola, G. Zaza, G. Gambaro, M. Franchi and M. Onisto, *Semin. Cancer Biol.*, 2020, **62**, 86–98.

30 M. Toyoshima and M. Nakajima, *J. Biol. Chem.*, 1999, **274**, 24153–24160.

31 L. Wu, C. M. Viola, A. M. Brzozowski and G. J. Davies, *Nat. Struct. Mol. Biol.*, 2015, **22**, 1016–1022.

32 L. Wu, J. Jiang, Y. Jin, W. W. Kallemeijn, C.-L. Kuo, M. Artola, W. Dai, C. van Elk, M. van Eijk, G. A. van der Marel, J. D. C. Codée, B. I. Florea, J. M. F. G. Aerts, H. S. Overkleeft and G. J. Davies, *Nat. Chem. Biol.*, 2017, **13**, 867–873.

33 S. Winkler, D. Schweiger, Z. Wei, E. Rajkovic and A. J. Kungl, *Carbohydr. Res.*, 2014, **389**, 72–77.

34 A. Pennacchio, A. Capo, S. Caira, A. Tramice, A. Varriale, M. Staiano and S. D'Auria, *Biotechnol. Appl. Biochem.*, 2018, **65**, 89–98.

35 J. Warwicker, S. Charonis and R. A. Curtis, *Mol. Pharm.*, 2013, **11**, 294–303.

36 R. M. Kramer, V. R. Shende, N. Motl, C. N. Pace and J. M. Scholtz, *Biophys. J.*, 2012, **102**, 1907–1915.

37 P. Chan, R. A. Curtis and J. Warwicker, *Sci. Rep.*, 2013, **3**, 3333.

38 S. Simizu, K. Ishida, M. K. Wierzba and H. Osada, *J. Biol. Chem.*, 2004, **279**, 2697–2703.

39 F. Levy-Adam, H. Q. Miao, R. L. Heinrikson, I. Vlodavsky and N. Ilan, *Biochem. Biophys. Res. Commun.*, 2003, **308**, 885–891.

40 D. Li, A. M. Damry, J. R. Petrie, T. Vanhercke, S. P. Singh and C. J. Jackson, *Biochemistry*, 2020, **59**, 1398–1409.

41 N. Amin, A. D. Liu, S. Ramer, W. Aehle, D. Meijer, M. Metin, S. Wong, P. Gualfetti and V. Schellenberger, *Protein Eng., Des. Sel.*, 2004, **17**, 787–793.

42 J. H. Whitfield, W. H. Zhang, M. K. Herde, B. E. Clifton, J. Radziejewski, H. Janovjak, C. Henneberger and C. J. Jackson, *Protein Sci.*, 2015, **24**, 1412.

43 M. A. Spence, J. A. Kaczmarski, J. W. Saunders and C. J. Jackson, *Curr. Opin. Struct. Biol.*, 2021, **69**, 131–141.

44 B. Borgo and J. J. Havranek, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 1494–1499.

45 J. Schymkowitz, J. Borg, F. Stricher, R. Nys, F. Rousseau and L. Serrano, *Nucleic Acids Res.*, 2005, **33**(2), 382–388.

46 A. Goldenzweig, M. Goldsmith, S. E. Hill, O. Gertman, P. Laurino, Y. Ashani, O. Dym, T. Unger, S. Albeck, J. Prilusky, R. L. Lieberman, A. Aharoni, I. Silman, J. L. Sussman, D. S. Tawfik and S. J. Fleishman, *Mol. Cell*, 2016, **63**, 337–346.

47 J. W. Lamppa, S. A. Tanyos and K. E. Griswold, *J. Biotechnol.*, 2013, **164**, 1–8.

48 F. Georgescauld, K. Popova, A. J. Gupta, A. Bracher, J. R. Engen, M. Hayer-Hartl and F. U. Hartl, *Cell*, 2014, **157**, 922–934.

49 S. Haldar, R. Tapia-Rojo, E. C. Eckels, J. Valle-Orero and J. M. Fernandez, *Nat. Commun.*, 2017, **8**, 668.

50 J. Lobstein, C. A. Emrich, C. Jeans, M. Faulkner, P. Riggs and M. Berkmen, *Microb. Cell Fact.*, 2012, **11**, 56.

51 X. Brazzolotto, A. Igert, V. Guillon, G. Santoni and F. Nachon, *Molecules*, 2017, **22**, 1828.

52 O. Khersonsky, R. Lipsh, Z. Avizemer, Y. Ashani, M. Goldsmith, H. Leader, O. Dym, S. Rogotner, D. L. Trudeau, J. Prilusky, P. Amengual-Rigo, V. Guallar, D. S. Tawfik and S. J. Fleishman, *Mol. Cell*, 2018, **72**, 178–186.e5.

53 I. Campeotto, A. Goldenzweig, J. Davey, L. Barfod, J. M. Marshall, S. E. Silk, K. E. Wright, S. J. Draper, M. K. Higgins and S. J. Fleishman, *Proc. Natl. Acad. Sci. U. S. A.*, 2017, **114**, 998–1002.

54 M. Sternke, K. W. Tripp and D. Barrick, *Proc. Natl. Acad. Sci. U. S. A.*, 2019, **116**, 11275–11284.

55 T. A. Whitehead, A. Chevalier, Y. Song, C. Dreyfus, S. J. Fleishman, C. De Mattos, C. A. Myers, H. Kamisetty, P. Blair, I. A. Wilson and D. Baker, *Nat. Biotechnol.*, 2012, **30**, 543–548.

56 E. Hammond, C. P. Li and V. Ferro, *Anal. Biochem.*, 2010, **396**, 112–116.

57 C. Freeman and R. C. Parish, *Biochem. J.*, 1998, **330**, 1341–1350.

58 E. Jurrus, D. Engel, K. Star, K. Monson, J. Brandi, L. E. Felberg, D. H. Brookes, L. Wilson, J. Chen and K. Liles, *Protein Sci.*, 2018, **27**, 112–128.

59 G. Raghunathan, S. Sokalingam, N. Soundrarajan, B. Madan, G. Munussami and S.-G. Lee, *Mol. BioSyst.*, 2013, **9**, 2379–2389.

60 I. Campeotto, A. Goldenzweig, J. Davey, L. Barfod, J. M. Marshall, S. E. Silk, K. E. Wright, S. J. Draper, M. K. Higgins and S. J. Fleishman, *Proc. Natl. Acad. Sci. U. S. A.*, 2017, **114**, 998–1002.

61 A. R. Lambert, J. P. Hallinan, R. Werther, D. Głów and B. L. Stoddard, *Structure*, 2020, **28**, 760–775.e8.

62 E. Campbell, M. Kaltenbach, G. J. Correy, P. D. Carr, B. T. Porebski, E. K. Livingstone, L. Afriat-Jurnou, A. M. Buckle, M. Weik, F. Hollfelder, N. Tokuriki and C. J. Jackson, *Nat. Chem. Biol.*, 2016, **12**, 944–950.

63 D. G. Gibson, in *Synthetic Biology, Part B*, ed. C. B. T.-M. in E. Voigt, Academic Press, 2011, vol. 498, pp. 349–361.

64 D. Aragão, J. Aishima, H. Cherukuvada, R. Clarken, M. Clift, N. P. Cowieson, D. J. Ericsson, C. L. Gee, S. Macedo, N. Mudie, S. Panjikar, J. R. Price, A. Riboldi-Tunnicliffe,

R. Rostan, R. Williamson and T. T. Caradoc-Davies, *J Synchrotron Radiat.*, 2018, **25**, 885–891.

65 W. Kabsch, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2010, **66**, 125–132.

66 M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. W. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin and K. S. Wilson, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2011, **67**, 235–242.

67 P. A. Karplus and K. Diederichs, *Science*, 2012, **336**, 1030–1033.

68 P. V. Afonine, R. W. Grosse-Kunstleve, N. Echols, J. J. Headd, N. W. Moriarty, M. Mustyakimov, T. C. Terwilliger, A. Urzhumtsev, P. H. Zwart and P. D. Adams, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2012, **68**, 352–367.

69 P. Emsley and K. Cowtan, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2004, **60**, 2126–2132.

70 V. B. Chen, W. B. Arendall, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson and D. C. Richardson, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2010, **66**, 12–21.

71 P. V. Afonine, R. W. Grosse-Kunstleve, V. B. Chen, J. J. Headd, N. W. Moriarty, J. S. Richardson, D. C. Richardson, A. Urzhumtsev, P. H. Zwart and P. D. Adams, *J. Appl. Crystallogr.*, 2010, **43**, 669–676.

72 *The PyMOL Molecular Graphics System, Version 2.0 Schrödinger*, LLC.

73 Schrodinger Release 2020-3, *Maestro,Schrödinger*, LLC, New York, NY, 2020.

74 M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess and E. Lindah, *SoftwareX*, 2015, **1–2**, 19–25.

75 S. Piana, K. Lindorff-Larsen and D. E. Shaw, *Biophys. J.*, 2011, **100**, L47–L49.

76 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926–935.

77 U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee and L. G. Pedersen, *J. Chem. Phys.*, 1995, **103**, 8577–8593.

78 G. Bussi, D. Donadio and M. Parrinello, *J. Chem. Phys.*, 2007, **126**, 014101.

79 H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola and J. R. Haak, *J. Chem. Phys.*, 1984, **81**, 3684.

80 M. Parrinello and A. Rahman, *Phys. Rev. Lett.*, 1980, **45**, 1196–1199.

81 B. Hess, *J. Chem. Theory Comput.*, 2008, **4**, 116–122.

82 R. T. McGibbon, K. A. Beauchamp, M. P. Harrigan, C. Klein, J. M. Swails, C. X. Hernández, C. R. Schwantes, L. P. Wang, T. J. Lane and V. S. Pande, *Biophys. J.*, 2015, **109**, 1528–1532.

83 F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot and É. Duchesnay, *J. Mach. Learn. Res.*, 2011, **12**, 2825–2830.

© 2022 The Author(s). Published by the Royal Society of Chemistry

RSC Chem. Biol., 2022, **3**, 341–349 | **349**

Computational design and experimental characterisation of a stable human heparanase variant

Cassidy Whitefield[a†], Nansook Hong [a†], Joshua A. Mitchell[a] and Colin J. Jackson [a*]

a. Research School of Chemistry, Australian National University, Canberra, ACT, 2601, Australia.
† Contributed equally to this work. *To whom correspondence should be addressed: colin.jackson@anu.edu.au
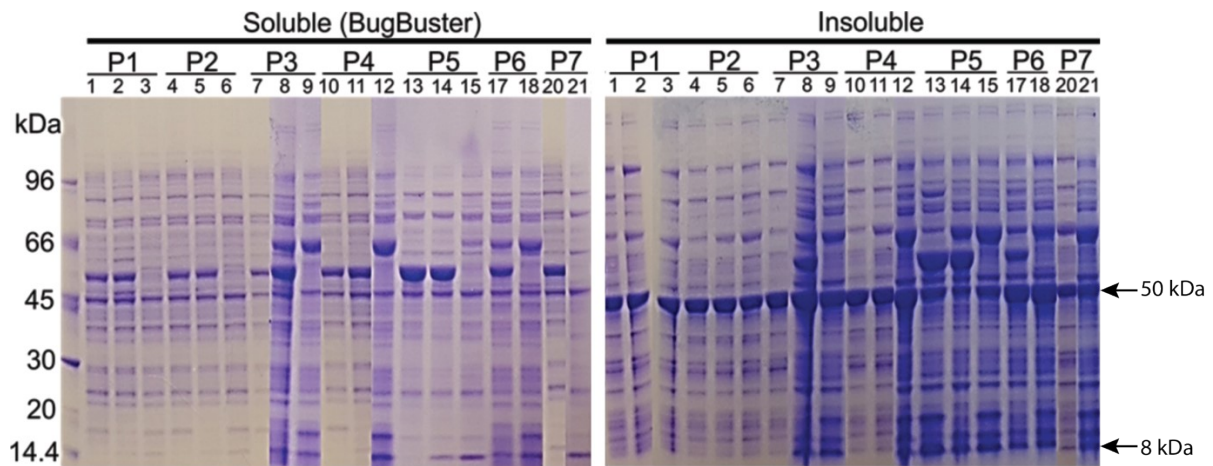
Supplementary Material
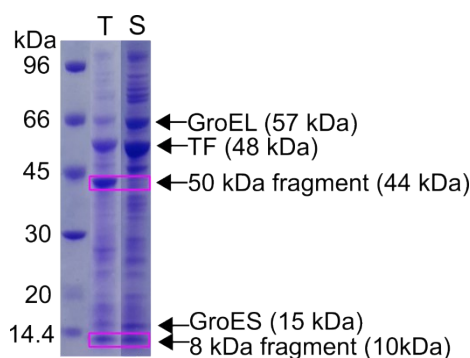
Supplementary Figure 1. Wild type human heparanase expression in *E. coli*. The 50 and 8 kDa subunits and chaperones[28] in total (annotated as T) and in soluble fraction (annotated as S) are shown with LMW protein marker (GE healthcare). 8 kDa subunit with N-terminal 6xHis-tag and 50 kDa subunit were amplified by PCR using gBlocks (IDT, **Supplementary Table 1**) and inserted into multiple cloning sites of pETDuet-1 coexpression vector (Novagen). The plasmid DNA was transformed into Shuffle T7 Express competent cells (NEB) together with chaperones in pACYC vector. Overnight seed culture from single colony was inoculated by 1% into LB media supplemented with ampicillin and chloramphenicol antibiotics. The culture was incubated at 37 °C for about 3-5 hours until the OD was about 0.8. 0.4 mM IPTG was added and further incubated for about 3 hours at 37 °C. The cell pellet was resuspended in 50 mM acetate buffer at pH 5 and lysed using BugBuster protein extraction reagent (Merck) for the solubility measurement.
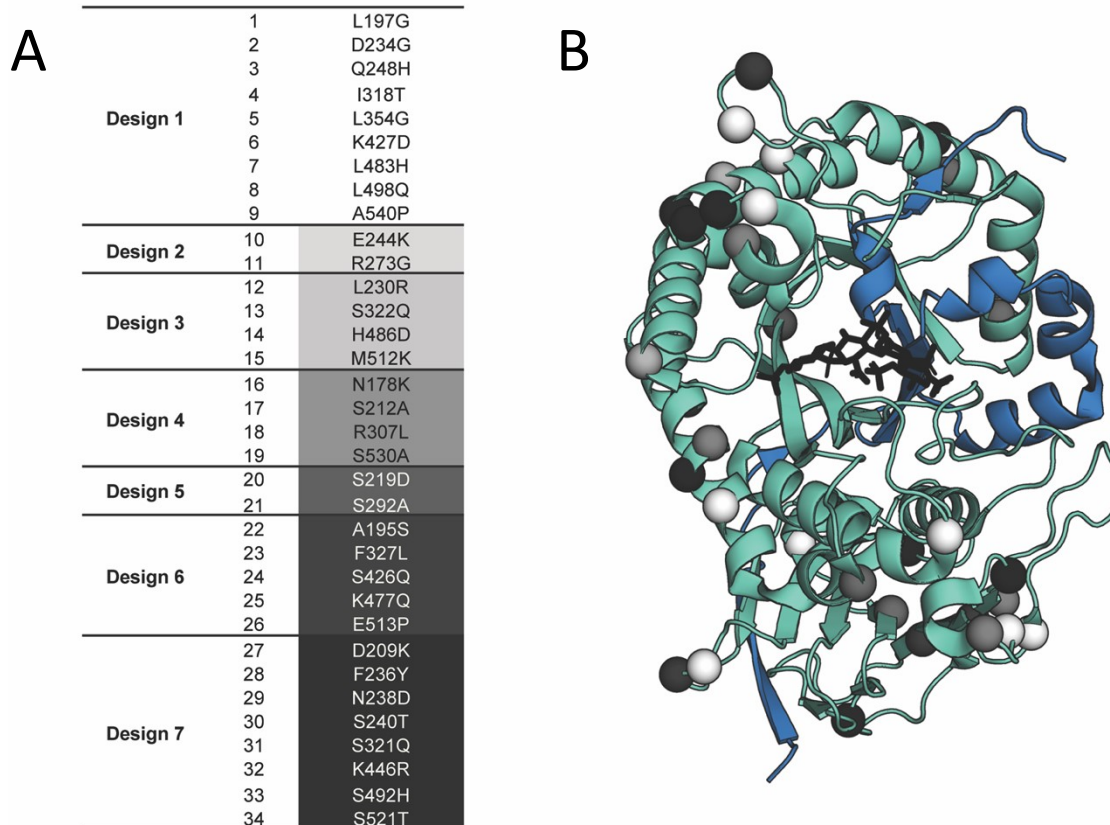
Supplementary Figure 2. Design of stable heparanase mutants from PROSS (**A**) the resulting mutations are shown for each design. (**B**) The crystal structure of heparanase (PDB ID: 5E9C) was used as template whereby 50 kDa subunit (cyan cartoon) was targeted with restriction on the residues contacting the ligand (black sticks) and the 8 kDa subunit (blue cartoon).

Supplementary Figure 3. Pross variant expression trials in *E. coli* SHuffle cells. P1-P7 plasmid DNA was transformed into Shuffle T7 Express competent cells (NEB) together with varying chaperones in pACYC vector. Skp+TF was tested in 1, 4, 7, 10 and 13. GroEL/ES+TF was tested in 2, 5, 8, 11, 14, 17, and 20. Skp + GroEL/ES was tetesd in 3, 6, 9, 12, 15, 18 and 21. LMW protein marker (GE healthcare) is shown on the left. Overnight seed culture from single colony was inoculated by 1% into LB media supplemented with ampicillin and chloramphenicol antibiotics. The culture was incubated at 37 °C for about 3-5 hours until the OD was about 0.8. 0.4 mM IPTG was added and further incubated for about 3 hours. The cell pellet was resuspended in 50 mM acetate buffer at pH 5 and lysed using BugBuster protein extraction reagent (Merck) for the solubility measurement.
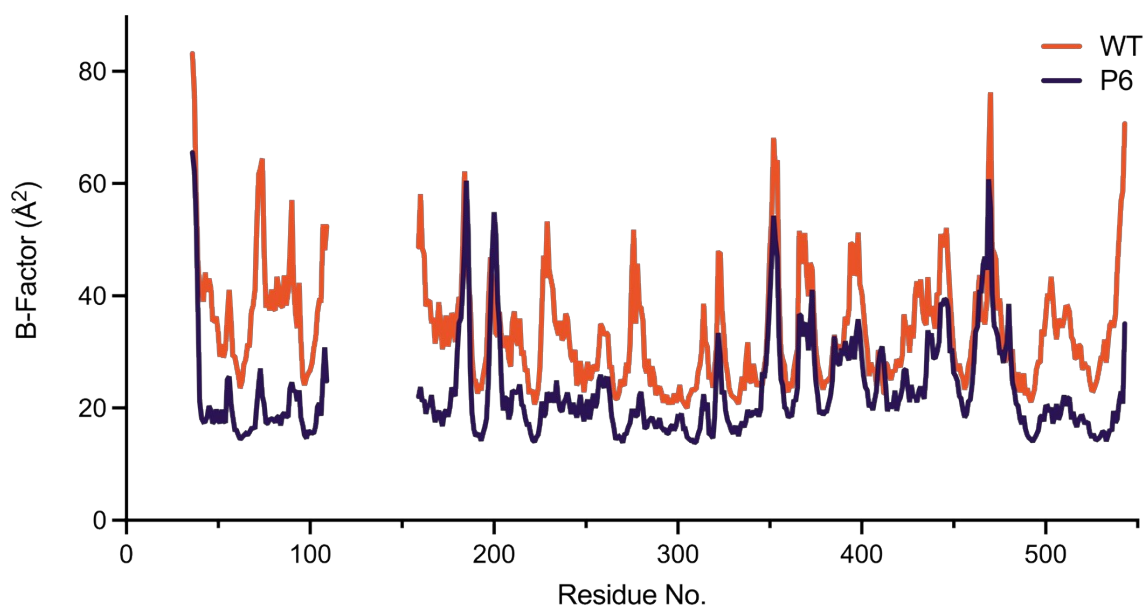
Supplementary Figure 4: Crystallographic B-Factors of WT HSPE and HSPE P6. Comparison between the B factors from the two structures (with different Wilson B-factors and resolution) show an overall decrease in magnitude of the B factors in the Mutant P6, yet the overall trend in terms of regions of high/low B-factors is very similar.
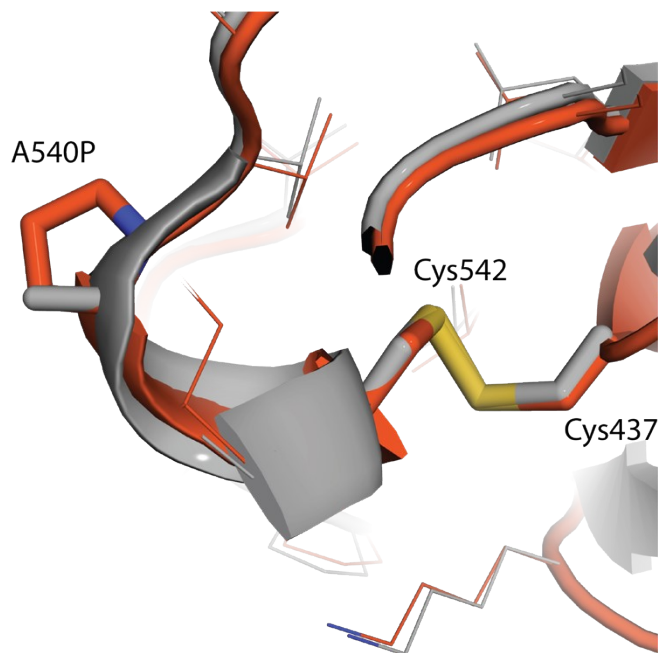
Supplementary Figure 5: The disulfide bond between Cys542 and Cys437 that is stabilised by A540P mutation, which is in a favourable conformation is the kinked loop preceding Cys542.
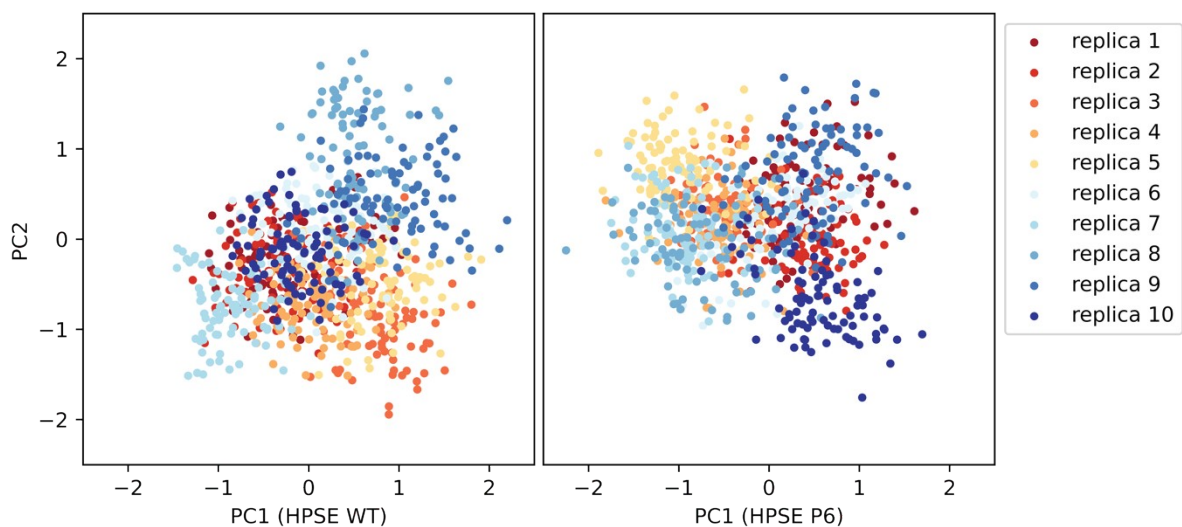
Supplementary Figure 6: Principal component analysis (PCA) sampling quality for WT and P6 simulations. Sampling shows that the 10 simulations sample a large range of conformational space for both WT and mutant replica simulations, and using all replica simulations is viable in the analysis.

Supplementary Table 1. Nucleotide sequences of heparanase variants used for cloning

Name	Sequences
8 kDa subunit of wild type (gBlock)	CAGGATGTGGTCGATCTTGATTTCTTCACGCAGGAGCCTTTGCACCTTGTATCCCCCTCATTTTTAAGCGTTACCATAGACGCAAACCTTGCCACTGACCCGCGCTTCTTAATCTTGCTTGGTAGTCCAAAGTTAAGAACGCTGGCGCGGGGGCTTAGTCCAGCATATCTGCGCTTTGGCGGAACGAAAACAGACTTCCTGATTTTTGATCCTAAAAAGGAA
50 kDa subunit of wild type (gBlock)	AAAAAATTTAAGAACTCCACTTATAGCCGCAGTTCAGTCGACGTGCTGTACACCTTCGCGAACTGTTCGGGATTGGATTTGATATTTGGATTAAATGCATTGTTGCGCACGGCGGATCTGCAGTGGAACTCTAGTAACGCGCAATTATTGTTAGATTATTGTAGTTCGAAGGGCTACAATATATCGTGGGAATTGGGTAATGAGCCGAACAGCTTTTTGAAGAAAGCCGACATCTTTATTAATGGGTCTCAGCTGGGCGAAGATTTTATACAACTTCACAAGCTGTTACGCAAATCAACATTTAAGAACGCGAAGTTATATGGACCAGATGTTGGGCAGCCACGTAGAAAGACCGCCAAGATGCTGAAAAGCTTCCTTAAAGCAGGAGGTGAAGTGATTGACTCGGTGACCTGGCATCACTACTACTTAAACGGAAGAACAGCAACTCGTGAGGATTTCTTGAACCCGGATGTCCTTGATATATTTATTTCATCTGTACAAAAAGTCTTCCAAGTTGTAGAATCCACCAGACCTGGCAAAAAGTGTGGTTAGGAGAGACTTCAAGCGCTTACGGCGGTGGTGCACCTCTTTTGTCCGACACCTTCGCGGCAGGCTTCATGTGGCTGGACAAATTGGGCTTAAGCGCGCGTATGGGGATCGAAGTGGTGATGCGGCAGGTATTCTTTGGCGCCGGGAACTATCACCTGGTCGATGAAAATTTTGATCCTTTACCTGATTATTGGTTGTCATTACTGTTTAAAAAGTTGGTCGGGACAAAGGTCCTTATGGCTTCTGTCCAGGGGAGTAAAAGAAGAAAATTGAGAGTTTACTTGCATTGCACCAATACGGACAACCCGAGATATAAGGAAGGAGACTTGACCTTATACGCTATCAATTTGCACAATGTTACGAAATATTTGCGTTTACCTTACCCATTCTCCAACAAACAAGTTGACAAATACTTGCTGCGCCCTCTGGGTCCGCATGGCTTATTATCCAAATCGGTTCAGTTGAATGGCTTAACTCTGAAAATGGTAGATGATCAGACATTGCCACCATTGATGGAGAAA
50 kDa subunit of design (gBlock)	ccccatcttagtatattagttaagtataagaaggagatatacatATGAAAAAATTCAAAAACTCGACGTATAGCCGGTCTTCTGTGGATGTGCTCTATACTTTTGCGAAGTGTTCGGGCCTGGACTTAATCTTCGGCTTAAATGCACTGCTTCGGACTTCAGATGGGCAGTGGAATTCTAGCAATGCTCAGCTCCTGCTCGATTACTGTGCCTCTAAAGGGTATAACATCGACTGGGAGTTGGGCAACGAGCCAAATAGCTTCCGTAAAAAGGCTGGGATCTTCATCAACGGGTCGCAATTAGGCAAGGACTTCATTCACCTTCACAAACTGCTCCGGAAATCGACATTTAAGAATGCGAAACTGTATGGCCCTGATGTAGGTCAACCGCGCGGGAAAACGGCCAAAATGCTTAAATCGTTCCTGAAGGCGGGCGGCGAAGTCATTGATGCAGTAACATGGCACCATTACTATTTGAATGGTCGCACCGCCACCTTAGAAGATTTCCTGAATCCGGACGTATTGGACACGTTTATTTCTCAGGTTCAAAAGGTCTTGCAAGTTGTCGAATCG

ACCCGGCCTGGGAAGAAAGTTTGGCTCGGGGAGACAAGTTCCGCCTATGGCGGT
GGCGCTCCTGGCCTCTCAGATACCTTTGCTGCTGGTTTCATGTGGCTTGATAAACT
GGGCCTCTCCGCTCGCATGGGGATCGAAGTCGTGATGCGCCAAGTATTTTTTGGC
GCTGGCAACTACCACCTCGTCGACGAAAACTTCGATCCATTGCCTGACTACTGGCT
GAGCCTCCTTTTCAAAAAGTTAGTTGGTACAAAGGTGTTGATGGCAAGTGTTCAGG
GTCAGGATCGCCGCAAACTTCGCGTTTATCTCCATTGCACAAATACGGATAATCCT
CGCTACAAAGAAGGCGACCTGACGCTCTATGCTATCAACCTCCATAACGTCACCA

AGTATCTCCGCCTGCCATATCCTTTTAGTAATAAACAGGTGGATCAATATCTCTTGC
GCCCTCATGGCCCTGATGGTTTACTGTCCAAGAGCGTGCAGTTGAATGGCCAGAC
CCTCAAGATGGTTGATGACCAGACTTTGCCTCCTTTGAAGCCAAAACCACTGCGTC
CGGGGAGCAGTCTTGGCCTGCCTGCCTTCTCCTACGCATTTTTTGTAATTCGTAAC
GCAAAGGTCCCAGCCTGCATCTGATAACTCGAGTCTGGTAAAGAAACCG

Supplementary Table 2. Amino acid sequences of heparanase variants

8 kDa subunit	MGSSHHHHHHSQDPNSSSQDVVDLDFFTQEPLHLVSPSFLSVTIDANLATDPRFLIL LGSPKLRTLARGLSPAYLRFGGTKTDFLIFDPKKE*
50 kDa subunit of wild type	MKKFKNSTYSRSSVDVLYTFANCSGLDLIFGLNALLRTADLQWNSSNAQLLLDYCSS KGYNISWELGNEPNSFLKKADIFINGSQLGEDFIQLHKLLRKSTFKNAKLYGPDVGQP RRKTAKMLKSFLKAGGEVIDSVTWHHYYLNGRTATREDFLNPDVLDIFISSVQKVFQ VVESTRPGKKVWLGETSSAYGGGAPLLSDTFAAGFMWLDKLGLSARMGIEVVMRQ VFFGAGNYHLVDENFDPLPDYWLSLLFKKLVGTKVLMASVQGSKRRKLRVYLHCTN TDNPRYKEGDLTLYAINLHNVTKYLRLPYPFSNKQVDKYLLRPLGPHGLLSKSVQLN GLTLKMVDDQTLPPLMEKPLRPGSSLGLPAFSYSFFVIRNAKVAACI*
50 kDa subunit of design	MKKFKNSTYSRSSVDVLYTFAKCSGLDLIFGLNALLRTSDGQWNSSNAQLLLDYCA SKGYNIDWELGNEPNSFRKKAGIFINGSQLGKDFIHLHKLLRKSTFKNAKLYGPDVG QPRGKTAKMLKSFLKAGGEVIDAVTWHHYYLNGRTATLEDFLNPDVLDTFISQVQKV LQVVESTRPGKKVWLGETSSAYGGGAPGLSDTFAAGFMWLDKLGLSARMGIEVVM RQVFFGAGNYHLVDENFDPLPDYWLSLLFKKLVGTKVLMASVQGQDRRKLRVYLH CTNTDNPRYKEGDLTLYAINLHNVTKYLRLPYPFSNKQVDQYLLRPHGPDGLLSKSV QLNGQTLKMVDDQTLPPLKPKPLRPGSSLGLPAFSYAFFVIRNAKVPACI*

Supplementary Table 3. Data collection and refinement statistics

PDB ID	7RG8
Data collection	
Space group	P 2_1 2_1 2_1
Cell dimensions	
a, *b*, *c* (Å)	59.78 76.09 124.43
α, β, γ (°)	90.00 90.00 90.00
Resolution (Å)	47.01-1.30 (1.346-1.3)
R_{merge}	0.07 (1.11)
I / σ*I*	11.82 (0.33)
$CC_{1/2}$	0.998 (0.882)
Completeness (%)	99.78% (99.0%)
Redundancy	13.2 (11.3)
Refinement	
Resolution (Å)	47.01-1.30 (1.346-1.3)
No. reflections	139516 (13656)
R_{work} / R_{free}	0.142/ 0.164
No. atoms	
Protein	3660
Ligand/ion	11
Water	363
B-factors	
Protein	24.89
Ligand/ion	21.94
Water	41.23
R.m.s. deviations	
Bond lengths (Å)	0.007
Bond angles (°)	1.01
Ramachandran plot	
Preferred (%)	98.68
Allowed (%)	1.32
Outliers (%)	0.00

*X-ray data were collected from single crystals. *Values in parentheses are for highest-resolution shell.

Chapter 3

Understanding the binding mechanisms of glycomimetics with heparanase

3.1 Declaration

All experimental work except where otherwise stated, was carried out by the author under the supervision of Professor Colin Jackson The contribution of other authors are as follows: Dr. Yen Vo and Dr. Brett Schwartz developed methods and synthesised the sulfated xylooligosaccaride analogue, Xyl_4S for this paper, and the corresponding supplementary material. Hideki Onagi contributed to separation sciences required for this synthesis. Dr. Hafna Ahmed helped the author separate and assay PPS fractions. Caryn Hepburn collected and processed HDX-MS experiments. The author conceived the experiments with support from Professor Colin J. Jackson, Dr. Brett Schwartz, Professor Martin Banwell and Keats Nelms. Colin Jackson helped with experimental design and interpretation of the data. The author prepared the manuscript with input from other authors. The following paper has been published with ACS Biochemistry on June 27, 2023 and is available online

"Complex Inhibitory Mechanism of Glycomimetics with Heparanase"

DOI: 10.1021/acs.biochem.3c00038

Biochemistry

Article

Complex Inhibitory Mechanism of Glycomimetics with Heparanase

Cassidy Whitefield, Yen Vo, Brett D. Schwartz, Caryn Hepburn, F. Hafna Ahmed, Hideki Onagi,
Martin G. Banwell, Keats Nelms, Lara R. Malins, and Colin J. Jackson*

Cite This: https://doi.org/10.1021/acs.biochem.3c00038

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Heparanase (HPSE) is the only mammalian *endo-β*-
glucuronidase known to catalyze the degradation of heparan sulfate.
Dysfunction of HPSE activity has been linked to several disease
states, resulting in HPSE becoming the target of numerous
therapeutic programs, yet no drug has passed clinical trials to
date. Pentosan polysulfate sodium (PPS) is a heterogeneous, FDA-
approved drug for the treatment of interstitial cystitis and a known
HPSE inhibitor. However, due to its heterogeneity, characterization
of its mechanism of HPSE inhibition is challenging. Here, we show
that inhibition of HPSE by PPS is complex, involving multiple
overlapping binding events, each influenced by factors such as
oligosaccharide length and inhibitor-induced changes in the protein
secondary structure. The present work advances our molecular
understanding of the inhibition of HPSE and will aid in the development of therapeutics for the treatment of a broad range of
pathologies associated with enzyme dysfunction, including cancer, inflammatory disease, and viral infections.
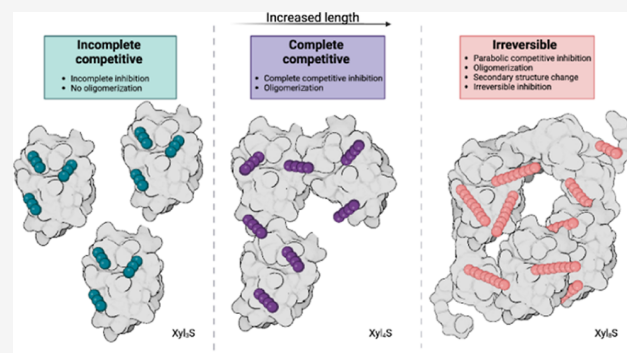
INTRODUCTION

Heparan sulfate proteoglycans (HSPGs) are ubiquitous
macromolecules associated with the cell surface and extrac-
ellular matrix (ECM) of animal tissues, where they mediate
critical interactions between cells and their environment.[1]
HSPGs consist of pericellular or extracellular core proteins to
which several heparan sulfate (HS) chains are covalently
bound via oxygen-based linkages.[2] The enormous structural
variation available within HS chains allows HSPGs to interact
with a wide range of proteins, thereby influencing many
biological processes. These processes include ECM homeo-
stasis, signaling, developmental patterning, cell adhesion,
barrier formation, and endocytosis.[3−8] HS also regulates the
activity of bioactive molecules such as growth factors,
cytokines, chemokines, and coagulation factors by providing
low-affinity storage within the ECM.[9−12] Cleavage of HS side
chains therefore not only alters the integrity of the ECM but
also leads to the release of such HS-bound bioactive molecules.

Human heparanase (HPSE) is the only known *endo-β*-
glucuronidase that cleaves HS side chains, producing shorter
oligosaccharides.[13−16] HPSE is a heterodimeric protein with an
8 kDa subunit spanning residues Gln36−Glu109 and a 50 kDa
subunit comprising Lys159-Ile543. HPSE consists of two
domains: the $(\beta/\alpha)_8$ domain, which contains the active site,
flanked by a smaller *β*-sandwich domain (Figure 1A). Baseline
HPSE activity is highly regulated and is only seen at low levels
in platelets, immune cells, and the placenta. Increased HPSE
expression is often observed in disease states, perhaps most
notably in cancer and viral infections.[17,18] The increased

activity can significantly alter cell motility by weakening the
structural HSPG networks within the ECM and basal
membranes,[19] facilitating angiogenesis,[20] inflammation,[21] and
invasion of the surrounding tissue.[22,23] In addition to
modifying the structure of the ECM, the breakdown of HS
chains by HPSE can release latent pools of growth factors.[24]
This can have a cumulative effect of altering the cell
proliferation, motility, and activation of important intracellular
signaling pathways, including those involving pro-inflammatory
cytokines.[25,26]

Due to the role HPSE plays in promoting tumor growth and
metastasis, it has been a target of many therapeutic
development programs. HPSE inhibitors in development
include HS-glycomimetics,[27−30] synthetically produced small-
molecule compounds,[31−33] nucleic acid-based inhibitors,[34]
covalent inhibitors,[35] monoclonal antibodies,[36] proteins,[37]
and natural products.[38] HS-glycomimetics and certain
derivatives are the only such compounds to advance to clinical
trials. These include muparfostat (PI88) (Figure 1B),[27]
pixatimod (PG545),[28] roneparstat (SST0001),[29] and necupar-
anib (M402),[30] which have been shown to exert anti-cancerous
and anti-metastatic effects in animal models and in early
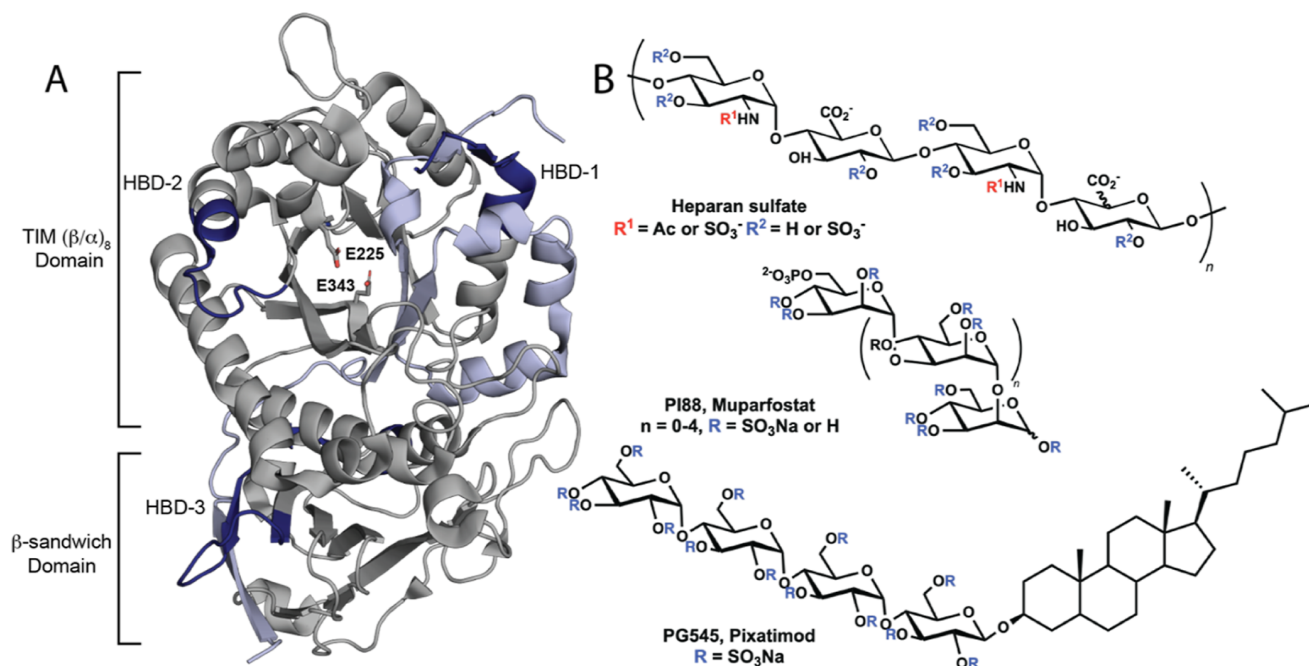
Figure 1. Crystal structure of HPSE and known inhibitors. (A) Crystal structure of HPSE (PDB ID: 7RG8). Chain-A is shown in gray and chain-B in light blue. Heparin binding domains (HBDs) 1−3 are depicted in dark blue, and the catalytic residues Glu225 (acid/base) and Glu343 (nucleophile) in the enzyme-binding cleft are shown in stick representation in gray. (B) Structure of heparan sulfate (top) compared to the studied HS mimetic inhibitors PI88 (middle) and PG545 (bottom).

clinical trials.[39,40] Although initially promising, clinical trials of all but PG545, which is currently in phase II trials,[41] have either been paused or terminated due to poor efficacy (SST0001 and M402) and adverse side effects (PI88).[42−44]

Pentosan polysulfate (PPS; Elmiron) is an FDA-approved, semi-synthetic mixture of polysulfated xylans now deployed as an orally administered treatment for bladder pain or discomfort associated with interstitial cystitis.[45,46] Currently, PPS is undergoing clinical trials for the treatment of non-infectious arthritis,[47] while recent research has shown that it might also be effective against alphavirus-induced arthritis. PPS is a known cytokine-binding molecule that possesses biological neutralization capacities and exerts broad anti-inflammatory effects.[48] However, despite the interest in PPS and related sulfated oligosaccharides as HPSE inhibitors,[49] the molecular basis for these inhibitory effects is not comprehensively understood.

In this work, we describe the mechanisms by which PPS analogues of different lengths inhibit HPSE. HPSE inhibition was examined in combination with X-ray protein crystallography and hydrogen/deuterium exchange mass spectrometry (HDX-MS), leading to the identification of three oligosaccharide binding sites, including a less-studied remote site. Enzyme kinetic and inhibition studies, alongside circular dichroism (CD), were used to reveal the complex nature of HS-mimetic-induced inhibition of HPSE. This revealed that larger oligosaccharide molecules, including PPS, caused aggregation and loss of the HPSE secondary structure, a phenomenon that emphasizes the challenges confronting the clinical deployment of such compounds. These data advance our understanding of HS-glycomimetic inhibitors and will aid future drug development directed toward inhibition of the enzyme.

■ METHODS

HPSE P6 was obtained following the methods of Whitefield et al.,[50] while the chemical synthesis of Xyl$_4$S is reported in detail in the Supporting Information. The preparation of Xyl$_3$S and Xyl$_8$S has been described previously.[51] Samples of PI-88 were donated by Professor Chris Parish of the John Curtin School of Medical Research at the Australian National University.

Pentosan Separation. Size-exclusion chromatography (Superdex S30 pg GE Healthcare) was used to separate PPS into its individual components. Thus, PPS (500 μL of 50 mg/mL) was injected onto a column equilibrated with 300 mM aqueous ammonium bicarbonate and the eluting compounds were detected through their absorbance at 254 nm. Molecular weights were determined by plotting retention volumes against a standard curve for sulfated sugars of defined composition, the line of best fit being defined by

$$y = -0.021x + 4.7$$

where $y = \log(MW)$ and x = retention volume.

Samples were collected and freeze-dried before being reconstituted in MQ water at defined concentrations.

HPSE P6 Mutagenesis. The cloning of HPSE P6 is shown in Whitefield et al.[50] Genes for most HPSE P6 mutants (F159A+K160A, K161A, R272A, K274A, K411Q, and G415E) were synthesized by Twist Bioscience. The original 50 kDa subunit was removed and linearized using NdeI and XhoI restriction enzymes (Fast Digest, Thermo) and mutant genes were inserted into the multiple cloning site 2 by Gibson assembly.[52] K417A and R428A mutations were introduced through PCR, amplified with mid_duet and mutagenesis primers (Table S1), and ligated with Gibson assembly. All ligated DNA was transformed in *Escherichia coli* TOP10 cells, and the plasmid DNA was extracted and sent to the Garvan Institute (Sydney) for Sanger sequencing conformation.

Protein Expression and Purification. HPSE P6 expression was conducted following previously described methods.[50] HPSE P6 was transformed into *E. coli* SHuffle T7 Express cells (NEB), together with GroEL/ES + Trigger factor chaperones, in a pACYC vector and spread on an agar plate with ampicillin and chloramphenicol. 1% overnight seed culture from a single colony was inoculated into 1 L of LB medium supplemented with ampicillin (100 mg L^{-1}) and chloramphenicol (34 mg L^{-1}) and then incubated at 37 °C for 5 h. Overexpression was induced by adding IPTG to a final concentration of 0.05 mM, and the culture was further incubated for 3 h at 37 °C. The derived cell pellet was resuspended in buffer A (20 mM HEPES pH 8, 300 mM NaCl, 5 mM β-mercaptoethanol, 10% (v/v) glycerol, and 20 mM imidazole) with Turbonuclease (Sigma) then lysed by sonication (Omni Sonic Ruptor 400 Ultrasonic homogenizer). The lysate was filtered (0.45 μm) and loaded onto a Ni-NTA column (GE healthcare) and eluted with 100% buffer B (buffer A + 500 mM imidazole). The peak eluent was diluted 5 times with buffer C (20 mM HEPES pH 7.4, 200 mM NaCl, 5 mM β-mercaptoethanol, and 10% (v/v) glycerol) and loaded onto a heparin affinity column (GE Healthcare) and eluted with 100% buffer D (buffer C + 1.5 M NaCl). The peak eluent was loaded onto a size-exclusion column (HiLoad 26/600 Superdex 200 pg, GE Healthcare) and eluted into buffer E (20 mM sodium acetate pH 5, 200 mM NaCl, 10% (v/v) glycerol, and 1 mM tris(2-carboxyethyl)phosphine (TCEP)). The final concentration of the monomeric HPSE from the gel filtration was estimated by absorbance at 280 nm using NanoDrop One (Thermo).

HPSE Colorimetric Assays. Assays were conducted using the colorimetric assay designed by Hammond et al.[53] Bovine serum albumin-coated 96 well microplates were used for all assays and were prepared by incubation of the plates with 1% BSA dissolved in phosphate-buffered saline (PBS) with 0.05% Tween-20 (PBST) at 37 °C for 75 min. The plates were then washed three times with PBST, dried, and stored at 4 °C. Assay mixtures contained 40 mM sodium acetate buffer (pH 5.0), 0.8 nM HPSE in 0.01% Tween 20 sodium acetate buffer, and 100 μM fondaparinux (GlaxoSmithKline) with or without increasing concentrations of inhibitor. Plates were incubated at 37 °C for 2−20 h before the reaction was terminated using 100 μL of 1.69 mM 4-[3-(4-iodophenyl)-2-(4-nitrophenyl)-2*H*-5-tetrazolio]-1,3-benzene disulfonate (WST-1) in 0.1 M aqueous NaOH. The plates were resealed and developed at 60 °C for 1 h, and the absorbance was measured at 584 nm (Synergy 2). Kinetics were carried out with a standard curve constructed with ᴅ-galactose as the reducing sugar standard, prepared in the same buffer and volume over the range of 0−2 μM.

Data Analysis. The Hill-type model reported by Cao et al.[54] was used to plot parabolic inhibition as it proved a better fit compared to the standard parabolic competitive equation[55] as has been demonstrated for other HPSE inhibitors.

$$v = \frac{V_{max} \cdot S}{K_m \cdot \left(1 + \frac{I^n}{K_{IC}^n}\right) + S}$$

(1)

This equation can be rearranged for analysis of the slope data (Figure 3F)

$$slope = \frac{K_m \cdot \left(1 + \frac{I^n}{K_{IC}^n}\right)}{V_{max}}$$

(2)

The competitive inhibition equation (eq 3) was used for Xyl$_4$S and Xyl$_8$S (as they exhibited the same mode of binding) and fitted to the velocity data using global nonlinear regression.

$$v = \frac{V_{max} \cdot S}{K_m \cdot \left(1 + \frac{I}{K_{IC}}\right) + S}$$

(3)

All curve fittings to calculate IC$_{50}$ values and Michaelis−Menten constants were carried out using GraphPad Prism software (v 9.4).

Aggregation Assays. Assay mixtures contained 10 μM HPSE in buffer E and inhibitor were added to a concentration of between 1.25 and 100 μM in a volume of 200 μL. Samples were agitated every 5 s between reads, for 60 min, and absorbance was measured at 300 nM. Negative controls were measured by testing HPSE with no inhibitor present. Curves were fitted using GraphPad Prism software (v 9.4).

Isothermal Titration Calorimetry. Isothermal titration calorimetry (ITC) experiments were performed using a Nano ITC low-volume calorimeter (TA Instruments) and carried out at 25 °C with stirring at 350 rpm. Samples were prepared in SEC buffer (20 mM sodium acetate pH 5.0, 200 mM NaCl, 10% (v/v) glycerol, and 1 mM TCEP). Inhibitors dissolved in MQ water to 2.5 mM were diluted into SEC buffer with the same MQ water dilution matched precisely to the protein buffer. Both forward and reverse titrations were undertaken, with forward titrations involving 50 μL of various inhibitor concentrations injected continuously into 500 μL of 20−25 μM protein and reverse titrations involving 50 μL of various protein concentrations injected continuously into 500 μL of 30−80 μM inhibitor over 150 second injection intervals. NITPIC (version 1.2) was used to integrate the thermograms. The data were serially integrated and placed into a single SEDPHAT configuration file for global analysis.

The integrated ITC data were analyzed using SEDPHAT (version 12.1b). The "A + B + B ↔ B + AB ↔ BA + B ↔ ABB with two non-symmetric sites, microscope K" model was used (A defined as HPSE and B as one of synthetic variants) for the forward reactions, and reverse reactions used the A + B ↔ AB model.

Since NITPIC provides error estimates for all integrated data points, the SEDPHAT option to use these as weights in the fitting sessions was activated. Switching between the Simplex and Marquardt−Levenberg optimization routines was necessary to achieve convergence of the parameter set. Standard deviations were calculated for K$_D$ for both the forward and reverse reactions using GraphPad Prism software (version 9.4).

Circular Dichroism. CD spectra for HPSE were obtained using a 1 mm quartz cuvette and an Applied Photophysics Chirascan spectrometer. Purified enzymes in SEC purification buffer (20 mM sodium acetate pH 5, 200 mM NaCl, 10% (v/v) glycerol, and 1 mM TCEP) were measured at a protein concentration of 0.2 mg/mL along with the addition of inhibitors at a 1:10 molar ratio. Samples were scanned at 20 °C between 200 and 260 nm using a band width of 1 nm, and a scan rate of 1 s (with adaptive sampling enabled). Spectra were recorded in triplicate, and a buffer blank was subtracted from the results. Molar ellipticity (deg·cm^2/dmol) was used for all subsequent analyses.

Protein Crystallography. Well-diffracting single crystals were obtained by the hanging-drop vapor-diffusion method at 18 °C by combining the protein in 20 mM sodium acetate pH
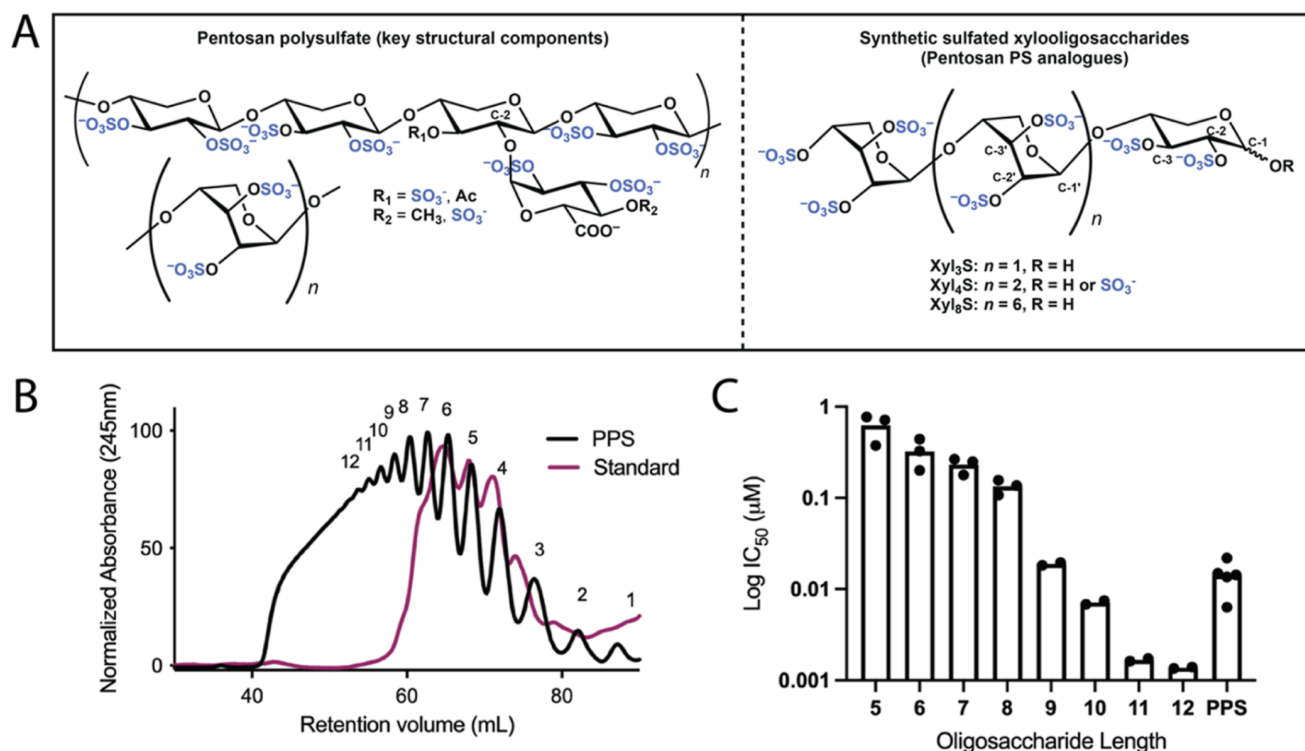
Figure 2. Structure, separation, and activity of the individual components of PPS. (A) (upper left-hand panel) representative repeating unit of PPS, the likely favored conformation (lower left panel) of the xylose disulfate residues of the PPS backbone lacking a glucuronic acid branch, and the established structures (right-hand panel) of the synthetically derived PPS analogues Xyl_3S, Xyl_4S, and Xyl_8S; (B) size-exclusion chromatogram of PPS (in black) overlaid with that of a synthetically derived[51] reference mixture of straight-chain sulfated xylooligosaccharides (in purple)[51] (the numerals represent the average number of monosaccharide residues associated with each component); (C) dose−response profile (established using a fondaparinux-based assay) for each component of the PPS mixture vs that of PPS itself (all replicates are shown, and the data are presented on a log_{10} scale).

5, 200 mM NaCl, 10% (v/v) glycerol, 1 mM TCEP (6−8 mg mL^{-1}), and the well solution [1.9 M $(NH_4)_2SO_4$] in a ratio of 1.5:1.5 μL. Crystals appeared within a week and continued to grow for 1−2 months. Small amounts of ligand compound were added to the crystal drop and left to soak for 4 h to 2 days. Crystals were then frozen directly in liquid nitrogen without a cryoprotectant. Crystallographic data were collected at 100 K at the Australian Synchrotron (MX2,[56] 0.9537 Å). The derived diffraction data were indexed and integrated with DIALS.[57] Resolution estimation and data truncation were performed using AIMLESS as implemented in CCP4.[58,59] All structures were solved by molecular replacement using the MOLREP program in CCP4[58] and using the structure deposited under the PDB accession code 7RG8 as a starting model. The models were refined using phenix.refine,[60] and the model was subsequently optimized by iterative model building with the program COOT v0.9.[61] Alternative conformations were modeled based on mFo−DFc density, and the occupancies and B-factors were determined using phenix.refine.[60] Ligands were optimized via elBOW and fitted into the structure via LigandFit. The structures were then evaluated using MolProbity[62] in Phenix. Details of the refinement statistics were produced by Phenix (version 1.19)[63] and are summarized in Table S2. The structures were visualized and analyzed using PyMol (version 2.5).[64] Electrostatic potential maps were generated using eF-surf.[65]

HDX-MS Deuterium Labeling and Quenching Conditions. A Trajan LEAP HDX Automation manager was used to automate labeling, quenching, and injection of samples.

Inhibitor-bound proteins were prepared at a concentration of 12 μM with a 10× molar ratio of inhibitor. A 3 μL protein sample was incubated in 57 μL of sodium acetate buffer pH 5.0 reconstituted in D_2O (99.90%, Sigma). Deuterium labeling was performed for 0.5, 1, 10, and 100 min, followed by quenching 50 μL of the deuterium exchange reaction mixture in 50 μL of a pre-chilled 50 mM sodium acetate quenching solution containing 2 M guanidine hydrochloride (Sigma) and 200 mM TCEP to lower the pH to 2.5 and lower the temperature to 0.1 °C. A post-quench reaction time of 30 s was used.

MS and Peptide Identification. Quenched samples (80 μL) were injected onto a chilled Trajan HDX Manager. Samples were subjected to online digestion using an immobilized Waters Enzymate BEH pepsin column (2.1 × 30 mm) in 0.1% formic acid in water at 100 μL/min. The proteolyzed peptides were trapped in a 2.1 × 5 mm C18 trap (ACQUITY BEH C18 VanGuard pre-column, 1.7 μm, Waters, Milford, MA). The proteolyzed peptides were eluted using acetonitrile and 0.1% formic acid gradient (5 to 35% 6 min, 35 to 40% 1 min, 40 to 95% 1 min, and 95% 2 min) at a flow rate of 40 μL/min using an ACQUITY UPLC BEH C18 Column (1.0 × 100 mm, 1.7 μm, Waters, Milford, MA) pumped by a UPLC I-Class Binary Solvent Manager (Waters, Milford, MA). A positive electrospray ionization source fitted with a low flow probe was used to ionize peptides sprayed into a SYNAPT G2-Si mass spectrometer (Waters, Milford, MA). Data were acquired in MassLynx version 4.2 in the MS^E acquisition mode using 200 pg/μL leucine enkephalin and 100 fmol/μL [Glu1]-fibrinopeptideB ([Glu1]-Fib). Lockspray was introduced by

Table 1. HPSE Binding Data for the Three Synthetic PPS Analogues, Including Maximum Inhibition and IC$_{50}$ Values (Derived from Inhibition Curves)[a]

	Xyl$_3$S	Xyl$_4$S	Xyl$_8$S	PPS
max inhibition (%)	66.6	95.7	97.0	100
IC$_{50}$ (μM)	34.6 ± 4.2	4.01 ± 1.0	0.128 ± 0.006	0.01 ± 0.002
K_i (μM)	ND	1.26	0.021	0.0022
K_D combined (μM)	ND	5.44 ± 1.97	2.17 ± 0.552	ND
K_D Site 1 (μM)	ND	0.430 ± 0.486	1.58 ± 0.21	ND
K_D Site 2 (μM)	ND	5.87 ± 0.857	10.3 ± 1.21	ND
unit of binding (N)	ND	2.7	1.8	ND

[a]K_i values obtained from Lineweaver−Burk enzyme kinetics experiments. K_D values and stoichiometry of binding were obtained from ITC analysis. All experiments were conducted under the same buffer conditions. Errors are one standard deviation from the mean.
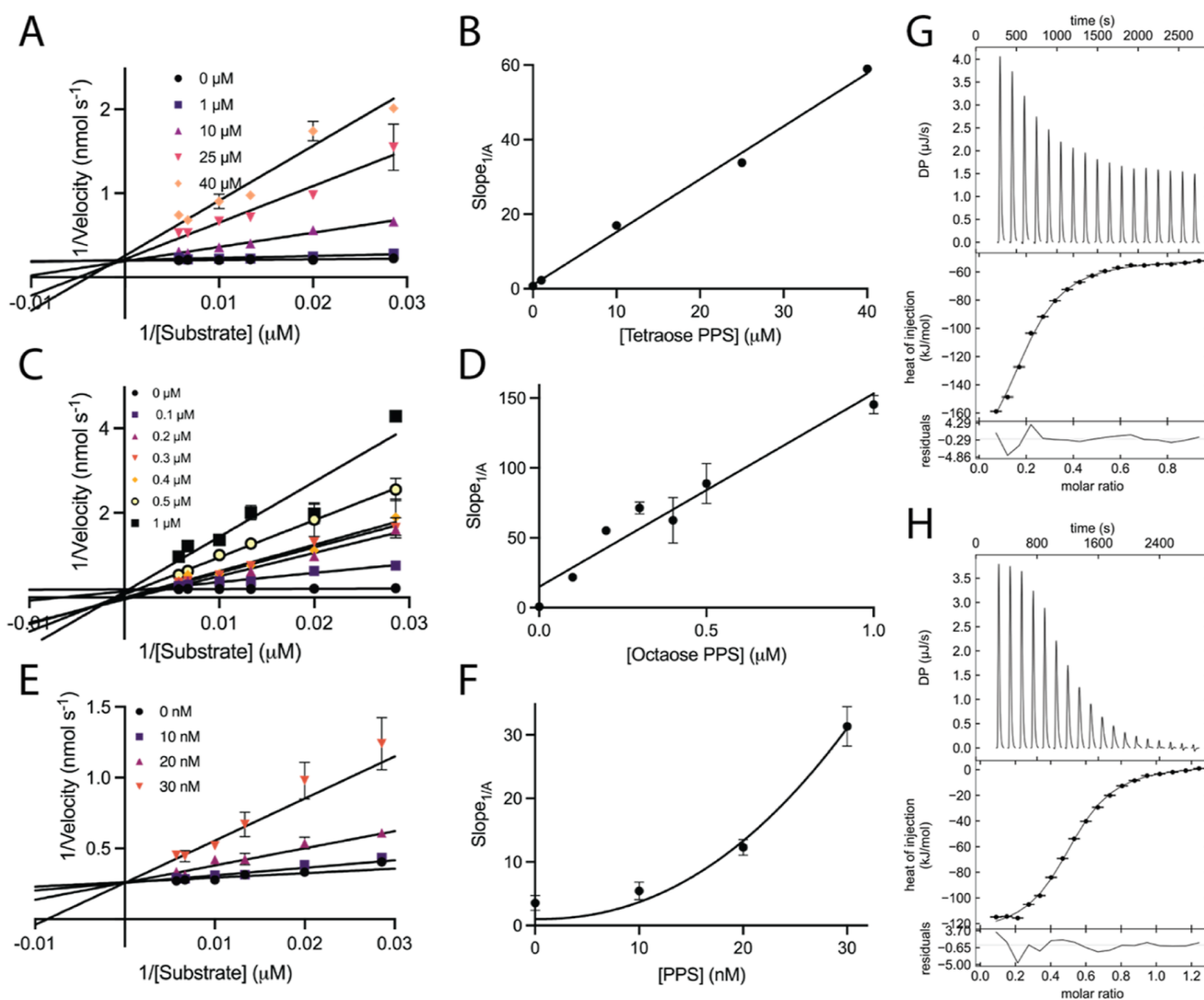
Figure 3. Binding analysis of pentosan analogues. (A,C,E) Double-reciprocal analysis of HPSE inhibition by Xyl$_4$S, Xyl$_8$S, and PPS. Data are means of two measurements, while errors are shown as standard deviations (SD). The slopes from panels (A,C,E) were replotted as a function of inhibitor concentration in panels (B,D,F), respectively. The plots in panels (B,D) represent the global fit of the competitive inhibitor equation (eq 3). The plot in panel F represents the fit to the Hill-type slope (eq 2) ($R^2 = 0.959$). (G,H) ITC reverse titration data for Xyl$_4$S and Xyl$_8$S, respectively, using the A + B ↔ AB binding model. Stoichiometry is shown as the molar ratio of proteins binding to each ligand, reversed to provide ligand/protein stoichiometry.

infusion at a flow rate of 5 μL/min into the mass spectrometer. Protein Lynx Global Server (PLGS) (version 3.0) was used to identify peptides in non-deuterated protein samples. The identified peptides were further filtered in DynamX (version 3.0) using a minimum intensity cut-off of 10,000 for product

and precursor ions, a minimum products per amino acid of 0.3, and a precursor ion mass tolerance of 5 ppm using DynamX (version 3.0) (Waters, Milford, MA). Deuterium exchange, relative deuterium exchange, and difference plots were generated. All deuterium exchange experiments were per-

formed in triplicate, and the reported values are not corrected for deuterium back exchange.

■ RESULTS

Pentosan length affects HPSE inhibition. PPS is composed of a heterogeneous mixture of β-1-4 linked and sulfated xylooligo-saccharides that include occasional sulfated 4-O-methyl-α-D-glucuronic acid residues attached to the C-2 position of the xylose backbone (Figure 2A). Deconvoluting the heterogeneity of PPS is therefore an essential step toward establishing which features of its constituents are the most important for activity. To investigate the molecular composition of PPS, size-exclusion chromatography was conducted, revealing twelve discrete peaks, with each successive peak representing a homologue incorporating an additional sulfated xylan unit (Figure 2B). Each such average component was subjected to an inhibition assay using mammalian HPSE, which revealed that the larger species were the more potent (Figure 2C). Oligosaccharides with nine or more sugar units proved to be better inhibitors than bulk PPS, while those shorter than this displayed weaker activity. This suggests that an increased oligosaccharide length is correlated with greater HPSE inhibition, even when surpassing the discrete active site binding capacity of HPSE, which is limited to four oligosaccharide residues.[66] These results also indicate that the majority of inhibition resulting from heterogeneous (bulk) PPS derives from the very large oligosaccharide species (9+ oligosaccharide units) present in PPS.

Due to the importance of PPS length in HPSE inhibition, three PPS analogues of defined structure and lacking glucuronic acid branches (side-chains) were synthesized to better understand the binding interactions.[51] The compounds prepared for this purpose were Xyl_3S, Xyl_4S, and Xyl_8S, comprising three, four, and eight xylose-backbone residues, respectively. As WT HPSE cannot express at the yield required for these experiments, HPSE-P6 was used for assays and subsequent studies as it has been shown to have essentially identical overall structure, dynamics, and activity as mammalian WT HPSE while yielding high levels of soluble expression in *E. coli*, being cost-effective, and producing protein crystals that diffract to high resolution (<2 Å) within ~24 h.[50] Importantly, HPSE-P6 shows an identical inhibitory response with PPS, i.e., the conservative mutations (remote from any binding site) in HPSE-P6 do not appear to affect its function or interaction with substrates or inhibitors.[50] The results of inhibition measurements indicate that PPS and the three short-chain analogues all inhibit HPSE activity in a concentration-dependent manner (Table 1; Figure 1). PPS was found to be the most effective, followed by Xyl_8S, Xyl_4S, and Xyl_3S (Table 1). A correlation between increased oligosaccharide length and increased binding interactions and inhibition was observed. PPS, Xyl_4S, and Xyl_8S were all able to completely inhibit HPSE at saturating concentrations, while Xyl_3S only displayed 66.6% of enzyme activity at the highest concentration tested (125 μM; Table 1). While competitive inhibition of HPSE by HS-mimetics is typically assumed, the incomplete inhibition by Xyl_3S is consistent with observations made on other triose-based oligosaccharide inhibitors of HPSE.[67] We hypothesize that because Xyl_3S does not occupy the full length of the binding cleft, there is only partial rather than complete blockage of substrate binding (notably, the substrate analogue, fondaparinux, consists of five monosaccharide units).

Full-Length PPS and Shorter Oligosaccharides Inhibit HPSE via Different Mechanisms. Given the significant differences between the inhibition of HPSE by bulk PPS and the short-chain oligosaccharides, the mechanistic basis for HPSE inhibition by the synthetically derived compounds was investigated. Since Xyl_3S did not result in complete inhibition and given its close similarity to Xyl_4S, it was omitted from these experiments. Double reciprocal (Lineweaver−Burk) enzyme kinetic analyses conducted at varying substrate and inhibitor concentrations indicated that Xyl_4S, Xyl_8S, and PPS are all competitive inhibitors of HPSE (Figure 3A,C,E). However, when the gradients of the double−reciprocal graphs were plotted against inhibitor concentration, it became clear that while the responses of Xyl_4S and Xyl_8S were linear, long-chain PPS fits to a Hill-like model (Figure 3F).[54,68] These results suggest that, unlike the linear competitive inhibition seen with Xyl_4S and Xyl_8S, PPS is a parabolic competitive inhibitor, as also observed for the HPSE inhibitors PG545[67] and SST0001.[68] The Hill-type kinetic behavior (eq 2, Methods) displayed in the PPS inhibition of HPSE revealed a K_i of 2.2 nM and suggests that more than one inhibitor molecule is binding to the enzyme or that PPS causes aggregation of HPSE (or both).[69,70] The Hill coefficient reflects the number of inhibitor binding sites on the enzyme, and a value of n greater than 1 suggests that multiple binding sites are involved in the interaction between the inhibitor and enzyme. Aggregation of the enzyme also leads to Hill-type kinetics as the sequestered enzyme can no longer interact with substrate. In contrast, a linear competitive inhibition model (eq 3, Methods) was used for Xyl_4S and Xyl_8S. From this analysis, K_i values for Xyl_4S and Xyl_8S were determined to be 1.26 μM and 21 nM, respectively. The large difference in K_i values between Xyl_4S and Xyl_8S relative to the number of monosaccharide residues in the two inhibitors, combined with the parabolic binding profile of PPS, suggests that the binding of these inhibitors to HPSE is a complex process.

Thermodynamics of Inhibitor Binding. The binding interactions between HPSE and Xyl_4S and Xyl_8S were probed further using isothermal titration calorimetry (ITC). ITC allows for greater understanding of the thermodynamics of ligand binding, as well as the stoichiometries involved. Both forward and reverse titrations (ligand into protein and protein into ligand) were used to gain insights into the binding stoichiometry. Forward titrations, wherein the ligand is titrated into protein, should reveal the affinity of a ligand for an individual binding site (Figure S2) while reverse titrations, wherein protein is titrated into the ligand solution, should enable a global analysis of the affinity of the ligands for HPSE and the stoichiometry involved (Figure 3G,H). Combined affinity calculations derived from the reverse ITC titrations showed that K_D decreased (affinity increased) with increasing oligosaccharide length (Table 1) with Xyl_4S and Xyl_8S showing combined K_D values of 5.44 ± 1.97 and 2.17 ± 0.55 μM, respectively. These same titrations also established that Xyl_4S and Xyl_8S bind to HPSE with ~3:1 and ~2:1 stoichiometries, respectively (Figure 3G,H). Forward titration measurements with both compounds reveal that there are likely multiple binding events/sites with different affinities (Figure S2). The correlation between length and binding affinity and the magnitude of change of the K_D values for the individual sites are more consistent with the length of the oligosaccharides than the enzyme kinetic inhibition results (Table 1). The comparison between the K_i and K_D values for Xyl_4S support
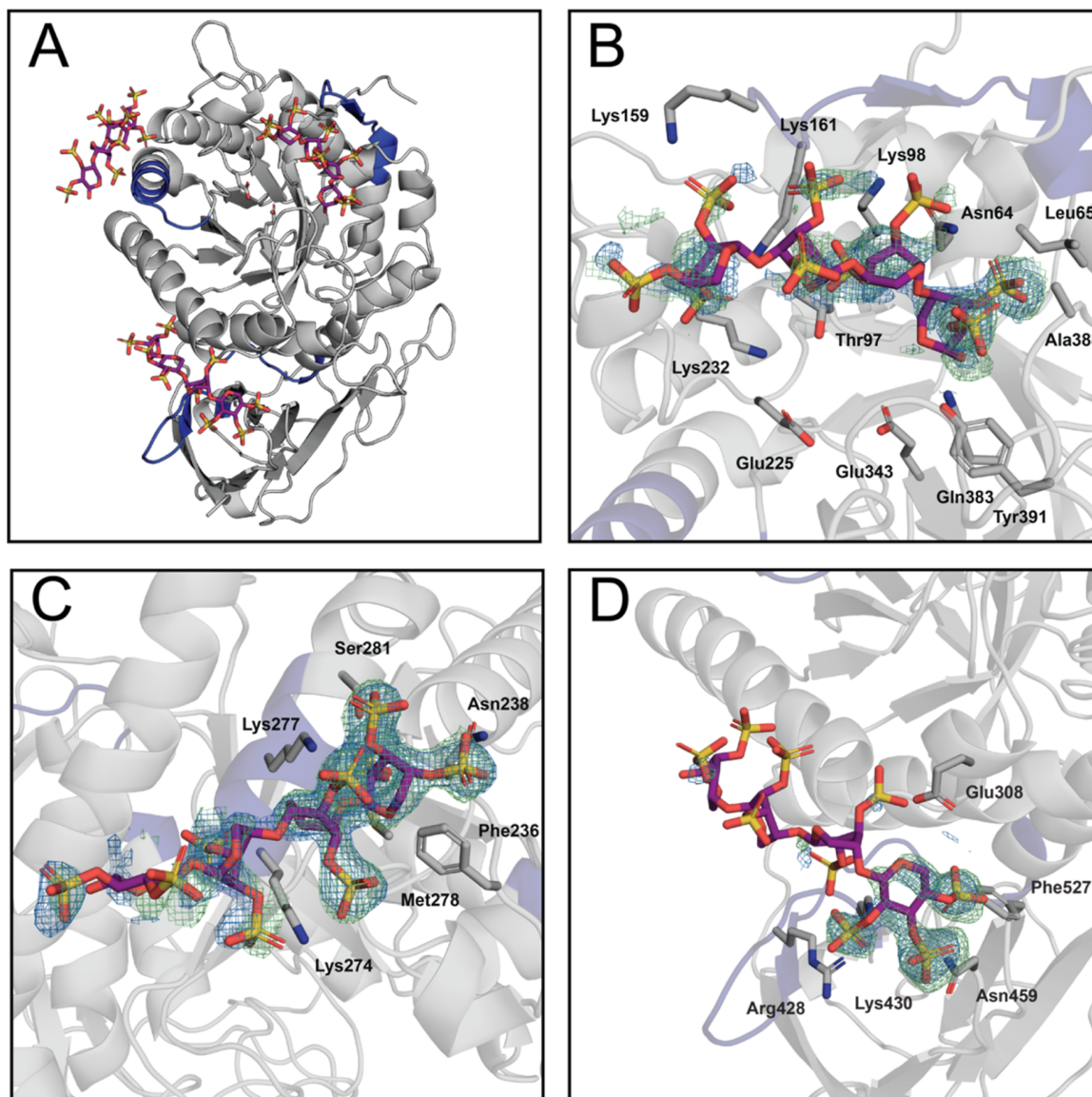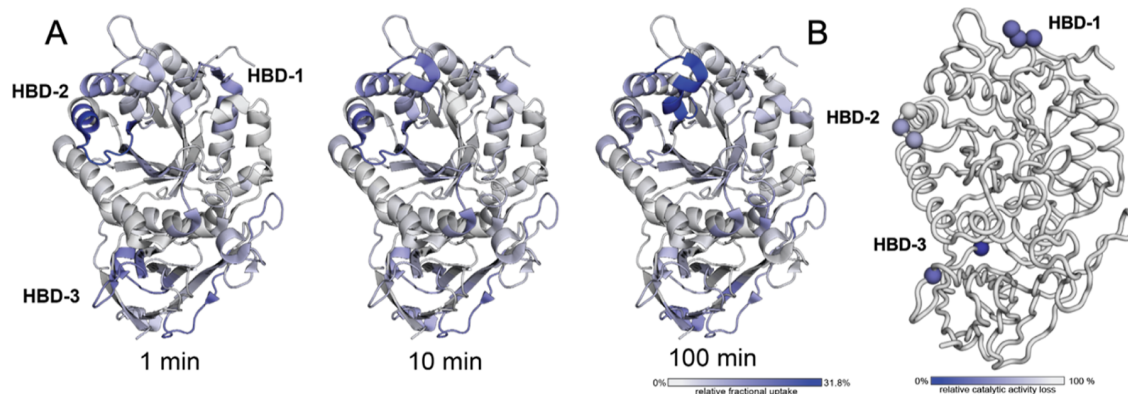
Figure 4. Crystal structure of the Xyl_4S-bound HPSE (8E08). (A) Crystal structure of Xyl_4S bound to HPSE. HBDs shown in blue and catalytic residues Glu225 and Glu343 shown as sticks and so highlighting the active site. (B) Zoomed-in view of the Xyl_4S ligand with density at HBD-1 with nearby catalytic residues shown, (C) HBD-2, (D) and HBD-3. mFo-DFc maps are represented in green (2.5σ), and 2mFo-DFc maps are in blue (1σ).

competitive binding (K_i is of the same magnitude as K_D). However, this is not the case for Xyl_8S, where the K_i is approximately 100-fold lower than the K_D ($K_i \ll K_D$). These results suggest that the inhibition of HPSE by Xyl_8S (and longer-chain components in PPS) is a more complex process than one involving simple competitive mechanisms. Given that ITC measurements are recorded instantaneously, while the kinetic assays take place over a longer timescale, the difference in K_i and K_D values could be due to the longer PPS species affecting the HPSE structure.

Structural Interactions of PPS Showing Multiple Binding Sites to HPSE. To further probe the interactions of PPS oligosaccharides with HPSE at the molecular level, we solved crystal structures of HPSE-P6 in complex with Xyl_3S and Xyl_4S (Table S3). Crystals of HPSE were soaked with ligand before flash-cooling and data collection. Electron density corresponding to Xyl_3S and Xyl_4S ligands were

identified in the crystal structures. In contrast, it was not possible to obtain a complex of HPSE with Xyl_8S, due to structural disruption of the protein (crystal deterioration and loss of diffraction). Two molecules of Xyl_3S and Xyl_4S were seen around the active site, namely at the previously described heparan binding domains (HBDs) 1 (Lys158-Lys162) and 2 (Pro271-Met278) (Figures 4 and S3). The electron density for some regions of the oligosaccharides were weak, indicating that either the binding is relatively dynamic or the ligands were bound at less than full occupancy (Figures 4 and S3). However, the electron density was unambiguous in certain regions, confirming that the analogues were bound at these sites. At HBD-1, there are interactions between the Xyl_3S sulfate moieties and Asn64, Lys98, Lys159, Lys161, and Tyr391. For Xyl_4S, further interactions are also seen with Tyr97, Lys232, and Gln383. Although the two PPS analogues bind at similar locations, Xyl_3S faces out of the binding site

Figure 5. HDX-MS difference data and effect of HBD mutations. (A) Representation of the HDX-MS difference data for apo- and Xyl$_3$S-bound HPSE at 5 °C at the 1, 10, and 100 min timepoints overlaid on the HPSE P6 crystal structure (7RG8), highlighting relative fractional uptake. (B) Relative catalytic activity (%) of point mutations made on HBD1-3 compared to HPSE P6 activity.

toward the solvent, possibly still allowing for the substrate to continue to interact with HPSE and resulting in the partial inhibition observed in the assays (Table 1). In contrast, Xyl$_4$S has more extensive interactions with HPSE and so allowing it to sterically block more of the active site and so providing a structural explanation for the partial inhibition observed for Xyl$_3$S (Table 1). In both Xyl$_x$S/HPSE complexes, the HBD-1 residues of Lys159, Phe160, and Lys161 undergo a large conformational change compared to the apo and ligand bound forms of HPSE that have been observed in other crystallographic studies (PDB: 7RG8 and 5E9C). These residues rotate out toward the solvent, allowing for Xyl$_x$S to bind, whereas they usually face toward the binding site over the top of the substrate (Figure S4). These residues have previously been shown to play an important role in HPSE inhibition.[71] At HBD-2, the binding interactions of Xyl$_3$S and Xyl$_4$S are very similar (Figures 4C and S3C) with HBD-2 residues Asn238, Ser240, Lys274, Lys277, Met278, and Ser281 involved in binding.

Consistent with the predicted stoichiometry from the ITC measurements (Figure 3G,H), a third binding site can also be discerned in the crystal structure involving Xyl$_4$S (Figure 4D). HBD-3 was first proposed by Levy-Adam et al., but because of an inability to purify the properly folded protein with this domain deleted, it was suggested that this region was important for heterodimer formation and most likely did not bind to HS.[71] Here, we observe electron density for the non-reducing end of Xyl$_4$S, indicating that it is bound at the HBD-3 site (involving residues Lys411-Lys417, Lys427-Arg432),[71] while Asn459, Phe527, and Glu308 are also shown to interact. The remaining units of this oligosaccharide are not well resolved in the crystal structure, most likely due to the high conformational flexibility of the oligosaccharide in this region. Nevertheless, it is evident that a third oligosaccharide binding domain likely exists in HPSE, clarifying ambiguities in the literature surrounding this site.

The observed conformations of the sulfated xylooligosaccharides in the crystal structures are consistent with those of the isolated Xyl$_{2-8}$S systems recently reported by Vo et al.[51] Specifically, the sulfated, non-reducing xylose residues preferentially adopt a $^1C_4$ conformation. However, when bound to HSPE, the internal sulfated xylose residue of Xyl$_3$S adopts a $^4C_1$ conformation, suggesting that hydrogen bonding to Lys98 and Lys161 stabilizes this over the $^1C_4$ conformation observed in solution.[51] Likewise, the reducing end of Xyl$_4$S

adopts a $^1C_4$ conformation, and, as a result, hydrogen bonds between Tyr391 and the β-C-1 hydroxyl are observed while Asn64 and Lys98 interact with the sulfate residues at C-2 and C-3, respectively.

HDX-MS Measurements and Mutagenesis Support Multiple Binding Sites. HDX-MS was employed to confirm that the in-solution binding of Xyl$_3$S occurs in the regions suggested by X-ray protein crystallography. Xyl$_3$S was the only analogue studied using this method due to the high level of aggregation seen with the other analogues. In-line pepsin digestion and filtering yielded 206 peptides with 98.7% coverage of the primary sequence. In the presence of Xyl$_3$S, several regions of HPSE were shielded from solvent exchange, with the most noticeable regions spanning residues Lys159−Val170 (including HBD-1), Gly265−Leu283 (including HBD-2), and, to a lesser extent, Ser422−Leu435 (including HBD-3) (Figures 5A and S7). These regions of reduced deuterium incorporation are consistent with the HPSE-Xyl$_4$S co-crystal structure showing binding at HBD-1, HBD-2, and HBD-3. It is likely that Xyl$_3$S was not observed at HBD-3 in the crystal structure, owing to a lower affinity at this site compared with Xyl$_4$S. Other regions of HPSE also exhibited minor shielding from deuterium exchange, such as the region spanning residues Asn224−Ile237 at the active site of HPSE, encapsulating one of the catalytic residues Glu225 as well as the neighboring HBD-2 loop. Surface-exposed random coils of the β-sandwich domain, encompassing residues Tyr468−Gln477 and Val503−Ser521, also experience variations in deuterium uptake. These variations may be explained by the considerable structural dynamics of HPSE, suggesting that ligand binding could affect conformational sampling.

Site-directed mutagenesis of residues at the three heparin binding domains led to a reduction in activity and so confirming the importance of these regions (Figure 5, Table S2). All mutations localized at HBD-1 (Lys159Ala, Phe160Ala, and Lys161Ala) caused a 35−40% reduction in activity, thus highlighting the importance of this domain. The HBD-2 mutation Arg272Ala caused a decrease in activity to 57%, but Lys274Ala had no effect. Lastly, HBD-3 mutations Lys417Ala and Arg428Ala resulted in a reduction of activity to 17 and 35%, respectively. These results are notable in that they suggest that the remote HBD-3 site can affect catalytic activity allosterically.

Larger Oligosaccharides Induce Loss of the Secondary Structure and Aggregation in HPSE. The structural
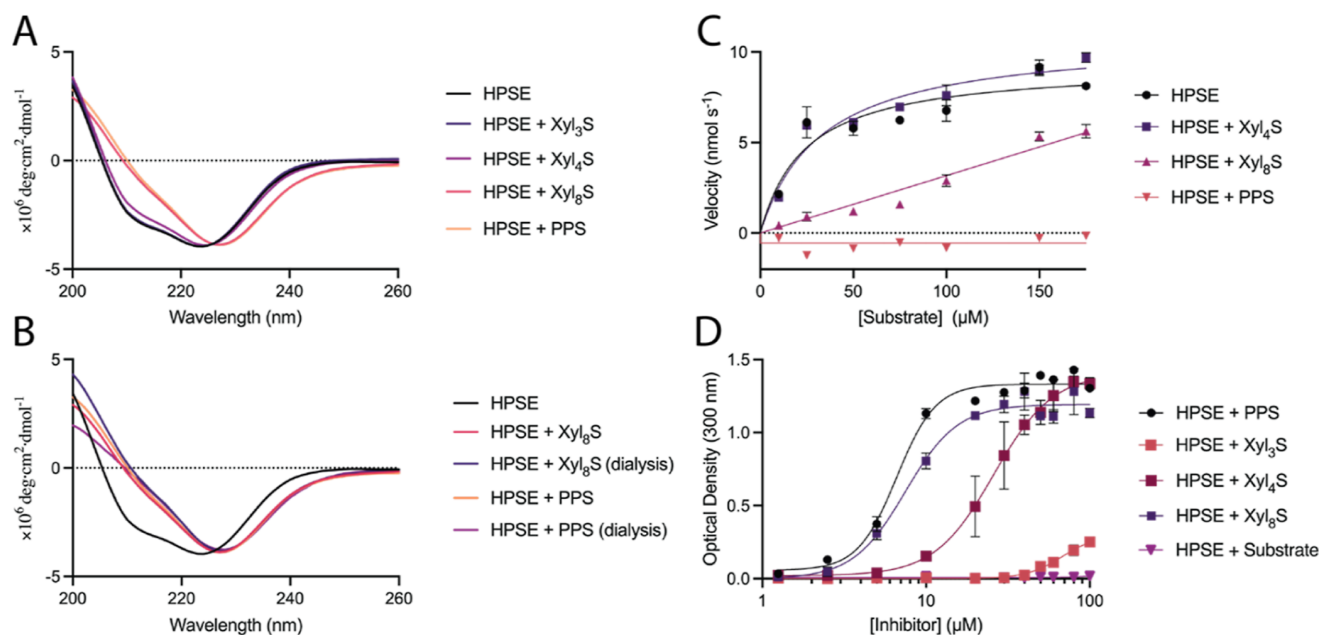
Figure 6. PPS-induced aggregation promotes secondary structure changes and reversibility. (A) CD of HPSE in the presence of PPS and the synthetic analogues shows that secondary structure is altered with longer oligosaccharide lengths. (B) CD of HPSE before and after dialysis shows the secondary structure change is not reversible upon attempted removal of the inhibitor. (C) Michaelis−Menten kinetics of HPSE after removal of PPS and synthetic analogues (by dialysis), showing an increased trend between length and irreversible loss of activity. (D) Optical density assay showing macroscopic aggregation of HPSE with PPS and its synthetic analogues, as well as substrate (fondaparinux), showing that with increased PPS length, there is increased aggregation. A 1:1 molar ratio occurs at 10 μM inhibitor concentrations.

effects of larger oligosaccharides binding to HPSE could not be assessed using X-ray crystallography or HDX-MS techniques due to protein aggregation and structural disruption. To further investigate and probe the link between this and the anomalous enzyme inhibition effects exerted by long-chain oligosaccharides, CD, enzyme kinetic assays, and non-specific aggregation assays were conducted to establish whether PPS and its analogues affect HPSE function other than by direct competitive inhibition.

CD analysis showed that 10 min after the addition of a 10:1 molar ratio of Xyl$_3$S or Xyl$_4$S with HPSE, there was no significant change in the protein secondary structure, as compared to the apo-protein. In contrast, addition of Xyl$_8$S and PPS in the same molar ratio resulted in a large shift in the secondary structure distribution with loss of the α-helix troughs at 210 and 222 nm as well as a shift to a minimum at 228 nm, a change which may be due to aggregation[72] (Figure 6A). Notably, upon removal of Xyl$_8$S and PPS through dialysis, the secondary structure remained altered (Figure 6B). These results are suggestive of a loss of a typical secondary structure for an α/β protein (210 and 222 nm) and gain in CD signatures corresponding to an aberrant and likely aggregated state. We observed the same effect on the protein structure when we tested another long-chain, high-affinity oligosaccharide HPSE inhibitor, namely the drug candidate PI-88 (Figure S5). Electrostatic potential maps generated for HPSE and the two crystal structures of Xyl$_3$S and Xyl$_4$S show a decrease in the positively charged surface, which could potentially play a role in aggregation.

Activity assays were conducted with the dialyzed HPSE samples to understand how changes in secondary structure affected the catalytic activity. HPSE incubated with Xyl$_4$S was able to regain full activity after dialysis (6 h) to remove the bound ligand (Figure 6C). In contrast, Xyl$_8$S-incubated HPSE

only partially regained catalytic activity (15% of WT), and HPSE incubated with PPS did not regain any catalytic activity after dialysis. The observation that the Xyl$_8$S-treated HPSE regains some activity while PPS effects irreversible inactivation suggests that the length and branching variations within the components of PPS could make it impossible to completely remove it from the system due to the increased complexity of the aggregation process(es) or the high affinity of the interaction.

Finally, assays were conducted to confirm the aggregation of HPSE, a feature that may be contributing to long-term inhibitory effects. Optical density measurements were used to identify the level of inhibitor-induced aggregation of HPSE at a concentration of 10 μM (Figure 6). At a 1:1 stoichiometry, (10 μM concentration), PPS and Xyl$_8$S caused aggregation of 80% and 70% HPSE, respectively. In contrast, at 1:1 stoichiometry, Xyl$_4$S resulted in just over 10% aggregation, while Xyl$_3$S resulted in a very low level of aggregation even at 10:1 stoichiometry. The high levels of aggregation detailed above are consistent with the observed changes in secondary structure induced by Xyl$_8$S and PPS (Figure 6A), and even low concentrations of Xyl$_8$S and PPS affect protein structure. The results from the Xyl$_4$S-based studies suggest that although high concentrations of this ligand induce aggregation, the binding interaction is weaker, not producing a change in secondary structure, with the net result being that aggregation is reversible. Importantly, fondaparinux, a synthetic substrate of HPSE, does not induce aggregation (Figure 6D), suggesting that the HS mimetics under study here interact differently than HS or that the ability to cleave the substrate can prevent aggregation. The observation that PI-88 also results in macro-aggregation of HPSE at low stoichiometries suggests that this effect could be a general property of long-chain glycomimetic inhibitors (Figure S5).
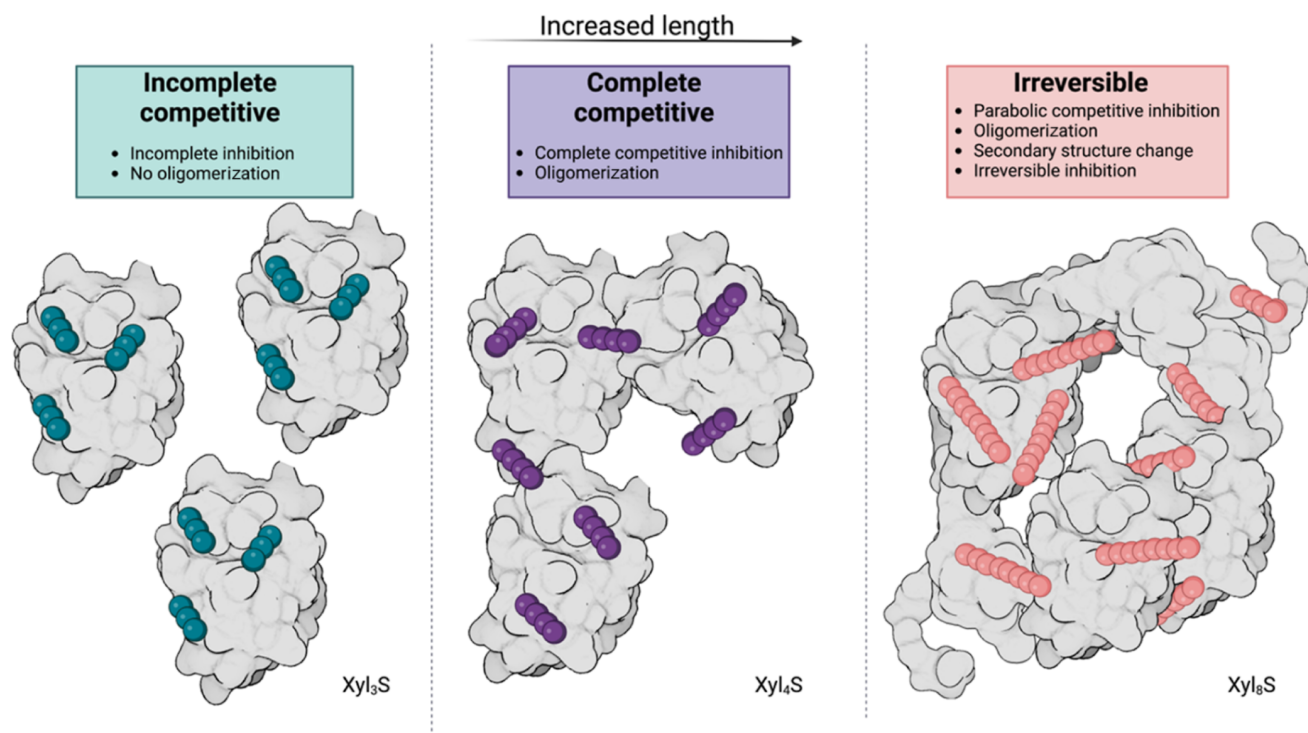
Figure 7. Schematic representation of the HPSE-ligand binding modes with increasing oligosaccharide length. HPSE is shown in light gray, with pentosan glycol units (colored spheres) represented in teal for Xyl_3S, purple for Xyl_4S, and pink for Xyl_8S.

These results explain the anomalous kinetic and thermodynamic data presented in Figures 3 and 4. Specifically, the parabolic inhibition of HPSE by PPS is likely due to the aggregation and simultaneous binding of multiple HPSE enzymes.[69,70] Furthermore, the 100-fold difference between the K_D and K_i for Xyl_8S derives from a time-dependent effect: the affinity for Xyl_8S is likely only on the order of 2 μM, but over the duration of the assay, the protein is inactivated through aggregation and structurally related events, resulting in an apparent K_i of 0.02 μM.

◼ DISCUSSION

HPSE has proven to be a challenging therapeutic target, with few inhibitors reaching clinical trials and, despite decades of effort, no approved HPSE-specific drug. The most promising candidates have been HS glycomimetics but, despite potent inhibition in vitro, these often present undesired side-effects and/or limited benefits have complicated their clinical deployment. This gap between in vitro potency and poor clinical outcomes has been difficult to rationalize based on the current understanding of HPSE. Indeed, undesirable effects of HS-glycomimetics have been noted, such as the observation that chronic use of PPS is associated with a novel pigmentary maculopathy that is attributed to primary retinal pigment epithelium (RPE) injury and toxicity, an effect potentially connected to the drug's interaction with HPSE.[73,74] Our understanding of PPS/HPSE interactions (as well as other HS-glycomimetics) has been limited by a lack of data on the binding interactions with HPSE. This study was undertaken with chemically and structurally well-defined oligosaccharides to further understand the binding interactions of such compounds with HPSE and help overcome the issues presented by the heterogeneity of many HS-glycomimetics.

Our analysis of the PPS binding mechanisms reveals that there are various modes of HPSE inhibition, and these increase in complexity with increasing oligosaccharide length (Figure 7). Three potential binding sites on HPSE have been identified for PPS, overlapping with previously identified (and putative) HS binding domains. While Xyl_3S, which cannot completely span the active site, showed incomplete competitive inhibition, Xyl_4S behaved as a relatively straightforward competitive inhibitor ($K_D \approx K_i$). As the length and interactions increase, as seen with Xyl_8S ($K_D \gg K_i$), a significant increase in protein aggregation is observed. This is followed by a loss of native secondary structure, leading to irreversible inhibition. This effect increases with full-length PPS, which exhibits parabolic competitive inhibition consistent with the binding of multiple proteins by a single PPS molecule.

This multiple protein-binding mode appears to be somewhat generalizable, as other sulfated oligosaccharides, such as PI-88, show similar mechanisms of protein aggregation and secondary structure disruption upon incubation with HPSE. Indeed, the parabolic mode of inhibition observed here for PPS has also been observed in other HPSE inhibitors deployed in clinical trials.[67,68] For example, it has been suggested that Roneparstat (SST0001) binding could be consistent with oligomerisation of HPSE, with molecular dynamics suggesting that Roneparstat binds at the heparan binding domains surrounding the active site and orientating toward the solvent.[68] NMR spectroscopic and MD studies suggest that Pixatimod (PG545) also undergoes multi-inhibitor binding, with two such molecules occupying the area around HBD-1 and -2.[75]

◼ CONCLUSIONS

Overall, this study demonstrates that the length of the oligosaccharide backbones in PPS plays an important role in the inhibition of HPSE and that this is not due to an increased

affinity at the binding site but, rather, results from the ability to interact with numerous HPSE molecules, indirectly affecting inhibition by deactivating HPSE through aggregation. This complex inhibitory mechanism is likely shared by other glycomimetic compounds that have entered clinical trials based on the high potency observed during in vitro studies. Besides aggregation, these inhibitors could also exhibit non-specific binding to other HS-binding proteins, causing off-target effects. The irreversible aggregation of HPSE by these compounds could account for many of the undesirable or limited clinical effects of these drug candidates and highlights the need for potent and specific small-molecule inhibitors. This study should help future drug development programs by allowing for the design of more effective inhibitors of this enzyme.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.biochem.3c00038.

Inhibition dose response curves for pentosan and synthetic analogues; isothermal titration calorimetry for pentosan analogues; crystal structure of Xyl_3S-bound HPSE; shift of HBD-1 residues due to binding of PPS-type inhibitors; aggregation results; electrostatic potential maps; mutagenesis primers; effects of mutations on activity; single-crystal X-ray data and refinement statistics; hydrogen/deuterium exchange heatmap; hydrogen/deuterium exchange data; general synthesis methods; specific transformations; and ^1H, ^{13}C and/or 2D NMR spectra of compounds (PDF)

Accession Codes

HPSE: Q9Y251. Crystal structures: 8E07 (Xly_3S) and 8E08 (Xly_4S).

AUTHOR INFORMATION

Corresponding Author

Colin J. Jackson − *Research School of Chemistry, Australian National University, Canberra, Australian Capital Territory 2601, Australia; Australian Research Council Centre of Excellence for Innovations in Peptide and Protein Science, Australian National University, Canberra, Australian Capital Territory 2601, Australia;* orcid.org/0000-0001-6150-3822; Email: colin.jackson@anu.edu.au

Authors

Cassidy Whitefield − *Research School of Chemistry, Australian National University, Canberra, Australian Capital Territory 2601, Australia; Australian Research Council Centre of Excellence for Innovations in Peptide and Protein Science, Australian National University, Canberra, Australian Capital Territory 2601, Australia;* orcid.org/0000-0002-3387-7744

Yen Vo − *Research School of Chemistry, Australian National University, Canberra, Australian Capital Territory 2601, Australia*

Brett D. Schwartz − *Research School of Chemistry, Australian National University, Canberra, Australian Capital Territory 2601, Australia; Australian Research Council Centre of Excellence for Innovations in Peptide and Protein Science, Australian National University, Canberra, Australian Capital Territory 2601, Australia*

Caryn Hepburn − *Waters Australia Pty Ltd, Rydalmere, New South Wales 2116, Australia*

F. Hafna Ahmed − *Research School of Chemistry, Australian National University, Canberra, Australian Capital Territory 2601, Australia;* orcid.org/0000-0002-9153-2699

Hideki Onagi − *Research School of Chemistry, Australian National University, Canberra, Australian Capital Territory 2601, Australia*

Martin G. Banwell − *Institute for Advanced and Applied Chemical Synthesis, College of Pharmacy, Jinan University, Guangzhou, Guangdong 510632, China*

Keats Nelms − *Beta Therapeutics Pty. Ltd. Level 6, Canberra, Australian Capital Territory 2601, Australia*

Lara R. Malins − *Research School of Chemistry, Australian National University, Canberra, Australian Capital Territory 2601, Australia; Australian Research Council Centre of Excellence for Innovations in Peptide and Protein Science, Australian National University, Canberra, Australian Capital Territory 2601, Australia;* orcid.org/0000-0002-7691-6432

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.biochem.3c00038

Author Contributions

C.W., C.J.J., M.G.B., B.D.S., and K.N. conceived the study. B.D.S. and Y.V. developed synthetic pathways, chemically synthesized the sulfated xylooligosaccharide analogues, and prepared the corresponding supplementary material. C.W. and H.A. separated and tested PPS fractions. C.W. purified and crystallized protein. C.W. collected, processed, and refined X-ray crystallography data with assistance from C.J.J. C.W. performed assays, isothermal titration calorimetry, and circular dichroism studies. C.H. collected and processed HDX-MS data. H.O. provided invaluable assistance in separation science. C.W. and C.J.J. analyzed the results. C.W. prepared the manuscript with input from other authors. C.J.J., M.G.B., and L.R.M. supervised the project.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

REFERENCES

(1) Sarrazin, S.; Lamanna, W. C.; Esko, J. D. Heparan Sulfate Proteoglycans. *Cold Spring Harbor Perspect. Biol.* **2011**, 3, a004952.

(2) Kjellén, L.; Lindahl, U. Proteoglycans: Structures and Interactions. *Annu. Rev. Biochem.* **1991**, 60, 443−475.

(3) Perrimon, N.; Bernfield, M. Specificities of Heparan Sulphate Proteoglycans in Developmental Processes. *Nature* **2000**, 404, 725−728.

(4) Bode, L.; Murch, S.; Freeze, H. H. Heparan Sulfate Plays a Central Role in a Dynamic in Vitro Model of Protein-Losing Enteropathy. *J. Biol. Chem.* **2006**, 281, 7809−7815.

(5) Soares, M. A.; Teixeira, F. C. O. B.; Fontes, M.; Arêas, A. L.; Leal, M. G.; Pavão, M. S. G.; Stelling, M. P. Heparan Sulfate Proteoglycans May Promote or Inhibit Cancer Progression by Interacting with Integrins and Affecting Cell Migration. *BioMed Res. Int.* **2015**, *2015*, e453801–e453808.

(6) Christianson, H. C.; Belting, M. Heparan Sulfate Proteoglycan as a Cell-Surface Endocytosis Receptor. *Matrix Biol.* **2014**, *35*, 51–55.

(7) Knelson, E. H.; Nee, J. C.; Blobe, G. C. Heparan Sulfate Signaling in Cancer. *Trends Biochem. Sci.* **2014**, *39*, 277–288.

(8) Poulain, F. E.; Yost, H. J. Heparan Sulfate Proteoglycans: A Sugar Code for Vertebrate Development? *Development* **2015**, *142*, 3456–3467.

(9) Yayon, A.; Klagsbrun, M.; Esko, J. D.; Leder, P.; Ornitz, D. M. Cell Surface, Heparin-like Molecules Are Required for Binding of Basic Fibroblast Growth Factor to Its High Affinity Receptor. *Cell* **1991**, *64*, 841–848.

(10) Bitan, M.; Weiss, L.; Reibstein, I.; Zeira, M.; Fellig, Y.; Slavin, S.; Zcharia, E.; Nagler, A.; Vlodavsky, I. Heparanase Upregulates Th2 Cytokines, Ameliorating Experimental Autoimmune Encephalitis. *Mol. Immunol.* **2010**, *47*, 1890–1898.

(11) Amara, A.; Lorthioir, O.; Valenzuela, A.; Magerus, A.; Thelen, M.; Montes, M.; Virelizier, J. L.; Delepierre, M.; Baleux, F.; Lortat-Jacob, H.; Arenzana-Seisdedos, F. Stromal Cell-Derived Factor-1α Associates with Heparan Sulfates through the First β-Strand of the Chemokine. *J. Biol. Chem.* **1999**, *274*, 23916–23925.

(12) Liu, J.; Pedersen, L. C. Anticoagulant Heparan Sulfate: Structural Specificity and Biosynthesis. *Appl. Microbiol. Biotechnol.* **2007**, *74*, 263–272.

(13) Vlodavsky, I.; Friedmann, Y.; Elkin, M.; Aingorn, H.; Atzmon, R.; Ishai-Michaeli, R.; Bitan, M.; Pappo, O.; Peretz, T.; Michal, I.; Spector, L.; Pecker, I. Mammalian Heparanase: Gene Cloning, Expression and Function in Tumor Progression and Metastasis. *Nat. Med.* **1999**, *5*, 793–802.

(14) Hulett, M. D.; Freeman, C.; Hamdorf, B. J.; Baker, R. T.; Harris, M. J.; Parish, C. R. Cloning of Mammalian Heparanase, an Important Enzyme in Tumor Invasion and Metastasis. *Nat. Med.* **1999**, *5*, 803–809.

(15) Kussie, P. H.; Hulmes, J. D.; Ludwig, D. L.; Patel, S.; Navarro, E. C.; Seddon, A. P.; Giorgio, N. A.; Bohlen, P. Cloning and Functional Expression of a Human Heparanase Gene. *Biochem. Biophys. Res. Commun.* **1999**, *261*, 183–187.

(16) Toyoshima, M.; Nakajima, M. Human Heparanase: Purification, Characterization, Cloning, and Expression. *J. Biol. Chem.* **1999**, *274*, 24153–24160.

(17) Arvatz, G.; Weissmann, M.; Ilan, N.; Vlodavsky, I. Heparanase and Cancer Progression: New Directions, New Promises. *Hum. Vaccines Immunother.* **2016**, *12*, 2253–2256.

(18) Buijsers, B.; Yanginlar, C.; de Nooijer, A.; Grondman, I.; Maciej-Hulme, M.-H.; Jonkman, J.; Janssen, N. A. F.; Rother, N.; et al. Increased Plasma Heparanase Activity in COVID-19 Patients. *Front. Immunol.* **2020**, *11*, 575047.

(19) Bar-Ner, M.; Mayer, M.; Schirrmacher, V.; Vlodavsky, I. Involvement of Both Heparanase and Plasminogen Activator in Lymphoma Cell-Mediated Degradation of Heparan Sulfate in the Subendothelial Extracellular Matrix. *J. Cell. Physiol.* **1986**, *128*, 299–306.

(20) Elkin, M.; Ilan, N.; Ishai-Michaeli, R.; Friedmann, Y.; Papo, O.; Pecker, I.; Vlodavsky, I. Heparanase as Mediator of Angiogenesis: Mode of Action. *FASEB J.* **2001**, *15*, 1661–1663.

(21) Digre, A.; Singh, K.; Åbrink, M.; Reijmers, R. M.; Sandler, S.; Vlodavsky, I.; Li, J.-P. Overexpression of Heparanase Enhances T Lymphocyte Activities and Intensifies the Inflammatory Response in a Model of Murine Rheumatoid Arthritis. *Sci. Rep.* **2017**, *7*, 46229.

(22) Sasaki, N.; Higashi, N.; Taka, T.; Nakajima, M.; Irimura, T. Cell Surface Localization of Heparanase on Macrophages Regulates Degradation of Extracellular Matrix Heparan Sulfate. *J. Immunol.* **2004**, *172*, 3830–3835.

(23) Gallagher, J. T. Heparan Sulfate: Growth Control with a Restricted Sequence Menu. *J. Clin. Invest.* **2001**, *108*, 357–361.

(24) Ishai-Michaeli, R.; Eldor, A.; Vlodavsky, I. Heparanase Activity Expressed by Platelets, Neutrophils, and Lymphoma Cells Releases Active Fibroblast Growth Factor from Extracellular Matrix. *Cell Regul.* **1990**, *1*, 833–842.

(25) Gutter-Kapon, L.; Alishekevitz, D.; Shaked, Y.; Li, J. P.; Aronheim, A.; Ilan, N.; Vlodavsky, I. Heparanase Is Required for Activation and Function of Macrophages. *Proc. Natl. Acad. Sci. U.S.A.* **2016**, *113*, E7808–E7817.

(26) Bitan, M.; Weiss, L.; Reibstein, I.; Zeira, M.; Fellig, Y.; Slavin, S.; Zcharia, E.; Nagler, A.; Vlodavsky, I. Heparanase Upregulates Th2 Cytokines, Ameliorating Experimental Autoimmune Encephalitis. *Mol. Immunol.* **2010**, *47*, 1890–1898.

(27) Parish, C. R.; Freeman, C.; Brown, K. J.; Francis, D. J.; Cowden, W. B.; Hampson, I. N. Identification of Sulfated Oligosaccharide-Based Inhibitors of Tumor Growth and Metastasis Using Novel in Vitro Assays for Angiogenesis and Heparanase Activity. *Cancer Res.* **1999**, *59*, 3433–3441.

(28) Ferro, V.; Liu, L.; Johnstone, K. D.; Wimmer, N.; Karoli, T.; Handley, P.; Rowley, J.; Dredge, K.; Li, C. P.; Hammond, E.; Davis, K.; Sarimaa, L.; Harenberg, J.; Bytheway, I. Discovery of PG545: A Highly Potent and Simultaneous Inhibitor of Angiogenesis, Tumor Growth, and Metastasis. *J. Med. Chem.* **2012**, *55*, 3804–3813.

(29) Naggi, A.; Casu, B.; Perez, M.; Torri, G.; Cassinelli, G.; Penco, S.; Pisano, C.; Giannini, G.; Ishai-Michaeli, R.; Vlodavsky, I. Modulation of the Heparanase-Inhibiting Activity of Heparin through Selective Desulfation, Graded N-Acetylation, and Glycol Splitting. *J. Biol. Chem.* **2005**, *280*, 12103–12113.

(30) Zhou, H.; Roy, S.; Cochran, E.; Zouaoui, R.; Chu, C. L.; Duffner, J.; Zhao, G.; Smith, S.; Galcheva-Gargova, Z.; Karlgren, J.; Dussault, N.; Kwan, R. Y. Q.; Moy, E.; Barnes, M.; Long, A.; Honan, C.; Qi, Y. W.; Shriver, Z.; Ganguly, T.; Schultes, B.; Venkataraman, G.; Kishimoto, T. K. M402, a Novel Heparan Sulfate Mimetic, Targets Multiple Pathways Implicated in Tumor Progression and Metastasis. *PLoS One* **2011**, *6*, No. e21106.

(31) Xu, Y.-J.; Miao, H.-Q.; Pan, W.; Navarro, E. C.; Tonra, J. R.; Mitelman, S.; Camara, M. M.; Deevi, D. S.; Kiselyov, A. S.; Kussie, P.; Wong, W. C.; Liu, H. N-(4-{[4-(1H-Benzoimidazol-2-Yl)-Arylami-no]-Methyl}-Phenyl)-Benzamide Derivatives as Small Molecule Heparanase Inhibitors. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 404–408.

(32) Courtney, S. M.; Hay, P. A.; Buck, R. T.; Colville, C. S.; Phillips, D. J.; Scopes, D. I. C.; Pollard, F. C.; Page, M. J.; Bennett, J. M.; Hircock, M. L.; McKenzie, E. A.; Bhaman, M.; Felix, R.; Stubberfield, C. R.; Turner, P. R. Furanyl-1,3-Thiazol-2-Yl and Benzoxazol-5-Yl Acetic Acid Derivatives: Novel Classes of Heparanase Inhibitor. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 2295–2299.

(33) Madia, V. N.; Messore, A.; Pescatori, L.; Saccoliti, F.; Tudino, V.; De Leo, A.; Bortolami, M.; Scipione, L.; Costi, R.; Rivara, S.; Scalvini, L.; Mor, M.; Ferrara, F. F.; Pavoni, E.; Roscilli, G.; Cassinelli, G.; Milazzo, F. M.; Battistuzzi, G.; Di Santo, R.; Giannini, G. Novel Benzazole Derivatives Endowed with Potent Antiheparanase Activity. *J. Med. Chem.* **2018**, *61*, 6918–6936.

(34) Simmons, S. C.; McKenzie, E. A.; Harris, L. K.; Aplin, J. D.; Brenchley, P. E.; Velasco-Garcia, M. N.; Missailidis, S. Development of Novel Single-Stranded Nucleic Acid Aptamers against the Pro-Angiogenic and Metastatic Enzyme Heparanase (HPSE1). *PLoS ONE* **2012**, *7*, No. e37938.

(35) de Boer, C.; Armstrong, Z.; Lit, V. A. J.; Barash, U.; Ruijgrok, G.; Boyango, I.; Weitzenberg, M. M.; Schröder, S. P.; Sarris, A. J. C.; Meeuwenoord, N. J.; Bule, P.; Kayal, Y.; Ilan, N.; Codée, J. D. C.; Vlodavsky, I.; Overkleeft, H. S.; Davies, G. J.; Wu, L. Mechanism-Based Heparanase Inhibitors Reduce Cancer Metastasis in Vivo. *Proc. Natl. Acad. Sci. U.S.A.* **2022**, *119*, No. e2203167119.

(36) He, X.; Brenchley, P. E. C.; Jayson, G. C.; Hampson, L.; Davies, J.; Hampson, I. N. Hypoxia Increases Heparanase-Dependent Tumor Cell Invasion, Which Can Be Inhibited by Antiheparanase Antibodies. *Cancer Res.* **2004**, *64*, 3928–3933.

(37) Temkin, V.; Aingorn, H.; Puxeddu, I.; Goldshmidt, O.; Zcharia, E.; Gleich, G. J.; Vlodavsky, I.; Levi-Schaffer, F. Eosinophil Major

Basic Protein: First Identified Natural Heparanase-Inhibiting Protein. *J. Allergy Clin. Immunol.* **2004**, *113*, 703−709.

(38) Shiozawa, H.; Takahashi, M.; Takatsu, T.; Kinoshita, T.; Tanzawa, K.; Hosoya, T.; Furuya, K.; Takahashi, S.; Furihata, K.; Seto, H. Trachyspic Acid, a New Metabolite Produced by Talaromyces Trachyspermus, That Inhibits Tumor Cell Heparanase: Taxonomy of the Producing Strain, Fermentation, Isolation, Structural Elucidation, and Biological Activity. *J. Antibiot.* **1995**, *48*, 357−362.

(39) Jia, L.; Ma, S. Recent Advances in the Discovery of Heparanase Inhibitors as Anti-Cancer Agents. *Eur. J. Med. Chem.* **2016**, *121*, 209−220.

(40) Heyman, B.; Yang, Y. Mechanisms of Heparanase Inhibitors in Cancer Therapy. *Exp. Hematol.* **2016**, *44*, 1002−1012.

(41) Davar, D.. Phase IIA Basket Study of Pixatimod (PG545) in Combination With Nivolumab in PD-1 Relapsed/Refractory Metastatic Melanoma and NSCLC and Pixatimod (PG545) in Combination With Nivolumab and Low-Dose Cyclophosphamide in MSS Metastatic Colorectal Carcinoma (MCRC); *Clinical Trial Registration NCT05061017; clinicaltrials.gov*, 2022. https://clinicaltrials.gov/ct2/show/NCT05061017 (accessed 2022-11-28).

(42) Galli, M.; Chatterjee, M.; Grasso, M.; Specchia, G.; Magen, H.; Einsele, H.; Celeghini, I.; Barbieri, P.; Paoletti, D.; Pace, S.; Sanderson, R. D.; Rambaldi, A.; Nagler, A. Phase I Study of the Heparanase Inhibitor Roneparstat: An Innovative Approach for Ultiple Myeloma Therapy. *Haematologica* **2018**, *103*, e469−e472.

(43) O'Reilly, E. M.; Barone, D.; Mahalingam, D.; Bekaii-Saab, T.; Shao, S. H.; Wolf, J.; Rosano, M.; Krause, S.; Richards, D. A.; Yu, K. H.; Roach, J. M.; Flaherty, K. T.; Ryan, D. P. Randomised Phase II Trial of Gemcitabine and Nab-Paclitaxel with Necuparanib or Placebo in Untreated Metastatic Pancreas Ductal Adenocarcinoma. *Eur. J. Cancer* **2020**, *132*, 112−121.

(44) Khasraw, M.; Pavlakis, N.; McCowatt, S.; Underhill, C.; Begbie, S.; de Souza, P.; Boyce, A.; Parnis, F.; Lim, V.; Harvie, R.; Marx, G. Multicentre Phase I/II Study of PI-88, a Heparanase Inhibitor in Combination with Docetaxel in Patients with Metastatic Castrate-Resistant Prostate Cancer. *Ann. Oncol.* **2010**, *21*, 1302−1307.

(45) Schuchman, E. H.; Ge, Y.; Lai, A.; Borisov, Y.; Faillace, M.; Eliyahu, E.; He, X.; Iatridis, J.; Vlassara, H.; Striker, G.; Simonaro, C. M. Pentosan Polysulfate: A Novel Therapy for the Mucopolysaccharidoses. *PLoS One* **2013**, *8*, No. e54459.

(46) Anger, J. T.; Zabihi, N.; Clemens, J. Q.; Payne, C. K.; Saigal, C. S.; Rodriguez, L. V. Treatment Choice, Duration, and Cost in Patients with Interstitial Cystitis and Painful Bladder Syndrome. *Int. UrogynEcol. J.* **2011**, *22*, 395−400.

(47) Kongtawelert, P.; Brooks, P. M.; Ghosh, P. Pentosan Polysulfate (Cartrophen) Prevents the Hydrocortisone Induced Loss of Hyaluronic Acid and Proteoglycans from Cartilage of Rabbit Joints as Well as Normalizes the Keratan Sulfate Levels in Their Serum. *J. Rheumatol.* **1989**, *16*, 1454−1459.

(48) Sanden, C.; Mori, M.; Jogdand, P.; Jönsson, J.; Krishnan, R.; Wang, X.; Erjefält, J. S. Broad Th2 Neutralization and Anti-Inflammatory Action of Pentosan Polysulfate Sodium in Experimental Allergic Rhinitis. *Immun., Inflammation Dis.* **2017**, *5*, 300−309.

(49) Parish, C. R.; Coombe, D. R.; Jakobsen, K. B.; Bennett, F. A.; Underwood, P. A. Evidence That Sulphated Polysaccharides Inhibit Tumour Metastasis by Blocking Tumour-Cell-Derived Heparanases. *Int. J. Cancer* **1987**, *40*, 511−518.

(50) Whitefield, C.; Hong, N.; Mitchell, J. A.; Jackson, C. J. Computational Design and Experimental Characterisation of a Stable Human Heparanase Variant. *RSC Chem. Biol.* **2022**, *3*, 341−349.

(51) Vo, Y.; Schwartz, B. D.; Onagi, H.; Ward, J. S.; Gardiner, M. G.; Banwell, M. G.; Nelms, K.; Malins, L. R. A Rapid and Mild Sulfation Strategy Reveals Conformational Preferences in Therapeutically Relevant Sulfated Xylooligosaccharides. *Chem.—Eur. J.* **2021**, *27*, 9830−9838.

(52) Gibson, D. G. Enzymatic Assembly of Overlapping DNA Fragments. In *Synthetic Biology, Part B*; Voigt, C., Ed.; Academic Press, 2011; Vol. *498*, Chapter 15, pp 349−361.

(53) Hammond, E.; Li, C. P.; Ferro, V. Development of a Colorimetric Assay for Heparanase Activity Suitable for Kinetic Analysis and Inhibitor Screening. *Anal. Biochem.* **2010**, *396*, 112−116.

(54) Cao, R.; Zeidan, A. A.; Rådström, P.; van Niel, E. W. J. Inhibition Kinetics of Catabolic Dehydrogenases by Elevated Moieties of ATP and ADP − Implication for a New Regulation Mechanism in Lactococcus Lactis. *FEBS J.* **2010**, *277*, 1843−1852.

(55) Hyperbolic and Parabolic Inhibition. In *Comprehensive Enzyme Kinetics*; Leskovac, V., Ed.; Springer US: Boston, MA, 2003; pp 95−110.

(56) Aragão, D.; Aishima, J.; Cherukuvada, H.; Clarken, R.; Clift, M.; Cowieson, N. P.; Ericsson, D. J.; Gee, C. L.; Macedo, S.; Mudie, N.; Panjikar, S.; Price, J. R.; Riboldi-Tunnicliffe, A.; Rostan, R.; Williamson, R.; Caradoc-Davies, T. T. MX2: A High-Flux Undulator Microfocus Beamline Serving Both the Chemical and Macromolecular Crystallography Communities at the Australian Synchrotron. *J Synchrotron Radiat.* **2018**, *25*, 885−891.

(57) Winter, G.; Waterman, D. G.; Parkhurst, J. M.; Brewster, A. S.; Gildea, R. J.; Gerstel, M.; Fuentes-Montero, L.; Vollmar, M.; Michels-Clark, T.; Young, I. D.; Sauter, N. K.; Evans, G. DIALS: Implementation and Evaluation of a New Integration Package. *Acta Crystallogr., Sect. D: Struct. Biol.* **2018**, *74*, 85−97.

(58) Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G. W.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S. Overview of the CCP4 Suite and Current Developments. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2011**, *67*, 235−242.

(59) Karplus, P. A.; Diederichs, K. Linking Crystallographic Model and Data Quality. *Science* **2012**, *336*, 1030−1033.

(60) Afonine, P. V.; Grosse-Kunstleve, R. W.; Echols, N.; Headd, J. J.; Moriarty, N. W.; Mustyakimov, M.; Terwilliger, T. C.; Urzhumtsev, A.; Zwart, P. H.; Adams, P. D. Towards Automated Crystallographic Structure Refinement with Phenix.Refine. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2012**, *68*, 352−367.

(61) Emsley, P.; Cowtan, K. Coot: Model-Building Tools for Molecular Graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *60*, 2126−2132.

(62) Chen, V. B.; Arendall, W. B.; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C. MolProbity: All-Atom Structure Validation for Macromolecular Crystallography. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 12−21.

(63) Afonine, P. V.; Grosse-Kunstleve, R. W.; Chen, V. B.; Headd, J. J.; Moriarty, N. W.; Richardson, J. S.; Richardson, D. C.; Urzhumtsev, A.; Zwart, P. H.; Adams, P. D. *phenix.model_vs_data*: a high-level tool for the calculation of crystallographic model and data statistics. *J. Appl. Crystallogr.* **2010**, *43*, 669−676.

(64) DeLano, W. L. *The PyMOL Molecular Graphics System*. DeLano Scientific, 2002.

(65) Kinoshita, K.; Murakami, Y.; Nakamura, H. EF.-S. eF-seek: prediction of the functional sites of proteins by searching for similar electrostatic potential and molecular surface shape. *Nucleic Acids Res.* **2007**, *35*, W398−W402.

(66) Wu, L.; Viola, C. M.; Brzozowski, A. M.; Davies, G. J. Structural Characterization of Human Heparanase Reveals Insights into Substrate Recognition. *Nat. Struct. Mol. Biol.* **2015**, *22*, 1016−1022.

(67) Hammond, E.; Handley, P.; Dredge, K.; Bytheway, I. Mechanisms of Heparanase Inhibition by the Heparan Sulfate Mimetic PG545 and Three Structural Analogues. *FEBS Open Bio* **2013**, *3*, 346−351.

(68) Pala, D.; Rivara, S.; Mor, M.; Milazzo, F. M.; Roscilli, G.; Pavoni, E.; Giannini, G. Kinetic Analysis and Molecular Modeling of the Inhibition Mechanism of Roneparstat (SST0001) on Human Heparanase. *Glycobiology* **2016**, *26*, 640−654.

(69) Acker, M. G.; Auld, D. S. Considerations for the Design and Reporting of Enzyme Assays in High-Throughput Screening Applications. *Perspect. Sci.* **2014**, *1*, 56−73.

(70) Prinz, H.; Schönichen, A. Transient Binding Patches: A Plausible Concept for Drug Binding. *J. Chem. Biol.* **2008**, *1*, 95−104.

(71) Levy-Adam, F.; Abboud-Jarrous, G.; Guerrini, M.; Beccati, D.; Vlodavsky, I.; Ilan, N. Identification and Characterization of Heparin/ Heparan Sulfate Binding Domains of the Endoglycosidase Heparanase. *J. Biol. Chem.* **2005**, *280*, 20457−20466.

(72) Mu, X.; Eckes, K. M.; Nguyen, M. M.; Suggs, L. J.; Ren, P. Experimental and Computational Studies Reveal an Alternative Supramolecular Structure for Fmoc-Dipeptide Self-Assembly. *Biomacromolecules* **2012**, *13*, 3562−3571.

(73) Van Bergen, T.; Etienne, I.; Jia, J.; Li, J.-P.; Vlodavsky, I.; Stitt, A.; Vermassen, E.; Feyen, J. H. M. Heparanase Deficiency Is Associated with Disruption, Detachment, and Folding of the Retinal Pigment Epithelium. *Curr. Eye Res.* **2021**, *46*, 1166−1170.

(74) Pearce, W. A.; Chen, R.; Jain, N. Pigmentary Maculopathy Associated with Chronic Exposure to Pentosan Polysulfate Sodium. *Ophthalmology* **2018**, *125*, 1793−1802.

(75) Chhabra, M.; Wilson, J. C.; Wu, L.; Davies, G. J.; Gandhi, N. S.; Ferro, V. Structural Insights into Pixatimod (PG545) Inhibition of Heparanase, a Key Enzyme in Cancer and Viral Infections. *Chem.— Eur. J.* **2022**, *28*, No. e202104222.

Supporting Information for:

Complex Inhibitory Mechanism of Glycomimetics with Heparanase

Cassidy Whitefield[ab], Yen Vo[a], Brett D. Schwartz[ab], Caryn Hepburn[c], F. Hafna Ahmed[a], Hideki Onagi[a], Martin G. Banwell[d], Keats Nelms[e], Lara R. Malins[ab], Colin J. Jackson[ab]*

[a] Research School of Chemistry, Australian National University, Canberra, ACT, 2601, Australia
[b] Australian Research Council Centre of Excellence for Innovations in Peptide and Protein Science, Australian National University, Canberra, ACT 2601, Australia
[c] Waters Australia Pty Ltd, 38-46 South Str, Rydalmere, NSW, 2116, Australia
[d] Institute for Advanced and Applied Chemical Synthesis, College of Pharmacy, Jinan University, Guangzhou, Guangdong, 510632 China
[e] Beta Therapeutics Pty. Ltd. Level 6, 121 Marcus Clarke Street, Canberra, ACT 2601 Australia
* E-mail: colin.jackson@anu.edu.au

Table of Contents

Specific Transformations **S-20**

References **S-25**

**¹H, ¹³C and/or 2D NMR Spectra of Compounds** **S-26**

All sulfated compounds are sodiated but the sodium counter-ions have been omitted for clarity.

Inhibition dose response curves for Pentosan and synthetic analogues
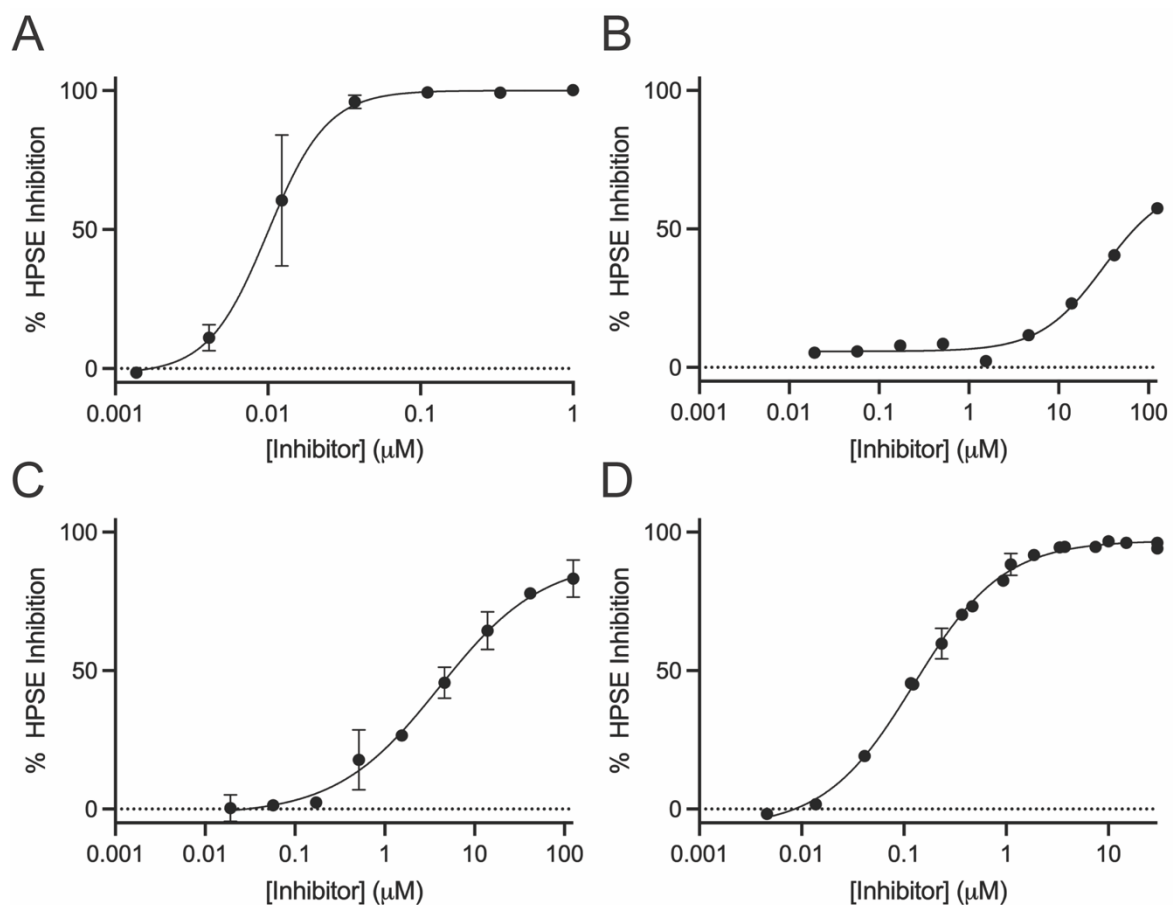

Figure S1 Inhibition dose response curves for Pentosan and its synthetic analogues. A) Pentosan dose response curve. **B)** Xyl_3S dose response curve. **C)** Xyl_4S dose response curve. **D)** Xyl_8S dose response curve. Error bars show standard error of 2-4 replicates

Isothermal titration calorimetry
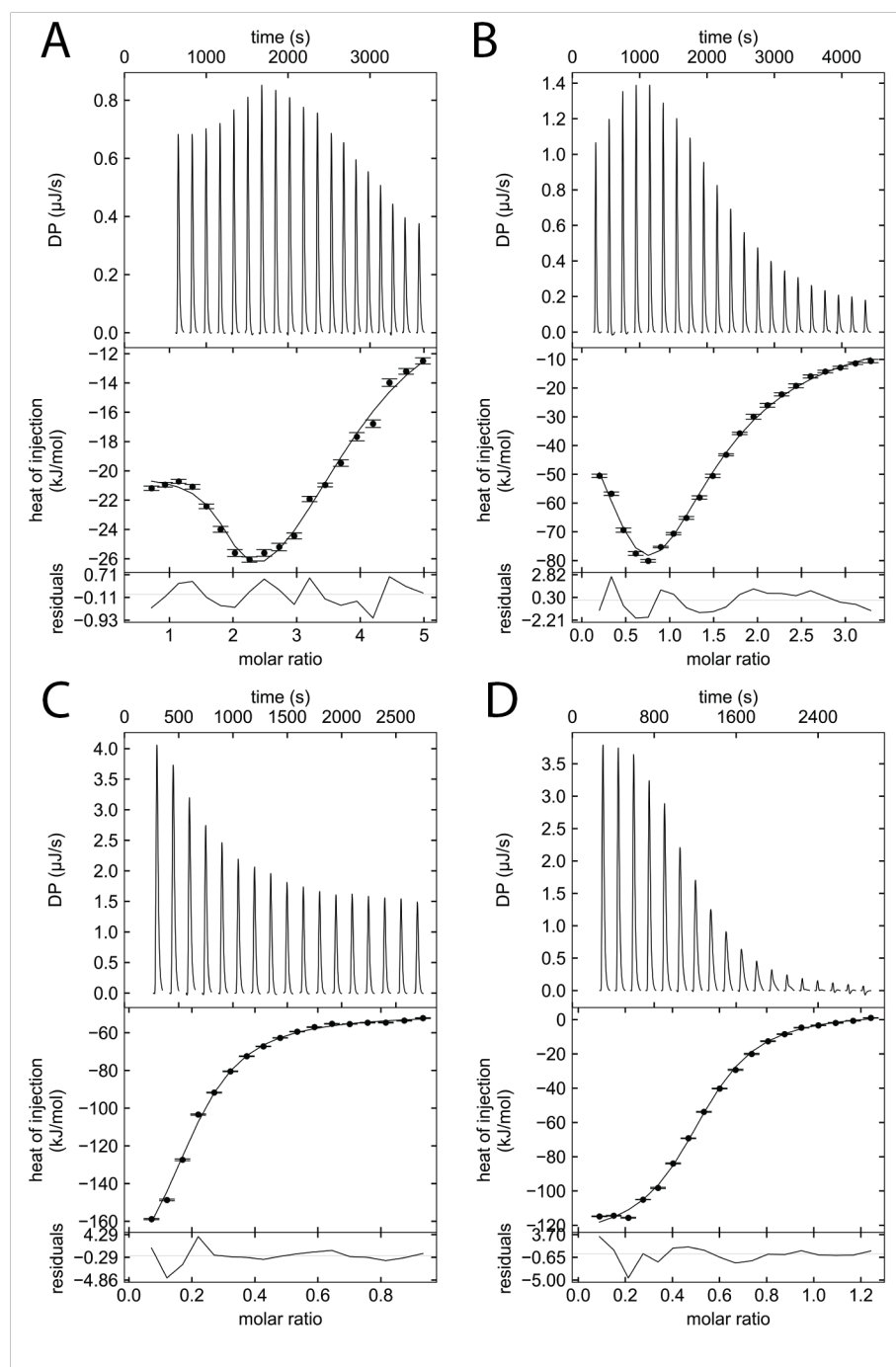
Figure S2 Isothermal titration calorimetry for Pentosan analogues. A-B) showing a single forward titration of Xyl_4SO_4 and Xyl_8SO_4. **C-D)** show a single reverse titration for Xyl_4SO_4 and Xyl_8SO_4, respectively.

Crystal structure of Xyl₃S-bound HPSE
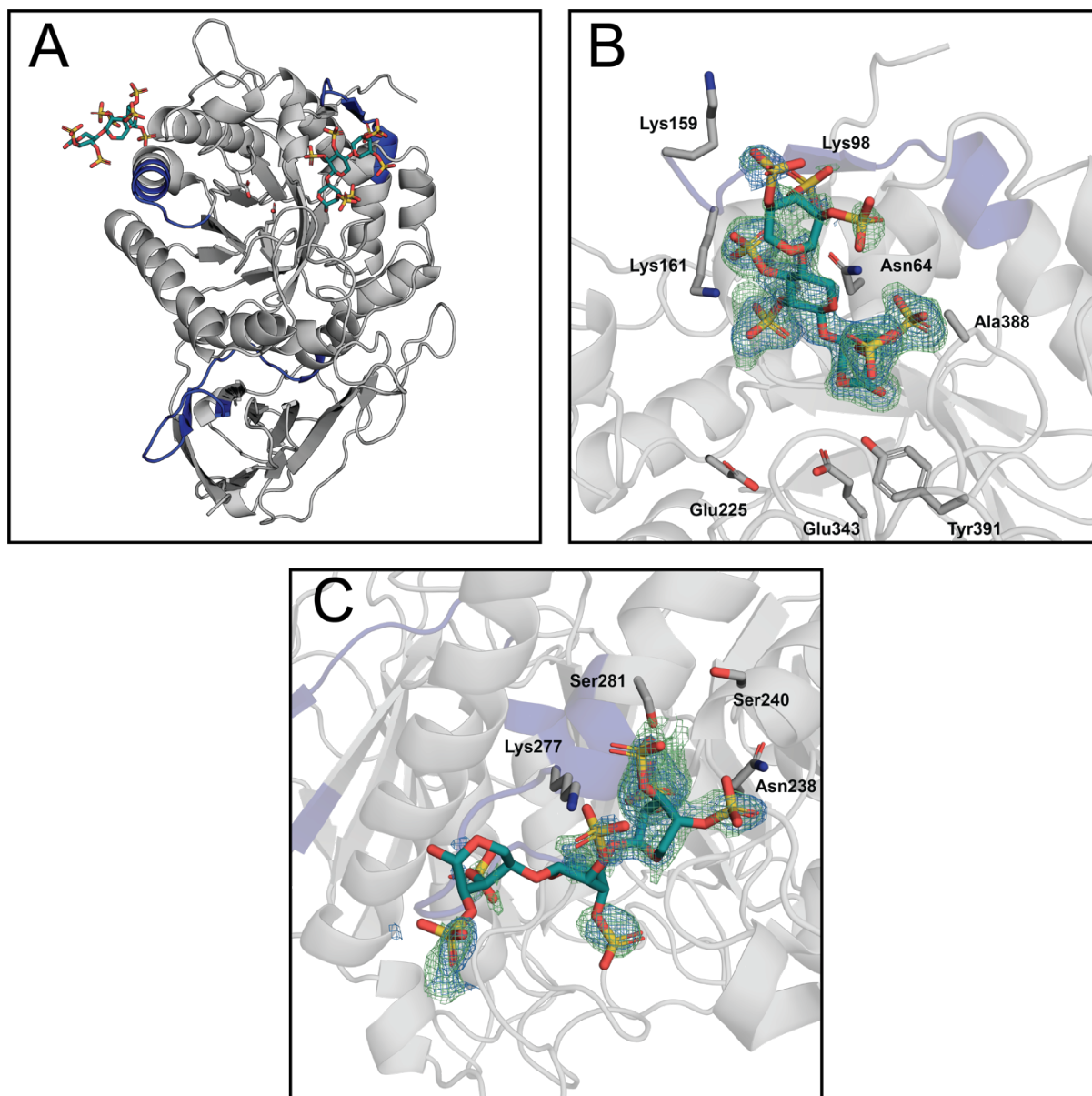
Figure S3 Crystal structure of the Xyl₃S-bound HPSE (8E07). **A)** Crystal structure of Xyl₃S bound to HPSE, with HBDs highlighted in blue **B)** Shows a zoomed in view of the Xyl₃S ligand with density at HBD-1, **C)** HBD-2. Fo-Fc maps represented in green (2.5σ) and the 2Fo-Fc map in blue (1σ).

**Shift of HBD-1 residues due to binding of PPS-type inhibitors**
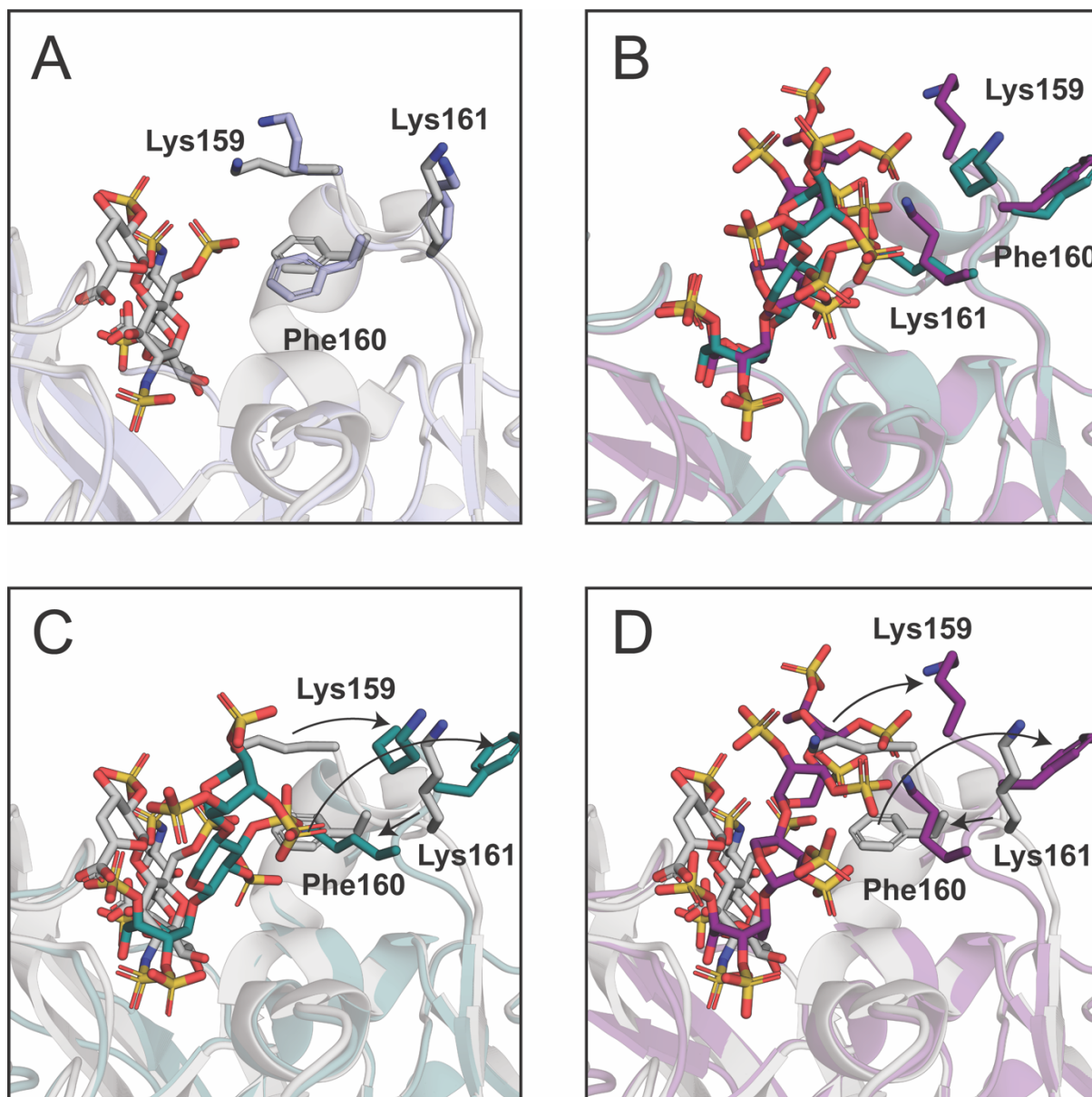
Figure S4. Shift of HBD-1 residues upon binding of PPS-type inhibitors. **A)** Binding of the substrate analogue DP4 in grey (PDB: 5E9C) compared to apo form in light blue (PDB: 7RG8) with residues of HBD-1 highlighted. **B)** Binding of Xyl_3S in teal (PDB: 8E07) and Xyl_4S in purple (PDB: 8E08) with residues of HBD-1 highlighted. **C)** Binding of Xyl_3S in teal and **D)** of Xyl_4S in purple, compared to the substrate bound form of DP4, showing large shifts of HBD-1 residue conformation. Amino acids are labelled in their original position before Inhibitor binding.

Aggregation Results
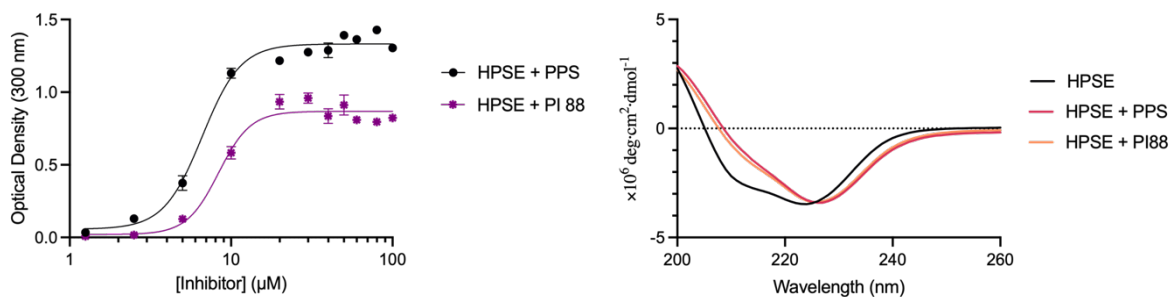
Figure S5. Aggregation assays of other known heparan sulfate mimetic inhibitors. **A)** aggregation response curves of PI88 and PPS. **B)** Circular dichroism of HPSE in the presence of PPS and another HS-mimetic, PI 88, demonstrating that secondary structure is altered by such mimetics.

Electrostatic Potential Maps
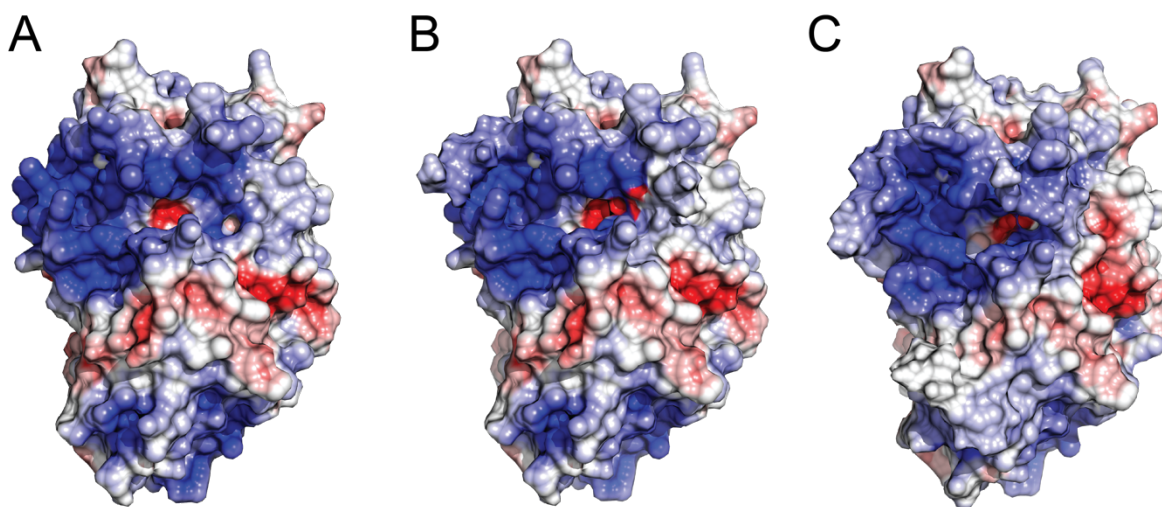
Figure S6. Electrostatic potential maps of HPSE and Xyl$_x$S structures. **A**, **B**, and **C** show the map of HPSE, HPSE + Xyl$_3$S and HPSE + Xyl$_4$S respectively showing a change in the electropositive patches around the HS binding domains.

Mutagenesis primers

Table S1: Primers for HPSE P6 point mutations.

Name	Sequence
K417A_F	CCTCCTTTTCAAAAAGTTAGTTGGTACAGCGGTGTTGATGGCAAGTG
K417A_R	CCCTGAACACTTGCCATCAACACCGCTGTACCAACTAACTTTTTG
R428A_F	GTGTTCAGGGTCAGGATGCGCGCAAACTTCGC
R428A_R	GATAAACGCGAAGTTTGCGCGCATCCTGACCCTGAAC
Mid_Duet_F	CCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGC
Mid_Duet_R	GCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGG

Hydrogen/Deuterium Exchange Heatmap
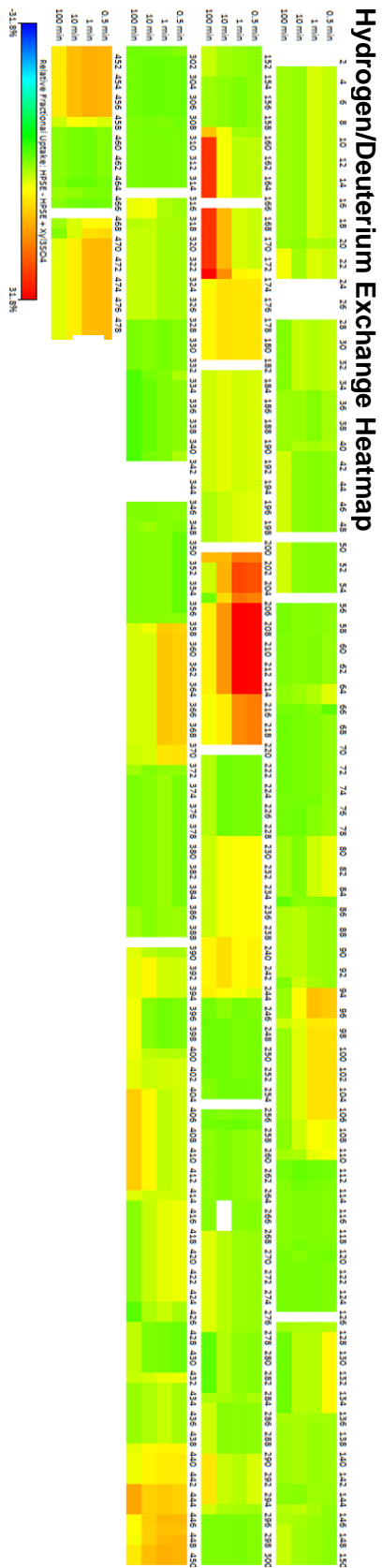
Figure S6. Hydrogen/deuterium-exchange mass spectrometry heat map representing the relative fractional uptake of individual peptides of HPSE relative to Xyl$_3$S after 0.5-, 1-, 10-, and 100-minute incubation. Numbering of 1 starts from his-tagged 8 kDa subunit, with the 50 kDa subunit starting at 93.

Hydrogen/Deuterium Exchange Data
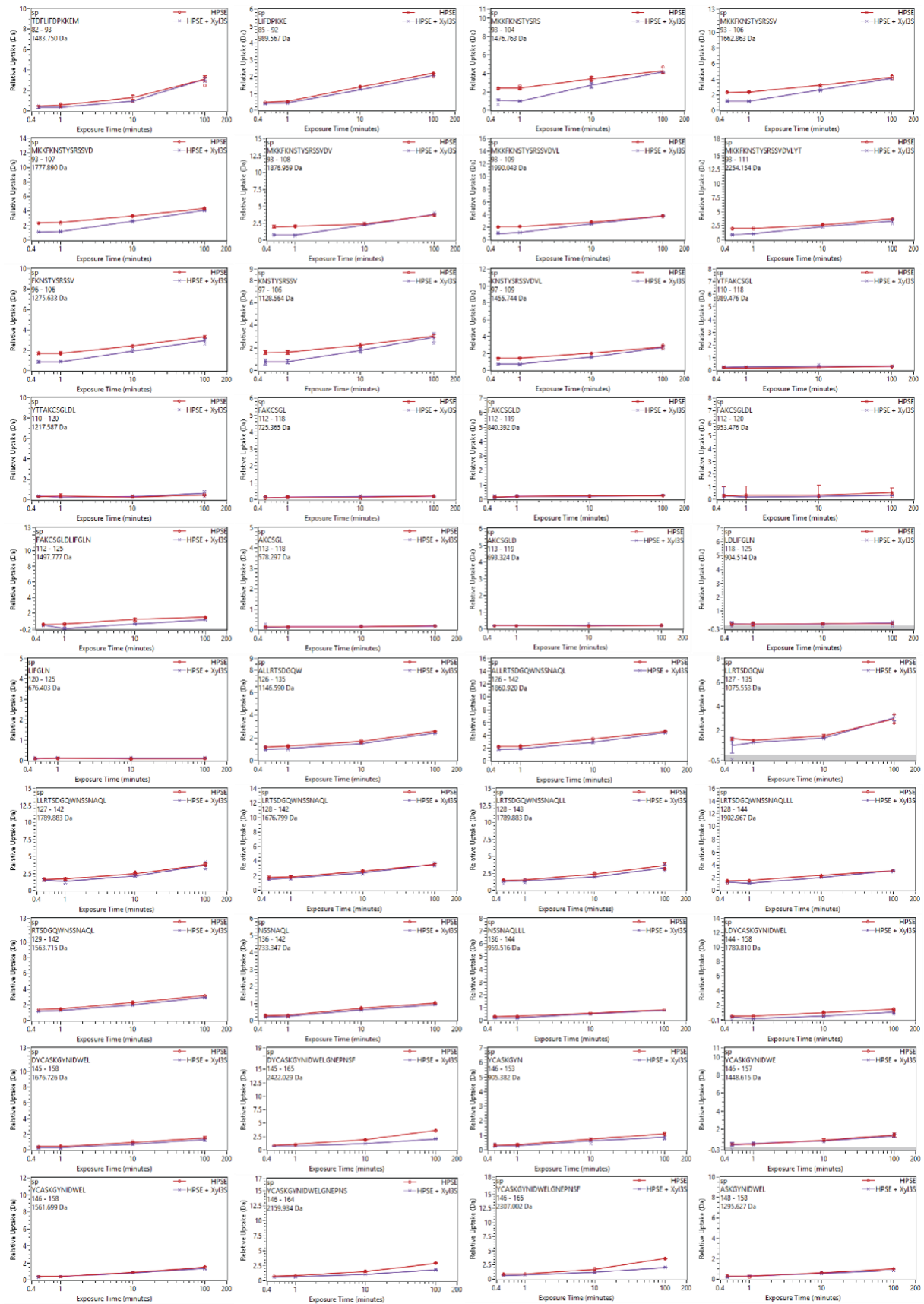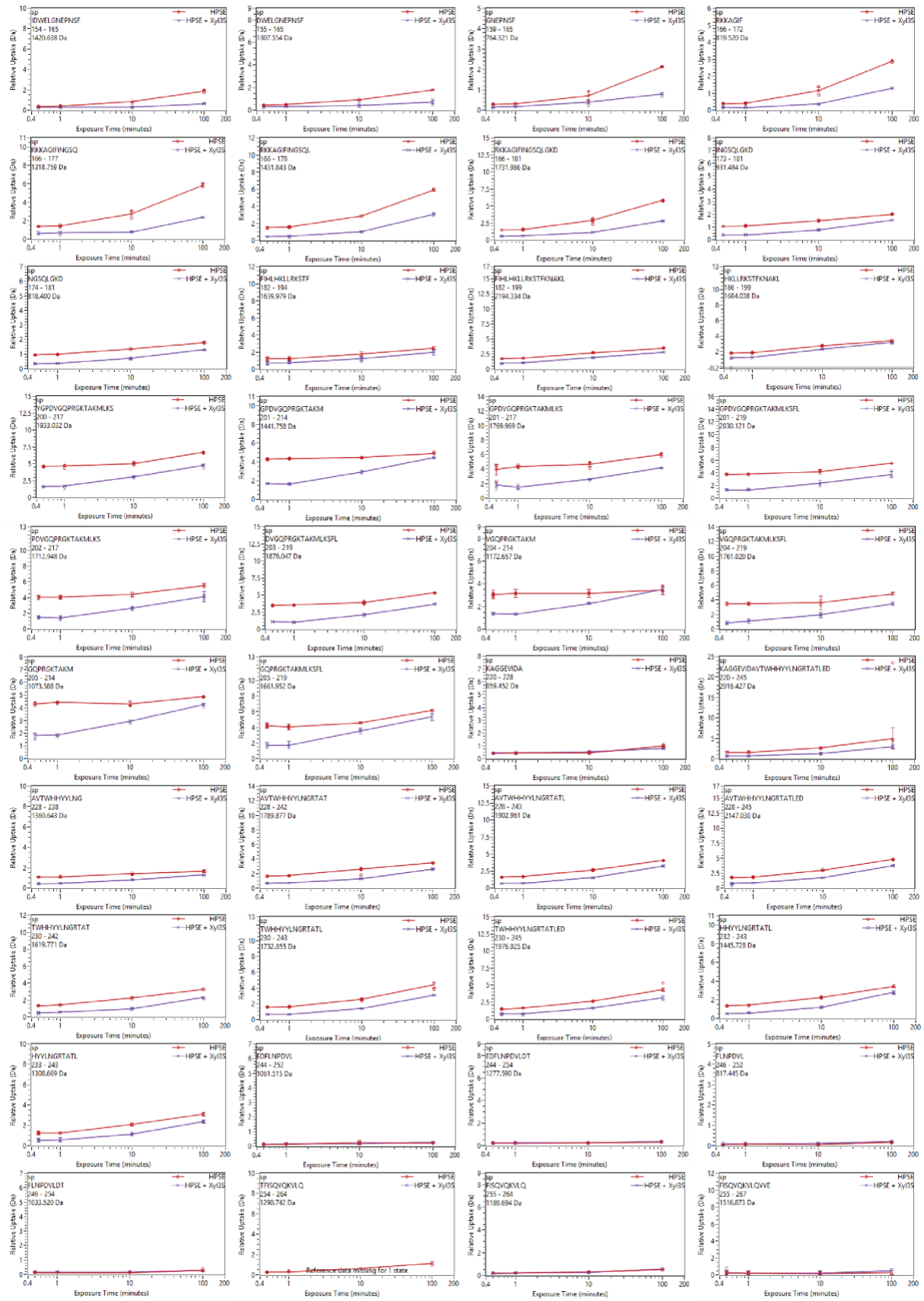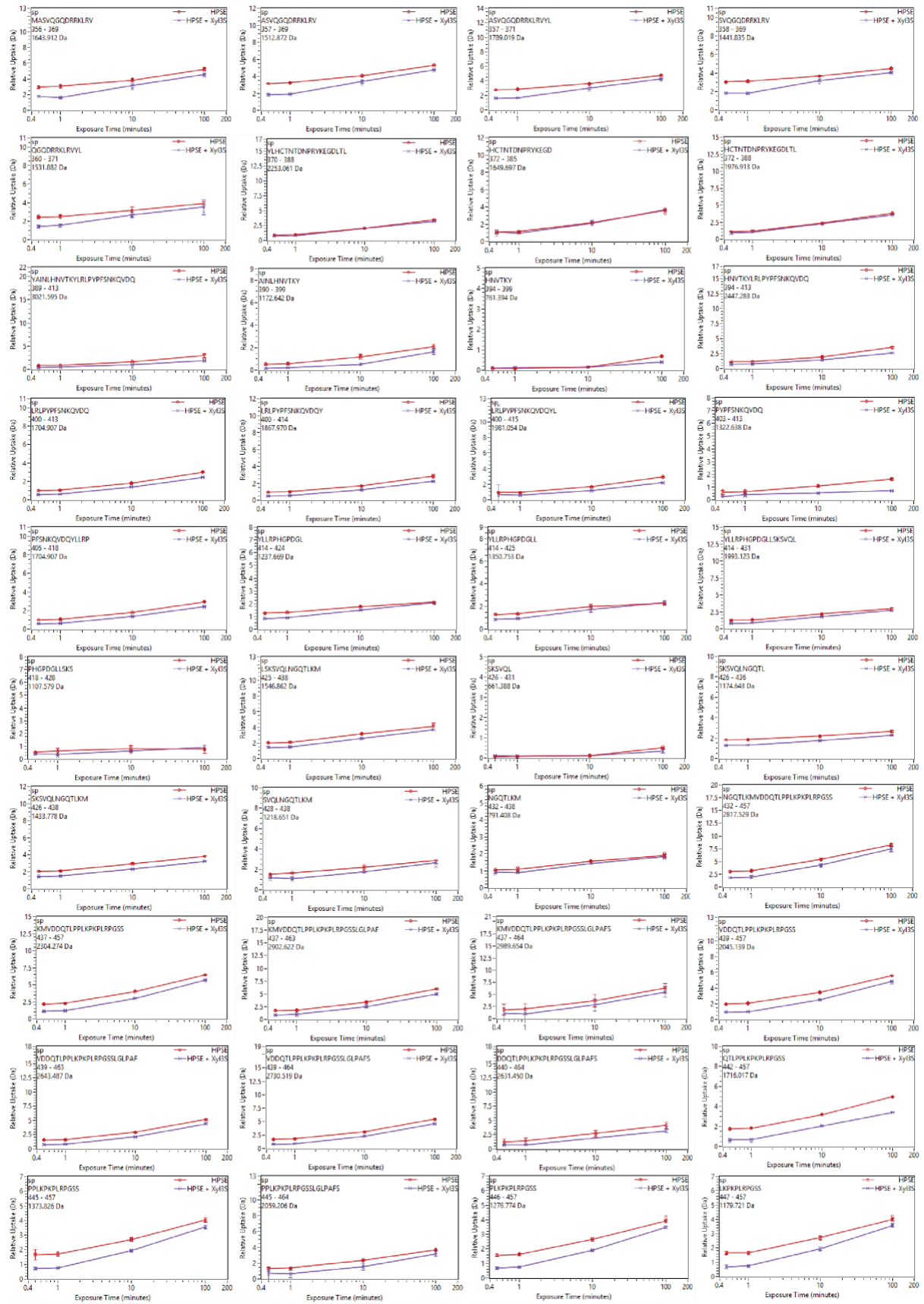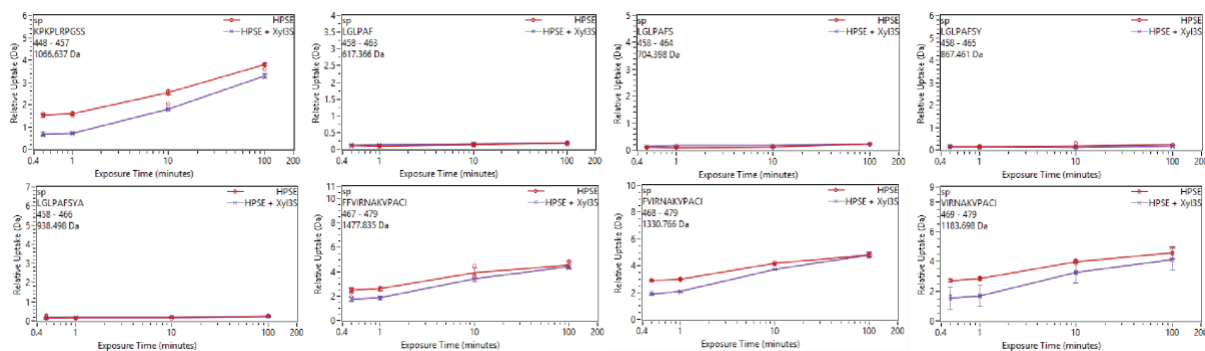
Figure S7. Hydrogen/deuterium exchange data of apo and Xyl₃S bound HPSE. Deuterium uptake plots of the indicated peptides. Apo and Xyl₃S bound HPSE are indicated by red and blue lines, respectively. Error bars represent the standard deviation from triplicate experiments.

General Synthesis Methods

Optical rotations were measured in the solvent specified using a Rudolph Research Analytical Autopol I Automatic Polarimeter (sodium D line, 589 nm). Unless otherwise noted, proton (^{1}H) and carbon (^{13}C) NMR spectra were recorded at 18 °C in CDCl$_3$ (treated with K$_2$CO$_3$ and then stored over 4 Å molecular sieves) or in D$_2$O (99.9 atom % D) on Bruker Ascend 400, 600, 700 MHz or Varian MR400 400 MHz spectrometers. Chemical shifts are recorded in parts per million (ppm) downfield from tetramethylsilane (δ TMS = 0). For ^{1}H NMR spectra, signals arising from the residual protio-forms of the solvent were used as the internal standards. Accordingly, residual CHCl$_3$ appearing at δ$_H$ 7.26, and the HDO resonance at δ$_H$ 4.79, were used to reference ^{1}H NMR spectra. For ^{13}C{^{1}H} NMR spectra, the central resonance of the CDCl$_3$ "triplet" appearing at δ$_C$ 77.16 and for samples recorded in D$_2$O, ca. 1 μL of CH$_3$CN was added and the "CH$_3$" signal arising at δ$_C$ 1.47 was used as a reference. When required for the purposes of differentiation, signals appearing in ^{13}C NMR spectra are reported to two decimal places. ^{1}H NMR data are recorded as follows: chemical shift δ (ppm) [multiplicity, coupling constant(s) *J* (Hz), relative integral] where multiplicity is defined as: s = singlet; d = doublet; t = triplet; q = quartet; m = multiplet; b = broad, app. = apparent, obsc. = obscured, or combinations of the above. Infrared spectra were recorded on a Perkin-Elmer UATR Two spectrometer as a thin film or solid. Low-resolution mass spectrometry (LRMS) and high-resolution mass spectrometry (HRMS) were performed using positive or negative mode electrospray ionisation (ESI$^{+/-}$) techniques on a Micromass ZMD ESI-Quadrupole, a Waters LCT Premier XE, a Thermo-Fischer Scientific Orbitrap Elite^{TM} Hybrid Ion Trap-Orbitrap or an Orbitrap QExactive mass spectrometer. A VG Fisons Auto-Spec mass spectrometer was used to obtain low- and high-resolution electron impact (EI) mass spectra. Analytical thin layer chromatography (TLC) was performed on aluminum-backed Merck silica gel 60 F$_{254}$ plates. Eluted plates were visualized using a 254 nm UV lamp, and/or charring the dried plate with a blow-torch or by treatment with a suitable dip followed by heating. These dips included phosphomolybdic acid : ceric sulfate : sulfuric acid (conc.) : water (37.5 g : 7.5 g : 37.5 g : 720 mL) or potassium permanganate : potassium carbonate : 5% sodium hydroxide aqueous solution : water (3 g : 20 g: 5 mL : 300 mL). Flash chromatographic separations were carried out following protocols defined by Still *et al.*[1] with Merck silica gel 60 (40-63 μm). THF,

Et$_2$O, and DCM were dried using a Glass Contour solvent purification system that is based upon a technology originally described by Grubbs *et al.*[2] Solvent compositions were mixed in the *v/v* ratios indicated. All solvents and reagents were used as supplied unless otherwise specified. Petroleum spirit in the boiling range of 40 – 60 °C was used. Reactions were conducted in oven-dried glassware open to the atmosphere unless otherwise specified. Desalting was performed with a BioGel P2 column (50 mm × 450 mm) and eluted with reverse osmosis purified water. The xylo-oligomeric mixture used (product number XOS-95P, food grade) was obtained from Shandong Longlive Bio-technology Co., Ltd., Qingdao, China. All sulfated compounds are sodiated but the sodium counter-ions have been omitted for clarity. Reduced pressure short-path vacuum distillation, when utilised, was conducted using a still-head connected to a receiving flask *via* a vacuum receiver adapter. Xylotetraose was purified on a Waters Alliance 2695 separation module coupled with an ACQUITY TQD detector via Waters flow splitter 15:1. MassLynx version 4.1 Software was used for data acquisition and processing.

Specific Transformations

For the preparation of Xyl$_3$S and Xyl$_8$S see Malins et al.[3]

Preparation of Xyl$_4$S
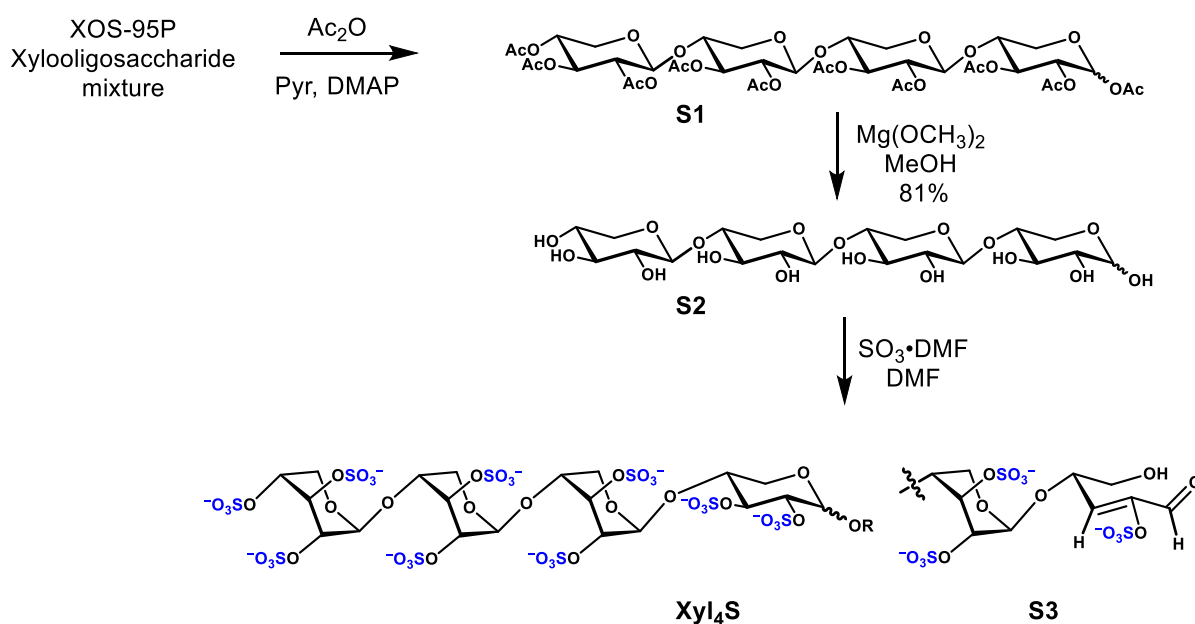
*Deca-O-acetyl-D-xylotetraose (**S1**)*
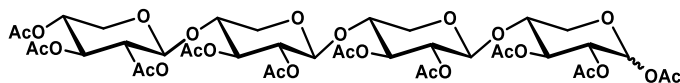
Following a procedure analogous to that reported by Malins[3] and Kosma,[4] a mixture of xylooligosaccharides (8.0 g, XOS-95P food-grade material) in anhydrous pyridine (40 mL), and maintained under an atmosphere of nitrogen, was magnetically stirred and after ca. 1 h a clear solution was obtained. Thereafter, the solution was cooled in an ice-water bath (0 – 5 °C) and treated sequentially with 4-dimethylaminopyridine (DMAP) (50 mg) and acetic anhydride (100 mL). Stirring was continued and the reaction was warmed to 18 °C over 24 h then cooled in an ice-water bath (0 – 5 °C) and treated cautiously with methanol (200 mL), portion-wise, over 0.5 h. After 1 h the mixture was co-evaporated with toluene (3 × 100 mL) and so affording a viscous pale-yellow syrup that was subjected to flash column chromatography (silica, 1:1 → 3:2 → 7:3 *v/v* EtOAc/pet. spirit gradient elution) and so affording three fractions, A, B and C.

Concentration of fraction A (**R**$_f$ = 0.4 in 4:3 *v/v* EtOAc/pet. spirit) afforded hexa-*O*-acetyl-D-xylobiose (4.90 g) as a colorless foam. See Malins[3] and Kosma[4] for further details and spectral data.

Concentration of fraction B (**R**$_f$ = 0.3 in 4:3 *v/v* EtOAc/pet. spirit) afforded octa-*O*-acetyl-D-xylotriose (3.80 g) as a colorless foam. See Malins[3] and Kosma[4] for further details and spectral data.

Concentration of fraction C (**R**$_f$ = 0.2 in 4:3 *v/v* EtOAc/pet. spirit) afforded deca-*O*-acetyl-D-xylotetraose (**S1**) (1.20 g) as a colorless foam.
[α]$_D$ = −81.0 (*c* = 1.0, CDCl$_3$)
^{1}H NMR (400 MHz, CDCl$_3$) *1.0β : 0.7α mixture of C-1 anomers* δ 6.18 (d, *J* = 3.7 Hz, 1H), 5.62 (d, *J* = 7.2 Hz, 1H), 5.35 (dd, *J* = 10.3, 8.8 Hz, 1H), 5.11 (app. t, *J* = 8.4 Hz, 1H), 5.09 – 5.00 (m, 6H), 4.95 – 4.91 (m, 2H), 4.88 – 4.83 (m, 2H), 4.79 – 4.76 (m, 2H), 4.74 – 4.70 (m, 4H), 4.54 – 4.52 (m, 2H), 4.47 – 4.44 (m, 4H), 4.09 – 4.05 (m, 2H), 3.97 (dd, *J* = 11.9, 5.1 Hz, 1H), 3.94 – 3.89 (m, 4H), 3.77 – 3.72 (m, 7H), 3.63 (app. t, *J* = 10.5 Hz, 1H), 3.44 (dd, *J* = 12.0, 9.1 Hz, 1H), 3.37 (dd, *J* = 12.0, 7.7 Hz, 2H), 3.33 – 3.27 (m, 4H), 2.15 – 1.99 (20 × CH$_3$, 60H).

13C NMR (101 MHz, CDCl₃) *1.0β : 0.7α mixture of C-1 anomers, many overlapping signals; all observed signals are reported here* δ 170.04, 169.99, 169.94, 169.92, 169.91, 169.83, 169.71, 169.66, 169.48, 169.44, 169.27, 169.25, 169.12, 100.78, 100.37, 100.34, 100.32, 99.56, 92.26, 89.31, 75.80, 74.78, 74.74, 74.72, 74.23, 72.04, 71.99, 71.93, 71.92, 70.99, 70.87, 70.51, 70.38, 70.05, 69.85, 69.52, 68.37, 63.47, 62.64, 62.55, 61.62, 61.43, 21.00, 20.89, 20.87, 20.86, 20.84, 20.76, 20.73, 20.70, 20.61.

LRMS (ESI+): *m/z* (rel. int.) 989 (100) [M+ Na]⁺.

HRMS (ESI+) found 989.2755, calcd for $C_{40}H_{54}NaO_{27}^+$ [M+ Na]⁺ 989.2745.

IR ν_max = 1748, 1368, 1216, 1043 cm⁻¹.

*Xylotetraose (**S2**)*
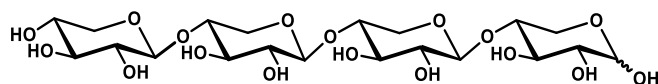
A magnetically stirred solution of deca-*O*-acetyl-D-xylotetraose (**S1**) (300 mg, 0.31 mmol) in anhydrous methanol (10 mL) was treated, at 18 °C, with a solution of magnesium methoxide in methanol (8.20 mL, 4.65 mmol, 6-10 wt. %). The resulting mixture was stirred for 48 h then quenched with Dowex 50WX4-100 (H⁺-form) to pH 2. The resin was filtered off and washed with methanol and the combined filtrates then concentrated under reduced pressure to afford xylotetraose (**S2**) (137 mg, 81%) as an amorphous white powder that was used without further purification. A portion of the ensuing product (ca. 20 mg) was subjected to purification by HPLC (see details below) for characterisation purposes which provided, after concentration of the appropriate fractions, xylotetraose, as an amorphous white powder.

[α]_D = −53.9 (*c* = 0.5, H₂O), Lit⁵ −59.1 (*c* = 1.0, H₂O)

1H NMR (400 MHz, D₂O) *1.0β : 0.55α mixture of C-1 anomers* δ 5.20 (d, *J* = 3.7 Hz, 1H), 4.60 (d, *J* = 7.9 Hz, 1H), 4.51 – 4.46 (m, 6H), 4.12 (dd, *J* = 11.8, 5.3 Hz, 4H), 4.07 (dd, *J* = 11.8, 5.4 Hz, 1H), 3.98 (dd, *J* = 11.6, 5.5 Hz, 2H), 3.85 – 3.74 (m, 8H), 3.67 – 3.60 (m, 3H), 3.60 – 3.53 (m, 6H), 3.47 – 3.34 (m, 8H), 3.34 – 3.23 (m, 8H).

13C NMR (101 MHz, D₂O) *1.0 β : 0.55 α mixture of C-1 anomers, many overlapping signals; all observed signals are reported here* δ 102.38, 102.20, 97.04, 92.56, 77.09, 76.93, 76.90, 76.87, 76.14, 74.53, 74.44, 74.20, 74.17, 73.31, 73.23, 71.92, 71.48, 69.72, 65.75, 63.51, 63.49, 59.36.

LRMS (ESI+): *m/z* (rel. int.) 569 (100) [M+ Na]$^+$.

HRMS (ESI+) found 569.1686 calcd for $C_{20}H_{34}NaO_{17}^+$ [M+ Na]$^+$ 569.1688.

HPLC purification of Xylotetraose: TR = 8.38 minutes (Column: Waters Carbohydrate Analysis, 3.9 x 300 mm (Part# WAT084038). Solvent A – Water + 0.1% (v/v) formic acid; Solvent B – Acetonitrile + 0.1% (v/v) formic acid. Solvent gradient: Solvent B 80% (0 min), 50% (from 0 to 7 min), 50% (from 7 to 10 min). Flow: 1 mL/min. Column temperature: 30 °C. Mass scanned from 200 to 800 (*m/z*); Ionisation mode ES-; Capillary voltage 3.60 kV; Cone voltage 45 V; Source temperature 150 °C; Desolvation Temperature 450 °C: SIR channel was also set at 545.4658 (*m/z*).

*Sulfation of xylotetraose with SO$_3$•DMF, Preparation of **Xyl$_4$S***

R = H or SO$_3^-$ **Xyl$_4$S**   **S3**

A magnetically stirred solution of SO$_3$•DMF (448 mg, 2.93 mmol) in DMF (1.5 mL) maintained under an atmosphere of nitrogen was cooled to –20 °C (using a CO$_2$(s)/EtOH/H$_2$O-based cryogenic system) then treated dropwise with a solution of xylotetraose (**S2**) (80 mg, 146 µmol) in DMF (2 mL). After 0.7 h the mixture was warmed to 0 °C (ice-bath) and stirring continued for 0.5 h before the reaction was cautiously quenched with aqueous sodium carbonate (10% w/v) until effervescence had ceased (pH ~7 – 8). The mixture was concentrated under reduced pressure by short-path distillation (1 mmHg, water bath at 40 °C) to remove volatiles and afford, as the residue, a white powder. The residue was desalted by passing it through a BioGel P2 column and fractions were analyzed by spotting on a TLC plate and charring. The appropriate fractions were collected and water removed by freeze-drying. The powder thus obtained was dissolved in water (ca. 2 mL) and subjected to ion-exchange with Dowex 50WX4-100 (Na$^+$-form) followed by freeze-drying to afford the desired product(s) (**Xyl$_4$S**) (77 mg) as a white amorphous powder. At the time of ^{13}C NMR data acquisition, **Xyl$_4$S** contained a mixture of R = H : R = SO$_3$: ΔXyl$_{red}$(CHO) (**S3**) of 11.0 : 7.7 : 2.3. The spectral data and decomposition/hydrolysis profiles are in accord with those reported for sulfated xylooligosaccharides incorporating an unprotected reducing terminus.[3,6,7]

¹H NMR (400 MHz, D_2O) *mixture of C-1 anomers* δ 9.38 (s, 1H, ΔXyl$_{red}$(CHO)), 6.76 (d, J = 8.8 Hz, 1H, ΔXyl$_{red}$(CHO)), 5.93 (d, J = 3.4 Hz, 1H, α-Xyl$_{red}$-OSO$_3$), 5.77 (app. s, 1H, β-Xyl$_{red}$-OSO$_3$), 5.36 (d, J = 2.9 Hz, 1H, α-Xyl$_{red}$), 5.24 (app. s, 1H, β-Xyl$_{red}$), 5.18 – 5.13 (m), 5.10 – 5.08 (m), 4.84 – 4.82 (m), 4.79 – 4.78 (m), 4.77 – 4.70 (m), 4.57 – 4.54 (m), 4.54 – 4.42 (m), 4.42 – 4.38 (m), 4.35 (dd, J = 7.6, 2.9 Hz), 4.00 – 3.77 (m).

¹³C NMR (101 MHz, D_2O) *many overlapping signals; all observed signals are reported here* δ 190.91, 147.52, 138.71, 99.74, 98.93, 98.31, 98.12, 96.07, 92.98, 91.26, 88.70, 76.23, 75.96, 75.52, 75.22, 74.94, 74.81, 74.73, 74.66, 74.43, 73.94, 73.31, 73.13, 73.07, 72.98, 72.84, 72.77, 72.39, 71.87, 71.63, 61.54, 61.36, 59.83, 59.73, 59.45, 58.92.
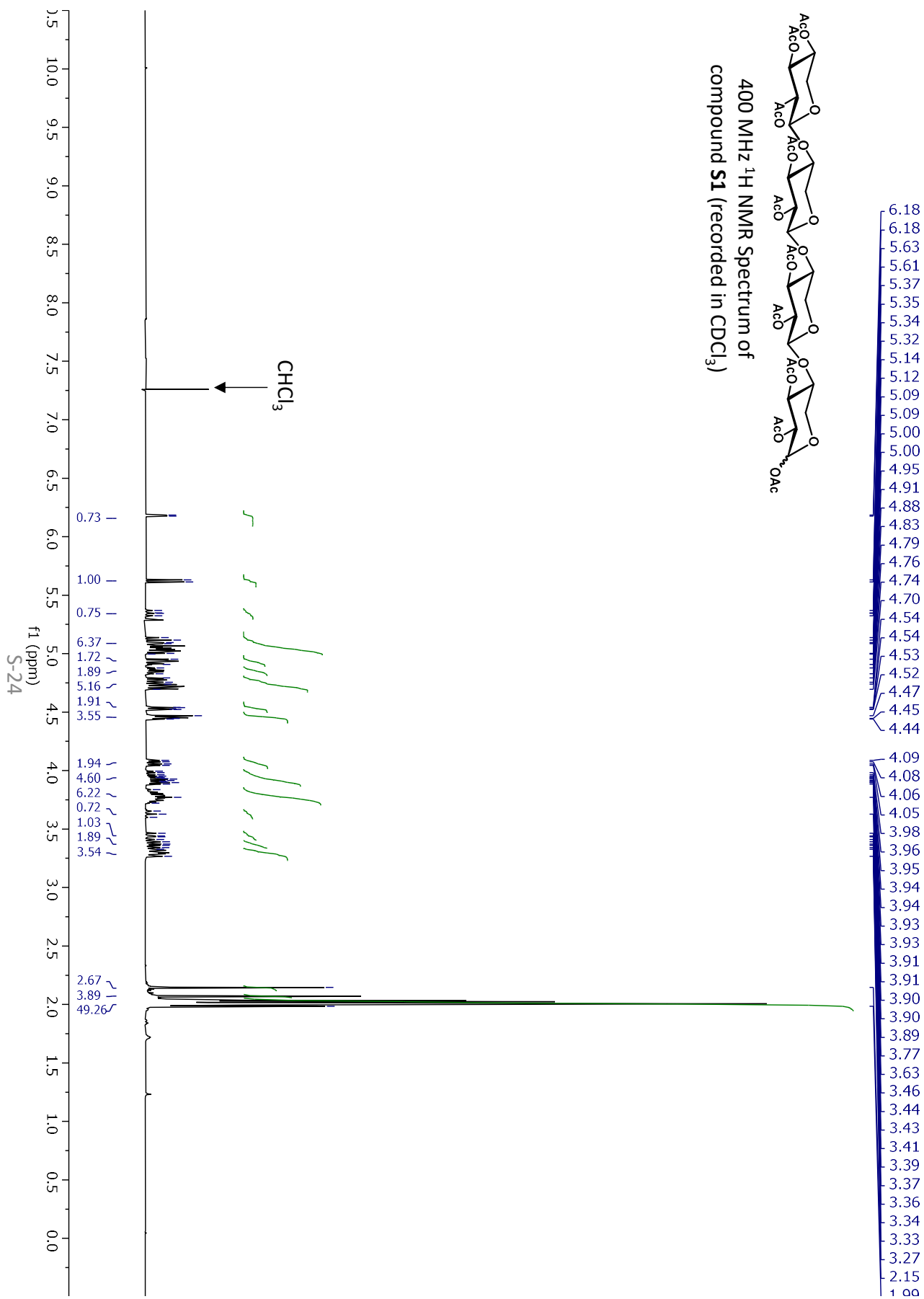
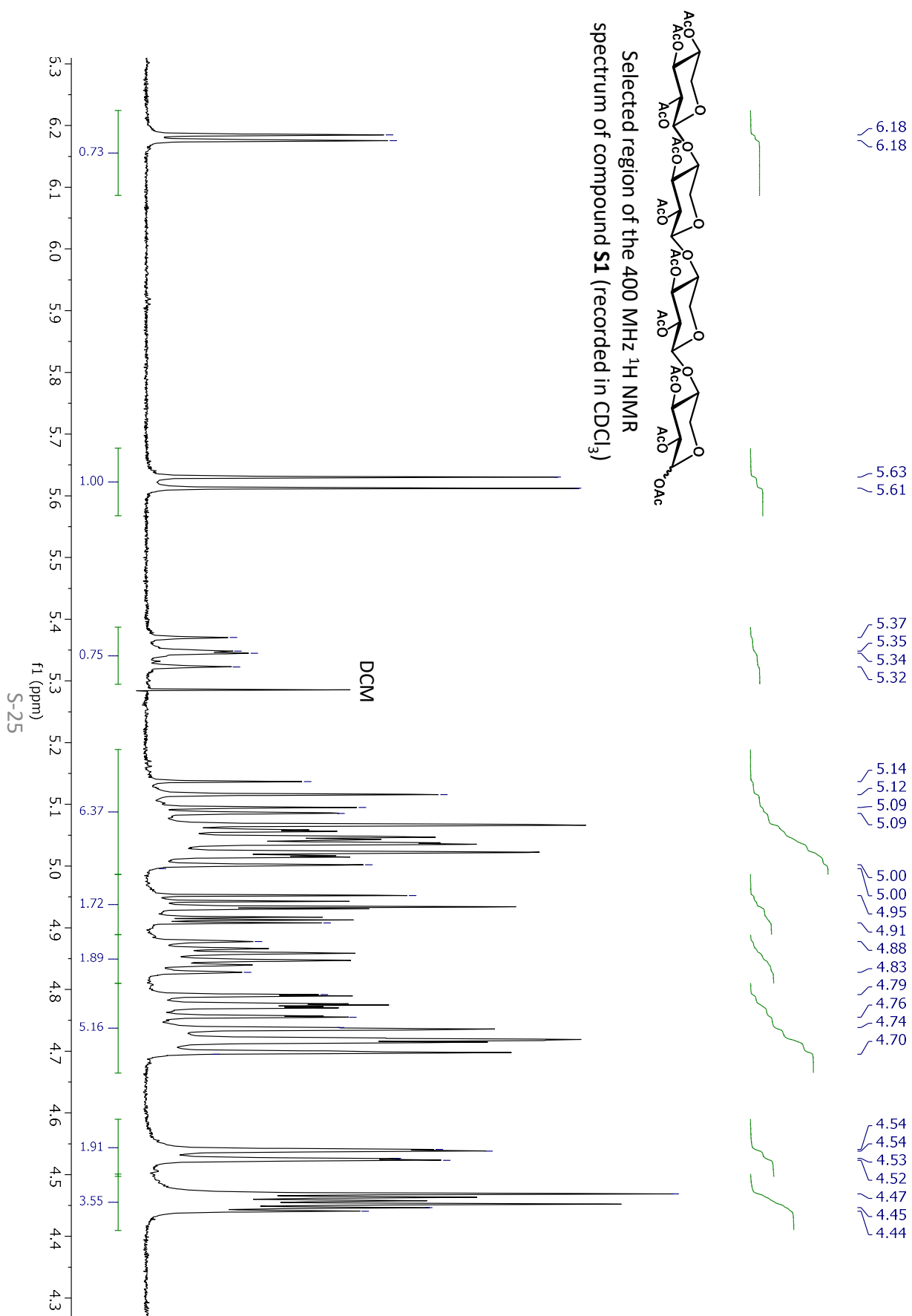**LRMS** (ESI−): *m/z* 465 [M − 3Na]$^{3-}$ (R=H), [M − 3Na]$^{3-}$ (R=SO$_3$)

HRMS (ESI+) found 1486.6209, calcd for $C_{20}H_{25}Na_{10}O_{44}S_9^+$ (R=H) [M+ Na]$^+$ 1486.6177.

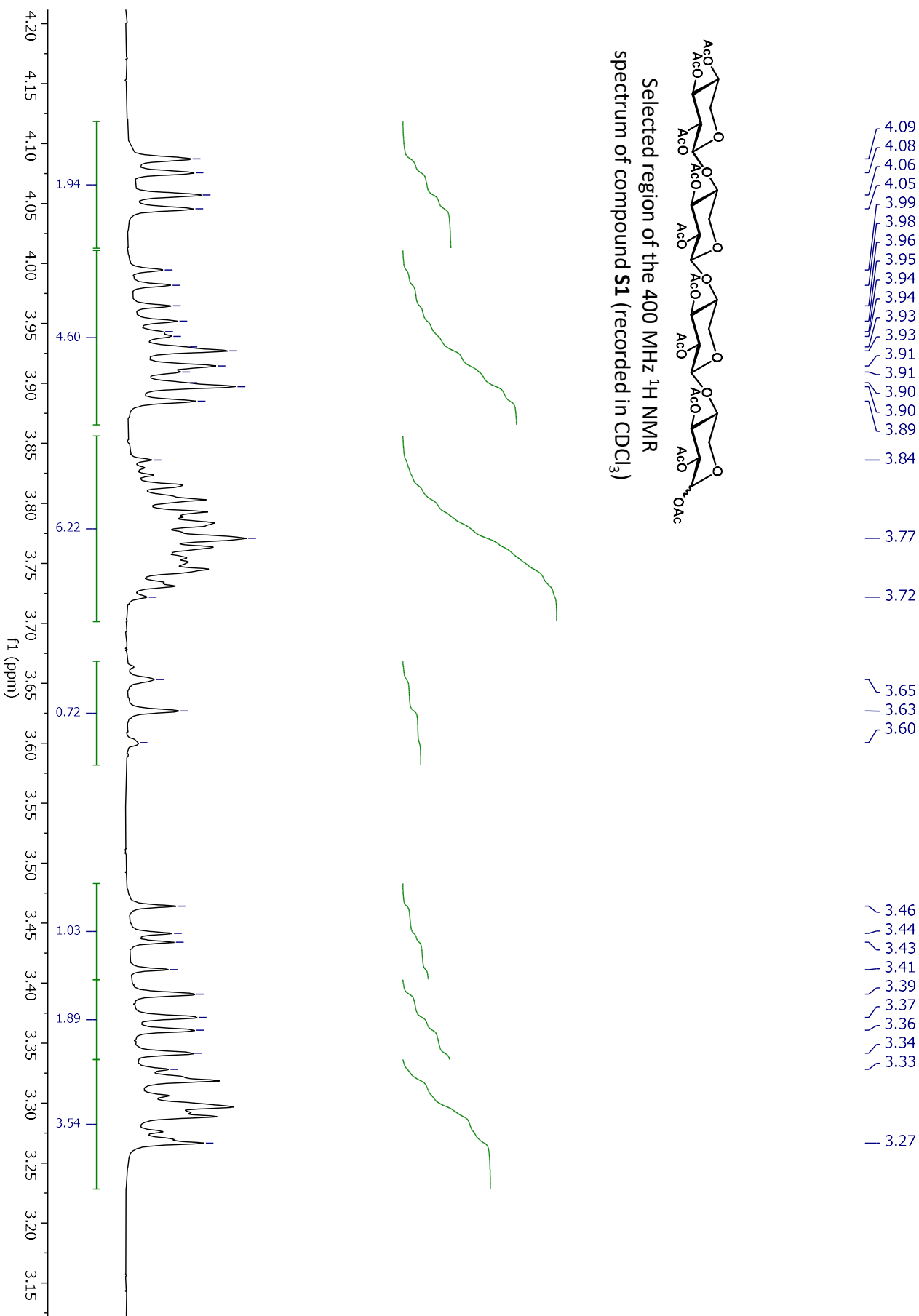**HRMS** (ESI+) found 1588.5597, calcd for $C_{20}H_{24}Na_{11}O_{47}S_{10}^+$ (R=SO$_3$) [M+ Na]$^+$ 1588.5564.
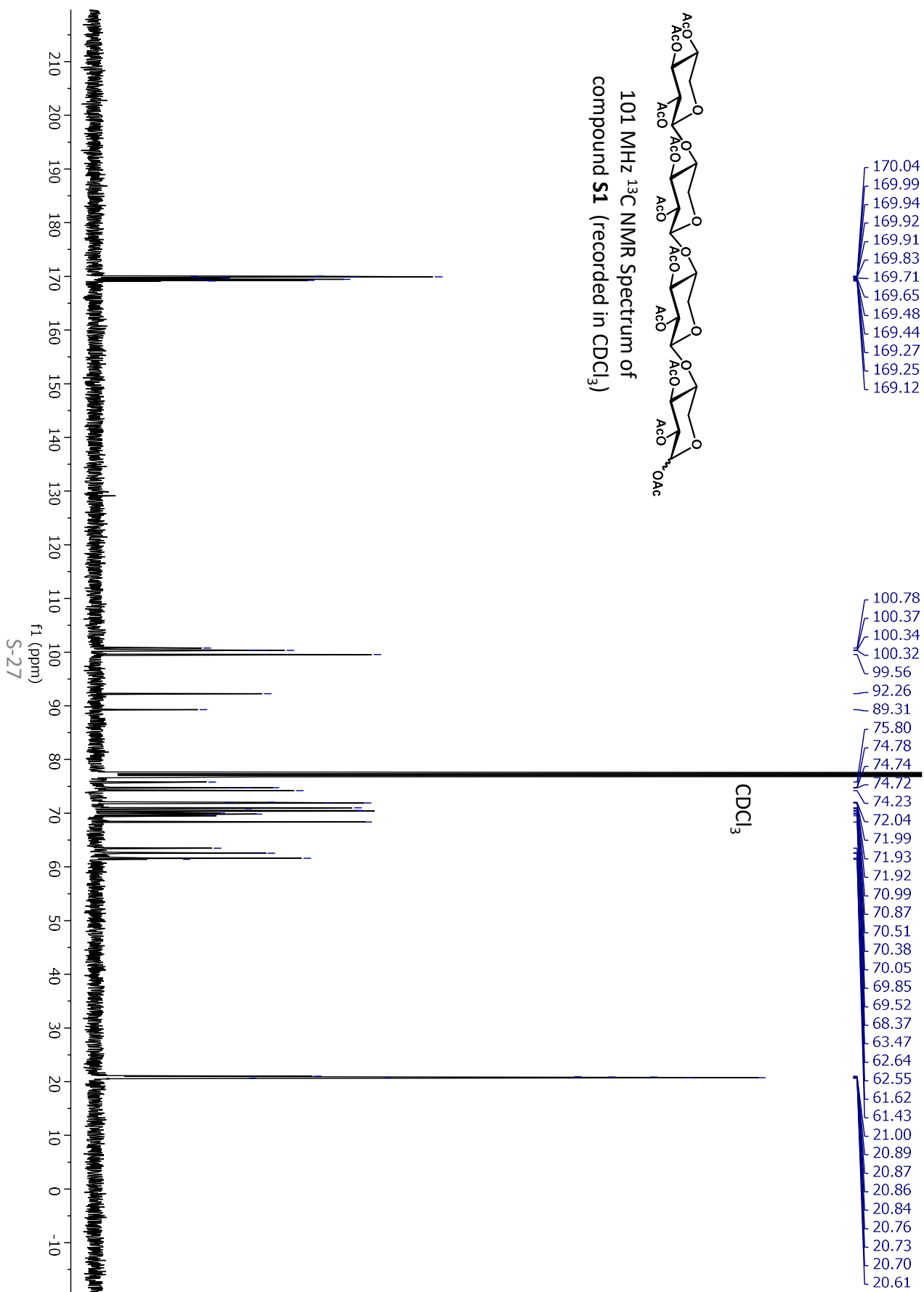
References

(1) Still, W. C.; Kahn, M.; Mitra, A. Rapid Chromatographic Technique for Preparative Separations with Moderate Resolution. *J. Org. Chem.* **1978**, *43* (14), 2923–2925. https://doi.org/10.1021/jo00408a041.

(2) Pangborn, A. B.; Giardello, M. A.; Grubbs, R. H.; Rosen, R. K.; Timmers, F. J. Safe and Convenient Procedure for Solvent Purification. *Organometallics* **1996**, *15* (5), 1518–1520. https://doi.org/10.1021/om9503712.

(3) Vo, Y.; Schwartz, B. D.; Onagi, H.; Ward, J. S.; Gardiner, M. G.; Banwell, M. G.; Nelms, K.; Malins, L. R. A Rapid and Mild Sulfation Strategy Reveals Conformational Preferences in Therapeutically Relevant Sulfated Xylooligosaccharides. *Chem. Eur. J.* **2021**, *27* (38), 9830–9838. https://doi.org/10.1002/chem.202100527.

(4) Abad-Romero, B.; Mereiter, K.; Sixta, H.; Hofinger, A.; Kosma, P. Synthesis of Regioselectively Sulfated Xylodextrins and Crystal Structure of Sodium Methyl β-d-Xylopyranoside 4-O-Sulfate Hemihydrate. *Carbohydr. Res.* **2009**, *344* (1), 21–28. https://doi.org/10.1016/j.carres.2008.09.018.

(5) Takeo, K.; Ohguchi, Y.; Hasegawa, R.; Kitamura, S. Synthesis of (1 → 4)-β-d-Xylo-Oligosaccharides of Dp 4–10 by a Blockwise Approach. *Carbohydr. Res.* **1995**, *278* (2), 301–313. https://doi.org/10.1016/0008-6215(95)00259-6.

(6) Itzstein, M. V.; Chang, C.-W. Sulfation Method. WO2019113646A1, June 20, 2019.

(7) Alekseeva, A.; Raman, R.; Eisele, G.; Clark, T.; Fisher, A.; Lee, S. (Larry); Jiang, X.; Torri, G.; Sasisekharan, R.; Bertini, S. In-Depth Structural Characterization of Pentosan Polysulfate Sodium Complex Drug Using Orthogonal Analytical Tools. *Carbohydr. Polym.* **2020**, *234*, 115913. https://doi.org/10.1016/j.carbpol.2020.115913.

400 MHz ^{1}H NMR Spectrum of
compound **S1** (recorded in CDCl$_3$)

CHCl$_3$

S-24

Selected region of the 400 MHz ^{1}H NMR spectrum of compound **S1** (recorded in CDCl$_3$)

Selected region of the 400 MHz ^{1}H NMR spectrum of compound **S1** (recorded in CDCl$_3$)

101 MHz ^{13}C NMR Spectrum of compound **S1** (recorded in CDCl$_3$)

170.04
169.99
169.94
169.92
169.91
169.83
169.71
169.65
169.48
169.44
169.27
169.25
169.12

100.78
100.37
100.34
100.32
99.56
92.26
89.31
75.80
74.78
74.74
74.72
74.23
72.04
71.99
71.93
71.92
70.99
70.87
70.51
70.38
70.05
69.85
69.52
68.37
63.47
62.64
62.55
61.62
61.43
21.00
20.89
20.87
20.86
20.84
20.76
20.73
20.70
20.61

CDCl$_3$

f1 (ppm)

S-27

Selected region of the 101 MHz ^{13}C NMR spectrum of compound **S1** (recorded in CDCl$_3$)

400 MHz ^{1}H NMR Spectrum of compound **S2** (recorded in D_2O)

HDO

CH_3CN (calibrant)

5.20
5.19
4.61
4.50
4.50
4.48
4.48
4.46
4.14
4.13
4.11
4.10
4.08
4.06
4.05
4.00
3.99
3.97
3.96
3.85
3.73
3.67
3.54
3.46
3.35
3.32
3.24

1.00
2.29
9.20
8.57
3.56
13.14
12.86
24.56

f1 (ppm)

Selected region of the 400 MHz ¹H NMR spectrum of compound **S2** (recorded in D_2O)

101 MHz ^{13}C NMR Spectrum of
compound **S2** (recorded in D$_2$O)

CH$_3$CN (calibrant)

CH$_3$CN (calibrant)

S-31

f1 (ppm)

102.38
102.20
97.04
92.56
77.09
76.93
76.90
76.87
76.14
74.53
74.44
74.20
74.17
73.31
73.23
71.92
71.48
69.72
65.75
63.51
63.49
59.36

Selected region of the HSQC-edited 2D experiment of compound **S2** (recorded in D$_2$O)

400 MHz ^{1}H NMR Spectrum of compound **Xyl$_4$S** (recorded in D$_2$O)

Xyl$_4$S

R = H or SO$_3^-$

Elimination product **S3**

HDO

CH$_3$CN (calibrant)

H$_b$

H$_a$

H$_a$

H$_b$

9.38
6.77
6.75
5.94
5.93
5.77
5.40
5.36
5.36
5.28
5.24
5.16
5.10
5.08
4.84
4.83
4.82
4.79
4.79
4.78
4.77
4.70
4.57
4.54
4.54
4.42
4.38
4.36
4.35
4.34
4.33
4.00
3.77

2.26
6.86
0.82
10.00
0.95

f1 (ppm)

Selected region of the 400 MHz ^{1}H NMR spectrum of compound(s)

Xyl$_4$S (recorded in D$_2$O)

101 MHz ^{13}C NMR Spectrum of compound **Xyl$_4$S** (recorded in D$_2$O)

R = H or SO$_3^-$ **Xyl$_4$S**

CH$_3$CN (calibrant)

CH$_3$CN (calibrant)

Selected region of the 101 MHz ^{13}C NMR spectrum of compound **Xyl$_4$S** (recorded in D$_2$O)

Xyl$_4$S   R = H or SO$_3^-$

Chapter 4

Allosteric small molecule inhibitors of human heparanase

4.1 Declaration

All experimental work except where otherwise stated, was carried out by the author under the supervision of Professor Colin Jackson The contribution of other authors are as follows: Junming He, and Dr. Brett D. Schwartz from Martin Banwell's and Lara Malin's lab groups synthesised the small molecule inhibitors. Dr. Ben Clifton worked on non-specific binding and non-specific aggregation assays. Dr. Nansook Hong alongside the author cloned HPSE P6 mutants. Dr. F. Hafna Ahmed and Dr. Elaff Mohamed assayed some small molecule inhibitors not presented in this current manuscript. Caryn Hepburn from Waters ran and processed hydrogen/deuterium exchange mass spectrometry data. All authors contributed to conception of experiments. The Chapter was written by the author with support Professor Colin Jackson. This manuscript is drafted for submission, and missing data associated with the small molecule structure-activity relationship.

Allosteric small molecule inhibitors of human heparanase.

Cassidy Whitefield[ab], Junming He[ab], Brett D Schwartz[ab], Ben Clifton[a], Nansook Hong[a], Hafna Ahmed[a], Elaaf Mohamed[a], Keats Nelms[d], Martin Banwell[c], Lara R Malins[ab], Colin J Jackson.[ab]*

[a] Research School of Chemistry, Australian National University, Canberra, ACT, 2601, Australia
[b] Australian Research Council Centre of Excellence for Innovations in Peptide and Protein Science, Australian National University, Canberra, ACT 2601, Australia
[c] Institute for Advanced and Applied Chemical Synthesis, College of Pharmacy, Jinan University, Guangzhou, Guangdong, 510632 China
[d] Beta Therapeutics Pty. Ltd. Level 6, 121 Marcus Clarke Street, Canberra, ACT 2601 Australia
* E-mail: colin.jackson@anu.edu.au

Abstract

Heparanase (HPSE) is an endo-β-glucuronidase that catalyses the degradation of heparan sulfate and is the only known enzyme with this activity in mammals. This process is involved in many essential biological processes as well as the development of various diseases including cancer, inflammatory disorders and viral infections, such as Covid-19. Despite numerous efforts to develop therapies targeting HPSE, no drugs targeting this enzyme have yet passed clinical trials. Small molecule inhibitors of HPSE have been in development for some time, although considerably more work is needed to understand their mechanism of inhibition. In this study, we have conducted a comprehensive structure-activity relationship analysis of quinazoline analogues as inhibitors of HPSE and their mode of inhibition, showing that these molecules behave as allosteric inhibitors. We have also tested a total of 71 mutations of HPSE to better understand the interaction between the small molecule and HPSE, as well as how remote structural changes can affect activity. These findings may inform the development of new therapeutic approaches targeting HPSE and increase our understanding of the mechanisms of HPSE inhibition and activity.

Introduction

Human heparanase (HPSE) is an endo-β-glucuronidase that cleaves the glycosidic bonds of heparan sulfate (HS) side chains from heparan sulfate proteoglycans (HSPG). HSPGs are a major component of the extracellular matrix, cell surface, and basement membranes.[1] HSPGs contribute to the structural integrity of tissue, while also acting as a reservoir for biological signalling molecules, such as growth factors, cytokines, and chemokines, which control cell growth, angiogenesis, coagulation, and inflammation.[1,2] HPSE-mediated modification of HSPGs involves the release of regulatory molecules, while released HS chains can regulate protein-protein interactions and structural modification of the extracellular matrix can allow immune cell penetration and tissue remodelling.[1,3–6]

Dysregulation of HPSE can result in altered cell motility through the weakening of structural HSPG networks within the ECM and basal membranes,[7] facilitating angiogenesis,[8] inflammation[9] and invasion into surrounding tissue.[10,11] These effects have resulted in HPSE being associated with a number of diseases, including diabetes,[12] liver fibrosis,[13] cancers,[14] age-related macular degeneration,[15] and diabetic retinopathy,[16] while also supporting the invasion of viruses, such as Covid-19.[17,18] Due to the crucial role of HPSE in these diseases and many others, HPSE has long been an important drug target for a wide range of diseases.

HPSE inhibitors in development include HS analogues[19–22] synthetically produced small molecule compounds,[23–25] nucleic acid-based drugs,[26] vaccines,[27] monoclonal antibodies,[28] proteins,[29] covalent drugs,[30] and natural products.[31] However, to date, there are only four HPSE drugs that have reached clinical trials, all being polysaccharide mimetics: Muparfostat (PI-88),[21] pixatimod (PG545),[22] roneparstat (SST0001),[19] and necuparanib (M-402).[20] These polysaccharides have been shown to have anti-cancerous and anti-metastatic properties in animal models and early clinical trials.[32,33] However, these oligosaccharides have potential safety concerns as well as limited effects *in vivo,* which have limited their use as effective therapeutics.[34,35] The heterogeneity of these oligosaccharides also has limited the standardisation of their active components, and their size requires parenteral administration.

Small molecule inhibitors, which can be optimised for oral administration and have favourable pharmacokinetics, are a desirable alternative to polysaccharide mimetics for the treatment of HPSE. However, few small molecule inhibitors for HPSE have been developed and even fewer have progressed to clinical trials due to limited *in vitro* and *in vivo* studies.[23,36,37] This may be

due to the large exposed binding cleft in the active site of HPSE, which poses challenges for specific binding and drug development. Recently, polyanionic saccharide-based covalent inhibitors have garnered interest to bridge the gap between heparan sulfate mimetics and standard small molecule inhibitors[30] Understanding how the traditional reversible small molecules bind is important for future drug discovery, as most lack kinetic and structural information.

Because HPSE has such an important biological role, with little biological redundancy, and very high local substrate concentration, allosteric inhibitors are of particular interest. Allosteric inhibitors allow the modulation of protein activity by the binding of a ligand at an allosteric site that is topographically distinct from the orthosteric functional site. This may affect conformational changes of the protein, or its active site, to initiate a change in activity.[38,39] Thus, allosteric inhibitors for HPSE may be able to overcome many limitations of competitive inhibitors and allow modulation of the activity in a non-competitive fashion. Allosteric drugs may also reduce non-specific side effects as they typically bind to more specific sites[38] (compared to current HS mimetics) and interact with multiple glycan-binding proteins.[34] However, identifying allosteric sites on HPSE is challenging as there is limited research into the location of potential allosteric sites, and how they could be targeted.

In this work, we describe the identification of quinazoline-based compounds as HPSE inhibitors. Through an initial high-throughput screen, the compound BT2012 was identified with an IC$_{50}$ of 24.9 µM towards HPSE, which led to the testing of the structurally related drug doxazosin, which is used to treat hypertension. Using BT2012 and doxazosin as lead compounds, a series of analogues were synthesised to explore the chemical space and structure-activity relationship (SAR) studies, which identified a compound, BT2162, with reasonable pharmacokinetic properties and an IC$_{50}$ value in the order of ~15 µM. Interestingly, analysis of the mechanism of inhibition of this molecule revealed it to be a non-competitive, allosteric inhibitor; a novel mechanism of action for HPSE inhibitors. Further mutagenesis work confirmed that structural changes remote from the active site can have significant effects on activity, confirming that allosteric inhibition is possible. These results should provide a foundation for future work on allosteric HPSE inhibitors and a deeper understanding of the structure, function and inhibition of this important enzyme.

Results

Structure Activity Relationships (SAR) of Quinazoline Compounds and HPSE

We initially undertook an automated high-throughput chemical screen (HTCS) of over 300,000 potential HPSE inhibitor using a miniaturized HPSE assay that is described in two patents (WO2018107201 and WO2018107226A1) that disclose the structural diversity and synthesis of small molecule inhibitors of HPSE. In this study, we explored the SAR of these small molecule inhibitors. The compounds BT2012 and BT2071 were identified from the initial screen and form the lead compounds of the first series of molecules (**Figure 1**). BT2012 exhibited inhibition of HPSE (IC$_{50}$ 16.4 µM), and its analogue BT2071 was synthesised soon thereafter (IC$_{50}$ 7.0 µM). Following the development of Series I, an alternative scaffold was adopted based on the clinical therapeutic doxazosin mesylate, a selective α$_1$-adrenergic receptor blocker, which exhibited an IC$_{50}$ of 57.4 µM, but had much better pharmacokinetic properties (**SI Table 1**). This led to the development of BT2162 as part of Series II, with an IC$_{50}$ of 15.2 µM (**Figure 1**).
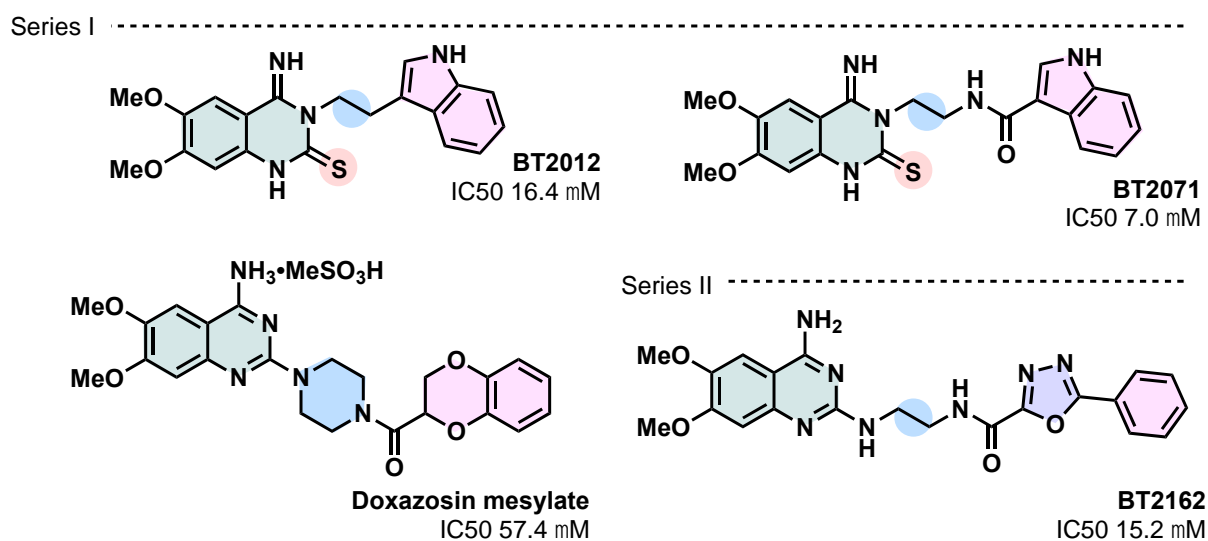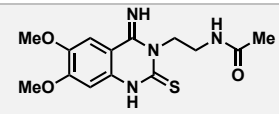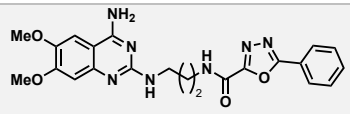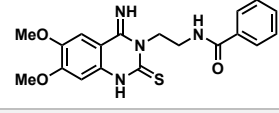

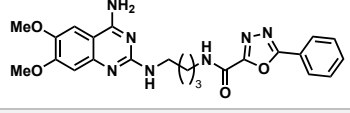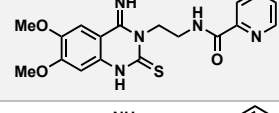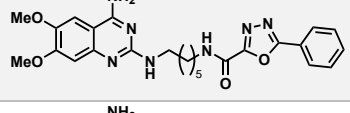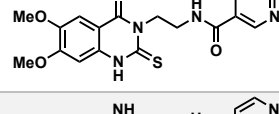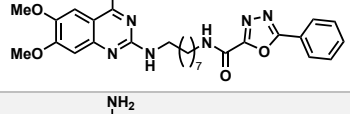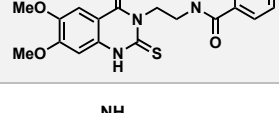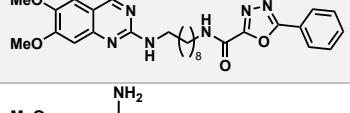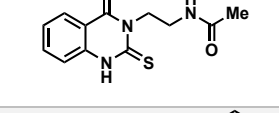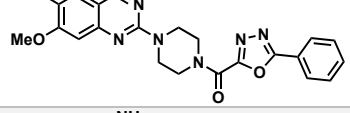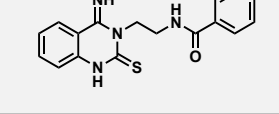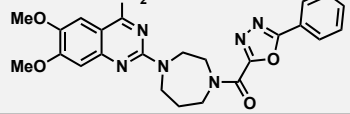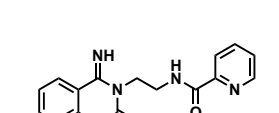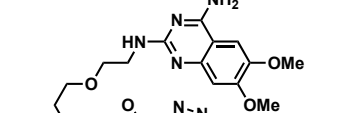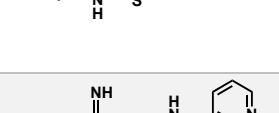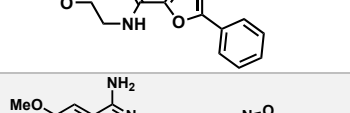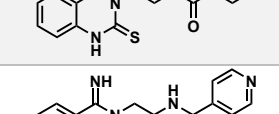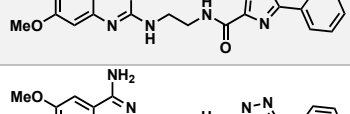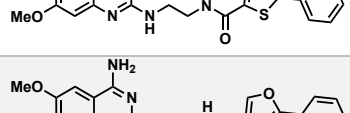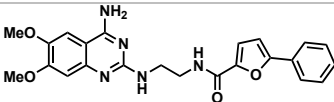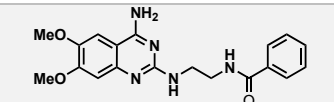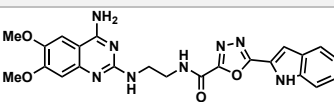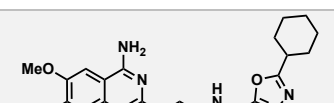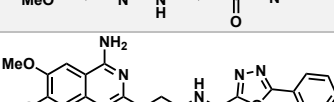**Figure 1** Structures of Series I and II lead compounds, highlighting the similar scaffolds between the compounds.

The IC$_{50}$ was determined for all molecules in Series I and Series II (**Table 1**). Several compounds exhibit unusually steep Hill coefficients, (**SI Figure 1**) which may suggest a complex binding mechanism since Hill coefficients of 1 signify a 1:1 binding mechanism, while values higher than 1:1 could potentially indicate non-specific aggregation, tight binding, or high stoichiometry.[40] We selected BT2162 as a lead compound in order to understand the binding mechanism of this Series because it displayed an IC$_{50}$ value of 15.2 ± 2.3 µM, which is similar to other small molecule inhibitors of HPSE,[23,37] and because it possessed the best

pharmacokinetic properties of the series (**SI Table 1**). Although this IC_{50} value is considerably higher than those typically reported for glycomimetic compounds,[41] our previous research using the HS-mimetic Pentosan (described in **Chapter 3**) suggests that the binding affinity of these sugars may not necessarily be in the nM range and be convoluted by multiple factors *in vitro*, including aggregation.

Table 1. Summary of small molecules designed for the SAR around BT2162. IC_{50} values provided with error given as standard deviation.

Structure (Series I)	Name	IC_{50}	Structure (Series II)	Name	IC_{50}
	JHBT003	>150		JHBT017	122 ± 27
	JHBT001	>150		JHBT019	55.9 ± 7
	JHBT005	>150		JHBT021	16.9 ± 3.8
	JHBT007	>150		JHBT018	13.4 ± 7.2
	JHBT009	>150		JHBT022	2.7 ± 0.6
	JHBT004	>150		JHBT023	11.5 ± 0.1
	JHBT002	>150		JHBT024	>150
	JHBT006	>150		JHBT025	>150
	JHBT008	>150		BT2236	46 ± 9
	JHBT010	>150		BT2229	20.4 ± 3.7
				BT2228	104 ± 9

				Name	IC50
				BT2238	>150
				BT2239	>150
				BT2169	4
				JHBT026	>150
				BT2162	15.2 ± 2.6

Binding mode analysis

In order to gain a better understanding of the inhibitory mechanism of BT2162 on HPSE, we conducted Michaelis-Menten kinetics studies. Double reciprocal (Lineweaver-Burk) enzyme kinetic analysis at various substrate and inhibitor concentrations showed that BT2162 exhibits non-competitive inhibition (**Figure 1A**). This non-competitive inhibition mechanism for HPSE inhibitors has only been observed once before, with the large synthetic inhibitor Suramin (IC50 of 46 µM).[42] This mode of inhibition is consistent with allosteric inhibition, potentially occurring at a site separate from the active site of HPSE. The location of the allosteric site was not identified in the case of Suramin.[42] Kinetic analysis of small molecule inhibitors of HPSE is infrequently performed due to the complexity of the assay. Thus, without similar data for other superficially similar small molecules, it is not possible to conclude whether the allosteric inhibition caused by BT2162 is generalisable to other small molecule inhibitors of HPSE.[23–25,37]

Since non-competitive inhibition is uncommon for HPSE, we sought to confirm that BT2162 is not a non-specific aggregator, which can confound the search for new drugs by forming aggregates in solution and non-specifically binding and denaturing proteins, thereby appearing to inhibit activity *in vitro*.[43] To test this, we performed two additional assays. The first was an inhibition assay with different concentrations of detergent in the system (**Figure 1C**). These reagents disrupt aggregated structures, and a change in IC50 would be expected with increasing detergent concentrations. However, we observed no significant change in the IC50 of BT2162

across four different detergent concentrations, indicating that BT2162 is unlikely to be a non-specific aggregator. The second assay involved testing the effect of enzyme concentration and aggregation on the activity of BT2162. An aggregator would be expected to show a change in IC_{50} with increasing enzyme concentration or upon removal of aggregation via centrifugation (**Figure 1D**). However, neither increasing the enzyme concentration nor centrifuging BT2162 resulted in a change in the IC_{50} of BT2162, providing further evidence that BT2162 is not an aggregator. Finally, we tested the compound with other, unrelated enzymes to see if BT2162 binds specifically to HPSE. S2 is a mutated bacterial phosphotriesterase from *Agrobacterium radiobacter*[44] and αE7 is an α-carboxylesterase from *Lucilia cuprina*.[45] Activity for these enzymes were tested against 4-nitrophenyl acetate hydrolysis with and without BT2162, and it was seen that BT2162 does not interfere with their catalytic activity, supporting the conclusion that BT2162 is a specific inhibitor of HPSE.

We also used isothermal titration calorimetry (ITC) to directly measure the thermodynamics of BT2162 binding to HPSE, identify the binding stoichiometry and compare the IC_{50} to K_D. Optimising the solubility of BT2162, which increases with DMSO concentration, and HPSE, which decreases with DMSO concentration, required several trials. Eventually, 2.5 % DMSO was used, which required "reverse" titration (protein into ligand) to be used for analysis because sufficiently high BT2612 concentrations could not be achieved for "forward" titration. The results demonstrated that the binding affinity of BT2162 to HPSE is 10.2 ± 1.87 µM, with a stoichiometry of ~1 (**Figure 2D**). This firstly confirms that the small molecule is binding directly to the protein and that there is a single binding site. Secondly, the observation that the K_D value is similar to the IC_{50} of 15.2 ± 2.6 suggests that the inhibitor is able to bind and inhibit HPSE without being affected by the competitive binding of the substrate.
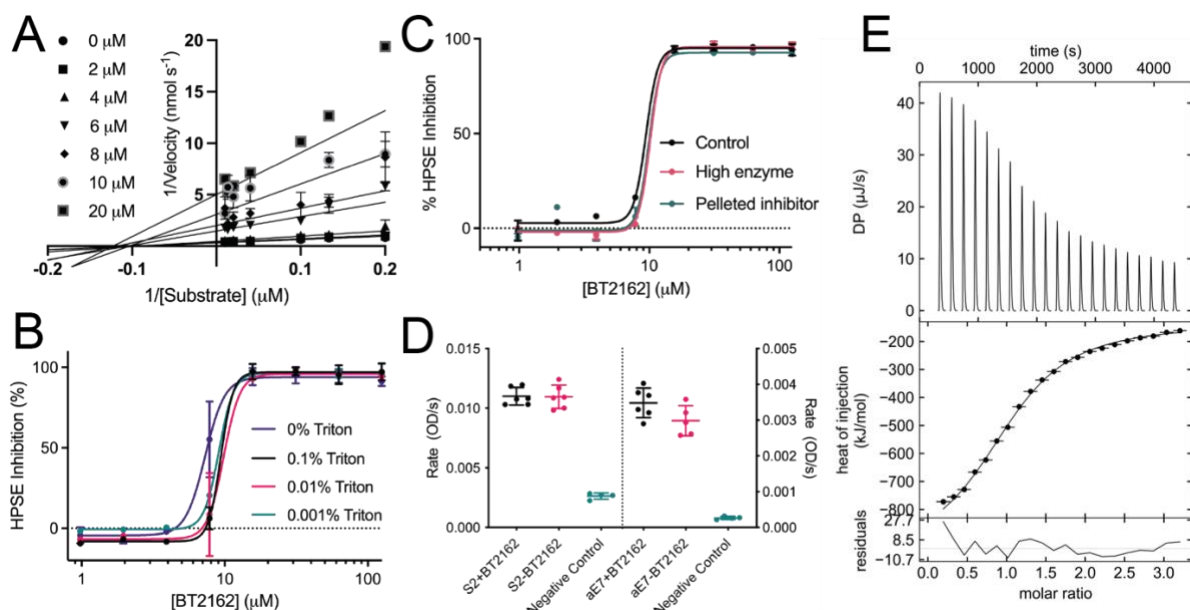
Figure 2: Binding data of lead compound, BT2162 confirming non-competitive inhibition. **A)** dose-response curve of BT2162. **B)** Double reciprocal plot of a range of BT2162 concentrations supporting a non-competitive binding mode. **D)** BT2162 inhibition of HPSE with an increasing range of detergent, supporting that inhibition isn't due to inhibitor aggregation. **D)** Aggregation test assays with high enzyme concentration, and centrifuged inhibitor, to confirm BT2162 is not an aggregator **E)** ITC dose-response curve of BT2162 binding to HPSE showing a 1:1 binding with an affinity of 10 mM

Altogether, these results support the conclusion that BT2162 is a non-competitive (allosteric), reversible inhibitor of HPSE. We also confirmed that BT2162 is not a non-specific aggregator and that it binds to a single site on HPSE, HS-mimetic inhibitors of HPSE that bind either at the active site or on the HS binding domains present as competitive inhibitors, therefore the non-competitive inhibition mechanism seen here is a strong indication that this inhibitor binds at an allosteric site, remote from the active site. It has been suggested that because of the various activities of HPSE, and the poor correlation between the structure and activity of various small molecule inhibitors, that allosteric sites may exist.[36] Indeed, we show in **Chapter 3** that structural changes at the remote heparan binding domain 3 affect activity.[46] However, the exact location of the binding site targeted by BT2162 remains unknown.

Identifying the allosteric binding site (i): protein X-ray crystallography

To determine the binding site of BT2162, protein X-ray crystallography was performed using crystals of HPSE P6, which we tested to ensure was inhibited (and therefore bound to BT2162 in the same manner as HPSE WT; **SI Figure 2**). Firstly, various high-throughput crystal screens with different protein:ligand concentrations were attempted to find an ideal condition that facilitates ligand binding. This yielded several crystals of HPSE grown in the presence of BT2162, although none exhibited clear electron density corresponding to the compound.

Secondly, using the best crystal conditions, co-crystallization with different concentrations of BT2162 and crystal soaking approaches were used (**Table 2**). Again, no discernible electron density corresponding to the small molecule was apparent in the electron density maps.

Table 2. Summary of crystallography conditions that were tested for the identification of BT2162 electron density

Condition (Co-crystallisation)	BT2162 condition	Resolution (Å)
0.2M $(NH_4)_2SO_4$, 0.1M Na Acetate pH 4.6, 25% PEG 4K	2.5 mM	1.88
0.2M $(NH_4)_2SO_4$, 30% PEG 8K	2.5 mM	1.95
0.2M $(NH_4)_2SO_4$, 30% PEG 4K	2.5 mM	1.8
0.2M $(NH_4)_2SO_4$, 0.1M Na Acetate pH 4.6, 30% PEG MME 2K	2.5 mM	1.6
1M LiCl, 0.1M Na_3Citrate pH 4.2, 20% PEG 6K	2.5 mM	2.2
20% PEG 3350, 0.2M $(NH_4)_2$H Citrate	2.5 mM	2.3
0.1M Na Acetate pH 4.5, 30% Peg 8K, 0.2M $LiSO_4$	2.5 mM	2.2
0.2M $(NH_4)_2SO_4$, 25% PEG 3350, 0.1M Bis-Tris pH 5.5	2.5mM	1.88
25% PEG 3350, 0.1M BisTris pH 5.5, 0.2M $LiSO_4$	2.5mM	1.36
2M NaCl, 10% w/v PEG 6K	2.5mM	1.6
0.2M $(NH_4)_2SO_4$, Na Acetate pH 5.5, 20% PEG 4K	2.5mM	1.6
0.2M $(NH_4)_2SO_4$, Na Acetate pH 5.5, 20% PEG 4K	2.5 mM in 10% dioxane	1.8
0.2M $(NH_4)_2SO_4$, Na Acetate pH 5.5, 15% PEG 4K	2.5 mM in 10% dioxane	1.38
20% w/v PEG 3350, 0.2M $LiSO_4$	2.5 mM	1.86
2M NaCl, 10% w/v PEG 6K	2.5 mM	1.52
0.17M $(NH_4)_2SO_4$, 15% glycerol, 25.5% PEG 4K	5 mM	1.37
25% w/v PEG 3350, 0.2M $LiSO_4$, 0.1M Bistris pH 5.5	2.5 mM	1.48
2M $(NH_4)_2SO_4$	2.5 mM	1.4
LI_2SO_4, $(NH_4)_2SO_4$, Peg3350	2.5mM	1.8
0.2M $(NH_4)_2SO_4$, 25% PEG 3350, 0.1M Tris pH 8.5	2.5 mM in dioxane	2
0.2M $(NH_4)_2SO_4$, 0.2M K Na tartrate, 0.1M Na_3Citrate pH 5.6	2.5 mM	2.5
0.17M $(NH_4)_2SO_4$, 15% v/v glycerol, 25.5% w/v PEG 4K	2.5 mM	1.9
0.2M $(NH_4)_2SO_4$, 25% PEG 3350, 0.1M Bis-tris pH 5.5	2.5 mM	2.28
2M $(NH_4)_2SO_4$, 0.1M Bis-tris pH 5.5	2.5 mM with 5% DMSO	1.89
0.2M $(NH_4)_2SO_4$, 0.1M Na Acetate pH 4.6, 30% w/v PEG MME 2K	2.5 mM	2.1
Condition (Soaking)	**BT2162 condition**	**Resolution (Å)**
0.2M $(NH_4)_2SO_4$, Na Acetate pH 4.6, 20% PEG 4K	5 mM 20 min	2.5
0.2M $(NH_4)_2SO_4$, Na Acetate pH 4.6, 25% PEG 4K	5 mM 4 hr	2.8
0.2M $(NH_4)_2SO_4$, Na Acetate pH 5, 20% PEG 4K	5 mM 20 min	2.45
0.2M $(NH_4)_2SO_4$, Na Acetate pH 4.8, 20% PEG 4K	5 mM 8 hr	3
0.2M $(NH_4)_2SO_4$, Na Acetate pH 4.6, 20% PEG 4K	5 mM 30 min	2.6
0.2M $(NH_4)_2SO_4$, Na Acetate pH 4.6, 25% PEG 4K	5 mM 8 hr	2.3
0.2M $(NH_4)_2SO_4$, Na Acetate pH 4.4, 20% PEG 4K	5 mM 30 min	2.2
0.2M $(NH_4)_2SO_4$, Na Acetate pH 4.6, 20% PEG 4K	5 mM 30 min	2.2

0.2M $(NH_4)_2SO_4$, Na Acetate pH 4.6, 20% PEG 4K	12.5 mM 5 min	2.1
0.2M $(NH_4)_2SO_4$, Na Acetate pH 4.6, 20% PEG 4K	5 mM 40 min	2.8
0.2M $(NH_4)_2SO_4$, Na Acetate pH 4.6, 20% PEG 4K	5 mM 20 min	2.5
Lys mutation HPSE condition (Cocrystals)	**BT2162 condition**	**Resolution (Å)**
0.2M $(NH_4)_2SO_4$, 30% w/v PEG 4K	2.5 mM filtered with 5% DMSO	2.14
0.2M $(NH_4)_2SO_4$, 25% w/v PEG 3350, 0.1M HEPES pH 7.5	2.5 mM filtered with 5% DMSO	1.73
0.2M $(NH_4)_2SO_4$, 30% w/v PEG 8K	2.5 mM filtered with 5% DMSO	2.25
0.2M $(NH_4)_2SO_4$, 25% w/v PEG 4K	2.5 mM filtered with 5% DMSO	2.1
0.2M $(NH_4)_2SO_4$, 25% w/v PEG 3350, 0.1M Bis-Tris pH 6.5	2.5 mM filtered with 5% DMSO	3.7
0.2M $(NH_4)_2SO_4$, 25% w/v PEG 3350, 0.1M Bis-Tris pH 6.5	2.5mM filtered with 5% DMSO	2.9
0.2M $(NH_4)_2SO_4$, 0.1M Na Acetate pH 4.6, 30% w/v PEG MME 2K	2.5mM filtered with 5% DMSO	2.7
Lys mutation HPSE Condition (Soaking)	**BT2162 condition**	**Resolution (Å)**
HEPES pH 7.9, 2M $(NH_4)_2SO_4$, 20% PEG 3350	5 mM 30 min	2.8
HEPES pH 7.9, 2M $(NH_4)_2SO_4$, 25% PEG 3350	2.5 mM 30 min	2.95
0.2M Li_2SO_4, 30% PEG 3350, 0.1 M HEPES pH 7.5	5 mM 5 hr	2.25
0.2M Li_2SO_4, 30% PEG 3350, 0.1 M HEPES pH 7.5	5 mM 2 hr	2.22
0.2M Li_2SO_4, 30% PEG 8K, 0.1 M HEPES pH 8	5 mM 5 hr	2.95
0.2M Li_2SO_4, 30% PEG 8K, 0.1 M HEPES pH 8	5 mM 2 hr	3.15
0.2M Li_2SO_4. 19% PEG 3350, 0.1 M HEPES pH 7.5	5 mM 2 hr	3

Assuming that the compound might be binding at low occupancy given the solubility of BT2162 (4 μM) is similar to the K_D, the Pan-Dataset Density Analysis (PanDDA) method was used.[47] PanDDA highlights density changes in the protein (*e.g.* from bound ligands or mobile regions) by averaging the electron density maps of multiple (over 30) apo-crystals in their ground state and subtracting this density from the density of the bound state structures. This produces a partial-difference density map for the bound state, allowing the identification of significant, large regions of unmodeled density in the sample datasets. In this case, 34 apo datasets were processed and used to generate a reference dataset for PanDDA **(SI Table 2)**. However, no low occupancy binding modes were observed in the previously collected co-crystallization or soaking trials. Although this work did not result in a crystal structure of the BT2162:HPSE complex, the reference apo-HPSE PanDDA dataset will be a valuable resource for future protein crystallography studies with HPSE. It has already been used in a fragment-based drug design screen, where it helped identify low occupancy fragments that may have been missed otherwise.[48]

We hypothesised that the binding site may be located at a crystal-packing interface, where the proximity of two neighbouring proteins could block the binding site and prevent the density from being identified. To test this hypothesis, three lysine mutations (G76K, S213K, G336K) were introduced on the surface of HPSE at these crystal-packing interfaces in an effort to alter the crystal packing of the protein. Although HPSE crystallized in the same space group, the mutations allowed for a crystal packing arrangement by increasing the asymmetric unit from one HPSE protein per unit cell to two HPSE proteins per unit cell (**Table 3**) However, despite this change in packing, BT2162 could not be observed in any crystals through soaking or co-crystallization.

Table 3: Collection data of HPSE P6 and HPSE P6 with three Lys mutations

PDB ID	HPSE P6	HPSE P6 Lys$_3$
Data collection		
Space group	P 2_1 2_1 2_1	P 2_1 2_1 2_1
Cell dimensions		
a, b, c (Å)	59.78 76.09 124.43	71.494 83.904 179.962
α, β, γ (°)	90.00 90.00 90.00	90 90 90
Resolution (Å)	47.01-1.30 (1.346-1.3)	40.86 - 2.1 (2.175 - 2.1)
R_{merge}	0.07 (1.11)	0.04 (0.26)
I / σI	11.82 (0.33)	9.96 (2.54)
$CC_{1/2}$	0.998 (0.882)	0.999 (0.874)
Completeness (%)	99.78% (99.0%)	99.87 (99.91)
Redundancy	13.2 (11.3)	2.0 (2.0)
Refinement		
Resolution (Å)	47.01-1.30 (1.346-1.3)	40.86 - 2.1 (2.175 - 2.1)
No. reflections	139516 (13656)	63934 (6311)
R_{work} / R_{free}	0.142/ 0.164	0.186/0.230
No. atoms		
Protein	3660	7308
Ligand/ion	11	133
Water	363	418
B-factors		
Protein	24.89	32.21
Ligand/ion	21.94	58.19
Water	41.23	38.42

R.m.s. deviations		
Bond lengths (Å)	0.007	0.009
Bond angles (°)	1.01	1.10
Ramachandran plot		
Preferred (%)	98.68	96.48
Allowed (%)	1.32	3.19
Outliers (%)	0.00	0.33

In summary, despite extensive efforts, we were unable to obtain a co-crystal structure of BT2162 and HPSE. There are several potential reasons why a complex could not be obtained, including:

1. Poor solubility of the compound. BT2162 has poor solubility and may not be able to reach a high enough concentration to obtain detectable density

2. Dynamic binding. The binding site or the binding mode of BT2162 may be highly dynamic, which could cause a lack of density at the binding site.

3. Crystal packing. The packing of HPSE in the crystal lattice may still cause limitations if the binding site remains inaccessible.

Identifying the allosteric binding site (ii): hydrogen-deuterium exchange mass spectrometry

Because the binding site of BT2162 could not be identified through crystallography, we tested whether binding could be observed in the solution phase. Hydrogen-deuterium exchange mass spectrometry (HDX-MS) allows this by tracking the increased mass of peptides as deuterium exchanges on the amide backbone,[49,50] which a bound ligand would block. We were able to successfully use this for the binding of pentosan, supporting the binding at an allosteric binding domain, and attempted the same for BT2162.[46]
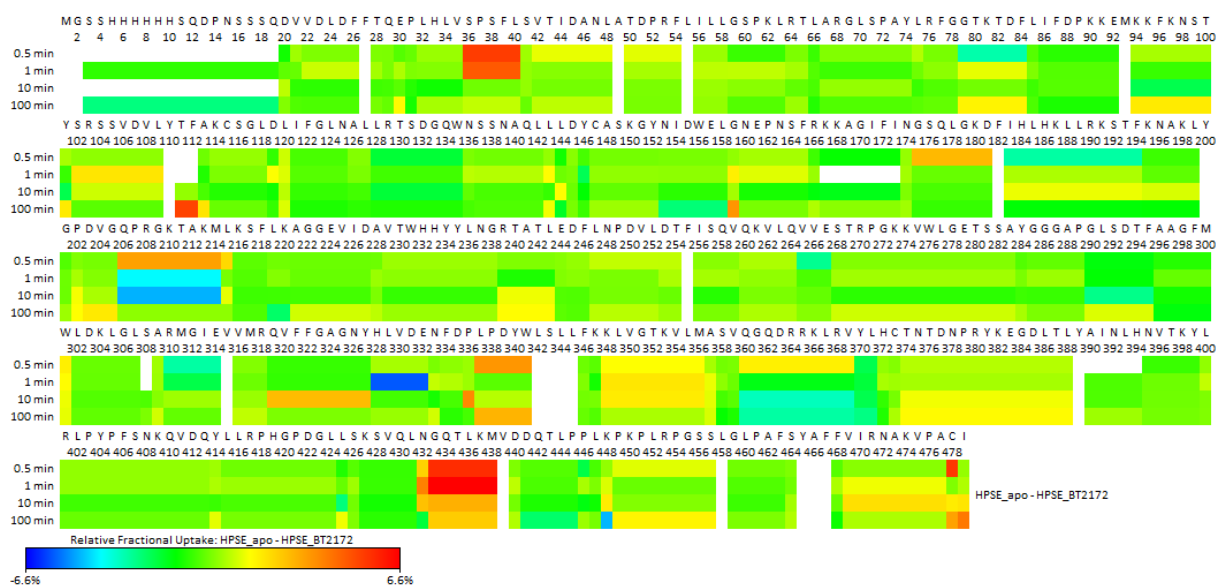
Figure 3 Hydrogen/deuterium-exchange mass spectrometry heat map representing the relative fractional uptake of individual peptides of HPSE relative to HPSE+BT2162 after 0.5-, 1-, 10-, and 100-minute incubation. Numbering of 1 starts from his-tagged 8 kDa subunit, with the 50 kDa subunit starting at 94.

The poor solubility of BT2162 gave inconclusive results, with limited change in deuterium uptake between the two systems (maximum 6.6%). Areas of HPSE with the largest change between the two systems were residues Ser53-Leu57 (36-40 in **Figure 3**), albeit only for the first minute, which sit on the TIM barrel domain opposite the active site, and residues Asn496-Met502 (432-438 in **Figure 3**), which are located on a loop that comes up from the β-sandwich domain to sit beside the TIM barrel domain. However, these results are only marginally significant and unlikely to be robust (because of the poor-quality data, only a single replicate was performed). Unfortunately, the low fractional uptake of deuterium uptake, probably due in part to the limited solubility of BT2162, was not convincing enough to identify a potential binding region.

Identifying the allosteric binding site (iii): mutagenesis

Attempts to identify the binding site of BT2162 through various co-crystal and soaking methods for X-ray protein crystallography were unsuccessful and inconclusive results were obtained from hydrogen/deuterium exchange mass spectrometry. In the absence of direct biophysical/structural data that could be used to identify the binding site of BT2162, we performed a large-scale mutagenesis study of HPSE, from the active site to remote pockets, to identify (i) whether any mutations significantly affected BT2612 binding thereby revealing the binding site and (ii) investigating how structural perturbation at remote (allosteric) regions can affect HPSE activity.

Firstly, we tested previously-made mutations of HPSE. In **Chapter 2**, we describe the mutagenesis of the large 50 kDa subunit of HPSE as part of a computational design strategy, using the server Protein Repair One Stop Shop (PROSS)[51] to generate a soluble HPSE variant that could be used for heterologous *E. coli* expression. Through this, 26 mutations were introduced, (**Table 2, Figure 4A, B**). When compared to HPSE WT, both enzymes exhibit essentially identical activity and inhibition with BT2162. This establishes that HPSE P6 is a suitable model system to use for binding analysis as it behaves identically to WT HPSE with regard to inhibition by BT2162 (**SI Figure 2**) and suggests that none of these mutations perturbs the BT2162 binding site.
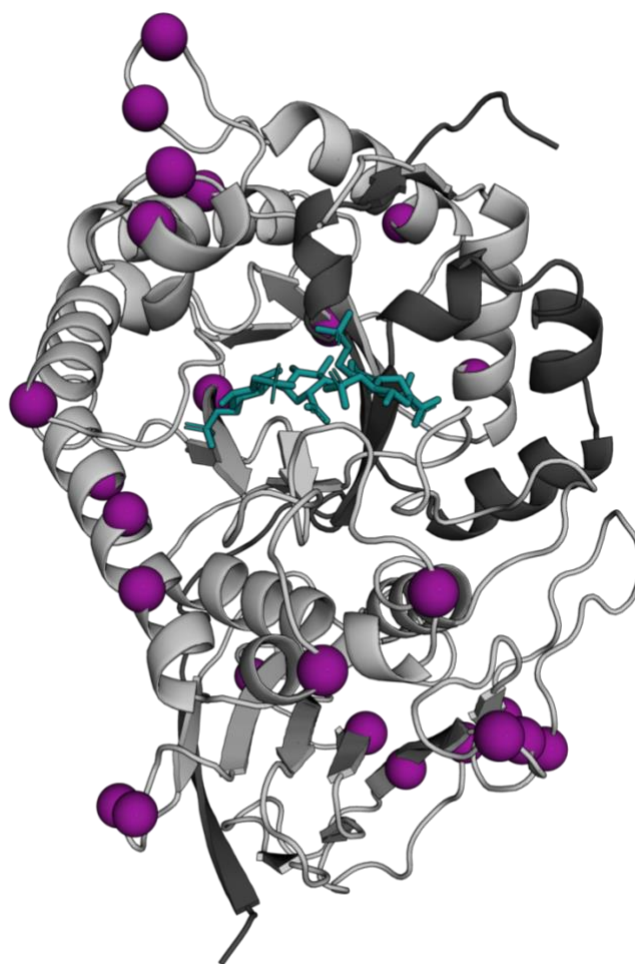
Figure 4 Mutations introduced on HPSE WT to create the HPSE P6 variant. All mutations made on the large 50 kDa subunit (purple spheres). The substrate shown in teal and the small 8 kDa subunit in dark grey.

Secondly, mutations around the active site of HPSE were tested to investigate their role in substrate binding and activity and to confirm that the binding of BT2162 was not occurring around the active site of HPSE. Specifically, mutations were primarily made to alanine residues

to remove potential hydrogen bonding interactions with substrate and/or BT2162. Of the 14 mutations made at and surrounding the binding site, several did have an effect on the maximum catalytic activity compared to the wild-type protein, which is expected due to the importance of the binding domain and active site residues for heparan sulfate affinity and binding (**Table 2, Figure 5**). The mutations that had the greatest effect on activity were K159A, F160A and K161A, which are part of HBD-1 and have previously been shown to be most important for substrate binding,[52] and have demonstrated importance per one of our previous papers.[46] Residues surrounding the active site, such as K231A, Y348A and E396A, also had a significant impact, whereas residues further away from the active site did not affect activity. With regard to BT2162 binding, there was no significant change in the IC_{50} values for any of these mutants, compared to WT protein. The lack of change in IC_{50} from mutations at the active site supports our biophysical and kinetic analyses which indicates that BT2162 is a potential allosteric non-competitive inhibitor, not binding in the substrate binding cleft of HPSE.
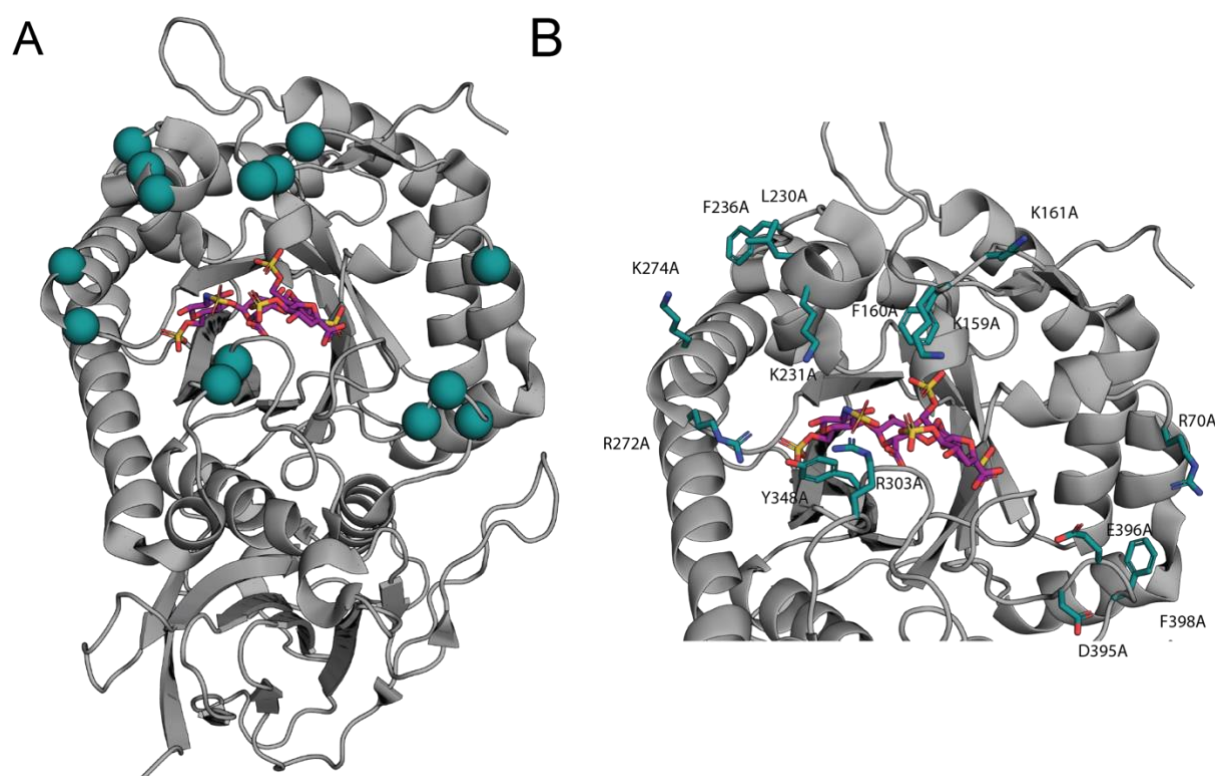
Figure 5: Active site mutations targeting BT2162 binding. **A)** Heparanase structure highlighting the region of mutations made around the HS binding site. Substrate bound to demonstrate binding region in purple. Mutations in teal. **B)** Mutations made at the active site, highlighting mutations shown to play a role in HS binding, predominantly Arg and Lys residues.

As it is unlikely that BT2162 binds at the binding cleft of HPSE, and there is limited knowledge of allosteric sites on the surface of HPSE, a series of HPSE P6 point mutants were tested to

hopefully disrupt BT2162 binding and inhibition on the surface of the protein. These mutations were based on the investigation of heparan binding domains, potential allosteric binding pockets and potential cavities identified over the protein surface. These binding pockets were identified computationally with SiteMap, from the Schrödinger drug design package (Version 18.3). This program identifies possible allosteric sites and assesses their "druggability" independent of the ligand.[53] A site is identified by its size and exposure to solvent, and the site's hydrophobic and hydrophilic character. HPSE has five potential binding sites, with one of those being the substrate binding site (**SI Table 2**).
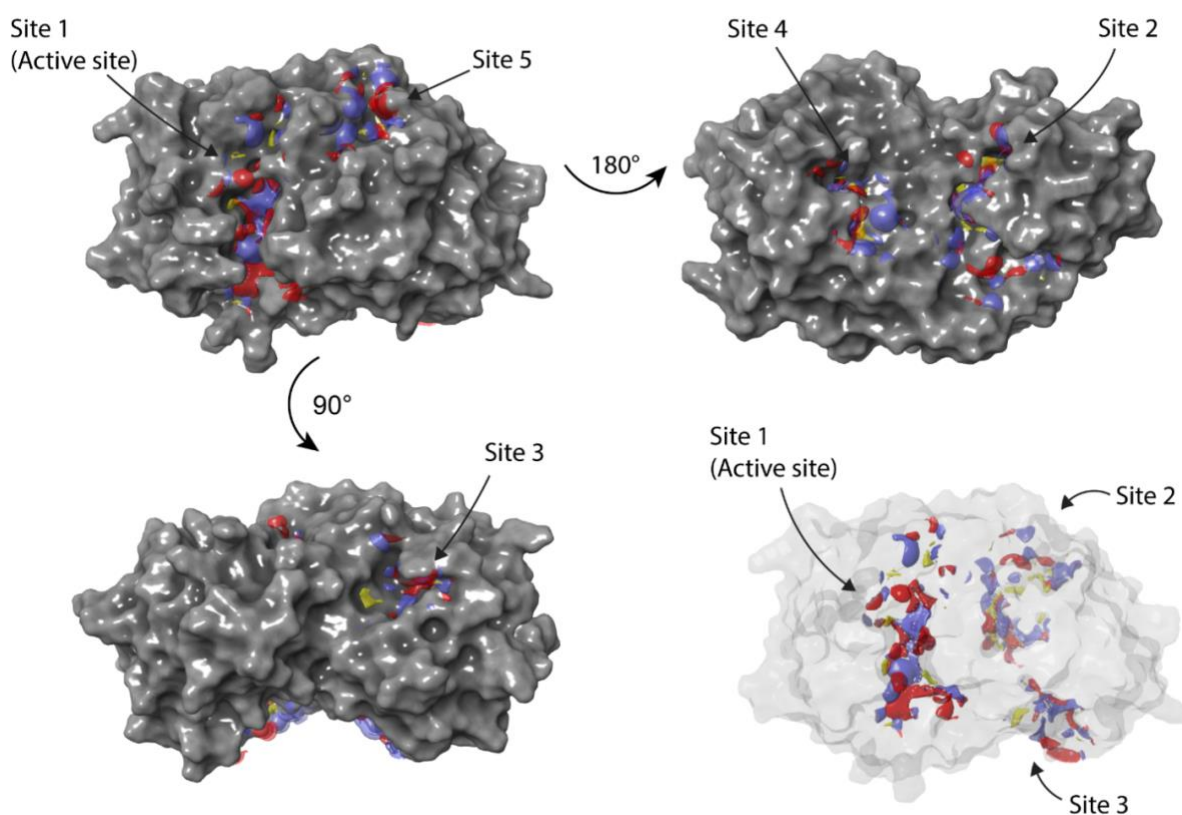
Figure 6 Identification of binding sites of HPSE identified by SiteMap. A) Site 1 (substrate binding site) and Site 4. B) Site 2 and Site 4. C) Site 3. D) Relative positioning of Sites 1, 2 and 3. Hydrophobic residues are yellow, hydrogen bond donors are blue and hydrogen bond acceptors are red.

Based on the binding site predictions, five different sites were identified on the surface of the protein (**Figure 4, SI Figure 3**). The substrate binding site was identified as the top site, which as expected had both a high site and druggability score. All other sites except Site 5 reached the threshold of a potential binding site. Mutations around the HPSE surface were also tested to obtain a broad screen over the surface of the protein in hope to find the allosteric site.

Table 2. Data summary of HPSE P6 point mutants against BT2162 inhibition, highlighting IC_{50} and Max activity, relative to HPSE P6. Error is presented in standard deviation.

	Mutation	IC_{50} (µM)	Max activity (%)
HPSE P6	**N178K, A195S, L197G, S212A, S219D, L230R, D234G, E244K, Q248H, R273G, S292A, R307L, I318T, S322Q, F327L, L354G, S426Q, K427D, K477Q, L483H, H486D, L498Q, M512K, S513P, S530A, A540P**	15.2 ± 2.3	97.5 ± 8.1
Binding site 1 (active site)	**R70A**	14 ± 1	98.5 ± 2.5
	K159A, F160A	13.6 ± 4.1	39.4 ± 3.5
	K161A	13.8 ± 2.3	35.7 ± 1.1
	R230A, F236A	12.1 ± 5.4	101 ± 1
	K231A	11.8 ± 2.7	17.3 ± 2
	R272A	12.7 ± 1.9	56.6 ± 1.6
	K274A	14.4 ± 1.8	100 ± 2
	R303A	17.5 ± 4.3	110 ± 2
	Y348A	15.5 ± 1.2	21 ± 1
	D395A	14 ± 1	52.4 ± 2.8
	E396A	11.3 ± 1.9	31.1 ± 6.2
	F398A	17.3 ± 0.7	50.3 ± 2.4
Binding site 2	**F56M**	21.1 ± 2.7	77.2 ± 4.7
	K411Q	12.2 ± 1.4	94.4 ± 2.2
	G415E	11.5 ± 2.8	104 ± 5
	N439E	9.1 ± 1.9	104 ± 1
	N439E, D441H	11 ± 2	93.6 ± 3.6
	D441H	10 ± 4	103 ± 1
	N442A	14.5 ± 1.1	33.8 ± 1.5
Binding site 3	**R428A**	24.5 ± 8.8	35.2 ± 1.8
Binding site 4	**R93N**	14.4 ± 0.8	41.7 ± 1.8
	D183A	9.83 ± 1.3	102 ± 3
	N217A	13.8 ± 1.5	54 ± 2
Binding site 5	**K501A**	15.5 ± 6.5	68.2 ± 11.4
	D505A, Q506A	24 ± 12	23 ± 1
Other surface mutations	**D40A**	15 ± 1	68.1 ± 3.4
	H50A	13.7 ± 6.4	100 ± 4
	S55A	29.3 ± 2.8	93.6 ± 1.7
	L57A	17.4 ± 1.6	11.8 ± 0.5

S58A	15.7 ± 1.2	53.8 ± 2.1
R81A	15.6 ± 1.4	71.6 ± 5.7
Y91A	12.9 ± 1.3	22.7 ± 1.78
K108A	13.8 ± 1.1	28.6 ± 1.5
S169A	14.9 ± 2.7	66.9 ± 3.1
N203A	11.2 ± 4	8.9 ± 3.8
H250A	19.9 ± 0.6	102.6 ± 2.1
R254A	20.5 ± 0.4	97.4 ± 1
D291A	22.6 ± 0.9	106 ± 5
K338A	13.9 ± 0.4	87.4 ± 1.6
E378A	12.6 ± 3.9	17.2 ± 5.1
F398A	17.3 ± 0.7	50.3 ± 2.4
K417A	20 ± 4	17.1 ± 1.4
R465A	13.6 ± 3.2	104 ± 5

For this study, 45 mutations were tested for BT2162 inhibition and catalytic activity, including the nine mutations having being tested previously (**Chapter 3, Figure 5**). With the addition of the 26 mutations from the HPSE P6 variant, a total of 71 mutations were tested to find the location of BT2162 binding. No mutation showed a large change in BT2162 binding, although some showed small shifts. For instance, S55A and F56M showed the biggest difference, with IC_{50} values of 29.3 μM and 21.1 μM respectively (with little effect on maximum activity), where S55A is nearly double the IC_{50} of HPSE P6 of 15.2 μM. Nearby residues of S58A, Y91A, L57A, and D183A showed no change, although this may indicate that S55 is on the periphery of the binding site.

Heparanase activity is perturbed by remote structural changes

Although a ligand binding site was not unambiguously identified, this mutagenesis screen revealed that some mutations, a considerable distance away from the active site, had a large effect on catalytic activity, highlighting the potential importance of allosteric networks within HPSE (**Table 2**; **Figure 7**). As previously mentioned, mutations around the active site had a significant effect on activity, but residues near the core of the TIM barrel domain, such as L57A, S58A, Y91A, R93N, and E378A, all significantly reduced activity, possibly by disrupting the folding of a highly structurally conserved TIM barrel domain.[54] As seen in **Chapter 2**, this area was quite conserved with limited mutations and showed very limited fluctuations. Interestingly, residues within the β-sandwich domain, which plays no role in the catalytic activity of HPSE, also had a significant effect on catalysis, specifically the mutations

D40A, K417A, R428A, N442A, K501A, D505A, and Q506A. This could suggest that these domains are dynamically coupled, as is common in many enzymes.[55–57]
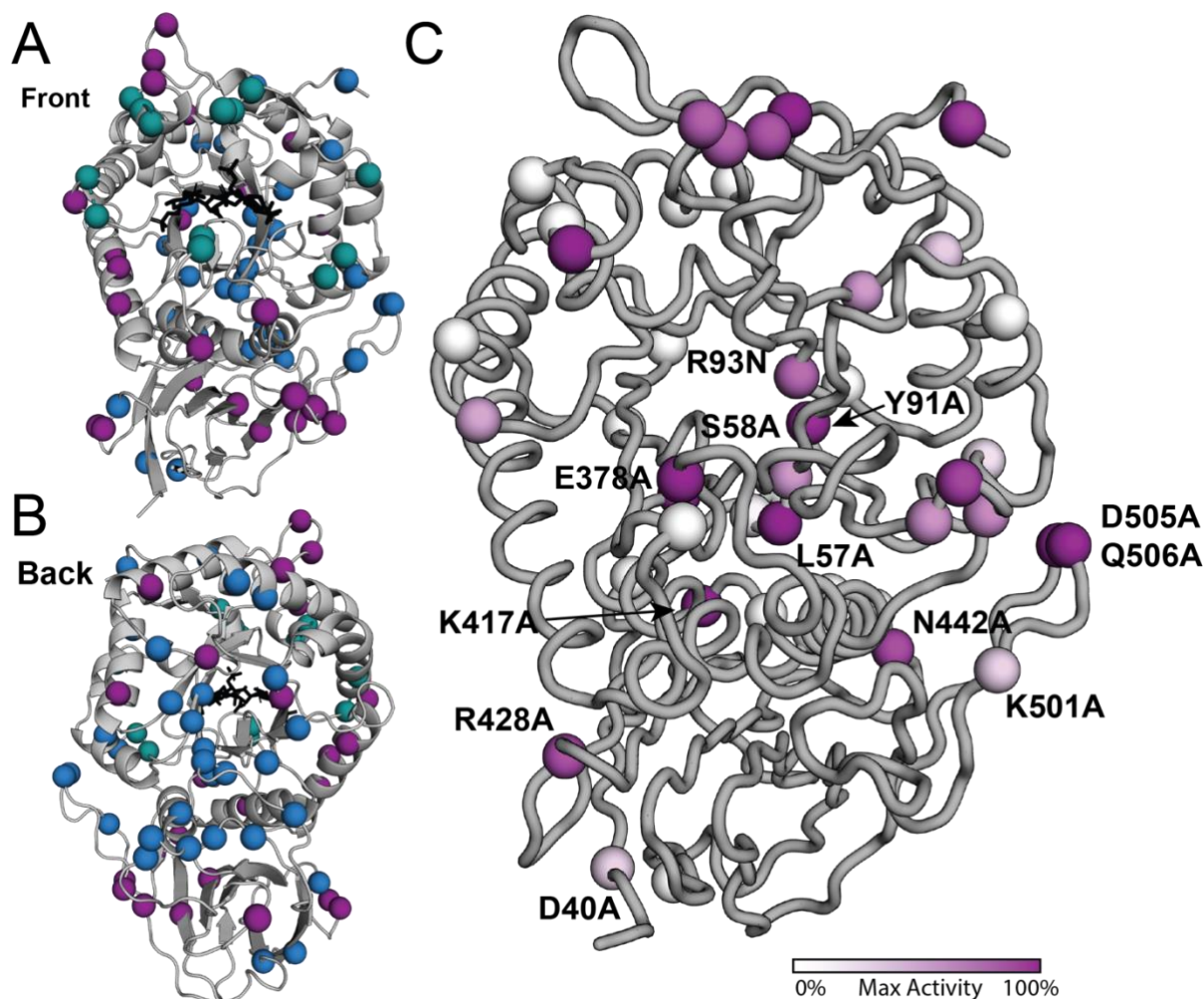
Figure 7: Location and effect of HPSE P6 point mutations on catalytic activity. **A-B)** All mutations that were made in the process of scanning for the BT2162 binding site. PROSS mutations for HPSE P6 are in purple, active site mutations in green and all other mutations are in blue. The focus was on the back of HPSE, with some mutations made also in the b-Sandwich domain. **C)** The effect of point mutations on the maximum activity of HPSE compared to HPSE P6 activity.

Discussion

In recent years, numerous experimental studies have demonstrated the potential of HPSE as a drug target for the treatment of various cancers and diseases. HPSE has been validated as a target for anticancer therapies, with a growing body of evidence supporting its role in cancer growth, angiogenesis, and metastasis.[58] Clinical trials have also provided support for the use of HPSE inhibitors, particularly oligo- and polysaccharide-derived inhibitors, in chemotherapy regimens.[34,35] These findings suggest that HPSE may represent a promising target for the

development of new and effective cancer treatments.[59] One limitation of competitive oligosaccharide inhibitors is their inability to be administered orally due to their chemical nature. These inhibitors, which are often heterogeneous mixtures of molecules with different molecular weights that bind to multiple targets, can also be difficult to characterize and standardise. While previously thought to be the most potent inhibitors of HPSE (in the nM range), work described in **Chapter 3** suggests that these inhibitors may have more complex inhibitory mechanisms and a reduced affinity for HPSE.[46] In contrast, small molecule inhibitors have the advantage of being easily modified to improve their metabolic and pharmacokinetic properties, making them suitable for oral administration.

Many small molecule inhibitors designed to inhibit HPSE are assumed to bind to the enzyme's active site, with limited studies being conducted to prove whether or not this is the case. These competitive binding, small molecule inhibitors have had limited luck reaching clinical trials. This may have due to the large binding cleft making it difficult to create specific binders, potentially causing them to also be non-specific. Allosteric non-competitive inhibitors, on the other hand, have the advantage of being more specific at protein-specific sites, potentially reducing the incidence of side effects. This has been shown useful for kinases, where competitive inhibitors showed poor selectivity as well as acquired resistance, now having allosteric inhibitors in clinical trials.[60] Active site binding is often assumed for small molecule compounds, but a complete small molecule SAR has not been established, suggesting that there may be multiple binding mechanisms for small molecules.

Our analysis of the binding mechanism of the compound BT2162, which serves as a lead compound, demonstrates binding affinity in the low micromolar range. The kinetic and ITC data support that BT2162 is a reversible, non-competitive inhibitor that binds to HPSE with a 1:1 molar ratio, making it the second known non-competitive HPSE inhibitor to date.[42] As non-competitive inhibition is rare, and allosteric sites of HPSE is unknown, we also confirmed that BT2162 isn't a non-specific aggregator. However, despite our strong evidence for non-competitive inhibition of HPSE, we were unable to identify the binding site of this compound using X-ray crystallography and hydrogen-deuterium exchange mass spectrometry.

The binding site of BT2162 has proven difficult to identify using various techniques. While structures of HS-mimetics and a few small molecules have been solved, they have all been found at the active site. The only exception is the structure presented in **Chapter 3**, where we show pentosan binding to the remote heparan binding domain 3 (**Chapter 3, Figure 4**). It is

possible that BT2162, if it binds remote from the active site, may be excluded from crystal packing due to factors such as solubility issues, dynamic binding causing a lack of density, or inaccessibility in the crystal lattice. Hydrogen-deuterium exchange mass spectrometry may also be limited by BT2162's solubility, resulting in a decreased signal. Despite mutating a large percentage of HPSE in an attempt to identify the binding site, no region has been identified with confidence. While residue S55A showed a 2-fold decrease in inhibition by BT2162 with only a small change in the amino acid's structure, no allosteric site was suggested for that region, and surrounding residues did not exhibit a change in affinity. The lack of an identified allosteric site hinders our understanding of the full mechanism of these inhibitors. It is possible that other small molecule inhibitors of HPSE may also exhibit this non-competitive mechanism, but the extent of this mode of inhibition is unknown due to a lack of kinetic and structural analysis on most inhibitors. The uncertain binding mechanism of these compounds may have impeded their development and led to the abandonment of potentially more potent drugs due to the possibility of binding to other sites.

In conclusion, we have conducted a thorough structure-activity relationship analysis of quinazoline analogues and identified a lead compound with good inhibition in the low μM range. The comprehensive kinetic and mechanism studies carried out support that BT2162 is a specific, non-competitive inhibitor. While we determined through enzyme kinetics that the inhibition is non-competitive, mutagenesis and X-ray crystallography were unable to identify the binding site. This study will hopefully help other small molecule drug development programs in the future to identify the mode of inhibition seen with these inhibitor classes, which will allow for more specific and potent drugs for this complex drug target.

Methods

Cloning

The cloning of HPSE P6 is shown in Whitefield *et al.*[61] Cloning of F159A+K160A, K161A, R272A, K274A, K411Q and G415E, K417A and R428A were shown previously.[46] Genes for other HPSE P6 point mutants (F56M, R93N, D183A, R230A+F236A, K231A, R303A, Y348A, N439E, D441H, N439E-D441H) were *E. coli* codon optimised and synthesised by Twist bioscience. The original 8kDa or 50kDa subunit was removed and linearized using NdeI and XhoI restriction enzymes (Fast Digest, Thermo) and mutant genes were inserted into the multiple cloning site 2 by Gibson assembly.[62] All other mutations were introduced though PCR, amplified with Pet_duet_mid and mutagenesis primers (**SI Table 3**) and ligated with Gibson assembly. All ligated DNA was transformed to *E. coli* TOP10 cells and the plasmid DNA was extracted and sent to Garvan Institute for the Sanger sequencing to confirm the sequences.

Protein expression and purification

HPSE P6 expression was conducted following previously described methods.[61] HPSE P6 was transformed into *E. coli* SHuffle T7 Express cells (NEB), together with GroEL/ES + Trigger factor chaperones in a pACYC vector and spread on an Agar plate with ampicillin and chloramphenicol. 1% overnight seed culture from a single colony was inoculated into 1 L of LB medium supplemented with ampicillin (100 mg L^{-1}) and chloramphenicol (34 mg L^{-1}), then incubated at 37 °C for 5 hours. Overexpression was induced by adding IPTG to a final concentration of 0.05 mM and the culture was further incubated for 3 hours at 37 °C. The cell pellet was resuspended in buffer A (20 mM HEPES pH 8, 300 mM NaCl, 5 mM β-mercaptoethanol, 10% (v/v) glycerol, 20 mM Imidazole) with Turbonuclease (Sigma) and lysed by sonication (Omni Sonic Ruptor 400 Ultrasonic homogenizer). The lysate was filtered (0.45 μm) and loaded onto Ni-NTA column (GE healthcare) and eluted with 100% buffer B (buffer A + 500 mM Imidazole). The peak eluent was diluted 5 times with buffer C (20 mM HEPES pH 7.4, 200 mM NaCl, 5mM β-mercaptoethanol, 10% (v/v) glycerol) and loaded to heparin affinity column (GE healthcare) and eluted with 100% buffer D (buffer C + 1.5 M NaCl). The peak eluent was loaded onto a size exclusion column (HiLoad 26/600 Superdex 200 pg, GE healthcare) and eluted in a buffer E (20 mM Sodium Acetate pH 5, 200 mM NaCl, 10% (v/v) glycerol, 1 mM TCEP). The final concentration of the monomeric HPSE from the gel filtration was estimated by absorbance at 280 nm using NanoDrop One (Thermo).

Colorimetric assay using fondaparinux

Assays were conducted using the colorimetric assay designed by Hammond et al.[63] Bovine serum albumin-coated 96 well microplates were used for all assays and were prepared by incubation of the plates with 1 % BSA dissolved in phosphate-buffered saline (PBS) with 0.05 % Tween-20 (PBST) at 37 °C for 75 minutes. The plates were then washed three times with PBST, dried and stored at 4 °C. Assay mixtures contain 40 mM sodium acetate buffer (pH 5.0), 0.8 nM HPSE in 0.01 % Tween 20 sodium acetate buffer and 100 mM fondaparinux (Arixtra) with or without increasing concentrations of inhibitor. Plates were incubated at 37 °C for 2-20 hours before the reaction was stopped with 100 μL of 1.69 mM 4-[3-(4-Iodophenyl)-2-(4-nitrophenyl)-2H-5-tetrazolio]-1,3-benzene Disulfonate (WST-1) in 0.1 M NaOH. The plates were resealed and developed at 60 ºC for 60 minutes, and the absorbance was measured at 584 nm. Kinetics were carried out with a standard curve constructed with D-galactose as the reducing sugar standard, prepared in the same buffer and volume over the range of 0–2 mM. All curve fitting to calculate IC_{50} values and Michaelis-Menten constants, was done using GraphPad Prism software (version 8.1)

Off target assays were conducted by testing the rate of 4-nitrophenyl butyrate hydrolysis by aE7 and S2 in the presence and absence of 62.5 mM BT2162. 2.5μM protein was incubated with BT2162 for 10 minutes prior to adding 100μM of substrate. Absorbance was measured at A_{405}

Isothermal titration calorimetry

ITC experiments were performed using a Nano ITC low-volume calorimeter (TA Instruments). ITC experiments were carried out at 25°C with stirring at 350 rpm. Samples were prepared in SEC buffer (20 mM Sodium acetate pH 5.0, 200 mM NaCl, 10% (v/v) glycerol, 1 mM TCEP). Inhibitors dissolved in MQ to 2.5mM were diluted into SEC buffer with the same MQ dilution matched precisely to the protein buffer. Both forward and reverse titrations were completed, with forward titrations involving 50 μL of various inhibitor concentrations injected continuously into 500 μL of 20-25 μM protein, and reverse titrations involving 50 μL of various protein concentrations injected continuously into 500 μL of 30-80 μM inhibitor over with 150 second injection intervals. NITPIC version 1.2 Was used to integrate the thermograms. The data was serially integrated and placed into a single SEDPHAT configuration file for global analysis.

Isothermal Titration Calorimetry data analysis

The integrated ITC data was analysed using SEDPHAT version 12.1b. using the A + B ↔ AB model. Because NITPIC provides error estimates for all integrated data points, the SEDPHAT option to use these as weights in the fitting sessions was activated. Switching between the Simplex and Marquardt-Levenberg optimization routines was necessary to achieve the convergence of the parameter set. Std Deviation was calculated for K_D for both the forward and reverse reactions using GraphPad Prism software (v 9.4).

Protein Crystallography

Protein crystals were prepared as per the results provided. Crystals were then frozen directly in liquid nitrogen without cryoprotectant. Crystallographic data were collected at 100 K at the Australian Synchrotron (MX2,[64] 0.9537 Å). The derived diffraction data were indexed and integrated with DIALS.[65] Resolution estimation and data truncation were performed using the AIMLESS as in implemented in CCP4[66,67] All structures were solved by molecular replacement using the MOLREP program in CCP4[66] using the structure deposited under PDB accession code 7RG8 as a starting model. The models were refined using phenix.refine,[68] and the model was subsequently optimized by iterative model building with the program COOT v0.9.[69] Alternative conformations were modelled based on mFo–DFc density and the occupancies and B-factors were determined using phenix.refine.[68] Ligands were optimised *via* elBOW and fitted into the structure *via* LigandFit. The structures were then evaluated using MolProbity[70] in Phenix.

PanDDA

The ground-state data set was made up of 32 solved HPSE crystal structures that were auto-refined using the DIMPLE pipeline in CCP4 before further refinement using Phenix.refine.[71,72] The sample data set was processed and organised as outlined in section 6.1.4 and coordinates for each of the fragments were generated using Phenix's eLBOW.[73]

The PanDDA protocol as described by Pearce *et al.* (2017) was followed with appropriate parameters changed to suit the data set being analysed.[47] The input datasets each had an RMSD of less than 0.6 Å to the reference structure. The PanDDA.analyse function was used to generate the mean density map for the ground-state dataset and identify unmodelled density in the sample data set that differs significantly from the ground-state. Identified unmodelled areas of density were then visually examined in Coot, BT2162 structure was modelled in and merged

with the protein if the density looked appropriate using the PanDDA.inspect function.[74] Radar plots for the five relevant validation metrics for each of the merged models were generated (see supplementary figure 2 and 4). If the validation scores were within the acceptable ranges, merged models can be exported through the PanDDA.export script for subsequent rounds of refinement to minimise rotamer and Ramachandran outliers using Phenix.refine.

HDX-Mass Spectrometry Deuterium labelling and quench conditions

A Trajan LEAP HDX Automation manager was used to automate labelling, quenching and injection of samples. Inhibitor-bound proteins were prepared at a concentration of 12 μM with a 10x molar ratio of inhibitor. 3 μL of protein sample was incubated in 57μL sodium acetate buffer pH 5.0 reconstituted in D_2O (99.90%, Sigma). Deuterium labelling was performed for 0.5 min, 1 min, 10 min and 100 min, followed by quenching of 50μL the deuterium exchange reaction mixture in 50μL pre-chilled 50mM sodium acetate quench solution containing 2 M guanidine hydrochloride (Sigma) and 200 mM TCEP, to lower the pH to 2.5 and lower temperature to 0.1 °C. A post-quench reaction time of 30 sec was used.

Mass Spectrometry and peptide identification

Quenched samples (80 μL) were injected onto chilled Trajan HDX Manager. Samples were subjected to online digestion using an immobilized Waters Enzymate BEH pepsin column (2.1 × 30 mm) in 0.1% formic acid in water at 100 μL/min. The proteolyzed peptides were trapped in a 2.1 × 5 mm C18 trap (ACQUITY BEH C18 VanGuard Pre-column, 1.7 μm, Waters, Milford, MA). The proteolyzed peptides were eluted using acetonitrile and 0.1 % formic acid gradient (5 to 35 % 6 min, 35% to 40% 1min, 40% to 95% 1 min, 95% 2 min) at a flow rate of 40 μL/min using an ACQUITY UPLC BEH C18 Column (1.0 × 100 mm, 1.7 μm, Waters, Milford, MA) pumped by UPLC I-Class Binary Solvent Manager (Waters, Milford, MA). A positive electrospray ionization source fitted with a low flow probe was used to ionize peptides sprayed into a SYNAPT G2-Si mass spectrometer (Waters, Milford, MA). Data was acquired in Masslynx 4.2 in MS^E acquisition mode using 200pg/μL Leucine enkephalin and 100 fmol/μL [Glu1]-fibrinopeptideB ([Glu1]-Fib). Lockspray was introduced by infusion at a flow rate of 5 μL/min into the mass spectrometer. Protein Lynx Global Server (PLGS) (version 3.0) was used to identify peptides in non-deuterated protein samples. The identified peptides were further filtered in DynamX (version 3.0) using a minimum intensity cut-off of 10,000 for product and precursor ions, minimum products per amino acids of 0.3 and a precursor ion mass tolerance

of 5 ppm using DynamX (version 3.0) (Waters, Milford, MA). Deuterium exchange plots, relative deuterium exchange and difference plots were generated. All deuterium exchange experiments were performed in triplicate and reported values are not corrected for deuterium back exchange.

Binding site identification

Binding sites of the HPSE structure were identified using SiteMap v4.8. Sites were determined to contain a minimum of 15 site points, using a restrictive definition of hydrophobicity, and a fine grid. The top five sites were selected for further analysis. Docking grids were generated using grid generator from Glide v8.0, using the SiteMap site points. The default options for the scaling factor (1.0) of the Van der Waals radii of receptor atoms and the partial charge cut off (0.25) were used. Advanced settings were not changed. The site allowed ligands with a length of less than 20 Å to be docked, and the centre of the SiteMap entry had to contain the ligand. Scaling of Van der Waals radii for the ligand were left as default (0.80 scaling factor and 0.15 partial charge cut off).

References

(1)   Kim, S.-H.; Turnbull, J.; Guimond, S. Extracellular Matrix and Cell Signalling: The Dynamic Cooperation of Integrin, Proteoglycan and Growth Factor Receptor. *The Journal of endocrinology* **2011**, *209* (2), 139–151. https://doi.org/10.1530/JOE-10-0377.

(2)   Meirovitz, A.; Goldberg, R.; Binder, A.; Rubinstein, A. M.; Hermano, E.; Elkin, M. Heparanase in Inflammation and Inflammation-Associated Cancer. *FEBS Journal* **2013**, *280* (10), 2307–2319. https://doi.org/10.1111/febs.12184.

(3)   Ishai-Michaeli, R.; Eldor, A.; Vlodavsky, I. Heparanase Activity Expressed by Platelets, Neutrophils, and Lymphoma Cells Releases Active Fibroblast Growth Factor from Extracellular Matrix. *Cell regulation* **1990**, *1* (11), 833–842.

(4)   Gutter-Kapon, L.; Alishekevitz, D.; Shaked, Y.; Li, J.-P.; Aronheim, A.; Ilan, N.; Vlodavsky, I. Heparanase Is Required for Activation and Function of Macrophages. *Proceedings of the National Academy of Sciences of the United States of America* **2016**, *113* (48), E7808–E7817. https://doi.org/10.1073/pnas.1611380113.

(5)   Caruana, I.; Savoldo, B.; Hoyos, V.; Weber, G.; Liu, H.; Kim, E. S.; Ittmann, M. M.; Marchetti, D.; Dotti, G. Heparanase Promotes Tumor Infiltration and Antitumor Activity of CAR-Redirected T Lymphocytes. *Nature Medicine* **2015**, *21* (5), 524–529. https://doi.org/10.1038/nm.3833.

(6)   Zcharia, E.; Metzger, S.; Chajek-Shaul, T.; Aingorn, H.; Elkin, M.; Friedmann, Y.; Weinstein, T.; Li, J.-P.; Lindahl, U.; Vlodavsky, I. Transgenic Expression of Mammalian Heparanase Uncovers Physiological Functions of Heparan Sulfate in Tissue Morphogenesis, Vascularization, and Feeding Behavior. *The FASEB Journal* **2004**, *18* (2), 252–263. https://doi.org/10.1096/fj.03-0572com.

(7) Bar-Ner, M.; Mayer, M.; Schirrmacher, V.; Vlodavsky, I. Involvement of Both Heparanase and Plasminogen Activator in Lymphoma Cell-Mediated Degradation of Heparan Sulfate in the Subendothelial Extracellular Matrix. *Journal of Cellular Physiology* **1986**, *128* (2), 299–306. https://doi.org/10.1002/jcp.1041280223.

(8) Elkin, M.; Ilan, N.; Ishai-Michaeli, R.; Friedmann, Y.; Papo, O.; Pecker, I.; Vlodavsky, I. Heparanase as Mediator of Angiogenesis: Mode of Action. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* **2001**, *15* (9), 1661–1663.

(9) Digre, A.; Singh, K.; Åbrink, M.; Reijmers, R. M.; Sandler, S.; Vlodavsky, I.; Li, J.-P. Overexpression of Heparanase Enhances T Lymphocyte Activities and Intensifies the Inflammatory Response in a Model of Murine Rheumatoid Arthritis. *Scientific Reports* **2017**, *7* (1), 46229. https://doi.org/10.1038/srep46229.

(10) Sasaki, N.; Higashi, N.; Taka, T.; Nakajima, M.; Irimura, T. Cell Surface Localization of Heparanase on Macrophages Regulates Degradation of Extracellular Matrix Heparan Sulfate. *J Immunol* **2004**, *172* (6), 3830–3835.

(11) Gallagher, J. T. Heparan Sulfate: Growth Control with a Restricted Sequence Menu. *J Clin Invest* **2001**, *108* (3), 357–361. https://doi.org/10.1172/JCI13713.

(12) Parish, C. R.; Freeman, C.; Ziolkowski, A. F.; He, Y. Q.; Sutcliffe, E. L.; Zafar, A.; Rao, S.; Simeonovic, C. J. Unexpected New Roles for Heparanase in Type 1 Diabetes and Immune Gene Regulation. *Matrix Biol* **2013**, *32* (5), 228–233. https://doi.org/10.1016/j.matbio.2013.02.007.

(13) Zcharia, E.; Jia, J.; Zhang, X.; Baraz, L.; Lindahl, U.; Peretz, T.; Vlodavsky, I.; Li, J. P. Newly Generated Heparanase Knock-out Mice Unravel Co-Regulation of Heparanase and Matrix Metalloproteinases. *PLoS One* **2009**, *4* (4), e5181. https://doi.org/10.1371/journal.pone.0005181.

(14) Nakajima, M.; Irimura, T.; Di Ferrante, D.; Di Ferrante, N.; Nicolson, G. L. Heparan Sulfate Degradation: Relation to Tumor Invasive and Metastatic Properties of Mouse B16 Melanoma Sublines. *Science (New York, N.Y.)* **1983**, *220* (4597), 611–613. https://doi.org/10.1126/SCIENCE.6220468.

(15) Keenan, T. D. L.; Pickford, C. E.; Holley, R. J.; Clark, S. J.; Lin, W.; Dowsey, A. W.; Merry, C. L.; Day, A. J.; Bishop, P. N. Age-Dependent Changes in Heparan Sulfate in Human Bruch's Membrane: Implications for Age-Related Macular Degeneration. *Investigative Opthalmology & Visual Science* **2014**, *55* (8), 5370. https://doi.org/10.1167/iovs.14-14126.

(16) Abu El-Asrar, A. M.; Alam, K.; Nawaz, M. I.; Mohammad, G.; Van den Eynde, K.; Siddiquei, M. M.; Mousa, A.; De Hertogh, G.; Geboes, K.; Opdenakker, G. Upregulated Expression of Heparanase in the Vitreous of Patients With Proliferative Diabetic Retinopathy Originates From Activated Endothelial Cells and Leukocytes. *Investigative Opthalmology & Visual Science* **2015**, *56* (13), 8239. https://doi.org/10.1167/iovs.15-18025.

(17) B, B.; C, Y.; A, de N.; I, G.; ML, M.-H.; I, J.; NAF, J.; N, R.; M, de G.; P, P.; M, K.; LAB, J.; T, N.; MG, N.; L, H.; FL, van de V.; R, D.; Q, de M.; J, van der V. Increased Plasma Heparanase Activity in COVID-19 Patients. *Frontiers in immunology* **2020**, *11*. https://doi.org/10.3389/FIMMU.2020.575047.

(18) Drost, C. C.; Rovas, A.; Osiaevi, I.; Rauen, M.; van der Vlag, J.; Buijsers, B.; Salmenov, R.; Lukasz, A.; Pavenstädt, H.; Linke, W. A.; Kümpers, P. Heparanase Is a Putative Mediator of Endothelial Glycocalyx Damage in COVID-19 – A Proof-of-Concept Study. *Frontiers in Immunology* **2022**, *13*.

(19) Alekseeva, A.; Mazzini, G.; Giannini, G.; Naggi, A. Structural Features of Heparanase-Inhibiting Non-Anticoagulant Heparin Derivative Roneparstat. *Carbohydr Polym* **2017**, *156*, 470–480. https://doi.org/10.1016/j.carbpol.2016.09.032.

(20) Zhou, H.; Roy, S.; Cochran, E.; Zouaoui, R.; Chu, C. L.; Duffner, J.; Zhao, G.; Smith, S.; Galcheva-Gargova, Z.; Karlgren, J.; Dussault, N.; Kwan, R. Y. Q.; Moy, E.; Barnes, M.; Long, A.; Honan, C.; Qi, Y. W.; Shriver, Z.; Ganguly, T.; Schultes, B.; Venkataraman, G.; Kishimoto, T. K. M402, a Novel Heparan Sulfate Mimetic, Targets Multiple Pathways Implicated in Tumor Progression and Metastasis. *PLoS ONE* **2011**, *6* (6), e21106. https://doi.org/10.1371/journal.pone.0021106.

(21) Parish, C. R.; Freeman, C.; Brown, K. J.; Francis, D. J.; Cowden, W. B.; Hampson, I. N. Identification of Sulfated Oligosaccharide-Based Inhibitors of Tumor Growth and Metastasis Using Novel in Vitro Assays for Angiogenesis and Heparanase Activity. *Cancer research* **1999**, *59* (14), 3433–3441. https://doi.org/10.1158/0008-5472.can-03-2718.

(22) Dredge, K.; Hammond, E.; Davis, K.; Li, C. P.; Liu, L.; Johnstone, K.; Handley, P.; Wimmer, N.; Gonda, T. J.; Gautam, A.; Ferro, V.; Bytheway, I. The PG500 Series: Novel Heparan Sulfate Mimetics as Potent Angiogenesis and Heparanase Inhibitors for Cancer Therapy. *Invest New Drugs* **2010**, *28* (3), 276–283. https://doi.org/10.1007/s10637-009-9245-5.

(23) Xu, Y.-J.; Miao, H.-Q.; Pan, W.; Navarro, E. C.; Tonra, J. R.; Mitelman, S.; Camara, M. M.; Deevi, D. S.; Kiselyov, A. S.; Kussie, P.; Wong, W. C.; Liu, H. N-(4-{[4-(1H-Benzoimidazol-2-Yl)-Arylamino]-Methyl}-Phenyl)-Benzamide Derivatives as Small Molecule Heparanase Inhibitors. *Bioorganic & Medicinal Chemistry Letters* **2006**, *16* (2), 404–408. https://doi.org/10.1016/J.BMCL.2005.09.070.

(24) Courtney, S. M.; Hay, P. A.; Buck, R. T.; Colville, C. S.; Phillips, D. J.; Scopes, D. I. C.; Pollard, F. C.; Page, M. J.; Bennett, J. M.; Hircock, M. L.; McKenzie, E. A.; Bhaman, M.; Felix, R.; Stubberfield, C. R.; Turner, P. R. Furanyl-1,3-Thiazol-2-Yl and Benzoxazol-5-Yl Acetic Acid Derivatives: Novel Classes of Heparanase Inhibitor. *Bioorganic & Medicinal Chemistry Letters* **2005**, *15* (9), 2295–2299. https://doi.org/10.1016/J.BMCL.2005.03.014.

(25) Madia, V. N.; Messore, A.; Pescatori, L.; Saccoliti, F.; Tudino, V.; De Leo, A.; Bortolami, M.; Scipione, L.; Costi, R.; Rivara, S.; Scalvini, L.; Mor, M.; Ferrara, F. F.; Pavoni, E.; Roscilli, G.; Cassinelli, G.; Milazzo, F. M.; Battistuzzi, G.; Di Santo, R.; Giannini, G. Novel Benzazole Derivatives Endowed with Potent Antiheparanase Activity. *Journal of Medicinal Chemistry* **2018**, *61* (15), 6918–6936. https://doi.org/10.1021/acs.jmedchem.8b00908.

(26) Simmons, S. C.; McKenzie, E. A.; Harris, L. K.; Aplin, J. D.; Brenchley, P. E.; Velasco-Garcia, M. N.; Missailidis, S. Development of Novel Single-Stranded Nucleic Acid Aptamers against the Pro-Angiogenic and Metastatic Enzyme Heparanase (HPSE1). *PLoS ONE* **2012**, *7* (6), e37938. https://doi.org/10.1371/journal.pone.0037938.

(27) Fu, J.; Zhao, B.; Dong, Z.; Sun, Y.; Luan, H.; Shen, X.; Gao, X.; Gong, F.; Li, S.; Song, H. Heparanase DNA Vaccine Delivered by Electroporation Induces Humoral Immunity and Cytoimmunity in Animal Models. *Vaccine* **2012**, *30* (12), 2187–2196. https://doi.org/10.1016/J.VACCINE.2012.01.002.

(28) He, X.; Brenchley, P. E. C.; Jayson, G. C.; Hampson, L.; Davies, J.; Hampson, I. N. Hypoxia Increases Heparanase-Dependent Tumor Cell Invasion, Which Can Be Inhibited by Antiheparanase Antibodies. *Cancer Research* **2004**, *64* (11), 3928–3933. https://doi.org/10.1158/0008-5472.CAN-03-2718.

(29) Temkin, V.; Aingorn, H.; Puxeddu, I.; Goldshmidt, O.; Zcharia, E.; Gleich, G. J.; Vlodavsky, I.; Levi-Schaffer, F. Eosinophil Major Basic Protein: First Identified Natural

Heparanase-Inhibiting Protein. *The Journal of allergy and clinical immunology* **2004**, *113* (4), 703–709. https://doi.org/10.1016/j.jaci.2003.11.038.

(30) de Boer, C.; Armstrong, Z.; Lit, V. A. J.; Barash, U.; Ruijgrok, G.; Boyango, I.; Weitzenberg, M. M.; Schröder, S. P.; Sarris, A. J. C.; Meeuwenoord, N. J.; Bule, P.; Kayal, Y.; Ilan, N.; Codée, J. D. C.; Vlodavsky, I.; Overkleeft, H. S.; Davies, G. J.; Wu, L. Mechanism-Based Heparanase Inhibitors Reduce Cancer Metastasis in Vivo. *Proceedings of the National Academy of Sciences* **2022**, *119* (31), e2203167119. https://doi.org/10.1073/pnas.2203167119.

(31) Shiozawa, H.; Takahashi, M.; Takatsu, T.; Kinoshita, T.; Tanzawa, K.; Hosoya, T.; Furuya, K.; Takahashi, S.; Furihata, K.; Seto, H. Trachyspic Acid, a New Metabolite Produced by Talaromyces Trachyspermus, That Inhibits Tumor Cell Heparanase: Taxonomy of the Producing Strain, Fermentation, Isolation, Structural Elucidation, and Biological Activity. *The Journal of Antibiotics* **1995**, *48* (5), 357–362. https://doi.org/10.7164/antibiotics.48.357.

(32) Jia, L.; Ma, S. Recent Advances in the Discovery of Heparanase Inhibitors as Anti-Cancer Agents. *European Journal of Medicinal Chemistry* **2016**, *121*, 209–220. https://doi.org/10.1016/J.EJMECH.2016.05.052.

(33) Heyman, B.; Yang, Y. Mechanisms of Heparanase Inhibitors in Cancer Therapy. *Experimental hematology* **2016**, *44* (11), 1002–1012. https://doi.org/10.1016/j.exphem.2016.08.006.

(34) Khasraw, M.; Pavlakis, N.; McCowatt, S.; Underhill, C.; Begbie, S.; de Souza, P.; Boyce, A.; Parnis, F.; Lim, V.; Harvie, R.; Marx, G. Multicentre Phase I/II Study of PI-88, a Heparanase Inhibitor in Combination with Docetaxel in Patients with Metastatic Castrate-Resistant Prostate Cancer. *Annals of Oncology* **2010**, *21* (6), 1302–1307. https://doi.org/10.1093/annonc/mdp524.

(35) O'Reilly, E. M.; Barone, D.; Mahalingam, D.; Bekaii-Saab, T.; Shao, S. H.; Wolf, J.; Rosano, M.; Krause, S.; Richards, D. A.; Yu, K. H.; Roach, J. M.; Flaherty, K. T.; Ryan, D. P. Randomised Phase II Trial of Gemcitabine and Nab-Paclitaxel with Necuparanib or Placebo in Untreated Metastatic Pancreas Ductal Adenocarcinoma. *European Journal of Cancer* **2020**, *132*, 112–121. https://doi.org/10.1016/j.ejca.2020.03.005.

(36) Rivara, S.; Milazzo, F. M.; Giannini, G. Heparanase: A Rainbow Pharmacological Target Associated to Multiple Pathologies Including Rare Diseases. *Future Medicinal Chemistry* **2016**, *8* (6), 647–680. https://doi.org/10.4155/fmc-2016-0012.

(37) Pan, W.; Miao, H.-Q.; Xu, Y.-J.; Navarro, E. C.; Tonra, J. R.; Corcoran, E.; Lahiji, A.; Kussie, P.; Kiselyov, A. S.; Wong, W. C.; Liu, H. 1-[4-(1H-Benzoimidazol-2-Yl)-Phenyl]-3-[4-(1H-Benzoimidazol-2-Yl)-Phenyl]-Urea Derivatives as Small Molecule Heparanase Inhibitors. *Bioorganic & Medicinal Chemistry Letters* **2006**, *16* (2), 409–412. https://doi.org/10.1016/J.BMCL.2005.09.069.

(38) Nussinov, R.; Tsai, C.-J. Allostery in Disease and in Drug Discovery. *Cell* **2013**, *153* (2), 293–305. https://doi.org/10.1016/j.cell.2013.03.034.

(39) *Allosteric Methods and Their Applications: Facilitating the Discovery of Allosteric Drugs and the Investigation of Allosteric Mechanisms | Accounts of Chemical Research*. https://pubs-acs-org.virtual.anu.edu.au/doi/10.1021/acs.accounts.8b00570 (accessed 2023-01-10).

(40) Shoichet, B. K. Interpreting Steep Dose-Response Curves in Early Inhibitor Discovery. *J. Med. Chem.* **2006**, *49* (25), 7274–7277. https://doi.org/10.1021/jm061103g.

(41) Hammond, E.; Handley, P.; Dredge, K.; Bytheway, I. Mechanisms of Heparanase Inhibition by the Heparan Sulfate Mimetic PG545 and Three Structural Analogues. *FEBS Open Bio* **2013**, *3* (1), 346–351. https://doi.org/10.1016/j.fob.2013.07.007.

(42) Nakajima, M.; DeChavigny, A.; Johnson, C. E.; Hamada, J.; Stein, C. A.; Nicolson, G. L. Suramin. A Potent Inhibitor of Melanoma Heparanase and Invasion. *Journal of Biological Chemistry* **1991**, *266* (15), 9661–9666. https://doi.org/10.1016/S0021-9258(18)92871-1.

(43) Auld, D. S.; Inglese, J.; Dahlin, J. L. Assay Interference by Aggregation. In *Assay Guidance Manual*; Markossian, S., Grossman, A., Brimacombe, K., Arkin, M., Auld, D., Austin, C., Baell, J., Chung, T. D. Y., Coussens, N. P., Dahlin, J. L., Devanarayan, V., Foley, T. L., Glicksman, M., Gorshkov, K., Haas, J. V., Hall, M. D., Hoare, S., Inglese, J., Iversen, P. W., Kales, S. C., Lal-Nag, M., Li, Z., McGee, J., McManus, O., Riss, T., Saradjian, P., Sittampalam, G. S., Tarselli, M., Trask, O. J., Wang, Y., Weidner, J. R., Wildey, M. J., Wilson, K., Xia, M., Xu, X., Eds.; Eli Lilly & Company and the National Center for Advancing Translational Sciences: Bethesda (MD), 2004.

(44) Campbell, E. C.; Grant, J.; Wang, Y.; Sandhu, M.; Williams, R. J.; Nisbet, D. R.; Perriman, A. W.; Lupton, D. W.; Jackson, C. J. Hydrogel-Immobilized Supercharged Proteins. *Advanced Biosystems* **2018**, *2* (7), 1700240. https://doi.org/10.1002/adbi.201700240.

(45) Jackson, C. J.; Liu, J.-W.; Carr, P. D.; Younus, F.; Coppin, C.; Meirelles, T.; Lethier, M.; Pandey, G.; Ollis, D. L.; Russell, R. J.; Weik, M.; Oakeshott, J. G. Structure and Function of an Insect α-Carboxylesterase (AEsterase7) Associated with Insecticide Resistance. *Proceedings of the National Academy of Sciences* **2013**, *110* (25), 10177–10182. https://doi.org/10.1073/pnas.1304097110.

(46) Whitefield, C.; Vo, Y.; Schwartz, B. D.; Hepburn, C.; Ahmed, F. H.; Onagi, H.; Banwell, M. G.; Nelms, K.; Malins, L. R.; Jackson, C. J. The Complex Inhibitory Mechanism of Glycomimetics with Human Heparanase. bioRxiv January 5, 2023, p 2023.01.05.522817. https://doi.org/10.1101/2023.01.05.522817.

(47) Pearce, N. M.; Krojer, T.; Bradley, A. R.; Collins, P.; Nowak, R. P.; Talon, R.; Marsden, B. D.; Kelm, S.; Shi, J.; Deane, C. M.; Von Delft, F. A Multi-Crystal Method for Extracting Obscured Crystallographic States from Conventionally Uninterpretable Electron Density. *Nature Communications* **2017**, *8*. https://doi.org/10.1038/ncomms15123.

(48) Davies, L. Fragment-Based Drug Design for Inhibitors of Heparanase, 2021.

(49) Konermann, L.; Pan, J.; Liu, Y.-H. Hydrogen Exchange Mass Spectrometry for Studying Protein Structure and Dynamics. *Chem Soc Rev* **2011**, *40* (3), 1224–1234. https://doi.org/10.1039/c0cs00113a.

(50) Masson, G. R.; Burke, J. E.; Ahn, N. G.; Anand, G. S.; Borchers, C.; Brier, S.; Bou-Assaf, G. M.; Engen, J. R.; Englander, S. W.; Faber, J.; Garlish, R.; Griffin, P. R.; Gross, M. L.; Guttman, M.; Hamuro, Y.; Heck, A. J. R.; Houde, D.; Iacob, R. E.; Jørgensen, T. J. D.; Kaltashov, I. A.; Klinman, J. P.; Konermann, L.; Man, P.; Mayne, L.; Pascal, B. D.; Reichmann, D.; Skehel, M.; Snijder, J.; Strutzenberg, T. S.; Underbakke, E. S.; Wagner, C.; Wales, T. E.; Walters, B. T.; Weis, D. D.; Wilson, D. J.; Wintrode, P. L.; Zhang, Z.; Zheng, J.; Schriemer, D. C.; Rand, K. D. Recommendations for Performing, Interpreting and Reporting Hydrogen Deuterium Exchange Mass Spectrometry (HDX-MS) Experiments. *Nat Methods* **2019**, *16* (7), 595–602. https://doi.org/10.1038/s41592-019-0459-y.

(51) Goldenzweig, A.; Goldsmith, M.; Hill, S. E.; Gertman, O.; Laurino, P.; Ashani, Y.; Dym, O.; Unger, T.; Albeck, S.; Prilusky, J.; Lieberman, R. L.; Aharoni, A.; Silman, I.; Sussman, J. L.; Tawfik, D. S.; Fleishman, S. J. Automated Structure- and Sequence-Based Design of Proteins for High Bacterial Expression and Stability. *Molecular Cell* **2016**, *63* (2), 337–346. https://doi.org/10.1016/j.molcel.2016.06.012.

(52) Levy-Adam, F.; Abboud-Jarrous, G.; Guerrini, M.; Beccati, D.; Vlodavsky, I.; Ilan, N. Identification and Characterization of Heparin/Heparan Sulfate Binding Domains of the Endoglycosidase Heparanase. *The Journal of biological chemistry* **2005**, *280* (21), 20457–20466. https://doi.org/10.1074/jbc.M414546200.

(53) Halgren, T. A. Identifying and Characterizing Binding Sites and Assessing Druggability. *Journal of Chemical Information and Modeling* **2009**, *49* (2), 377–389. https://doi.org/10.1021/ci800324m.

(54) Tiwari, S. P.; Reuter, N. Similarity in Shape Dictates Signature Intrinsic Dynamics Despite No Functional Conservation in TIM Barrel Enzymes. *PLOS Computational Biology* **2016**, *12* (3), e1004834. https://doi.org/10.1371/journal.pcbi.1004834.

(55) Ma, B.; Tsai, C.-J.; Haliloğlu, T.; Nussinov, R. Dynamic Allostery: Linkers Are Not Merely Flexible. *Structure* **2011**, *19* (7), 907–917. https://doi.org/10.1016/j.str.2011.06.002.

(56) Hartwell, L. H.; Hopfield, J. J.; Leibler, S.; Murray, A. W. From Molecular to Modular Cell Biology. *Nature* **1999**, *402* (6761 Suppl), C47-52. https://doi.org/10.1038/35011540.

(57) Goodey, N. M.; Benkovic, S. J. Allosteric Regulation and Catalysis Emerge via a Common Route. *Nat Chem Biol* **2008**, *4* (8), 474–482. https://doi.org/10.1038/nchembio.98.

(58) Jayatilleke, K. M.; Hulett, M. D. Heparanase and the Hallmarks of Cancer. *Journal of Translational Medicine* **2020**, *18* (1), 453. https://doi.org/10.1186/s12967-020-02624-1.

(59) Coombe, D. R.; Gandhi, N. S. Heparanase: A Challenging Cancer Drug Target. *Frontiers in Oncology* **2019**, *9*.

(60) Lu, X.; Smaill, J. B.; Ding, K. New Promise and Opportunities for Allosteric Kinase Inhibitors. *Angewandte Chemie International Edition* **2020**, *59* (33), 13764–13776. https://doi.org/10.1002/anie.201914525.

(61) Whitefield, C.; Hong, N.; Mitchell, J. A.; Jackson, C. J. Computational Design and Experimental Characterisation of a Stable Human Heparanase Variant. *RSC Chem. Biol.* **2022**, *3* (3), 341–349. https://doi.org/10.1039/D1CB00239B.

(62) Gibson, D. G. Chapter Fifteen - Enzymatic Assembly of Overlapping DNA Fragments. In *Synthetic Biology, Part B*; Voigt, C. B. T.-M. in E., Ed.; Academic Press, 2011; Vol. 498, pp 349–361. https://doi.org/10.1016/B978-0-12-385120-8.00015-2.

(63) Hammond, E.; Li, C. P.; Ferro, V. Development of a Colorimetric Assay for Heparanase Activity Suitable for Kinetic Analysis and Inhibitor Screening. *Analytical Biochemistry* **2010**, *396* (1), 112–116. https://doi.org/10.1016/j.ab.2009.09.007.

(64) Aragão, D.; Aishima, J.; Cherukuvada, H.; Clarken, R.; Clift, M.; Cowieson, N. P.; Ericsson, D. J.; Gee, C. L.; Macedo, S.; Mudie, N.; Panjikar, S.; Price, J. R.; Riboldi-Tunnicliffe, A.; Rostan, R.; Williamson, R.; Caradoc-Davies, T. T. MX2: A High-Flux Undulator Microfocus Beamline Serving Both the Chemical and Macromolecular Crystallography Communities at the Australian Synchrotron. *urn:issn:1600-5775* **2018**, *25* (3), 885–891. https://doi.org/10.1107/S1600577518003120.

(65) Winter, G.; Waterman, D. G.; Parkhurst, J. M.; Brewster, A. S.; Gildea, R. J.; Gerstel, M.; Fuentes-Montero, L.; Vollmar, M.; Michels-Clark, T.; Young, I. D.; Sauter, N. K.; Evans, G. DIALS: Implementation and Evaluation of a New Integration Package. *Acta Cryst D* **2018**, *74* (2), 85–97. https://doi.org/10.1107/S2059798317017235.

(66) Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G. W.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S. Overview of the CCP4 Suite and Current Developments. *Acta*
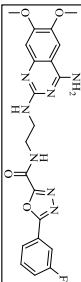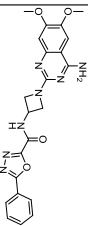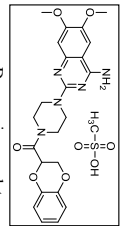
Crystallographica Section D: Biological Crystallography, 2011, *67*, 235–242. https://doi.org/10.1107/S0907444910045749.

(67) Karplus, P. A.; Diederichs, K. Linking Crystallographic Model and Data Quality. *Science* **2012**, *336* (6084), 1030–1033. https://doi.org/10.1126/science.1218231.

(68) Afonine, P. V.; Grosse-Kunstleve, R. W.; Echols, N.; Headd, J. J.; Moriarty, N. W.; Mustyakimov, M.; Terwilliger, T. C.; Urzhumtsev, A.; Zwart, P. H.; Adams, P. D. Towards Automated Crystallographic Structure Refinement with Phenix.Refine. *Acta Crystallographica Section D: Biological Crystallography* **2012**, *68* (4), 352–367. https://doi.org/10.1107/S0907444912001308.

(69) Emsley, P.; Cowtan, K. Coot: Model-Building Tools for Molecular Graphics. *Acta Crystallographica Section D: Biological Crystallography* **2004**, *60* (12 I), 2126–2132. https://doi.org/10.1107/S0907444904019158.

(70) Chen, V. B.; Arendall, W. B.; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C. MolProbity: All-Atom Structure Validation for Macromolecular Crystallography. *Acta Crystallographica Section D: Biological Crystallography* **2010**, *66* (1), 12–21. https://doi.org/10.1107/S0907444909042073.

(71) Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G. W.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S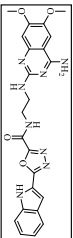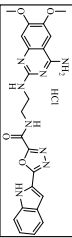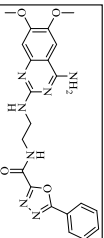. Overview of the CCP4 Suite and Current Developments. *Acta 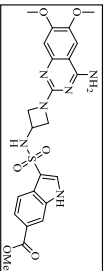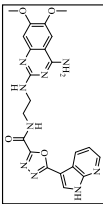Crystallographica Section D: Biological Crystallography*. International Union of Crystallography April 18, 2011, pp 235–242. https://doi.org/10.1107/S0907444910045749.

(72) Afonine, P. V.; Grosse-Kunstleve, R. W.; Echols, N.; Headd, J. J.; Moriarty, N. W.; Mustyakimov, M.; Terwilliger, T. C.; Urzhumtsev, A.; Zwart, P. H.; Adams, P. D. Towards Automated Crystallographic Structure Refinement with Phenix.Refine. *Acta Crystallographica Section D: Biological Crystallography* **2012**, *68* (4), 352–367. https://doi.org/10.1107/S0907444912001308.

(73) Liebschner, D.; Afonine, P. V.; Baker, M. L.; Bunkoczi, G.; Chen, V. B.; Croll, T. I.; Hintze, B.; Hung, L. W.; Jain, S.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R. D.; Poon, B. K.; Prisant, M. G.; Read, R. J.; Richardson, J. S.; Richardson, D. C.; Sammito, M. D.; Sobolev, O. V.; Stockwell, D. H.; Terwilliger, T. C.; Urzhumtsev, A. G.; Videau, L. L.; Williams, C. J.; Adams, P. D. Macromolecular Structure Determination Using X-Rays, Neutrons and Electrons: Recent Developments in Phenix. *Acta Crystallographica Section D: Structural Biology* **2019**, *75* (10), 861–877. https://doi.org/10.1107/S2059798319011471.

(74) Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. Features and Development of Coot. *Acta Crystallographica Section D Biological Crystallography* **2010**, *66* (4), 486–501. https://doi.org/10.1107/S0907444910007493.
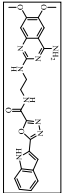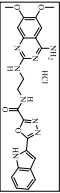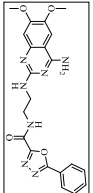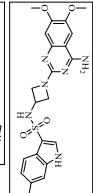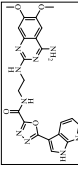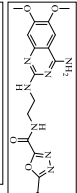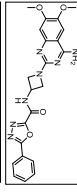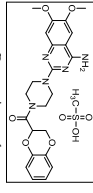
Supporting information for:

Quinazoline analogues, allosteric inhibitors of Heparanase.

Cassidy Whitefield, Junming He, Brett D Schwartz, Ben Clifton, Nansook Hong, Hafna Ahmed, Elaaf Mohamed, Martin Banwell, Lara R Malins, Colin J Jackson.

Table S1: Absorption, Distribution, Metabolism and Excretion (ADME) studies for selected HPSE inhibitors.

Structure	Heparanase IC50 (uM)	BTCODE	Neut. Comp Mol_weight	Mol_weight	(Rat HW Hep Metab) CLint (µl/min/1E6)	(Rat HW Hep Metab) t ½ (min)	(MDCK-MDR1 ABBA) A to B Papp (1E-6,cm/s)	LogD	(Solubility DMSO) Solubility (µM)	(Solubility Dried Clint) CLint (µl/min/1E6)	(human Hep Clint) t ½ (min)	(Rat HW Plasma Bind) Prot binding (% free)
	3.3	BT2169	474.48	474	18.2	38.1	0.35	2.40	0.7	3.07	226	9.14
	4.3	BT2173	474.48	511	19.4	35.7	<0.046 - <0.042	2.50	0.7	2.62	265	8.03
	15.2	BT2162	435.44	435	33.8	20.5	2.34 - 2.71	1.70	4	<1	>693	31.2
	20.5	BT2148	512.54	513	25.9	26.8	<0.045 - <0.053	2.60	0.2	8.07	85.9	-
	21	BT2177	475.47	475	6.1	-	-	1.3	9	24.3	28.5	-
	22	BT2185	453.43	453	29.4	-	-	1.7	18	23.8	29.1	30.500
	33	BT2167	447.46	447	6.1	-	-	2.4	9	3	230	0.23
Doxazosin mesylate	65	BT2013	451.48	547.58	28.7	-	-	2.9	167	32	21.66	7.55

Structure	BTCODE	(Hu Plasma Bind) Prot binding (% free)	(Mu C57/BL6 Plasma Bind) Prot binding (% free)	(CYP2A6 HLM) IC50 (µM)	(CYP3A4 HLM LC/MSMS) IC50 (µM)	(CYP2E1 HLM LC/MS) IC50 (µM)	(CYP2B6 HLM LC/MSMS) IC50 (µM)	(CYP2C8 HLM LC/MSMS) IC50 (µM)	(Hu Caco2 ABBA) A to B Papp (1E–6,cm/s)	(CYP2D6 Hu) IC50 (µM)	(CYP2C9 Hu) IC50 (µM)	(CYP1A2 Hu) IC50 (µM)	(CYP2C19 Hu) IC50 (µM)
	BT2169	7.10	12.9	>30	>30	>30	>30	>30	0.06 - 0.08				
	BT2173	8.70	11.7	>30	>30	>30	>30	>30	0.12 - 0.13				
	BT2162	29	42.5	>30	>30	>30	>30	>30	2.18 - 2.34				
	BT2148	4.70	9.95	>30	18.2 - 21.4	>30	>30	7.77 - 10.1	0.069 - 0.073				
	BT2177	24	66.3	-	>30	-	-	-	-	>30	>30	>30	>30
	BT2185	23	43.5	-	>30	-	-	-	-	>30	>30	>30	>30
	BT2167	0.67	4.97	-	>30	-	-	-	-	>30	>30	>30	>30
Doxazosin mesylate	BT2013	4.1	8	-	>30	-	-	-	-	>30	>30	>30	19.1

Figure S1: Dose-response curves of small molecule inhibitors from the SAR. Curves based of 2-4 replicates. Error represented as standard deviation.

Table S2 Data collection and refinement statistics for the HPSE PanDDA dataset

Apo structure number:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Data collection																
Space group	P 21 21 21	P 21 21 21	P 21 21 21	P 21 21 21	P 21 21 21	P 21 21 21	P 21 21 21	P 21 21 21	P 21 21 21	P 21 21 21	P 21 21 21	P 21 21 21	P 21 21 21	P 21 21 21	P 21 21 21	P 21 21 21
Cell Dimensions																
a, b, c (Å)	59.89 74.935 124.121	59.428 75.221 123.993	59.62 75.219 124.24	60.186 74.22 123.974	59.681 74.787 124.024	59.149 74.431 124.149	60.398 74.198 124.005	59.574 75.088 124.043	60.188 74.634 124.104	59.519 75.446 124.185	59.142 75.636 124.095	59.054 75.715 124.417	60.221 74.529 124.128	59.047 75.43 124.434	59.029 75.503 124.21	59.172 75.379 124.052
a, b, g (°)	90 90 90	90 90 90	90 90 90	90 90 90	90 90 90	90 90 90	90 90 90	90 90 90	90 90 90	90 90 90	90 90 90	90 90 90	90 90 90	90 90 90	90 90 90	90 90 90
Resolution (Å)	2.281 (2.281)	1.68 (1.68)	1.583 (1.583)	1.78 (1.78)	1.651 (1.651)	1.611 (1.611)	1.71 (1.71)	1.66 (1.66)	1.77 (1.77)	1.701 (1.701)	1.64 (1.64)	1.732 (1.732)	1.814 (1.814)	1.69 (1.69)	1.873 (1.873)	1.63 (1.63)
R_{merge}	0.4499 (0.458)	0.2014 (1.205)	0.3263 (1.515)	0.5323 (1.066)	0.2829 (1.537)	0.4302 (1.405)	0.2264 (1.394)	0.2243 (1.272)	0.3108 (1.073)	0.2595 (1.37)	0.3388 (1.313)	0.4004 (1.481)	0.2089 (1.054)	0.1747 (1.379)	0.2442 (1.647)	0.2134 (1.22)
R_{pim}	0.128 (0.129)	0.05631 (0.330)	0.09151 (0.429)	0.1498 (0.299)	0.07893 (0.423)	0.1211 (0.398)	0.06323 (0.379)	0.06271 (0.344)	0.08737 (0.288)	0.07319 (0.375)	0.09554 (0.364)	0.1127 (0.405)	0.05822 (0.284)	0.04861 (0.377)	0.06881 (0.457)	0.05982 (0.341)
I/sI	22.55 (5.09)	14.39 (0.94)	13.63 (0.88)	10.90 (0.87)	13.78 (0.87)	12.87 (0.93)	13.30 (0.89)	13.47 (0.91)	12.77 (0.82)	12.37 (0.82)	12.36 (0.90)	10.35 (0.81)	10.94 (0.87)	11.83 (0.90)	9.06 (0.70)	13.60 (0.90)
$CC_{1/2}$	0.864 (0.886)	0.993 (0.862)	0.962 (0.742)	0.842 (0.711)	0.987 (0.735)	0.949 (0.687)	0.993 (0.823)	0.991 (0.83)	0.979 (0.718)	0.985 (0.779)	0.965 (0.742)	0.95 (0.691)	0.996 (0.797)	0.997 (0.793)	0.991 (0.708)	0.992 (0.821)
Completeness (%)	98.32 (83.55)	99.88 (99.15)	99.10 (91.19)	99.34 (93.65)	99.13 (91.57)	99.32 (93.44)	98.98 (91.15)	99.10 (92.23)	99.46 (94.76)	99.20 (92.24)	99.97 (99.96)	99.50 (95.17)	99.32 (94.09)	99.91 (99.31)	98.85 (88.56)	99.95 (99.77)
Redundancy	13.3 (12.6)	13.7 (14.1)	13.7 (13.3)	13.6 (13.9)	13.7 (14.0)	13.7 (13.2)	13.7 (14.1)	13.8 (14.2)	13.6 (14.0)	13.7 (14.1)	13.7 (13.8)	13.7 (14.1)	13.6 (14.0)	13.7 (14.1)	13.6 (13.8)	13.7 (13.5)
Refinement																
Resolution (Å)	2.281 (2.281)	1.68 (1.68)	1.583 (1.583)	1.78 (1.78)	1.651 (1.651)	1.611 (1.611)	1.71 (1.71)	1.66 (1.66)	1.77 (1.77)	1.701 (1.701)	1.64 (1.64)	1.732 (1.732)	1.814 (1.814)	1.69 (1.69)	1.873 (1.873)	1.63 (1.63)

No. reflections	25697 (2128)	64047 (6295)	76182 (6913)	53625 (4986)	66807 (6080)	72125 (6706)	60373 (5508)	65891 (6050)	54950 (5132)	61639 (5659)	68935 (6788)	58510 (5494)	51001 (4760)	62911 (6155)	45901 (4058)	69943 (6897)
R_{work}	0.1622 (0.216)	0.1703 (0.279)	0.1750 (0.283)	0.1780 (0.268)	0.1767 (0.304)	0.1753 (0.280)	0.1714 (0.283)	0.1723 (0.277)	0.1729 (0.330)	0.1723 (0.275)	0.1730 (0.261)	0.1723 (0.270)	0.1724 (0.316)	0.1738 (0.286)	0.1748 (0.292)	0.1728 (0.275)
R_{free}	0.1945 (0.267)	0.1897 (0.309)	0.1904 (0.276)	0.2061 (0.263)	0.1939 (0.298)	0.1927 (0.293)	0.1947 (0.289)	0.1898 (0.272)	0.1943 (0.329)	0.1894 (0.283)	0.1915 (0.284)	0.1918 (0.275)	0.1961 (0.302)	0.1964 (0.314)	0.2009 (0.271)	0.1896 (0.291)
No. atoms	4112	4112	4112	4112	4112	4112	4112	4112	4112	4112	4112	4112	4112	4112	4112	4112
Protein	3761	3761	3761	3761	3761	3761	3761	3761	3761	3761	3761	3761	3761	3761	3761	3761
Ligand/ion	49	49	49	49	49	49	49	49	49	49	49	49	49	49	49	49
Water	302	302	302	302	302	302	302	302	302	302	302	302	302	302	302	302
B-factors (overall)	32.38	29.94	25.99	29.07	25.6	29.83	27.65	31.1	31.22	30.1	26.24	27.27	34.46	29.5	33.66	28.48
Protein	31.29	28.7	24.73	27.82	24.33	28.83	26.37	29.95	29.83	28.18	24.95	25.98	33.16	28.18	32.42	27.2
Ligand/ion	70.48	67.61	62.39	66.3	61.49	68.17	64.27	68.56	66.17	63.9	62.31	72.59	65.21	67.53	65.19	65.19
Water	39.81	39.35	35.78	38.7	35.56	40.89	37.68	41.06	40.01	36.17	37.67	44.57	40.15	43.69	38.5	38.5
R.m.s. deviations																
Bond lengths (Å)	0.014	0.015	0.015	0.016	0.015	0.015	0.015	0.015	0.015	0.016	0.015	0.015	0.016	0.015	0.014	0.015
Bond angles (Å)	1.91	1.92	1.87	1.95	1.87	1.86	1.92	1.92	1.91	1.92	1.86	1.9	1.9	1.88	1.87	1.86

Apo structure number:	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
Data Collection																
Space group	P 21 21 21	P 21 21 21	P 21 21 21	P 21 21 21	P 21 21 21	P 21 21 21	P 21 21 21	P 21 21 21	P 21 21 21	P 21 21 21	P 21 21 21	P 21 21 21	P 21 21 21	P 21 21 21	P 21 21 21	P 21 21 21
Cell Dimensions																
a, b, c (Å)	59.603 / 75.47 / 124.474	59.003 / 75.74 / 124.425	58.996 / 75.692 / 124.318	60.604 / 73.763 / 124.111	58.94 / 75.819 / 124.474	59.289 / 75.407 / 124.203	59.237 / 75.397 / 124.302	59.139 / 75.405 / 124.157	59.23 / 75.716 / 124.206	59.079 / 75.41 / 124.063	59.007 / 75.656 / 124.195	58.901 / 75.438 / 124.392	59.205 / 75.481 / 124.16	59.079 / 75.802 / 124.2	58.799 / 75.583 / 124.015	59.603 / 75.47 / 124.474
α, β, γ (°)	90 90 90	90 90 90	90 90 90	90 90 90	90 90 90	90 90 90	90 90 90	90 90 90	90 90 90	90 90 90	90 90 90	90 90 90	90 90 90	90 90 90	90 90 90	90 90 90
Resolution (Å)	1.875 (1.875)	2.214 (2.214)	1.863 (1.863)	1.66 (1.66)	2.095 (2.095)	1.86 (1.86)	1.854 (1.854)	2.03 (2.03)	2.192 (2.192)	1.655 (1.655)	1.7 (1.7)	1.63 (1.63)	1.81 (1.81)	2.36 (2.36)	1.642 (1.642)	1.72 (1.72)
R_{merge}	0.4443 (1.54)	0.04455 (0.434)	0.4161 (1.624)	0.1762 (1.298)	0.05719 (0.514)	0.3689 (1.347)	0.2203 (1.441)	0.5299 (1.349)	0.543 (1.369)	0.4011 (1.391)	0.03976 (0.328)	0.3915 (1.277)	0.537 (1.053)	0.08399 (0.433)	0.3029 (1.297)	0.3574 (1.524)
R_{pim}	0.1252 (0.426)	0.04455 (0.434)	0.1174 (0.448)	0.04951 (0.354)	0.05719 (0.514)	0.1059 (0.378)	0.06204 (0.397)	0.1498 (0.386)	0.1533 (0.388)	0.1132 (0.382)	0.03976 (0.328)	0.1101 (0.357)	0.152 (0.303)	0.08399 (0.433)	0.08506 (0.359)	0.1029 (0.423)
I/sI	9.28 (0.87)	10.47 (1.47)	9.25 (0.78)	15.80 (0.89)	6.88 (1.20)	11.98 (0.88)	8.87 (0.75)	7.69 (0.77)	6.84 (0.81)	10.13 (0.81)	9.08 (1.44)	11.14 (0.87)	8.87 (0.81)	7.03 (1.55)	10.56 (0.84)	10.06 (0.78)
$CC_{1/2}$	0.903 (0.611)	0.998 (0.62)	0.925 (0.635)	0.995 (0.738)	0.997 (0.56)	0.955 (0.702)	0.994 (0.736)	0.917 (0.574)	0.919 (0.605)	0.947 (0.735)	0.998 (0.87)	0.945 (0.792)	0.931 (0.768)	0.989 (0.585)	0.967 (0.765)	0.95 (0.73)
Completeness (%)	99.08 (90.98)	98.76 (88.08)	99.27 (93.00)	99.74 (97.62)	99.31 (93.40)	99.43 (94.57)	99.33 (93.75)	99.95 (99.89)	98.35 (90.53)	99.46 (94.79)	96.34 (66.65)	99.63 (96.60)	99.77 (98.16)	99.64 (97.88)	99.43 (94.41)	98.44 (84.55)
Redundancy	13.6 (13.8)	1.9 (1.9)	13.6 (13.8)	13.7 (14.1)	1.9 (1.9)	13.5 (13.8)	13.6 (13.9)	13.6 (13.1)	13.8 (13.7)	13.7 (14.1)	1.9 (1.9)	13.7 (13.5)	13.6 (13.9)	1.9 (1.9)	13.7 (13.8)	13.7 (14.0)
Refinement																
Resolution (Å)	1.875 (1.875)	2.214 (2.214)	1.863 (1.863)	1.66 (1.66)	2.095 (2.095)	1.86 (1.86)	1.854 (1.854)	2.03 (2.03)	2.192 (2.192)	1.655 (1.655)	1.7 (1.7)	1.63 (1.63)	1.81 (1.81)	2.36 (2.36)	1.642 (1.642)	1.72 (1.72)
No. reflections	46445 (4224)	28067 (2467)	47023 (4329)	66849 (6392)	33209 (3083)	47365 (4437)	47699 (4427)	36725 (3630)	28733 (2611)	66975 (6261)	59476 (4059)	69813 (6696)	51168 (4950)	23487 (2260)	68442 (6386)	58534 (4974)

R_{work}	0.1734 (0.265)	0.1739 (0.259)	0.1728 (0.274)	0.1716 (0.278)	0.1814 (0.270)	0.1718 (0.288)	0.1728 (0.276)	0.1774 (0.253)	0.1867 (0.292)	0.1776 (0.279)	0.1865 (0.298)	0.1743 (0.247)	0.1730 (0.270)	0.1735 (0.285)	
R_{free}	0.1986 (0.260)	0.2076 (0.289)	0.1969 (0.260)	0.1884 (0.278)	0.2160 (0.275)	0.1945 (0.268)	0.1981 (0.289)	0.2033 (0.279)	0.2217 (0.276)	0.1924 (0.290)	0.2075 (0.306)	0.2066 (0.304)	0.2080 (0.278)	0.1964 (0.295)	
No. atoms	4112	4112	4112	4112	4112	4112	4112	4112	4112	4112	4112	4112	4112	4112	
Protein	3761	3761	3761	3761	3761	3761	3761	3761	3761	3761	3761	3761	3761	3761	
Ligand/ion	49	49	49	49	49	49	49	49	49	49	49	49	49	49	
Water	302	302	302	302	302	302	302	302	302	302	302	302	302	302	
B-factors (overall)	26.8	36.98	28.98	30.74	31.18	33.18	30.94	32.27	36	24.79	21.47	25.2	27.27	27.01	25.35	28.13
Protein	25.54	35.98	27.73	29.47	29.95	29.62	30.84	34.77	23.54	20.33	24.03	26.03	26.04	24.09	26.91
Ligand/ion	63.24	73.14	64.8	66.33	68.9	69.85	69.09	69.87	61.96	57.06	60.5	63.65	62.13	62.31	64.13
Water	36.54	43.52	38.63	40.77	40.38	43.21	41.57	44.16	45.74	34.32	29.82	34.08	36.85	33.43	35.08	37.5
R.m.s. deviations															
Bond lengths (Å)	0.015	0.014	0.015	0.015	0.014	0.015	0.014	0.014	0.014	0.015	0.015	0.015	0.014	0.015	0.015
Bond angles (°)	1.92	1.91	1.89	1.87	1.88	1.9	1.87	1.88	1.88	1.9	1.87	1.92	1.92	1.85	1.91	1.9
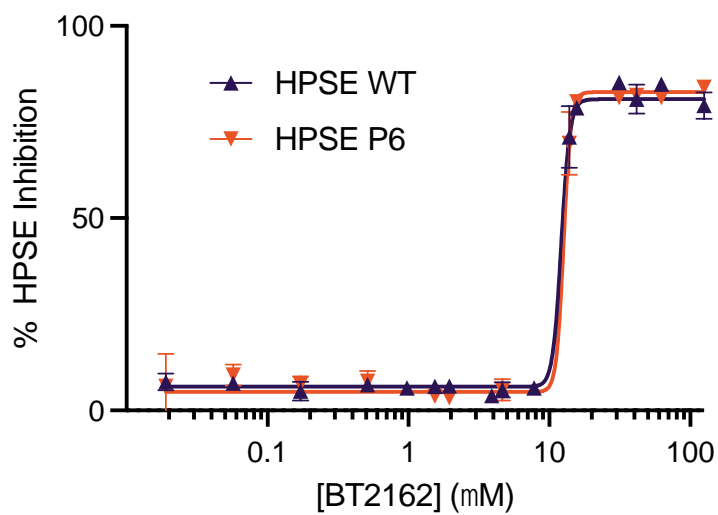
Figure S2: Dose-response curves with BT2162 tested against HPSE WT expressed in mammalian cells and HPSE P6 with fondaparinux as a substrate. Error bars represent standard error from a minimum of four measurements.
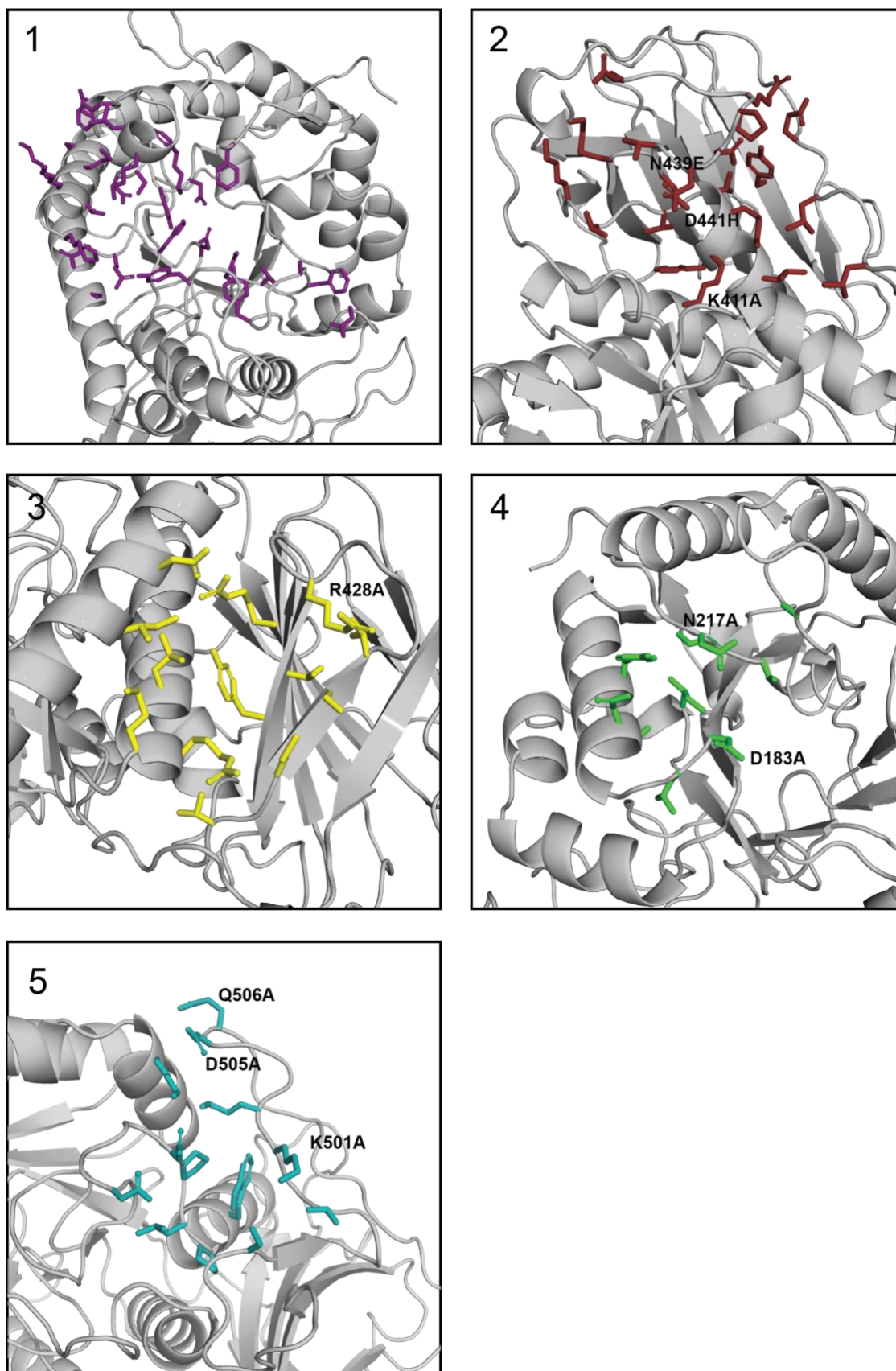
Figure S3 Predicted binding sites on the surface of heparanase from 1-5. Some of the mutated residues highlighted.

Table S3: Primers for HPSE P6 point mutations.

Name	Sequence
D40A_F	GAGCTCGCAGGATGTGGTCgcgCTTGATTTCTTCAC
D40A_R	CTGCGTGAAGAAATCAAGcgcGACCACATCCTGCG
H50A_F	CTTGATTTCTTCACGCAGGAGCCTTTGgcgCTTGTATCC
H50A_R	CTTAAAAATGAGGGGGATACAAGcgcCAAAGGCTCCTGCG
S55A_F	GCCTTTGCACCTTGTATCCCCCgcgTTTTTAAGCGTTAC
S55A_R	GTCTATGGTAACGCTTAAAAAcgcGGGGGATACAAGGTGC
L57A_F	CCTTGTATCCCCCTCATTTgcgAGCGTTACCATAGACGC
L57A_R	GTTTGCGTCTATGGTAACGCTcgcAAATGAGGGGGATAC
S58A_F	GCACCTTGTATCCCCCTCATTTTTAgcgGTTACCATAGACGCAAACC
S58A_R	GTGGCAAGGTTTGCGTCTATGGTAACcgcTAAAAATGAGGGGGATACAA
R70A_F	CAAACCTTGCCACTGACCCGgcgTTCTTAATCTTGCTTGG
R70A_R	CTACCAAGCAAGATTAAGAAcgcCGGGTCAGTGGCAAGG
R81A_F	TGCTTGGTAGTCCAAAGTTAgcgACGCTGGCGC
R81A_R	CTAAGCCCCCGCGCCAGCGTcgcTAACTTTG
Y91A_F	GGGCTTAGTCCAGCAgcgCTGCGCTTTG
Y91A_R	GCCAAAGCGCAGcgcTGCTGGACTAAGC
K108A_F	CTTCCTGATTTTTGATCCTAAAgcgGAATAAAAGCTTGCGGCCGC
K108A_R	GCGGCCGCAAGCTTTTATTCcgcTTTAGGATCAAAAATCAGG
S169A_F	CTCGACGTATAGCCGGTCTgcgGTGGATGTGCTCTATAC
S169A_R	GTATAGAGCACATCCACcgcAGACCGGCTATACGTCGAG
N203A_F	GGCAGTGGAATTCTAGCgcgGCTCAGCTCCTG
N203A_R	CGAGCAGGAGCTGAGCcgcGCTAGAATTCCACTG
N217A_F	CTGTGCCTCTAAAGGGTATgcgATCGACTGGGAGTTGGG
N217A_R	GCCCAACTCCCAGTCGATcgcATACCCTTTAGAGGC
H250A_F	GCAAGGACTTCATTCACCTTgcgAAACTGCTCCGGAAATCGAC
H250A_R	GTCGATTTCCGGAGCAGTTTcgcAAGGTGAATGAAGTCCTTGC
R254A_F	TTCACCTTCACAAACTGCTCGCGAAATCGACATTTAAGAATGC
R254A_R	GCATTCTTAAATGTCGATTTCGCGAGCAGTTTGTGAAGGTGAATG
D291A_F	AGGCGGGCGGCGAAGTCATTgcgGCAGTAACATGGCACCATTAC
D291A_R	TAATGGTGCCATGTTACTGCcgcAATGACTTCGCCGCCCGCCTTC
K338A_F	GAATCGACCCGGCCTGGGAAGgcgGTTTG
K338A_R	CTTGTCTCCCCGAGCCAAACcgcCTTCCCAG
E378A_F	CCGCTCGCATGGGGATCgcgGTCGTGATG
E378A_R	CTTGGCGCATCACGACcgcGATCCCCATG
D395A_F	CTGGCAACTACCACCTCGTCgcgGAAAACTTCGATCCATTG
D395A_R	CAATGGATCGAAGTTTTCcgcGACGAGGTGGTAGTTGCCAG
E396A_F	GCAACTACCACCTCGTCGACgcgAACTTCGATCCATTG
E396A_R	CAGGCAATGGATCGAAGTTcgcGTCGACGAGGTGGTAG
F398A_F	CACCTCGTCGACGAAAACgcgGATCCATTGCCTG

F398A_R	CAGTAGTCAGGCAATGGATCcgcGTTTTCGTCGACGAG
N442A_F	CCATTGCACAAATACGGATgcgCCTCGCTACAAAGAAGGC
N442A_R	CGCCTTCTTTGTAGCGAGGcgcATCCGTATTTGTGCAATG
R465A_F	CAACCTCCATAACGTCACCAAGTATCTCgcgCTGCCATATCCTTTTAG
R465A_R	CCACCTGTTTATTACTAAAAGGATATGGCAGcgcGAGATACTTGGTGACG
K501A_F	GTTGAATGGCCAGACCCTCgcgATGGTTGATGACCAG
K501A_R	GTCTGGTCATCAACCATcgcGAGGGTCTGGCCATTC
D505A_Q506A_F	CCTCAAGATGGTTGATgcggcgACTTTGCCTCCTTTG
D505A_Q506A_R	CTTCAAAGGAGGCAAAGTcgccgcATCAACCATCTTGAG
G76K_F	CACTGACCCGCGCTTCTTAATCTTGCTTaaaAGTCCAAAGTTAAGAACGC
G76K_R	GCGCCAGCGTTCTTAACTTTGGACTtttAAGCAAGATTAAGAAGCGC
S213K_F	CTCAGCTCCTGCTCGATTACTGTGCCaaaAAAGGGTATAACATCGACTG
S213K_R	CAACTCCCAGTCGATGTTATACCCTTTtttGGCACAGTAATCGAGCAGG
G336K_F	GTCGAATCGACCCGGCCTaaaAAGAAAGTTTGGCTCGGG
G336K_R	CCCGAGCCAAACTTTCTTtttAGGCCGGGTCGATTCGAC
Mid_Duet_F	CCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGC
Mid_Duet_R	GCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGG
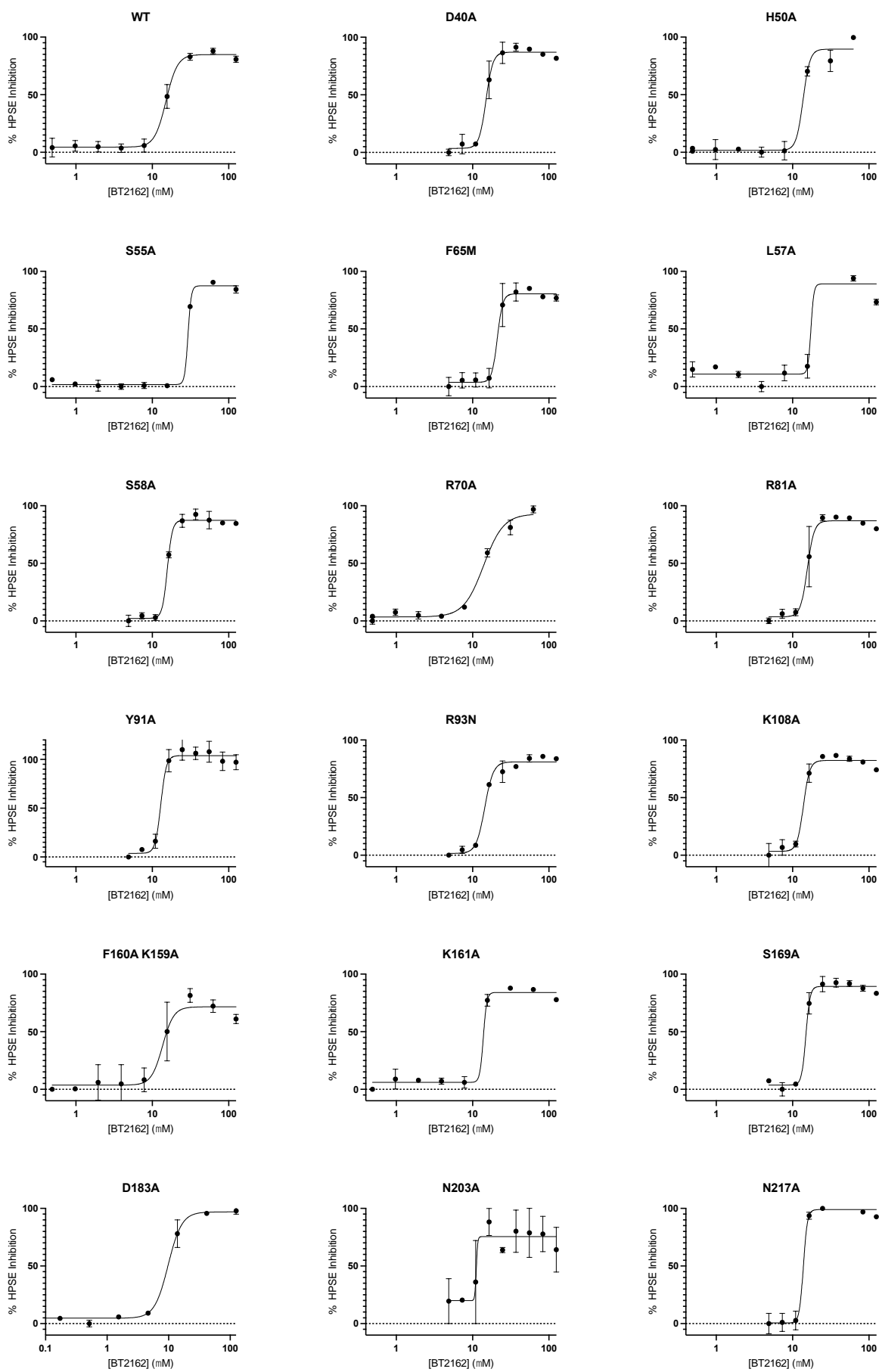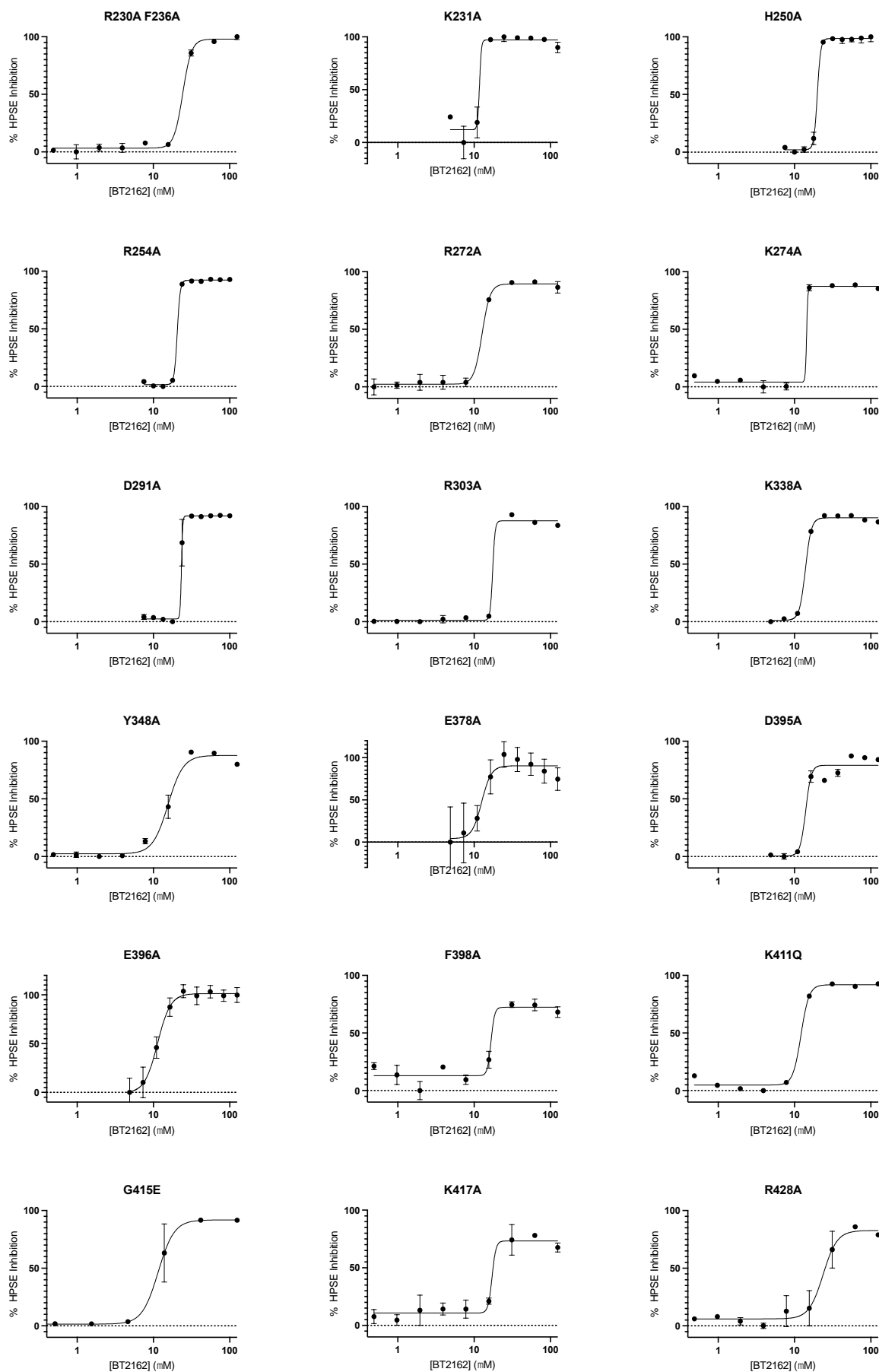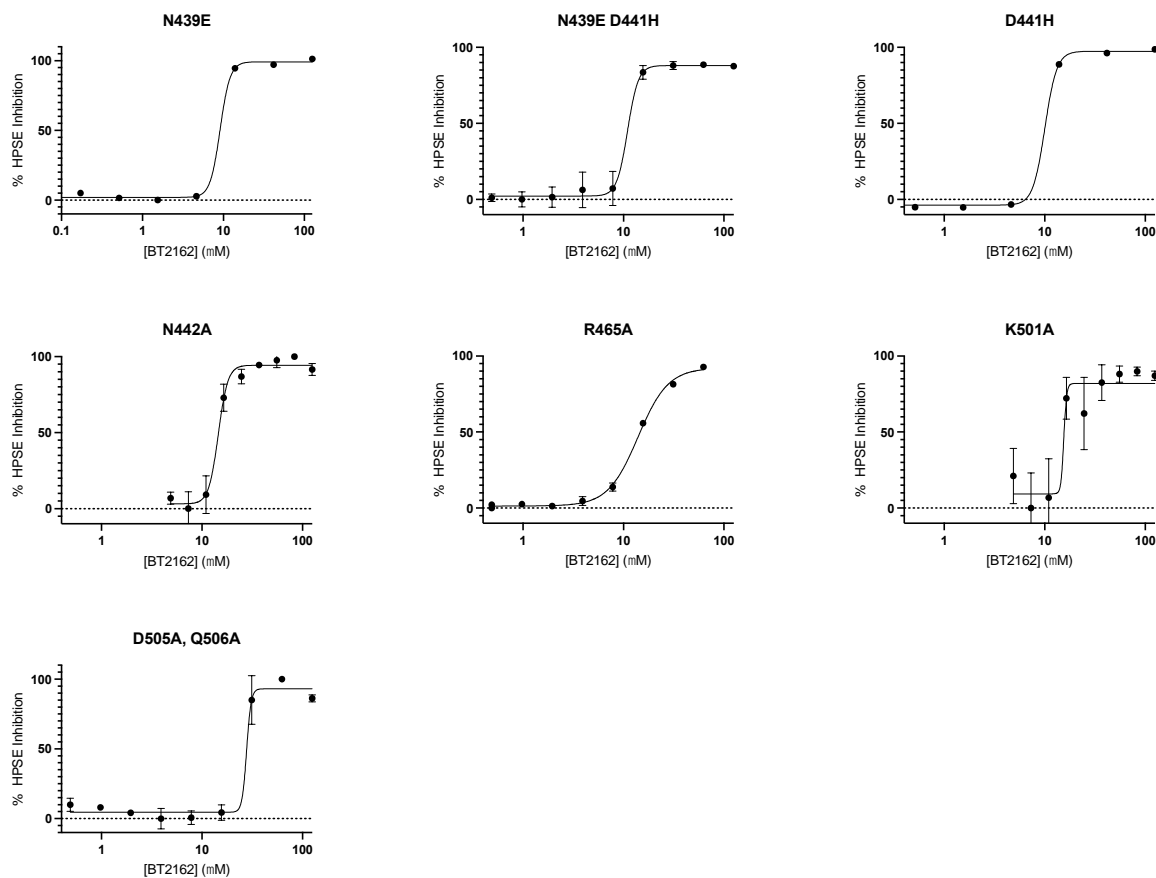
Figure S4: Dose-response curves of HPSE P6 point mutants against BT2162 inhibition. Curves based off 2-4 replicates. Error represented as standard deviation.

Table S3: Scores for each HPSE site identified from Sitemap, Schrodinger.

Site	Site score	Druggability score
Site 1 (substrate binding site)	1.02	0.89
Site 2	0.97	0.97
Site 3	0.93	0.84
Site 4	0.81	0.80
Site 5	0.77	0.63

Chapter 5

Conclusions, Implications and Future research

The work presented in this thesis has focused on using biophysical and kinetic techniques to gain a better understanding of the diverse binding mechanisms of heparanase inhibitors. Through the use of protein design techniques, we were able to generate a variant of heparanase that exhibited wild type activity and inhibition and allowed us to conduct structural and biophysical analysis of heparanase inhibitors, which had been previously been limited by the challenges involved in producing this protein recombinantly. This allowed us to understand the binding mechanisms of various classes of heparanase inhibitors, providing insight into their complex binding mechanisms. The results of this work have the potential to inform the design of improved drug candidates in the future. Additionally, this thesis highlights the complexity of inhibitor development and the various approaches that can be employed to understand these inhibitors, including the role of multi-compound binding in complex inhibition mechanisms (Chapter 3) and the potential for new allosteric inhibitors with their own benefits (Chapter 4). Our findings have already been utilized within our research group and more broadly, with the heparanase mutant HPSE P6 (Chapter 2) being shared with multiple groups for use in drug development and research on heparan sulfate and heparan sulfate proteoglycans.

5.1 Key Findings

Designing a stable heparanase variant for drug development purposes: In Chapter 2, I presented the engineering and experimental characterization of a stable heparanase variant. Previous research into bacterial heterologous expression of recombinant heparanase had used the wild-type sequence to express and purify the two subunits of heparanase, resulting in low yields that were not reproducible. We instead mutated the large 50 kDa subunit to increase solubility. This computationally designed heparanase mutant could be expressed in soluble form in *E. coli* and was found to be structurally and kinetically similar to wild-type heparanase, despite the introduction of 26 mutations to the large subunit of the protein. This protein can now be expressed in *E. coli* in a soluble, folded and active state in relatively large quantities, whereas previously it was only obtainable from expensive expression systems using mammalian or insect cell lines. Through protein X-ray crystallography and molecular dynamics, we demonstrated that the 26 changes to heparanase do not alter the protein structure, but do slightly contribute to the rigidification of surface loops, which may affect stability. In summary, the kinetic, structural and computational characterization of this new protein variant provides a reliable starting point for studying protein-ligand interactions for heparanase therapeutic development.

Structural insights of heparanase inhibition by Pentosan polysulfate In Chapter 3, I presented work that provided a detailed understanding of how a heparan sulfate mimetic, PPS, binds to heparanase. This work showed that the binding mechanism was not the expected competitive inhibition model that has been supported for decades, but a complex mechanism influenced by various changes to the protein structure. We observed changes in maximal inhibition with increased oligosaccharide length, which is unexpected for competitive inhibitors, leading to changes in the kinetic model with increased length, where PPS showed a parabolic binding mechanism. Isothermal titration calorimetry, protein X-ray crystallography and hydrogen-deuterium exchange mass spectrometry supported the presence of multiple binding sites, with PPS binding at each of the three heparan binding domains. This binding mechanism contributes to the ability of these inhibitors to induce protein oligomerization, crosslinking multiple proteins and causing macroaggregation and irreversible inhibition. We also showed in this paper that this inhibition mechanism occurs with other sulfated oligosaccharides that are currently in clinical trials.

Development of new allosteric inhibitors of heparanase In Chapter 4, I presented a manuscript that explores the complex structure-activity relationship (SAR) of a new class of small molecule heparanase inhibitors - quinazoline analogues. Through various assays and binding studies, we demonstrated that the binding model for the lead compound was non-competitive with an IC_{50} in the low μM range. We confirmed the binding affinity through isothermal titration calorimetry (ITC), which was consistent with the IC_{50}, supporting the non-competitive binding mechanism and allosteric inhibition. We also demonstrated that this inhibitor had specific activity towards heparanase and was not a non-specific aggregator. While we were unable to obtain definitive structural information about the binding site of these small molecules, we conducted a mutagenesis screen that revealed complex allosteric networks that affect catalytic activity.

5.2   Future Directions

While significant progress has been made in understanding the binding of inhibitors to heparanase, there is still much work to be done in the development of effective therapeutics for commercial use. The HPSE P6 variant meets the requirements for the compounds examined in this study, but it is uncertain how it will interact with other inhibitor classes binding at potential new allosteric sites. Therefore, the HPSE P6 mutation sites should be taken into consideration when exploring new allosteric inhibitor classes. Many researchers in the field of glycobiology also use heparanase as

a means to process heparan sulfate and there is considerable potential in this variant in this context. The effects of these mutations were only evaluated in relation to the specific inhibitor series investigated in this study. Further research is needed to determine if these mutations have any broader effects on specificity and to determine the potential utility of this protein in other research areas. It is currently unknown the effect of each mutation in the HPSE-P6 design and further studies could be undertaken to understand the effect of these mutations in expression and if a variant can be designed with reduced changes in sequence, reducing the effect of altering possible binding sites.

The research presented in this thesis represents a significant advance in understanding the binding mechanism of PPS and other sulfated oligosaccharide drug candidates. In Chapter 3, we found that the secondary structure shift of heparanase upon inhibitor binding is not reversible, yet in some cases, the enzyme can still maintain catalytic activity. Further work is needed to determine the nature of this structural shift and its role in enzyme activity, as it may be key to the development of additional HS mimetic inhibitors. It is also important to understand the effect of secondary structure change and protein aggregation induced by these inhibitors *in vivo*, as these effects may help to explain the limited benefits observed in clinical trials for cancer treatment. PPS and other sulfated oligosaccharide inhibitors are structurally similar to HS, but they interact with heparanase differently as no denaturation is observed with the native substrate. This could be as simple as HS being rapidly hydrolyzed before the "wrapping" of the enzyme can affect the structure. Understanding the reasons for this difference is essential for the development of competitive glycomimetic inhibitors that do not irreversibly denature their target.

Despite comparatively less interest in the development of synthetic small molecules for heparanase inhibition, these inhibitors have many advantages. Small molecule inhibitors can overcome many of the limitations associated with sulfated oligosaccharides, such as promiscuous binding and membrane permeability. The main challenge for this class of inhibitors is to understand their binding mechanism to heparanase. Some may bind at the active site as competitive inhibitors, while others may bind at potential allosteric sites, such as the compounds discussed in Chapter 5. Identifying and understanding the binding site, as well as any allostery at play, will be an important future step in the development of more targeted heparanase inhibitors. Now that it is easier to undertake structural and biophysical analysis on HPSE, Understanding the mechanisms of previously studied highly active compounds, such as OGT 2115 will aid in the further development of HPSE inhibitors.

Heparanase is a complex enzyme with various functions and locations in the body. Drug development efforts for heparanase have investigated a range of inhibitor classes in an effort to effectively target heparanase while addressing the limitations of other inhibitor classes. One class of inhibitors that has yet to be explored is peptide therapeutics, which is an area of growing interest with promising results. Given the intricate binding mechanisms of HS mimetics and the challenges of studying small molecule inhibitors, it is worth considering the potential utility of this class of inhibitors in further research on heparanase.

A significant challenge in the field of heparanase therapeutic design is the lack of a high-throughput assay for evaluating dose-response curves and kinetics. The assay that is most widely used, which was also utilized in this thesis, only allows for activity endpoint measurements and was found to interact with many of the compounds designed for heparanase inhibition. This complexity and the labor-intensive nature of the assay has also hindered kinetic studies, which are important for understanding the various modes of inhibition exhibited by heparanase inhibitors. The development and availability of a more suitable assay should be a priority, as the lack of an ideal assay has contributed to the limited understanding of many heparanase inhibitors published to date. Work going into this field currently has identified FRET and fluorescence based assays as alternative assay methods, but limitations on substrate homogeneity and affinity are issues that need to be overcome with these methods.

5.3 Concluding Remarks

Like many previous papers and theses, the work presented here highlights the complexities and ongoing gaps in our understanding of heparanase inhibition. The development of the heparanase variant HPSE P6 has allowed for greater insight into these inhibitor types, but there are still many challenges to overcome. Large sulfated oligosaccharides, once thought to be competitive inhibitors due to their structural similarity to HS, have been shown to exhibit highly complex binding mechanisms that may pose difficulties for use in vivo. Small molecule inhibitors may offer benefits such as easily manipulatable solubility and activity, but their design can be challenging if the binding location is unknown. This thesis contributes to an increase in our understanding of heparanase inhibition mechanisms and it is hoped that this progress will continue in the future.

Chapter 6

References

(1) *Essentials of Glycobiology*, 3rd; Varki, A., Cummings, R. D., Esko, J. D., Stanley, P., Hart, G. W., Aebi, M., Darvill, A. G., Kinoshita, T., Packer, N. H., Prestegard, J. H., Schnaar, R. L. and Seeberger, P. H., Eds.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor (NY), 2015.

(2) Iozzo, R. V. and Schaefer, L. (2015). Proteoglycan form and function: A comprehensive nomenclature of proteoglycans. *Matrix biology : journal of the International Society for Matrix Biology 42*, 11–55.

(3) Parish, C. R. (2006). The role of heparan sulphate in inflammation. *Nature Reviews Immunology 6*, 633–643.

(4) Shriver, Z., Capila, I., Venkataraman, G. and Sasisekharan, R. (2012). Heparin and Heparan Sulfate: Analyzing Structure and Microheterogeneity. *Handbook of experimental pharmacology*, 159–176.

(5) Sarrazin, S., Lamanna, W. C. and Esko, J. D. (2011). Heparan sulfate proteoglycans. *Cold Spring Harb Perspect Biol 3*, DOI: 10.1101/cshperspect.a004952.

(6) Sugahara, K. and Kitagawa, H. (2002). Heparin and Heparan Sulfate Biosynthesis. *IUBMB Life 54*, 163–175.

(7) Kusche-Gullberg, M. and Kjellén, L. (2003). Sulfotransferases in glycosaminoglycan biosynthesis. *Current Opinion in Structural Biology 13*, 605–611.

(8) Lyon, M. and Gallagher, J. T. (1998). Bio-specific sequences and domains in heparan sulphate and the regulation of cell growth and adhesion. *Matrix Biology 17*, 485–493.

(*9*) Ilan, N., Elkin, M. and Vlodavsky, I. (2006). Regulation, function and clinical significance of heparanase in cancer metastasis and angiogenesis. *The International Journal of Biochemistry & Cell Biology 38*, 2018–2039.

(*10*) Bame, K. J. (2001). Heparanases: endoglycosidases that degrade heparan sulfate proteoglycans. *Glycobiology 11*, 91R–98R.

(*11*) Vreys, V. and David, G. (2007). Mammalian heparanase: what is the message? *Journal of Cellular and Molecular Medicine 11*, 427–452.

(*12*) Goldshmidt, O., Nadav, L., Aingorn, H., Irit, C., Feinstein, N., Ilan, N., Zamir, E., Geiger, B., Vlodavsky, I. and Katz, B.-Z. (2002). Human Heparanase Is Localized within Lysosomes in a Stable Form. *Experimental Cell Research 281*, 50–62.

(*13*) Fairbanks, M. B., Mildner, A. M., Leone, J. W., Cavey, G. S., Mathews, W. R., Drong, R. F., Slightom, J. L., Bienkowski, M. J., Smith, C. W., Bannow, C. A. and Heinrikson, R. L. (1999). Processing of the Human Heparanase Precursor and Evidence That the Active Enzyme Is a Heterodimer*. *Journal of Biological Chemistry 274*, 29587–29590.

(*14*) Nardella, C., Lahm, A., Pallaoro, M., Brunetti, M., Vannini, A. and Steinkühler, C. (2004). Mechanism of Activation of Human Heparanase Investigated by Protein Engineering. *Biochemistry 43*, Publisher: American Chemical Society, 1862–1873.

(*15*) Fux, L., Feibish, N., Cohen-Kaplan, V., Gingis-Velitski, S., Feld, S., Geffen, C., Vlodavsky, I. and Ilan, N. (2009). Structure-Function Approach Identifies a COOH-Terminal Domain That Mediates Heparanase Signaling. *Cancer Research 69*, 1758–1767.

(*16*) Levy-Adam, F., Abboud-Jarrous, G., Guerrini, M., Beccati, D., Vlodavsky, I. and Ilan, N. (2005). Identification and characterization of heparin/heparan sulfate binding domains of the endoglycosidase heparanase. *The Journal of biological chemistry 280*, 20457–66.

(*17*) Crispel, Y., Ghanem, S., Attias, J., Kogan, I., Brenner, B. and Nadir, Y. (2017). Involvement of the heparanase procoagulant domain in bleeding and wound healing. *Journal of Thrombosis and Haemostasis 15*, 1463–1472.

(*18*) Wu, L., Viola, C. M., Brzozowski, A. M. and Davies, G. J. (2015). Structural characterization of human heparanase reveals insights into substrate recognition. *Nat Struct Mol Biol 22*, 1016–1022.

(19) Simizu, S., Ishida, K., Wierzba, M. K. and Osada, H. (2004). Secretion of Heparanase Protein Is Regulated by Glycosylation in Human Tumor Cell Lines*. *Journal of Biological Chemistry 279*, 2697–2703.

(20) Hulett, M. D., Hornby, J. R., Ohms, S. J., Zuegg, J., Freeman, C., Gready, J. E. and Parish, C. R. (2000). Identification of active-site residues of the pro-metastatic endoglycosidase heparanase. *Biochemistry 39*, 15659–15667.

(21) Gingis-Velitski, S., Zetser, A., Kaplan, V., Ben-Zaken, O., Cohen, E., Levy-Adam, F., Bashenko, Y., Flugelman, M. Y., Vlodavsky, I. and Ilan, N. (2004). Heparanase Uptake Is Mediated by Cell Membrane Heparan Sulfate Proteoglycans*. *Journal of Biological Chemistry 279*, 44084–44092.

(22) Goldberg, R., Meirovitz, A., Hirshoren, N., Bulvik, R., Binder, A., Rubinstein, A. M. and Elkin, M. (2013). Versatile role of heparanase in inflammation. *Matrix biology : journal of the International Society for Matrix Biology 32*, 234–240.

(23) Peterson, S. and Liu, J. (2012). Deciphering Mode of Action of Heparanase Using Structurally Defined Oligosaccharides*. *Journal of Biological Chemistry 287*, 34836–34843.

(24) Pikas, D. S., Li, J.-p., Vlodavsky, I. and Lindahl, U. (1998). Substrate Specificity of Heparanases from Human Hepatoma and Platelets*. *Journal of Biological Chemistry 273*, 18770–18777.

(25) Speciale, G., Thompson, A. J., Davies, G. J. and Williams, S. J. (2014). Dissecting conformational contributions to glycosidase catalysis and inhibition. *Current Opinion in Structural Biology 28*, 1–13.

(26) De Pasquale, V., Quiccione, M. S., Tafuri, S., Avallone, L. and Pavone, L. M. (2021). Heparan Sulfate Proteoglycans in Viral Infection and Treatment: A Special Focus on SARS-CoV-2. *International Journal of Molecular Sciences 22*, 6574.

(27) Wasteson, A., Höök, M. and Westermark, B. (1976). Demonstration of a platelet enzyme, degrading heparan sulphate. *FEBS letters 64*, 218–221.

(28) Benhamron, S., Nechushtan, H., Verbovetski, I., Krispin, A., Abboud-Jarrous, G., Zcharia, E., Edovitsky, E., Nahari, E., Peretz, T., Vlodavsky, I. and Mevorach, D. (2006). Translocation of Active Heparanase to Cell Surface Regulates Degradation of Extracellular Matrix Heparan Sulfate upon Transmigration of Mature Monocyte-Derived Dendritic Cells1. *The Journal of Immunology 176*, 6417–6424.

(*29*) Gutter-Kapon, L., Alishekevitz, D., Shaked, Y., Li, J. P., Aronheim, A., Ilan, N. and Vlodavsky, I. (2016). Heparanase is required for activation and function of macrophages. *Proc Natl Acad Sci U S A 113*, E7808–E7817.

(*30*) Savion, N., Vlodavsky, I. and Fuks, Z. (1984). Interaction of T lymphocytes and macrophages with cultured vascular endothelial cells: Attachment, invasion, and subsequent degradation of the subendothelial extracellular matrix. *Journal of Cellular Physiology 118*, 169–178.

(*31*) Matzner, Y., Bar-Ner, M., Yahalom, J., Ishai-Michaeli, R., Fuks, Z. and Vlodavsky, I. (1985). Degradation of heparan sulfate in the subendothelial extracellular matrix by a readily released heparanase from human neutrophils. Possible role in invasion through basement membranes. *The Journal of clinical investigation 76*, 1306–13.

(*32*) Bashkin, P., Razin, E., Eldor, A. and Vlodavsky, I. (1990). Degranulating Mast Cells Secrete an Endoglycosidase That Degrades Heparan Sulfate in Subendothelial Extracellular Matrix. *Blood 75*, 2204–2212.

(*33*) Laskov, R., Michaeli, R.-I., Sharir, H., Yefenof, E. and Vlodavsky, I. (1991). Production of heparanase by normal and neoplastic murine B-lymphocytes. *International Journal of Cancer 47*, 92–98.

(*34*) Temkin, V., Aingorn, H., Puxeddu, I., Goldshmidt, O., Zcharia, E., Gleich, G. J., Vlodavsky, I. and Levi-Schaffer, F. (2004). Eosinophil major basic protein: first identified natural heparanase-inhibiting protein. *Journal of Allergy and Clinical Immunology 113*, 703–709.

(*35*) Vlodavsky, I., Eldor, A., Haimovitz-Friedman, A., Matzner, Y., Ishai-Michaeli, R., Lider, O., Naparstek, Y., Cohen, I. R. and Fuks, Z. (1992). Expression of heparanase by platelets and circulating cells of the immune system: possible involvement in diapedesis and extravasation. *Invasion & Metastasis 12*, 112–127.

(*36*) Hulett, M. D., Freeman, C., Hamdorf, B. J., Baker, R. T., Harris, M. J. and Parish, C. R. (1999). Cloning of mammalian heparanase, an important enzyme in tumor invasion and metastasis. *Nature medicine 5*, 803.

(*37*) Escobar Galvis, M. L., Jia, J., Zhang, X., Jastrebova, N., Spillmann, D., Gottfridsson, E., van Kuppevelt, T. H., Zcharia, E., Vlodavsky, I., Lindahl, U. and Li, J.-P. (2007). Transgenic or tumor-induced expression of heparanase upregulates sulfation of heparan sulfate. *Nature Chemical Biology 3*, 773–778.

(38) Shteingauz, A., Boyango, I., Naroditsky, I., Hammond, E., Gruber, M., Doweck, I., Ilan, N. and Vlodavsky, I. (2015). Heparanase Enhances Tumor Growth and Chemoresistance by Promoting Autophagy. *Cancer Research 75*, 3946–3957.

(39) Roucourt, B., Meeussen, S., Bao, J., Zimmermann, P. and David, G. (2015). Heparanase activates the syndecan-syntenin-ALIX exosome pathway. *Cell Research 25*, 412–428.

(40) Thompson, C. A., Purushothaman, A., Ramani, V. C., Vlodavsky, I. and Sanderson, R. D. (2013). Heparanase regulates secretion, composition, and function of tumor cell-derived exosomes. *The Journal of Biological Chemistry 288*, 10093–10099.

(41) Schubert, S. Y., Ilan, N., Shushy, M., Ben-Izhak, O., Vlodavsky, I. and Goldshmidt, O. (2004). Human heparanase nuclear localization and enzymatic activity. *Laboratory Investigation; a Journal of Technical Methods and Pathology 84*, 535–544.

(42) Masola, V., Bellin, G., Gambaro, G. and Onisto, M. (2018). Heparanase: A Multitasking Protein Involved in Extracellular Matrix (ECM) Remodeling and Intracellular Events. *Cells 7*, DOI: `10.3390/cells7120236`.

(43) Putz, E. M., Mayfosh, A. J., Kos, K., Barkauskas, D. S., Nakamura, K., Town, L., Goodall, K. J., Yee, D. Y., Poon, I. K., Baschuk, N., Souza-Fonseca-Guimaraes, F., Hulett, M. D. and Smyth, M. J. (2017). NK cell heparanase controls tumor invasion and immune surveillance. *The Journal of Clinical Investigation 127*, 2777.

(44) Digre, A., Singh, K., Åbrink, M., Reijmers, R. M., Sandler, S., Vlodavsky, I. and Li, J.-P. (2017). Overexpression of heparanase enhances T lymphocyte activities and intensifies the inflammatory response in a model of murine rheumatoid arthritis. *Scientific Reports 7*, 46229.

(45) He, Y. Q., Sutcliffe, E. L., Bunting, K. L., Li, J., Goodall, K. J., Poon, I. K., Hulett, M. D., Freeman, C., Zafar, A., McInnes, R. L., Taya, T., Parish, C. R. and Rao, S. (2012). The endoglycosidase heparanase enters the nucleus of T lymphocytes and modulates H3 methylation at actively transcribed genes via the interplay with key chromatin modifying enzymes. *Transcription 3*, 130–145.

(46) Mayfosh, A. J., Goodall, K. J., Nguyen, T., Baschuk, N. and Hulett, M. D. (2022). Heparanase is a regulator of natural killer cell activation and cytotoxicity. *Journal of Leukocyte Biology 111*, 1211–1224.

(*47*) Schmidt, E. P. et al. (2012). The pulmonary endothelial glycocalyx regulates neutrophil adhesion and lung injury during experimental sepsis. *Nature medicine 18*, 10.1038/nm.2843.

(*48*) Shafat, I., Vlodavsky, I. and Ilan, N. (2006). Characterization of Mechanisms Involved in Secretion of Active Heparanase*. *Journal of Biological Chemistry 281*, 23804–23811.

(*49*) Reiland, J., Sanderson, R. D., Waguespack, M., Barker, S. A., Long, R., Carson, D. D. and Marchetti, D. (2004). Heparanase Degrades Syndecan-1 and Perlecan Heparan Sulfate: FUNCTIONAL IMPLICATIONS FOR TUMOR CELL INVASION*. *Journal of Biological Chemistry 279*, 8047–8055.

(*50*) Fux, L., Ilan, N., Sanderson, R. D. and Vlodavsky, I. (2009). Heparanase: busy at the cell surface. *Trends in Biochemical Sciences 34*, 511–519.

(*51*) Mohan, C. D., Hari, S., Preetham, H. D., Rangappa, S., Barash, U., Ilan, N., Nayak, S. C., Gupta, V. K., Basappa, Vlodavsky, I. and Rangappa, K. S. Targeting Heparanase in Cancer: Inhibition by Synthetic, Chemically Modified, and Natural Compounds, 2019.

(*52*) Jayatilleke, K. M. and Hulett, M. D. (2020). Heparanase and the hallmarks of cancer. *Journal of Translational Medicine 18*, 453.

(*53*) Arvatz, G., Weissmann, M., Ilan, N. and Vlodavsky, I. (2016). Heparanase and cancer progression: New directions, new promises. *Human Vaccines & Immunotherapeutics 12*, 2253–2256.

(*54*) Masola, V., Maran, C., Tassone, E., Zin, A., Rosolen, A. and Onisto, M. (2009). Heparanase activity in alveolar and embryonal rhabdomyosarcoma: implications for tumor invasion. *BMC Cancer 9*, 304.

(*55*) Zhang, W., Chan, H., Wei, L., Pan, Z., Zhang, J. and Li, L. (2013). Overexpression of heparanase in ovarian cancer and its clinical significance. *Oncology Reports 30*, 2279–2287.

(*56*) Zhang, Y., Li, L., Wang, Y., Zhang, J., Wei, G., Sun, Y. and Shen, F. (2007). Downregulating the expression of heparanase inhibits the invasion, angiogenesis and metastasis of human hepatocellular carcinoma. *Biochemical and Biophysical Research Communications 358*, 124–129.

(*57*) Lerner, I., Baraz, L., Pikarsky, E., Meirovitz, A., Edovitsky, E., Peretz, T., Vlodavsky, I. and Elkin, M. (2008). Function of heparanase in prostate tumorigenesis: potential for therapy. *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research 14*, 668–676.

(58) Gomes, A. M., Stelling, M. P. and Pavão, M. S. G. (2013). Heparan sulfate and heparanase as modulators of breast cancer progression. *BioMed Research International 2013*, 852093.

(59) Liu, X., Zhou, Z.-h., Li, W., Zhang, S.-k., Li, J., Zhou, M.-J. and Song, J.-W. (2019). Heparanase Promotes Tumor Growth and Liver Metastasis of Colorectal Cancer Cells by Activating the p38/MMP1 Axis. *Frontiers in Oncology 9*.

(60) Shafat, I., Pode, D., Peretz, T., Ilan, N., Vlodavsky, I. and Nisman, B. (2008). Clinical significance of urine heparanase in bladder cancer progression. *Neoplasia (New York, N.Y.) 10*, 125–130.

(61) Kundu, S., Xiong, A., Spyrou, A., Wicher, G., Marinescu, V. D., Edqvist, P.-H. D., Zhang, L., Essand, M., Dimberg, A., Smits, A., Ilan, N., Vlodavsky, I., Li, J.-P. and Forsberg-Nilsson, K. (2016). Heparanase Promotes Glioma Progression and Is Inversely Correlated with Patient Survival. *Molecular cancer research: MCR 14*, 1243–1253.

(62) Koliopanos, A., Friess, H., Kleeff, J., Shi, X., Liao, Q., Pecker, I., Vlodavsky, I., Zimmermann, A. and Büchler, M. W. (2001). Heparanase expression in primary and metastatic pancreatic cancer. *Cancer Research 61*, 4655–4659.

(63) Purushothaman, A. and Sanderson, R. D. (2020). Heparanase: A Dynamic Promoter of Myeloma Progression. *Advances in experimental medicine and biology 1221*, 331–349.

(64) Nobuhisa, T., Naomoto, Y., Ohkawa, T., Takaoka, M., Ono, R., Murata, T., Gunduz, M., Shirakawa, Y., Yamatsuji, T., Haisa, M., Matsuoka, J., Tsujigiwa, H., Nagatsuka, H., Nakajima, M. and Tanaka, N. (2005). Heparanase expression correlates with malignant potential in human colon cancer. *Journal of Cancer Research and Clinical Oncology 131*, 229–237.

(65) Hammond, E., Khurana, A., Shridhar, V. and Dredge, K. (2014). The Role of Heparanase and Sulfatases in the Modification of Heparan Sulfate Proteoglycans within the Tumor Microenvironment and Opportunities for Novel Cancer Therapeutics. *Frontiers in Oncology 4*, 195.

(66) Edovitsky, E., Elkin, M., Zcharia, E., Peretz, T. and Vlodavsky, I. (2004). Heparanase gene silencing, tumor invasiveness, angiogenesis, and metastasis. *Journal of the National Cancer Institute 96*, 1219–1230.

(67) Zong, F., Fthenou, E., Wolmer, N., Hollósi, P., Kovalszky, I., Szilák, L., Mogler, C., Nilsonne, G., Tzanakakis, G. and Dobra, K. (2009). Syndecan-1 and FGF-2, but Not FGF Receptor-1, Share a Common Transport Route and Co-Localize

with Heparanase in the Nuclei of Mesenchymal Tumor Cells. *PLOS ONE 4*, e7346.

(*68*) Doweck, I., Kaplan-Cohen, V., Naroditsky, I., Sabo, E., Ilan, N. and Vlodavsky, I. (2006). Heparanase Localization and Expression by Head and Neck Cancer: Correlation with Tumor Progression and Patient Survival. *Neoplasia 8*, 1055–1061.

(*69*) Yang, Y., Gorzelanny, C., Bauer, A. T., Halter, N., Komljenovic, D., Bäuerle, T., Borsig, L., Roblek, M. and Schneider, S. W. (2015). Nuclear heparanase-1 activity suppresses melanoma progression via its DNA-binding affinity. *Oncogene 34*, 5832–5842.

(*70*) Amin, R., Tripathi, K. and Sanderson, R. D. (2020). Nuclear Heparanase Regulates Chromatin Remodeling, Gene Expression and PTEN Tumor Suppressor Function. *Cells 9*, 2038.

(*71*) Bandari, S. K., Purushothaman, A., Ramani, V. C., Brinkley, G. J., Chandrashekar, D. S., Varambally, S., Mobley, J. A., Zhang, Y., Brown, E. E., Vlodavsky, I. and Sanderson, R. D. (2018). Chemotherapy induces secretion of exosomes loaded with heparanase that degrades extracellular matrix and impacts tumor and host cell behavior. *Matrix Biology 65*, 104–118.

(*72*) Bhattacharya, U., Gutter-Kapon, L., Kan, T., Boyango, I., Barash, U., Yang, S.-M., Liu, J., Gross-Cohen, M., Sanderson, R. D., Shaked, Y., Ilan, N. and Vlodavsky, I. (2020). Heparanase and chemotherapy synergize to drive macrophage activation and enhance tumor growth. *Cancer research 80*, 57–68.

(*73*) McKenzie, E., Tyson, K., Stamps, A., Smith, P., Turner, P., Barry, R., Hircock, M., Patel, S., Barry, E., Stubberfield, C., Terrett, J. and Page, M. (2000). Cloning and Expression Profiling of Hpa2, a Novel Mammalian Heparanase Family Member. *Biochemical and Biophysical Research Communications 276*, 1170–1177.

(*74*) Levy-Adam, F., Feld, S., Cohen-Kaplan, V., Shteingauz, A., Gross, M., Arvatz, G., Naroditsky, I., Ilan, N., Doweck, I. and Vlodavsky, I. (2010). Heparanase 2 Interacts with Heparan Sulfate with High Affinity and Inhibits Heparanase Activity*. *Journal of Biological Chemistry 285*, 28010–28019.

(*75*) Gross-Cohen, M., Feld, S., Doweck, I., Neufeld, G., Hasson, P., Arvatz, G., Barash, U., Naroditsky, I., Ilan, N. and Vlodavsky, I. (2016). Heparanase 2 Attenuates Head and Neck Tumor Vascularity and Growth. *Cancer Research 76*, 2791–2801.

(76) Liu, J., Knani, I., Gross-Cohen, M., Hu, J., Wang, S., Tang, L., Ilan, N., Yang, S. and Vlodavsky, I. (2021). Role of heparanase 2 (Hpa2) in gastric cancer. *Neoplasia (New York, N.Y.) 23*, 966–978.

(77) Vlodavsky, I., Gross-Cohen, M., Weissmann, M., Ilan, N. and Sanderson, R. D. (2018). Opposing functions of heparanase-1 and heparanase-2 in cancer progression. *Trends in biochemical sciences 43*, 18–31.

(78) Xu, Y.-J., Miao, H.-Q., Pan, W., Navarro, E. C., Tonra, J. R., Mitelman, S., Camara, M. M., Deevi, D. S., Kiselyov, A. S., Kussie, P., Wong, W. C. and Liu, H. (2006). N-(4-[4-(1H-Benzoimidazol-2-yl)-arylamino]-methyl-phenyl)-benzamide derivatives as small molecule heparanase inhibitors. *Bioorganic & Medicinal Chemistry Letters 16*, 404–408.

(79) Courtney, S. M., Hay, P. A., Buck, R. T., Colville, C. S., Phillips, D. J., Scopes, D. I., Pollard, F. C., Page, M. J., Bennett, J. M., Hircock, M. L., McKenzie, E. A., Bhaman, M., Felix, R., Stubberfield, C. R. and Turner, P. R. (2005). Furanyl-1,3-thiazol-2-yl and benzoxazol-5-yl acetic acid derivatives: novel classes of heparanase inhibitor. *Bioorganic & Medicinal Chemistry Letters 15*, 2295–2299.

(80) Madia, V. N. et al. (2018). Novel Benzazole Derivatives Endowed with Potent Antiheparanase Activity. *Journal of Medicinal Chemistry 61*, Publisher: American Chemical Society, 6918–6936.

(81) Simmons, S. C., McKenzie, E. A., Harris, L. K., Aplin, J. D., Brenchley, P. E., Velasco-Garcia, M. N. and Missailidis, S. (2012). Development of Novel Single-Stranded Nucleic Acid Aptamers against the Pro-Angiogenic and Metastatic Enzyme Heparanase (HPSE1). *PLoS ONE 7*, ed. by Antopolsky, M., e37938.

(82) Fu, J., Zhao, B., Dong, Z., Sun, Y., Luan, H., Shen, X., Gao, X., Gong, F., Li, S. and Song, H. (2012). Heparanase DNA vaccine delivered by electroporation induces humoral immunity and cytoimmunity in animal models. *Vaccine 30*, 2187–2196.

(83) He, X., Brenchley, P. E., Jayson, G. C., Hampson, L., Davies, J. and Hampson, I. N. (2004). Hypoxia increases heparanase-dependent tumor cell invasion, which can be inhibited by antiheparanase antibodies. *Cancer Research 64*, ISBN: 0008-5472, 3928–3933.

(84) Shiozawa, H., Takahashi, M., Takatsu, T., Kinoshita, T., Tanzawa, K., Hosoya, T., Furuya, K., Takahashi, S., Furihata, K. and Seto, H. (1995). Trachyspic acid, a new metabolite produced by Talaromyces trachyspermus, that inhibits tumor cell heparanase: taxonomy of the producing strain, fermentation, isolation,

structural elucidation, and biological activity. *The Journal of Antibiotics 48*, 357–362.

(*85*)   Parish, C. R., Coombe, D. R., Jakobsen, K. B., Bennett, F. A. and Underwood, P. A. (1987). Evidence that sulphated polysaccharides inhibit tumour metastasis by blocking tumour-cell-derived heparanases. *International Journal of Cancer 40*, 511–518.

(*86*)   Naggi, A., Casu, B., Perez, M., Torri, G., Cassinelli, G., Penco, S., Pisano, C., Giannini, G., Ishai-Michaeli, R. and Vlodavsky, I. (2005). Modulation of the Heparanase-inhibiting Activity of Heparin through Selective Desulfation, Graded N-Acetylation, and Glycol Splitting*. *Journal of Biological Chemistry 280*, 12103–12113.

(*87*)   Zhou, H. et al. (2011). M402, a Novel Heparan Sulfate Mimetic, Targets Multiple Pathways Implicated in Tumor Progression and Metastasis. *PLoS ONE 6*, ed. by Minna, J. D., e21106.

(*88*)   Hammond, E., Handley, P., Dredge, K. and Bytheway, I. (2013). Mechanisms of heparanase inhibition by the heparan sulfate mimetic PG545 and three structural analogues. *FEBS open bio 3*, 346–351.

(*89*)   Joyce, J. A., Freeman, C., Meyer-Morse, N., Parish, C. R. and Hanahan, D. (2005). A functional heparan sulfate mimetic implicates both heparanase and heparan sulfate in tumor angiogenesis and invasion in a mouse model of multistage cancer. *Oncogene 24*, 4037–4051.

(*90*)   Chhabra, M. and Ferro, V. (2018). The Development of Assays for Heparanase Enzymatic Activity: Towards a Gold Standard. *Molecules 23*, 2971.

(*91*)   Poon, I. K. H., Goodall, K. J., Phipps, S., Chow, J. D. Y., Pagler, E. B., Andrews, D. M., Conlan, C. L., Ryan, G. F., White, J. A., Wong, M. K. L., Horan, C., Matthaei, K. I., Smyth, M. J. and Hulett, M. D. (2014). Mice deficient in heparanase exhibit impaired dendritic cell migration and reduced airway inflammation. *European Journal of Immunology 44*, 1016–1030.

(*92*)   Zcharia, E., Zilka, R., Yaar, A., Yacoby-Zeevi, O., Zetser, A., Metzger, S., Sarid, R., Naggi, A., Casu, B., Ilan, N., Vlodavsky, I. and Abramovitch, R. (2005). Heparanase accelerates wound angiogenesis and wound healing in mouse and rat models. *The FASEB Journal 19*, 211–221.

(*93*)   Barash, U., Cohen-Kaplan, V., Arvatz, G., Gingis-Velitski, S., Levy-Adam, F., Nativ, O., Shemesh, R., Ayalon-Sofer, M., Ilan, N. and Vlodavsky, I. (2010). A

novel human heparanase splice variant, T5, endowed with protumorigenic characteristics. *The FASEB Journal 24*, 1239–1248.

(94) Crispel, Y., Axelman, E., Tatour, M., Kogan, I., Nevo, N., Brenner, B. and Nadir, Y. (2016). Peptides inhibiting heparanase procoagulant activity significantly reduce tumour growth and vascularisation in a mouse model. *Thrombosis and Haemostasis 116*, 669–678.

(95) Hammond, E., Li, C. P. and Ferro, V. (2010). Development of a colorimetric assay for heparanase activity suitable for kinetic analysis and inhibitor screening. *Analytical Biochemistry 396*, 112–116.

(96) McKenzie, E. A. (2007). Heparanase: a target for drug discovery in cancer and inflammation. *British journal of pharmacology 151*, 1–14.

(97) Rivara, S., Milazzo, F. M. and Giannini, G. (2016). Heparanase: a rainbow pharmacological target associated to multiple pathologies including rare diseases. *Future Medicinal Chemistry 8*, 647–680.

(98) Alekseeva, A., Mazzini, G., Giannini, G. and Naggi, A. (2017). Structural features of heparanase-inhibiting non-anticoagulant heparin derivative Roneparstat. *Carbohydrate Polymers 156*, 470–480.

(99) Pala, D., Rivara, S., Mor, M., Milazzo, F. M., Roscilli, G., Pavoni, E. and Giannini, G. (2016). Kinetic analysis and molecular modeling of the inhibition mechanism of roneparstat (SST0001) on human heparanase. *Glycobiology 26*, 640–654.

(100) Galli, M., Chatterjee, M., Grasso, M., Specchia, G., Magen, H., Einsele, H., Celeghini, I., Barbieri, P., Paoletti, D., Pace, S., Sanderson, R. D., Rambaldi, A. and Nagler, A. (2018). Phase I study of the heparanase inhibitor roneparstat: an innovative approach for ultiple myeloma therapy. *Haematologica 103*, e469–e472.

(101) O'Reilly, E. M., Barone, D., Mahalingam, D., Bekaii-Saab, T., Shao, S. H., Wolf, J., Rosano, M., Krause, S., Richards, D. A., Yu, K. H., Roach, J. M., Flaherty, K. T. and Ryan, D. P. (2020). Randomised phase II trial of gemcitabine and nab-paclitaxel with necuparanib or placebo in untreated metastatic pancreas ductal adenocarcinoma. *European Journal of Cancer 132*, 112–121.

(102) Saiki, I., Murata, J., Nakajima, M., Tokura, S. and Azuma, I. (1990). Inhibition by Sulfated Chitin Derivatives of Invasion through Extracellular Matrix and Enzymatic Degradation by Metastatic Melanoma Cells1. *Cancer Research 50*, 3631–3637.

(103) Miao, H. Q., Elkin, M., Aingorn, E., Ishai-Michaeli, R., Stein, C. A. and Vlo-davsky, I. (1999). Inhibition of heparanase activity and tumor metastasis by laminarin sulfate and synthetic phosphorothioate oligodeoxynucleotides. *International Journal of Cancer 83*, 424–431.

(104) Mohamed, S. and Coombe, D. R. (2017). Heparin Mimetics: Their Therapeutic Potential. *Pharmaceuticals (Basel, Switzerland) 10*, DOI: 10.3390/ph10040078.

(105) Liu, C. J. et al. (2014). Adjuvant heparanase inhibitor PI-88 therapy for hepato-cellular carcinoma recurrence. *World J Gastroenterol 20*, 11384–11393.

(106) Khasraw, M., Pavlakis, N., McCowatt, S., Underhill, C., Begbie, S., de Souza, P., Boyce, A., Parnis, F., Lim, V., Harvie, R. and Marx, G. (2010). Multicentre phase I/II study of PI-88, a heparanase inhibitor in combination with docetaxel in patients with metastatic castrate-resistant prostate cancer. *Annals of Oncology 21*, 1302–1307.

(107) Clinical study investigating the safety PG545 in combination with nivolumab in patients with advanced solid tumours and in patients with metastatic pancreatic cancer.

(108) Witczak, Z. J. In *Carbohydrate Drug Design*; ACS Symposium Series 932, Vol. 932, Section: 2; American Chemical Society: 2006, pp 25–46.

(109) Hawking, F. (1978). Suramin: with special reference to onchocerciasis. *Advances in Pharmacology and Chemotherapy 15*, 289–322.

(110) Nakajima, M., DeChavigny, A., Johnson, C., Hamada, J., Stein, C. and Nicolson, G. (1991). Suramin. A potent inhibitor of melanoma heparanase and invasion. *Journal of Biological Chemistry 266*, 9661–9666.

(111) de Boer, C. et al. (2022). Mechanism-based heparanase inhibitors reduce cancer metastasis in vivo. *Proceedings of the National Academy of Sciences 119*, e2203167119.

(112) Baslé, A. and Lewis, R. J. In *Biomolecular and Bioanalytical Techniques*; John Wiley & Sons, Ltd: 2019, pp 385–419.

(113) Maveyraud, L. and Mourey, L. (2020). Protein X-ray Crystallography and Drug Discovery. *Molecules 25*, 1030.

(114) Sun, Z., Liu, Q., Qu, G., Feng, Y. and Reetz, M. T. (2019). Utility of B-Factors in Protein Science: Interpreting Rigidity, Flexibility, and Internal Motion and Engineering Thermostability. *Chemical Reviews 119*, Publisher: American Chemical Society, 1626–1665.

(*115*) Leskovac, V., *Comprehensive Enzyme Kinetics*; Kluwer Academic Publishers: Boston, 2004.

(*116*) Freyer, M. W. and Lewis, E. A. In *Methods in Cell Biology*; Biophysical Tools for Biologists, Volume One: In Vitro Techniques, Vol. 84; Academic Press: 2008, pp 79–113.

(*117*) Tellinghuisen, J. (2008). Isothermal titration calorimetry at very low c. *Analytical Biochemistry 373*, 395–397.

(*118*) Velázquez-Campoy, A., Ohtaka, H., Nezami, A., Muzammil, S. and Freire, E. (2004). Isothermal Titration Calorimetry. *Current Protocols in Cell Biology 23*, 17.8.1–17.8.24.