Application of Activity-Based Probes in Disease

Elisha Moran

PhD

University of York

Chemistry

August 2024

Abstract

The enormous complexity and diversity of carbohydrates are matched by the range of carbohydrate-active enzymes involved in their processing. These enzymes include glycoside hydrolases, a widespread group of enzymes responsible for the hydrolytic cleavage of glycosidic bonds. As glycoside hydrolases play essential biological roles, there is great motivation for developing imaging agents and therapeutic strategies targeting these enzymes. Concerning this, the work discussed in the thesis focuses on developing small-molecule inhibitors for glycoside hydrolases, specifically cyclophellitol-based inhibitors and activity-based probes.

The work presented in **Chapter 2** describes a series of β-glucuronidase X-ray crystallography structures for the determination of the mechanism of inhibition of the broad-spectrum inhibitor siastatin B. Ultimately, the unveiling of the true inhibitors for the mode of siastatin B inhibition has established the mechanism by which the inhibitor acts a broad spectrum inhibitor for glycoside hydrolases with different substrate specificities. As highlighted in the COVID-19 pandemic, there is an urgent need for host-targeted antivirals. **Chapter 3** focuses on the development of cyclophellitol-based inhibitors targeting key proteins of the endoplasmic reticulum quality control pathway in the hope that the work presented validates the use of cyclophellitol-based inhibitors as new antiviral host-targeted agents. A series of N-acyl aziridine oseltamivir-based inhibitors were designed to further address the growing need for new antiviral therapeutics in response to the emergence of antiviral-resistant influenza strains. **Chapter 4** outlines the gene expression of recombinant viral neuraminidase followed by its characterisation with the N-acyl inhibitors. To support this, **Chapter 5** discussed the structural analysis of the covalent mechanism-based N-acyl aziridines inhibitors with viral neuraminidase by cryo-EM. Unexpectedly, a non-covalent species, produced by elimination following a covalent attack, was found in the active site, consequently formed by radiation damage. Overall, the work presented is hoped to support future developments of broad-spectrum inhibitors and cyclophellitol analogues as therapeutic agents (discussed in **Chapter 6**).

Acknowledgements

First and foremost, I would like to express my gratitude to Prof. Gideon Davies. His invaluable advice, continuous support, and patience have been instrumental in my PhD journey. Overall, thank you for being an exceptional supervisor. I also extend my thanks to the colleagues from the Davies group, past and present, for creating a welcoming and enjoyable atmosphere.

I want to thank the faculty and staff of the York Structural Biology Department, including Sam Hart, Dr Johan Turkenburg and Dr Jamie Blaza, for your invaluable assistance with X-ray crystallography and cryo-EM data collection and processing. I would also like to recognise the technicians for their technical support and seamless running of the labs. Concerning scientific support, I would like to thank Dr Zach Armstrong for providing me with a foundation of scientific skills to kickstart my PhD research and Dr Richard Meek for his guidance and willingness to answer my countless questions. Additionally, thank you to Dr. Alison Parkin and Dr. Lianne Willems for their support throughout my PhD.

I want to give a special thanks to my friends, both near and far, who have been my pillars of support throughout my journey. I also want to acknowledge the YSBL coffee club for providing much-needed laughter and entertainment during long lab days. Also, thanks to Isabelle, Megan, and Alex for the unforgettable and humorous times at conferences and to Ginevra, and Rowan for your friendship and support. Additionally, thank you to George for keeping me sane over the past three years.

Lastly, to my parents, thank you for your unwavering love, encouragement and support throughout my journey. Your constant belief and reassurance have been the cornerstone of my achievements, and I am deeply grateful for all you have done for me.

Authors Declaration

I declare that this thesis is a presentation of original work, and I am the sole author. All the work presented was performed by the author unless stated otherwise, this is outlined below. This work has not previously been presented for a degree or other qualification at this University or elsewhere. All sources are acknowledged as references.

- Researchers in Prof. Overkleeft's group (University of Leiden) synthesised inhibitors and activity-based probes discussed in **Chapters 2-5**.
- X-ray crystallography structures were collected externally at the Diamond Light Source facility, which was organised by Sam Hart and Dr Johan Turkenburg.
- Expression of HPSE (**Chapter 2**) was conducted by previous members of Prof. Gideon Davies's group.
- Some data presented in **Chapter 3** were obtained from collaborators Prof. Hermet and Prof. Overkleeft (University of Leiden). This is indicated where applicable in the chapters.
- Researchers from Prof. Overkleeft's and Prof. Rovira's groups conducted some of the data presented in **Chapters 4** and **5**. This is indicated where applicable in the chapters.
- Cryo-EM data collection was conducted at the Electron Biology Imaging Centre (eBIC) with Dr. Jamie Blaza's support.
- Mass spectrometry was conducted by the York Technological Facility.

Contents

Contents of Figure

Contents of Tables

1. Chapter 1- Introduction to Carbohydrate Processing and Glycoside Hydrolases.

1.1. Abstract

Carbohydrates are one of the most diverse and abundant biomolecules on earth. They are complex, exhibiting large stereochemical variation, with the hexasaccharides alone, there are a possible $>10^{12}$ isomers [1]. The enormous complexity and diversity of carbohydrates is mirrored in the range of carbohydrate-active enzymes (CAZymes) involved in their processing. CAZymes play pivotal roles in cell physiology and metabolism; their abnormal activities in disease underscore their importance. Owing to their significance, efforts have been made to develop treatments for CAZymes in disease. This chapter aims to give a brief overview (more detailed introductions precede each chapter) of carbohydrates' diverse roles and functions, including their role in cell physiology, metabolism, and disease. The CAZyme family glycoside hydrolases (GH) will be discussed alongside mechanisms and inhibition. Subsequent chapters will focus on GH's role in disease and the development of novel inhibitors for their treatment.

1.2. Carbohydrates

The most abundant biomolecule in nature is the carbohydrate molecule and is arguably the most diverse form of post-translational modification [2, 3]. The decorating of biomolecules with oligosaccharides through glycosylation is not limited to proteins but also includes glycosphingolipids, proteoglycans, glycosylphosphatidylinositol (GPI)- anchored proteins, free oligosaccharides and polysaccharides which all make up the glycome of the cell [3].

For many years, carbohydrates were mainly seen as building blocks or biochemical fuel, with their biological role majorly underplayed [2]. Glycosylation has since been recognised as one of the most complicated post-translational modifications due to its non-templated manner that goes beyond the genome [4]. Their immense complexity is owed to the diversity of monosaccharide units, which can be linked in various combinations to form oligosaccharides, polysaccharides and glycoconjugates. Various linkages can be formed between monosaccharide units; the glycosidic bond is the most fundamental linkage, as it is found in all oligosaccharides [5]. This bond forms between the anomeric carbon of one monosaccharide and a hydroxyl of another, this linkage can give rise to two stereoisomers α or β [5]. Further diversity can be added through substituents. The diversity and composition of a cell's glycome can change over time, as glycans are regulated by cell differentiation, metabolism and neoplastic transformation [6].

1.2.1. Glycosylation

Animal models of human disease have greatly expanded our knowledge of the human glycome, where the importance of glycosylation in mammalian physiology has been highlighted [3]. Unfortunately, our understanding of glycosylation within the functional proteome is limited, as we cannot dissect the roles of certain glycan-protein relationships [3]. Around 3-4% of the human genome encodes for proteins involved in glycan processing. These carbohydrate-active enzymes (CAZymes) regulate the diverse biological roles of glycans [3].

Simply put, mammalian glycan functions can be divided into four broad categories: structural and modulatory roles, interspecies recognition, intrinsic recognition and molecular mimicry

[7]. A prime example of their structural roles is cellulose, one of the most abundant biomolecules on earth that plays a prominent structural role in maintaining the strength and rigidity of plant cell walls [7]. As mentioned, glycans can function in intrinsic and extrinsic recognition, where immune cells express glycoproteins and glycolipids on their surface to communicate with cells and their environment [8]. Immune cells also recognise pathogen-associated glycans from invading microorganisms to activate the innate and adaptive immune response [8]. Enveloped viruses evade the immune response by deriving their membranes from host cells and utilising mammalian glycoprotein synthesis pathways for molecular mimicry of glycans [8].

1.2.1.1. *Process of Glycosylation*

Mammalian protein glycosylation occurs in the secretory pathway. Ten monosaccharides (Figure 1.1) are used to build the human glycome [3]. The assembled glycans can be attached to proteins by N-linked to asparagine, O-linked to a hydroxyl group of a serine, threonine or tyrosine, and C-linked to tryptophan and glypiation (adding a GPI to protein) [3]. As glycosylation is non-templated, the synthesis of glycans involves glycosylation machinery, modification and degradation.

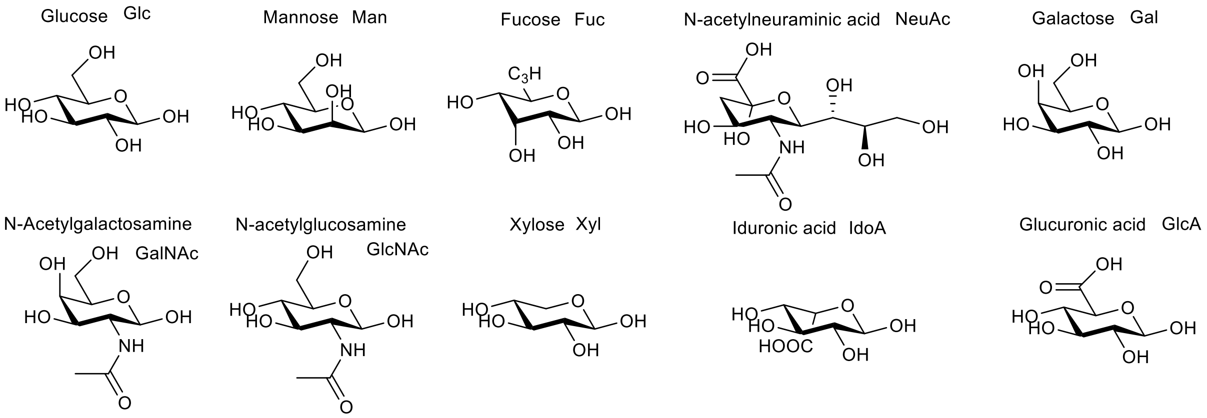


Figure 1.1- Monosaccharides that make up the human glycome.

Most N-glycosylation begins in the endoplasmic reticulum (ER), with a few exceptions that start in the cytoplasm, golgi, and nucleus [9]. Glycosylation is initiated by glycosyltransferases, a group of type II transmembrane proteins with catalytic domains orientated into the ER and Golgi lumen [9]. Here, they transfer glycans onto protein using high-energy donors (UDP or

CMP), which are then further processed by glycoside hydrolases (GHs) [9]. Glycan processing involves sequential steps, which include core extension, elongation and branching (Figure 1.2).

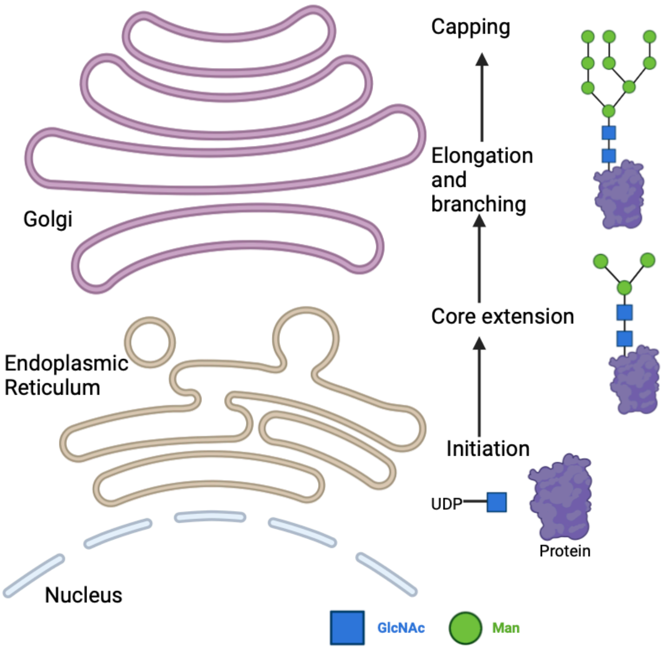


Figure 1.2- Process of N-glycosylation.

Glycan initiation occurs in the ER, where the first monosaccharide or a preassembled glycan is transferred onto the protein via the glycosyltransferases [3]. Further glycosylation processing (core extension, elongation and capping) occurs in the Golgi and trans-Golgi network [3].Image created in BioRender.

1.2.2. Glycans and Disease

Defects in glycosylation highlight their vital role. Congenital disorders of glycosylation (CDG) can cause embryonic defects due to abnormalities in oligosaccharide structure and function [8]. The severity of CDG depends on the glycosylation effects; some affect multiple pathways, altering activation, presentation or transport of carbohydrates and the expression/activity of CAZymes [8]. In N-glycan-related CDGs, the type of gene mutation determines the disease severity; mutations in PMM2 are commonly observed, causing impairment in the production of lipid-bound precursors of N-glycans where the enzyme is partially inactive but is lethal to embryos when inactive [8]. Strategies for CDG treatments include chaperones that stabilise the mutant enzyme, but the lack of knowledge on the underlying mechanism of CDG means there are no current therapies [8].

Changes in the cell glycome occur throughout neoplastic advancement. Cancer cells bypass cell cycle checkpoints to evade cell death, immune response evasion, and cell invasion [8]. Glycosylation patterns change when cells become cancerous and are among the first biomarkers for cancer diagnosis [8]. The cancer glycome changes with its metabolism throughout its development, where abnormal fucosylation, increased N-glycan branching, and presentation of the mucin-O-type glycan antigen are key features of cancer [8].

These increased understandings of how glycobiology is implicated in disease have allowed advancements in glycan-based therapeutics. Microorganisms' evasion of the immune response has informed the design of glycan-dependent epitopes for HIV vaccines to raise neutralising monoclonal antibodies [8]. Furthermore, sugars themselves have therapeutic value, as they can directly target CAZymes [8].

1.3. CAZymes

CAZymes are enzymes involved in the processing of carbohydrates. The CAZyme database provides information on these enzymes' structure, function, and mechanism by categorising them into enzyme families and subfamilies using amino acid sequence similarity [10, 11]. Enzyme families include Glycoside hydrolases, glycosyltransferases, esterases, lyases, and auxiliary activities. CAZyme database was first available in 1998 and today consists of individual CAZyme family pages containing protein names with their UniProt sequence codes, EC numbers, and PDBs [11]. The database was created on the basis that a detailed and accessible summary of CAZymes and their families would be an indispensable tool for glycoscience researchers [10].

1.3.1. Glycoside Hydrolases

The glycosidic bond is among the most stable in nature [12]. GHs are common enzymes with activities and applications that go beyond those of nature; for example, their activities in biomass degradation can be utilised to process biofuels. They are of great interest as they are among the most proficient enzymes in nature, accelerating reactions by 10^{17} fold [13].

There are two main mechanisms of GHs: retention of the anomeric centre or inversion (Figure 1.3) [14]. Both mechanisms occur via a general acid catalysis, including a nucleophile/base and proton donor. The position of the proton donor is identical in both reaction types, but the nucleophile is closer to the anomeric carbon of the sugar substrate for the retaining mechanism than the base is for the inverting mechanism [1]. Inversion of the stereochemistry of the anomeric centre uses a general base catalysis for the attack of water at the anomeric centre and a general acid for the departure of the aglycon leaving group [13]. Meanwhile, retention of the anomeric centre stereochemistry uses a double displacement mechanism to form a covalent intermediate [13]. Firstly, there is general acid/base catalysis for removing the leaving group, creating the covalent glycosyl intermediate with the nucleophile residue [13]. Secondly, water attacks the anomeric centre to break the covalent intermediate, facilitated by the general acid/base residue [13].

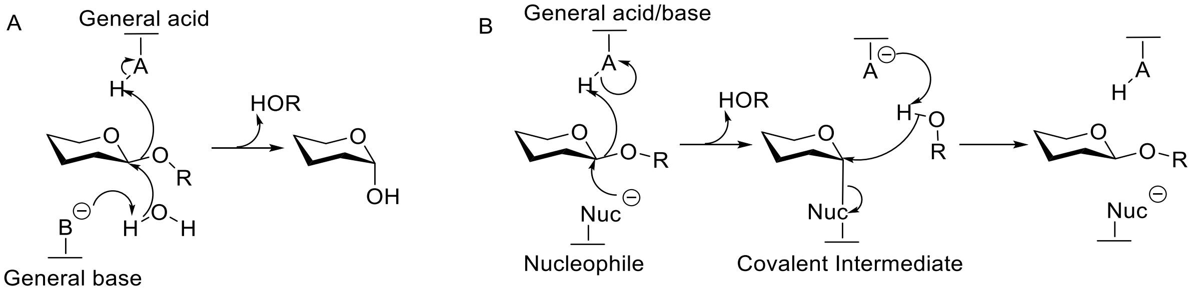


Figure 1.3- Glycoside hydrolase mechanisms.

(A) Inverting mechanism: a water molecule is activated by the general base residue, followed by the attack of the water molecule for the aglycon departure [1]. This single substitution creates a product with an opposite stereochemistry. (B) Retaining mechanism: double displacement mechanism where the stereochemistry of the substrate is maintained through the formation of a covalent intermediate [1].

During GH catalysis, sugars undergo complex conformational itineraries; prediction of these conformations is a significant research focus as this enables the design of inhibitors which mimic the GH transition state (TS) [12]. Previous work has demonstrated that sugars undergo complex conformational distortions, which can be predicted computationally or through direct observations via X-ray crystallography and cryo-EM [12]. For a nucleophilic attack to occur during retaining catalysis, the sugar must be distorted away from a chair conformation to enable a nucleophilic attack, as a nucleophilic attack on a low-energy chair conformation is not chemically feasible [12]. Over the past two decades, the discovery of new GH enzymes and their subsequent analysis have unveiled the different TS of their sugar substrates [12]. The numerous conformations of the six-membered ring are best visualised via a sphere to understand the numerous conformations (Figure 1.4) [12, 15].

In the example below (Figure 1.4), the pyranosides ring is described using a capital letter to define the conformation; this includes C chair, B boat, H half-chair, S skewed boat, and E envelope [16]. The super/subscripted numbers indicate the atoms outside the four-atom reference plane [16]. The sphere is created based on Cremer-Pople puckering parameters (*Q*, θ, and φ) to form a sphere where θ = 0, 180° represent the classical $^4C_1$and $^1C_4$ chairs at the North and South poles while the equator (θ = 90°) represents the skew-boat itinerary [16]. To better visualise, the sphere is projected to its "Stoddart" polar projection or "Mercator" projection and provides a visual framework for the interconversion of different confirmers during catalysis (Figure 1.4).

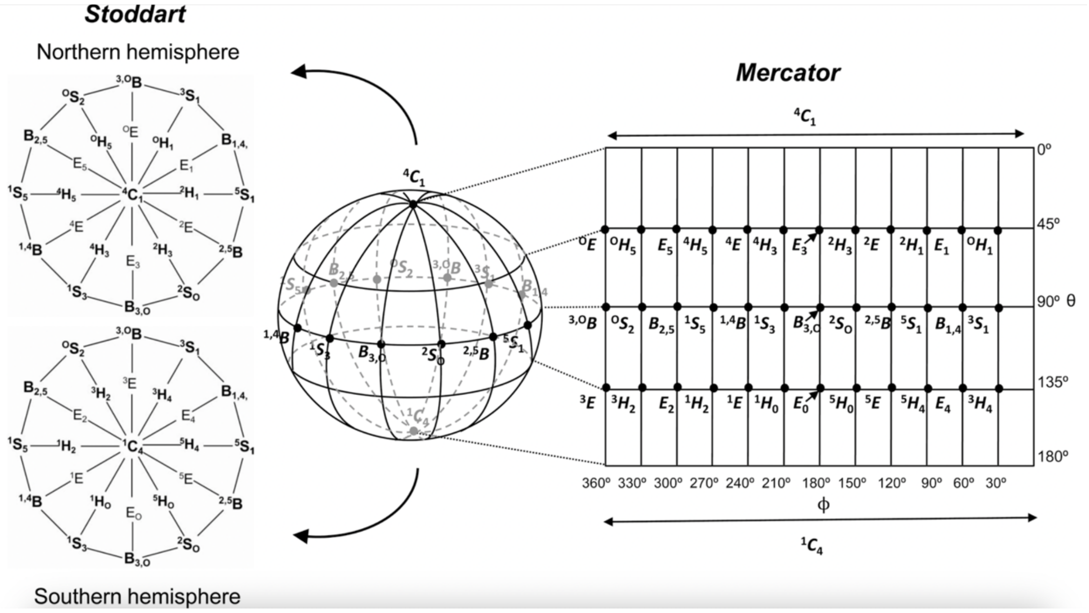


Figure 1.4- Cremer-Pople confirmation sphere with Stoddart (Nothern and southern hemisphere) and Mercator projection.

Image from [16].

GH surface topology also impacts the enzyme's catalytic properties and substrate specificity. The active sites of GHs are adapted to bind to specific substrates and cope with their stereochemical complexity. However, regardless of the enzyme specificity and mechanism, GHs topology fits into three general categories: Pocket, groove or tunnel [1, 12, 17]. The type of topology of a GH gives insight into the enzyme-substrate specificity, where *endo*-acting enzymes need access to internal glycosidic bonds as they act on elongated substrates and release products with varying degrees of oligomerisation [17]. Hence, they have a cleft/groove as their active site [1, 12]. Meanwhile, *exo*-acting enzymes hydrolyse glycosidic bonds at the end of oligo- or polysaccharides, so they typically have pockets or shallow tunnels [1, 12] (Figure 1.5).

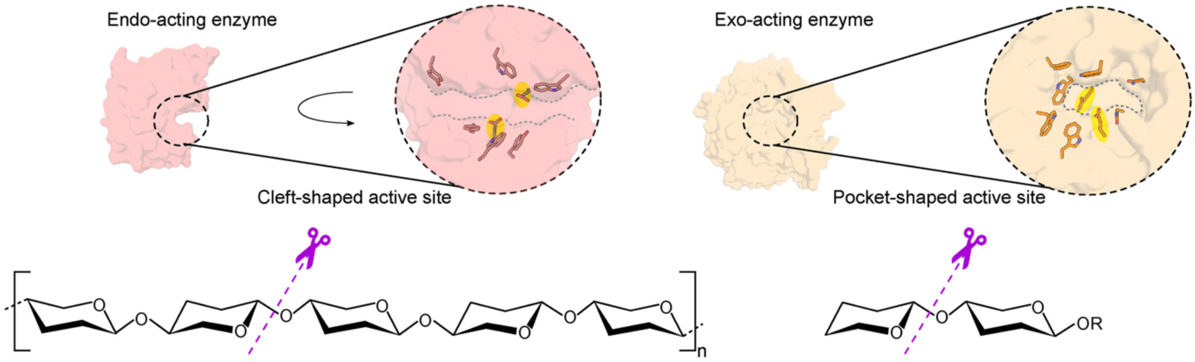


Figure 1.5- Endo- and exo-acting GHs.

Endo-acting GHs harbour a cleft/grove shaped active site for binding to polysaccharide substrates, while exo-acting GHs have shallow tunnels/pockets as active sites for the cleavage of terminal sugars [17]. Image from [17].

1.3.2. Inhibitors of Glycoside Hydrolases

GHs play crucial roles in many biological processes, including the processing of posttranslational modifications. These enzymes are diverse in their activities as they can cleave glycosidic bonds of di-, oligo- and polysaccharides, and glycoconjugates with 1% of the human genome encoding for GHs [18]. Their range in substrate specificities means they are implicated in many critical biological processes, most notably in the digestion of dietary carbohydrates. They also play regulatory roles in glycoprotein processing in the ER and Golgi and regulate the degradation of glycans in the lysosome [18]. As they are ubiquitous enzymes, their abnormal expression/activities are associated with a variety of diseases, such as cancer and lysosomal storage diseases (LSD) [18]. Moreover, microorganisms express their own GHs, and the most well-known example is the viral neuraminidase, which promotes the virus's release from infected host cells [18]. As GHs are involved in various diseases, the creation of small molecule therapeutics has become a highly active field in synthetic chemistry and glycoscience [18].

Structures of GHs have provided information on their catalytic mechanisms through their conformational itineraries; this has allowed for the design of inhibitors, many of which mimic the TS [19]. Covalent or non-covalent inhibitors can be used depending on the application. Non-covalent inhibitors bind reversibly; covalent inhibitors bind irreversibly, blocking the enzyme's activity by forming a stable glycosyl intermediate [19]. Mechanism-based inhibitors have been developed for the labelling of retaining glycosidases by aziridines and epoxides. These groups can specifically target active site residues through covalent labelling (Figure 1.6) [20].

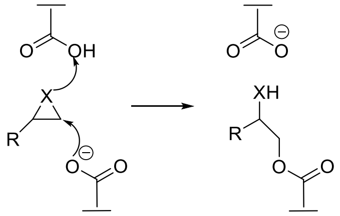


Figure 1.6- Mechanism of covalent binding for epoxides or aziridines. X= O or N.

One of the most characterised mechanism-based inhibitors is the conduritol B-epoxides (CBE) (*Figure 1.7*), a potent inhibitor of α- and β-glucosidases due to their structural symmetry [20-24]. CBE has proven to be a valuable inactivator of GHs. The rational design of other mechanism-based inhibitors includes a CBE derivative with a hydroxymethyl group at the C5 position to create cyclophellitol (*Figure 1.7*) [25]. Cyclophellitols have a structure similar to glucose, which makes the inhibitor more potent and selective for β-glucosidases [25]. Other cyclophellitol derivatives, such as the 1,6-*epi*-cyclophellitol (*Figure 1.7*), have been synthesised to inhibit α-glucosidases [20]. These small molecule mechanism-based inhibitors are potent tools for inactivating GHs and developing Activity-based probes (ABP); this will be explored further in future chapters.

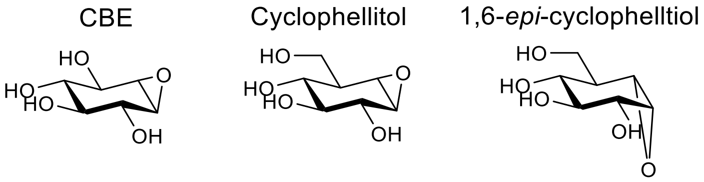


Figure 1.7- Structures of Conduritol B-epoxides, Cyclophellitol and 1,6-epi-cyclophellitol.

1.4. Activity-Based Probes

Advancements in modern genomics have allowed us to decipher the genetic code of any organism. Linking this genomic information with the functional proteome is challenging using just sequence alone [26]. Previous methods of proteome characterisation include transcriptional profiling, gene amplification, and RNA silencing; however, these methodologies rely on manipulating gene expression to determine protein function [27]. Endolysosmal proteases are a prime example of gene expression that does not correlate with protein activity [28]. Analysis by quantifying mRNA or protein levels would suggest that under normal physiological conditions, these proteases are expressed to levels which would induce apoptosis [28]. Still, upon analysis of the functional protease population, the only active forms of the protein are found in endosomes as they require an acidic environment for activity [28]. Thus, protein expression levels do not correlate with activity, as post-translational modifications, protein interactions and small molecules regulate proteins [28]. Activity-based probe profiling (ABPP) can instead be used to visualise only active forms of a specific enzyme to study its function under normal physiological or disease states [28].

ABPs are chemical probes which target the active site of a specific protein/protein family in a mechanism-based manner. This means only catalytically active enzyme forms are labelled as the probe typically reacts with an active site residue. ABP consist of three main components (Figure 1.8). A chemical warhead that mediates the covalent irreversible binding. The warhead is usually a reactive electrophilic moiety that targets a catalytic nucleophile [28]. The warhead is brought into proximity to the nucleophile by the recognition element, which is responsible for the probe's specificity [28]. The reporter tag allows for target characterisation; the tag type depends on the experimental analysis; tags include biotin, fluorophores, alkynes or azides [27]. Several analytical techniques can be used to identify probes. Polyacrylamide gel electrophoresis (PAGE) was initially utilised to detect probe-protein binding using in-gel fluorescence [27]. Other methodologies include Liquid chromatography-mass spectrometry (LC-MS), where proteome pull-downs can be analysed using a biotinylated probe [27]. The components of ABPs are joined by linkers that can be used to improve solubility, steric hindrance and membrane permeability [28].

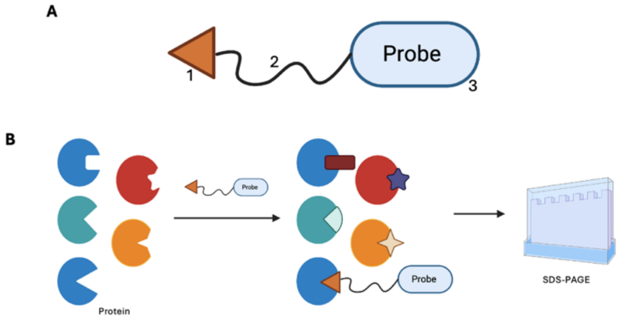


Figure 1.8- Activity-based probe development and applications. (Image created in Biorender)

(A) Scaffold of ABP, including a chemical warhead with recognition element (1), joined by a linker (2) to a reporter group (3). (B) The function of ABP: the probe only binds to the correct enzymes as determined by the specificity of the recognition element. The labelled protein can then be visualised PAGE.

An essential characteristic of ABPs is finding the precise balance between selectivity and reactivity. The rational design of an ABP warhead about specificity ensures that the electrophilic moiety is not too reactive; otherwise, off-target labelling occurs [28]. Recognition groups generally mimic the natural substrate or an irreversible inhibitor [28]. The correct balance between specificity and reactivity is achieved with the appropriate electrophile and recognition element, targeting the ABP to a specific protein/protein family. ABPP can be used for target discovery and to compare proteomes from healthy versus disease samples. Additionally, probes can be adapted to discover new inhibitors by identifying their ability to block probe labelling after adding inhibitors (Figure 1.9) [27].

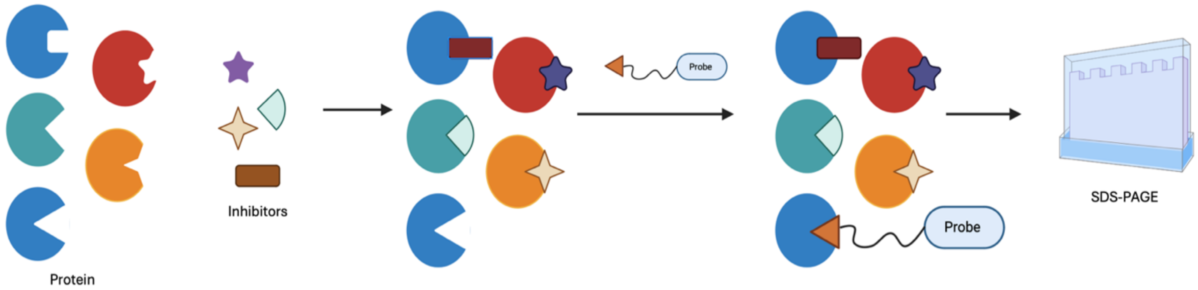


Figure 1.9- ABPP for target discovery. (Image created in BioRender).

Enzymes are incubated with an inhibitor before the adding an ABP. The probe will only bind if the inhibitor has not already, this can then be analysed by PAGE, no band will appear if the inhibitor was bound.

1.4.1. Activity-Based Probes in Glycosidases

As discussed earlier, GHs constitute a significant class of CAZymes implicated in many pathological manifestations. Mechanistically, GHs can perform a retaining or inverting catalysis where their substrates undergo complex conformational itineraries [28]. Retaining catalysis includes the formation of a covalent intermediate that can be applied in the design of GH ABPs, as the recognition group would act similarly to the natural substrate [28]. Both substrates or inhibitors can be utilised as the recognition element; a favourite is cyclophellitol-derived inhibitors, which typically adopt $^4H_3$, miming the TS of retaining glycosidases [28]. Other inhibitor-based ABP includes fluorosugars, which have been used to target β-xylanases, β-glucosaminidases and β-glucocerebrosidase [29-31]. Although fluorosugar-based ABP has been widely used with GH, it can be rapidly overturned by certain GHs such as α-glucosidases [32].

The development of GH ABPP has opened opportunities for drug discovery and *in vivo* imaging. ABPP offers a unique advantage over other diagnostic tools that require manipulating physiological conditions (temperature and pH) [31]. Comparatively, ABPP is cell-permeable and can be applied to whole tissues and organisms. Hence, ABPP has diagnostics and *in situ* imaging applications to analyse how pathological states affect protein levels trafficking and monitor therapeutic efficacy (Figure 1.10) [33]. Additionally, ABPP has applications in inhibitory discovery where probes can analyse compound libraries to discover new inhibitors (Figure 1.10)[33].

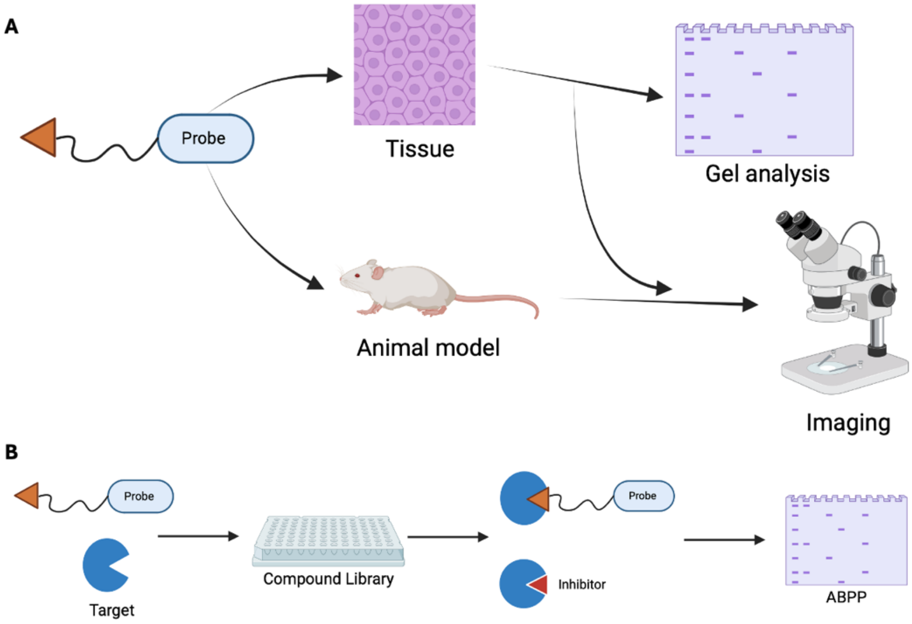


Figure 1.10- Application of ABP. (Image created in Biorender).

(A) Cell-permeable ABP can be used for in vivo imaging, such as the monitoring of active protein populations in disease states. (B) Compound libraries can be screened by reacting the chosen protein with an inhibitor; if a protein-inhibitor complex is formed, then the ABP cannot bind.

As discussed, retaining GHs hydrolyse through the formation of covalent enzyme-intermediate, and any enzymes that follow this mechanism are viable for ABPP [33]. However, *endo*-retaining glycosidases are more complicated; their design requires more complex recognition elements to mimic longer substrates [33]. Additionally, inverting GHs is less amendable for ABPP applications as it requires high-level synthetic chemistry to achieve specificity and reactivity [33].

1.5. This Work

In summary, the hydrolysis of glycosidic bonds by GHs is pivotal for maintaining cellular homeostasis, including glycosylation [18]. Their aberrant expression in disease states highlights their central roles in cell physiology. Hence, GHs are of pathological significance. As discussed in current and future chapters, many iminosugar derivatives and small-molecule inhibitors have been developed to research and inactivate GHs. In the coming chapters, specific GHs will be discussed alongside the development of novel inhibitors for these proteins.

1.5.1. Aims of this Work

In collaboration with Prof. Overkleeft at the University of Leiden and Prof. Rovira at the University of Barcelona as part of an ERC-funded "Synergy" consortium, this work aims to expand on the current ABP portfolio for retaining enzymes and the development of novel inhibitors. Specifically for my work, human heparanase, α- glucosidases I and II, human neuraminidase and viral neuraminidases will be investigated with inhibitors using biochemical characterisation and structural biology.

The summary of this work is as follows:

- Structural analysis of the mechanism of inhibition for the iminosugar siastatin B with HPSE and other glucuronidases. Collaboration with the Overkleeft lab (University of Leiden). -**Chapter 2**
- Gene expression of recombinant α-glucosidase I and II to structurally and biochemically analyse a range of 1,6 and 1,2 cyclophellitol-based inhibitors to design broad-spectrum antivirals. Collaboration with the Overkleeft lab (University of Leiden). -**Chapter 3**
- Recombinant expression of human and viral neuraminidases and their biochemical characterisation with N-acyl Oseltamivir aziridine derivatives. Collaboration with Overkleeft lab (University of Leiden) and Rovira group (University of Barcelona). – **Chapter 4**
- Structural analysis of viral neuraminidase with mechanism-based N-acyl Oseltamivir aziridine derivatives to inform the design of potent direct-acting antivirals. Collaboration with Overkleeft lab (University of Leiden) and Rovira group (University of Barcelona). – **Chapter 5**
- A final, **Chapter 6**, conclusions and perspectives chapter, provides a summary of the results and their broader context.

2. Chapter 2- Mechanism of Siastatin B Inhibition with HPSE: Anti-Cancer Inhibitors.

2.1. Abstract

The sulfated glycosaminoglycan polysaccharide known as heparan sulfate (HS) is a highly heterogeneous molecule central to biological and pathological processes, including its role as a ubiquitous extracellular matrix component [34] [35]. HS regulation is modulated by continual breakdown and biosynthesis. Heparanases are GH79 endo-β-glucuronidases involved in the breakdown of HS [35]. The human heparanase (HPSE) is involved in the cleavage of HS at GlcA residues. HPSE is a promising target for anti-cancer drugs due to its enzymatic and non-enzymatic activity; it is highly involved in cancer biology. Aberrant HPSE expression has been linked to the development of larger and highly metastatic tumours, consequently leading to poor prognosis [36]. Our work on HPSE activity-based probes led us to study new classes of HPSE inhibitor.

The iminosugar siastatin B is a known inhibitor of neuraminidases, N-acetyl-glucosaminidases and β-D-glucuronidase [37]. However, structures of HPSE with inhibitors and substrates suggest that the active site could not accommodate siastatin B due to the steric clashes of the *N*-acetyl group at the 2-position [38]. Here, we show that the inhibition of HPSE (and related enzymes) reflects not the parent compound but a pair of breakdown products[A]. We found a galacturonic-noeuromycin or 3-geminal diol derivative of galacturonic-isofagomine in the active sites of HPSE and homologs upon soaking crystals with siastatin B. These derivatives represent a new class of glycosidase inhibitors, which were subsequently re-synthesized, and their co-crystal structures with HPSE and homologs were investigated, where the compounds displayed submicromolar inhibition of exo-β-glucuronidases. Our results have demonstrated the mechanism by which siastatin B acts as a broad-spectrum glycosidase inhibitor. Therefore, the discovery of this new class of glycosidase inhibitors may offer new opportunities for developing prodrug-like inhibitors that can degrade into the pharmacologically active drug.

[A] Yurong Chen, Adrianus M. C. H. van den Nieuwendijk, Liang Wu, Elisha Moran, Foteini Skoulikopoulou, Vera van Riet, Hermen S. Overkleeft, Gideon J. Davies, and Zachary Armstrong. *Journal of the American Chemical Society* **2024** *146* (1), 125-133

2.2. Introduction

Proteoglycans are ubiquitous on the cell surfaces and extracellular matrix (ECM) of vertebrates and invertebrates. They are involved in many physiological processes due to their sizeable structural complexity [39]. Heparan sulfate proteoglycans (HSPGs) are constructed from a core protein covalently linked to heparan sulfate (HS), a sulfated glycosaminoglycan polysaccharide made up of repeating disaccharides of hexuronic acid (L-iduronic or glucuronic acid) and a D-glucosamine (*N*-acetyl glucosamine or *N*-sulfo glucosamine) to produce a linear polysaccharide that can vary in length from 5-70 kDa [40, 41]. Their structural complexity and diversity create a large polydispersity, the sugar residues can be modified with sulfation regions clustered along HS; the orientation and arrangement of the sulfated and carboxyl groups create ligand-binding sites (Figure 2.1) [42].

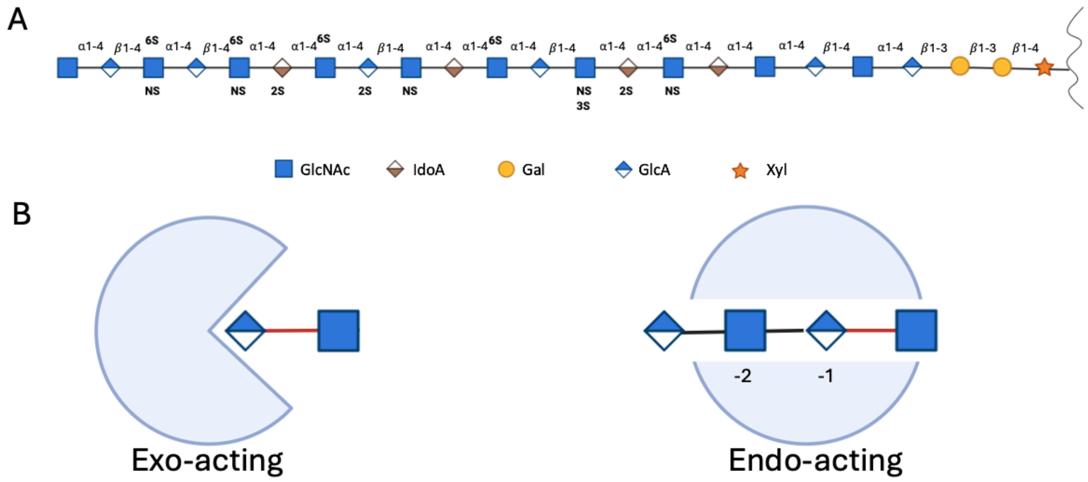


Figure 2.1- Sugar composition of heparan sulfate with mode of hydrolysis of exo- and endo-acting GH79 β-glucuronidases (Image created in BioRender).

(A) Heparan sulfate is assembled from GlcNAc, IdoA and GlcA and undergoes sulfation, epimerisation, and de-acetylation processing. (B) Difference between cleavage of an endo- and an exo-acting β-glucuronidases, red lines indicate where the hydrolysis occurs.

HS biosynthesis occurs in the endoplasmic reticulum and Golgi, where the non-sulfated precursor is modified in a non-templated manner through deacetylation, sulfation, and epimerisation [43]. These modifications create HSPGs large interactome where it modulates the availability and gradients of growth factors; hence, the polysaccharide plays a fundamental role in growth factor signalling, morphogenesis and pathophysiology [44] [45]. HSPGs interactome involves a range of proteins, including tyrosine kinase and chemokine receptors (Figure 2.2) [46].

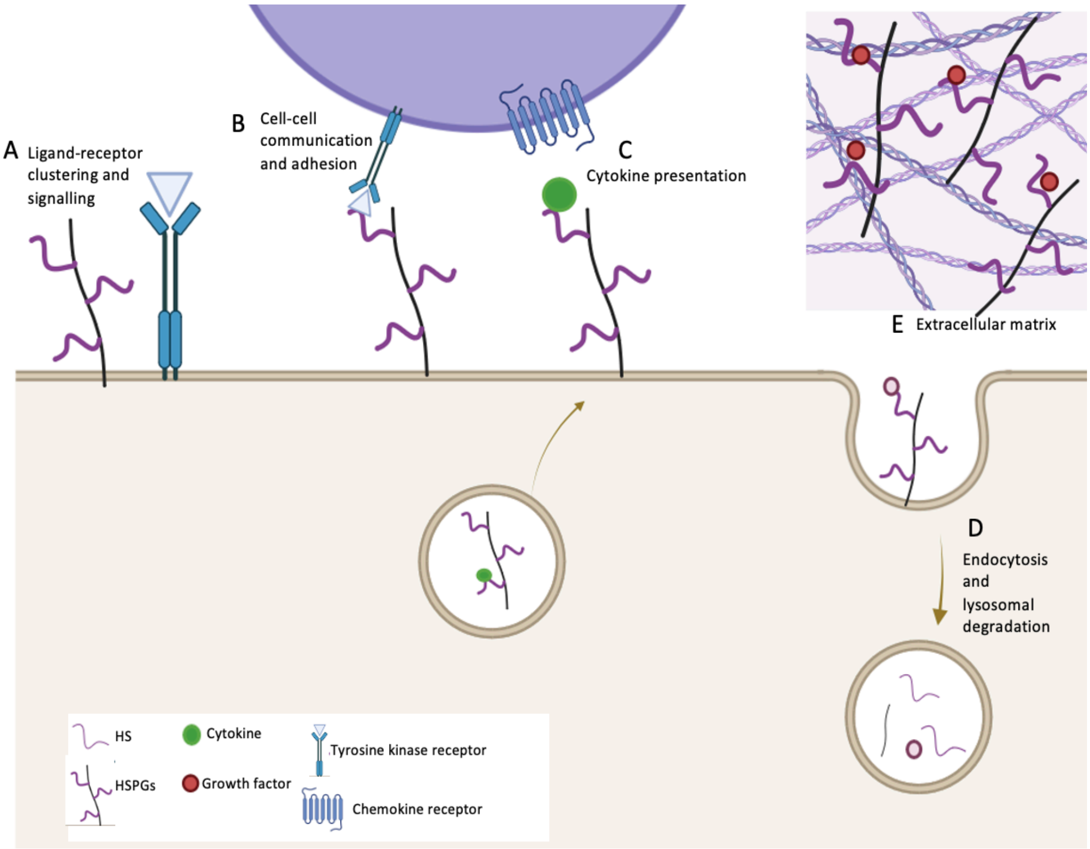


Figure 2.2- Role of heparan sulfate proteoglycans in cell physiology (image based on [46] and created in BioRender).

HSPGs act as co-receptors for ligand-tyrosine kinase receptors either in the same cell (A) or with adjacent cells where they are involved in cell-cell communication and cell adhesion (B). HSPGs bind to cytokines and are transported to the cell's surface, where they are presented to cytokine receptors (C). Endocytosis of HSPGs sequestering signalling molecules is taken into lysosomes where they are degraded and recycled (D). HSPGs are a major ECM component where HS sequesters growth factors and proteins (E).

In cultured cells, HSPGs have a 3-4 hour turnover due to the cell's response to external stimuli; this HS breakdown occurs through the endo-glucuronidases [43]. Glycoside hydrolase (GH) 79 β-glucuronidases are involved in the breakdown of HS; human heparanase (HPSE) is the only known mammalian heparanase involved in the cleavage of HS at GlcUA residues [35]. Aberrant expression of HPSE is detected in multiple pathological conditions, including cancer and amyloid diseases.

2.2.1. Human Heparanase

HPSE belongs to the GH 79 family (EC 3.2.1.166), which contains both endo-acting and exo-acting β-glucuronidases. The first structure for the GH79 family was the exo-acting β-glucuronidase of Acidobacterium capsulatum, which reported a (β/α) 8 TIM-barrel fold with glutamic acids acting as the acid/base and nucleophile, supporting the previously predicted retaining mechanism [47].

HPSE is the only human heparanase and is encoded by a single gene in the human genome, thus making it of interest physiologically, biochemically, and medically. HPSE genome has an 80% sequence identity to other animal heparanase's such as mice and rats [48]. Due to HPSE's pivotal role in cellular processes, its expression is upregulated in response to stimuli such as cell growth and immune cell activation through the action of several growth factors like cytokines. In turn, HPSE regulates the downstream expression of several genes and the bioavailability of signalling molecules [48]. Intracellularly, HPSE is found in late lysosomes and endosomes to process internalised HSPGs; extracellularly, it is involved in the processing of the ECM [34].

Initially, HPSE is translated as a pre-proenzyme (*Figure 2.3*). HPSE pre-proenzyme signal sequence M1-A35 is cleaved to produce the inactive 65 kDa proHPSE. Further proteolytic processing removes the S110-Q157 linker, freeing the 8 kDa N-terminus and the 50 kDa C-terminus, which fold to form the mature heterodimer HPSE [34]. The first 3D structure was solved by Wu and colleagues in 2015 (PDB: 5E8M) [34]. The two subunits associate non-covalently to form a heterodimer with a $(\beta/\alpha)_8$ TIM-barrel fold flanked by a small β-sandwich [34]. In the first structure of HPSE (5E8M) the active site is 10 Å in length [34], and contains the previously reported catalytic residues (E343 and E225) along with basic residues (A35, K158, K159, K161, K231, R272, R273 and R303) that interact with the negatively charged HS substrate. Additionally, there are six putative glycosylation sites. Still, in the first HPSE structure (5E8M), only five sites were structurally determined with N-linked GlcNAc residues found at N162, N200, N217, N238 and N459, which has additional α1→6-linked fucose-linked to GlcNAc.

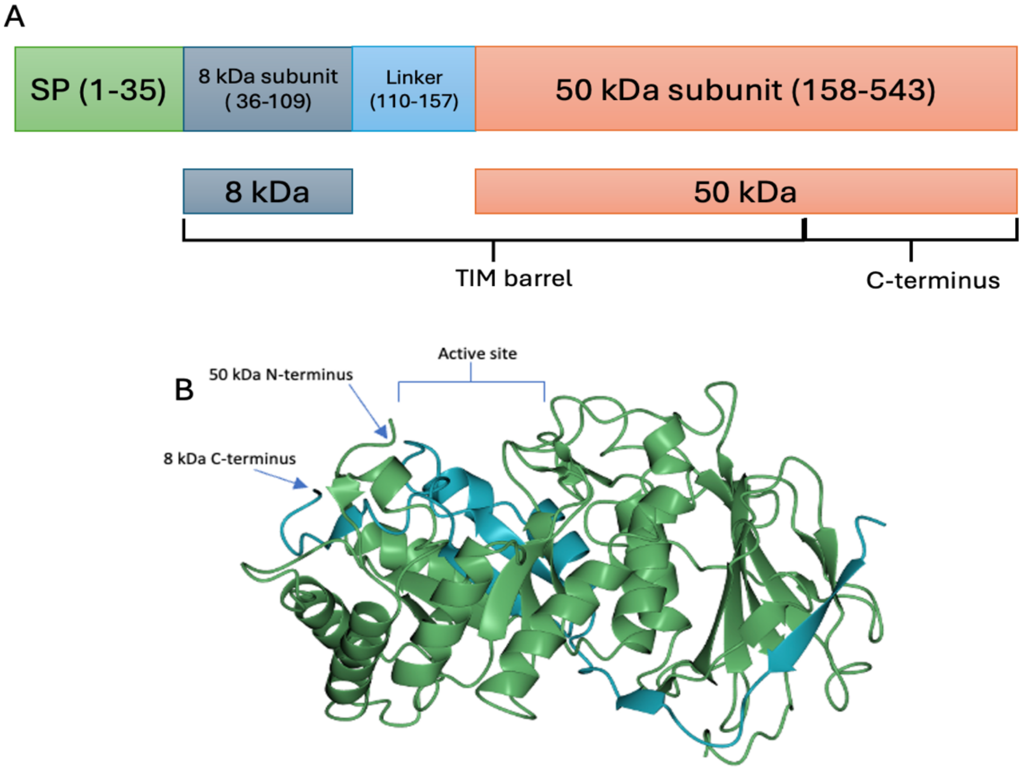


Figure 2.3- Configuration of the translated pre-proenzyme of HPSE, which undergoes proteolytic processing to produce the catalytically active HPSE.

(A) Translated pre-proenzyme HPSE, containing the cleaved signal peptide (SP) and linker region, which are proteolytically cleaved to yield the final HPSE structure. (B) Structure of HPSE, both the 8 kDa and 50 kDa domains contribute to the $(\beta/\alpha)_8$ TIM-barrel fold. The 8 kDa subunit contributes a small *β sandwich, and the first β-α-β fold and the remaining folds are contributed by the 50 kDa.*

Substrate recognition of HS by HPSE was initially investigated using defined oligosaccharides isolated from HS. Kinetic assays and NMR have confirmed that HPSE cleaves with net retention of the anomeric centre cleaving at the Glc (β-1,4) GlcNS linkages at sulfated regions of HS [34, 49, 50]. The minimum backbone required for cleavage is the trisaccharide GlcN-GlcUA-GlcN, where the glucosamine has a 6-O-sulfate group [50]. Catalysis occurs through Glu225 and Glu343 acting as the acid/base and nucleophile using a $^{1}S_{3}$- $^{4}H_{3}$- $^{4}C_{1}$ reaction itinerary (*Figure 2.4*)[51]; enzyme mutants of these catalytic residues render the complete loss of activity [52].

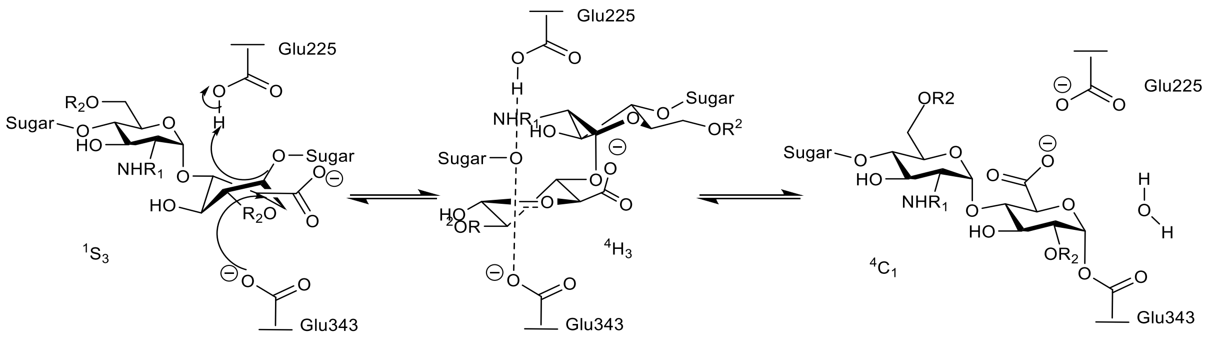


Figure 2.4- Conformational itinerary of HPSE. Retaining mechanism with $^1S_3$- $^4H_3$- $^4C_1$. R1= Ac, SO_3^- and R2= H, SO_3^-.

Later, the recognition of the sulfation pattern by HPSE was further confirmed by crystallography structures (Table 2.1) complexed with defined HS oligomers, verifying the recognition of a trisaccharide with N-sulfation at the -2 subsite and +1 subsite for 6-O-sulfation, as they form hydrogen bonds within the HPSE binding cleft (*Figure 2.5*) [34, 53]. The binding of the substrate is then further stabilised by electrostatic interactions due to the -2 6-O-sulfate and +1 N-sulfate [34].

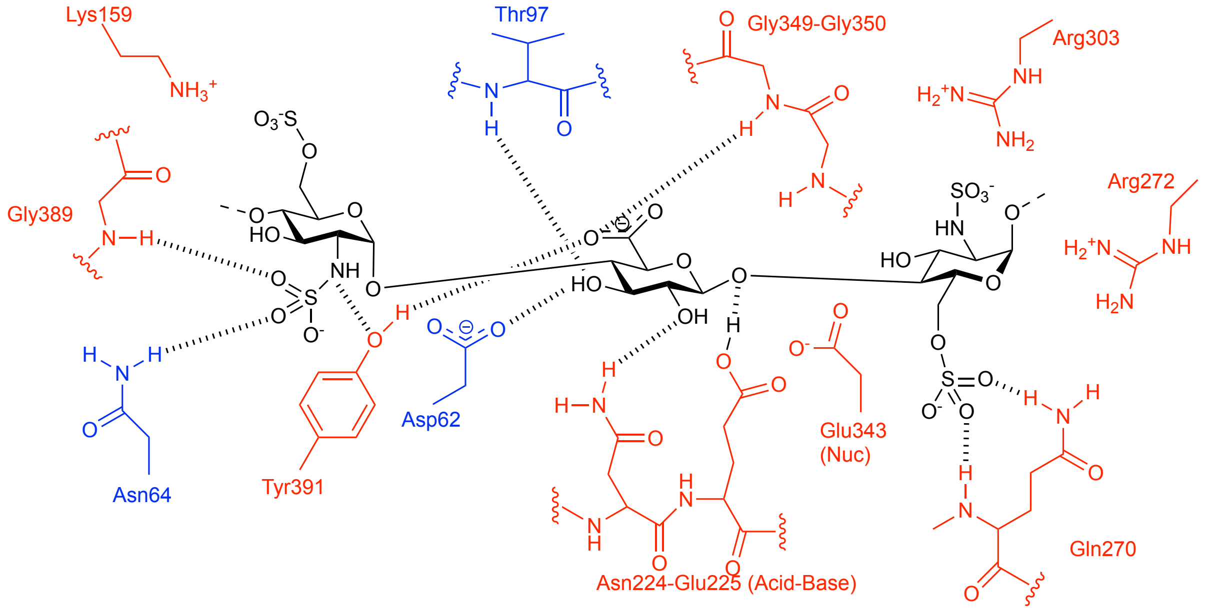


Figure 2.5- Substrate interactions of HS in HPSE binding cleft of the -2, -1 and +1 subsites (Image based on [34]).

Substrate interactions with HPSE in the binding cleft with annotated residues (blue residues are from the 8 kDa subunit and red residues from the 50 kDa subunit) based on the ligand interactions in PDBs 5E98 and 5E9C. GlcN-GlcUA-GlcN trisaccharide in the -2, -1 and +1 subsites with the acid/base Glu225 and nucleophile Glu343.

Table 2.1- PDB structures of HPSE.

PDB	Organism	Structure information	Res (Å)	Reference
5E8M, 5E97, 5E98, 5E9B,5E9C, 5L9Y, 5L9Z, 5LA4,5LA7	Human	"apo", Pro-HPSE and HPSE complexes with HepMers and ABP	1.63-1.94	Wu et al (2015)
7RG8	Human	Stable heparanase mutant	1.3	Whitefield et al (2022)
6ZDM	Human	Fluorogenic heparan sulfate disaccharide	1.7	Wu et al (2022)
7PR7, 7PR8, 7PRT	Human	HPSE complex with synthesised covalent inhibitors	1.52-1.70	De Boer et al (2022)
7YI7, 7YJC,	Human	HPSE in complex with inhibitor	2.30-2.80	Iami et al (2022)
8BOB, 8B0C	Human	HPSE in complex with inhibitor	1.95-2.1	Borlandelli et al 2023
8BAC	Human	HPSE in complex with inhibitor	1.62	Doherty et al, 2023
8E07, 8E08	Human	HPSE in complex with inhibitor	1.80-1.93	Whitefield et al , (2023)
8JYG	Human	HPSE in complex with inhibitor	2.00	Iami et al (2023)

2.2.2. HPSE and Disease

The ECM is a specialised and dynamic fibrillar cellular scaffold composed of collagen, fibronectin, elastin, and non-fibrillar molecules such as hyaluronan, glycoproteins, and proteoglycans [54]. It is a structural network in multicellular organisms responsible for the support and integrity of cells and tissues [48]. Growth factors are sequestered within the ECM due to their interactions with HS; hence, it is firmly rooted in cellular communication. The ECM controls the release of signalling molecules involved in cell proliferation, motility, angiogenesis, and apoptosis [53, 55, 56]. The tight regulation of ECM remodelling can be compromised in pathological conditions.

Under normal physiological conditions, the HPSE promoter is controlled by methylation silencing and transcription factors, but in disease states, single nucleotide polymorphisms at the promoter lead to altered HPSE expression (Figure 2.6)[48]. Different pathogens have been shown to hijack HPSE expression by targeting pro-survival pathways; upregulation of HPSE in herpes simplex, dengue, and cytomegalovirus has been observed, identifying a new therapeutic and diagnostic potential [57, 58]. HPSE upregulation is also implicated in certain bacterial infections such as *Helicobacter pylori (H.pylori) and* correlates with poor diagnosis due to the development of gastrointestinal cancer [59].

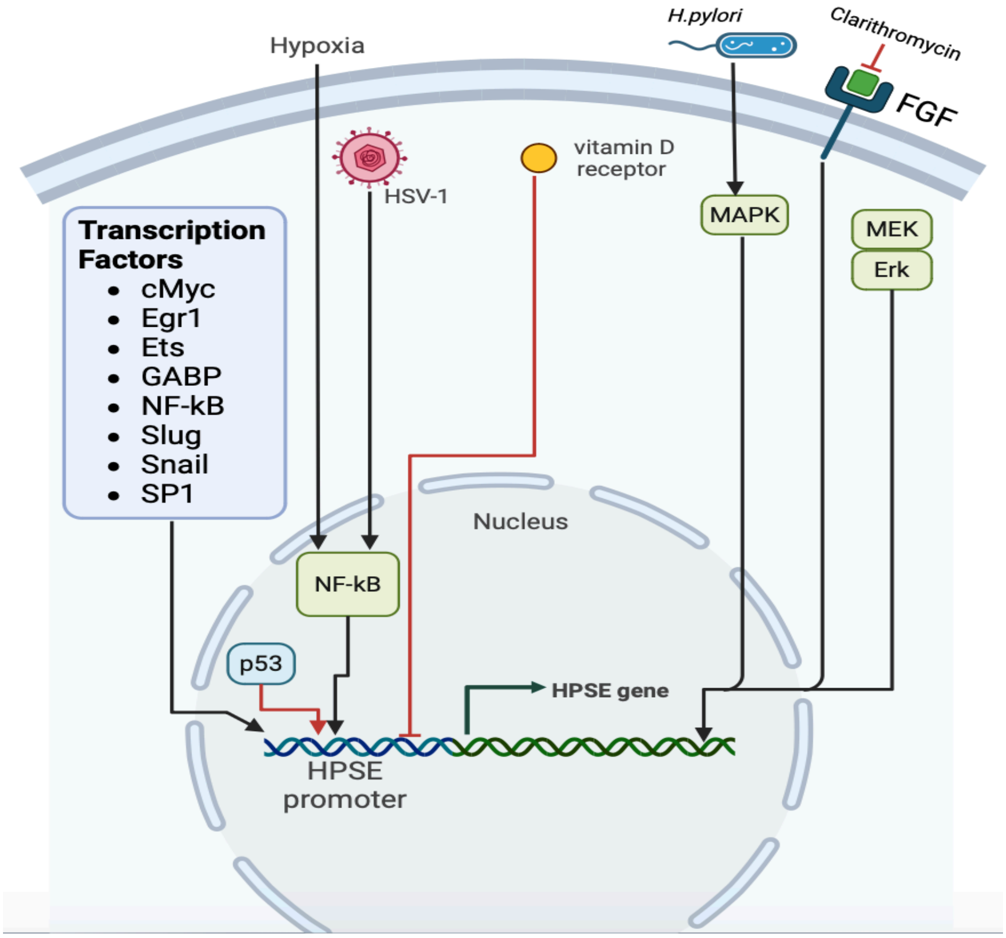


Figure 2.6- Regulation of HPSE promoter activation and deactivation by transcription factors and receptor signalling cascades Image created in BioRender)..

Image-based on [48]. Listed transcription factors activate the HPSE promoter, resulting in upregulated HPSE expression. HSV-1 infections and cell stress due to hypoxia can activate other transcription factors, such as NF-kB. Other transcription factors like p53 and vitamin D receptors inhibit the transcription of the HPSE gene.

2.2.3. HPSE and Cancer

HPSE activity correlates heavily with the development of multiple cancerous phenotypes. The master regulator p53, which regulates cell cycle apoptosis, controls HPSE expression through interaction with its promoter, causing transcription suppression [48]. Loss of the tumour suppressor p53 consequently causes the overexpression of HPSE, a key cancer characteristic [48]. The upregulation of HPSE leads to tumorigenic ECM remodelling, influencing changes in the cell's microenvironment due to the dysregulation of ECM modifying enzymes, causing chemical modifications at the post-translational level and proteolytic degradation [56]. The biochemical and biophysical changes of the ECM lead to effects on its spatial topology, increasing the bioavailability of signalling molecules involved in growth, inflammation, extravasation and angiogenesis, shaping the tumour microenvironment (Figure 2.7) [60].

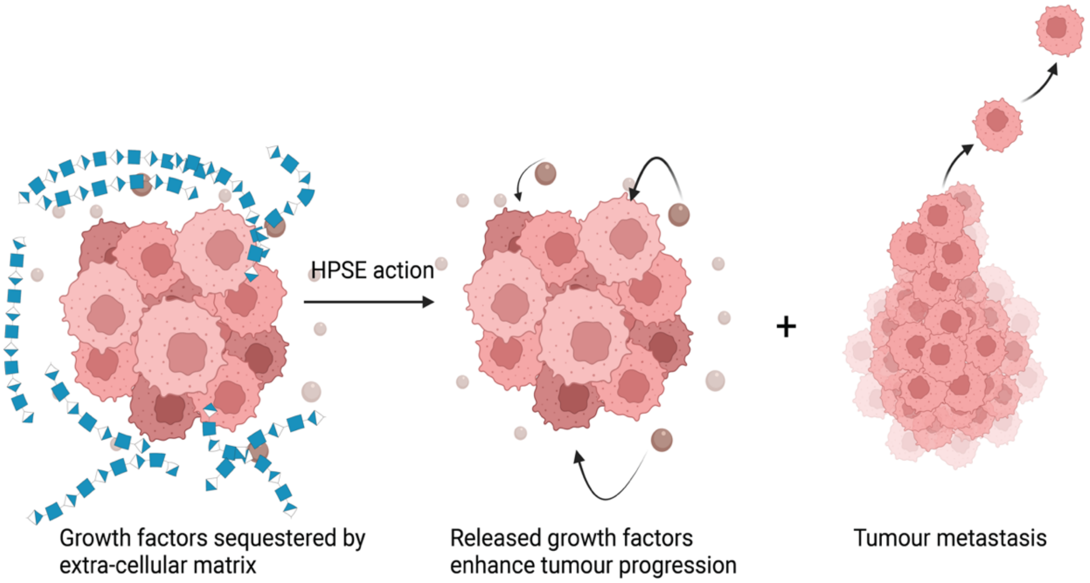


Figure 2.7- Effects of aberrant HPSE expression on the processing of the ECM leading to tumour progression and eventually cancer cell dissemination (Image created in BioRender).

Overexpression of HPSE in cancer cells leads to the deterioration of the ECM surrounding the tumour, releasing sequestered growth factors and signalling molecules residing in the ECM. This, in turn, amplifies tumour progression, leading to the progression of cancer and, eventually, metastasis.

HPSE promotes aggressive cancerous phenotypes, making it a promising target for anti-cancer drugs. It is the primary human enzyme involved in the breakdown of HS, linking aberrant HPSE activity to all the hallmarks of cancer (Figure 2.8) [60].

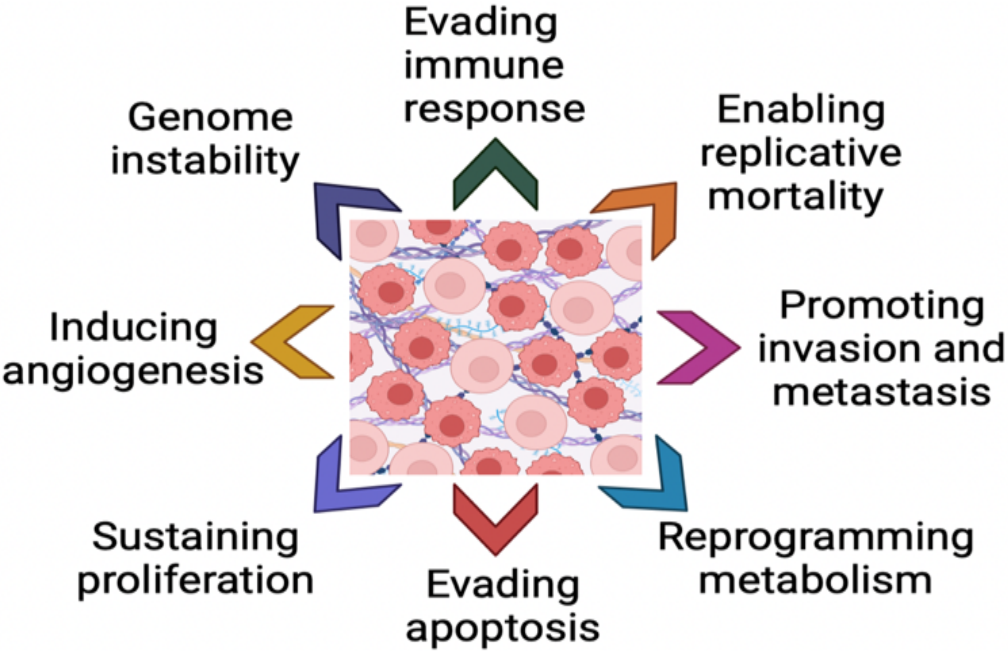


Figure 2.8- HPSE role in the hallmarks of cancer (Image created in BioRender).

The hallmarks of cancer consist of features that are successively acquired by cells during the multistep development of neoplastic disease [61], all of which HPSE is involved in. HPSE is involved in the original six hallmarks of cancer as well as the new emerging hallmarks through its enzymatic and non-enzymatic activities.

2.2.4. Human Heparanase Inhibitors

Many molecules have been developed to inhibit HPSE, including small molecules, antibody therapeutics, heparin mimetics, and neoproteoglycans [38]. Novel therapeutics such as antibodies have greatly improved the survival rate of cancer patients but with various adverse side effects; hence, there is a need for the development of novel molecules which prevent cancer progression without adverse side effects [62]. Recent acknowledgement that the tumour microenvironment drives aggressive cancerous phenotypes has led to the recognition that HPSE could be a potential strategy for targeting the tumour microenvironment due to the enzyme's involvement in various tumour types [61].

Only four HPSE targeting molecules have advanced to clinical trials, all of which are heparin mimetics. Heparin is a well-known HPSE inhibitor that has been shown to inhibit the degradation of the ECM via HPSE in various cancer types [19]. Heparin mimetics possess heparin-like properties and high structural heterogeneity, although they can have unwanted side effects such as anticoagulation and growth factor binding [19]. These molecules can be further chemically modified by adding sulfation to mimic the sulfation pattern involved in substrate recognition [19].

Four low anticoagulant polyanionic polysaccharide heparin mimetics have progressed to clinical trials: PG545, M402, SST0001, and PI-88 (Figure 2.9). These molecules act by downregulating pathways such as angiogenesis, cell growth, and cell dissemination through the inhibition of HPSE, and they can form binding interactions similar to those of the natural substrate [38]. These inhibitors were best-in-class due to their mimicry of the natural HPSE inhibitor heparin, which is closely related to the HS substrate [38]. However, none have been clinically approved, with PI-88 and M402 terminating at phase II and III clinical trials due to adverse side effects and insufficient efficacy, possibly caused by their highly charged heterogeneous nature that could produce unwanted side effects [38, 63-65].

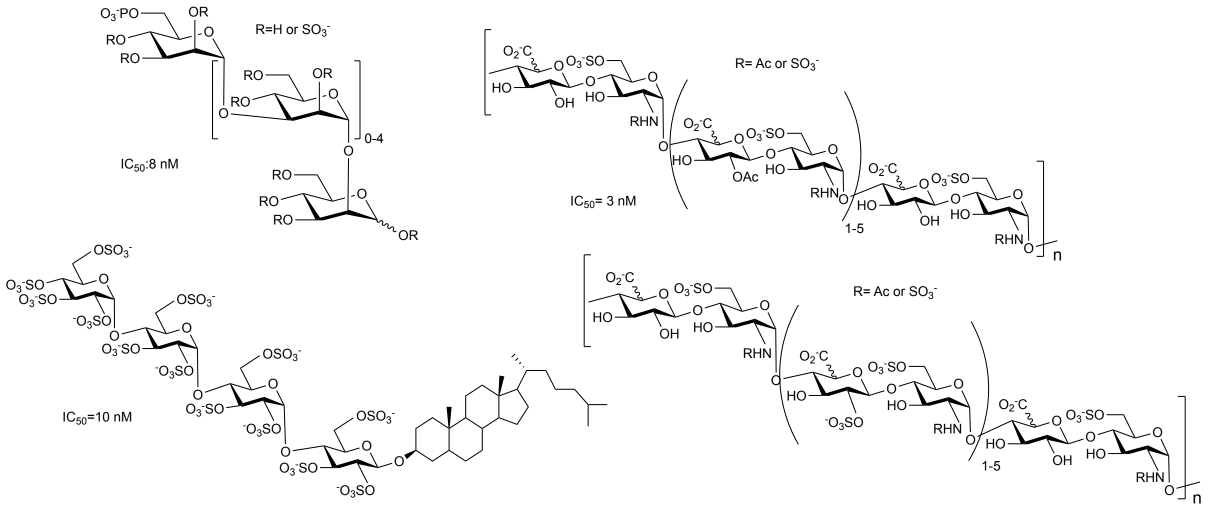


Figure 2.9- HPSE heparin mimetic inhibitors.

Image adapted from [38]. PI-88 (muparfostat) HPSE inhibitor for hepatocellular carcinoma failed phase III primary end survival, terminated at phase III clinical trials due to adverse side effects [63, 66]. M402 (necuparanib) is an HPSE inhibitor that stimulates apoptosis and anti-

angiogenesis in metastatic cancer, terminated at phase II clinical trials due to insufficient efficacy to warrant progress [65, 66]. PG545 (Pixatimod) is an angiogenesis inhibitor targeting solid tumours. Currently in phase II clinical trials for Non-small cell lung cancer, Metastatic Colorectal Carcinoma and Refractory Melanoma [66, 67]. SST001 (roneparstat) causes the downregulation of signal factors, such as vascular and human growth factors, by inhibiting the HPSE [66].

HPSE has evolved to bind HSPGs, due to the large polysaccharide substrate the active site consists of a large interaction surface [38]. This has led to HPSE historically being targeted by heparin mimetics instead of small-molecules. Small molecules have been developed with promising results, but none have progressed to clinical trials. High-throughput screening of potential HPSE inhibitors in 1999 unveiled the first set of small molecule inhibitors that selectively bind HPSE, based on the original hit of 2,3-dihydro-1,3-dioxo-1*H*-isoindole-5-carboxylic acid (Figure 2.10A) [66, 68]. The compound had moderate anti-anticoagulant activity and selectivity to HPSE (IC_{50} 8 µM) due to the carboxylic acid and arrangement of the benzoxazole ring, as their removal/rearrangement abolished activity [66, 68]. Using this as the initial scaffold, a set of compounds was created with improved selectivity and anti-angiogenesis activity in the sub-micromolar range (Figure 2.10B) [66, 68].

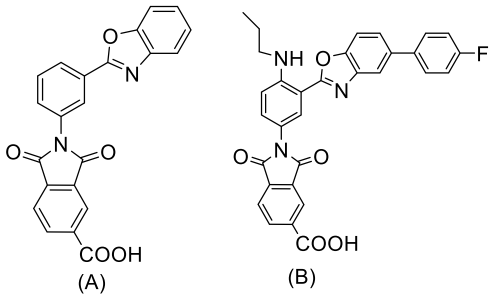


Figure 2.10- First set of HPSE small molecule inhibitors.

2,3-dihydro-1,3-dioxo-1H-isoindole-5-carboxylic acid (A), inhibitors designed based on this compound are potent inhibitors of HPSE (IC_{50} 8 µM). Compound B was created from the original scaffold from (A) with HPSE inhibition activity of IC_{50} 0.5 µM [66, 68].

Irreversible binding is not generally considered for drug development due to its covalent nature; however, it can be advantageous compared to competitive inhibitors due to their specificity and reduced off-target effects [38, 69]. Therefore, mechanism-based inhibitors could combat the problems of small molecules' initial weak/transient binding. Their highly selective nature is advantageous for targeting the catalytic nucleophile, as HS extensive interactions are hard to mimic for small molecules.

Cyclophellitols have been previously shown to be selective irreversible inhibitors of β-glucosidases; nucleophilic attack occurs through the trans-diaxial opening of the epoxide by the catalytic nucleophile, forming an ester linkage (Figure 2.11)[69]. Mechanism-based inhibitors for HPSE are challenging to design as *glucurono*-cyclophellitols have selectivity for other β-glucuronidases due to the 4-*O*-position (Figure 2.11B) [51]. The synthesis of a GlcNAc-*glucurono*-cyclophellitol (Figure 2.11C) created a more potent inhibitor with selectivity for HPSE by mimicking the $^4H_3$ transition state [51]. Substitution of the GlcNAc moieties at the 4-position was explored to create a more potent 4-*O*-alkyl glucuronic cyclophellitol micromolar inhibitor [51]. Comparatively, cyclophellitol epoxides with 1,4 glucosaminyl moiety have shown nanomolar potency as the electrophilic epoxide irreversibly targets the catalytic nucleophile (Figure 2.11) [38]. These derivatives were applied to U87 glioblastoma cells to assess their ability to inhibit cancer metastasis, where they showed a reduction in Matrigel invasion [38]. A comparison of compound (Figure 2.11C) with the HS mimetic for reducing B16 murine melanoma metastasis showed a remarkable reduction in the cancer comparable to the best clinical candidate [38].

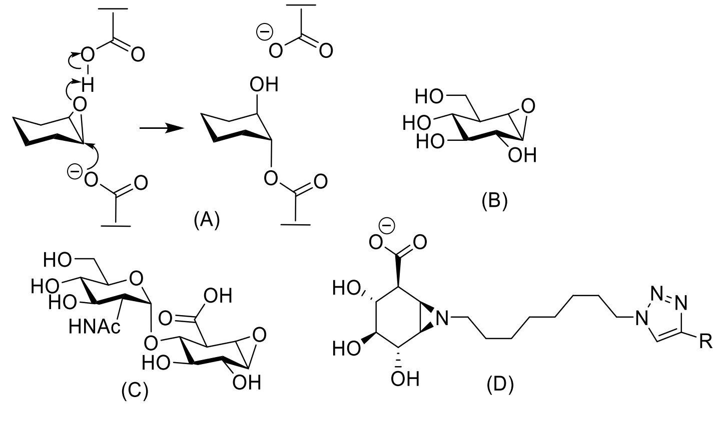


Figure 2.11- Small molecule inhibitors and ABP of HPSE with cyclophellitol epoxide mechanism.

(A) A covalent enzyme-inhibitor complex is formed through the ring opening of the epoxide via nucleophilic attack by the catalytic nucleophile. (B) 1,7-glucurono-cyclophellitol and (C) disaccharide 1,7-glucurono-cyclophellitol. (D) Activity-based probe for β- glucuronidase;

cyclophellitol aziridine with a spacer where R can be Cy5, BODIPY-TMR analogue, (BODIPY)-FL or biotin [70].

In addition to the development of small molecule inhibitors, ABP has also been developed for the quantitative visualisation of HPSE in biological samples [70]. ABP application allows for the deciphering of HPSE activity in normal and disease states; the fluorescent labelled ABP (Figure 2.11D) also targets the pre-proHPSE, allowing for the tracking of HPSE processing and how this is affected in disease [70].

Although multiple strategies of HPSE inhibition have been explored, to date, only four drugs, all of which are polyanionic oligo-/polysaccharides, have progressed to clinical trials. Both biological and small-molecule drugs have been developed with promising results, but they have yet to contend with HS mimetics [66]. Hence, the inhibition of HPSE is challenging mainly due to the extensive interactions between HPSE and substrate that are difficult to mimic [66].

2.2.5. Iminosugar inhibitors

Iminosugars are a group of carbohydrate mimetics that contain endocyclic nitrogen instead of endocyclic oxygen [71]. Owing to their potency, iminosugars have been explored as potential agents for antiviral development and inhibition of tumour metastasis [71]. Siastatin B, an iminosugar, has fascinated researchers due to its ability to inhibit HPSE and other β-glucuronidases, making it a promising candidate for tumour metastasis inhibition.

Siastatin B (Figure 2.12) is an iminosugar with a 6-acetamido-3-piperidine carboxylate structure with an N-acetyl group at the 2-position [72, 73]. Since its discovery in 1974, it was initially reported as a sialidase inhibitor but has since been shown to act as a broad-spectrum inhibitor of many GHs, including β- glucuronidases [70, 71]. The mode of binding for β-glucuronidases was unknown as siastatin B neither resembles sialic acid nor is it an N-acetylated sugar, but siastatin B can still inhibit β-glucuronidases [71]. Therefore, the mode of inhibition of siastatin B has been a mystery for a while; this chapter delves into the true mechanism of siastatin B inhibition of HPSE, representing a promising avenue for developing HPSE-targeted cancer therapies.

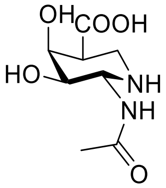
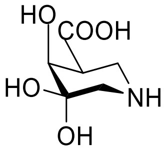


Figure 2.12- Siastatin B.

Siastatin B is a geminal diamine where an N replaces the anomeric carbon and a 2-N-acetyl group.

2.2.6. Research Aims

This work looks at determining the mode of inhibition for iminosugar siastatin B with HPSE and other glucuronidases; the *exo*-acting glucuronidase *Acidobaterium capsulatum* (AcGH79) and *endo*-acting *Burkholderia pseudomallei (BpHep)*. The mechanism of inhibition for siastatin B has long been enigmatic as the inhibitor is too big to be accommodated in the active site of HPSE. To do this, HPSE was structurally characterised with Siastatin B and the true inhibitor (*Figure 2.13*) using X-ray crystallography.



1

*Figure 2.13- Inhibitors of interest for HPSE. **1** 3-geminal diol breakdown products of siastatin B.*

Synthesised by the Overkleeft group (University of Leiden)

2.3. Materials and Methods

HPSE construct and protein was obtained from Dr Liang Wu and Dr Zach Armstrong from past work conducted in Prof. Gideon Davies group.

2.3.1. *Crystallisation of Heparanases.*

2.3.1.1. *Crystallisation Conditions with Compound 1*

HPSE- Well diffracting crystals were obtained using sitting-drop vapour diffusion from a well containing 16% w/v PEG 3350, 0.1M MES pH5.5 and 0.1M $MgCl_2$. Compound **1** was added to the well to give a final concentration of 5 mmol. Crystals were removed after 1 hour and placed in a solution of mother liquor and ethylene glycol as cryoprotectant. Data was collected at the Diamond light source to 2.1 Å, and data reduction was performed using AIMLESS, MOLREP for the molecular replacement and REFMAC5 for the model refinement and model building in ccp4mg [74, 75].

*Ac*GH79- Well-diffracting crystals were obtained using sitting-drop vapour diffusion from a well containing 1.1 M Li_2SO_4, 0.5 M $AmSO_4$ and 0.1 M trisodium citrate. Compound **1** was added to the well to give a final concentration of 5 mmol. Crystals were removed after 1 day and placed in a solution of 2 M Li_2SO_4 as a cryoprotectant. Data was collected at the Diamond light source to 2 Å; data reduction was performed using AIMLESS, MOLREP was used for the molecular replacement and REFMAC5 for the model refinement and model building in ccp4mg [74, 75].

*Bp*Hep- Well diffracting crystals were obtained using sitting-drop vapour diffusion from a well containing 12% w/v PEG6,000 and 0.1 M pH 5 citrate buffer. Compound **1** was added to the well to give a final concentration of 5 mmol. Crystals were removed after 1 day and placed in a solution of mother liquor and ethylene glycol as cryoprotectant. Data collected at the Diamond light source to 1.5 Å, data reduction performed using AIMLESS, MOLREP for the molecular replacement and REFMAC5 for the model refinement and model building in ccp4mg [74, 75].

*Table 2.2- Data collection and refinement statistics for X-ray crystallography structures with compound **1***

	HPSE-**1**	AcGH79-**1**	BpHep-**1**
Data collection			
Space group	$P2_1$	$I2_1$	$P2_1$
Cell dimensions			
a, *b*, *c* (Å)	46.8, 71.5, 79.3	82.9, 44.7, 137.02	76.5, 105.6, 114.3
β (°)	95.3	97.46	
Resolution (Å)	46.57-2.16 (2.16-2.10)	41.15-2.05 (2.05-2.00)	77.59-1.5 (1.53-1.50)
R_{merge}	0.127 (2.82)	0.055 (0.168)	0.14 (2.83)
R_{pim}	0.06 (1.2)	0.02 (0.07)	0.04 (0.79)
I / σI	10.8 (0.8)	25.1 (7.5)	9.6 (0.9)
CC1/2	0.997 (0.42)	0.99 (0.99)	0.99 (0.48)
Completeness (%)	99.7 (100)	99.8 (100)	99.7 (99)
Redundancy	6.4 (6.6)	6.4 (6.7)	13.2(13.5)
Refinement			
Resolution (Å)	46.62-2.10	41.18-2.00	63.65-1.50
No. reflections	30423	33995	147634
R_{work} / R_{free}	0.20/0.25	0.19/0.24	0.18/0.21
No. atoms			
Protein	3600	3487	6282
Ligand/ion	97/3	17/0	
Water	135	441	787
B-factors			
Protein	45.35	19.14	19.86
Ligand/ion	64.22/54.57	24.78/0	
Water	43.06	28.35	29.78
R.m.s. deviations			
Bond lengths (Å)	0.016	0.014	0.015
Bond angles (°)	2.00	1.88	1.82
Ramachandran			
Favored	96.7	96.93	97.24
Outliers	0.00	1.45	0
Wavelength	0.98	0.98	0.98

2.4. Results and Discussion

2.4.1. Previous Work

As mentioned previously, siastatin B has been shown to have efficacy in inhibiting HPSE and other β- glucuronidases, but its mode of inhibition has long been enigmatic. The bulkiness of the N-acetyl group implies that Siastatin B would not be accommodated in the small HPSE active site. Furthermore, the stereochemistry of siastatin B indicates that the inhibition of HPSE occurs unexpectedly due to its “galacto” configuration. At the same time, the natural substrate has a “gluco” configuration [71].

To establish the mode of siastatin B inhibition, Davies group and collaborators sought to determine if siastatin B was the true inhibitor (Appendix 7.3.1). Previous work had shown that upon soaking siastatin B with HPSE (and homologs), the processed structures revealed that siastatin B was not present in the active site, but instead, the inhibitory species were different derivatives of the iminosugar (as shown in PDBS: 8OHQ and 8OHT). The goal of the work was to investigate whether these inhibitory species were contaminants or degradation products of Siastatin B. Our approach to solve this was to study chemically synthesised siastatin B derivatives, establish their activity, and analyse their binding.

Siastatin B has always been regarded as a stable compound, but the creation of a trifluoroacetamide analogue of siastatin B offered insight into the mechanism of inhibition of siastatin B with β-glucuronidases [76, 77]. NMR analysis of the trifluoroacetamide analogue showed that they undergo pH-dependent decomposition into the potentially true inhibitors [78]. This was not shown to occur when siastatin B was analysed with NMR; however, high-resolution mass spectrometry showed minor contaminants which m/z corresponded to the noeuromycin-like breakdown product **2** and the 3-geminal diol (3-GDI) **1** (Figure 2.14). The same breakdown products were observed in the NMR of the trifluoroacetamide analogue [76, 77]. As previously mentioned, the active sites of HPSE and AcGH79 did not contain siastatin B or the same contaminant. The Overkleeft group synthesised compounds **1** and **2**, which were then investigated for their potency and assessed using X-ray crystallography.

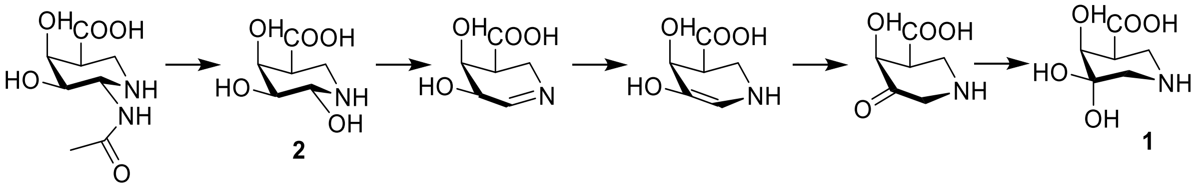
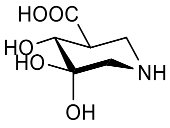


Figure 2.14- Proposed breakdown products of siastatin B into inhibitors of HPSE.

*Breakdown into galacturonic-noeuromycin (**2**) and 3-geminal diol iminosugars (**1**).*

The newly synthesised compound **1** was investigated for its potency on the *exo*-acting *Ac*GH79 and a GH2 family *E.coli* β-glucuronidase (*Ec*GusB) using a fluorogenic 4-methylumbelliferyl glucuronide (Figure 2.15) [71]. The breakdown *galacto*-configured **1** product observed in HPSE inhibited *Ac*GH79 (K_i =5.8 *± 0.5* µM) and *Ec*GUSB (137 *± 3* µM). A *glucose*-configured 1-*N*-iminosugar was also synthesised (**1G**) with the hope that the change in stereochemistry at the 4-position would better resemble the natural substrate and enhance inhibition (Figure 2.15). The *glucose*-configured 3-GDI (**1G**) was more potent than the *galacto*-configured as expected *Ac*GH79 (0.520 ± 0.030 nM) and *Ec*GUSB (28 ± 1 µM). The inhibitors were also examined using activity-based probe profiling (ABPP) competitive assay to measure the ability of the inhibitor to outcompete the probe, resulting in reduced fluorescence (Figure 2.15)[71]. Surprisingly, the *galacto*-configured 3-GDI (**1**) (27 *± 3* µM) was more potent than the *gluco*-configured 3-GDI (**1G**) (200 ± 80 µM) with HPSE, and all were more potent inhibitors for human GUSB (HsGUSB) (Figure 2.15); this is expected as they are monomer-like inhibitors acting on a multi-subsite *endo*-acting HPSE. The synthetically produced inhibitor **1** was more potent than previously reported for siastatin B with HPSE and HsGUSB [79]. Due to the promising results, I sought to characterise HPSE and other β-glucuronidases with the 3-GDI (**1**) so that they could be compared with previously solved structures produced by past members of the Davies group.



1G

A

Inhibitor	AcGH79 K_i (μM)	EcGusB K_i (μM)
1	5.8 ± 0.5	137 ± 3
1G	0.520 ± 0.030	28 ± 1

B

Inhibitor	HsGUSB IC_{50} (μM)	HPSE IC_{50} (μM)
1	4 ± 2	27 ± 3
1G	1.7 ± 0.2	200 ± 80

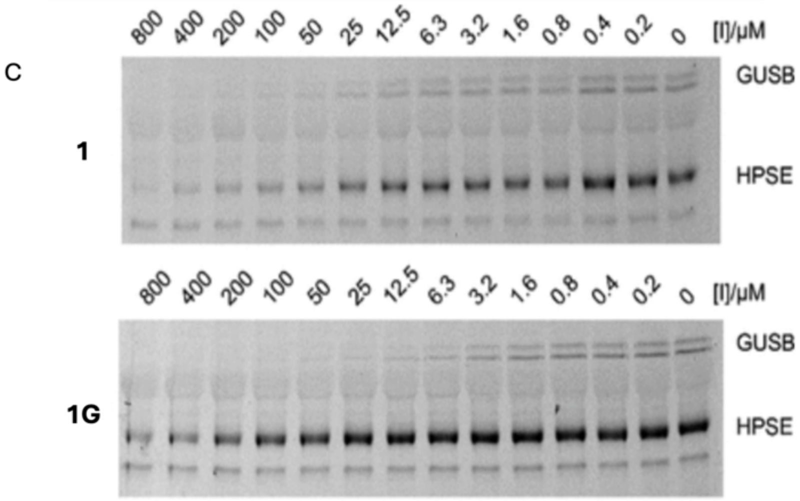


Figure 2.15- Competitive inhibition of Compound ***1*** *with Activity-based probe.*

Figure from (Appendix 7.3.1). ***1G*** *is the gluco-configured* ***1.*** *(A) Inhibition constants for compounds* ***1*** *and* ***1G*** *with AcGH79 and EcGUSB using 4-methylumbelliferyl glucuronide. (B) Inhibitory potency on platelet lysate containing HsGUSB and HPSE, derived from the competitive ABPP gel. (C) Competitive inhibition gels for HsGUSB and HPSE with compound* ***1*** *using platelet lysate. The fluorescence was quantified and used to determine IC50s. Where Compound* ***1*** *is bound, the bands are less visible as the ABP was unable to bind in the already occupied active site. This data was collected by a collaborator in the Overfkleeft group.*

2.4.2. *β-glucuronidase Structures with Siastatin B 3-Geminal Diol (3-GDI) Breakdown Product*

Previous structures of siastatin B showed a 3-GDI breakdown product present in the HPSE active site and a noeuromycin-like breakdown product found in *Ac*GH79 and *Bp*Hep. HPSE and

*Bp*Hep are *endo*-acting β-glucuronidases, and *Ac*GH79 is *exo*-acting; they have the same residues in the -1 subsite, as this is conserved among the GH79 family [34]. Although HPSE and *Bp*Hep are similar, they hydrolyse different HS residues, where HPSE favours GlcNS residues over GlcNAc for the *Bp*Hep [34]. The difference in substrate specificities between the three enzymes suggests why breakdown product **1** was only observed in the HPSE active site. Soaks were carried out to see if **1** can react with *Ac*GH79 and *Bp*Hep crystals and if similar interactions occur.

For HPSE with **1** (Figure 2.16), the NH group forms H-bonds with the catalytic acid/base Glu225 and nucleophile Glu343, where **1** is bound in a $^4C_1$ ring conformation at 2.1 Å. HPSE forms additional H-bonds with three active site waters, which create a network of interactions from the axial 3-OH to water molecules for **1** interactions. Hydrogen bonding occurs between D27, T62, E225, Y298, E343, G349, Q383 and Y391. The two additional bonds formed between 3-OH, Gln383 and Asp27 in HPSE may be why **2** was favoured in the bacterial homologs compared to **1**. The observation of 3-GDI with HPSE represents a new class of glycoside hydrolase inhibitors.

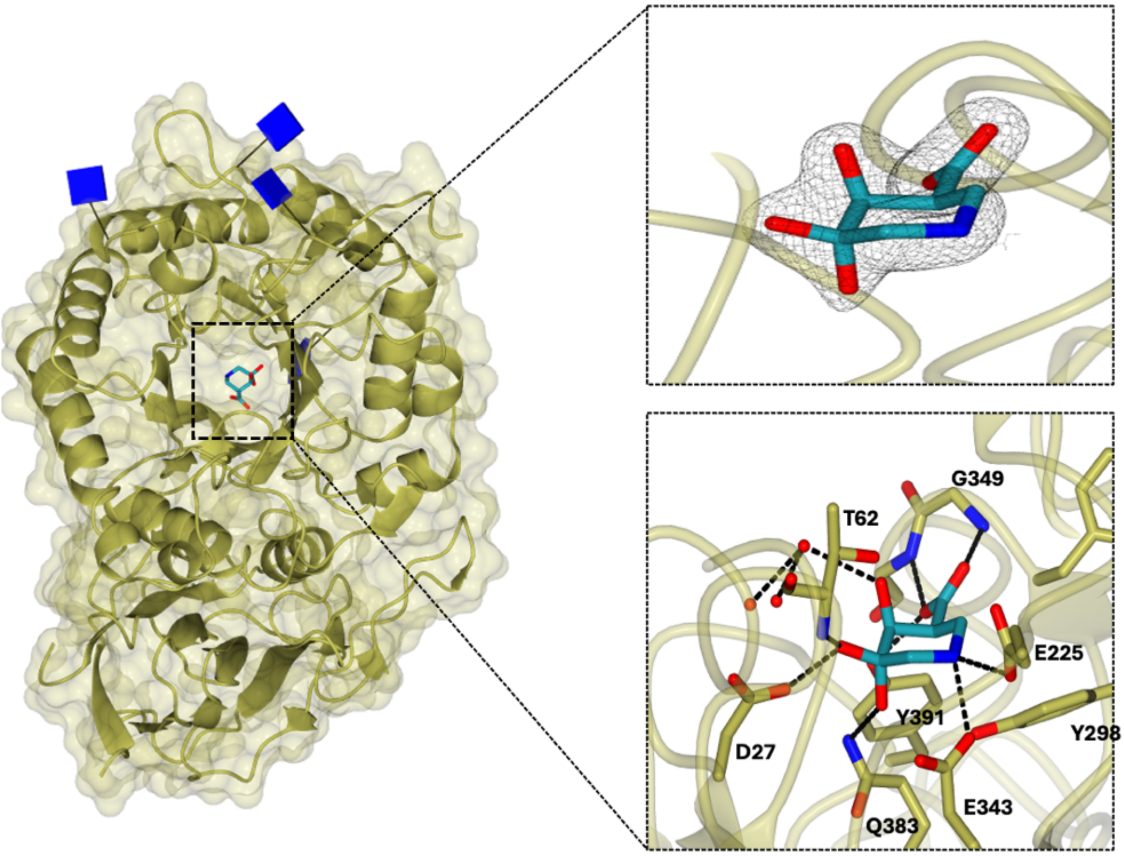


*Figure 2.16- Structure of HPSE with 3-geminal diol iminosugar non-covalent interactions in the active site***.**

Full length $(\beta/\alpha)_8$ TIM-barrel structure of HPSE (left panel) with compound **1** (blue) bound in the binding cleft (right panel), enhancing of that region shows the extensive H-bond interactions between the compound and active site residues as well as water molecules, refined to 2.1 *Å*. *Glycosylation NAG is found at Asn 162, 200, 217, 238, and 459.* Images created in CCP4mg (version 2.11.0) with F_o-F_c *maps (grey) have been displayed in chicken wire and contoured to 0.4* $e^-/Å^3$

Although *Ac*Gh79 and *Bp*Hep active sites contained **2** in the active site when soaked with siastatin B, it can also accommodate the 3-GDI **1** (*Figure 2.17*). *Ac*GH79 interacts with **1** through hydrogen bonding with E45, N80, D105, E173, H237, E287, G294 and G295. There is no coordination of waters with **1**, but there is an additional bond with the carbonyl of Asp105 and with water.

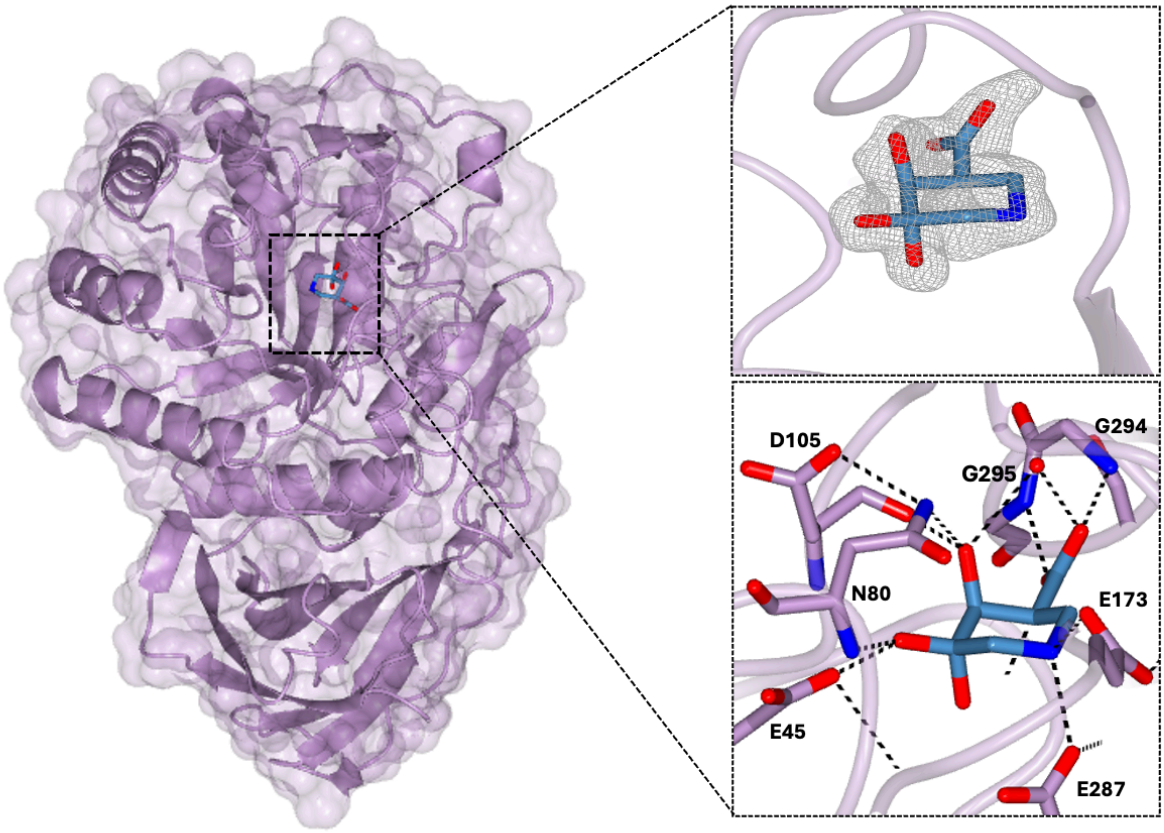


Figure 2.17- Structure of 3-geminal diol iminosugar non-covalent interactions in the AcGH79 active site.

Full-length structure of AcGh79 (left panel) with compound ***1*** *(blue) bound in the active site, along with a zoomed-in portion of the binding site (right panel), refined to 2.0 Å.* Images created in CCP4mg (version 2.11.0) with *F_o-F_c maps (grey) have been displayed in chicken wire and contoured to 0.3 $e^{-}/\text{Å}^{3}$.*

BpHep forms similar hydrogen bonds as AcGH79; E36, N69, E144, Y215, E255, G261, G262, H295, and Y302 (*Figure 2.18*). However, the aspartic acid interaction is not present in BpHep, although **1** still maintains the $^{4}C_{1}$ conformation and refined to 1.5 Å.

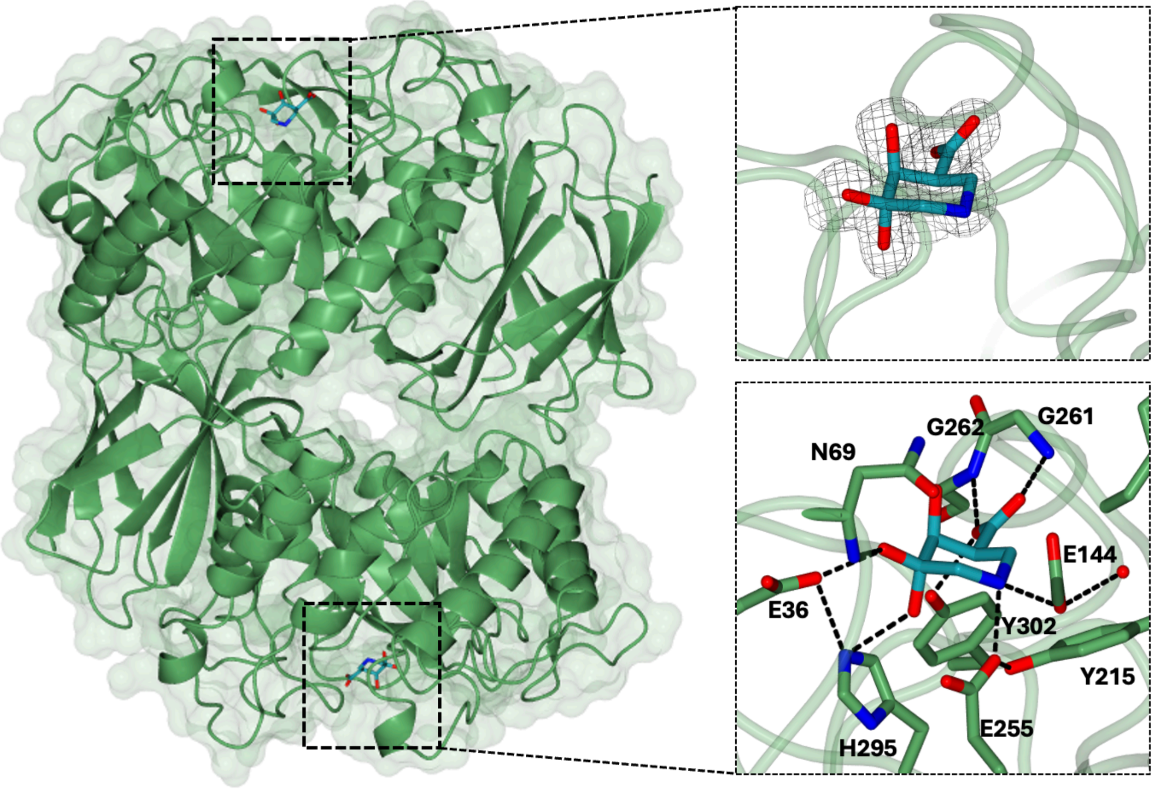


Figure 2.18- Structure of 3-geminal diol iminosugar non-covalent interactions in the BpHep active site.

Full-length structure of BpHep (left panel) with compound ***1*** *(blue) bound in the active site, along with a zoomed-in portion of the binding site (right panel), refined to 1.5 Å.* Images created in CCP4mg (version 2.11.0) with F_o-F_c *maps (grey) have been displayed in chicken wire and contoured to 0.35* $e^-/Å^3$. *Two molecules are contained within the unit cell.*

2.5. Summary

This work uncovered the mode of siastatin B inhibition with β-glucuronidases. Commercial siastatin B was initially analysed with high-resolution mass-spectrometry, where the true β-glucuronidase inhibitors were revealed (**1** and **2**). Surprisingly, inhibitory assays conducted with these purified breakdown-product peaks were less potent than commercialised siastatin B when tested on *exo*-acting *Ac*GH79 and *Ec*GusB [71]. Regardless, the 1-*N*-iminosugars were synthesised (**1** and **2**) and assessed kinetically and structurally to dissect their mechanism of inhibition [71]. Siastatin B breakdown products **1** and **2** were previously observed as breakdown products for the siastatin B derivative 2-trifluoroacetamido produced by Nishimura et al Figure 2.19 [76-78].

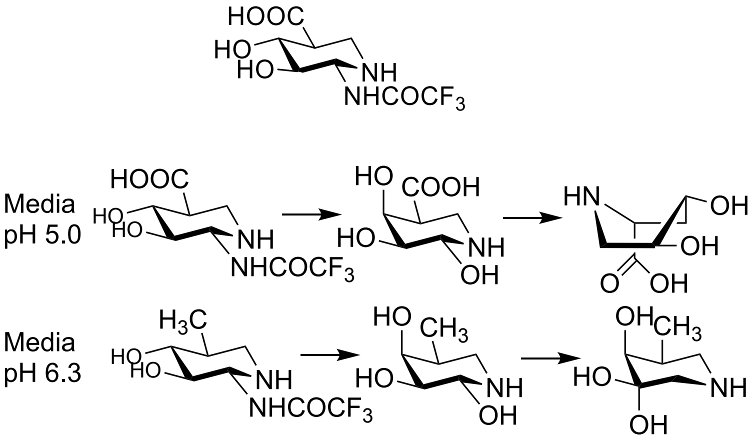


Figure 2.19- 2-trifluoroacetamido synthesised by [78].

2-trifluoroacetamido undergoes structural changes when in media pH 5.0-6.3 and also potentially inhibits glycosidases as well.

Although the 2-trifluoroacetamido derivatives were more potent than siastatin B derivatives, this is likely due to the increased elimination of the 2-trifluoroacetamide compared to the 2-*N*-acetyl group [71]. This rapid decay can be exploited to design prodrug-like inhibitors, where the compound is decayed into the pharmacologically active drug, allowing for the control of inhibitory activity and reducing off-target specificity. This typically requires the compound to be acted upon by another physiological enzyme. Optimisation of the *N*-acetyl at the 2-position of siastatin B for a better leaving group could generate a more rapidly decaying derivative [71].

Overall, the data presented has unveiled how siastatin B can act as a broad-spectrum inhibitor for glycoside hydrolases with different substrate specificities. Analysis of the inhibitors using X-ray crystallography showed that they are well accommodated in the HPSE active site and had improved potency compared to siastatin B. Collectively, the results will enable the future design of broad-spectrum inhibitors that target other GHs, beyond those targeted by siastatin B.

3. Chapter 3- Inhibitors for Endoplasmic Reticulum α-Glucosidases: Broad Spectrum Antivirals

3.1. Abstract

The emergence of treatment-resistant viruses is a significant health threat due to their risk of epidemic/pandemic potential [80]. As highlighted in the COVID-19 pandemic, vaccine development takes time; hence, an arsenal of antivirals would be helpful for urgent treatment in a health crisis. The development of broad-spectrum antivirals has been explored, as viruses rely on host cell machinery for their replication. Inhibition of host proteins by broad-spectrum antivirals would be beneficial as the virus has evolved to rely on mammalian systems and have no genetic control for developing resistance [80].

This chapter focuses on α-glucosidases I and II (ER-αGluI and ER-αGluII) in the endoplasmic reticulum (ER), where they are key proteins in the ER quality control pathway. These gatekeepers of glycoprotein folding are validated targets for broad-spectrum inhibition, as individuals with genetic defects in ER-αGluI showed no prior history of enveloped virus infection [81]. Concerning this, 1,6 and 1,2 cyclitols inhibitors were synthesised by collaborators, where I then sought to express soluble active ER-αGluI and ER-αGluII recombinantly using baculovirus and mammalian systems[B,C]. Inhibitors of ER-αGluI/ER-αGluII were analysed using thermal shift, ABP, viral reduction assays and docking studies were conducted. The most potent of the tested inhibitors was the 1,6-*epi*-cyclophellitol cyclosulfate, where it was shown that the inhibitor potential blocks the maturation of SARS-CoV-2 spike protein through inhibition of ER-αGluII, reducing the progeny virus infectivity. The reported antiviral activity of 1,6-*epi*-cyclophellitol cyclosulfate was comparable to other best-in-class α-glucosidase inhibitors. Here, we report a new strategy for host-targeted antivirals, which irreversible blocks ER-αGluII activity in a covalent manner.

B. Melissa Thaler, Tim P. Ofman, Ken Kok, Jurriaan J. A. Heming, Elisha Moran, Isabelle Pickles, Anouk A. Leijs, Adrianus M. C. H. van den Nieuwendijk, Richard J. B. H. N. van den Berg, Gijs Ruijgrok, Zachary Armstrong, Clarisse Salgado-Benvindo, Dennis K. Ninaber, Eric J. Snijder, Constant A. A. van Boeckel, Marta Artola, Gideon J. Davies, Herman S. Overkleeft, Martijn J. van Hemert. *ACS Cent. Sci.* **2024** 2374-7943

C. Tim P. Ofman, Jurriaan J. A. Heming, Dr. Alba Nin-Hill, Dr. Florian Küllmer, Elisha Moran, Megan Bennett, Roy Steneker, Anne-Mei Klein, Gijs Ruijgrok, Ken Kok, Zach W. B. Armstrong, Johannes M. F. G. Aerts, Gijsbert A. van der Marel, Carme Rovira, Gideon J. Davies, Marta Artola, Jeroen D. C. Codée, Herman S. Overkleeft. *Chem. Eur. J.* **2024**, 30

3.2. Introduction

The increasing level of antiviral resistance is becoming a global health risk. The growing number of viral outbreaks and the COVID-19 pandemic have highlighted the concern of time constraints when developing treatments [80]. Major viral epidemics in the 21st century include Ebola, severe acute respiratory syndrome (SARS), Influenza A (H1N1) and the recent COVID-19 pandemic [82]. Several approved antiviral agents targeting other viruses were repurposed for COVID-19 treatment. This included lopinavir, a nucleoside analogue used in the treatment of HIV, and numerous influenza antivirals such as zanamivir, oseltamivir, and peramivir [82]. Still, these drugs have side effects such as neurological toxicity [82]. However, even the repurposing of clinically approved antivirals still requires funding and time to investigate the specificity and toxicity of a new target; hence, approving a broad-spectrum antiviral would be beneficial.

Antiviral strategies have primarily focused on targeting a specific protein (direct-acting), but some antivirals have become ineffective due to the emergence of resistant strains. Although direct-acting antivirals have the advantage of reduced off-target specificity and increased efficacy, developing broad-spectrum antivirals would revolutionise the treatment of viral infections [80]. An effective strategy for targeting viruses by broad-spectrum antivirals would be using nucleoside analogues, as viruses replicate using DNA or RNA polymerases [80]. One example includes remdesivir, a C-nucleoside analogue, which was used in the treatment of the Ebola outbreak in 2013 and has shown efficacy in treating MERS-CoV and SARS-CoV-2 where the sugar analogue acts as a chain terminator [80, 83-85]. The above is a broad-spectrum virus-targeted antiviral; other broad-spectrum inhibitors are host-targeted antivirals that target the cellular machinery viruses rely upon for replicating (Figure 3.1).

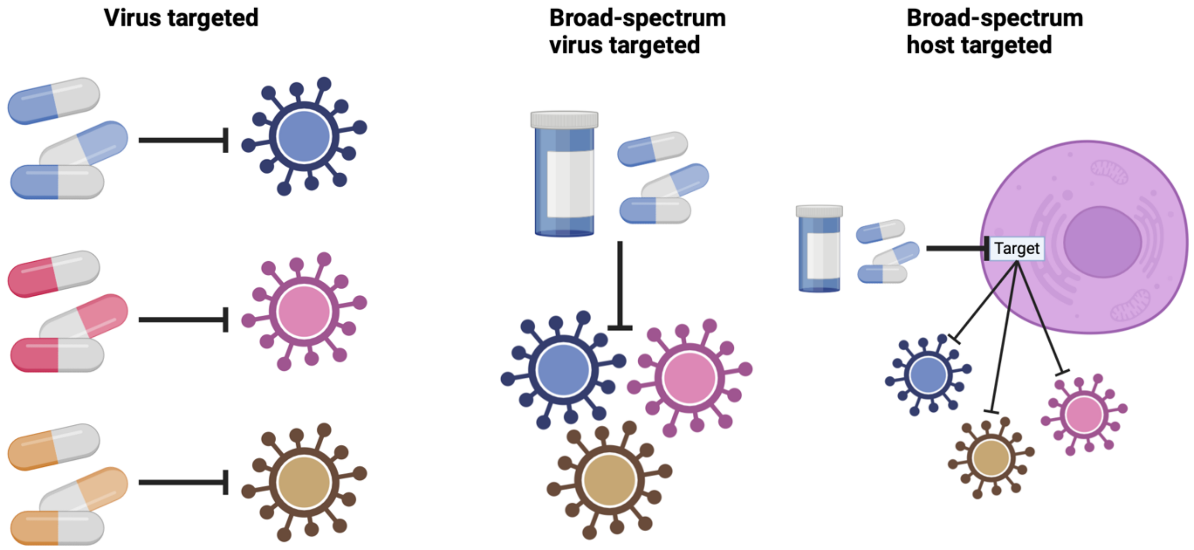


Figure 3.1- Types of antivirals (Image created in BioRender).
Virus-targeted antivirals (Direct-acting) are inhibitors that target a specific protein associated with a particular virus. Broad-spectrum acting are inhibitors that target viral processes/proteins that multiple viruses use in the replication cycle. Broad-spectrum host-targeted inhibitors target a mammalian protein essential for replicating or assembling various viruses. Image-based on [86].

The cellular machinery used in viruses' replication lifecycles is conserved across viral families, which have evolved to become reliant upon host cell proteins. This makes them appealing targets for broad-spectrum inhibitor development. Additionally, this expands the target repertoire as the human proteome encodes an estimated 20,000 proteins [86]. Moreover, compared to direct-acting antivirals, the barrier to antiviral resistance is substantially higher as the virus has evolved to utilise these mammalian pathways, and they have no genetic control [86]. Broad-spectrum inhibitor examples can include anticancer drugs, which target kinases used in the replication cycle of the Hepatitis C virus, Ebola and Dengue [87]. Cultured cells and murine models showed that the use of anticancer drugs sunitinib/erlotinib protected against morbidity and mortality through their inhibition of these kinases that are essential for these viruses' lifecycles [87]. Other antiviral targets include tyrosine kinase receptors and α-glucosidases (Figure 3.2) [87, 88]. The latter mentioned will be the subject of this chapter.

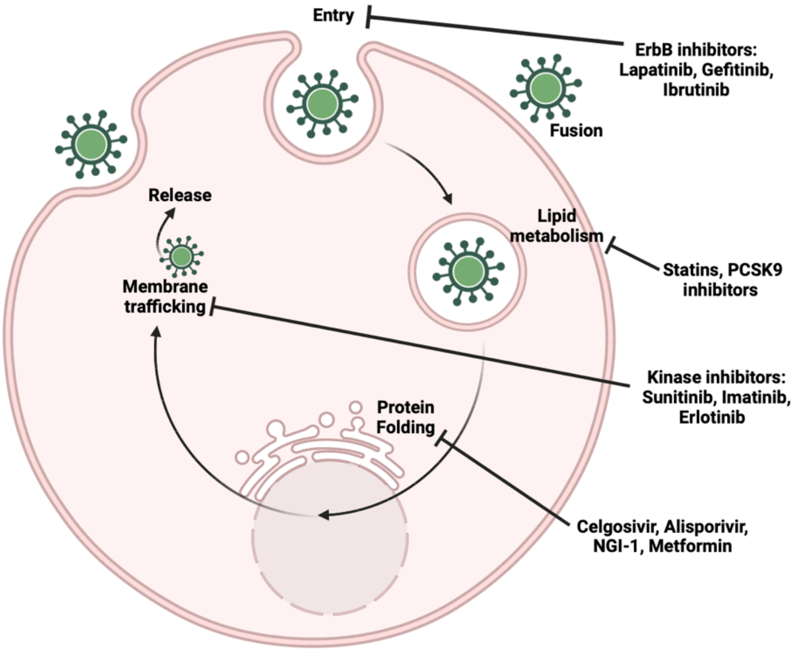


Figure 3.2- Experimental broad-spectrum host targeted inhibitor examples (Image created in BioRender).

Depiction of viral lifecycle and the cellular processes that can be targeted for inhibition. Image-based on [86].

3.2.1. Endoplasmic Reticulum Quality Control Pathway

N-linked oligosaccharides have various functions, as discussed in **Chapter 1**. Cell quality control pathways, such as the endoplasmic reticulum-associated quality control (ERQC) and endoplasmic reticulum-associated degradation (ERAD), regulate the N-linked glycosylation of glycoproteins at the luminal interface of the ER [89].

Nascent glycoproteins are directly translated into the ER by co-translational translocation via a ribosome. A preassembled N-linked glycan core ($Glc_3Man_9GlcNAc_2$) is transferred onto a nitrogen side chain of Asn from an Asn-X-Ser/Thr motif of the newly synthesised protein [89-91]. The N-linked oligosaccharide is processed by the sequential cleavage of terminal glucose by the resident ER proteins: ER-αGluI and ER-αGluII (Figure 3.3) [89]. Firstly, ER-αGluI cleaves the distal glucose residue (α1,2-glucose), producing $Glc_2Man_9GlcNAc_2$ [89]. ER-αGluI is the gatekeeper of the ERQC, as without its activity, proteins would not enter the pathway. Its catalytic activity produces the substrate for the following enzyme: ER-αGluII [92]. Generation of the $Glc_2Man_9GlcNAc_2$ allows the glycoprotein to associate with the ER-membrane-bound lectin or undergo further processing by ER-αGluII [89]. The monoglucosyalted glycan ($Glc_1Man_9GlcNAc_2$) production by ER-αGluII allows for its recognition by ER-resident chaperones (Calnexin/calreticulin) [89, 92]. Calnexin and calreticulin contain binding domains that recognise the monoglucosyalted glycan; the chaperones bind and recruit ATP, calcium, and protein disulfide isomerases to facilitate the proper folding of the glycoprotein [93]. Once the glycoprotein has appropriately folded, the last glucose residue is trimmed by ER-αGluII, producing $Man_9GlcNAc_2$, causing the disassociation of chaperones [92, 93]. These proteins are trafficked into the Golgi by trimming mannose by ER mannosidase I into $Man_8GlcNAc_2$, where the glycoprotein can bind to the lectin-like cargo receptor complex [93]. If the protein is not folded correctly, the misfolded protein is reglucosylated by UDP-glucosyltransferase for further attempts of refolding in the calnexin/calreticulin cycle [89, 92]. Glycoproteins that fail to fold are processed by mannosidases and targeted to ERAD.

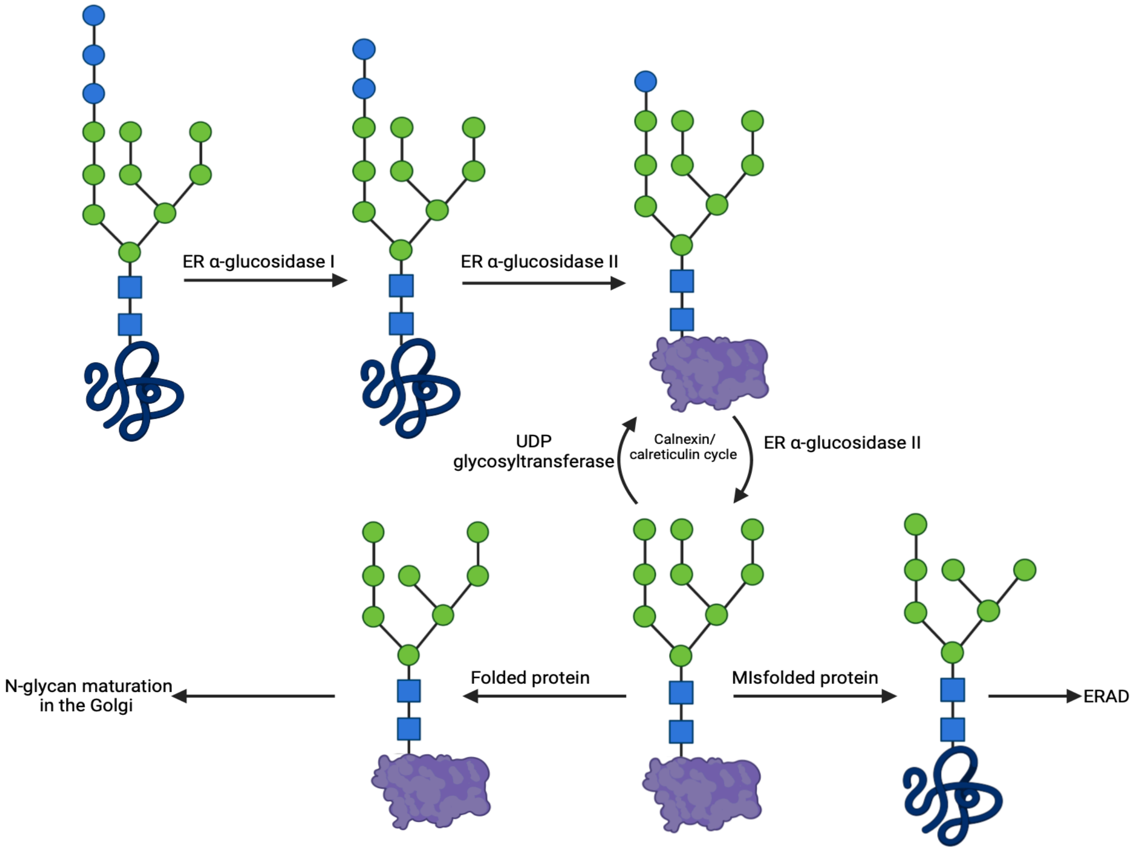


Figure 3.3- Endoplasmic reticulum quality control pathway.

Removal of the N-linked glycan is processed by ER-αGluI and ER-αGluII; once the protein has folded correctly, the protein is transported to Golgi, whereas improperly folded proteins enter the ER-associated protein degradation pathway (ERAD). Image based on [93] and made in BioRender.

The machinery of the ERQC is not only required to generate N-linked glycoproteins for resident mammalian proteins but is a critical pathway in the lifecycle of diverse viruses [89]. Viruses do not encode their glycosylation machinery [89, 93]. Instead, they hijack mammalian pathways for viral protein glycosylation. Glycosylation of viral proteins is essential for virion assembly and the function of their glycoproteins. This validates ERQC proteins as potential targets for host-directed broad-spectrum antivirals [92, 93].

3.2.2. *Mechanism and Structure of ER α-Glucosidases I*

ER-αGluI has a critical regulatory position in the ERQC, where it catalyses $Glc_3Man_9GlcNAc_2$ to $Glc_2Man_9GlcNAc_2$, with a terminal glucose-α(1,2)glucose specificity [94]. The GH63 family catalyses through an inverting mechanism in which the anomeric centre of the product is inverted through a nucleophilic attack by water (Figure 3.4)[12].

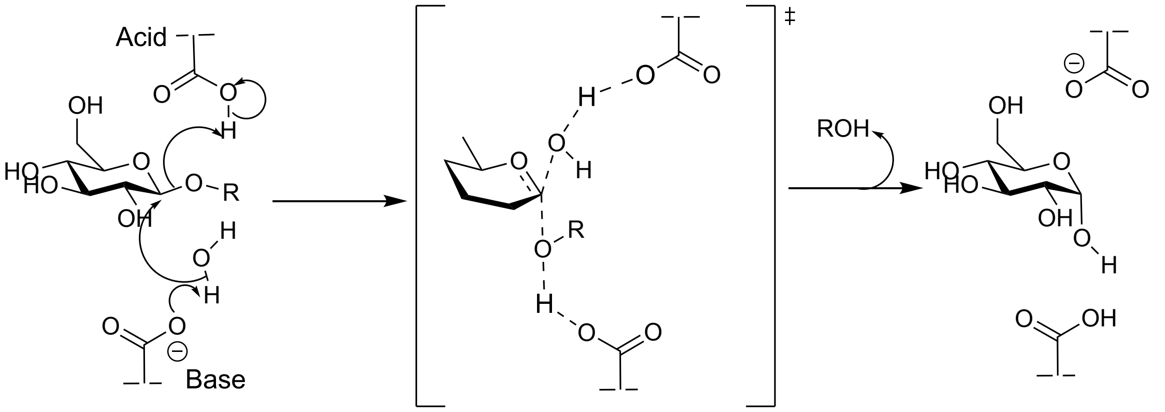


Figure 3.4- Inverting mechanism of Glycoside Hydrolase

Inversion of the anomeric centre through the protonation of the leaving group, typically by a glutamic or aspartic acid (general acid) and a second residue, again usually a glutamic or aspartic acid, acts as the general base to deprotonate the water molecule for nucleophilic attack.

The first structure of ER-αGluI was for *Saccharomyces cerevisiae (S. cerevisiae)*(Table 3.1)*. S. cerevisiae* shares substrate, inhibitors, and mechanisms similar to the human enzyme [94]. Most ER-αGluI is located in the ER lumen, a type II transmembrane protein with a C-terminus catalytic domain ($(\alpha/\alpha)_6$ barrel) and membrane-anchored N-terminus (Figure 3.5) [94, 95]. Although *S. cerevisiae* is an ideal model to study, the structure revealed that it was not a good model for studying inhibitor complexes as the C-terminal His-tag from the adjacent protein occludes the active site of another in the crystal lattice [94]. Removing the tag abolished crystallisation; instead, the *Mus musculus* (*Mm*) ER-αGluI structure was determined to study inhibitor complexes, but groups have reported difficulty repeating the complex crystallisation conditions [95, 96]. Instead, recently, the thermophilic fungus *Chaetomium thermophilum (Ct*) was investigated as a potential ER-αGluI crystallisation model [95]. Alignment of *Ct*, *Mm* and *S. cerevisiae* ER-αGluI revealed the conserved catalytic residues [95]. From predictions, the conserved active site residues are (with *Ct* numbering) F417, F421, W423, D424, P472, D586, Y725, Y726 and E783 [95]. This makes *Ct* a suitable substitute for inhibitor complex studies as they share the conserved catalytic residues of the GH63 family enzymes [95].

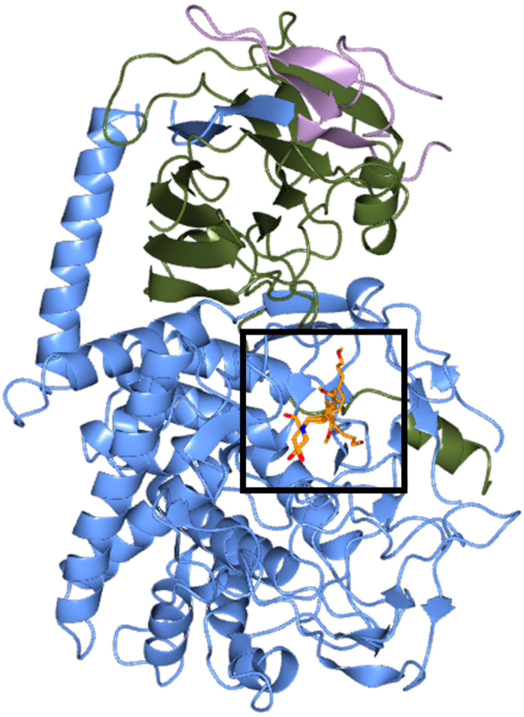


Figure 3.5- Structure of murine ER with N-9'-methoxynonyl-1-deoxynojirimycin.

The N-terminal is highlighted in green, and the C-terminal in blue. The active site is displayed within the black box where N-9'-methoxynonyl-1-deoxynojirimycin is bound in the active site with catalytic residues D580/E804. (PDB 5MHF), structure from [97].

Table 3.1- PDB structures of ER-αGluI

PDB{Warfield, 2020 #16}	**Organism**	**Ligand**	**Res (Å)**	**Reference**
5MHF	*Mus musculus*	N-9'-methoxynonyl-1-deoxynojirimycin	2.1	Warfield, et al (2020)
4J5T	*Saccharomyces cerevisiae*		2.04	Barker et al, (2013)
7T66, 7T68, 7T6W, 7T8V	*Chaetomium thermophilum*	"Apo" and co-crystal structures	2.19 -2.6	Karade et al (2022)
7R6J, 7RD2, 7REV, 8E3J, 8E3P, 8E4I, 8E4K, 8E4Z, 8E5U, 8E6G,8ECW, 8EGV, 8EHP, 8EID, 8EKN, 8ELE, 8EPJ, 8EPO, 8EPR, 8EQ7, 8EQX, 8ER4, 8ETL, 8ETO, 8EUD, 8EUR, 8EUT , 8EUX	*Chaetomium thermophilum*	With 25 DNJ and valiolamine derivatives	1.91 - 2.80	Karade et al (2023)

3.2.3. Mechanism and Structure of ER α-Glucosidases II

ER-αGluII is a part of the GH31 CAZy protein family. In searching for a soluble ER-αGluII, the *Mm* construct had the best sequence similarity to the human enzyme, with 92% identity to the α subunit (GANAB) and 87% to the β subunit (PRKSCH) (Table 3.2) [98].

Table 3.2- PDB structure of ER α-GluII

PDB	**Organism**	**Ligand**	**Res (Å)**	**Reference**
5F0E, 5H9O, 5HJO, 5HJR, 5IED, 5IEE, 5IEF, 5IEG	*Mus musculus*	"apo" and co-crystal structures with iminosugars	1.74-2.38	Caputo et al (2016)
7JTY, 7K9N, 7K9O, 7K9Q, 7K9T, 7KAD, 7KB6, 7KB8, 7KBJ, 7KBR, 7KRY, 7L9E	*Mus musculus*	Co-crystal structure with Valiolamine derivatives	2.07-2.55	Karade et al (2021)

ER-αGluII is a heterodimer (Figure 3.6) localised to the ER through an ER retention signal on the 60 kDa β subunit [99]. The enzyme cleaves two α-1,3-glucose residues sequentially to yield a Glc$_1$Man$_9$GlcNAc$_2$. The full-length structure of ER-αGluII has never been solved; to crystalise the protein, a trypsin digest is performed where most of the β subunit is removed [98]. The β subunit is presumed responsible for localising the heterodimer to the ER and contains high-affinity calcium-binding loops [100]. A recent cryo-EM structure of non-trypsinised ER-αGluII revealed that the regions that are removed by the trypsin digest are disordered, and were unable to be solved; these regions have been predicted in AlphaFold but have poor confidence due to the disorder (Figure 3.6) [60].

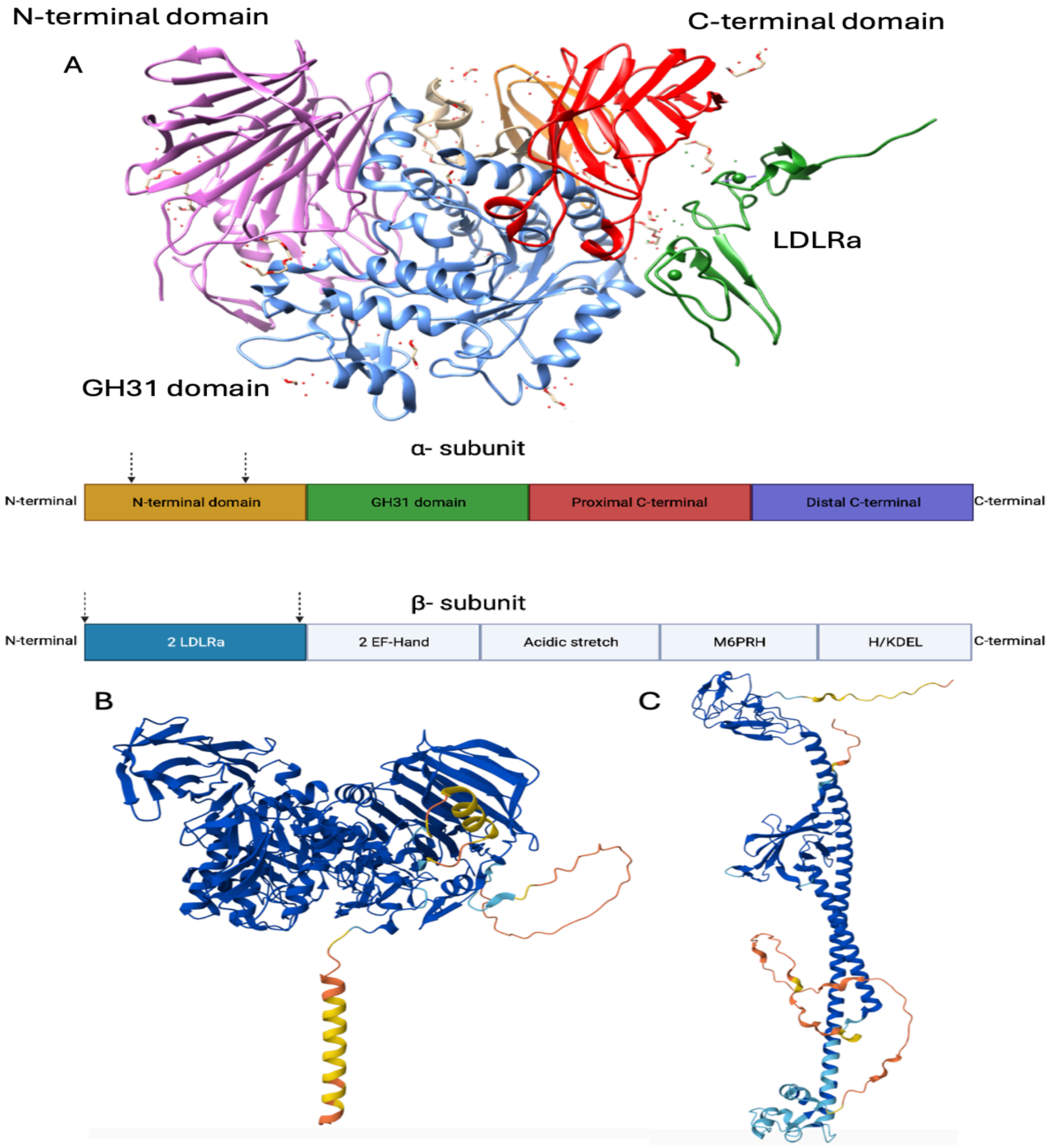


Figure 3.6- Crystal structure of Mm ER-αGluII (PDB 5F0E).

The structure of trypsinised ER-αGluII, N-terminal domain 33-392 (Purple), catalytic β8α8 barrel GH31 active site (blue), insertion subdomain (orange) and C-terminal (red) for the *α subunit and the N-terminal domain of the β subunit (green), which adopt two low-density lipoprotein receptor (LDLR) folds* [101]*. The regions that are removed during the trypsin digest are indicated by the arrows, with the majority of the PRKSCH β- subunit removed. (B, C) show the alpha fold predictions for UniProt accession nos. Q8BHN3-2 and O08795, the model has high confidence (blue regions) as these parts were previously solved by X-ray crystallography, other regions (yellow and orange) have low confidence as these regions are disordered.*

Crystal structures of ER-αGluII with iminosugar inhibitors such as DNJ show that the inhibitors bind in a $^4C_1$ chair conformation; this is in line with the conformation of the Michaelis complex for GH31 α-glucosidases; $^4C_1 \rightarrow {}^4H_3 \rightarrow {}^1S_3$ (Glycosylation itinerary: Michaelis complex→ TS→ covalent intermediate) [16]. Retaining glycosidases employ the Koshland double displacement mechanism where GH31 α-glucosidases hydrolyse acetals in the glycosidic linkage with the catalytic nucleophile D564 and acid/base D640 in *Mm* ER-αGluII (Figure 3.7) [102]. ER-αGluII has a $\beta_8\alpha_8$ barrel active site shared with other GH31 enzymes. Unique to ER-αGluII is the α N-terminus, which supports the 305-317 catalytic loop, creating the conserved GH31 +1 and -1 subsites [98]. The +2 and +3 subsites are less conserved compared to other GH31 enzymes [98].

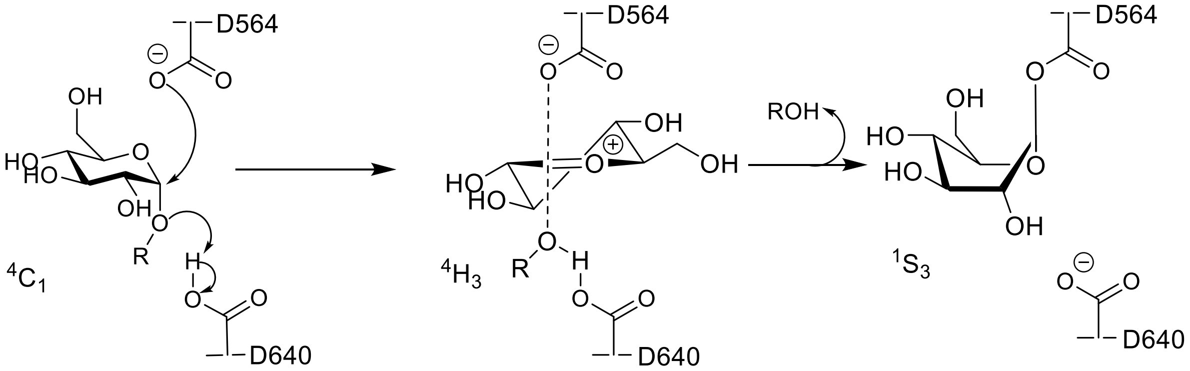


Figure 3.7- ER-αGluII retaining mechanism.

The substrate binds and forms the Michaelis complex $^4C_1$, where the leaving group is pseudo-axially positioned. Nucleophilic attack (D564) occurs along with the protonation of the glycosidic oxygen, creating a high-energy transition state and covalent intermediate $^1S_3$ [102].

The central role of ER-αGluI and ER-αGluII in the ERQC makes them appealing targets for designing broad-spectrum inhibitors. However, the iminosugars that target ER-αGluII are also effective against other GH31 human intestinal enzymes, so determining the full-length structure of ER-αGluII is vital for understanding the molecular basis of binding and designing specific inhibitors [98].

3.2.4. *α-Glucosidase Role in Disease*

ER-αGluI is the gatekeeper enzyme to the ERQC; the protein is also called the mannosyl-oligosaccharide glucosidase (MOGS) [103]. MOGS is implicated in the hereditary congenital disorders of glycosylation (CDG) known as MOGS-CDG. There are over 150 types of CDG, most of which affect N-glycosylation pathways [103]. Individuals with MOGS-CDG have deleterious mutations causing frameshifts in the MOGS gene [103]. Not only does the disorder result in the impaired catalytic activity of ER-αGluI, but individuals also display developmental delay, hypogammaglobulinaemia and dysmorphic features [103]. A study conducted on two siblings displaying severe hypogammaglobulinaemia due to MOGS-CDG were evaluated for their susceptibility to infectious disease [104]. The individuals had many health complications due to MOGS-CDG but displayed little history of viral infections [104]. Viral assays carried out on cells derived from the individuals revealed no change in susceptibility to non-enveloped viruses like adenoviruses. Still, enveloped viruses showed reduced infectivity in MOGS-CDG cells as these viruses are evolutionary glycosylation-dependent [104].

Enveloped viruses wrap their viral genome and core proteins with membranes derived from the infected host cell. These include viruses such as HIV, Influenza, Dengue, and Ebola, which have evolved to exploit the ERQC for extensive glycosylation of their viral proteins [105]. Glycosylation of their proteins plays functional and structural roles (*Figure 3.8*), where glycans can help shield the virus from the host immune system by decorating with 'Self' glycans [105]. The importance of N-linked glycans is highlighted in HIV, where Env protein folding and stabilisation depend on an N-linked glycosylation site [105].

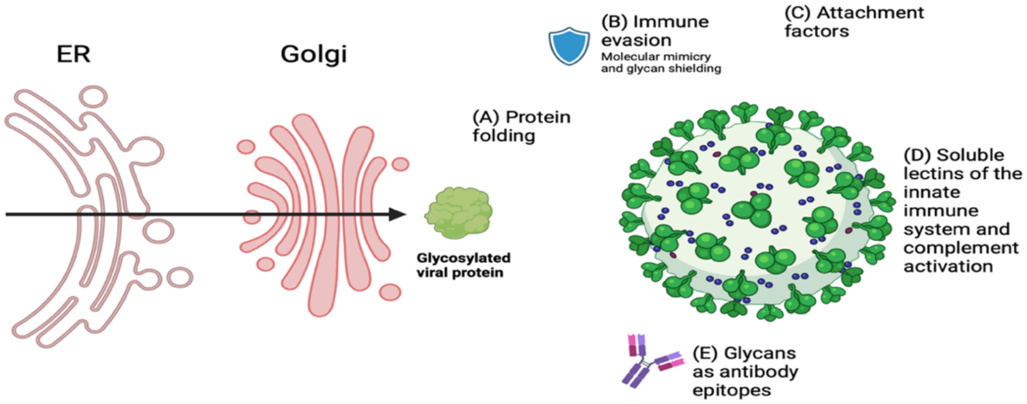


Figure 3.8- Role of glycosylation in the viral lifecycle (Image created in BioRender).
(A) Viral protein follows the host cells' secretory pathway, where the viral protein is glycosylated for proper folding protein and virus assembly. (B) Glycans help to shield and mimic the host cells' glycosylation to evade the host's immune response. (C) glycans on viral proteins help mediate contacts with host cell proteins. (D) If viral glycoproteins are under-processed, this can trigger an immune system response from the host. (E) Glycans can be epitopes in for antibodies.

The ERQC/ERAD has been implicated in various types of cancer, as cancer cells are heavily reliant on secreted glycoproteins, making them desirable targets for anti-cancer treatments [92]. Cancer types with a high secretory burden, such as multiple myeloma, heavily rely on the ER [92]. Survival of cancer cells under chronic ER stress is a crucial feature of aggressive cancers, as they rely heavily on the ERAD pathway; otherwise, the accumulated misfolded proteins would induce the unfolded protein response, eventually leading to cell death [92]. The inhibition of ER-αGluI and II could trigger apoptosis in cancer cells. Due to cancers' reliance on these proteins, their overexpression is linked to poor prognosis [106]. Where GANAB is found to be overexpressed in muscle-associated bladder cancer, GANAB can be used as a potential disease marker and further explored for therapeutic potential as well [106].

3.2.5. Inhibitors of Endoplasmic Reticulum α-Glucosidases I and II

The design of broad-spectrum antiviral inhibitors targeting ER-αGluI and II is desirable for treating diverse viral infections. Generally, the main challenge for antiviral drug development is the design of a compound that will remain efficacious by not causing the emergence of antiviral-resistant strains [89]. Creating host-targeted antivirals would avoid this, as resistant mutations would harm the viruses' lifecycle and assembly as they are critically dependent on host cell proteins [89]. In the case of host-targeted antivirals for the inhibition of ER-αGluI and II, there is an opportune window for short-term inhibition without significantly impacting mammalian protein production due to the long half-life of mammalian glycoproteins [107].

Inhibition of ER α-glucosidases leads to incomplete glucose trimming of N-linked oligosaccharides. This has been reported to affect the efficiency of virion assembly and secretion due to the degradation of enveloped viral glycoproteins [108-110]. However, occasionally, depending on the type of enveloped virus, the virion assembly is not affected; instead, the proteins have unprocessed glycans [111].

3.2.5.1. *Iminosugars for ER α-Glucosidases*

The first reported use of an iminosugar as an antiviral was in 1987 when compounds Castanospermine (CAST) and deoxynojiramycin (DNJ) were shown to inhibit ER-αGluI, blocking the replication of HIV [112, 113]. CAST and DNJ resemble glucose, allowing them to inhibit ER α-glucosidases as transition state (TS) mimics competitively [114]. Glucosidase inhibition changes virus morphology and misfolding of enveloped virus glycoproteins [115]. DNJ is a 6-membered-ring polyhydroxylated secondary amine, and CAST is an indolizidine iminosugar with a 5,6-ring system linked by ethylene (Figure 3.9) [114].

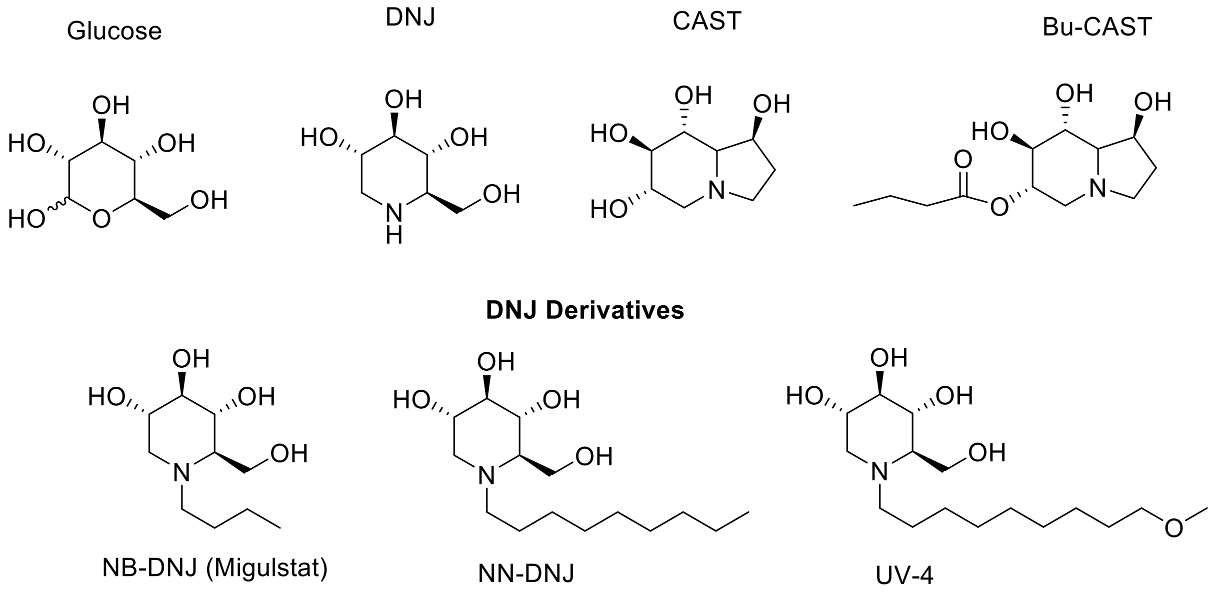


Figure 3.9- Iminosugar inhibitors and their derivatives.

DNJ and CAST are natural iminosugars. DNJ derivatives and Bu-CAST are iminosugar derivatives in preclinical development.

Chemical derivatives of DNJ have been synthesised, with over 200 reported [108, 114, 115]. The nitrogen of DNJ allows for chemical derivatisation, and adding an alkylated side chain significantly improves the EC_{50} to a low micromolar range in the case of *N*-nonyl-DNJ (NN-DNJ), which has a nine-carbon alkyl side chain. Although NN-DNJ was a more potent inhibitor than DNJ, there was a concerning increase in the cytotoxicity [116]. Adding alkyl side chains of >10 has been shown to increase cytotoxicity significantly, but when oxygen or methoxy groups are added, the toxicity is reduced [116]. In the case of an NN-DNJ derivative with a five-carbon side chain with a terminating cyclohexanol group, cytotoxicity was decreased due to the conformational restrictions of the cyclohexanol [116].

Multiple iminosugars have shown efficacy in treating various viruses in animal models [88, 117-119]. UV-4 was included in the treatment of Dengue infection in mouse models, where the iminosugar protected the mice from the lethal infection by reducing viremia and viral RNA in infected tissues [120]. This provided a foundation for developing iminosugar UV-4 as a potential antiviral to be studied in human clinical trials [120]. Bu-CAST (Celgosivir) was evaluated in clinical trials; the prodrug iminosugar is hydrolytically processed into its active form CAST, where it demonstrated potent activity against a broad range of viruses *in vivo* [115]. Unfortunately, when tested in phase I and II for HIV and hepatitis C virus, only modest reductions in viral load were observed again, with the patient exhibiting common iminosugar side effects of intestinal stress [115]. Furthermore, *N*-butyl-DNJ (Miglustat) antiviral activity was also evaluated for the treatment of HIV, but again, adverse side effects halted the trials [121]. Although the iminosugar was not licensed as an antiviral, the information collected from the clinical trial led to its approval as a glucosylceramidase and β- glucosidases inhibitor for treating Gaucher's disease [122]. As discussed, the main barrier to the approval of iminosugars as host-targeted antiviral is their poor specificity. To reduce their off-target effects, alternatives for administration could be explored as the oral application of these drugs is associated with intestinal stress; the parental route of administration or prodrug forms could instead be explored [114].

One of the main drivers for developing iminosugars as host-targeted antivirals is the high barrier to antiviral resistance. The generation of resistant mutations in viruses would likely impact the viruses' lifecycle; experiments with iminosugars have demonstrated if viruses have the potential to become resistant to these antivirals [89]. Plummer et al. investigated mice infected with Dengue virus and were treated with the iminosugar UV-4B [118]. Their samples were pooled after infection to evaluate for virus mutations [118]. 19 nonsynonymous mutations in Dengue virus proteins were identified after treatment with the iminosugar UV-4B [118]. However, despite the number of mutations, the viruses were still susceptible to UV-4B, suggesting a high barrier to resistance for host-targeted iminosugars [118]. Similar results were also seen when evaluating UV-4B resistance with influenza virus [89].

Alternatively, the design of inhibitors using structure-based design may produce inhibitors with reduced off-target effects. This could be done by targeting +2 and +3 subsites as these are less conserved in the GH31 family, and inhibitors would selectively target ER-αGluII instead of other GH31 enzymes [123]. Structures of ER-αGluII with iminosugars were determined by Caputo et al., where DNJ, NB-DNJ, CAST and UV-4 were captured in a complex with ER-αGluII [98]. All four bind in the -1 conserved GH31 subsite with their hydroxyl moieties having similar interactions to glucose [123]. Their endocyclic nitrogen is orientated towards the catalytic D564 [123]. Compared to other GH31 glucosidases, the F307 loop is unique to ER-αGluII; this could be exploited in the design of selective inhibitors, as this structural characteristic is not present in intestinal GH31 enzymes [123].

3.2.6. Small Molecule Inhibitors for ER α-Glucosidases

The majority of inhibitor designs for ER-αGluII have focused on the development of competitive iminosugars. Most retaining glycosidases employ a conserved mechanism, where the sugar substrate undergoes various conformational distortions. Mimicking these substrate conformational itineraries is a powerful method for designing irreversible covalent inhibitors [102]. Cyclophellitol analogues have been proven potent irreversible inhibitors that lock into the TS-like conformation of retaining glycosidases [19, 102]. Inspired by the potency of these analogues, Artola et al. sought to investigate the cyclophellitol derivative cyclic sulfate, an electrophilic equivalent to the cyclophellitol epoxides in organic synthesis [102]. The group verified that the 1,6-*epi*-cyclophelltiol cyclosulfate (**3**) is a rapid mechanism-based inactivator of retaining α-glucosidases (Figure 3.10) [102]. The 1,6-*epi*-cyclophelltiol cyclosulfate adopts the $^4C_1$ Michaelis complex-like conformation of the substrate, so it is readily positioned for nucleophilic attack and subsequent covalent adduct formation.

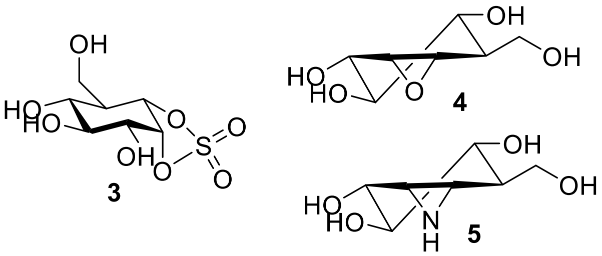


Figure 3.10- Structure of ER-αGluII cyclophellitol inhibitors.

*α-cyclosulfate (**3**), 1,6-epi-cyclophellitol (**4**), 1,6-epi-cyclophellitol aziridine (**5**). Synthesised by [102].*

In vitro inhibition assays revealed **3** had remarkable specificity for α-glucosidases over β-glucosidases such as GBA1 and GBA2, where the IC$_{50}$ reported were >100 µM [102]. Investigating its specificity with α-glucosidases, compound **3** showed more potency for ER-αGluII (IC$_{50}$ 29 nM) over lysosomal acid α-glucosidase (GAA) (82 nM) [102]. Artola et al. also reported on the permeability of their compounds; **3** was able to cross the cell membrane *in situ*, where inhibition of GAA (3µM IC$_{50}$) and GANAB (2.6 µM IC$_{50}$) was observed [102]. This work provided a good starting point for developing new agents for treating viral diseases.

3.2.7. Activity-Based Probes for ER α-Glucosidases

α-Glucosidases are implicated in diverse physiological processes due to their carbohydrate processing, which is involved in food processing, protein folding and glycogen storage [124]. Jiang et al. designed activity-based probes (ABPs) for the tissue and pH-specific reporting of α-glucosidases [124]. Using the Koshland double-displacement mechanism, their retaining mechanisms allow cyclophellitols for the design of ABP for retaining glucosidases. The probes were proven to selectively label ER-αGluII, reporting that the enzymes are optimally labelled at a neutral pH and can be reacted with cyclophellitol aziridine with Cy5 fluorescence (Figure 3.11) [124]. Other probes were also reported conjugated to a BODIPY-green dye, a BODIPY-red dye, or a biotin tag [124].

ABP for α-glucosidases has been proven a powerful tool for verifying the enzyme's optimum conditions in a tissue-specific manner to identify functional enzyme populations rather than just expression levels [124]. The ability to label only functional enzymes allows the probes to be used for diagnosis and studying disease states.

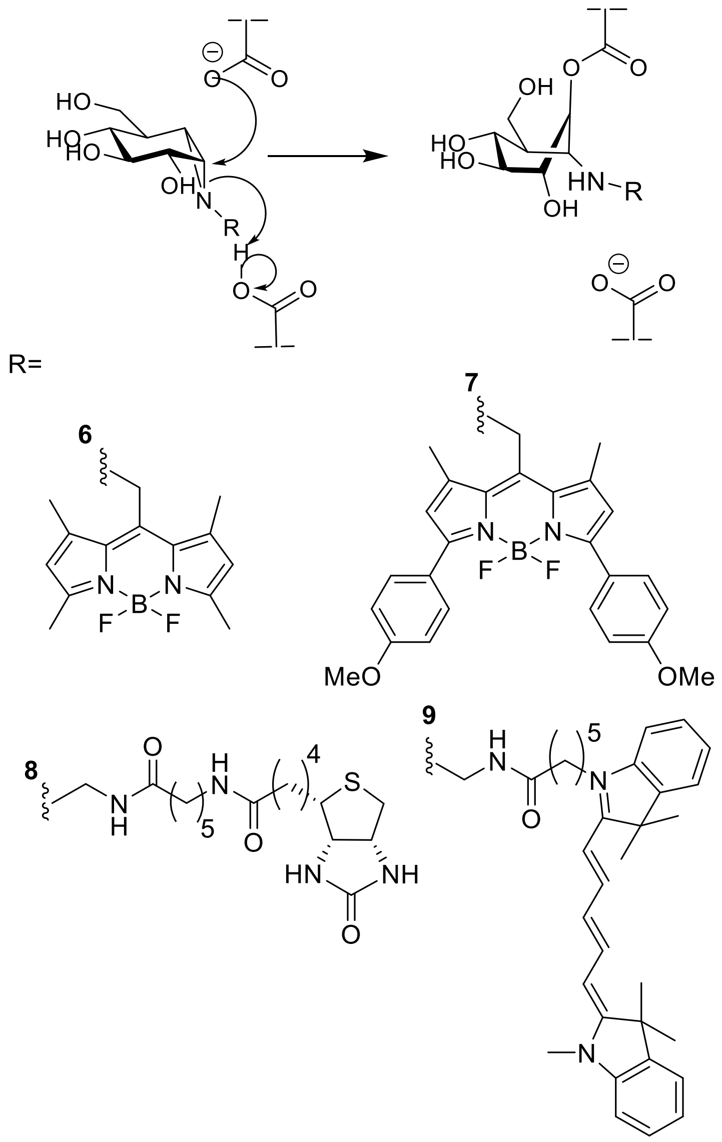


Figure 3.11- α-Glucose-configured N-alkyl cyclophellitol aziridine probes.

*Conjugated to BODIPY-green dye (**6**), a BODIPY-red dye (**7**), biotin tag (**8**) or Cy5 fluorescence (**9**).*

3.2.8. Research Aims

3.2.8.1. ER α-Glucosidases I Aims

This work aims to establish *Ct*ER-αGluI expression by testing different expression systems, such as baculovirus and *Pichia pastoris (P.pastoris)* so that structural and inhibitory work can be conducted in the future.

3.2.8.2. ER α-Glucosidases II Aims

This work aims to expand the portfolio of *Mm*ER-αGluII inhibitors to inform the design of broad-spectrum antivirals. The work presented in this chapter is in collaboration with the Overkleeft group (University of Leiden), who synthesised these inhibitors (Figure 3.12). The synthesised compounds include 1,6 and 1,2 cyclitols.

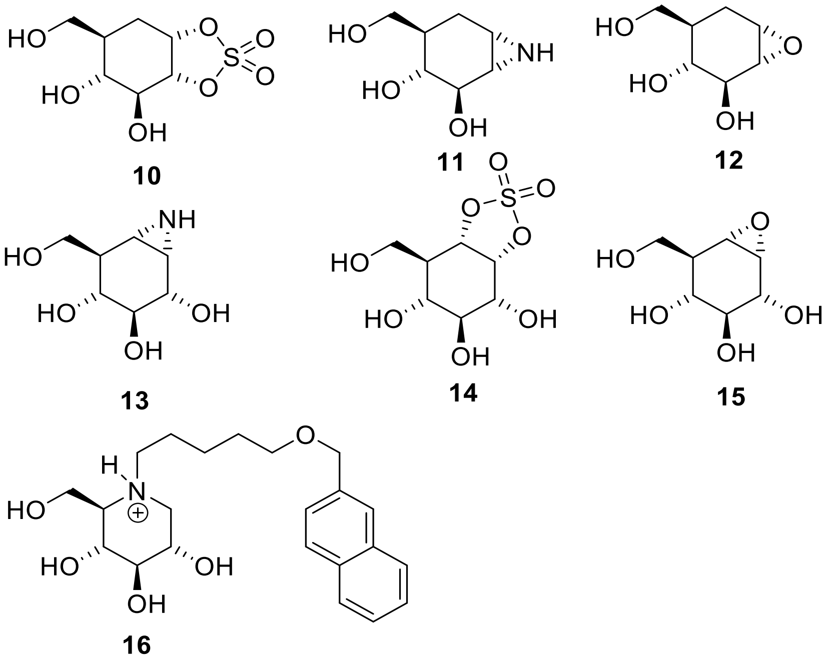


*Figure 3.12-. 1,2 cyclitol compounds **10-12,** 1,6 cyclitol compounds **13-15** and Miglustat analogue **16.***

3.3. Materials and Methods

3.3.1. *Chaetomium Thermophilum ER-α-Glucosidases I*

3.3.1.1. Pichia Pastoris Test Expression

*Ct*ER-αGluI pPICZαA construct (5 ng) was transformed into *E.coli* stellar cells for DNA isolation by heat shock at 42°C for 30 seconds and plated onto low salt LB agar (5 µg/mL NaCl, 5 µg/mL yeast extract and 10 µg/mL, tryptone 15 g/L agar) with 25 µg/mL Zeocin^{TM}. The streaked plate was left overnight at 37 °C before colonies were picked and then grown in 5 mL low salt LB (5 µg/mL NaCl, 5 µg/mL yeast extract and 10 µg/mL tryptone) with 25 µg/ml Zeocin^{TM} and glycerol stocks created when the culture reached OD_{600} = 0.5–0.7. The glycerol stocks were then used to create an overnight culture left at 37 °C before conducting DNA isolation using the PureLink^{TM} Quick Plasmid Miniprep Kit to isolate the pPICZαA construct.

The DNA was digested using the SacI restriction enzyme to linearise the construct. 50 µl reactions were set up (Table 3.3) and initially incubated for 30 minutes at 37 °C. SacI was denatured at 65 °C for 20 minutes before using a QIAquick PCR Purification Kit. The DNA was precipitated using ethanol to obtain 15-20 ng/µL of DNA.

Table 3.3- SacI restriction digest

COMPONENT	**50 µL REACTION**
DNA	1 µg
10X NEBuffer 1.1	5 µL (1X)
SacI	1.0 µL (20 units)
Nuclease-free Water	to 50 µL

To prepare the *P.pastoris* (X-33) cells for transformation, they are first streaked onto a YPD plate (1% yeast extract, 2% peptone, 2% dextrose (glucose), 2% agar) and grown for 2-3 days at 30 °C. A fresh colony was picked and cultured in 100 mL of YPD overnight at 30 °C and harvested at OD_{600} = 1-2, where the culture was then centrifuged at 3000 xg for 2 minutes at 4 °C and resuspended with sterile 1 M sorbitol, with this repeated two more times before resuspending in 1M sorbitol to 1-2 mL total.

5-10 µg of linearised DNA was added to 80 µL of the prepared X-33 cells and incubated for 5 minutes before adding to an ice-cold 0.2 cm electroporation cuvette and transformed by electroporation at 1.5 kV at a time constant 5-6 ms. The cells were immediately rescued with 1 mL ice-cold 50% 1M sorbitol 50% YPD to the cuvette; contents were then transformed to a sterile 2 mL Eppendorf and incubated at 30 °C at 800 rpm for 4 hours. The culture was then streaked onto a YPDS plate (1% yeast extract, 2% peptone, 1 M sorbitol, 2% agar, 2% dextrose (glucose)) with 100 µg/mL Zeocin™ and incubated at 30 °C for 3 days. Picked colonies were restreaked onto YPDS plates and then analysed using agarose gel after a colony PCR and glycerol stocks were made.

Test expressions were conducted in 20 mL cultures in BMGY media (1% Yeast extract 2% Peptone, 0.1M Potassium phosphate (pH 6.0), 1.34% YNB (Yeast Nitrogen Bases), 0.0004% Biotin, 1% Glycerol) grown at 30 °C at 300 rpm until the culture reached an OD_{600} = 2-6. The culture was centrifuged at 3000 xg for 5 minutes at room temperature before discarding the supernatant and placing the centrifuge tubes upside down to remove any residual BMGY media. Pellets were then washed with 20 mL of BMMY media (1% Yeast extract, 2% Peptone, 0.1 M Potassium phosphate (pH 6.0), 1.34% YNB (Yeast Nitrogen Bases), 0.0004% Biotin, 0.5% Methanol), the washing procedure was repeated twice, and cells were resuspended to an OD_{600} of 1.0. Cells were then cultured at 30 °C at 300 rpm, and sterilised methanol was added to 0.5% total every 24 hours for expression induction. Expression time points were taken every 24 hours over 4 days and centrifuged for 1300 xg for 5 minutes.

3.3.2. Chaetomium Thermophilum ER α-GluI Baculovirus Expression

The recombinant bacmid was produced using DH10Bac™ competent cells, which contained a tetracycline-resistant helper plasmid (bMON7142) and kanamycin-resistant baculovirus shuttle vector (bMON14272). The ER-αGluI construct was designed using a GenScript codon optimisation programme for High Five ™ (Trichoplusia ni, BTI-TN-5BI-4) cells. The plasmid was transformed by electroporation (1.8 kV) into DH10Bac™ competent cells before immediately adding SOC medium and incubating the cells for 4 hours at 37 °C. To confirm transformation, blue-white screening was conducted on IPTG containing agar plates (1 mM), Kanamycin (50 µg/ml), tetracycline (15 µg/ml) and gentamycin (15 µg/ml) containing LB agar plates and incubated for 2 days at 37 °C and the resulting white colonies restreaked.

3.3.2.1. Bacmid Verification

Selected restreaked colonies were incubated overnight in 10 mL of LB media at 37 °C, and the plasmid DNA was extracted using the PureLink^{TM} genomic DNA mini kit (Thermo Scientific). Verification of the bacmid transformation was carried out using colony PCR. PCR reactions (Table 3.4) were prepared for each restreaked colony to 20 µL containing DNA (1 µL), p55 (5'-CCCAGTCACGACGTTGTAAAACG -3') (0.5 µM), p56 (5'-AGCGGATAACAATTTCACACAGG-3') (0.5 µM) dNTPs (0.2 mM), Phusion enzyme (2 U) and Phusion HF buffer (1x) (Thermo scientific F530S). 1% agarose gel with SYBR-safe DNA gel stain (0.1x) (Thermo Scientific) was used to verify the formation of the bacmid.

Table 3.4- PCR reaction mix

Reaction Mix			**Cycle parameters**	
Component	Volume (µL)	10x stock	Step	Temperature
HF Buffer	4	40	ID	95
H20	11.8	118	D	95
DMSO	0.6	6	A	68
dNTPs	0.4	4	E	72
Primer_F (p55)	1	10	FE	72
Primer_R (p56)	1	10	H	4
Template	1		Total time	2hr 30
Phusion	0.2			

3.3.2.2. Generation of Recombinant Baculovirus in Sf9

A 60 mL Sf9 (derived from *Spodopetre frugiperda*) suspension culture was grown for 2 days at 28 °C in Insect-XPRESS protein-free insect cell medium (Lonza Bioscience) before splitting back to 0.45 x10^6 cells/mL. Suspended cells were transferred into each well of a 6-well culture plate and left in a humidified chamber at 28 °C until the cells became adherent. Each well of cells was transfected with 180 µL of Insect-XPRESS media, 50 µL DNA (~100 ng/µL) and 31.5 µL of FuGENE HD transfection reagent (Promega) before placing back into the humidified chamber for 4 days. Once the *Sf*9 cells are 95% fluorescent and 90 % viable, the culture is collected and centrifuged at 200 xg for 5 minutes to produce a clarified supernatant containing 2% fetal bovine serum: this is referred to as V1. A second-generation baculovirus stock (V2) was made to amplify virus titre using a 50 mL suspension of *Sf*9s at 1 x10^6 cells/mL and adding 1 mL of V1. The culture was harvested after incubation at 28 °C and 87 rpm until 95% fluorescence was achieved.

3.3.2.3. Expression and Purification of Chaetomium Thermophilum ER α-GluI

The suspended culture of High Five ™ (*Trichoplusia ni*) cells was expanded using Express Five ™ SFM supplemented with 20 mM media until the cells reached 3.6 L at 1.2 x10^6 cells/mL using 6 x 600 mL Corning Erlenmeyer shaker flasks. 2 mL of V2 stocks was added to each 600 mL culture and left for 3-4 days until >90% fluorescence and <10% death was reached at 28 °C and 87 rpm. The expression culture was harvested by pelleting the cells at 200 xg for 20 minutes at 4 °C; the clarified media was then further centrifuged for 20 minutes, 5000 xg at 4 °C and then supplemented with cOmplete ™ EDTA-free protease inhibitor (Sigma).

Clarified media was pH adjusted using 150 mL of 10 x PBS and 3.6 mL 1M imidazole to 3.6 L of the clarified media to a final pH of 7.12. Conditioned media loaded onto a 5 mL HisTrap excel column (GE Healthcare) equilibrated with binding buffer (1x PBS 5% glycerol 20 mM imidazole) and eluted using a buffer gradient 0-100% of elution buffer (1x PBS, 500 mM imidazole, 5% glycerol w/v) over 20 CVs at a flow rate of 5 mL/minute eluting 1.8 mL fractions. Fractions containing protein were pooled and concentrated using a 30 kDa Vivaspin ™ concentrator to 1 mL. Concentrated protein loaded onto size exclusion S200 column, 1 ml/minute flow rate with HEPES buffer (20 mM HEPES (pH 7.0) and 150 mM NaCl). Protein-containing fractions were pooled and concentrated using an Amicron 30K concentrator, and protein concentration was determined using absorbance at A280 nm, 4.65 mg/mL in 30 µL aliquots, flash frozen in liquid nitrogen and stored at -70 °C.

3.3.3. *Mus musculus ER-αGluII*

3.3.3.1. *Expression of Mus musculus ER-αGluII*

For the expression of *Mm* ER α-GluII; GANAB (α subunit) and PRKCSH (β- subunit) constructs, 4 L of LB media was inoculated with each construct and 1ml of ampicillin and left overnight at 37 °C, 250 rpm for 15 hours. A maxiprep of the 4L of LB was performed using the PureLink^{TM} HiPure plasmid filter Maxiprep kit and quantified using absorbance at A280 nm to obtain 750 µg of DNA for each construct. A suspension of Freestyle 293-F cells was prepared and cotransfected following the Freestyle 293 Expression system manufacturer protocol, scaling up to 600 mL with 750 ug of DNA in both constructs and incubated for 4 days at 37°C, 8% CO_2, 135 rpm. Cells were then pelleted at 200 xg, 20 minutes at 4°C, and the clarified media was then further centrifuged for 20 minutes, 5000 xg at 4°C.

3.3.3.2. *Purification of Mus musculus ER α-GluII*

Clarified media was pH adjusted using 150 mL of 10 x PBS and 3.6 mL 1M imidazole to 580 mL of the clarified media to a final pH 7.12. Conditioned media was loaded onto a 5 mL HisTrap excel column (GE Healthcare) equilibrated with binding buffer (1x PBS, 20 mM imidazole, 5% glycerol w/v) and eluted using a buffer gradient 0-100% of elution buffer (1x PBS, 500 mM imidazole, 5% glycerol w/v) over 20 CVs at a flow rate of 5 mL/minute eluting 1.8 mL fractions. Fractions containing protein were pooled and concentrated using a 30 kDa Vivaspin ^{TM} concentrator to 1 mL. Concentrated protein loaded onto size exclusion S200 column, 1 ml/minute flow rate with HEPES buffer (20 mM HEPES pH 7.5 and 150 mM NaCl). Protein-containing fractions were pooled and concentrated using Amicron 30K concentrator, and protein concentration was determined using absorbance at A280 nm, 0.94 mg/ml at 1ml added 9.4 µL of trypsin and 2 µL of 1M $CaCl_2$ incubated overnight at room temperature. Repeat the size exclusion purification and the resulting protein is pooled and concentrated using a 10 kDa concentrator to 4.4 mg/mL in ten 30 µL aliquots, flash frozen in liquid nitrogen and stored at -70 °C.

3.3.3.3. *Thermal Shift Assay-Nano Differential Scanning Fluorimetry (NanoDSF)*

Reaction mixtures of 30 µL were made up of 20 mM HEPES, pH 7.5 and 150 mM NaCl at 0.8 mg/mL ER α-GluII reaction mixtures. Reactions were set up with inhibitors for 5 hours at 1:10 (protein: inhibitor). Thermal unfolding of the protein was monitored using Prometheus NT.48

instrument by tryptophan absorbance at λ_{ex} 260 nm, λ_{em} 330 and 350 nm, and 40% gain at a temperature range of 20 -95 °C of 1 °C min^{-1}.

3.3.3.4. *ABP of Mus musculus ER α-GluII*

*Mm*ER-αGluII competitive gels were set up with 10 µM cyclophellitol inhibitors for 5 hours at 37 °C in reaction mixtures with 1 µM ER-αGluII 20 mM HEPES, pH 7.5 and 150 mM NaCl. After incubation with inhibitors the probe (**9**) was added for 1 hour 37 °C before samples were denatured with 5x Laemmli buffer (50 % (v/v) 1.0 M Tris−HCl, pH 6.8, 50% (v/v) 100 % glycerol, 10 % (w/v) DTT, 10 % (w/v) SDS, 0.01 % (w/v) bromophenol blue) by boiling for 5 min at 98 °C. Denatured samples were loaded onto a 10% SDS-PAGE gel at 200V for 30 minutes. Wet slab gels were scanned for ABP-emitted fluorescence using a Cy5 (λEX = 625 nm).

3.3.3.5. *Crystallisation of Mus musculus ER-αGluII*

Crystallisation trays in 48 well MAXI plates with 4.46 mg/ml *Mm*ER-αGluII, 0.1 M Tris/bicine buffer pH 8, 0.02 M alcohol mix (0.2M 1,6 Hexanediol, 0.2 M 1-butanol, 0.2 M 1,2-propanediol, 0.2M 2-propanol, 0.2M 1,4-butanediol, 0.2 M 1,3 propanediol) and Morpheus precipitant mix 1 (40% v/v PEG 500 and 20 % w/v Peg 20,000). 10 µL of 0.1 M tris/bicine buffer pH 8, 0.02M alcohol mix and Morpheus precipitant mix were added to each inhibitor vial, 0.8 µL of the same mixture was added to the sitting drop and 1.6 µL of ligand was added to the drop to create a 5 mmol final concentration of ligand (**10-12**). Other crystallisations were also explored: 0.03 M NPS (0.3 M sodium fluoride, 0.3 M sodium bromide, 0.3 M sodium iodide), 40% v/v ethylene glycol, 20 % w/v PEG800 and 1 M HEPES.

3.3.3.6. *Docking using Maestro software*

The crystal structure co-ordinates of *Mm*ER-αGluII in complex with D-glucose (PDB: 5H9O)[98] was imported into Maestro software. The protein was prepared for docking using the protein preparation module with conditions set to pH 7.4 [125]. A receptor grid was generated centred around the bound D-glucose molecule. The ligand was imported into the Maestro GUI and prepared using the LigPrep tool. Ligands were then docked using the generated receptor gird with the OPLS4 forcefield and Glide module [126-130]. XP mode (extra precision) and flexible ligand sampling were utilised, and epik state penalties were added to the docking scores [131]. Conformational, torsional and positional restraints were not used, except for stereochemical definitions in the ligand file. A maximum of 10 outputted poses was requested, and post-docking minimisation was carried out.

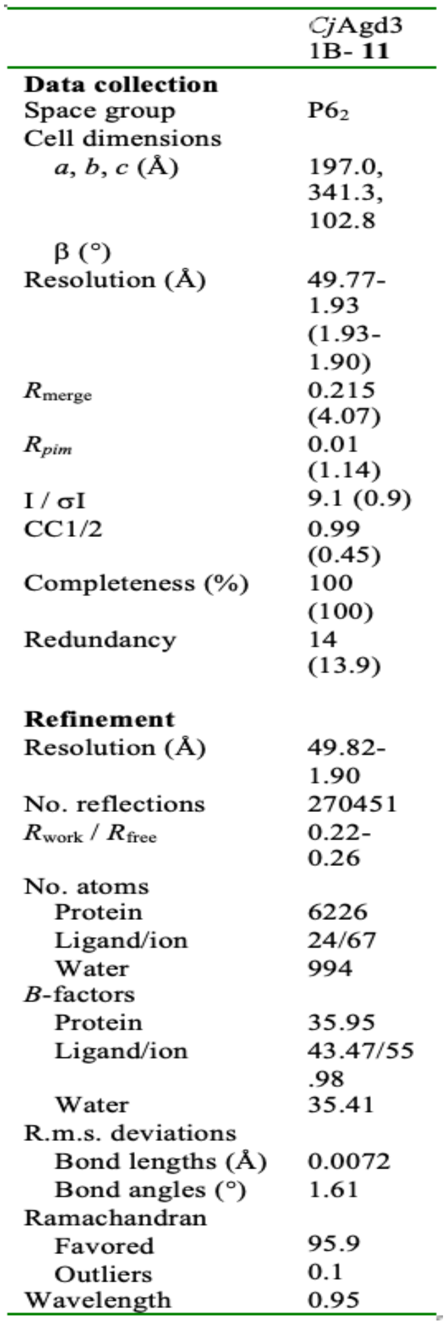
Covalent docking was also performed using the covalent docking module of Glide [132]. The nucleophile Asp564 was defined as the reactive residue, and the prepared ligands were docked into the previously generated receptor grid. A custom reaction type for the aziridine and cyclic sulfate was created, allowing for attack of the aspartate residue at position 1 of the ligand definition (cyclic sulfate: [C][O][S]=[O]; aziridine [C][N][C]), and bond breakage between positions 1 and 2 of the ligand definition. Docking was carried out in thorough 'pose prediction' mode, and MM-GBSA scoring was performed. Conformational, torsional and positional restraints were not used, except for stereochemical definitions in the ligand file.

The docking poses were exported as .pdb files and imported into ChimeraX, where figures were generated [133].

3.3.3.7. *Cellvibrio japonicus Agd31B X-ray Crystallography Structure*

CjAgd31B was obtained from previous recombinantly expressed stocks from Prof. Davies group [102]. Initially, *Cj*Agd31B was buffer exchanged into 5 mM Bis-Tris propane at 7 mg/ml and well-diffracting crystals were obtained by sitting drop vapour diffusion at 1.8 M ammonium sulfate, 0.1 M HEPES (pH 7), 2% PEG 400 at 20 °C. Inhibitor-ligand complexes were obtained by transferring the crystals into 2 M lithium sulfate, 0.1 M HEPES (pH 7), 2% PEG 400 cryoprotectant containing 5 mM ligand for 2 hours before flash freezing in LN2 for data collection. Data was collected at the Diamond light source to 1.9 Å; data reduction was performed using AIMLESS, MOLREP for molecular replacement and REFMAC5 for the model refinement and model building in ccp4mg [74, 75].

*Table 3.5- X-ray crystallography data collection and refinement statistics for Cellvibrio japonicus Agd31B with **11***



	*Cj*Agd31B- **11**
Data collection	
Space group	$P6_2$
Cell dimensions	
a, b, c (Å)	197.0, 341.3, 102.8
β (°)	
Resolution (Å)	49.77-1.93 (1.93-1.90)
R_{merge}	0.215 (4.07)
R_{pim}	0.01 (1.14)
I / σI	9.1 (0.9)
CC1/2	0.99 (0.45)
Completeness (%)	100 (100)
Redundancy	14 (13.9)
Refinement	
Resolution (Å)	49.82-1.90
No. reflections	270451
R_{work} / R_{free}	0.22-0.26
No. atoms	
Protein	6226
Ligand/ion	24/67
Water	994
B-factors	
Protein	35.95
Ligand/ion	43.47/55.98
Water	35.41
R.m.s. deviations	
Bond lengths (Å)	0.0072
Bond angles (°)	1.61
Ramachandran	
Favored	95.9
Outliers	0.1
Wavelength	0.95

3.4. Results and Discussion

3.4.1. *Chaetomium Thermophilum ER- αGluI Gene Expression and Purification*

Bacterial expression systems are predominately used to express recombinant proteins due to their rapid growth and easy expression with inexpensive media. However, bacteria expression has limitations, as more complex expression systems are required for mammalian proteins, which have more complex needs for soluble recombinant protein production, such as post-translational modifications. Therefore, for *Ct*ER-αGluI expression, an insect expression system was used. An ER-αGluI construct was designed using a pFastBac vector with *Ct*ER-αGluI flanked by an N-terminal honeybee melittin signal peptide for protein secretion and a C-terminal His-tag (Figure 3.13). *Ct*ER-αGluI sequence was derived from the previously reported production of *Ct* ER-αGluI [17], as the crystallisation of other homologs like *Mm* is challenging to replicate, as discussed in section 3.2.2.

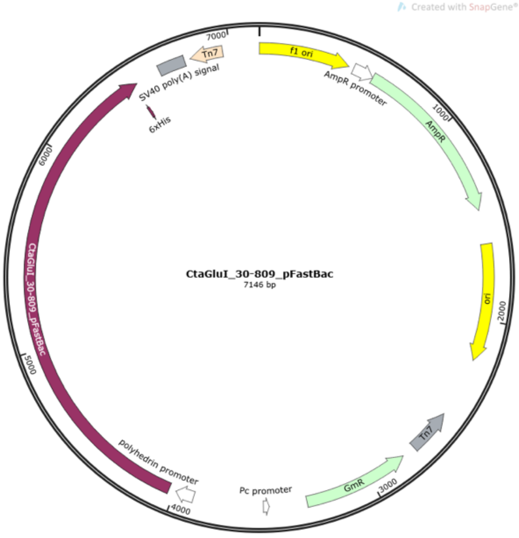


Figure 3.13- Ct ER-αGluI pFastBac construct.

Ct ER-αGluI is flanked by an N-terminal honeybee melittin signal peptide for secretion and a C-terminal hexahistidine tag, under the transcriptional control of the polyhedrin promoter. Plasmid contains gentamycin and ampicillin resistance along with a Tn7 transposition site for bacmid generation.

The first test expressions were conducted using Expi*Sf*9 cells. At the time, these cells were used to express protein instead of the commonly used Hi5 insect cells due to the supply issues of Hi5 media during this time period. Expi*Sf*9 expressions were carried out similarly, where recombinant baculovirus stocks were added to the scaled-up culture of Expi*Sf*9 cells. A high yield was expected as this cell line produces recombinant protein 19-fold higher than the traditional *Sf*9 system due to the cell's ability to achieve a higher cell density [134]. However, the expression yielded a large contaminant at ~40 kDa, compared to *Ct*ER- αGluI ~100 kDa after analysis of the HisTrap purification (Figure 3.14). Subsequent expressions were conducted but still yielded this contaminant; nonetheless, the protein-containing fractions were concentrated and further purified using SEC. However, the yield from the Expi*Sf*9 expression never yielded enough protein for X-ray crystallography studies (2.5 mg/mL).

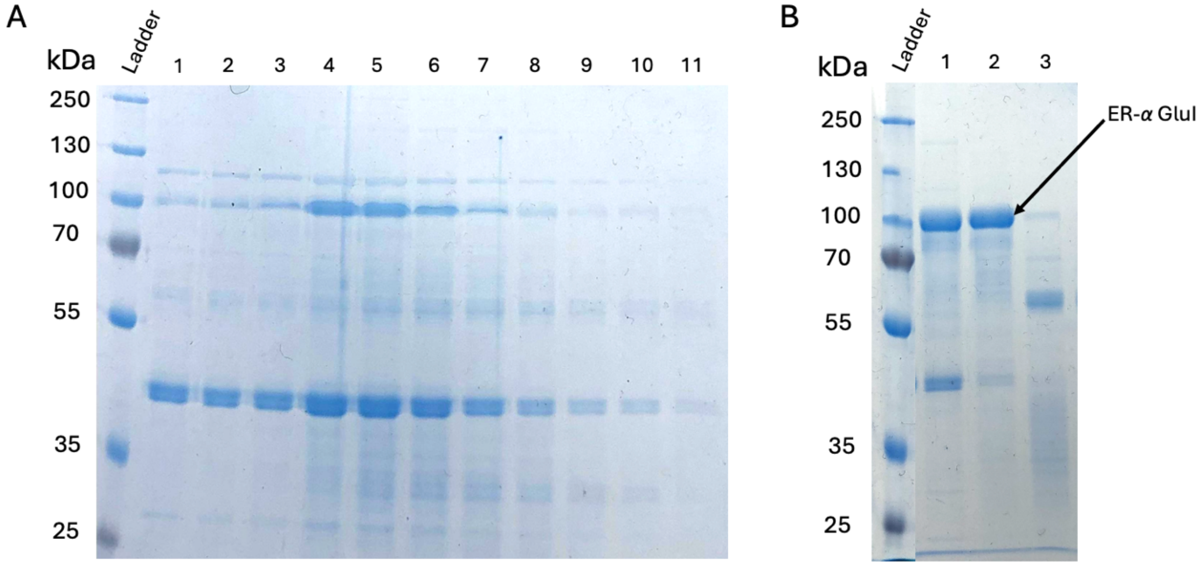


Figure 3.14- Purification of Ct ER α-GluI using ExpiSf9 expression system.

(A) SDS-PAGE of ER α-GluI protein fractions, which contained significant A280 absorbance from the 5 mL HisTrap column, band at ~40 kDa represents the contaminant. (B) SDS-PAGE of Ct ER α-GluI protein fractions, which contained significant A280 absorbance from the SEC.

Consequently, due to the unsuccessful attempts to express *Ct*ER-αGluI using the Expi*Sf*9 system, trial expressions of *Ct*ER-αGluI in *P.pastoris* were conducted. *P.pastoris* is an ideal expression system due to its similarity to mammalian expression systems; the cells can perform post-translational N- and O-linked glycosylation modifications and disulfide bond formation [135].

A *Ct* ER-αGluI construct was designed with an N-terminal α- factor signal sequence for secretion and a C-terminal His-tag into a pPICZα A plasmid with Zeocin^{TM} resistance (Figure 3.15). The linearised construct was transformed into X-33 *P.pastoris* cells to stimulate recombination during electroporation. Colonies with Zeocin^{TM} resistance were selected for small-scale test expressions. During expression, the production of the recombinant protein is induced by adding methanol every 24 hours, as *P.pastoris* is a methylotrophic yeast. Methylotrophic yeast can consume methanol as their sole carbon source; the transcription of the gene of interest is under transcriptional control of the AOX1 promoter induced by methanol [136].

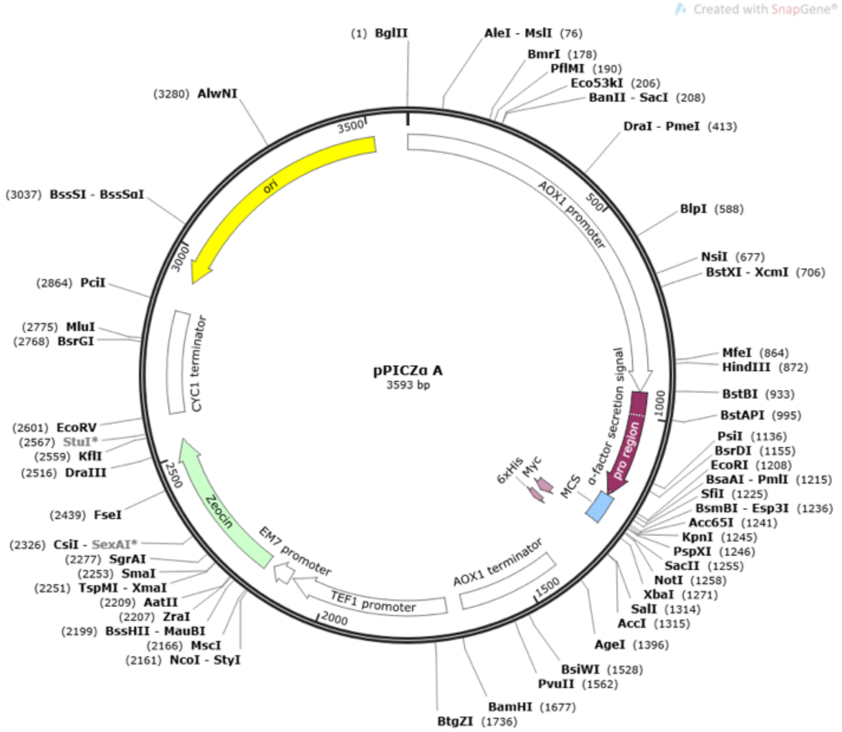


Figure 3.15- pPICZα A Ct ER-αGluI construct for P.pastoris expression.

CtER α-GluI was cloned into the multiple cloning site (MCS) using Xbal and Notl restriction sites. ER, α-GluI sequence contains an N-terminal α -factor signal sequence for secretion and a C-terminal hexahistidine tag with Zeocin^{TM} resistance.

Expression of *Ct*ER-αGluI in *P.pastoris* was unfortunately unsuccessful. Different X-33 transformations were performed, followed by multiple small-scale expressions. One of the desirable attributes of *P.pastoris* expression is their low-level production of endogenous secretory proteins [136]. However, the SDS-PAGE gels from the test expression showed a 'dirty' supernatant, which would not be expected from *P.pastoris*. Additionally, there is no band at

~100 kDa for *Ct* ER-αGluI (Figure 3.16); again, this was confirmed with a western blot where no His-tagged proteins were marked. The lack of expression is unknown, but potential reasons could be RNA instability, insoluble protein, methanol toxicity or challenges with secretion [137].

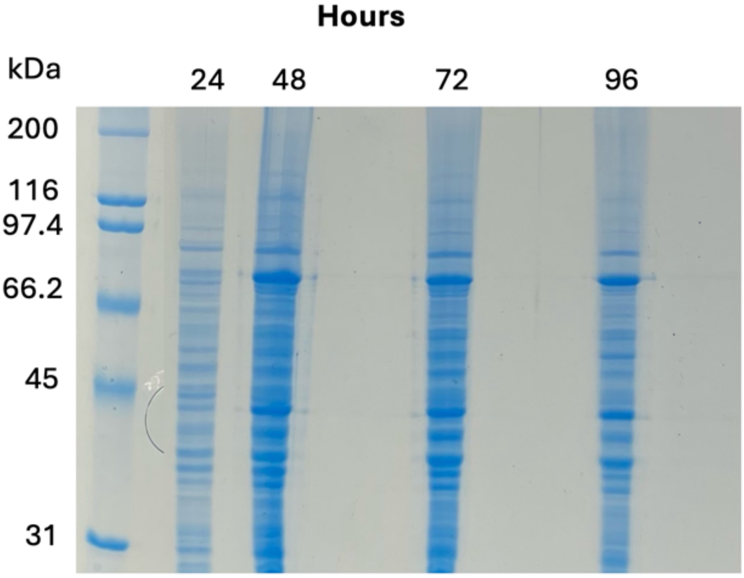


Figure 3.16- Ct *ER-αGluI P.pastoris expression.*
Expression supernatants of P.pastoris ER α-GluI were collected every 24 hours after methanol induction. No band at ~100 kDa (ER-α GluI).

Finally, *Ct* ER-αGluI was expressed using the Hi5 insect cell expression system. The same purification procedure was conducted. ER-αGluI eluted at 20 % imidazole (Figure 3.17) no contaminant was observed this time, and the protein was concentrated and further purified using SEC. The final yield was 5 mg/mL of *Ct* ER-αGluI, protein aliquots were flash frozen and crystallography work is expected to be carried out once ER-αGluI associated inhibitors have been synthesised.

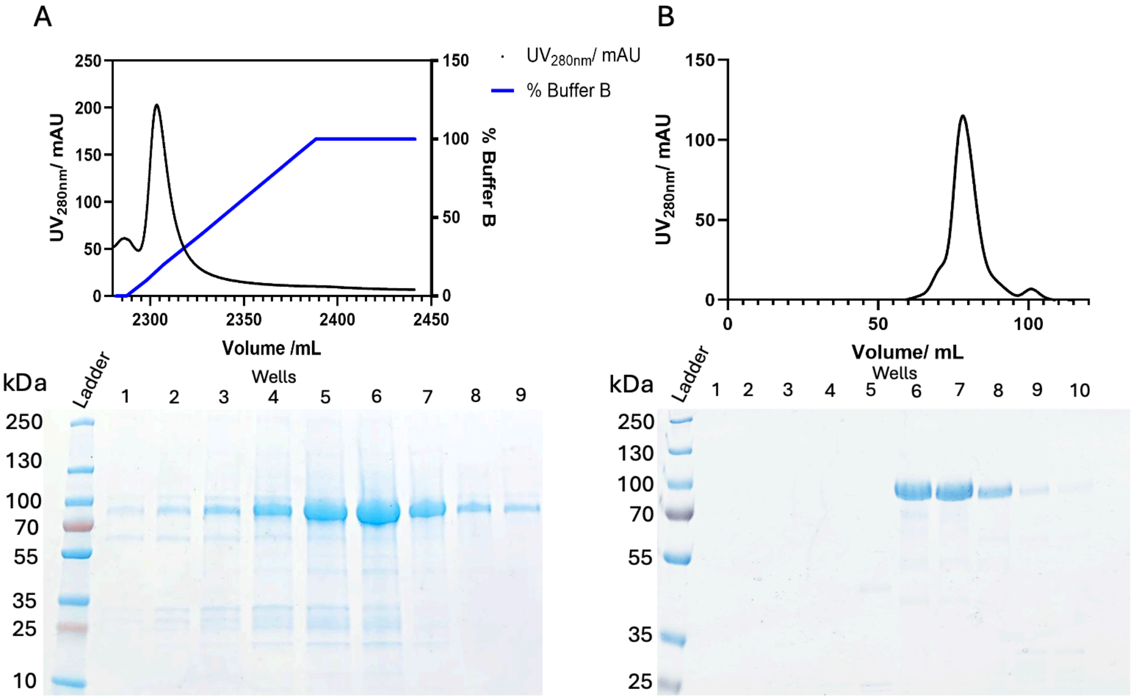


Figure 3.17- Purification of Ct *ER α-GluI using Hi5 expression system.*

(A) SDS-PAGE of ER α-GluI protein fractions, which contained significant A280 absorbance from the 5 mL HisTrap column, alongside the elution profile AKTA chromatogram (B) SDS-PAGE of ER α-GluI protein fractions, which contained significant A280 absorbance from the SEC, alongside the elution profile AKTA chromatogram.

3.4.2. Mus musculus ER α-glucosidase II Gene Expression and Purification

*Mm*ER-αGluII was expressed using the FreeStyle™ 293 Expression system using co-transfection. Two constructs were designed, one containing the α (GANAB) and the other the β (PRKSCH) subunit into pOPINGs plasmids with hexahistidine tags and secretion sequence. During co-transfection expression, the two subunits associate to form the *Mm* ER-αGluII heterodimer before secretion; after 5 days, the culture was harvested and clarified supernatant purified. The purification protocol reported previously was carried out [98]. Protein-containing fractions (Figure 3.18A) from the 5 mL HisTrap column were pooled and loaded onto a size exclusion. GANAB (110 kDa) and PRKSCH (90 kDa) (Figure 3.18B) fractions were again pooled, and a trypsin digest was performed. As discussed, previous papers have reported that crystallisation of *Mm* ER-αGluII does not occur unless the protein has undergone a trypsin digest (Figure 3.18), so there are no full-length structures of *Mm* ER-αGluII. A second size exclusion was performed on the trypsin-digested protein, yielding 5 mg/mL from a 600 mL culture.

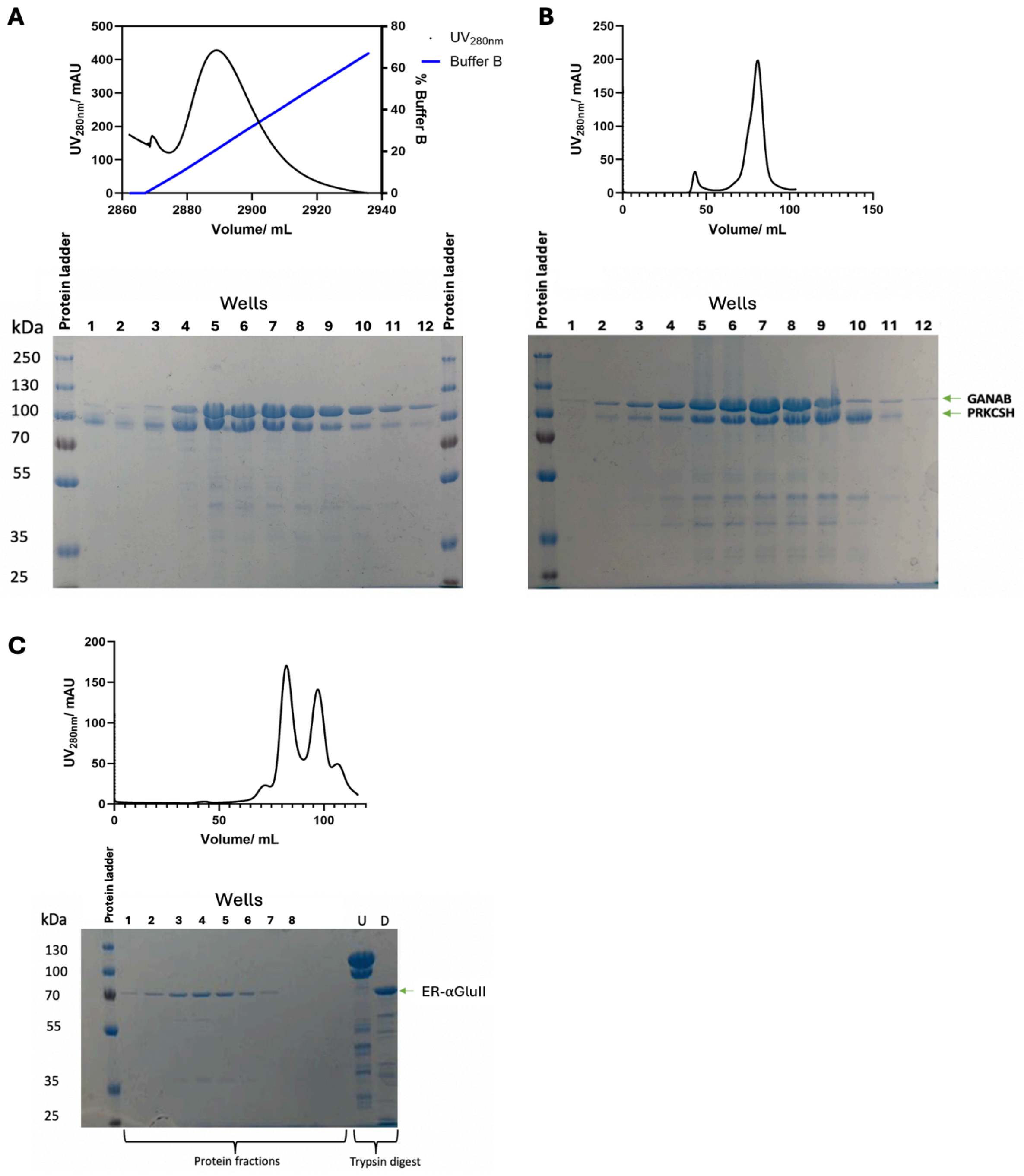


Figure 3.18- Purification of Mm ER α-GluII using baculovirus expression system.

(A) SDS-PAGE of Mm ER α-GluII protein fractions, which contained significant A280 absorbance from the 5 mL HisTrap column, alongside the elution profile AKTA chromatogram (B) SDS-PAGE of ER α-GluII protein fractions, which contained significant A280 absorbance from the SEC, alongside the elution profile AKTA chromatogram. (C) SDS-PAGE of Mm ER α-GluII protein

fractions, which contained A280 absorbance from the SEC after trypsin digest (D) and comparison of undigested (U).

3.4.3. Activity-Based Probe Analysis of Mus Musculus ER α-Glucosidase II

Competitive ABP (**9**) analysis was conducted to determine whether covalent inhibitors outcompete the probe. Reaction mixtures of *Mm* ER-αGluII (1 µM) with 1,2 cyclitols and 1,6 cyclitols (10 µM) were incubated for 5 hours at 37°C before adding **9** for 1 hour at 37°C and imaging the resulting gel for Cy5 fluorescence (Figure 3.19). Compared to the control, most of the inhibitors produced fainter bands of Cy5 fluorescence, suggesting that the inhibitors had bound covalently and were not competing with the probe. Compared to the other inhibitors, the 1,6-*epi*-cyclophellitol cyclosulfate (**14**) band was nearly completely diminished, implying that it could be the most potent of the inhibitors. However, the 1,2 cyclophellitol aziridine produced similar fluorescence to the control, suggesting that the inhibitor is potentially non-covalent.

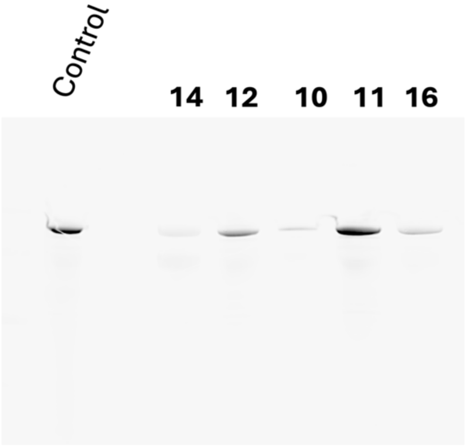


Figure 3.19- Competitive ABP SDS-PAGE gel for Mm ER α-GluII with ABP.

The protein ladder cannot be seen due to the intensity of the ABP bands, but the bands are at 70 kDa for the catalytic α- subunit of Mm ER-αGluII. The inhibitors (10 µM) were reacted with ER-αGluII (1 µM) before reacting with the ABP ***9*** *(10 µM)and visualising on an SDS-gel.*

3.4.4. Thermal Shift Analysis of Mus musculus ER α-Glucosidase II with Cyclitol Inhibitors

The proteins' fluorescence was monitored using the intrinsic fluorescence of tryptophan (330 nm and 350 nm) as a function of temperature to monitor protein denaturation. The thermal melting temperature (T_m) was determined by generating a fluorescence stability graph (Figure 3.20). The formation of an *Mm* ER-αGluII-inhibitor complex should increase the T_m of the protein compared to the unreacted control because of the inhibitor's stabilising effects. Depending on the type of ligand, non-covalent or covalent interactions can occur with *Mm* ER-αGluII. However, NanoDSF is not able to differentiate between these interactions. Nonetheless, all the ligands tested resulted in a shift of T_m compared to the control, suggesting that compounds had bound to *Mm* ER-αGluII. The most significant increase was with compound **12**, the 1,2 cyclophellitol epoxide (64.7 °C). The smallest shift in T_m was with the 1,2 cyclophellitol aziridine **11** (49.5 °C), further supporting that the inhibitor potentially reacts non-covalently, as this type of reactivity was suggested in ABPP (Figure 3.19).

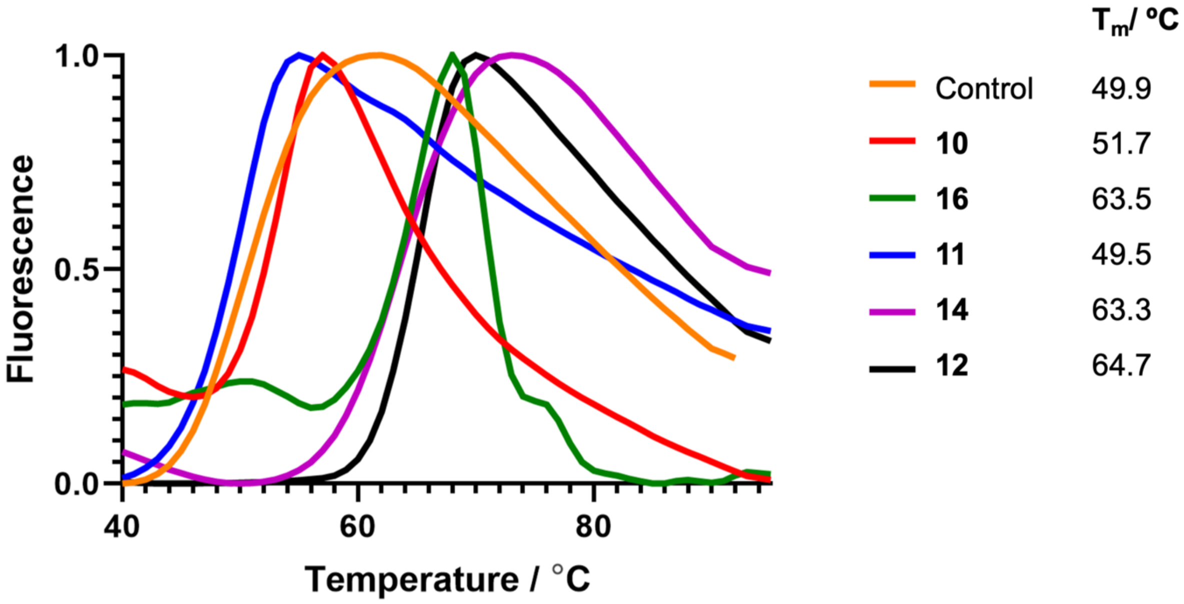


Figure 3.20- Thermal shift analysis of Mm ER α-GluII with inhibitors.

Thermal shift analyses of reaction of Mm ER-αGluII (0.8 mg/mL) after prior incubation with inhibitors(10 µM). The effect of inhibitor binding was monitored using the intrinsic fluorescence of tryptophan as the protein denatured with the increase in temperature.

3.4.5. Crystallisation of Mus musculus ER α-glucosidase II

Initial screening of crystallisation conditions for *Mm* ER-αGluII revealed that the protein only crystalises with Morpheus conditions after 3-4 weeks [138]. The crystals produced were tiny and needle-like. Optimisation plates were set up to obtain larger crystals based on the original hit conditions. Although larger crystals were made, they were thin, needle-like, and jagged, which were poor at diffracting, producing data with <50% completeness and >3Å resolution. In pursuit of better diffracting crystals, microseeding methods (streak seeding and seed beads) and macroseeding *in situ* were implemented to create larger crystals (Figure 3.21). Although larger crystals were collected again, the diffraction data was poor, with problems with completeness and resolution; this may be due to the jagged needle-like nature of the crystal.

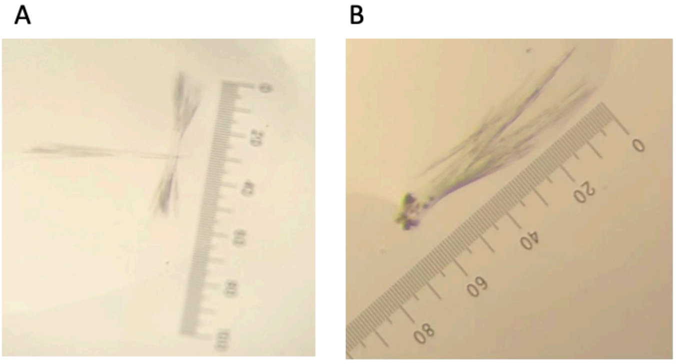


Figure 3.21- Crystals of Mm *ER- αGluII.*

Crystals were produced using crystallisation conditions described previously [3]; these larger but fractured crystals were produced after sequential seeding.

Conditions reported by Caputo et al. were also investigated, and the same kind of crystal was yielded. Previous groups also reported problems with obtaining good diffracting crystals from these conditions. Recently, Karade et al. described the crystallisation of *Mm* ER-αGluII with different crystal growth conditions and implemented glutaraldehyde cross-linking to create more stable crystals for fishing. However, replicating this was also tricky, as the crystals obtained were again needle-like and small, even after multiple rounds of microseeding and macroseeding. Other conditions were also tested, including prior treatment of *Mm*ER-αGluII with Endoglycosidase H for trimming N-linked oligosaccharides.

As the crystallisation of ER-αGluII was unsuccessful, other GH31 enzymes were investigated. *Cellvibrio japonicus* Agd31B (CjAgd31B), a GAA bacterial homolog, reacted with compounds **10-12**, but only a structure with **11** was obtained. Surprisingly, the aziridine remained unopened in a $^{1}S_{3}/^{4}H_{3}$conformation (Figure 3.22); this may be due to it lying in a low energy region as confirmed by the free energy landscape, where the aziridine sits 3 Å above the geometry that would be poised for nucleophilic attack [139].

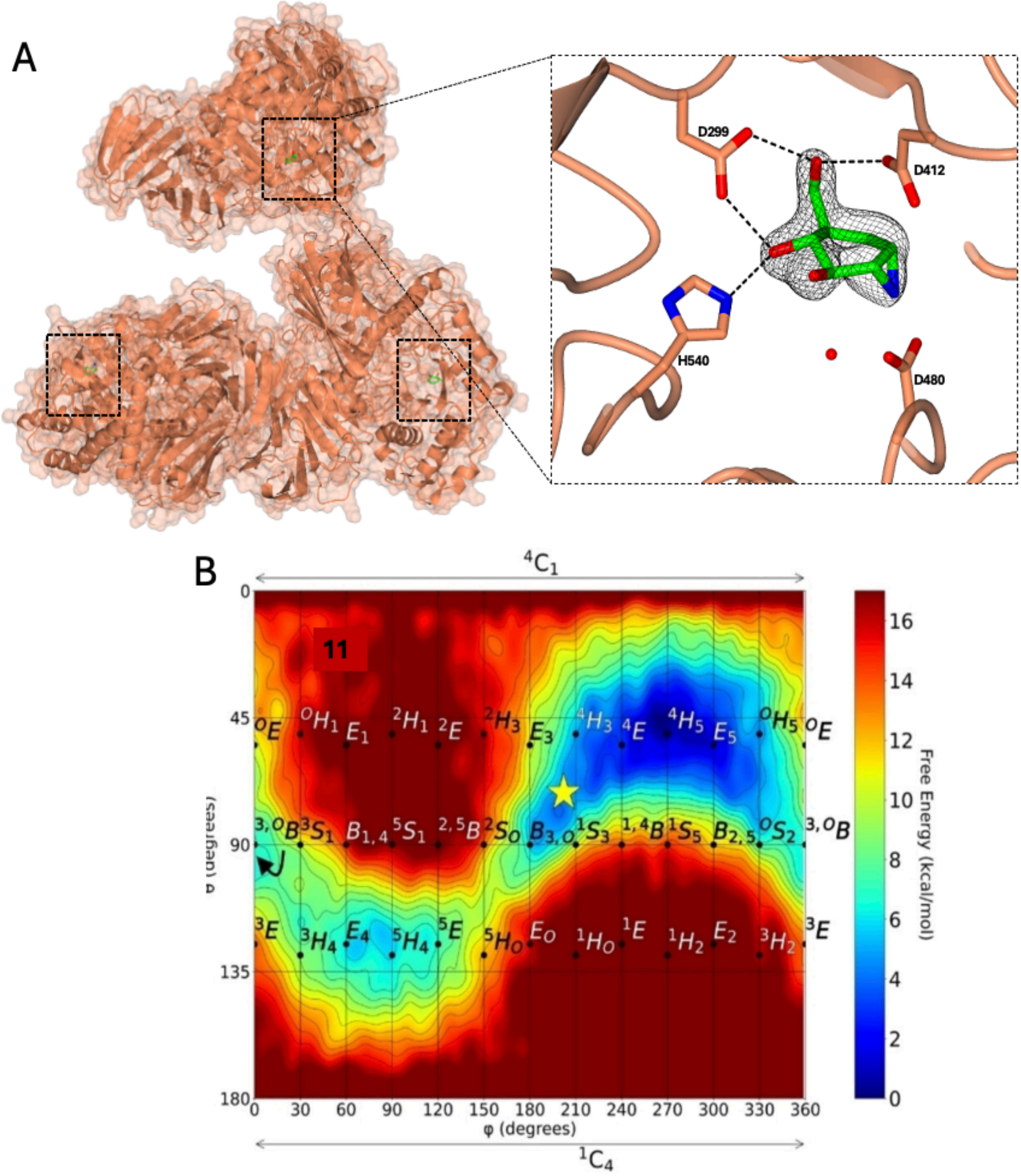


Figure 3.22- Structure of CjAgd31B with 1,2 cyclophellitol aziridine bound in the active site.

Full-length CjAgd31B with ***11*** *bound in the active site (resolved to 1.93 Å) of the catalytic* $(\beta/\alpha)_8$ *domain of the three protein molecules in the unit cell (left panel) [140].* ***11*** *bound in the active site -1 subsite of CjAgd31B forming hydrogen bonding with the catalytic nucleophile*

*D412, acid/base D480 and active site residues H540 and D299. Images created in CCP4mg (version 2.11.0) with F_o-F_c maps (grey) have been displayed in chicken wire and contoured to 0.22 $e^-/Å^3$. (B) Gas-phase conformational free energy landscapes of **11**, where the star indicates the state shown in (B). (B) was conducted by collaborators in the Rovira group.*

3.4.6. Docking Structures of Mus musculus ER α-glucosidase II with Cyclitol Inhibitors

As *Mm* ER-αGluII crystallisation attempts were unsuccessful, docking of inhibitors **10-16** using Schrödinger's Maestro was conducted*,* a computational tool for docking drug discovery. Docking of the 1,6 cyclitols was successful; identical hydrogen bond interactions occurred for each ligand (**13**, **14** and **15**) with D451, D564, R624, D640 and H698 (Figure 3.23A). The non-covalent inhibitors were overlayed with glucose- *Mm* ERαGluII structure (PDB 5H9O), adopting the same $^4C_1$ conformation of the Michaelis complex, suggesting the compounds are correctly orientated for covalent bond formation. The top covalent docking for each of the 1,6 compounds (Figure 3.23B-D) was overlaid with the covalent ER-αGluII complex with fluorinated glucose (PDB: 5HJR), where they all have the same $^1S_3$ conformation with a covalent bond formed with the catalytic nucleophile D564 and hydrogen bonds with D451, R624, D640 and H698.

The docking with 1,2 cyclitols was unsuccessful; in the example (Figure 3.23E), the aziridine/epoxide from compounds **13** and **15** are not positioned for nucleophilic attack, and the covalent docking is not shown as the orientations generated were unachievable. This may be due to the 1,2 position of the cyclitols, as although Maestro is a high-throughput docking software, it disregards protein flexibility [141]. In this case, the use of quantum mechanical molecular dynamics (QMMD) would be necessary as this theory accounts for the flexibility of the proteins as it utilises electrons and reality by accounting for all contributions to energy (electronic polarisation, charge transfer, halogen bonding and covalent-bond formation)[141].

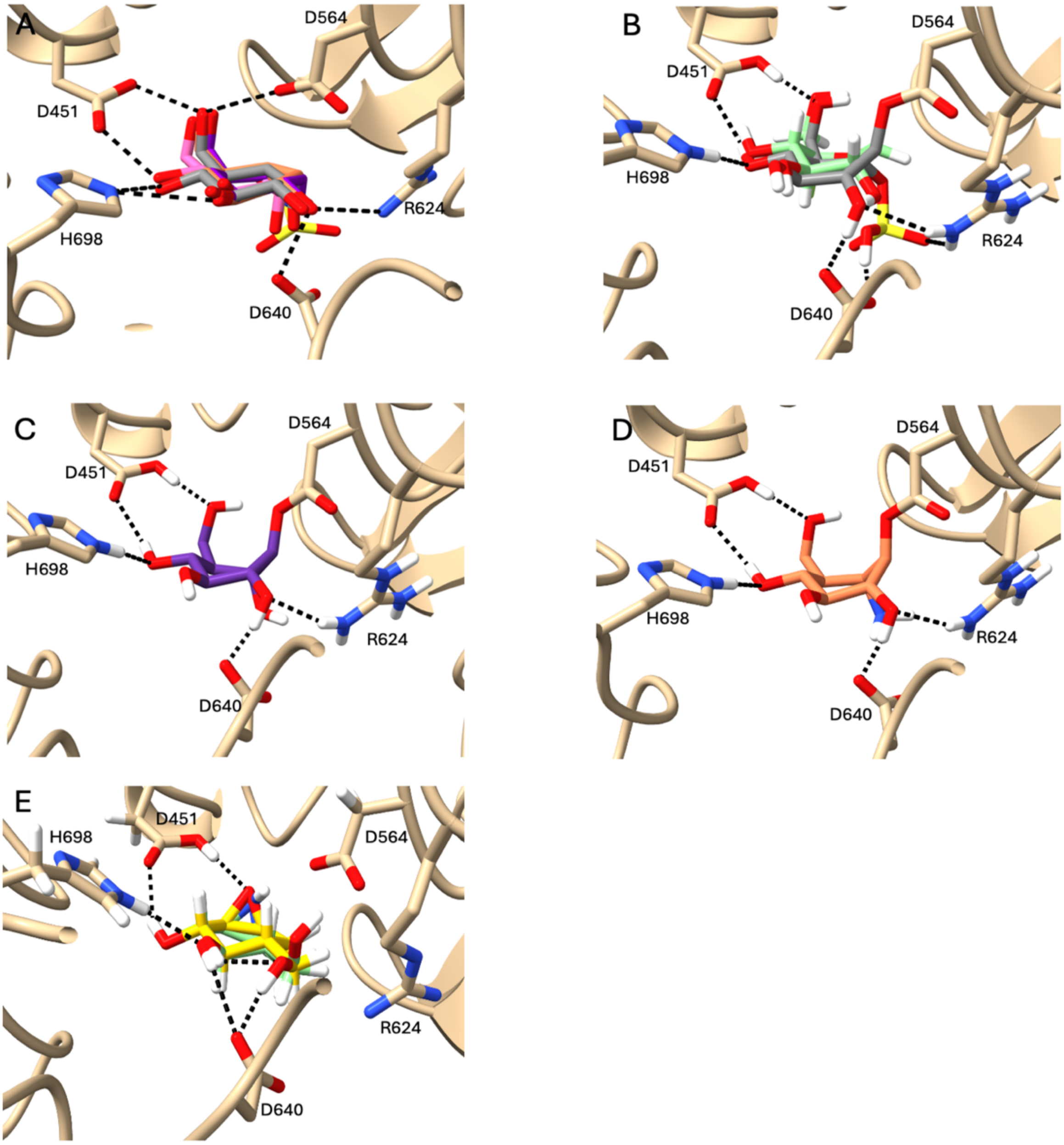
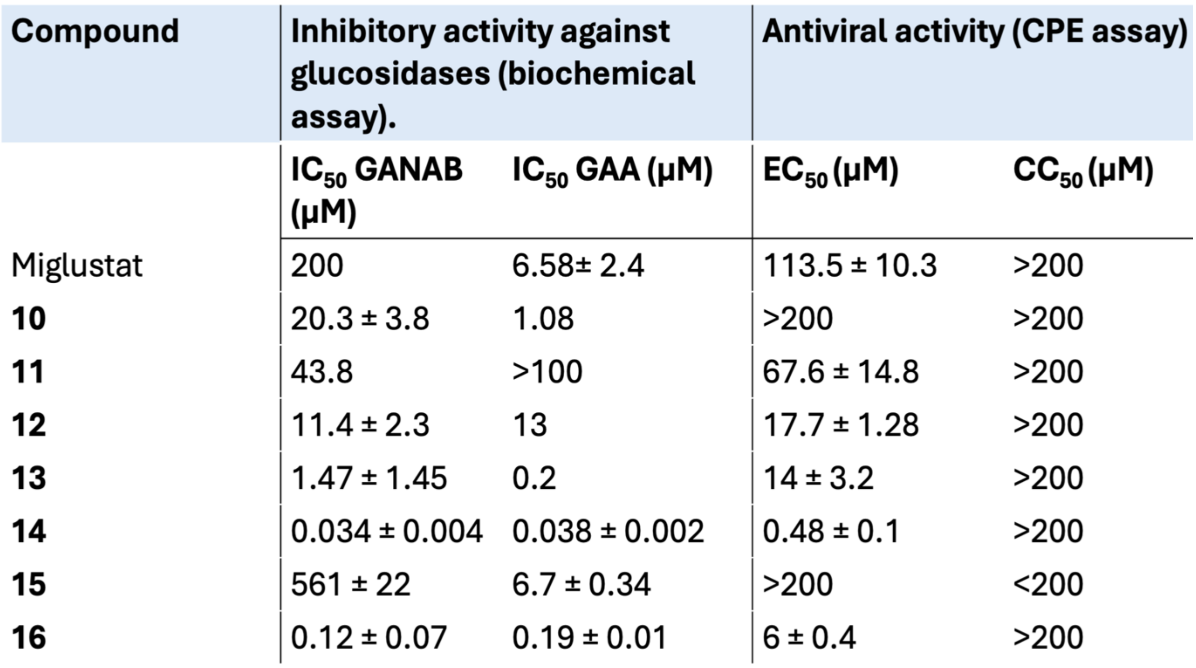


Figure 3.23- Docking of 1,6-cyclitols compounds into the Mm ER *α-glucosidase II active site. (A) Overlay of the top non-covalent dockings of compounds* ***13*** *(orange),* ***15*** *(purple) and* ***14*** *(grey) with glucose (pink PDB: 5H9O). Black lines indicating H-bond interactions with active site residues, including catalytic nucleophile D564.(B) Top covalent docking for compound* ***15*** *(grey), the covalent bond formed between nucleophile D564 along with H-bond interactions and overlaid with covalent 5F-Glc (PDB: 5HJR, green). (C) Top covalent docking of compound* ***14*** *covalently bound to catalytic nucleophile D564 along with H-bonds interactions (black dashes). (D) Top covalent docking of compound* ***13*** *covalently bound to catalytic nucleophile D564 along with H-bonds interactions (black dashes). (E) Top non-covalent docking of 1,2 cyclitols* ***11-12.***

3.5. Summary

In this work, recombinant *Ct* ER-αGluI and *Mm* ER-αGluII were expressed using insect and mammalian expression systems and docking studies of *Mm* ER α-GluII with inhibitors were carried out. Iminosugars have been extensively studied, but their downfall is their poor specificity; here, we considered mechanism-based cyclophellitol derivatives as an alternative that may alleviate off-target side effects.

Additional work conducted by our collaborators (Martijn Hermet, University of Leiden, Appendix 7.3.2) revealed that the lead candidate was the 1,6-*epi*-cyclophellitol cyclosulfate (**14**). In a series of experiments, the potency of the synthesised molecules and their mode of inhibition were investigated. Their inhibitory effect was determined by *in vitro* enzyme assays using 4-methylumbelliferyl-α-d-glucopyranoside (4-MU-α-Glc) as a fluorogenic substrate to measure inhibition with ER-αGluII (GANAB) and human α-glucosidase (GAA) (Figure 3.24)[142]. Modest inhibitory activity was observed for the 1,2 cyclitols (**10**-**12**) for GANAB, but surprisingly, the 1,2 cyclosulfate (**10**) was also a rather potent GAA inhibitor [142]. The most potent ER-αGluII inhibitor was the 1,6-epi-cyclophellitol cyclosulfate (**14**) with a submicromolar IC_{50} (0.034 µM) as well as for GAA (0.038 µM); this is not surprising as the compound was designed as a retaining α-glucosidase inhibitor, while the 1,6 cyclitol aziridine (**13**) and epoxide (**15**) were created for the inhibition of β-glucosidases [102]. The compounds were additionally analysed by cytopathic effect reduction assays (CPE), where Vero E6 cells were pre-treated with an inhibitor, then infected with SARS-CoV-2, and the post-infection viability was measured (Figure 3.24)[142]. Again, **14** was the lead candidate with a submicromolar EC_{50} (0.48 µM), compared to the commercialised inhibitor Miglustat (113.5 µM).



Compound	Inhibitory activity against glucosidases (biochemical assay).		Antiviral activity (CPE assay)	
	IC_{50} GANAB (μM)	IC_{50} GAA (μM)	EC_{50} (μM)	CC_{50} (μM)
Miglustat	200	6.58± 2.4	113.5 ± 10.3	>200
10	20.3 ± 3.8	1.08	>200	>200
11	43.8	>100	67.6 ± 14.8	>200
12	11.4 ± 2.3	13	17.7 ± 1.28	>200
13	1.47 ± 1.45	0.2	14 ± 3.2	>200
14	0.034 ± 0.004	0.038 ± 0.002	0.48 ± 0.1	>200
15	561 ± 22	6.7 ± 0.34	>200	<200
16	0.12 ± 0.07	0.19 ± 0.01	6 ± 0.4	>200

Figure 3.24- Inhibition of ER-αGluII by 1,2 and 1,6 cyclitols.

*Data conducted by collaborators [142]. IC_{50} values for inhibition of 1,6 (**13-15**) and 1,2 cyclitols (**10-12**) were determined using in vitro enzyme activity methods described previously in [102], inhibitors were tested with both with GANAB (ER-αGluII) and GAA (right panel). EC_{50} and CC_{50} (cytotoxic concentration) values from CPE reduction assay with SARS-CoV-2 (left panel). Vero E6 cells were infected with SARS-CoV-2 virus in the presence of 2-fold serial dilution of inhibitor and cell viability examined after 4 days .*

Additionally, the mode of action for **14** was assessed to determine whether inhibition causes changes in viral population, assembly, or neutralisation of SARS-CoV-2. Viral load reduction assays (conducted by Collaborators) confirmed that cells infected with SARS-CoV-2 in the presence of **14** impacted the virus's infectivity but did not affect the viral titre or entry. Therefore, it was suspected that the reduction of infectivity was due to the inhibition of the ERQC (ER-αGluII) affecting the maturation of SARS-CoV-2 glycoproteins, as the S2 fragment of its spike protein completely disappeared with the addition of **14** (Figure 3.25) [142]. The SARS-CoV-2 spike protein is an N-glycosylated protein dependent on the ERQC for its production/maturation [142]. Immunofluorescence microscopy further confirmed the lack of maturation for the spike protein, as its reduced trafficking to the Golgi was observed alongside the reduction of the protein in infected cells (Figure 3.25) [142].

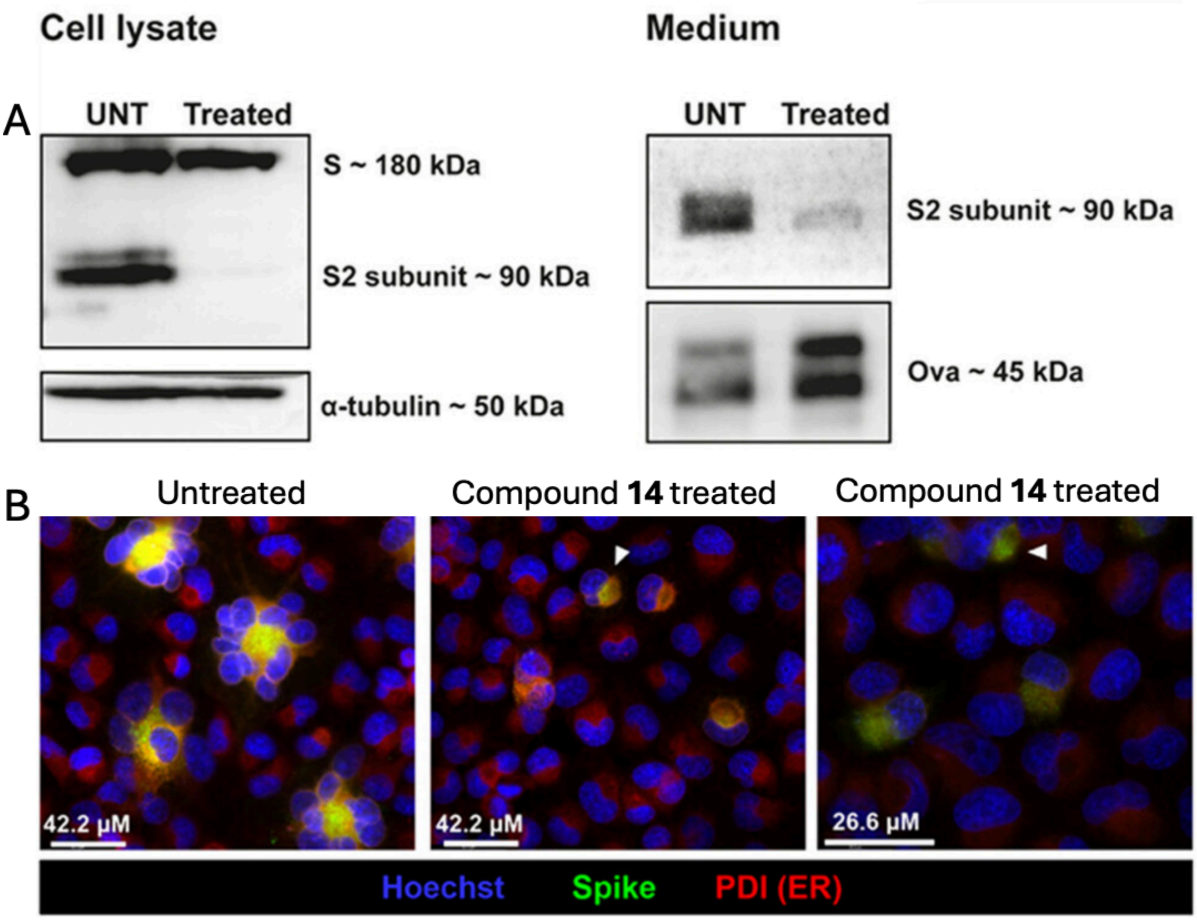


*Figure 3.25- 1,6-epi-cyclophellitol cyclosulfate mode of inhibition for SARS-CoV-2. Image/data generated by collaborators, from [142]. (A) Western blot analysis of SARS-CoV-2 spike protein in the medium and cell lysate of **14** treated (2 µM) H1299/ACE2 cells. The samples were recovered with ovalbumin as a recovery control and concentrated before analysis with the control α-tublin [142]. (B) Immunofluorescence microscopy analysis of SARS-CoV-2 infected H1299/ACE2 cells with compound **14** treatment. Cells were stained using human spike SARS-CoV-2 antibody (green), nuclei (Hoechst, blue) and mouse anti-PDI antibody for the ER (red) [142].*

In summary, iminosugars have been extensively studied for their applications as host-targeted broad-spectrum competitive inhibitors, whereas the mechanism-based cyclophellitol inhibitors discussed have not. The results described in this chapter validate these inhibitors as inactivators of the ERQC enzyme ER-αGluII. Most notably, the 1,6-*epi*-cyclophelltiol cyclosulfate **14** impaired the maturation of the spike protein in the ERQC, causing the loss of the S2 protein fragment and impairing its trafficking to the Golgi [142]. Additionally, the docking studies suggest that **14** reacts covalently with the catalytic nucleophile of ER-αGluII, where it binds in a $^4C_1$ conformation for the Michaelis complex before covalent adduct formation with D564 in a $^1S_3$ conformation similar to previous PDBs. Although the 1,6-*epi*-cyclophellitol cyclosulfate exhibited activity for other retaining α-glucosidases (GAA), compared to iminosugars like DNJ, which have selectivity for both inverting and retaining α-glucosidase, the cyclophellitol compounds have better specificity. In conclusion, the work discussed provides a good foundation for developing cyclophellitol analogues as antiviral agents.

4. Expression of Viral and Human Neuraminidase for the Design of Direct-Acting Antivirals

4.1. Abstract

Influenza is a virus that affects the respiratory system and causes annual epidemics. Entry of influenza virion into epithelial cells involves two major surface glycoproteins, hemagglutinin and neuraminidase (NA). NA hydrolyses sialic acid from the terminal end of glycans so that the newly assembled virus can infect surrounding cells. Targeting NA inhibition through antivirals traps virions on the cell's surface through HA-sialic acid interactions, preventing the virions from infecting surrounding cells. The development of new NA direct-acting antivirals is a highly active field due to the emergence of resistant strains.

Current NA antivirals mimic the sialic acid transition state. In relation to this, the Overkleeft Group (Leiden University) has synthesised a range of *N*-acyl derivatives of the antiviral Oseltamivir. I sought to recombinantly express different influenza NAs and human 2 NA using baculovirus and *E. coli* expression systems to obtain structural insights into the neuraminidase-inhibitor complexes. Biochemical analyses were also carried out, including Activity-based probe profiling, thermal shifts and viral plaque assays for the viral NA, while the human NA was successfully crystalised. The production of viral NA and subsequent analysis will enable its structural analysis, as described in **Chapter 5**. Analysis of these compounds' ability to inhibit viral infection in the viral plaque assays demonstrated that the *N*-acyl and unsubstituted aziridine derivatives outperform the commercialised inhibitor Oseltamivir. Surprisingly, the unsubstituted oseltamivir aziridine exhibited competitive inhibition instead of covalent inhibition and was the standout candidate of the inhibitor series. Our work has demonstrated the use of a new class of mechanism-based NA inhibitors that could combat the growing concern of antiviral resistance.

4.2. Introduction

Sialic acids are a diverse group of nine-carbon keto aldonic acids present in bacteria, viruses, fungi, protozoa, vertebrates and invertebrates, located on the terminal ends of glycoproteins and glycolipids on the cell's outer surface [143]. The most common variant is glycolylneuraminic acid (Neu5Gc), which is present in many animal species except for humans, instead of 5-acetamido-*D-glycero-D-galacto*-2-nonulosonic-acid (Neu5Ac) is the most abundant sialic acid in humans [144, 145]. Sialic acids have a significant structural diversity resulting from modification at C4- C9 (Figure 4.1) via *O*-acetyl, sulfate, methyl and lactate groups and further diversity is added through different alpha linkages (2,3/6/8) to galactose, N-acetyl galactosamine or N-acetyl glucosamine (Figure 4.1B) [144, 146]. This extensive diversity forms what is known as the sialome in the glycome.

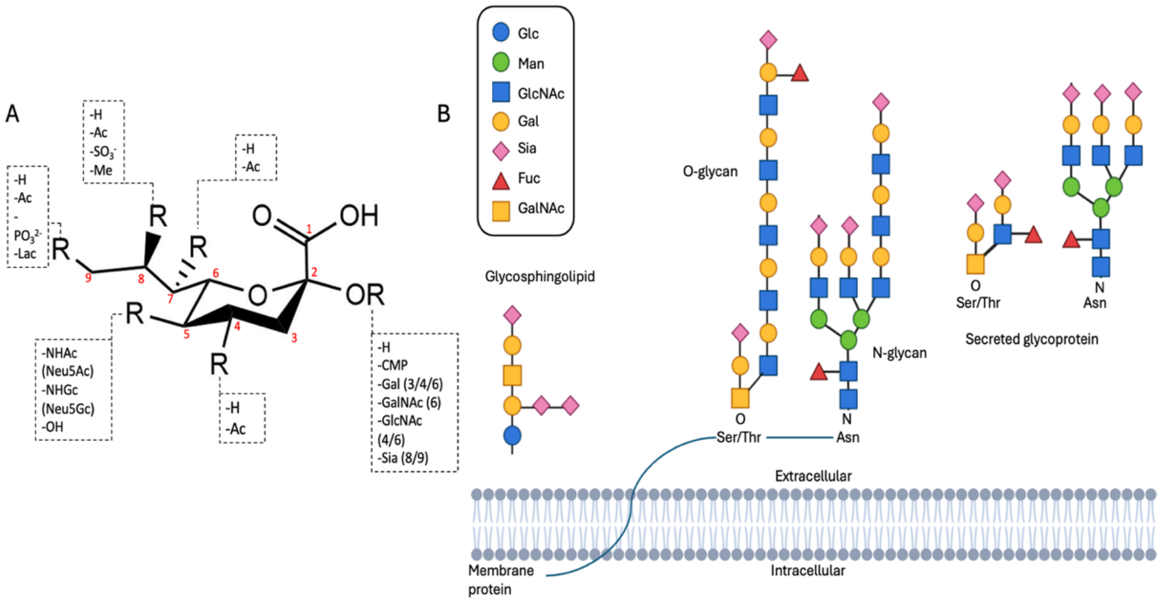


Figure 4.1- Diversity of sialic acid.

(A) Structure of Sialic acid, including the modification sites R2 and R4-9 with the natural modifications indicated. R5 includes modifications determining the sialic acid types N-acetylneuraminic acid (Neu5Ac) and N-glycolylneuraminic acid (Neu5Gc). (B) Diversity of sialic acid N- and O- linkages with glycoproteins and glycosphingolipids. Created in BioRender.

Sialic acid is produced by mammalian biosynthesis pathways or derived from dietary sources, where it is conjugated to a cytidine 5'monopohosphate (CMP) and transported to the Golgi, where the CMP-Neu5Ac is used as a donor to transfer sialic acid to a glycoprotein or glycolipid via sialyltransferases [145]. Modifications occur in the Golgi, where *O*-acetylation is the most abundant modification in humans. The type of modification gives sialic acid its function; for example, sulphated and acetylated sialic acids are receptors for viruses, whereas methylated is not [143].

Due to their diversity, sialic acids are implicated in various pathological and physiological processes because of their negative charge and hydrophobicity [144]. Their negative charge is utilised in many biological systems. For example, in the blood for cell-cell repulsion of erythrocytes [144]. Furthermore, sialic acids coat cell membranes, where their charge forms a highly viscous aqueous environment due to the trapping of water molecules, forming the glycocalyx [144]. The sialyation of circulating proteins and hormones increases their half-life, which is exploited in the biotherapeutic industry [144]. The subject of this chapter looks at sialic acids in viral disease, where sialic acid recognition initiates viral infection, and their degradation by viral neuraminidases (NA) propagates the disease.

4.2.1. Neuraminidase Role and Function

NAs are found in bacteria, viruses, and vertebrates, participating in the regulated degradation of Sia-glycoconjugates. The first step of sia-glycoconjugate degradation is removing sialic acid from the non-reducing termini catalysed by NAs [147, 148]. Viral and bacterial NAs have been well-characterised structurally compared to their mammalian counterparts. Despite the low 35% sequence identity between viral and bacterial NA, they all share the same tertiary structure, six-bladed β- propeller, conserved arginine triad and catalytic residues [147, 148].

Viral and mammalian NA (E.C 3.2.1.18) belong to the glycoside hydrolase (GH) families 34 and 33. The exact mechanism of catalysis for both families has long been speculated. In 2004, the reported structure for the GH33 *Trypanosoma cruzi* (*Tc*Ts) confirmed that NA catalysis occurs through a double displacement mechanism [149]. However, no carboxylate group was favourably positioned to act as the nucleophile. Many glycosidases act via a covalent glycosyl-intermediate using an acid/base and carboxylate nucleophile as the catalytic residues, but instead, NA uses a tyrosine [149]. The X-ray structure revealed a tyrosine favourably positioned; choosing a tyrosine as the nucleophile may be due to the charge of the substrate sialic acid. Using a carboxylate nucleophile would be unfavourable as its negative charge would repel the negatively charged sialic acid [150].

A retaining mechanism via tyrosine was further confirmed through the reaction of a 2,3 di fluorosilicic acid (DFSA) [151, 152], where LC-MS of the NA-DFSA complex showed that tyrosine and the glutamic acid acting as a nucleophile couple; this is the first confirmed example of a retaining glycosidase with an aryl glycosidase intermediate [150, 153]. The deprotonated glutamate acts as the base catalyst for tyrosine by relaying its charge onto nucleophilic tyrosine transiently to enhance its phenolate character through the formation of a hydrogen bond during the Michaelis complex; once the covalent intermediate is formed, hydrogen bonding occurs with water instead [7]. Therefore, the GH34 and GH33 families hydrolyse sialic acid via a double displacement mechanism by retaining the anomeric centre (Figure 4.2)[7]. Alongside X-ray crystallography structures and the creation of the deoxyfluorosugar, DFSA was trapped with tyrosine, forming an intermediate with a ${}^{2}C_{5}$ conformation [19, 149, 150]. This matches the conformational itinerary confirmed by the GH33 mentioned above *Tc*Ts ${}^{6}S_{2} \rightarrow {}^{4}H_{5}^{\ddagger} \rightarrow {}^{2}C_{5}$ [19, 149, 150].

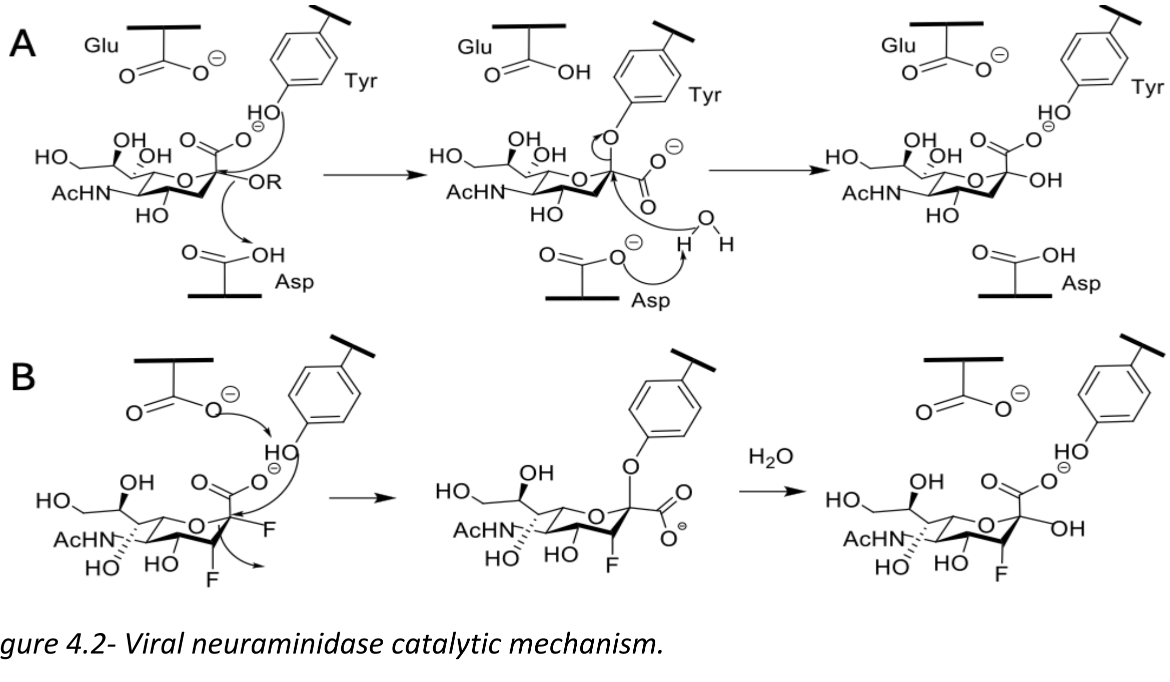


Figure 4.2- Viral neuraminidase catalytic mechanism.

(A) In the double displacement catalytic mechanism of GH34 enzymes, nucleophilic tyrosine attaches to the anomeric centre C2 and forms a covalent intermediate via a semi-planar oxocarbenium transition state. Water molecules are activated by Aspartic acid attacking the

anomeric centre C2 before forming a transition state and then releasing sialic acid. (B) Mechanism of DFSA with viral neuraminidase. The C-3 electronegative fluorine destabilises the oxocarbenium-like transition state, while the C-2 anomeric fluoride leaving group allows for the accumulation of the covalent intermediate [150, 153].

4.2.2. Human Neuraminidases and Disease

Human neuraminidases (NEU) are potential therapeutic targets for multiple modern-day diseases due to their involvement in many signalling pathways in neurodegeneration, cancers, developmental and cardiovascular disease. They are implicated in biological processes: antigenic expression, apoptosis, cell motility and signal transduction [148, 154]. Mammalian NEUs are exo-acting enzymes [72]. The mammalian NEUs have different substrate specificity, enzymatic properties, cellular expression and localisations due to their distinct roles in cell physiology [72, 155]. Their different localisations are as follows: NEU1 (lysosome), NEU2 (cytoplasm), NEU3 (plasma membrane-associated), and NEU4 (mitochondria and lysosome) (Figure 4.3)[154].

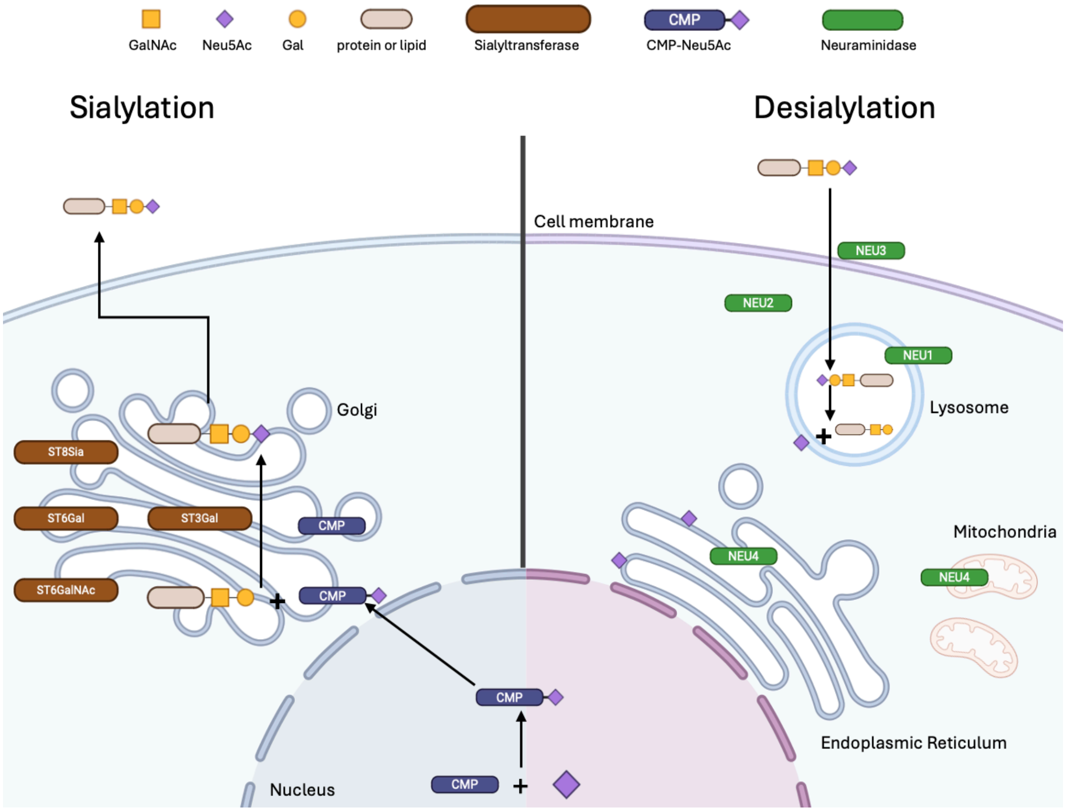


Figure 4.3- Sialylation and desialylation process of a protein or lipid in cells.

Two main types of enzymes regulate sialic acid levels: sialyltransferase and neuraminidases. Sialyltransferases perform sialylation of lipids and proteins using cytidine monophosphate

(CMP) – Neu5Ac donor, where Neu5Ac is transferred to the end of the glycan. For desialylation, neuraminidases remove the terminal sialic acid from the glycan [156]. (Image based on [156] and created in BioRender).

NEU2 is involved in the first step of degradation for sialic acids and is usually expressed at low cellular levels. Still, it has broad substrate specificity and acts on oligosaccharides, glycoproteins, and gangliosides [155]. Little is known about NEU2; its effects on various cellular pathways have implicated the protein in diseases such as cancer, but its exact role is still to be uncovered [155]. The first crystal structure of NEU2 was determined in 2005 and again showed that despite the low sequence homology with other GH33 NA, they shared the canonical six-bladed β-propeller [148]. Due to its low cellular expression, NEU2 has been challenging to study, but its activities have been implicated in cancer. A key characteristic of highly metastatic cancer is suppressed NEU2 expression, causing an increase in the cellular population of α2,3- and α2,6-linked sialic acids [157, 158]. Thus, many groups have looked into the transfection of NEU2 into cancerous cells; this resulted in the inhibition of metastasis as NEU2's increased activity causes decreased cell invasion and motility [157-159].

The abundance of structures for viral NA contributed to the design of many potent inhibitors, but this is not the case for NEU, as the only reported structures in the PDB are for NEU2 (Table 4.1). As there are no structures for NEU1, NEU3 and NEU4, structure-based information on the protein has had to be obtained via molecular modelling based on the NEU2 structure [160]. These models have revealed that NEU2, NEU3 and NEU4 have 34-40% sequence identity to each other, whereas NEU1, compared to the other human NEUs has 19-24% sequence identity [160]. The lack of structures is due to problems with the expression of NEU1, NEU3 and NEU4 as they tend to aggregate [160].

The lack of inhibitors for NEU2 is partly due to problems with expression, a lack of sailoside probes, and screening methods [154]. When designing inhibitors for NA, it's essential to look at how enzymes from different organisms accommodate their substrate, sialic acid. A comparison of viral, bacterial, and human NAs reveals that modification at the C5 position of Neu5Ac can be exploited to design selective NEU2 inhibitors as these derivatives would be well accommodated in their active site [148, 154].

Table 4.1- Structures of human neuraminidase (NEU2) from the PDB.

PDB	**Expression Source**	**Ligand**	**Res (Å)**	**Reference**
1SNT, 1SO7, 1VCU,	*E. coli*	"apo" and NEU2 in complex with 2-Deoxy-2,3-D hydro-N-acetyl-neuraminic acid and Maltose	1.49 - 2.85	Chavas et al (2005)
2F0Z, 2F1O, 2F11, 2F12, 2F13, 2F24, 2F25, 2F26, 2F27, 2F28, 2F29	*E. coli*	NEU2 in complex with inhibitors, including commercialised viral neuraminidase inhibitors for influenza treatment.	1.67 - 2.90	Chavas et al (to be published)
4NC5, 4NCS	*E. coli*	2,3-difluorosialic acid	2.2-2.51	Buchini et al 2014

4.2.3. Viral Neuraminidase and Influenza

Influenza infects the upper respiratory tract (nose, throat and lungs). Each year, the flu is responsible for a rise in mortality and morbidity rates during the autumn and winter seasons, where the virus is spread from person to person via airborne particles [161]. Typically, the immune system clears the virus after the individual has experienced symptoms such as high fever and head and body aches [161]. Vulnerable groups such as infants, elderly and immunocompromised are at higher risk for developing complications like pneumonia [161]. These individuals are advised to start treatment within 24 hours of symptoms for effective results and receive their annual flu vaccine [161]. The World Health Organisation (WHO) estimates that each year, there are still 3-5 million cases of severe infection and an estimated 290,000-650,000 deaths. Therefore, the development of antivirals is still essential for treatment, especially with the occurrence of treatment-resistant strains.

4.2.3.1. Influenza Virus

The influenza virus is a part of the *Orthomyxoviridae* family, a group of enveloped viruses with a segmented negative-sense single-stranded RNA genome [162]. The influenzas virion is encapsulated in a lipid membrane derived from the infected host cell, which protects its genome [163]. Its genome consists of eight single-stranded RNA genes that encode ten proteins [163]. These genes are released upon infection into the host cell, which initiates transcription and influenzas replication [163].

There are four main types of influenza. Types A and B are responsible for the annual seasonal epidemics as they are more virulent, while types C and D are not known to cause infection in humans. Influenza viruses are divided into subtypes based on their two surface glycoproteins, hemagglutinin (HA) and NA, which are further split into two phylogenetic groups (Figure 4.4), distinguished predominately by structural differences. Two groups of NAs are categorised based on their antigenic differences; group 1 consists of N1, N4, N5, and N8; group 2 is N2, N3, N6, and N7 [164]. In the corner of the active site is a '150 loop' (Figure 4.4), which is formed from residues 147-152 and can exist in two different conformations; in group 1 NAs, the loop forms an extra cavity next to the active site upon substrate binding, but this is not formed in group 2 [164]. There are over 130 influenza A subtype combinations of NA and HA, but this number is likely underestimated due to antigenic shift and drift [164].

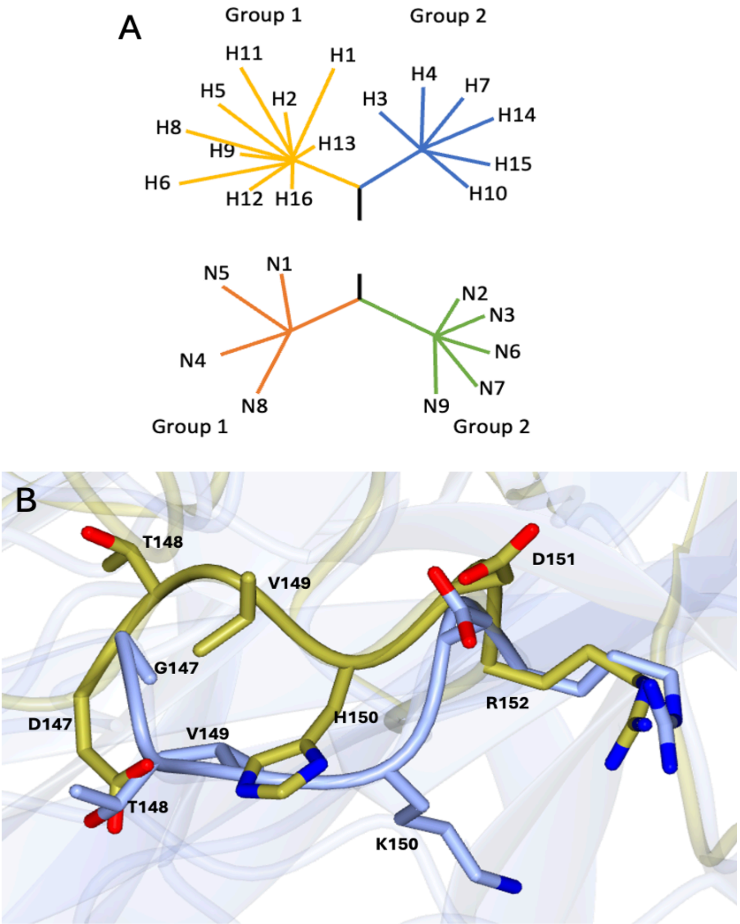


Figure 4.4- Phylogenetic groups of hemagglutinin (HA) and neuraminidase (NA) of Influenza A. (Image based on [164]). (A) HA are split into two phylogenetic groups based on their antigenic properties. NA groups are categorised based on their structural differences; group 1 contains NAs, which have a 150 loop that undergoes a conformational change upon substrate binding (B), forming a cavity next to the active site [164].

Entry of influenza into cells involves two major surface glycoproteins, HA and NA, which both recognise sialic acid on the host cell surface [164]. During the viral lifecycle, HA is involved in the virus binding to α2,6-linked sialic acid on the cell's surface. The tight binding of HA to sialic acid induces endocytosis, encapsulating the virus in an endosome upon entry [163]. The low pH of the endosome triggers the fusion of HA with the endosome membrane, releasing the eight RNA segments into the nucleus, where they are transcribed [163]. The host cells' machinery is then utilised to replicate, assemble and release new viruses. During virus replication, NA catalyses the removal of *N*-acetyl neuraminic acid from sialic acid moieties on the host cell's surface; this prevents the binding of newly assembled progeny viruses' HA's to sialic acid on the infected cell's surface [164]. Without NA activity, the progeny viruses become trapped on the cell's surface through HA-sialic acid interactions, preventing them from infecting surrounding cells (Figure 4.5) [164]. Current treatments of influenza A involve preventing virion release and replication, primarily through NA inhibition [165].

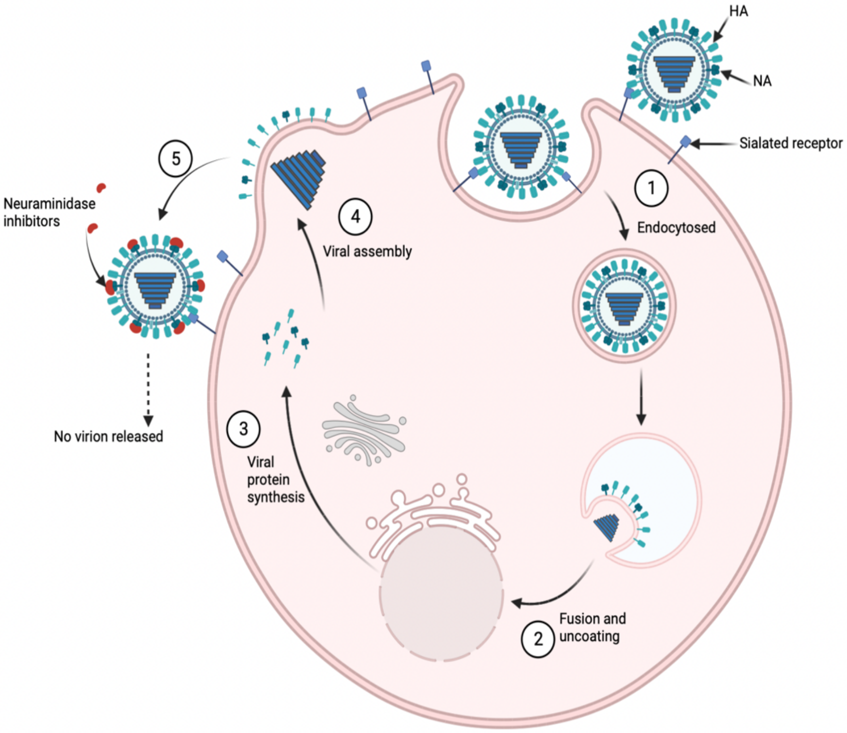


Figure 4.5- Influenza replication cycle (Image created in BioRender).

Firstly, HA recognises the sialated receptor on the host cell surface, allowing the virus to be endocytosed and trafficked into an endosome where the low pH triggers the release of viral RNA into the cytoplasm, where it undergoes transcription and translation of viral proteins. The virus assembles and is exocytosed, where HA will again bind to the sialated receptors on the *host cell's surface; in order for the virus to be released, NA cleaves sialic acid, releasing the virus. However, in the presence of NA antivirals, this cleavage does not occur, and the virus becomes trapped on the cell surface. (Image based on [166]).*

Historically, the 1918, 1957, and 1968 pandemics had high mortality rates due to evolutionary changes caused by antigenic shift or antigenic drift [7]. The 1918 influenza pandemic resulted from antigenic drift, a rare occurrence in which major mutations developed in HA, changing its linkage specificity and allowing previous animal influenza viruses to jump to humans [167]. Mutations in HA and NA due to antigenic drift can lead to the positive selection of immune-evading and antiviral-resistant viruses. Another rare form is antigenic shift, when two influenza strains enter the same cell. This leads to random reassortment of genome segments, which recombine to form a new and potentially more virulent strain (Figure 4.6)[167].

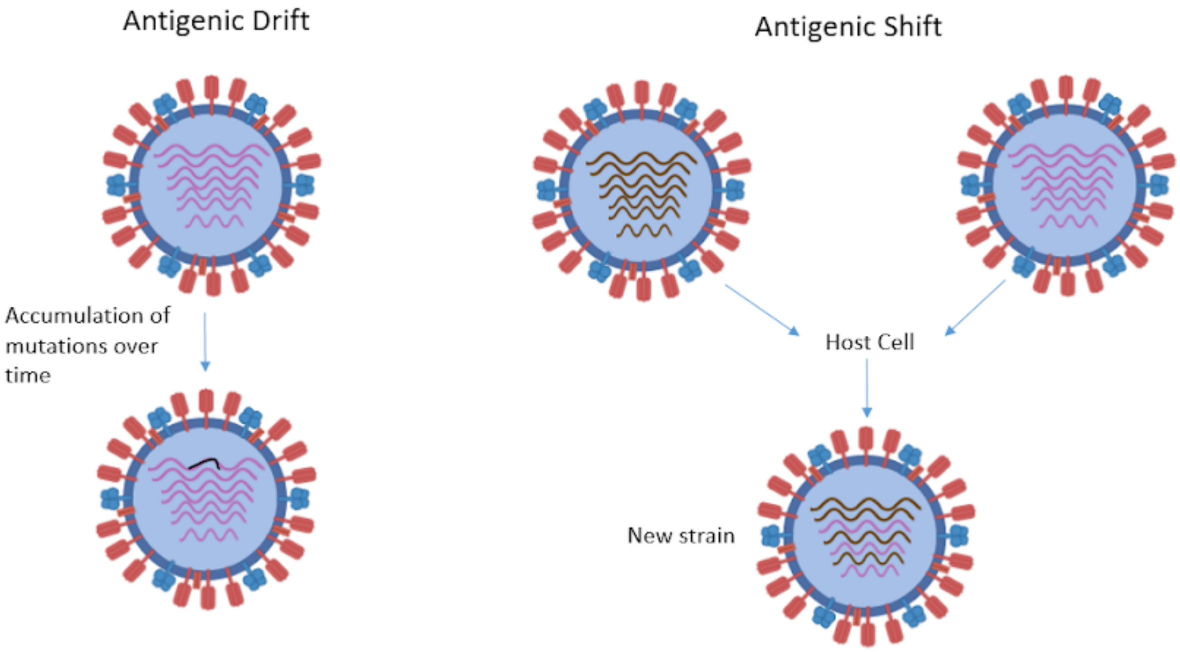


Figure 4.6- Mechanism of antigenic shift and drift resulting in new mutant strains.
The accumulation of mutations over time causes the occurrence of new strains (antigenic drift). New strains can be generated by combining two different genomes during a rare occurrence where two strains enter the same cell simultaneously (antigenic shift).

4.2.3.2. *Viral Neuraminidase*

Influenza NAs form non-covalent interactions between four identical NA subunits, creating a homotetrameric type II integral membrane protein embedded into the viral envelope, comprising four distinct structural domains (Figure 4.7) [168]. The N-terminal cytoplasmic domain is nearly 100% conserved across all NAs as it is involved in virus morphology; mutations here are associated with reducing virion yields [169]. The stalk domain varies in length depending on the NA subtype; the cysteine residues allow for disulphide bonds to form between the stalk and the tetramer [162]. The head domain consists of six-bladed β-propeller monomers, which associate to form tetramers. A large cavity forms the charged active site containing 9 conserved catalytic residues [170].

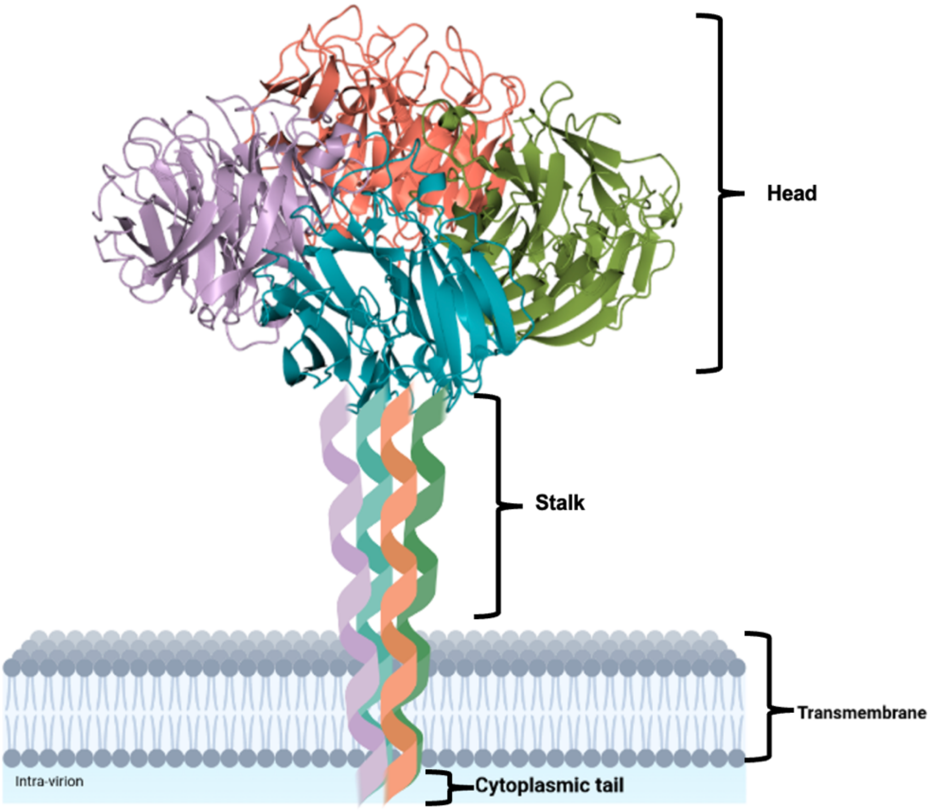


Figure 4.7-NA structure in the viral membrane.

NA is broken into four distinct structural areas. The head domain comprises four identical monomers that form a tetramer and are responsible for the hydrolysis of sialic acid. The head domain is attached to the stalk domain via disulphide bonds and is embedded in the cell membrane.

Recombinantly expressed viral NA's have been structurally resolved with X-ray crystallography and have consistently shown a tetramer with four-fold symmetry associating non-covalently. However, when recombinantly expressed without a tetramerisation domain, the tetramers can disassociate into monomers as the non-covalent interactions are weak [168]. Another phenomenon with recombinant NAs is that they will adopt an open or a closed conformation depending on the subtype. In the context of the viral membrane (Figure 4.8), NAs adopt a closed tetramer [171]; replicating this conformation recombinantly is important to understand NA antigens and stability for the structure-based design of vaccines and drugs. In this chapter, the gene expression of both the open and closed conformations of NA from H1N1 was investigated to structurally characterise the proteins with ligands, for past PDB structures see Table 4.2

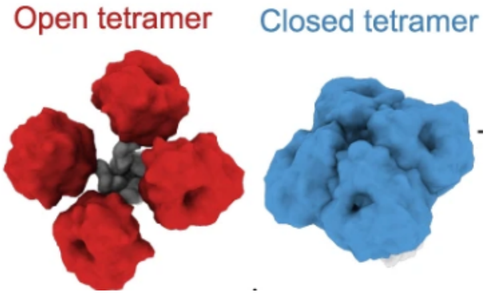


Figure 4.8- Open and closed conformations of viral neuraminidase. Image from [168].

Table 4.2- Structures of N1 from H1N1 influenza.

PDB	Expression Source	Ligand	Res (Å)	Reference
3TI3, 3TI4, 3TI5, 3TI6	Insect	N1 in complex with commercialised antiviral inhibitors.	1.7-1.9	Vavricka, et al
4B7J, 4B7M, 4B7N, 4B7Q, 4B7R	*Gallus*	N1 in complex with commercialised antiviral inhibitors.	1.9-2.8	Van Der Vries, et al

4.2.3.3. Influenza Prevention and Treatment

The severity of annual influenza epidemics varies with each different strain. At the beginning of each year, the World Health Organisation (WHO) holds yearly meetings where specialists review influenza and vaccine information for the northern and southern hemispheres to predict which strain populations should be immunised against. In case of an outbreak, we have an arsenal of antiviral treatments, but the occurrence of new resistant strains poses a global health risk as there are no curative agents for influenza.

4.2.3.3.1. Influenza Vaccines

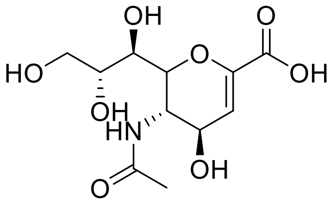
NA has been predominantly ignored in the creation of seasonal influenza vaccines due to the lack of knowledge of NA-directed humoral immunity. Influenza vaccines are primarily HA-based, even though over 25 years ago, a trial conducted on NA-based vaccines showed promising results compared to the HA-based vaccines at the time [168, 172]. This is mainly due to the disconnect between knowledge of structural and antigenic stability of NA outside of the viral membrane-bound state, as recombinantly produced NA can adopt different conformations. The use of HA in vaccines has been linked to a decrease in the extent of disease, whereas NA antibodies are linked to decreased virion release and disease severity [172].

4.2.3.3.2. Influenza Antivirals

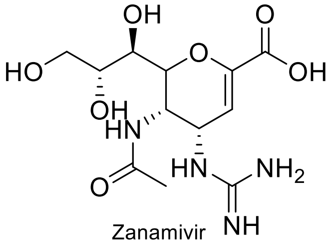
The first reported inhibitor for NA was 2,3-dihydro-2-deoxy-N-acetylneuraminic acid (DANA), an oxocarbenium ion-like transition state analogue inhibitor of NA within the micromolar range (Figure 4.9) [173]. Although DANA was the first inhibitor of influenza reported, it was never commercialised. Instead, DANA was used as a template for designing more potent inhibitors, some of which have become commercialised [174].

Zanamivir (4-guanidino-Neu5Ac2en) is a pyran scaffold analogue of DANA but instead with a guanidinium group at C4. Adding the charged group significantly increased inhibitor activity due to additional interactions [174]. However, zanamivir has low bioavailability; instead, carbocyclic scaffolds were investigated. Hydrophobic interactions can be employed for increased antiviral potency; these interactions are utilised in Oseltamivir through the action of the 3-pentyl ether side chain [174]. Again, an increase in affinity was achieved, but the development of antivirals still faced the bottleneck of bioavailability [174]. Instead, the prodrug inhibitor oseltamivir phosphate was designed, where hepatic carboxylases process oseltamivir phosphate into the active inhibitor oseltamivir carboxylate [174]. Remarkably, the DANA derivative peramivir achieved another significant increase in potency. This cyclopentane has both the guanidino moiety from zanamivir and the hydrophobic side chain from oseltamivir, allowing it to utilise polar and hydrophobic interactions [72].

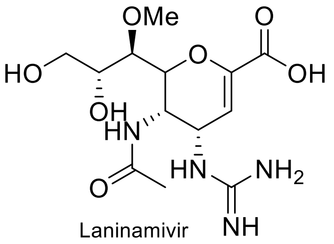
To date, only four inhibitors have been commercialised for use (Zanamivir, Oseltamivir, Peramivir, and Laninamivir) (Figure 4.9). Still, the design of other small molecules has been explored using these inhibitors as scaffolds. This is due to the emergence of resistant Influenza strains, such as the N1 mutant H275Y. The bulkiness of the tyrosine substitution caused a 100-fold decrease in the effectiveness of Oseltamivir treatment as the 3-pentyl side chain struggles to be accommodated in the mutant active site [175, 176].



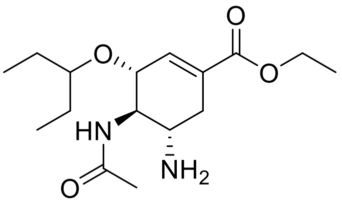
DANA
IC_{50} 8600 nM



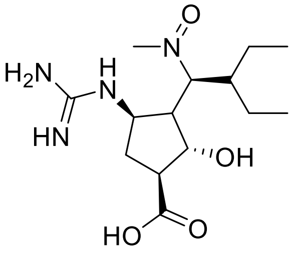
Zanamivir
IC_{50} 5nM
Oral inhalation



Laninamivir
IC_{50} 0.62 nM
Nasal inhalation



Oseltamivir
IC_{50} 2.5 nM
Oral administration



Peramivir
IC_{50} 0.16 nM
Intravenous administration

Figure 4.9- Transition state neuraminidase inhibitors.
Structures of transitions state mimic inhibitors and their IC_{50} compared to the first developed TS inhibitor for neuraminidases DANA [72]. Zanamivir and Laninamivir are TS mimetics with a pyran scaffold and Oseltamivir and Peramivir are TS mimetics with carbocyclic scaffold.

4.2.3.3.3. Small Molecule Inhibitors for Viral Neuraminidase

Many molecules have been developed to inhibit viral neuraminidases; many are analogues of the substrates sialic acid and commercialised antivirals. With the emergence of the Oseltamivir-resistant H275Y N1, a series of N-substituted Oseltamivir analogues have been developed [177-179]. Before the discovery of 150 loop in 2006, group 2, which did not include the loop, was used for the rational design of NA inhibitors. Since group 1 discovery, antivirals exploiting this structural difference have been developed, such as oseltamivir analogues (Figure 4.10) [177, 178]. N-substitutions of the guanidinium group at C5 include N-methyl, N-hydroxyl, N- biphenyl and N- phenyl thiophene, which results in inhibitory profiles similar to oseltamivir, except they are effective against Oseltamivir resistant strains [177, 178]. Other inhibitors have been developed that target the nucleophile and the arginine triad, where the addition of a phosphate group instead of the C1 carboxylate group on Neu5Ac improves inhibition due to the electrostatic interactions with the phosphate and the arginine triad [72].

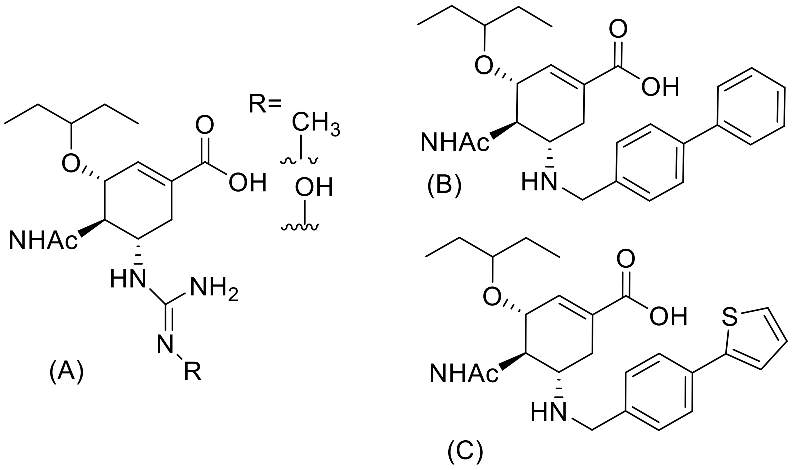


Figure 4.10- *Oseltamivir analogues.*

N-substitutions of the guanidinium group at the C5 position (A), where R is a methyl or hydroxyl group. The compound targets the 150 loop through the addition of biphenyl and phenyl thiophene. (B-C) N-substituted oseltamivir derivatives [177, 178].

NA exists as a homotetramer on the viral surface; this can be used to develop multivalent inhibitors linked together to react with all the active sites of the NA tetramer (*Figure 4.11*). Multiple dimeric, trimeric, and tetrameric inhibitors have been explored; some use zanamivir and are linked together through differently spaced linkers [180-182]. Using the correct linker length is essential so that it matches the molecular dimensions of the tetramer for binding all four active sites simultaneously [182]. The previously mentioned DFSA is a non-clinical inhibitor used to determine the mode of catalysis for viral neuraminidases; the inhibitor was further explored as a potential new class of antivirals (*Figure 4.11*) [183]. The electronegative fluorine at the C3 position destabilises the oxocarbenium-like transition state while the C2 fluoride rapidly forms a covalent bond with the nucleophilic tyrosine; the slow turnover causes the accumulation of a covalent species [183]. In animal models, DFSA shows an inhibitor profile comparable to zanamivir; initial binding was improved by substituting the C4 position with a guanidinium or amine group [183]. Adding these substituents allows the inhibitor to utilise charges for improved binding. Mechanism-based inhibitors based on Neu5Ac substrate offer a unique advantage over non-covalent ones as the barrier to resistance would be hard to overcome due to the inhibitors targeting the catalytic nucleophile.

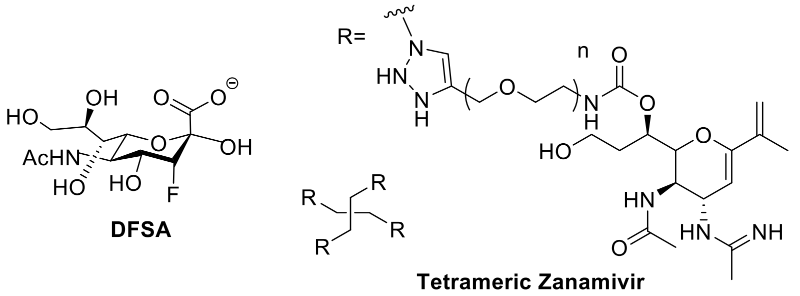


Figure 4.11- Chemical structures of 2,3-difluorosialic acid (DFSA). Tetrameric zanamivir linked by oligo ethylene glycol (OEG) were n=12.

4.2.3.3.4. Activity-Based Probes for Viral Neuraminidases

Current methods for quantifying Influenza include the detection of viral RNA using real-time (RT) PCR. However, using RT-PCR requires knowledge of sequence and primers, which would be a significant bottleneck for the development of vaccines during a global pandemic. In this case, using antibodies for detection would not be helpful as these are strain-specific; therefore, probes that work independently of viral strain are needed. The development of NA probes would be beneficial in this circumstance; as discussed in **Chapter 1**, ABPs have proven useful in proteomics studies, where they can selectively target and become activated upon cleavage, labelling the protein of interest. Various probes are designed to rapidly detect NAs, including a zanamivir-biotin conjugate that successfully captured influenza virions [184]. In addition to using an inhibitor as the specificity motif, other probes have utilised the substrate Neu5Ac conjugated to various reporter groups, such as difluoromethylphenyl, phenoxy-dioxetane (*Figure 4.12*), and 2-benzothiazole-2-yl-phenol [184-187].

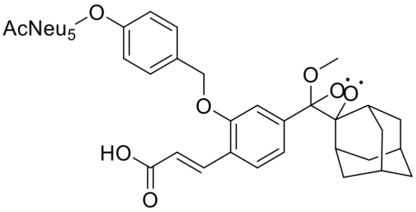


Figure 4.12- phenoxy-dioxetane chemiluminescent probe [184].

Neu5Ac conjugated via a linker to 1,2-adamantylidene-dioxetane luminophore with acrylic acid substituent. The probe undergoes hydrolysis via a 1,6 elimination process, releasing the phenolate intermediate. An excited benzoate is produced by a chemiexcitation process, releasing a green photon [184].

4.2.4. Research Aims

4.2.4.1. Human Neuraminidase Aims

This work aims to establish NEU2 expression and subsequent crystal generation for X-ray crystallography studies in York. The work will be continued when inhibitors/probes for NEU2 are synthesised.

4.2.4.2. Viral Neuraminidase Aims

This work aims to expand the portfolio of viral neuraminidase inhibitors to inform the design of more potent antivirals to treat emerging resistant influenza strains. The work presented in this chapter is in collaboration with the Overkleeft group (University of Leiden), who synthesised the inhibitors (Figure 4.13), and the Rovira group (University of Barcelona), who conducted quantum mechanics/molecular mechanics (QM/MM) simulations to inform the design and mechanism of inhibitor binding. The synthesised compounds include mechanism-based inhibitors, which selectively target NAs catalytic nucleophile; additionally, an ABP was synthesised to perform selective labelling of recombinant NA. I worked alongside our collaborators at The Francis Crick Institute to conduct a viral plaque assay to investigate the effects of these inhibitors on live influenza viruses.

Specifically, in this chapter, I report the expression and biochemical characterisation of the N-acyl aziridines (Figure 4.13) and free aziridine Oseltamivir **21** inhibitors with NA using the baculovirus expression system and activity-based probe profiling. **Chapter 5** will discuss the structural characterisation of the protein-ligand complexes using cryo-EM**.**

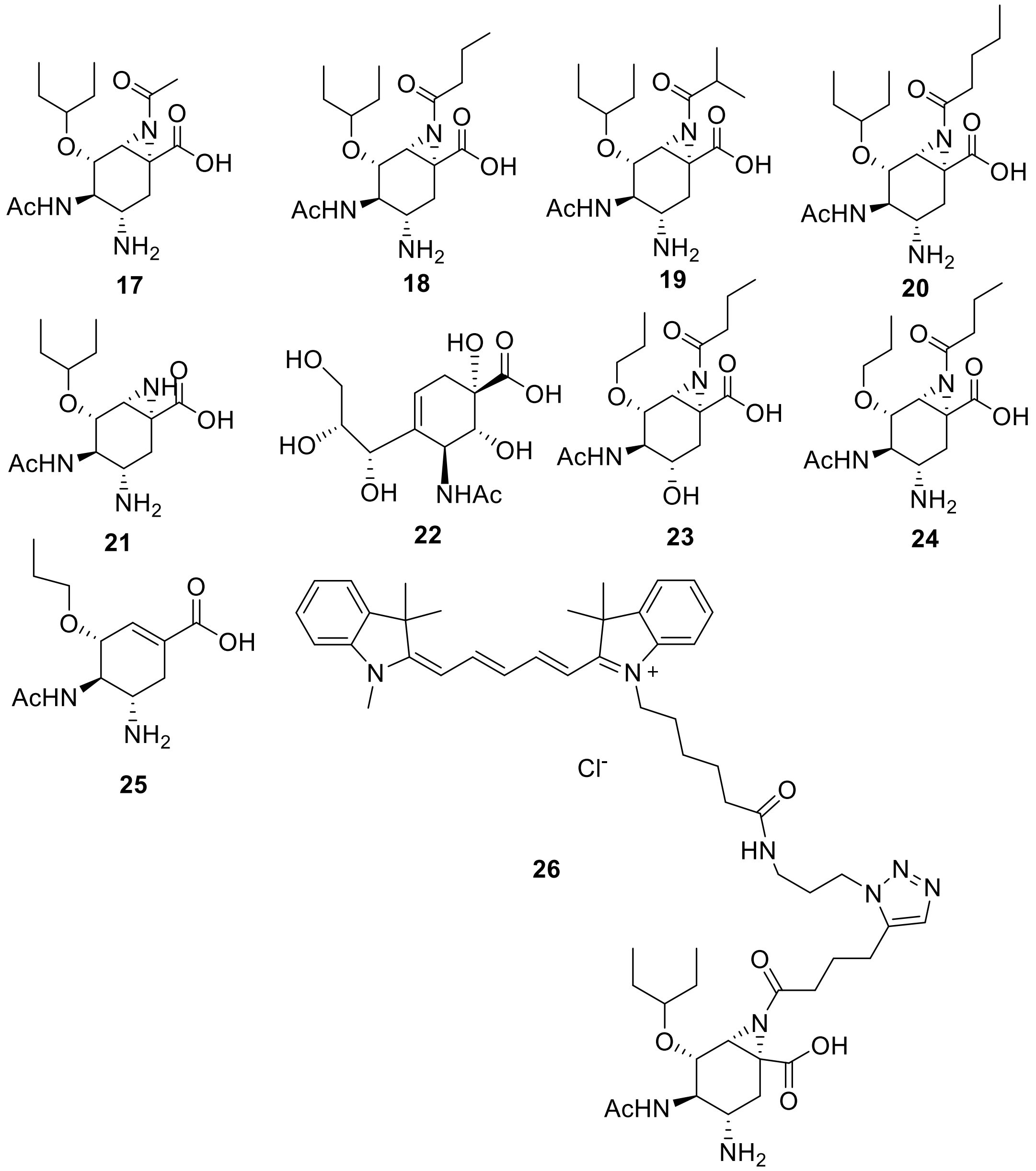


Figure 4.13- Viral neuraminidase inhibitors and activity-based probes.

*Compounds **17-25,** except for **22,** are based on the commercialised antiviral Oseltamivir. Activity-based probe **26** with cy5 group.*

4.3. Materials and Methods

4.3.1. Human NEU2

4.3.1.1. Expression of Human NEU2

pET_28 NEU2 plasmid was transformed into BL21(DE3) cells using a heat shock procedure. The protein was recombinantly expressed through the inoculation of 3 L of LBE-5052 autoinduction media supplemented with (2M $MgSO_4$, Trace metals (50 mM $FeCl_3$, 20 mM $CaCl_2$, 10 mM $MnCl_2$, 10 mM $ZnSO_4$, 2 mM $CoCl_2$, 2 mM $CuCl_2$, 2mM $NiCl_2$), 1M K_2HPO_4 , 30 µg/mL^{-1} kanamycin and 1 M KH_2PO_4). The expression was then incubated overnight at 37 °C, 250 rpm for 15 hours. NEU2 was not secreted so the pellet was harvested by centrifugation at 5000 rpm, 4°C for 40 minutes and resuspended using 12 mL elution buffer (20 mM HEPES, 100 mM NaCl and 500 mM imidazole), 196 mL binding buffer (20mM HEPES, 100 mM NaCl and 20 mM imidazole), 6 µL benzonase and lysozyme. The resuspended pellet was then sonicated for 5 minutes (1 second on and 2 seconds off) before centrifuging at 18,000 xg for 40 minutes at 4°C and collecting the supernatant.

4.3.1.2. Purification of NEU2

The supernatant was loaded onto a pre-equilibrated 5 mL HisTrap Excel column (Cytiva) and eluted using a buffer gradient 0-100% of elution buffer over 20 CVs at a flow rate of 5 mL/minute, eluting 1.8 mL fractions. Protein-containing fractions were confirmed with SDS-PAGE electrophoresis and concentrated using a 10 kDa concentrator to 8.1 mg and performing a TEV cleavage (1:1/50 TEV: protein) overnight to remove the hexahistidine tag. The His-tag affinity column procedure was repeated to remove any uncleaved NEU2. The flow-through from the column was collected and loaded onto an equilibrated MONO Q column with Buffer A (20 mM Tris/HCl pH 7.3) and eluted using buffer B (50 mM MES pH 6.7) at 1 mL/minute, collecting the flow-through and concentrating to 1 mL using 10 kDa Vivaspin concentrator. Concentrated protein was loaded onto an equilibrated HiTrap SPHP column with Buffer B and eluted with buffer C (50 mM MES 1M NaCl pH 6.7) with a buffer gradient of 0-100% at 1 ml/minute over 20 CV eluting 1.8 mL fractions. Fractions containing protein pooled and concentrated using Vivaspin 10 kDa concentrator to 100 µl at 9.4 mg/mL and then flash frozen using liquid nitrogen, stored at -70 °C.

4.3.1.3. X-ray crystallography for NEU2

48 – well MAXI crystallisation plates were set up using 9.4 mg/mL Neu2 with 100 mM sodium/potassium phosphate buffer pH 6.4 and 2.9 M NaCl with a 1:1 protein to precipitants. Data was collected at the Diamond light source to 1.5 Å, and data reduction was performed using AIMLESS, MOLREP for the molecular replacement and REFMAC5 for the model refinement and model building in ccp4i2 [75, 188].

4.3.2. Viral Neuraminidase N9

4.3.2.1. Generation of N9 Baculovirus

The DH10Bac™ competent cells are used to produce the recombinant bacmid as they contain the tetracycline-resistant helper plasmid (bMON7142) and kanamycin-resistant baculovirus shuttle vector (bMON14272). The N9 A/Anhui/1/201 (received from Dr Steve Gamblin's group, The Francis Crick Institute) was transformed into DH10Bac™ competent cells using electroporation at 1.8 kV briefly before adding SOC medium and incubating the cells for 4 hours at 37 °C. Blue-white screening was carried out on IPTG (1 mM), Kanamycin (50 µg/ml), tetracycline (15 µg/mL) and gentamycin (15 µg/mL) containing LB agar plates and incubated for 2 days at 37 °C and the resulting white colonies restreaked.

4.3.2.2. Bacmid Verification

Selected restreaked colonies were incubated overnight in 10 mL of LB media at 37 °C, and the plasmid DNA was extracted using the PureLink™ genomic DNA mini kit (Thermo Scientific). Verification of the bacmid transformation was carried out using colony PCR. PCR reactions (Table 3.4) were prepared for each restreaked colony to 20 µL containing DNA (1 µL), p55 (5’-CCCAGTCACGACGTTGTAAAACG -3’) (0.5 µM), p56 (5’-AGCGGATAACAATTTCACACAGG-3’) (0.5 µM) dNTPs (0.2 mM), Phusion enzyme (2 U) and Phusion HF buffer (1x) (Thermo scientific F530S). 1% agarose gel with SYBR-safe DNA gel stain (0.1x) (Thermo Scientific) was used to verify the formation of the bacmid.

Table 4.3- PCR reaction mix

Reaction Mix			**Cycle parameters**	
Component	Volume (µL)	10x stock	Step	Temperature
HF Buffer	4	40	ID	95
H20	11.8	118	D	95
DMSO	0.6	6	A	68
dNTPs	0.4	4	E	72
Primer_F (p55)	1	10	FE	72
Primer_R (p56)	1	10	H	4
Template	1		Total time	2hr 30
Phusion	0.2			

4.3.2.3. Generation of Recombinant N9 Viral Neuraminidase Baculovirus in Sf9

A 60 mL Sf9 (derived from *Spodopetre frugiperda,* IPLB-Sf-21-AE) suspension culture was grown for 2 days at 28 °C in Insect-XPRESS protein-free insect cell medium (Lonza Bioscience) before splitting back to 0.45 x10^6 cells/mL. An adherent cell culture was produced by transferring 2 mL of the suspended cells into each well of the 6-well culture plate and left to incubate in a humidifier chamber at 28 °C until the cells became adherent. A mixture of 1.05 mL Insect-XPRESS media, 50 µL N9 DNA (~100 ng/µL) and 31.5 µL of FuGENE HD transfection reagent (Promega) was created for transfection, and 180 µL added dropwise to each well before placing in the humidified chamber for 3 days. Once the *Sf*9 cells are 95% fluorescent and 90 % viable, the culture is collected and centrifuged at 200 xg for 5 minutes to produce a clarified supernatant and adding 2% fetal bovine serum: this is referred to as V1. V1 was amplified a 50 mL *Sf*9 suspension culture (1 x10^6 cells/mL) with 1 mL of V1 added and incubated at 28 °C and 87 rpm until 95% fluorescence was achieved. The culture was again harvested for centrifugation, and 2% fetal bovine serum was added.

4.3.2.4. *N9 Gene Expression and Purification*

The suspended culture of High Five TM (*Trichoplusia ni*, BTI-TN-5BI-4) cells was expanded using Express Five TM SFM supplemented with 20 mM media until the cells reached 3.6 L at 1.2 x10^6 cells/mL using 6 x 600 mL Corning Erlenmeyer shaker flasks. 1.5 mL of the previously prepared V2 was added to each 600 mL and left to incubate at 28 °C and 87 rpm until 95% fluorescence was achieved. Cells were then pelleted at 200 xg, 20 minutes at 4 °C and the clarified media was then further centrifuged for 20 minutes, 5000 xg at 4 °C and then supplemented with cOmplete TM EDTA-free protease inhibitor (Sigma).

Clarified media was pH adjusted using 150 mL of 50 mM Tris pH 8, 0.5M NaCl, and 20 mM imidazole to a final pH 7.12. Conditioned media loaded onto a 5 mL HisTrap excel column (GE Healthcare) equilibrated with binding buffer (50 mM Tris pH8, 0.5M NaCl, 20 mM imidazole) and eluted using a buffer gradient 0-100% of elution buffer (50mM Tris pH 8, 0.5 M NaCl, 500 mM imidazole) over 20 CVs at a flow rate of 5 mL/minute eluting 1.8 mL fractions. Fractions containing protein were pooled and concentrated using a 30 kDa Vivaspin TM concentrator to 1 mL. Concentrated protein was loaded onto a pre-equilibrated (25 mM Tris pH 7.5, 150 mM NaCl, 5% glycerol, 2 mM $CaCl_2$) superdex 200 increase 10/300 column (Cytiva).

4.3.3. *Viral Neuraminidase N1*

4.3.3.1. *Construct Design of Viral Neuraminidase Constructs*

A construct was designed using the sequence of N1 from influenza A/California/04/2009. The N-terminus contained a melittin signal sequence followed by a hexahistidine tag, tetramerisation domain and thrombin cleavage site into a pFastBacTM vector, which was produced, and codon optimised by GenScript (Appendix 7.1.3 and 7.1.4).

4.3.3.2. *Expression of Recombinant N1 using a Baculovirus Expression System*

The procedure of transformation into DH10BacTM cells and generation of the recombinant baculovirus was outlined in Section 4.3.2 Again, the expression was conducted in Hi5 cells, which were expanded to 3.6 L at 2.0 x10^6 cells/mL and infected with 2 mL of the generated P2 N1 baculovirus stock. The cells were cultured at 28 °C and 87 rpm until 95% fluorescence was achieved. Cells were then pelleted at 200 xg, 20 minutes at 4°C, and the clarified media was then further centrifuged for 20 minutes, 5000 xg at 4 °C and then supplemented with a cOmplete ^{TM} EDTA-free protease inhibitor (Sigma).

4.3.4. Purification of N1

The same purification procedure for N9 was conducted (4.3.2.4) but additionally, further purification procedures were investigated. Concentrated protein loaded onto an equilibrated 5 mL Mono Q™ (Cytiva) with 25 mM Tris pH 8, 5% glycerol and 50 mM NaCl, flow rate 1 ml/minute. Protein-containing fractions were eluted with 25 mM Tris pH 8, 5% glycerol and 1 M NaCl. Protein elutions were pooled and concentrated using an Amicron 30K concentrator to 1 mL and a 10 kDa concentrator to 2.7 mg/mL for N1 Closed conformation and 10 mg/ml for N1 open conformation.

4.3.5. Activity-Based Probe Labelling of Viral Neuraminidase N1

To test for N1 protein expression, the supernatant from a three-day baculovirus expression experiment was incubated for 1 hour at 37 °C with ABP **26**. After incubation, samples were denatured with 5x Laemmli buffer (50 % (v/v) 1.0 M Tris−HCl, pH 6.8, 50% (v/v) 100 % glycerol, 10 % (w/v) DTT, 10 % (w/v) SDS, 0.01 % (w/v) bromophenol blue) by boiling for 5 min at 98 °C. Denatured samples were loaded onto a 10% SDS-PAGE gel at 200V for 30 minutes. Wet slab gels were scanned for ABP-emitted fluorescence using a Cy5 (λEX = 625 nm).

To probe for the optimum reaction conditions (pH and [calcium]), purified N1 was incubated with **27** for 1 hour at 37 °C using various pHs (5-8) and calcium concentrations (0-15 µM). After incubation, samples were denatured with 5x Laemmli buffer by boiling for 5 min at 98 °C. Denatured samples were loaded onto a 10% SDS-PAGE gel at 200V for 30 minutes. Wet slab gels were scanned for ABP-emitted fluorescence using a Cy5 (λEX = 625 nm).

4.3.5.1. Competitive Inhibition ABPP for Viral Neuraminidase N1

Purified N1 (1 µM) was incubated with one of the inhibitors **17-25** (10 µM) overnight at 37 °C in 25 mM Tris pH 7 buffer. After the ABP (**26**, 10 µM) was incubated with each sample for 1 hour at 37 °C. After incubation, samples were denatured with 5x Laemmli buffer by boiling for

5 min at 98 °C. Denatured samples were loaded onto a 10% SDS-PAGE gel at 200V for 30 minutes. Wet slab gels were scanned for ABP-emitted fluorescence using a Cy5 (λEX = 625 nm).

4.3.6. *Thermal Shift Analysis of N1 using NanoDSF*

Reaction mixtures of 30 µL were made up in 25 mM Tris pH 7.5 at 1 mg/mL for both N1 open and closed conformation. Reactions, including 2 mM Calcium and 5 mM EDTA, were incubated for 7 hours. Those including inhibitors were also incubated for 7 hours at 1:10. Thermal unfolding of the protein was monitored using Prometheus NT.48 instrument by tryptophan absorbance at λ_{ex} 260 nm, λ_{em} 330 and 350 nm, and 40% gain at a temperature range of 20 - 95 °C of 1 °C min^{-1}.

4.3.7. *Viral Plaque Inhibition Assay*

This assay was carried out using a WHO compliant methodology [189]. MDCK-SIAT1 cells were aliquoted into a 96-well plate (200µL/well) using a 1:20 dilution factor and left to reach confluency at 37°C, 5% CO_2 over three days. Influenza virus A H3N2 was previously incubated with a 2-fold serial dilution (100-0.8 nM) of the five chosen inhibitors (**17**, **18, 19, 20, 21**) for four hours before adding to MDCK-SIAT1 cells [189]. After 3 hours of incubation, a 1.2 % Avicel overlay containing 2 µg/mL trypsin was added, and cells were incubated overnight at 37°C. The overlay was aspirated from the wells, and the cells were fixed with ice-cold 4% paraformaldehyde in PBS (200 µL/well) for 45 minutes and removed with PBS. Cells were permeabilised using 0.2% Triton X-100 in PBS for 30 minutes before washing with PBS and leaving to dry before assessing plaque formation using immune-staining. A monoclonal antibody for influenza A nucleoprotein was added to the fixed cells for 1 hour, followed by a 1-hour incubation with a secondary horseradish-peroxidase antibody [189]. True Blue^{TM} was added as a peroxidase substrate, and the reaction was allowed to proceed until blue plaques were observed; the plates were then washed with distilled water to stop the reaction from proceeding [189].

4.4. Results and Discussion

4.4.1. *Human NEU2*

4.4.1.1. NEU2 Purification and Crystallisation.

Human NEU2 was expressed in *E.coli* BL21(DE3) cells and was transformed with NEU2 pET_28 vector and incubated overnight in LBE-5052 autoinduction media. Instead of induction media, autoinduction was used so that bacteria growth didn't have to be monitored. Protein expression instead relies on the diauxic growth of *E. coli*, where bacteria can metabolise lactose after consuming more favourable carbon sources such as glucose [190]. The consumption of lactose induces the expression of the T7 *lac* promoter, which has transcriptional control over the NEU2 gene. Following the expression of NEU2, the media was harvested, and the pellet was resuspended in a binding buffer for affinity chromatography using a HisTrap Excel column (Figure 4.14). A sizeable broad peak was produced, and the fractions were pooled and concentrated using a 10 kDa concentrator. A TEV cleavage was performed overnight to remove the histidine tag, affinity chromatography was repeated to remove his-tagged NEU2. Again, the flowthrough was concentrated and then loaded onto an anion exchange column with the flowthrough collected to remove the 50 kDa contaminant. The concentrated protein was applied to a cation exchange column, yielding 13.8 mg/mL.

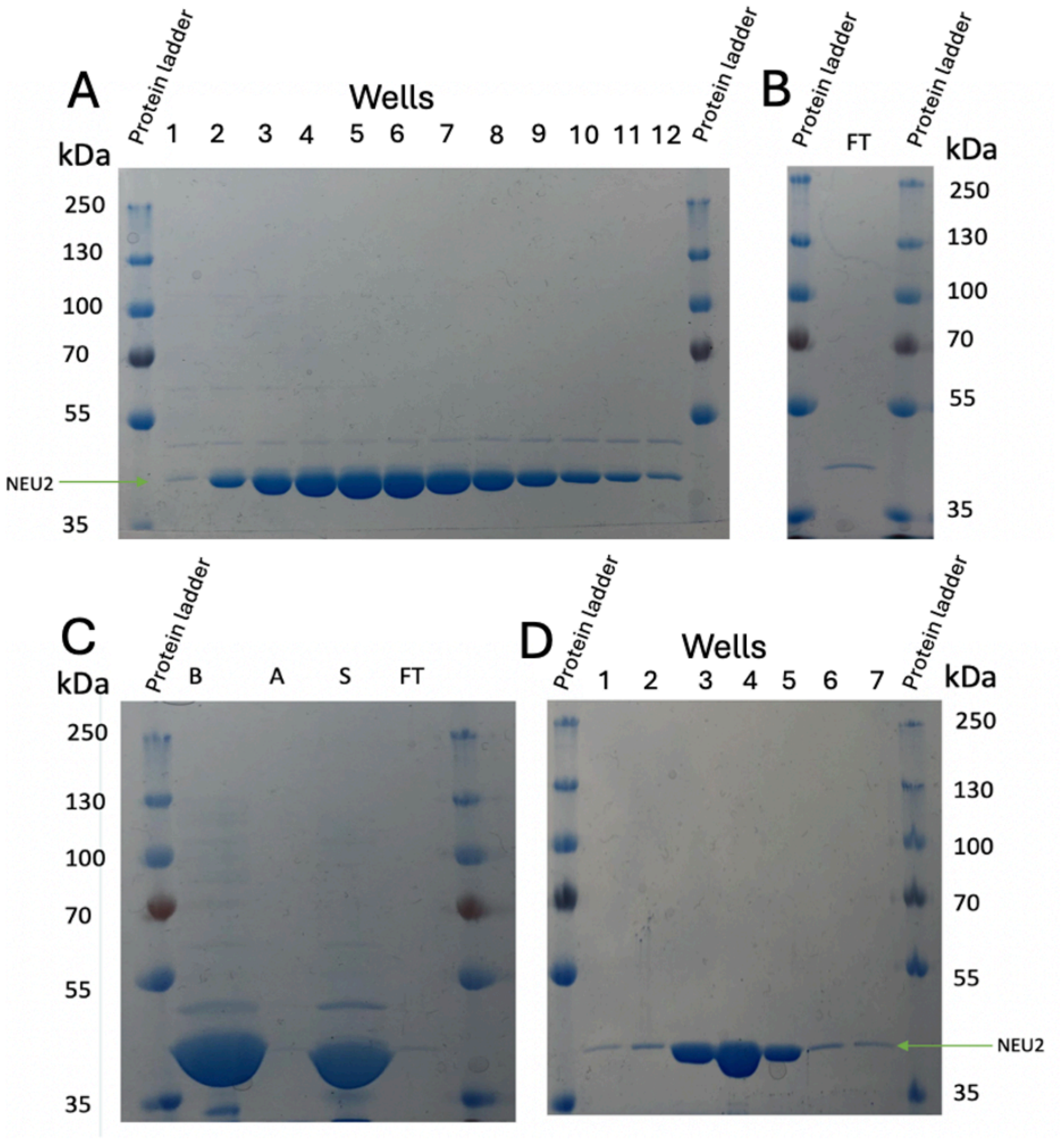


Figure 4.14- Purification of NEU2.

(A)The SDS-PAGE of NEU2 protein fractions contained significant A280 absorbance from the 5 mL HisTrap column. (B) SDS-PAGE gel from hexahistidine purification after TEV protease cleavage of the hexahistidine tag. The protein contained no hexahistidine tag and was eluted in the flowthrough (FT). (C) SDS-PAGE gel from MONO Q column, protein eluted in flowthrough (FT) indicated by the ~40 kDa, protein contaminants at 35 kDa and 50kDa removed by binding to the MONO Q column as indicated by the bands present in the before loading sample (B) and sample after loading (S). (D) SDS-PAGE gel from the HiTrap SP HP column produced only the NEU2 band at ~40 kDa, yielding 13.8 mg/mL of NEU2.

As our collaborators had yet to synthesise the compounds for the NEU2 project, no ligand complexes were obtained. Instead, the crystals were compared with the “apo”-structure of NEU2 (Figure 4.15) [148]. As seen in Figure 4.15; NEU2 comprises of 26 β-strands and 5 α-helices, similar to the reported viral and bacterial neuraminidases [148]. The activating core of the active site is found in the centre of the β-propeller, forming an acidic crevice with basic residues at the periphery to coordinate the sialic acid substrate [148]. In the “apo” structure, the loop which contains Glu111 has no electron density as the loop is disordered; once the inhibitors from our collaborators have been synthesised it is expected that this loop will become ordered as it does upon substrate binding [148].

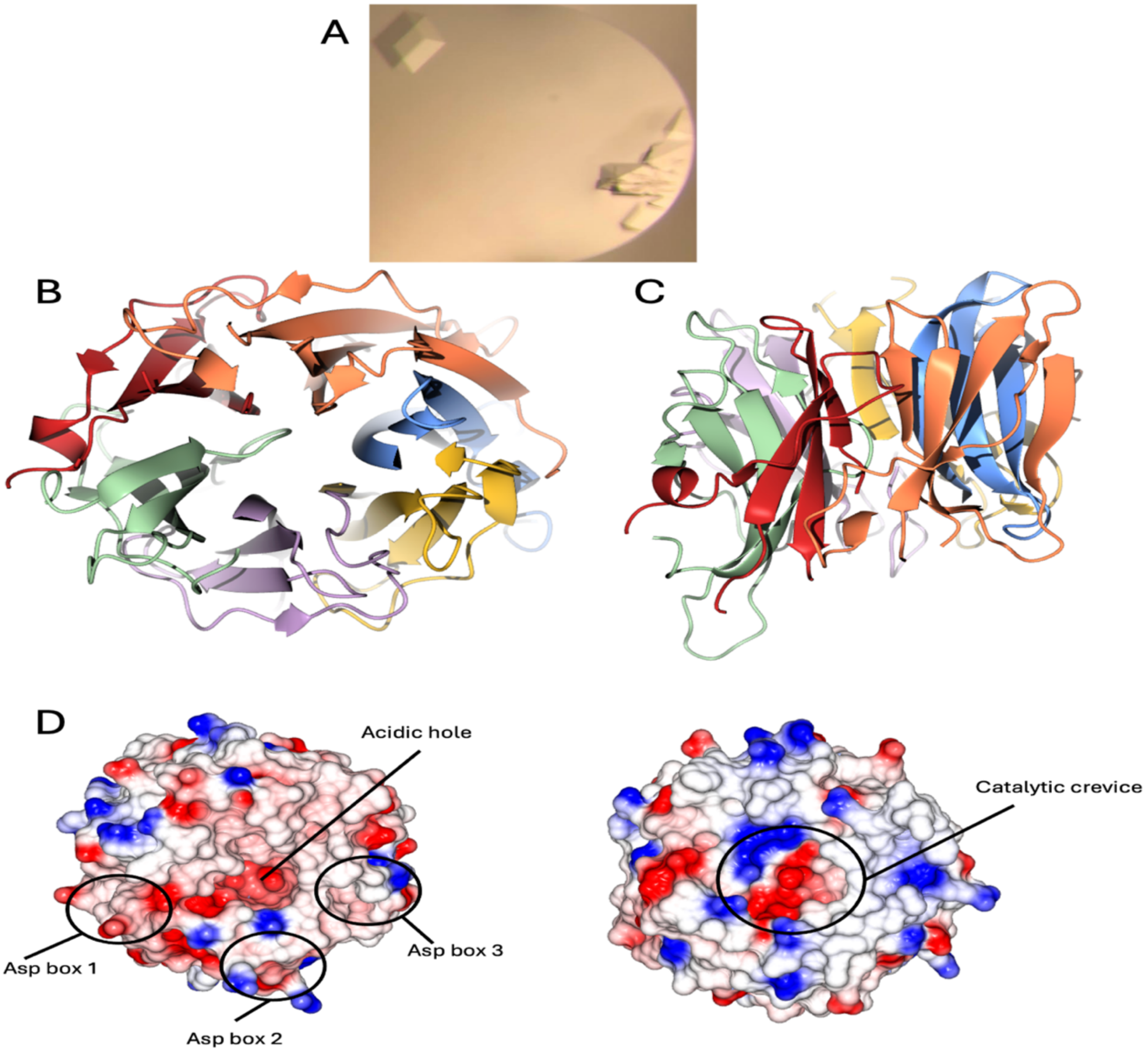


Figure 4.15- NEU2 crystallisation.

(A) NEU2 crystals. (B-C) Apo-structure of NEU2 (Top view) showing the canonical sialidase structure of six-bladed ß- propeller; Blade 1 (orange), Blade 2 (red), Blade 3 (green), Blade 4 (purple), Blade 5 (yellow), Blade 6 (blue). In the three Asp boxes in the 2nd, 3rd and 4th blades

of the six ß- propellers with an acidic crevice in the centre of the core for catalysis. (D) electrostatic surface of NEU2: positive potential blue and negative red.

4.4.2. Viral Neuraminidase

4.4.2.1. Gene Expression and Purification of N9

Initially, viral N9 (A/Anhui/1/2013) was expressed using the baculovirus expression system, but this was unsuccessful due to insufficient protein production. The construct was received from Dr. Steve Gamblin's group (The Francis Crick Institute) and contained an N-terminal hexahistidine tag, Vasodilator-stimulated phosphoprotein (VASP), and gp67 signal peptide. Numerous optimization tests were carried out, including testing V2 volumes and purification procedures. The low expression of N9 was problematic as its concentration was insufficient to bind to a 5 mL HisTrap column. Instead, a major contaminant at 100 kDa was observed (Figure 4.16A). Multiple expressions were conducted using different volumes of V2, but again, the same results were obtained. Alternatively, the harvested expression media was loaded onto a tangential flow to concentrate the media into a volume suitable for 1 mL HisTrap column. A band at 55 kDa representing N9 was shown after the 1 mL HisTrap column (Figure 4.16B). This did not yield enough protein, and after many unsuccessful attempts another construct was designed.

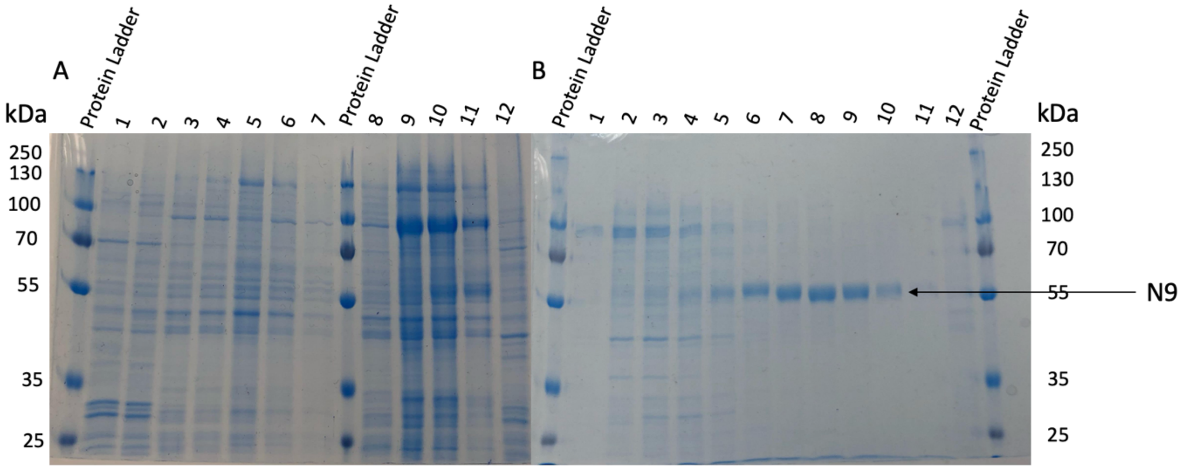


Figure 4.16- SDS-gel for N9 purification.

(A) 5 mL HisTrap purification was expected to see a prominent band at 55 kDa, which would indicate N9, but instead, a major contaminant was observed at 100 kDa. (B) Gel after 1 mL HisTrap column purification from concentrated expression by tangential flow, N9 indicated at 55 kDa.

4.4.2.2. Construct Design of Viral Neuraminidase N1

An N-terminal His-tagged N1 construct (Figure 4.17) was designed using GenScript with a pFastBac™ backbone containing gentamycin resistance and an SV40 polyadenylation signal [191]. A melittin signal sequence upstream of the N1 protein coding sequence directs the recombinantly produced N1 protein through the insect cell secretory pathway, resulting in secretion into the supernatant. The construct includes an hVASP tetramerisation domain and TEV cleavage site downstream of the hexahistidine tag. Two transposition sites (Tn7R and Tn7L) flank the polyhedrin promoter.

As mentioned, recombinant NAs can adopt an open or closed conformation, and N1 CA09-wildtype adopts an open conformation. Using the construct design from Ellis et al we expressed a closed conformation N1 [168]. This construct was designed using sequence alignments of non-bat NA subtypes and inserting homology-directed mutations to stabilise the tetrameric interface based on subtypes that adopt a closed conformation [168]. Mutations include Y100L, V117I, C161V, E165S, S172A, S196T, V205I, Q408M, and R419V *Figure 4.18*.

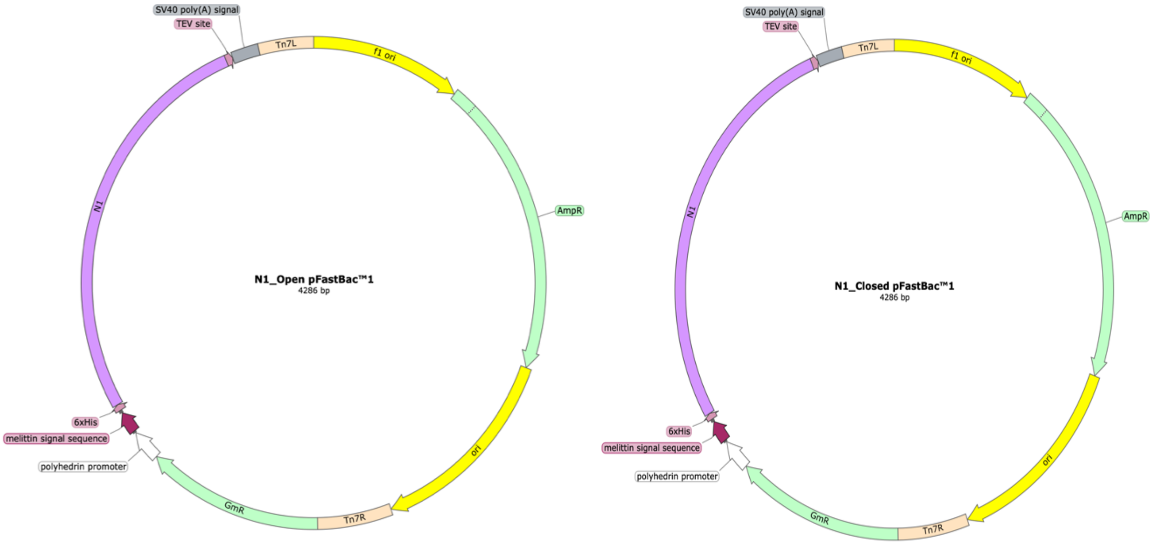


Figure 4.17- N1 pFastBac1 constructs.

Constructs of open and closed N1 inside a pFastBac ™ backbone with additional melittin signal sequence for the direction of the protein to the secretory pathway and a hexahistidine tag for purification.

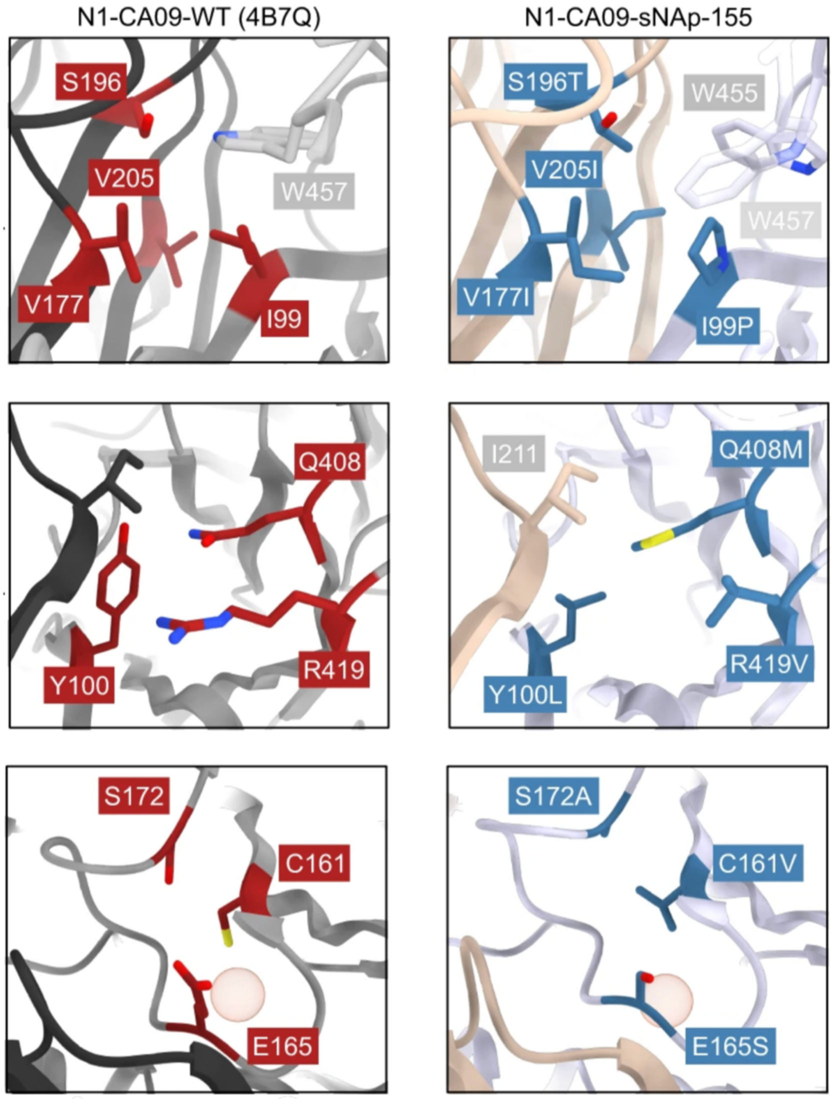


Figure 4.18- Mutations at tetrameric interface to forms a closed N1 structure. Image from [168].

Comparison of WT N1 CA09 with engineered N1 CA09 containing residues substitutions at the tetrameric interface to forma closed conformation.

4.4.2.3. Gene Expression and Protein Purification of N1

Initially, the N1 construct was transformed into DH10 EMBacY *E. coli* cells containing an EMBacY bacmid (bMON14272 (136 kb)) with a *lacZ* gene, mini-attTn7, yellow fluorescence protein (YFP), and kanamycin resistance. A helper plasmid (pMON7124 (13.2 kb)) encoding the transposase for the trans-Tn7 transposition function was also used to confer resistance to tetracycline [191]. Transposition occurs between the pFastBac™ vector into the bacmid (Figure 4.19). If this process is successful, the integration of the gene into the bacmid disrupts the *lacZ* gene, allowing the correct colonies to be determined using blue-white screening.

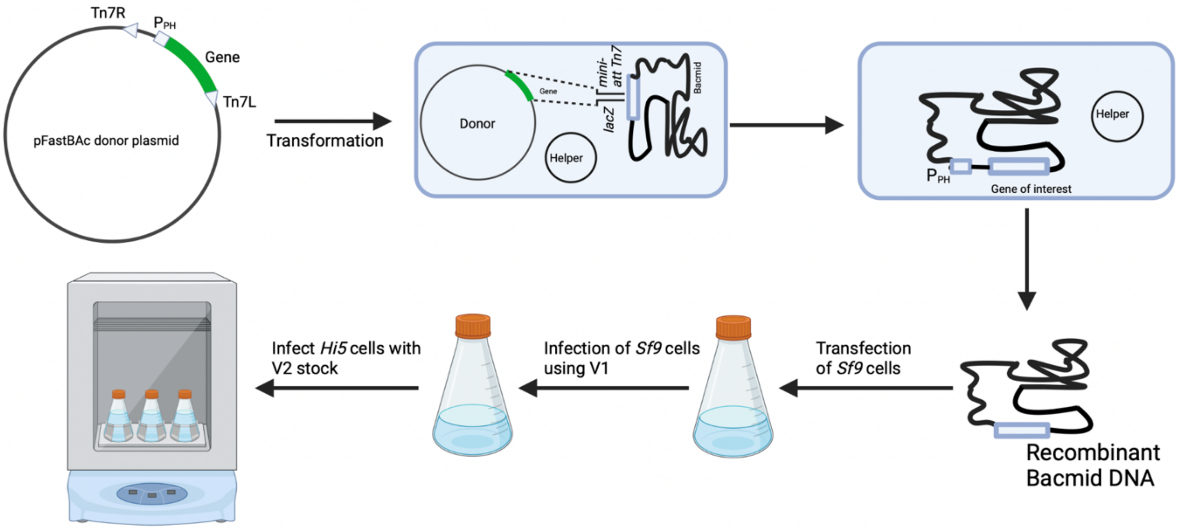


Figure 4.19- Process of baculovirus production (Image created in BioRender).

Firstly, the construct is transformed into DH10 EMBacY cells where transposition occurs with the assistance of the helper plasmid to form the recombinant bacmid DNA. Purified Recombinant bacmid is then transfected into Sf9 cells and undergoes two rounds of infection before yielding the resulting viral stock which is used for expression in Hi5 cells. Image made with Biorender.

The recombinant baculovirus was produced in *Sf*9 (*Spodopetre frugiperda)* cells, which have a 100-fold higher capacity for baculovirus production than other cell lines [192]. Once the *Sf*9 cells were cultured, they were transfected with the purified N1 bacmid using a lipid-based transfection reagent (Promega). As mentioned earlier, the EMBacY bacmid contains YFP, which monitors the cells producing the baculovirus. Once 95% total fluorescence was achieved, the recombinant virus was harvested and amplified to make the final V2 viral stock. Confirmation of protein expression was conducted using a small-scale expression of 30 mL, which was analysed by western blot.

The resulting stock was then used to infect 3.6 L of High Five (*Trichoplusia ni)* cell lines instead of continuing in *Sf*9's; this is due to the increased susceptibility of High Five cells to baculovirus infection [192]. The cells were harvested once 95% total fluorescence was achieved, and the clarified media was loaded onto a 5 mL HisTrap column. Per previous protocols [168], a size exclusion was conducted, although SDS-PAGE gel analysis suggested a large amount of contaminant (Figure 4.20B). To remove the contaminants still present after SEC, an anion exchange (Figure 4.20C) was performed, yielding a significant improvement in purity, making it suitable for Cryo-EM and crystallographic studies. For future purifications, the SEC has been removed. The open and closed conformations are expressed and purified by the same method but produce different concentrations 10 mg/mL (N1 open) and 2.6 mg/mL (N1 closed).

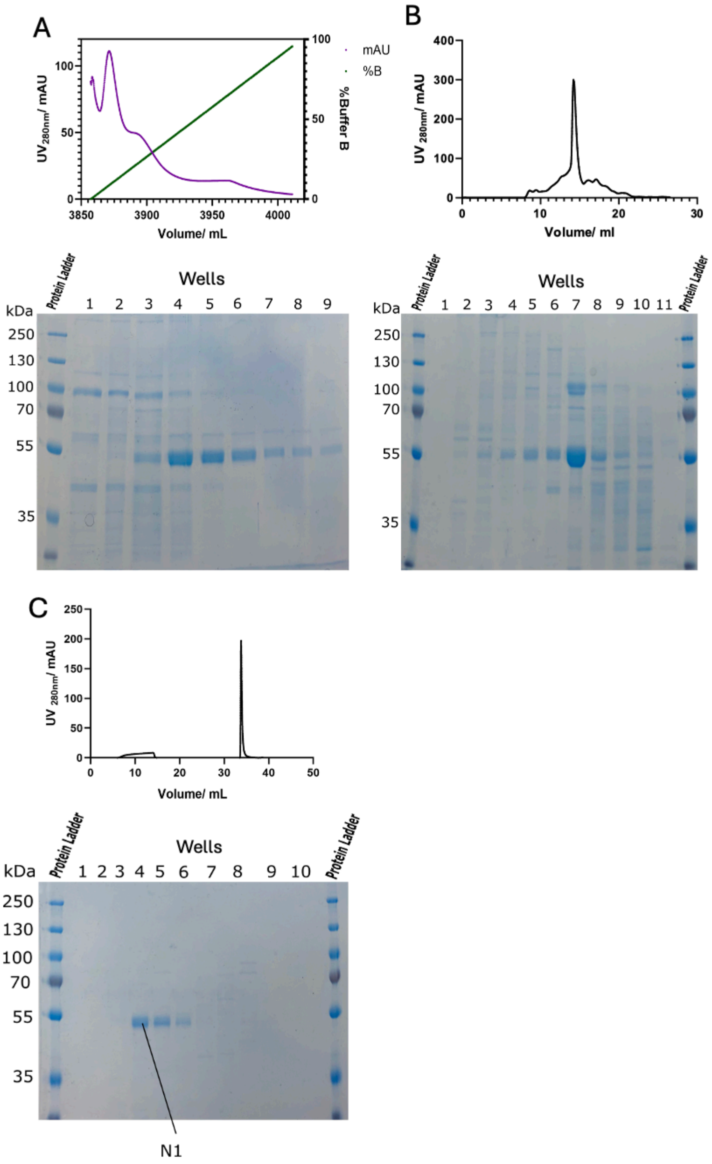


Figure 4.20- Purification gels of N1 closed.

(A) SDS-PAGE of N1 protein fractions which contained significant A280 absorbance from the 5 mL HisTrap column. (B) SDS-PAGE of N1 protein fractions, which contained significant A280 absorbance from the SEC. (C) SDS-PAGE of N1 protein fractions which contain significant A280 absorbance from anion exchange column. Arrow indicates what is thought to be N1 (~54 kDa).

(D) Chromatograph from AKTA anion exchange purification. N1 eluted early in the elution process, this is indicated by the large peak.

4.4.2.4. Activity based labelling of N1

The ABP **27** uses Oseltamivir as the specificity motif with an aziridine warhead conjugated to the reporter group Cy5 for fluorescence. ABP labelling was investigated to determine the optimum conditions for expression and inhibitor reactions. Each sample was analysed by SDS-PAGE, and the resulting gel was scanned for Cy5 fluorescence. To determine reaction conditions (Figure 4.21), supernatant expression samples were taken from days 1-3. Typically, western blots are used to analyse the recombinant expression of proteins, where an anti-his tag antibody is used to identify the tagged proteins. However, a western blot does not detect only active forms of the enzyme, unlike with ABPs (Figure 4.21). Figure B demonstrates that the ABP only binds to catalytically active forms of the enzyme, where at pH 2 there is no band as the protein has denatured. At optimum conditions pH 7, a fluorescent band is produced, signifying the catalytically active N1. At pH 9, some N1 activity persists as a faint band is present, suggesting suboptimal reaction conditions.

Calcium concentration and pH for reaction conditions were probed using **26** before cryo-EM grid preparation. The optimum conditions were selected based on the fluorescent readout of (Figure 4.21), and pH 7 was chosen as the optimum condition. As discussed in the above section, calcium contributes to catalytic activity. N1 is still active when calcium is removed as a faint band is produced, but a more prominent band is formed at 2 mM $CaCl_2$, signifying increased catalytic activity. The information from this set of experiments was used to determine the optimum conditions for cryo-EM grid preparation (buffer pH 7 and 2 mM $CaCl_2$).

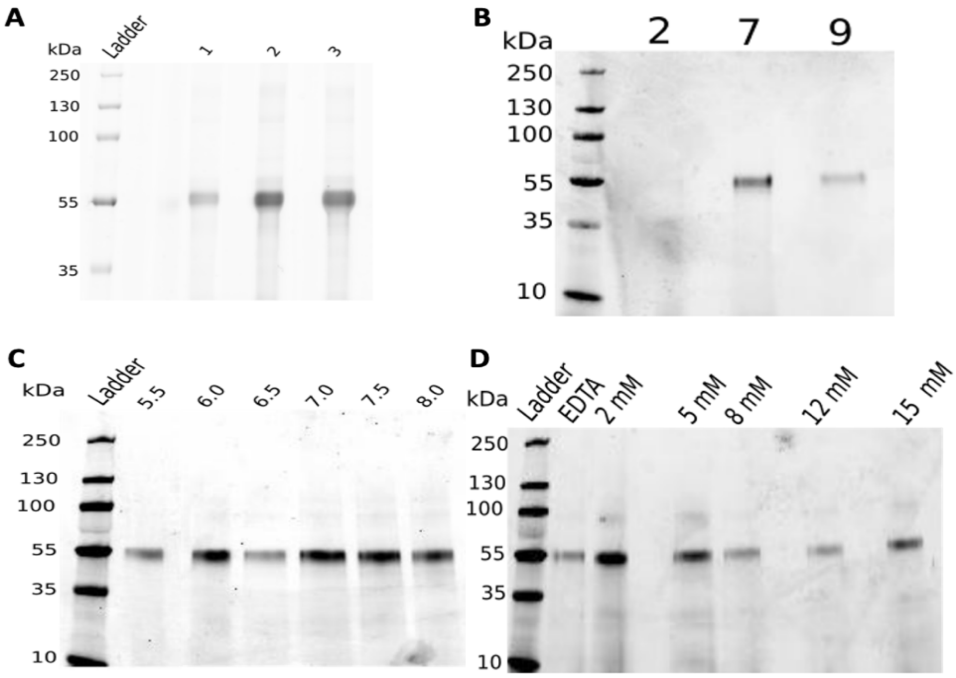


Figure 4.21- Activity-based probe gels for probing for the optimum reaction conditions for N1.

(A) various concentrations (1uM, 5uM and 10 uM) of ***26*** *reacted with the supernatant from an N1 baculovirus expression. (B) N1 reacted with* ***26*** *using a wide range of McIlvaine buffer pHs. No band is present at pH 2, a faint band at pH 9 and a prominent band indicating N1 at pH 7. (C) Narrow range of buffer pHs with N1 reacted with* ***26****; bands appear in all lanes, with pH 7 appearing the most prominent. (D) Calcium concentrations (0-15 mM) of N1 buffer reacted with* ***26****; a prominent band indicates the optimum condition at 2 mM.*

The reaction conditions determined through ABP may not represent the true optimum conditions for N1 but instead the optimum conditions for the probe reaction with N1. To test for this, a competitive ABP assay was run to investigate if the probe still labels after reaction with inhibitors **17-25**. Reaction mixtures of N1 (1 μM) were set up for each compound (10 μM) and left overnight at 37 °C before adding **26** (10 μM) and imaging the resulting gel for Cy5 fluorescence. Compared to the control, all of the inhibitors produced significantly fainter bands (*Figure 4.22*), suggesting that the inhibitors out-competed the probe. Again, this indicates that these are the optimum conditions for these inhibitors.

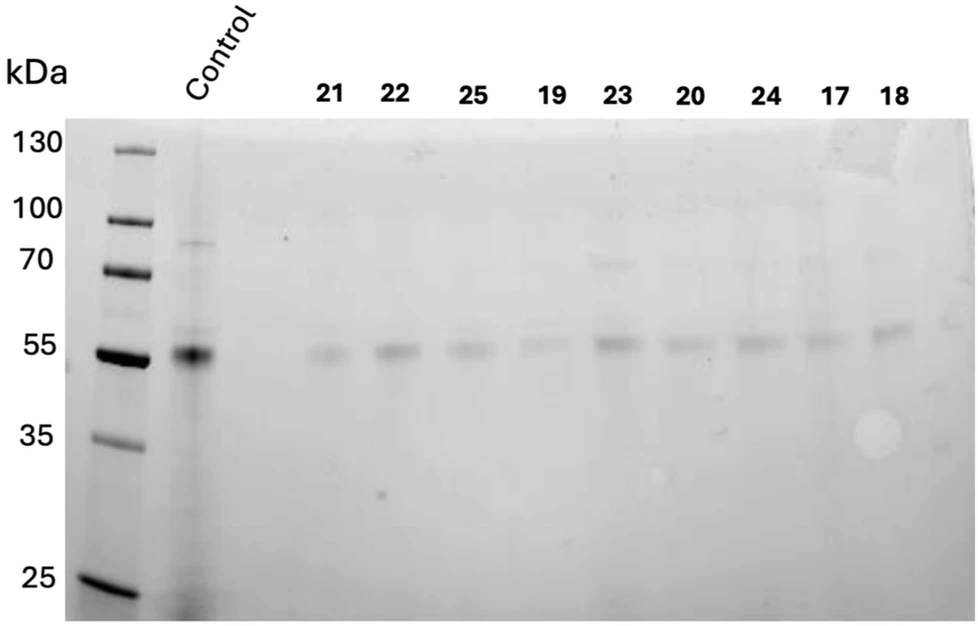


Figure 4.22- N1 competitive gel.

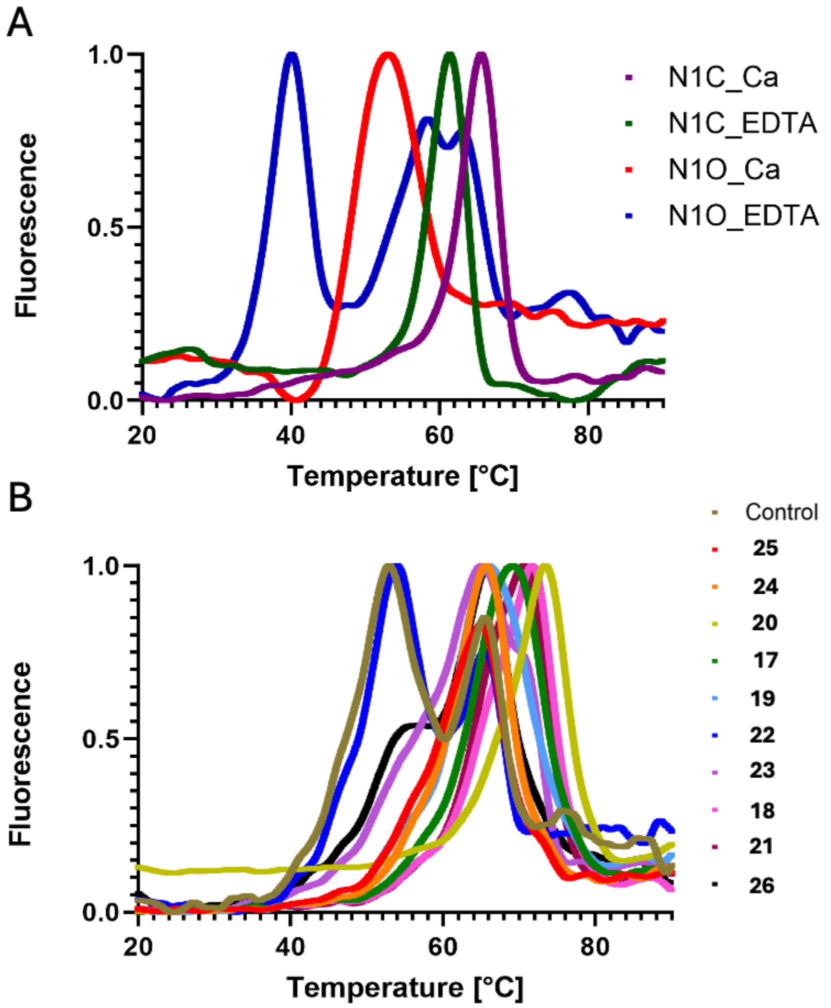
Purified N1 (1 µM) was incubated with one of the inhibitors (10 µM) at 37 ℃ before adding ABP ***26****. The reactions were analysed by SDS-PAGE electrophoresis and examined for Cy5 fluorescence of the ABP (λEX = 625 nm). Compared to the control, each inhibitor had outcompeted the ABP due to the faint bands on the gel.*

4.4.2.5. Thermal Shift Analysis of N1 with Inhibitors

Crystallography structures of different NA subtypes have identified conserved Ca^{2+} binding sites supporting the periphery of the active site [193]. Analysis of NA structures reveals five additional Ca^{2+} binding sites in each tetramer depending on the NA subtype [194]. Calcium binding has been shown to increase the initial reaction rate and decrease protein flexibility [168, 195, 196].

To examine the effect of Ca^{2+} on the different N1 conformations, the chelating agent Ethylenediaminetetraacetic acid (EDTA) was added to N1 to create a Ca^{2+}-free control. Their thermal stability was analysed using NanoDSF (*Figure 4.23*). For the open conformation, a 13.5 °C shift in thermal melting temperature (T_m) was observed with calcium. However, a second peak was observed in the control dataset, suggesting that there are two populations of N1 present in the sample, possibly due to the inadequate removal of calcium, the use of a better chelating agent such as ethylene glycol tetraacetic acid (EGTA) may resolve this. The closed conformation N1 showed a marginal increase of 4.3 °C in T_m. The increase in thermal stability observed in both constructs with calcium present suggests that the ion contributes to a conformational change in the protein, causing the burying of tryptophan residues.

Thermal shifts were also carried out for N1-inhibitor complexes (*Figure 4.23*). N1 (15 µM) was previously incubated with N-acyl aziridine inhibitors (75 µM) overnight at 37 °C. The sample was analysed using thermal shift, and increases in Tm were observed for all the compounds compared to the control, suggesting that they had bound to N1 and increased the protein's stability; however, compound **22** produced a negligible increase in T_m.



Sample	**T_m (°C)**
N1C_Calcium	65.4
N1C_EDTA	61.1
N1O_Calcium	53.5
N1O_EDTA	40
Control	52.95
17	68.95
18	71.25
19	66.2
20	73.15
21	70.05
22	53.90
23	65.05
38	65.45
25	64.50
26	65.5

Figure 4.23- Thermal shift analysis of N1 using intrinsic tryptophan fluorescence.

*(A) Thermal shift analysis of N1 open (N1O) and closed conformations (N1C) with and without calcium (N1C_EDTA and N1O_EDTA). (B) Effects of inhibitors **17-26** on the T_m of N1 from 1:10 (protein:ligand) reactions.*

4.4.2.6. *Viral Plaque Reduction Assay*

As a final series of experiments, the efficacy of compounds **17-25** was tested using a viral plaque reduction assay. This assay looks at the ability of the inhibitors to block the release of influenza virions using a WHO-compliant methodology [189]. To this end, H3N2 influenza A was incubated with a serial dilution of 100-0.8 nM of the five compounds before adding to cultured MDCK-SIAT1 cells. These cells are specialised for testing influenza susceptibility to inhibitors due to their overexpression of α 2,6 sialic acid linked to galactose moieties [189]. After 3 hours of infection, a 1.2% Avicel overlay was added to each well, and the formed viral plaques were assessed after 24 hours by immune-staining. As shown (Table 4.4), the non-acylated aziridine **21** was the most potent compared to the *N*-acyl aziridine derivatives. Compared to the inhibitor Oseltamivir, previous reports of a similar viral plaque assay state an EC_{50} of 420 nM against influenza A H3N2, suggesting that our compounds outperform the commercialised inhibitor [197].

Table 4.4- Viral plaque assay.

Compound	**21**	**17**	**19**	**20**	**18**
EC_{50} (nM)	3.1	50	6.3	3.1	25

4.5. Summary

In this work, recombinant viral neuraminidase N1 was expressed using a baculovirus expression system, and its N1-inhibitor complexes were analysed using thermal shifts and activity-based probe profiling. The complexes were also analysed by cryo-EM, which will be discussed in the subsequent **Chapter 5**.

The developed cyclophellitol N-acyl aziridines were shown to inhibit via a covalent mechanism, but the standout of the series was the non-covalent inhibitor **21**. This is the oseltamivir aziridine, a potential alternative to the competitive commercialised inhibitors. Viral plaque assays (Conducted by Dr Ruth Harvey's group, The Francis Crick Institute) revealed that **21** has a low EC_{50} (3.1 nM) superior to that reported for Oseltamivir [197]. Analysis of protein stability using thermal shift assays showed the inhibitors' effect on protein stability. However, thermal shift analysis cannot distinguish between covalent and non-covalent binding. Surprisingly, the competitive inhibitor **21** has the second highest T_m compared to the mechanism-based inhibitors. The competitive inhibition of **21** instead of covalent inhibition may be due to the effect of the electron-withdrawing groups present on the N-acyl inhibitors. These withdrawing groups decrease the activation energy of the reaction by polarising the filled σ-orbitals away from the nucleophile, causing the aziridine group to be more reactive [198]. Hence, the free aziridine **21** may have a higher activation energy due to the lack of electron-withdrawing groups, so covalent bond formation cannot be achieved [198].

Having both competitive and non-competitive inhibitors available means that the appropriate inhibitor for the experiment methodology can be chosen. For example, mechanism-based inhibitors permanently block catalytic activity and can help expand our understanding of NAs' mechanism and physiological role. The aziridine design also offers the opportunity for chemical derivatisation, which is not available for other inhibitors, such as the commercialised competitive inhibitors; this allows for additional designs for treating antiviral resistance.

In conclusion, I have shown the expression and purification of human and viral NAs in bacterial and insect culture systems, respectively. Their analysis, as discussed above, alongside their expression, will enable viral NAs structural analysis with related inhibitors. These experiments will be discussed in **Chapter 5**.

5. Chapter 5- Structural Analysis of Viral Neuraminidase with Mechanism-Based Inhibitors.

5.1. Abstract

Over the past decade, technological and software advancements in cryo-electron microscopy (cryo-EM) single particle analysis (SPA) have allowed for the determination of protein structures by cryo-EM to near-atomic resolution. Cryo-EM offers unique advantages over the leading structural biology technique, X-ray crystallography, as proteins can be imaged in their native state and don't have to be crystallised, so small amounts of sample can be used.

This chapter expands on the work discussed in **Chapter 4**, where I sought to analyse the recombinantly expressed N1 viral neuraminidase with the N-acyl aziridine derivatives of Oseltamivir. Analysis of the mechanism-based covalent N-acyl aziridine inhibitors by cryo-EM unexpectedly revealed a non-covalent species bound in N1's active sites for each of the investigated inhibitors. Further investigations by ABP profiling, mass spectrometry and Quantum mechanics/molecular mechanics (QM/MM) simulations suggested that the non-covalent species present in the cryo-EM structures are, in fact, an elimination product of subsequent covalent reactivity. In the past, the formation of an eliminated species in X-ray crystallography structures of viral neuraminidases in complex with similar inhibitors has been observed where the elimination product is formed due to X-ray radiation damage [183]. Together with the data presented in **Chapter 4**, the N-acyl aziridines are potent mechanism-based inhibitors, as demonstrated in solution work. Therefore, the non-covalent species observed in the cryo-EM structures were modelled as eliminated products formed due to electron damage during cryo-EM data acquisition.

5.2. Introduction

Determination of the 3D structures of macromolecules is essential for understanding their mechanism and function to further characterise their biological role and involvement in disease for drug discovery [199]. Early structural biology techniques include X-ray crystallography and nuclear magnetic resonance (NMR), which are used to determine the 3D structure of proteins at an atomic level. X-ray crystallography is the most broadly used method, with the first structure determined in 1958 for myoglobin [200]. Although still widely used, this method has drawbacks: large quantities of homogenous protein samples are required, and the protein must be crystallisable. The technique depends entirely on the protein's ability to crystallise, which is a considerable bottleneck for some proteins. However, over the past two decades, significant advancements in cryo electron-microscopy (cryo-EM) have allowed for the structural determination of these non-crystallisable protein structures.

Technological advancements in instrumentation and software for cryo-EM have pushed the achievable resolution over the past decade, often called the ‘resolution revolution’ [201, 202]. The technique involves obtaining electron microscopy images of particles embedded in a thin layer of vitrified buffer to capture the biomolecule’s various conformations for three-dimensional (3D) reconstruction. In contrast, X-ray crystallography locks proteins into their most stable orientation, whereas cryo-EM allows for the native state of the protein to be captured, leading to a deeper biological understanding of the protein’s structure and function [203].

5.2.1. Anatomy of a Transmission Electron Microscope

Cryo-EM uses electrons to produce images of macromolecules that can be averaged and processed to atomic-level detail. Electrons are scattered by proteins 10^6 times stronger than X-rays [204]. Some electrons pass through the sample, while others are scattered by atoms [205]. Scattered electrons can be elastically scattered where there is no energy loss or inelastically scattered where energy is transferred into the specimen [206]. The transfer of energy into the sample by inelastic scattering causes radiation damage where atoms can be ionised and specimen structures altered through chemical bond rearrangement and the generation of free radicals [205]. Electrons can also be scattered by air. Therefore, the microscope is operated under a vacuum, which is essential but problematic for biological samples that prefer aqueous environments [205]. To overcome this, the biological samples are vitrified in a cryogen such as liquid ethane to protect them from the vacuum and radiation damage, as well as minimising the particle's motion and preserving their high-resolution features [207, 208].

In short, the transmission electron microscope (TEM) cryo-EM apparatus consists of an electron source, a series of lenses, a sample and an image-detecting system (Figure 5.1) [205]. In high-performance cryo-EM, a field emission gun is used as the electron source, where thermally emitted electrons are extracted by a strong potential gradient at the emitter surface and then accelerated (100-300 kV) [205, 209]. A series of condenser lenses and microcondensor focuses the diverging electrons, creating a parallel beam of electrons illuminating the sample [205]. Behind the sample is an objective aperture where broadly scattered electrons are prevented from reaching the image plane [205]. The electrons then pass through an objective lens, which provides magnification (20-50x); the intermediate and projective lenses further magnify the image before detection [205].

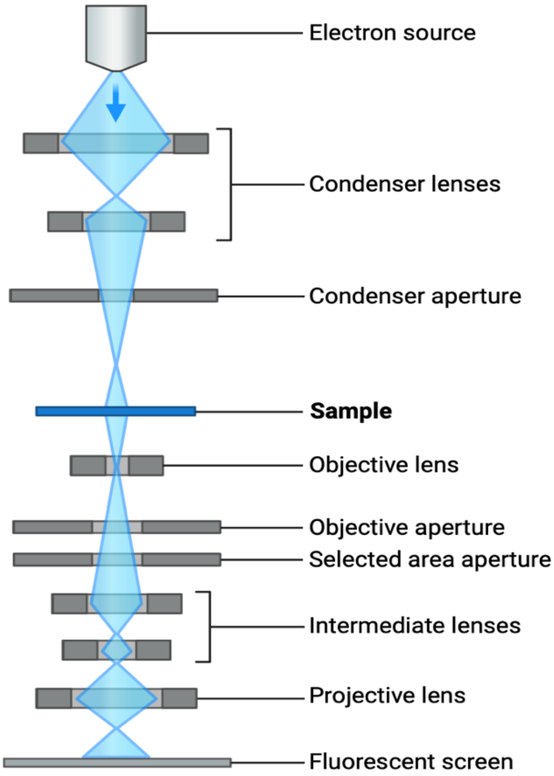


Figure 5.1- Anatomy of a typical Cryo-EM microscope.

Image based on [205], made in BioRender.

Early cryo-EM image generation used photographic film, a tedious technique for obtaining high-resolution structures [203]. Later, charge-coupled devices (CCD) were used, which converted electrons into photons for image generation, but this limited resolution; currently, direct-electron detectors (DED) are used for the sensitive detection of electrons directly [203]. DEDs generate images in the form of movies by using longer exposures where the total electron dose is fractionated across the movie images [210]. These movie image frames can be averaged and aligned to correct beam-induced motion and optimise the signal-to-noise ratio (SNR) for radiation-damaged images [210]. Phase plates and energy filters can also be implemented to improve SNR [203]. Energy filters prevent the inelastically scattered electrons from contributing to the final image, while phase plates introduce a phase shift for stronger scattering compared to the unscattered beam, providing image contrast [211].

5.2.2. Grid Preparation for Cryo-EM

For single-particle cryo-EM (SPA), thousands of particles with different orientations are averaged, classified and aligned to produce a 3D reconstruction [212]. Therefore, the ideal SPA sample is homogenous with particles embedded in a thin layer of vitrified buffer without preferred orientation, aggregation or denaturation [202, 211]. To achieve this, sample purification and grid preparation have to be optimised.

Appropriate protein purification methods are required to reduce sample heterogeneity. Some proteins are more challenging than others; membrane proteins are tricky to purify and image using SPA compared to their soluble cytosolic counterparts [202]. To combat this, detergents are typically added to mimic the lipid bilayer and stabilise the protein [202]. Not only does this aid in protein purification, but it also reduces the particles' interactions with the air-water interface (AWI) during gird preparation [202].

Grid preparation aims to preserve the protein particles in their native state and protect them from radiation damage and dehydration from the vacuum under which the microscope operates [212]. The typical grid preparation procedure includes applying 3-5 µL of sample onto the EM grid before blotting and removing excess liquid [211]. Samples are then rapidly frozen in a cryogen; this technique was created by Dubochet et al. but has since been developed into a semi-automated technique [207, 212, 213]. Commercially available grids are made of gold, copper or nickel and have a range of mesh sizes (200-400 grid bars) [202]. The choice of the grid can impact the sample's behaviour, where the gold mesh of the UltraAuFoil grids increases sample stability due to its increase in electric conduction and heat dissipation compared to copper grids [214]. A perforated foil support is overlaid onto the metal support, where the specimen is imaged [211]. Continuous carbon films can also combat problems such as preferred orientation [211].

Blotting is the most commonly used technique to create a thin film. The sample is applied onto a glow-discharged grid, and the droplet is then blotted with filter paper to wick any excess liquid [202]. This process occurs in a chamber with high humidity and low temperatures, as the thin sample film is especially susceptible to evaporation due to its high surface-to-volume ratio [202]. The blotting time and force determine the layer thickness; thinner ice is desirable as this increases electron transmission and decreases background noise [202]. During grid preparation, protein particles mainly interact with the AWI rather than the bulk of the solution [202]. The AWI is the boundary between sample buffer and the surrounding air during grid preparation, the association of particles with AWI effects sample quality as the AWI is a hostile environment. Interactions can result in preferred orientations and the unfolding and denaturation of particles; this can all occur in a millisecond timescale (Figure 5.2)[202]. To maintain the sample's integrity during grid preparation, the time between applying the sample and its vitrification is reduced [202]. Vitrification involves rapidly cooling the sample to obtain a glass-like state with preserved particle features for high-resolution imaging [202]. The cryogen determines the cooling rate. Liquid ethane is commonly used where the grid is plunge-freeze to obtain a vitrified layer of <1 um [202].

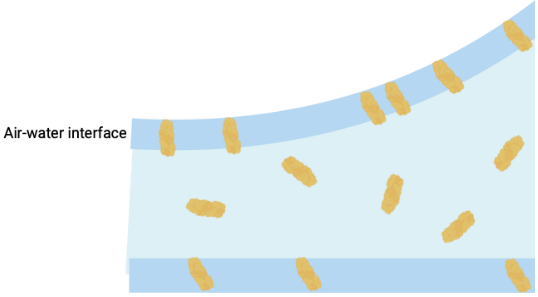


Figure 5.2- Protein particles interact with the air-water interface during grid preparation. The particles interact with the AWI (dark blue) within milliseconds of application to the cryo-EM grid, which can result in preferred orientation. Image made in BioRender.

5.2.3. TEM Imaging- Data Acquisition

SPA requires high-contrast images and high-resolution information to reconstruct a 3D model [210]. Currently, the record for the highest resolution structure from SPA is 1.25 Å [215]. However, achieving this resolution is difficult as obtaining good-contrast images from vitrified samples is challenging [210]. Therefore, the appropriate imaging conditions must be used. These settings include defocus values, selection of objective aperture, magnification and electron dose [210]. Magnification dictates the pixel size of the image. A high magnification produces an image with a fine pixel size [216]. Pixel size determines the attainable resolution; the Nyquist theorem describes the theoretically attainable resolution, where resolution is limited to twice the pixel size [210].

Images are typically formed by variations in specimen transmission due to absorption of the incident beam, creating amplitude contrast [205]. However, biological molecules produce poor amplitude contrast as they scatter electrons at varying angles with different path lengths, resulting in phase contrast [205]. Although proteins are weak phase objects, so to combat the weak scattering, the objective lens is defocused to introduce differences between the scattered and unscattered electrons, creating phase contrast [217]. The defocus values are defined before data collection. The aim is to achieve the right balance between contrast to identify particles while retaining high-resolution information [210].

The image generated contains a high level of noise because it primarily contains phase contrast instead of amplitude contrast; phase contrast is essential as it’s needed for extracting the maximum amount of information from the sample [218, 219]. TEM imaging is described by the contrast transfer function (CTF), a mathematical description of the image distortions caused by defocus and spherical aberration concerning spatial frequencies [220]. The amplitude and phase contrast are affected by CTF, causing variations in the image contrast at different spatial frequencies [220]. CTF is a sinusoidal oscillating wave where greater defocus values generate higher CTF oscillations, causing a loss of high-resolution information [217]. Due to the high spatial frequency oscillations of CTF, the information of the image is spread; this is described as the point spread function (PSF) [221]. Simply put, the PSF is equal to the Fourier transform of CTF [221].To obtain high-resolution details, the optimal defocus is determined. This is the minimum value at which particles are still visible, as cryo-EM imaging is a trade-off between contrast and resolution [217]. As CTF is an oscillating function, it can cross zero multiple times, causing loss of information; therefore, a range of defocus values is used to sample the whole reciprocal space [210].

5.2.4. Cryo-EM Image Processing and Reconstruction

To improve the SNR of the 2D projection images, many particles are needed for averaging [203]. Ideally, the particles will have various orientations/conformations of the protein, as this is necessary for 3D reconstruction [203]. Images are aligned and merged to generate an initial 3D model subjected to iterative rounds of refinement before the protein is fitted into the 3D map to build a 3D model (Figure 5.3) [203].

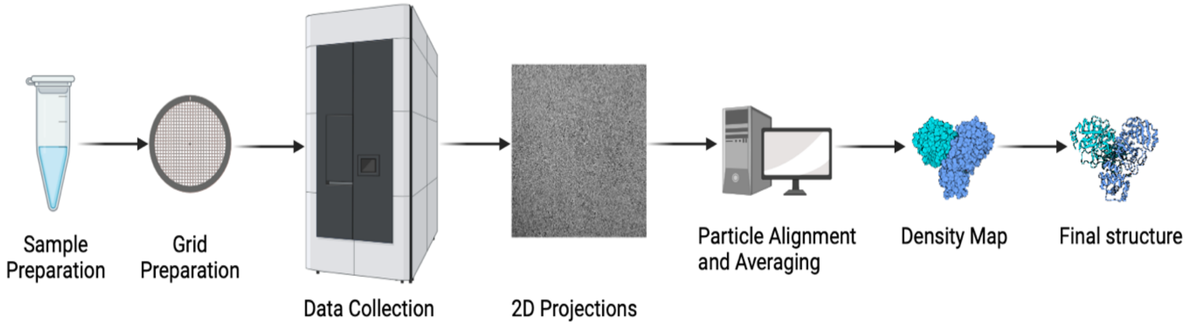


Figure 5.3- Simplified workflow for single-particle cryo-EM processing from sample preparation to 3D reconstruction. Image-based on [203] and made in BioRender.

5.2.4.1. Processing of Micrographs

During cryo-EM imaging, the total electron dose is fractionated into a series of image frames known as movies. Images become blurred due to beam-induced motion from exposure to electrons; this dramatically limits resolution [210]. Unfortunately, earlier movie frames which contain high-resolution information have the most beam-induced motion [210]. In contrast, later frames experience less motion but have a higher accumulated dose, causing a lack of high-resolution information [210]. The beam-induced motion causes a shift in particles between subsequent frames during exposure; this can be corrected in software such as RELION [210]. The movie frames are realigned during motion correction to create an unblurred micrograph. Corrected micrographs are then used for CTF estimation. Algorithms such as CTTFIND are widely used programmes for CTF estimation and are essential for providing an initial evaluation of micrograph quality [210, 222]. For CTF estimation, a few parameters are needed; the instrument used defines the wavelength (accelerated voltage) and spherical abbreviation (lens design) [210]. The defocus values for each image have to be known for CTF estimation, but the defocus values set by the user are approximate; accurate defocus values are obtained by fitting the calculated CTF pattern to the oscillations induced by CTF in the image power spectra (Thon rings) [210]. The CTF oscillation patterns are aligned with the Thon rings by adjusting the defocus values [210]. Once these parameters are known, the mathematical form of CTF is used to correct each image [210].

Picking quality particles is essential for downstream processing and obtaining high-resolution structures, as micrographs can contain frost, aggregates and artefacts [223]. Automated picking uses a reference-free Laplacian of Gaussian filter where particles with high SNR are picked [205, 212, 224]. Semi-automated methods use manually picked particles to train for automated picking, but the manual selection of particles can lead to bias for the more visible particles and certain orientations [210]. At the end of particle picking, the particles are assigned a set of coordinates and assembled into a stack.

The stack of images is then used for 2D classification, where raw particles are compared against each other and sorted into 2D classes with the same orientation/conformation [216, 218]. This generates classifications with reinforced structural features; at this stage in the pipeline, data can be assessed for quality, such as preferred orientation or structural heterogeneity [216]. Additionally, many particles will be discarded at this stage as artefacts or empty fields are removed. The 2D classifications can be repeated in an iterative process to obtain a stack of refined 2D classes [216].

5.2.4.2. Structure Refinement and Resolution

As described, 3D reconstruction of macromolecules in SPA is generated by combining 2D projection images of randomly orientated particles [225]. The Fourier slice theorem explains how these 2D projections can be combined [223, 225]. The Fourier transform of a 2D projection of a 3D object has the same amplitude and phases as the central slice of the 3D Fourier transform of the same object [223, 225]. Hence, collecting many 2D projection images of the 3D object from different directions would correlate to various cross-sections of the 3D Fourier transform of the object filling up the Fourier space (Figure 5.4A) [223, 225]. Once enough projections of multiple angles have been collected, the inverse Fourier transform can be used to calculate the object's 3D structure [223, 225]. However, the projection vectors of the randomly orientated particles are unknown. Typically, projection-matching is employed where an initial 3D model is back-projected, and the computed projections are compared with the experimental dataset relative to the 3D structure before using the projection slice theorem to estimate the unknown projection vectors (Figure 5.4B) [205, 223, 225]. This process relies on experimental coverage of all projection angles/directions through randomly orientated particles embedded in the vitrified buffer during grid preparation.

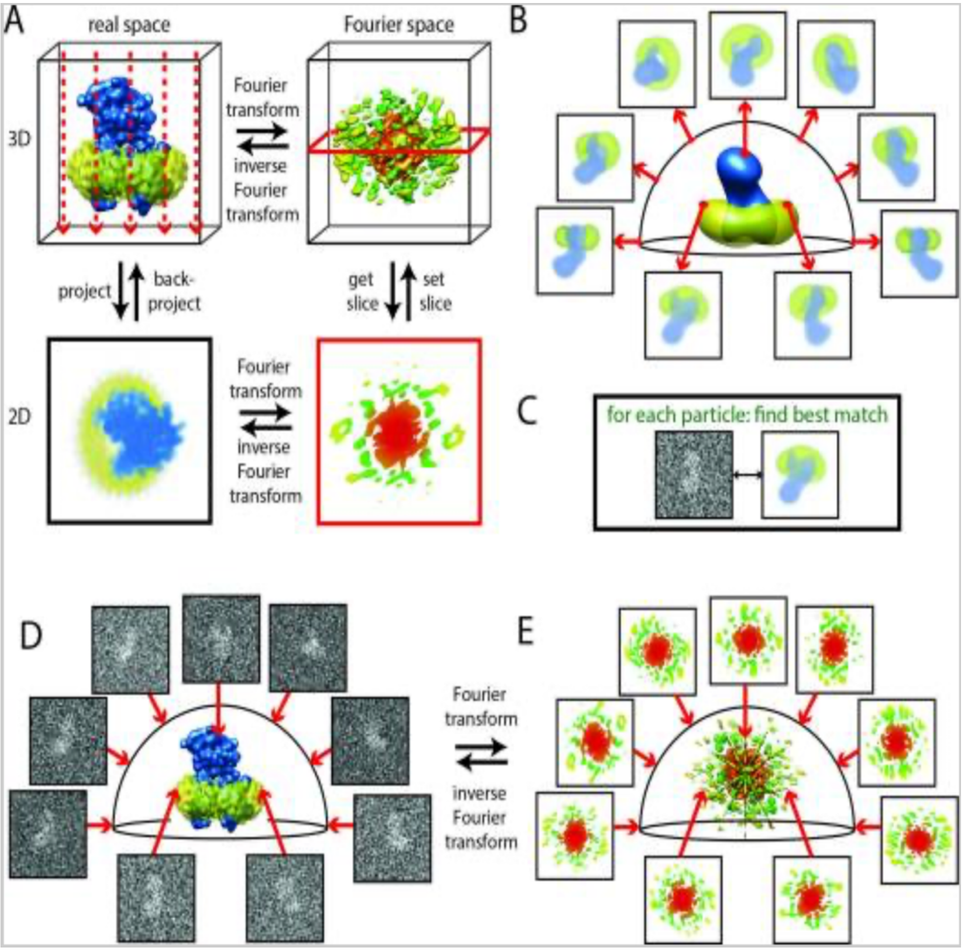


Figure 5.4- Principles of SPA cryo-EM processing.

(A) Schematic of the projection-slice-theorem where the central 2D projection of a 3D particle is the same as the central slice of the 3D Fourier transform [225]. (B-C) Projection matching of an initial 3D model is back-projected, and the computed projections are compared to the experimental dataset [225]. (D-E) The projection slice theorem suggests that the 3D structure can be determined by positioning the 2D slices into the 3d transform and calculating the inverse transform for 3D reconstruction [225]. Image from [225].

The iterative application of projection matching eventually yields a better resolution structure than the initial model [205, 223, 225]. After each round of refinement, the generated model is used as the reference for the next cycle to improve orientation assignments until improvements are no longer seen. The cycle of refinements includes 3D classification run without symmetry, and exhaustive angular searches are conducted to select particles with high-resolution features [226].

Several indicators can be monitored during the iterative rounds of refinement, including the Fourier shell correlation curve (FSC), which presents the Fourier shell correlation as a function of spatial frequency [210]. FSC curves are generated by refinement 3D, where the dataset is split into two halves [210]. The two halves are processed separately by calculating a volume for each half and determining their correlation coefficients within resolution shells extracted from the Fourier transforms [210].

Maps generated by 3D refinements should always be post-processed, where masks improve resolution and the FSC curve by excluding noisy solvent regions [205]. Bayesian polishing and CTF refinements for per-particle defocus, magnification, and higher-order aberration estimation can further enhance post-process resolution [227]. Firstly, CTF refinement estimates and corrects the beam tilt asymmetrical and symmetrical aberrations and re-estimates defocus values on a per-particle basis [227]. Bayesian polishing produces weighted averages of the aligned movie frames by tracking particle motion on a per-particle level [227]. These approaches are conducted with higher-resolution models from the 3D refinement; again, an iterative processing procedure (3D auto-refinement, post-process, CTF refinement and Bayesian polishing) would be beneficial until no further improvements in resolution are seen.

5.2.4.3. Model Building and Refinement

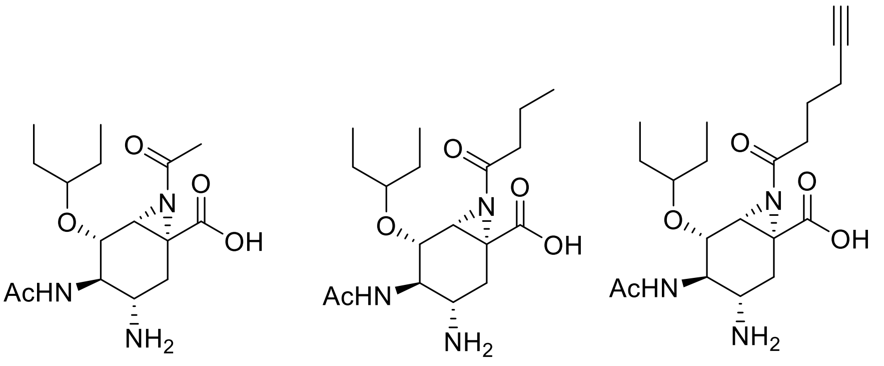
Overall, the quality of an EM map is determined by the resolution reached, where >10 Å is a low-resolution map as only the overall shape is confirmed, whereas <4 Å map (high-resolution) reveals residues and secondary structures [210, 228]. At 3 Å, most side chains can be modelled; hence, the higher the resolution, the more distinct features [228]. Typically, prior knowledge is required for cryo-EM model building, where an initial template is refined and fitted into the

EM map using software such as UCSF Chimera [133, 228]. The initial template can be taken from previously resolved PDB structures or model prediction methods like AlphaFold [228, 229]. Following template fitting, the model is manually refined to fit within the EM map using software packages such as Coot before a final automated real-space refinement using packages like Phenix [64, 75, 228].

In summary, forming images with cryo-EM involves preparing and vitrifying biological samples, followed by TEM data collection and image processing to produce a 3D reconstruction of the macromolecule. Cryo-EM's ability to provide high-resolution structures for more problematic proteins, such as membrane proteins, which are typically difficult to crystalise, makes it a powerful tool for visualising proteins in their native state for deeper biological understanding.

5.2.5. Aims of This Work

This chapter aims to structurally characterise the viral neuraminidase N1 (expression and purification procedures described in **Chapter 4**) using cryo-EM for compounds **17**, **18** and **20** (Figure 5.5) from **Chapter 4**. Results discussed in this chapter include compounds (**17** ,**18** and **20**) synthesised by the Overkleeft group (University of Leiden) and molecular dynamic simulations conducted by the Rovira group (University of Barcelona).



17 **18** **20**

Figure 5.5- N-Acyl aziridines based on the antiviral Oseltamivir for cryo-Em analysis with N1. Synthesised by the Overkleeft group (University of Leiden).

5.3. Materials and Methods

Expression and purification of Viral neuraminidase N1 influenza A/California/04/2009 see **Chapter 4**.

5.3.1. Cryo-EM Data Collection

For enzyme-inhibitor complexes with **17**, **18** and **20** reaction mixtures were prepared with 1.5 mg/mL of N1 to a final inhibitor concentration of 300 µM in 25 mM Tris, 2 mM $CaCl_2$ pH 7.0 buffer. The mixture was incubated overnight at 37 °C prior to grid preparation. For each reaction mixture, 3 µL was applied to glow-discharged UltraAuFoil (Quantifioil) 300 mesh R1.2/1.3 grids and subsequently blotted for 2 seconds with a blot force of -10 using the Mark IV Vitrobot (Thermo Fisher Scientific) at 100 % humidity and 4 °C. Data was collected remotely on a Titan Krios 1 (ThermoFisher Scientific) housed at eBIC Diamond Light Source UK. The microscope operated at 300 kV with a GATAN K3 detector. Each dataset was acquired at a dose rate of 18.3 e^-/pixel^2/s, pixel size 0.635 Å, and defocus values of -1.4, -1.2, -1.0, -0.8, -0.6 µm with a total dose of 43.99 e^-/Å^2.

5.3.2. Image Processing and 3D Reconstruction

Movie frames for each inhibitor dataset were motion-corrected, and CTF estimations were executed with CTFFIND-4.1 [222]. Using RELION 3 [230] , Laplacian of Gaussian automated picking detected particles, and the particles extracted. These were run into 2D classification and 2D classes with high-resolution features were selected to build an initial 3D model without symmetry constraints. 3D classification without symmetry restraint was run, and the class with defined structural features was selected for 3D refinement with C4 symmetry imposed. Most of the subsequent processing steps were carried out using RELION 5 [231], where CTF-refinement was performed using per-particle and per micrograph for astigmatism [230]. Followed by Bayesian polishing [232] based on an 8,000-particle training dataset. Iterative rounds of Auto-Refinement and CTF-Refinement for aberration corrections with per-particle and per micrograph for astigmatism yielded maps at 2.0 Å (**17**), 2.0 Å (**18**) and 2.25 Å (**20**).

5.3.3. Model Building and Refinement

UCSF Chimera [133] was used to fit the models [PDB:3TI3] into cryo-EM maps as well as changing the pixel size from 0.889 Å to 0.9 Å to better fit the model before conducting a final post-process in RELION. Point mutations were added manually, and the structure was refined in Coot [233]. Models were refined and validated with Phenix [234] using real-space refinement with secondary structure restraints (Table 5.1). Structural figures were rendered in Chimera [133].

*Table 5.1- Cryo-EM processing and refinement statistics for N1 complexed with **17, 18** and **20**.*

	N1-18	**N1-20**	**N1-17**
Data collection and processing			
Microscope	Krios 1	Krios 1	Krios 1
Voltage (kV)	300	300	300
Electron fluence (e^-/Å^2)	43.99	43.99	43.99
Nominal defocus values (µm)	-1.4, -1.2, -1.0, -0.8,-0.6	-1.4, -1.2, -1.0, -0.8,-0.6	-1.4, -1.2, -1.0, -0.8,-0.6
Calibrated pixel size (Å)	0.635	0.635	0.635
Symmetry imposed	C4	C4	C4
Initial particles	4763583	4261747	5464419
Final particle images	216383	229896	235807
Map resolution (Å)	2.00	2.25	2.00
FSC threshold	0.143	0.143	0.143
Refinement			
Map sharpening *B* factor (Å^2)	-49.69	-62.47	-51.69
Model composition			
Non-H atoms	11870	11864	11845
Residues	1476	1476	1476
Ligands	12	12	12
Waters	382	392	377
Mean *B* factors (Å^2)			
Protein	12.79	18.26	13.09
Ligands	26.09	36.48	26.31
Water	16.68	20.03	15.32
RMS deviations			
Bond length (Å)	0.003	0.004	0.005
Bond angles (°)	0.678	0.681	0.786
Validation			
MolProbaility score	1.23	1.36	1.25
Clashscore	2.66	2.31	1.31
Poor rotamers %	0.8	1.52	1.36
Ramachandran plot			
Outliers %	0.28	0.28	0.28
Allowed %	2.75	3.03	3.93
Favoured %	96.97	96.69	95.80

5.3.4. Activity Based-Probe Profiling

N1 (1 µM) reactions were incubated with inhibitor **18** (10 µM) along with a control N1 (1 µM) in 25 mM Tris pH 7 buffer at 37 °C. Multiple time points were taken between 3.5 and 7 hours before incubating the reactions with ABP **26** 10 µM for 30 minutes at 37 °C. After incubation, samples were denatured with 5x Laemmli buffer by boiling for 5 min at 98 °C. Denatured samples were loaded onto a 10% SDS-PAGE gel at 200V for 30 minutes. Wet slab gels were scanned for ABP-emitted fluorescence using a Cy5 (λEX = 625 nm).

5.3.5. Intact Mass-Spectrometry

A reaction mixture of **18** (200 µM) with N1 (50 µM) was incubated at 37 °C overnight before preparing the sample for mass spectrometry by acetone precipitation. To remove contaminants, acetone precipitation was conducted by adding 4 volumes of ice-cold acetone to the sample and freezing for 2 hours before centrifuging at 10,000 xg at 4 °C for 10 minutes. The pellet was resuspended, and the procedure was repeated before removing the supernatant and resuspending the pellet in ice-cold methanol through sonication before centrifuging again. The methanol wash step was repeated, and the methanol was removed, allowing the sample to air dry for 5 minutes before resuspending in 20 µL ammonium acetate. The sample was submitted to the University of York Department of Biology technological facilities for mass spectrometry analysis.

5.4. Results and Discussion

For each of the N1-inhibitor (**17**, **18**, and **20**) complexes, micrographs were acquired through multi-grid data collection on a Titan Krios 1 (ThermoFisher Scientific) housed at eBIC Diamond Light Source UK. The microscope operated at 300 kV, calibrated pixel size 0.635 Å and a total dose of 43.99 $e^-/Å^2$ with defocus values -1.4, -1.2, -1.0, -0.8, -0.6 µm. All datasets were processed using RELION.

5.4.1. Cryo-EM structures of Viral neuraminidase

Initially, micrographs were motion corrected, and CTF estimation was carried out using RELION integrations MotionCor2 and CTFFIND4.1 [222, 230, 231]. Particles from the corrected micrographs were then picked using automated picking with the Laplacian of Gaussian filter; for the number of particles picked for each dataset, see Table 5.2. As a consequence of automated picking, junk particles can be extracted, and these are filtered out by iterative rounds of 2D classifications.

*Table 5.2- Number of particles used in RELION processing for compounds **17**, **18** and **20** with N1.*

	17	**18**	**20**
No. Micrographs	9519	8236	7298
No. Particles picked	4744877	4763583	4261747
No. Particles after final 2D classification	473421	408221	336195
No. Particles fed into iterative 3D refinement	235807	216385	229896

For each N1-inhibitor complex, iterative rounds of 2D classifications were conducted. The number of particles after each iteration decreased as the 'junk' particles were removed to produce 2D classes with high-quality features. The SNR improved over the multiple rounds of 2D classification, revealing the N1 tetrameric morphological features. Secondary structure characteristics are observed in the selected 2D classes, where the six-bladed beta-propeller monomers are defined (Figure 5.6)

.

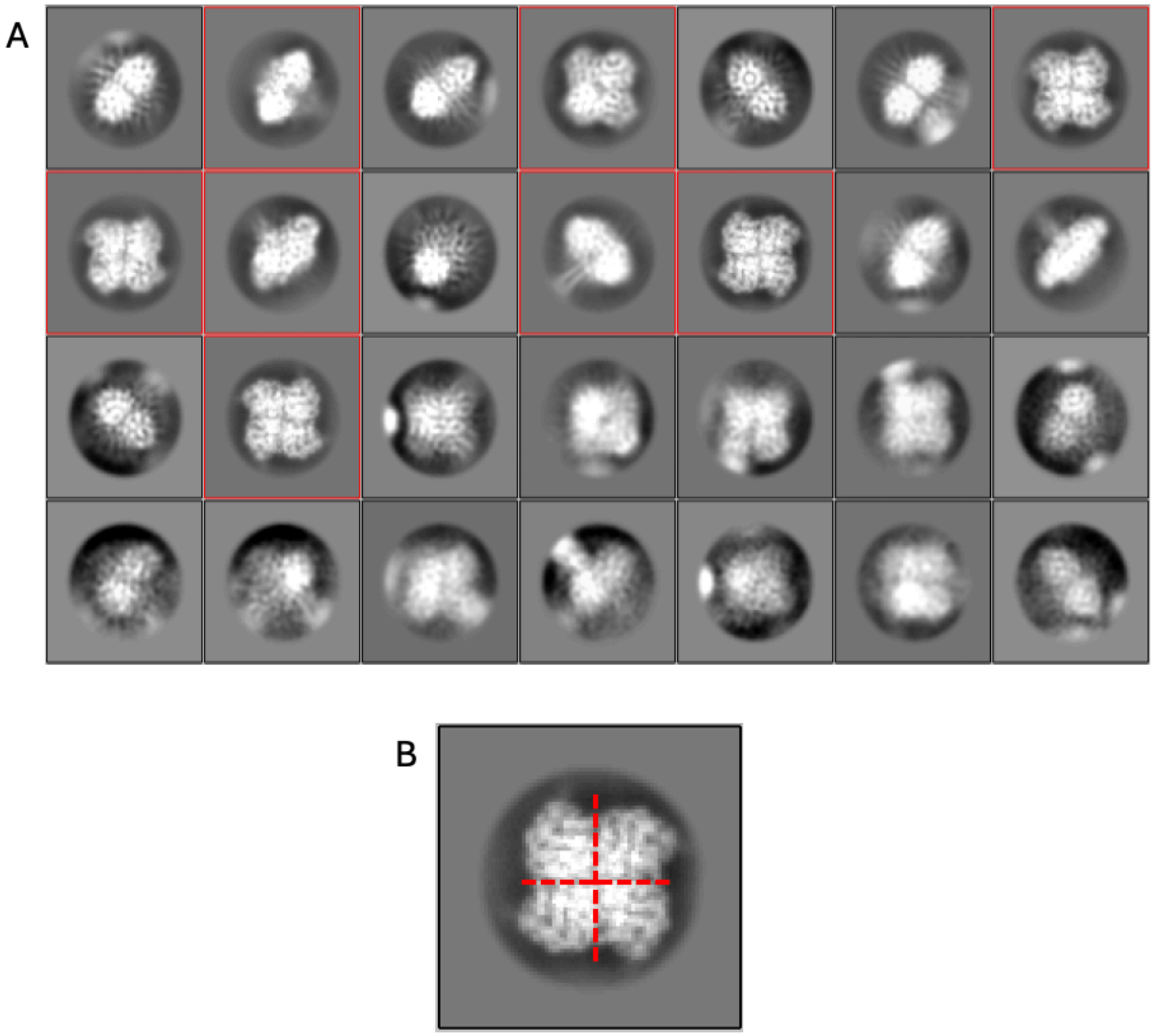


Figure 5.6- N1 2D classifications reveal C4 symmetry.

*These example 2D classes are taken from the N1-**17** complex dataset, starting with 4744877 particles for the initial 2D classification and 473421 particles after iterative rounds of 2D classifications. (A) Selected 2D classes of N1-**17** in red boxes are ordered by In Class Distribution with the most populated view starting from the left. (B) Red dots show the C4 symmetry of N1, where there are four monomers to create a tetramer.*

The particles from the high-quality 2D classes were fed into iterative rounds of refinement to generate the final 3D electron density map. Firstly, an initial model was generated without symmetry imposed using the 2D classes for reference-free projection modelling. After confirming symmetry in the UCSF Chimera, C4 symmetry was imposed using the RELION image handler. The final maps were generated through iterative rounds of 3D auto-refinement, post-process, CTF refinement and Bayesian polishing using a mask to obtain 2.0 Å, 2.0 Å and 2.25 Å resolutions for **17**, **18** and **20**, respectively (Figure 5.7). The maps were fitted into the PDB 3TI3 with mutations added manually. The pixel size was changed from 0.889 Å to 0.9 Å to better fit the model in UCSF chimera and then post-processed in RELION [133, 230]. the structures were refined in Coot and validated in Phenix. For processing details, see (Figure 5.8, Figure 5.9, Figure 5.10).

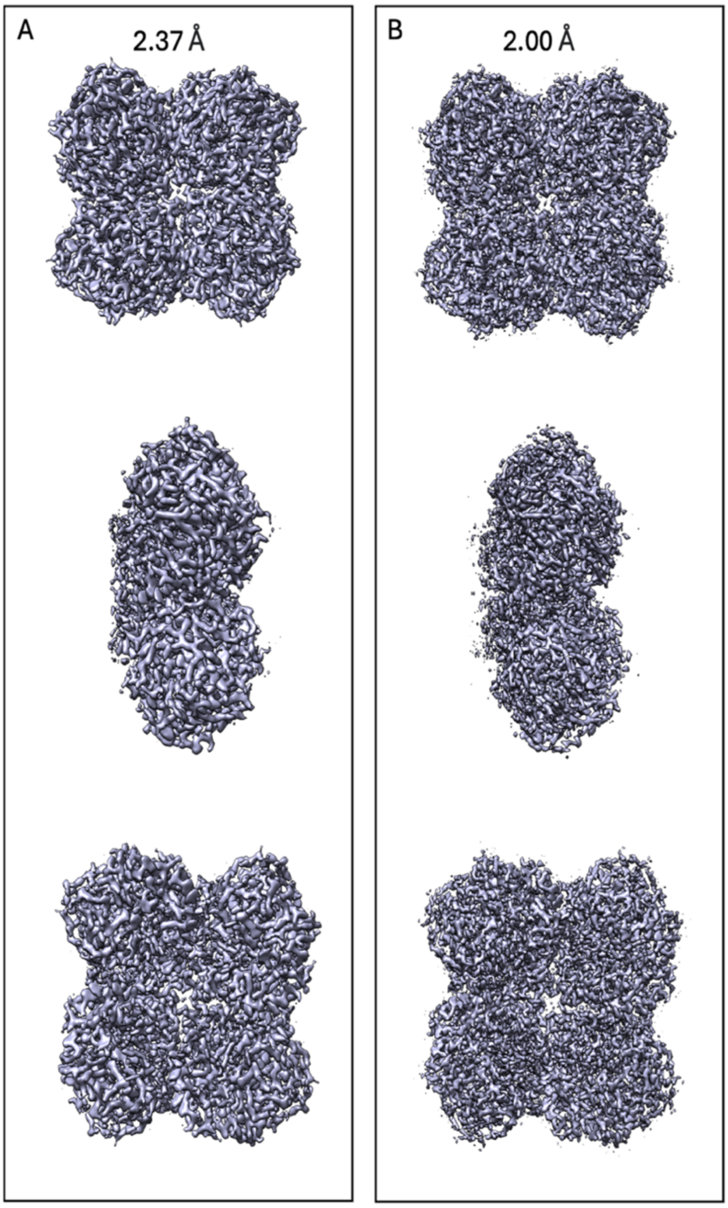


Figure 5.7- Refined electron density of N1 ligand complex.

*Refine 3D of N1 reacted with **17** employing C4 symmetry showing front, side and back views of the N1 tetramer. (A) The initial 3D refinement of N1-**17** starting at 2.37 Å. (B) Final refinement of N1-**17** at 2.00 Å from 235807 particles.*

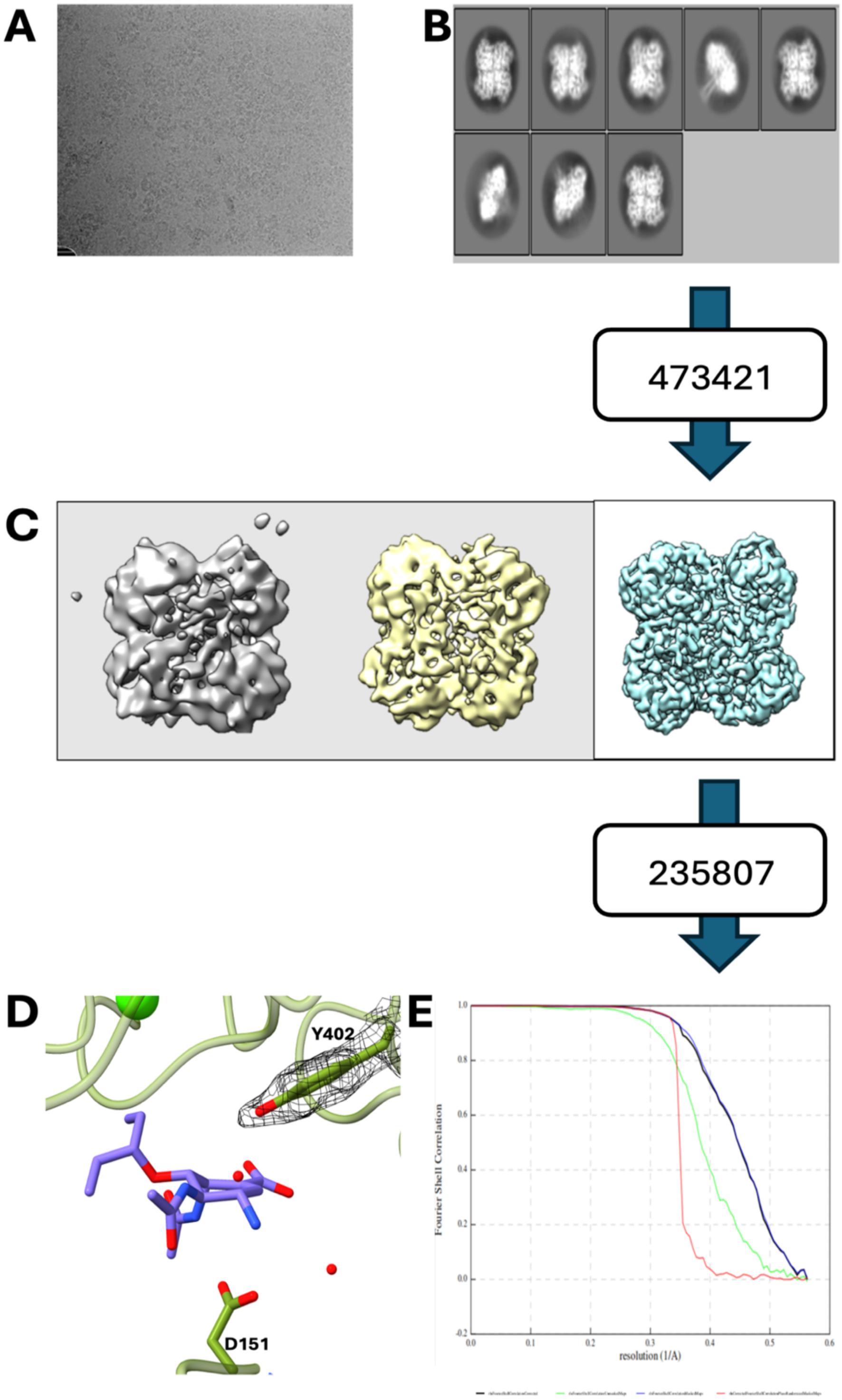


Figure 5.8- Cryo-EM reconstruction of N1 with compound **17** solved to 2.00 Å.

(A) Example micrograph. (B) Selected 2D classes were used for initial 3D reconstruction, and the selected class was highlighted. (C) Initial models, with the one select highlighted in white (D) An example of residue density is shown in the nucleophile Y402 contoured to 0.032. (E) Fourier shell correlation plot: Corrected FSC (Black), unmasked FSC (green), Masked FSC (Blue), and phase-randomised mask FSC (Red).

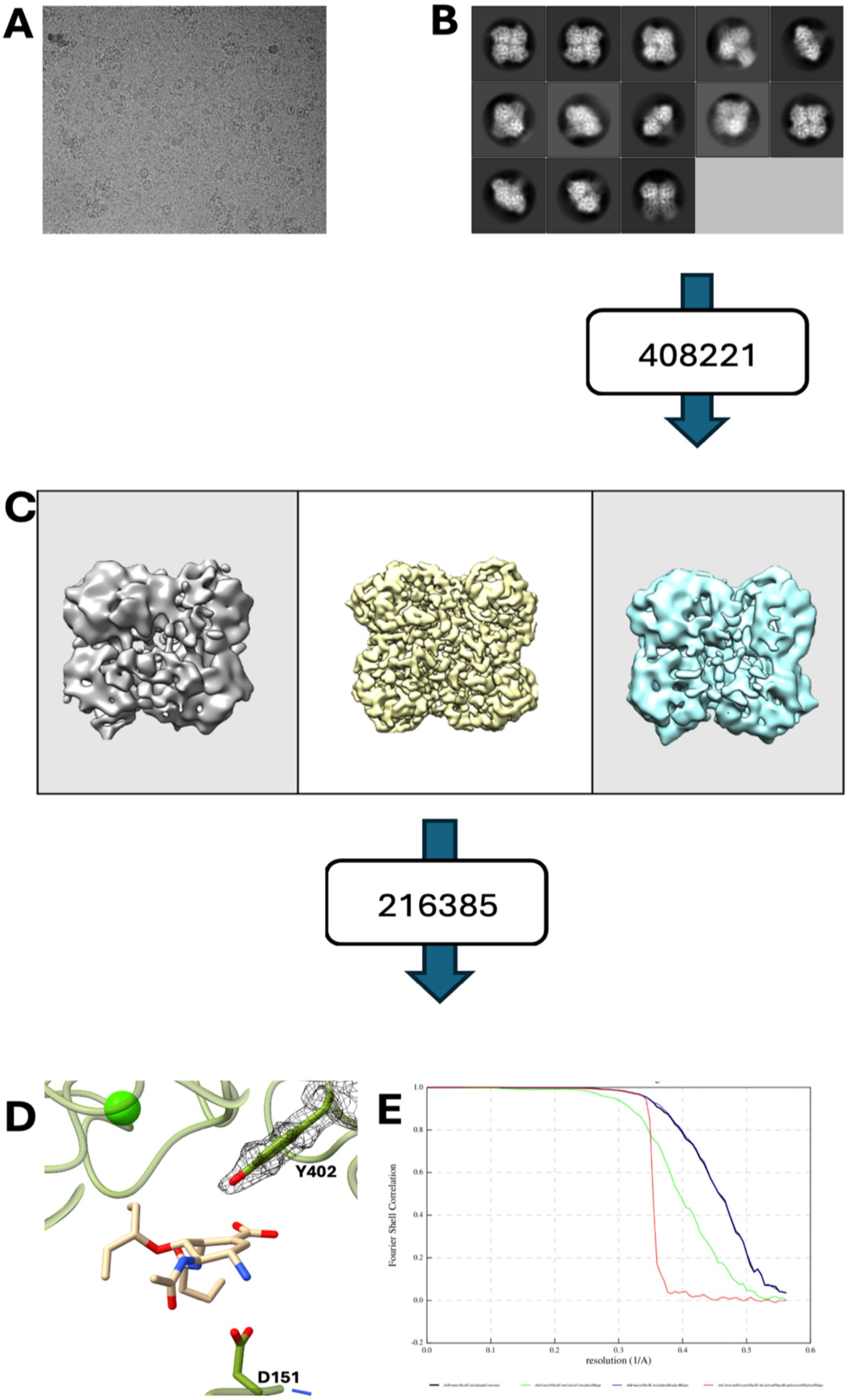


*Figure 5.9- Cryo-EM reconstruction of N1 with compound **18** solved to 2.00 Å.*

(A) Example micrograph. (B) Selected 2D classes were used for initial 3D reconstruction, and the selected class was highlighted. (C) Initial models, with the one select highlighted in white (D) An example of residue density is shown in the nucleophile Y402 contoured to 0.036. (E) Fourier shell correlation plot: Corrected FSC (Black), unmasked FSC (green), Masked FSC (Blue), and phase-randomised mask FSC (Red).

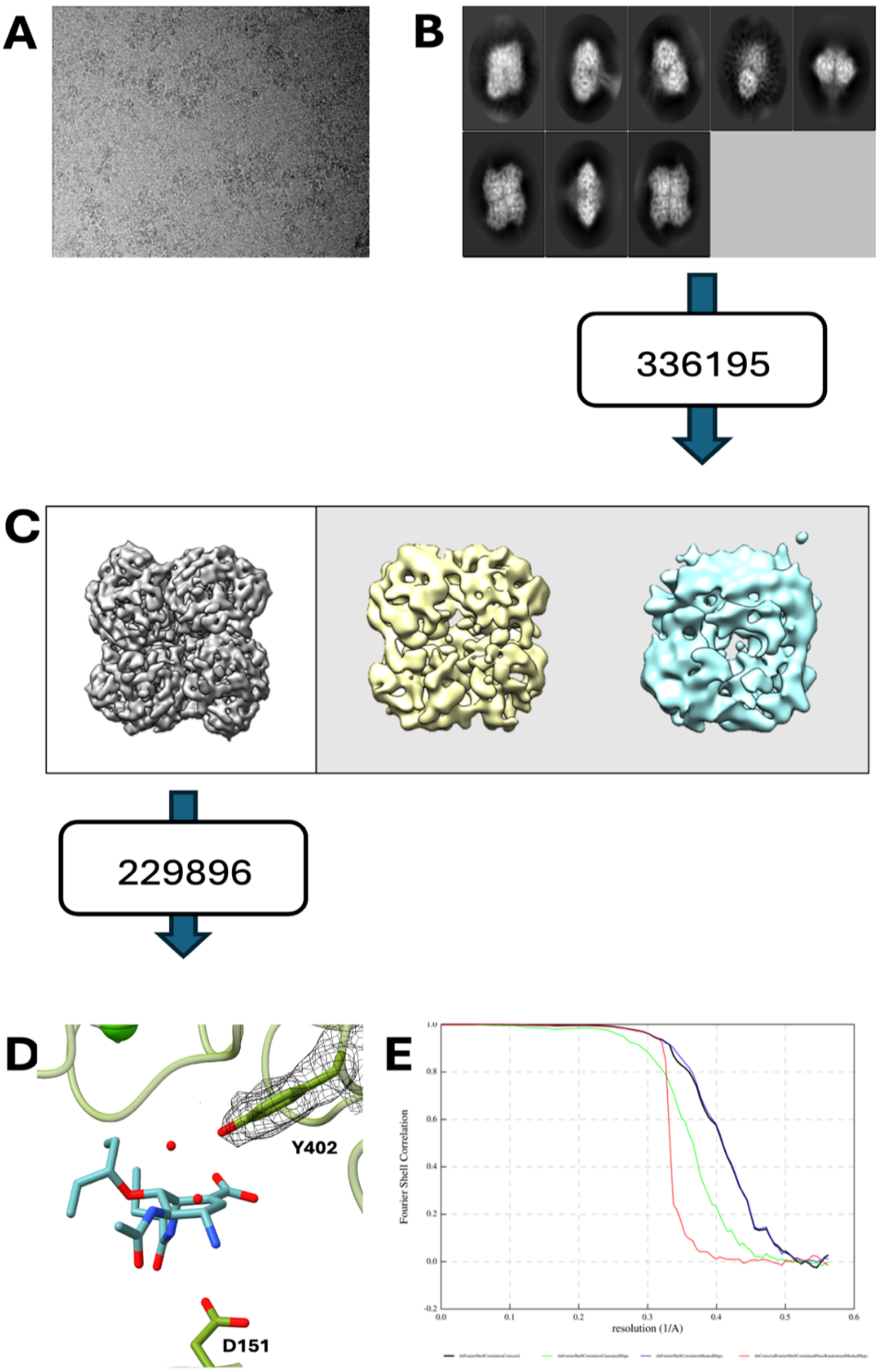


Figure 5.10- Cryo-EM reconstruction of N1 with compound ***20*** *solved to 2.25 Å.*

(A) Example micrograph. (B) Selected 2D classes were used for initial 3D reconstruction, and the selected class was highlighted. (C) Initial models, with the one select highlighted in white (D) The nucleophile Y402 contoured to 0.037 shows an example of residue density. (E) Fourier shell correlation plot: Corrected FSC (Black), unmasked FSC (green), Masked FSC (Blue), and phase-randomised mask FSC (Red).

As expected, the high-resolution cryo-EM structures revealed the N1 homotetramer, composed of monomers with six-bladed beta-propeller folds stabilised by disulphide bonds. Firstly, the structure for **18** was resolved, and the inhibitor was bound in the acidic active site, coordinated by the nucleophile Tyr402 and general acid/base Asp151 [168]. Surprisingly, there was no density for a covalent adduct in any of the inhibitor structures with the N1 nucleophile Tyr402, suggesting that the inhibitors were either unreacted (non-covalent) or eliminated. Initially, **18** was modelled as an unreacted non-covalent species. To further confirm that covalent bond formation is, in fact, feasible, our collaborators, the Rovira group (University of Barcelona), conducted Molecular dynamics (MD) and Quantum mechanics/molecular mechanics (QM/MM) simulations/calculations. The formation of the covalent adduct was modelled using on-the-fly probability-enhanced sampling (OPES) QM/MM MD calculations where a free energy landscape (FEL) was produced. The FEL suggests that the covalent adduct formation depends on the protonation and orientation of the acid/base Asp151 (Figure 5.11). The opening of the aziridine ring via nucleophilic attack is initiated by the protonated Asp151, forming a hydrogen bond with the nitrogen of the aziridine (TS1). Once the aziridine is protonated (INT) (Figure 5.11), the reaction proceeds with a nucleophilic attack to open the aziridine ring via the TS2, which has an energy barrier of 11.8 kcal/mol, forming the covalent adduct. As the initiation of covalent adduct formation depends on the protonation of Asp151, cryo-EM collections with N1-**18** were repeated using different reaction conditions to ensure the protonation of Asp151. Nevertheless, again, this data set yielded no covalent adduct, leading to the investigation of whether the inhibitors are, in fact, covalent in solution work or if the inhibitor is eliminated as a consequence of radiation damage during cryo-EM data collection.

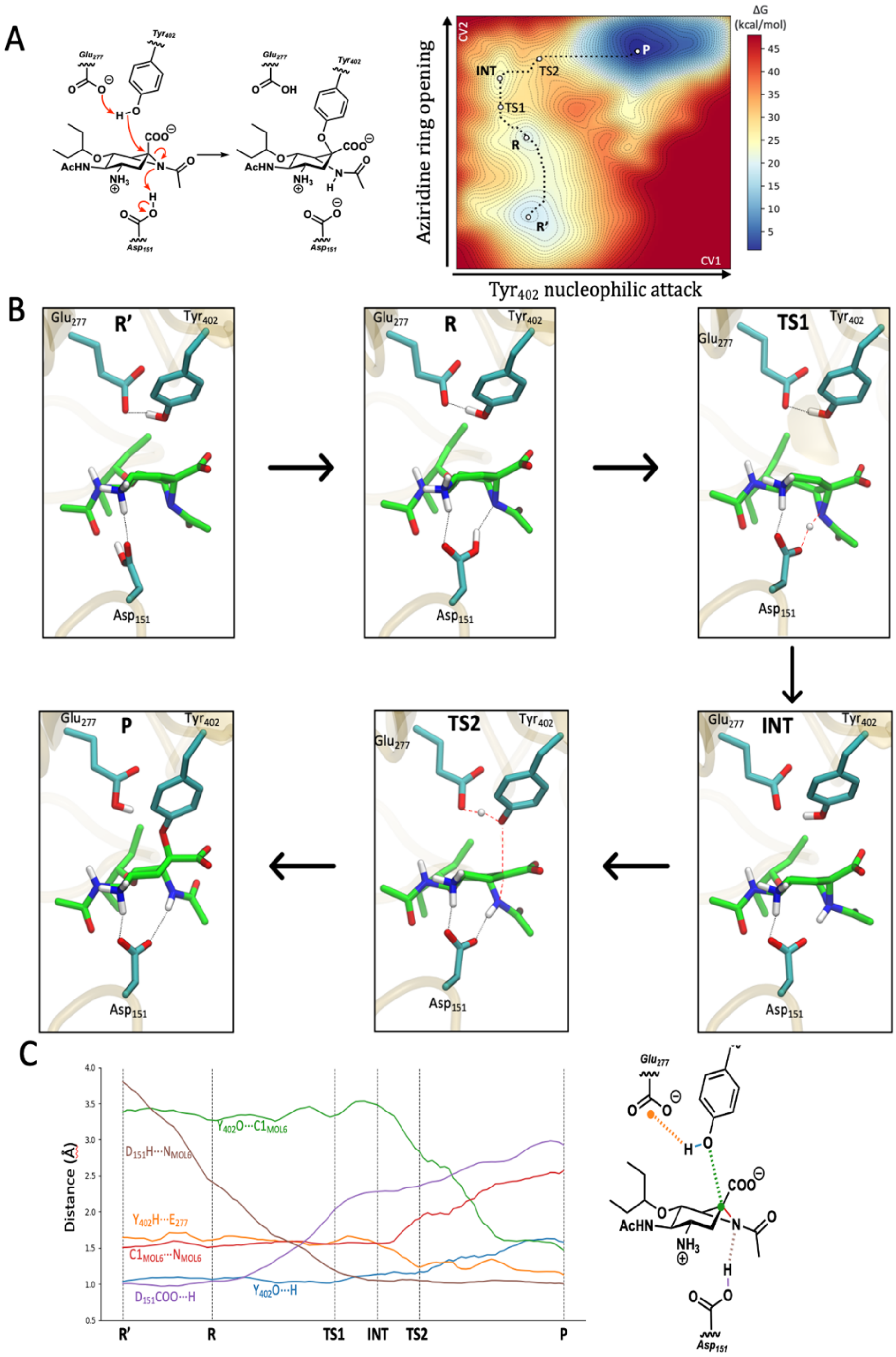


*Figure 5.11- Quantum mechanics/molecular mechanic (QM/MM) molecular dynamics simulations for the covalent reaction of N-acyl aziridine (**17**) inhibitors with N1 (PDB:3CL2). (A) Reaction of acid/base Asp151 and nucleophile Tyr402 with N-acyl aziridine through the aziridine ring opening via nucleophilic attack represented by the free energy landscape. (B) Simulated structures of the leading states observed in forming the covalent adduct between Tyr402 and N-acyl aziridine. (C) Evolution of the distances between catalytic residues and inhibitors along the minimum path energy for forming a covalent species. The simulation was conducted by the Rovira group (University of Barcelona).*

In reference to **Chapter 4**, Figure 4.2, the competitive ABP gel suggests that the N-acyl aziridine inhibitors react covalently, as they all produced significantly fainter bands compared to the control as they out-competed the covalent ABP **26**. To investigate if potentially the inhibitor is eliminated before cryo-EM grid preparation, an additional ABP assay was conducted (Figure 5.12A), where N1 (1 µM) was reacted with **18** (10 µM) at 37 °C and time points taken as indicated before reacting with ABP **27** and visualising on a SDS-PAGE gel. As no bands are formed for reactions, which contained **18**, the inhibitor outcompetes the probe. Suggesting the elimination reaction occurs during cryo-EM preparation/data collection. The covalency of **18** was further confirmed by Intact mass spectrometry (MS). Intact MS is ideal for investigating covalent adduct formation, as there will be no shift in mass from non-covalent interactions. Initially, MS analysis was challenging as the N1-inhibitors complexes did not fly well in electrospray ionisation (ESI) MS. Instead, low-resolution matrix-assisted laser desorption/ionisation (MALDI) MS was used to analyse the shift in mass from the covalent reaction of N1 with **18**. The shift between the control and reacted N1-**18** is ~320 Da, similar to the expected mass of **18** (369.46 Da), suggesting that the inhibitor reacts covalently (Figure 5.12B). Therefore, the non-covalent species in the cryo-EM structure is an eliminated product produced from electron damage during TEM data collection. Other cryo-EM datasets were collected and processed with inhibitors 1**7** and **20**, which yielded the same non-covalent species.

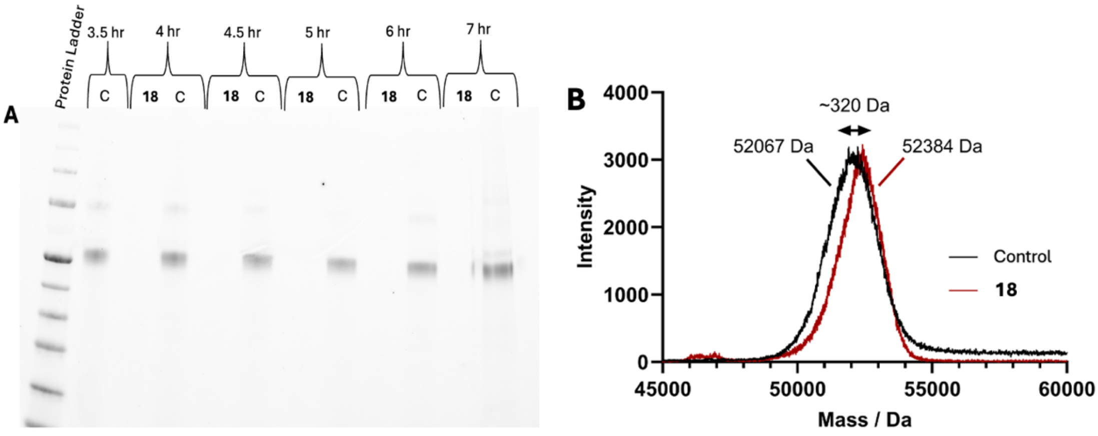
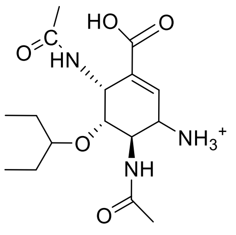


*Figure 5.12- Solution work of N1-**18** to prove covalent reaction.*

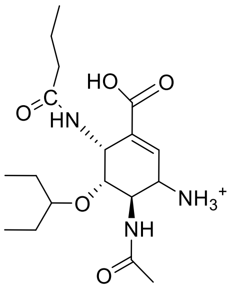
*(A) Time course-labelling of N1 by **18** analysed by ABP **26** gel. Timepoint samples of the N1-8 reactions were taken at the indicated times before reacting with ABP **26** to assess if an elimination reaction occurred. (B) MALDI intact MS of N1-**18** using cryo-EM grid preparation reaction conditions.*

5.4.2. Elimination of inhibitors during Cryo-EM acquisition

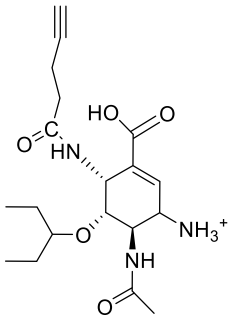
In light of the solution work confirming the covalent reaction of inhibitors with N1, the inhibitor species were instead modelled as elimination products of subsequent covalent reactivity (Figure 5.13). The eliminated inhibitors are present in the charged active site of N1, where they are coordinated by hydrogen bonding with residues (E278, E227, E119, D151 and Y402)nd the arginine triad (R118, R293 and R368).



17E



18E



20E

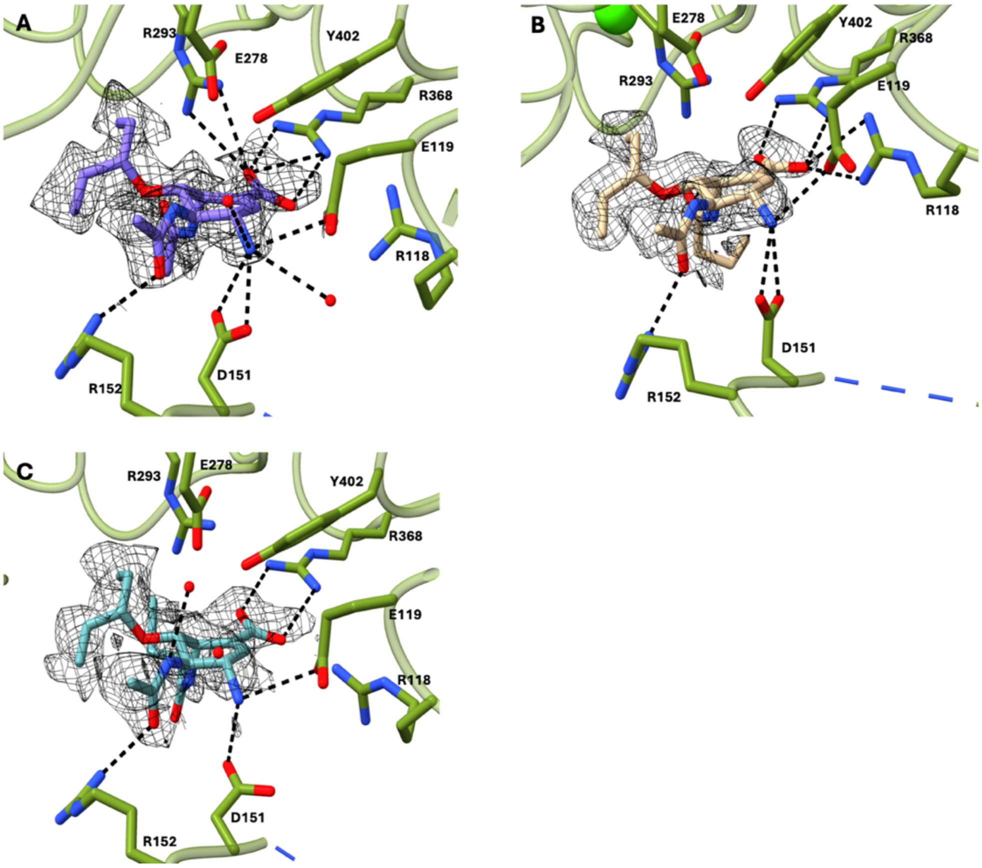


Figure 5.13- Cryo-EM structures of N1 with N-acyl aziridines inhibitors.

*Cryo-EM structures of **17E** (A), 1**8E** (B), and 20E (C) bound in the N1 active site with catalytic residues Y402 and D151. Compounds 7E, 8E, and 10E are shown with density at a threshold of 0.029, 0.0358, and 0.0278, respectively, and solved to 2.0 Å, 2.0 Å, and 2.25 Å.*

This is not the first circumstance of neuraminidases catalysing elimination reactions, most notably, the synthesis of its own inhibitor, DANA, from its natural substrate, sialic acid, via an elimination reaction for influenza B virus neuraminidase [235]. Similar to our compounds, the covalent DFSA (described in **Chapter 4**) showed complete covalency in solution work, but when solved by X-ray crystallography in complex with NA, an eliminated species was observed; its presence was described as a consequence of X-ray radiation damage contributing to the elimination reaction of the covalent DFSA [183]. Thus, we have also presumed that the cryo-EM structures of **17**,**18** and **20** contain eliminated species **17E**, **18E** and **20E** due to electron damage during cryo-EM data collection, as the compounds have otherwise been shown as covalent in solution work.

Furthermore, the feasibility of an elimination reaction was confirmed by additional QM/MM MD simulations. The same computational methods as the covalent adduct reactions were used. The calculations conducted by the Rovira group suggested that the elimination reaction proceeds through one TS with a barrier of 15.2 kcal/mol where there is the protonation of the Asp151 from the aziridine of the inhibitor and the protonation of Tyr402 assisted by Glu277; breaking the covalent bond forming the eliminated aziridine (Figure 5.14). The formation of an eliminated product is plausible, as during cryo-EM data acquisition, it was shown that radiation damage can break covalent bonds [212]. Additionally, tyrosine residues have also been implicated in elimination reactions; this is the case for the CAZymes polysaccharide lyases [236].

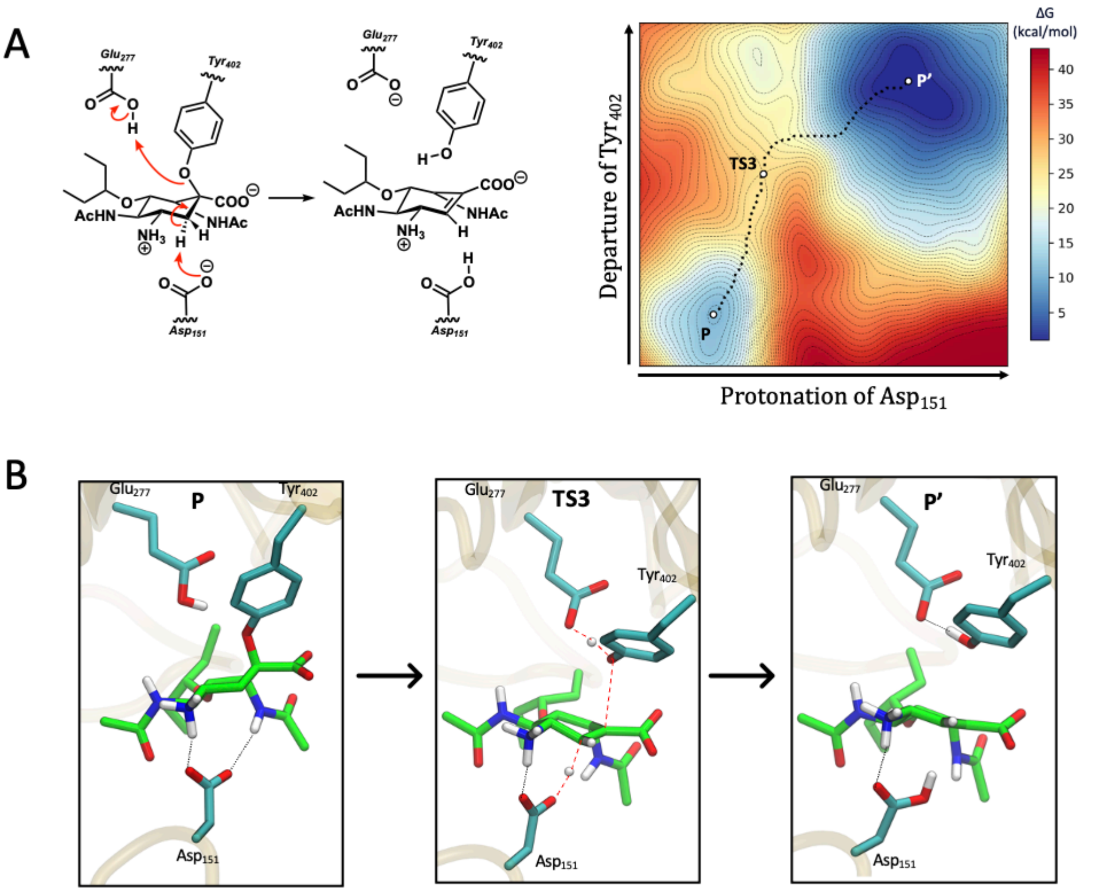


*Figure 5.14- Quantum mechanics/molecular mechanics (QM/MM) molecular dynamics simulations (MD) for the elimination reaction of N-acyl aziridine (**17**).*

(A) The Free energy surface of the N-acyl aziridine elimination reaction. (B) The Main states observed in the elimination reaction. Simulations conducted by the Rovira group (University of Barcelona).

5.5. Summary

In this work, recombinant viral N1 was solved by cryo-EM in complex with N-acyl aziridine inhibitors **17**, **18** and **20**. As discussed, the electron density map contained no signs of covalent adduct formation; instead, there was an eliminated species. The formation of an eliminated species by viral NAs has been previously observed by reactions with the natural substrate and other similar covalent inhibitors; the latter formed as a consequence of X-ray damage [183, 235].

Radiation damage by X-rays during X-ray crystallography has been extensively studied [237-239]. The main consequences of radiation damage are the breakage of bonds, decarboxylation of residues, and chemical reduction of metal ions [237]. Radiation damage caused by electron beams manifests similarly where the breakage of disulphide bonds and residue modifications have been observed [237]. However, the collection of cryo-EM datasets requires a high dose of electrons to obtain high-resolution structures without much consideration for sample radiation damage [212]. To circumvent this, the earlier frames of micrographs can be used, as these have less radiation damage but the most beam-induced motion; through processing, high-resolution structures can still be obtained from these frames [212]. However, when this methodology was applied to this dataset, the eliminated species was still observed; this could be because the covalent bond between Y402 and the inhibitor is particularly susceptible to electron damage even in the earlier frames, or a better detector would be needed to capture the covalent species [212].

Furthermore, the ABPP assays conducted by our collaborators (Overkleeft group, University of Leiden) demonstrated the utility of ABP (**26**) for the rapid and quantitative visualisation of viral neuraminidases. The labelling of N1 by the ABP can be inhibited by pre-incubation of the enzyme with an inhibitor in a concentration-dependent manner (Figure 5.15). The ABPP implies that the **27** ABP can be utilised for the rapid and quantitative visualisation of viral neuraminidases; this can also be applied to biological and vaccine samples, with the potential for *in vivo* imaging.

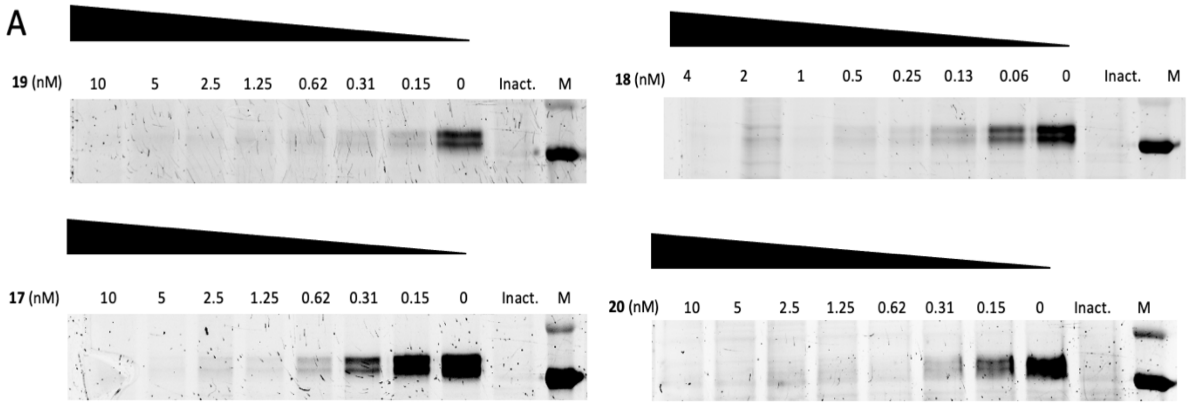


*Figure 5.15—Labelling of N1 using ABPP. The inhibition of ABP (**26**) labelling by inhibitors **17**, **18**, **19**, and **20** was investigated in a concentration-dependent manner.*

*N1 was incubated with the indicated concentrations for each inhibitor at 37 ℃ overnight before adding 1 uM of **26** at 37℃ and incubating for 1 hour. The samples were prepared and loaded onto an SDS-PAGE gel before visualising the cy5 fluorescence. This work was conducted by the Overkleeft group, University of Leiden.*

Together with the data obtained in **Chapter 4** and **Chapter 5**, such as the viral plaque assays, the cryo-EM structures, and ABPP, we can conclude that the cyclophellitol aziridine-based inhibitors are potent covalent inactivators of viral neuraminidases. Furthermore, these inhibitors serve as alternatives to the fluoroneuraminic acid derivatives that have been reported before, and their mechanism-based nature is advantageous as they permanently block catalytic activity [183]. Lastly, the N-acyl and free aziridine inhibitors are potential alternatives to the commercialised antiviral oseltamivir, for which resistant strains have emerged.

6. Chapter 6- Conclusions and Future Perspectives

6.1. This Work

As discussed in **Chapter 1,** the huge diversity of carbohydrate structures (oligosaccharides, polysaccharides and glycoconjugates) reflects their many roles in biological processes and is further mirrored by the numerous enzymes involved in their synthesis and degradation [102]. The work discussed in this thesis focused on the class of carbohydrate-processing enzymes, glycoside hydrolases, where their numerous roles in human physiological processes, cancer and viral disease were highlighted. One aim of this work, in collaboration with Prof. Overkleeft (University of Leiden) and Prof. Rovira (University of Barcelona), was to design mechanism-based inhibitors to inhibit influenza infectivity. Specifically, I sought to characterise the protein-inhibitor complexes using cryo-EM to inform the subsequent design of future inhibitors.

6.1.1. Mechanism of Siastatin B Inhibition with HPSE

The overexpression of HPSE strongly drives aggressive cancerous phenotypes and facilitates the growth of malignancies; because of this, HPSE is considered an attractive target for the development of cancer therapies. Currently, there are no licensed HPSE inhibitors; in **Chapter 2**, the work aimed to expand on the current portfolio of HPSE small molecule inhibitors and inform the design of more potent anticancer compounds through the elucidation of the mechanism of siastatin B inhibition with HPSE. Siastatin B is a known broad-spectrum iminosugar inhibitor of multiple glycosidase hydrolases, including neuraminidases and β-glucuronidases. The mechanism of inhibition for the β-glucuronidase HPSE by siastatin B had been a mystery, as the N-acetyl group of the compound appears too bulky to be accommodated in the HPSE active site. This chapter uncovered the true inhibitors through X-ray crystallography structures of HPSE and homologs with siastatin B [71]. Further analysis suggested that the siastatin B breakdown products, a hemiaminal and a 3-GDI, were the true inhibitors. Subsequently, the degradation products were synthesised by the Overkleeft group, but the synthesis of the 3-GDI inhibitor was difficult, and there was no increase in potency compared to the siastatin B breakdown product [71]. In summary, the mechanism of siastatin B activity as a broad-spectrum inhibitor was uncovered, where the 3-GDI represented a new class of glycosidase inhibitors. This information, in turn, could enable the future design of a "pro-inhibitor" design for glycosidase inhibitors. To delay the rapid decay of siastatin B into the 3-GDI, the 2-position can be optimised [71]. This "pro-inhibitor" strategy for the decay of molecules into their iminosugar active form could also be applied to developing enzymes involved in lysosomal storage disease [71].

Although uncovering the true mode of inhibition of siastatin B is useful for the development of future broad-spectrum inhibitors, there is still a requirement for HPSE-specific inhibitors to reduce cancer metastasis. In the case of the 3-GDI, a competitive small molecule inhibitor, an initial high affinity is required for binding of the inhibitor to the binding site of HPSE, which normally accommodates large HS oligosaccharide substrates through extensive interactions [51]. This high affinity is difficult to access for competitive inhibition, but this is not true for mechanism-based, covalent irreversible inhibitors [51]. Cyclophellitols are mechanism-based inhibitors which have been explored before [38, 51, 69]. The cyclophellitol-derived epoxides and aziridines have been proven to be powerful inactivators of β-glucosidases as they can selectively label the catalytic nucleophile [69, 70]. In relation to this, a 1,6 cyclophellitol disaccharide was synthesised (Overkleeft group, University of Leiden), and the structure of **28** was solved in complex with HPSE by X-ray crystallography. The electrophilic epoxide warhead of **28** selectively targeted the inhibitor to HPSE nucleophile Glu343, forming a covalent bond (Figure 6.1). These types of mechanism-based inhibitors are more specific and easier to synthesise than compound **1** and may provide a better avenue for the development of anticancer compounds.

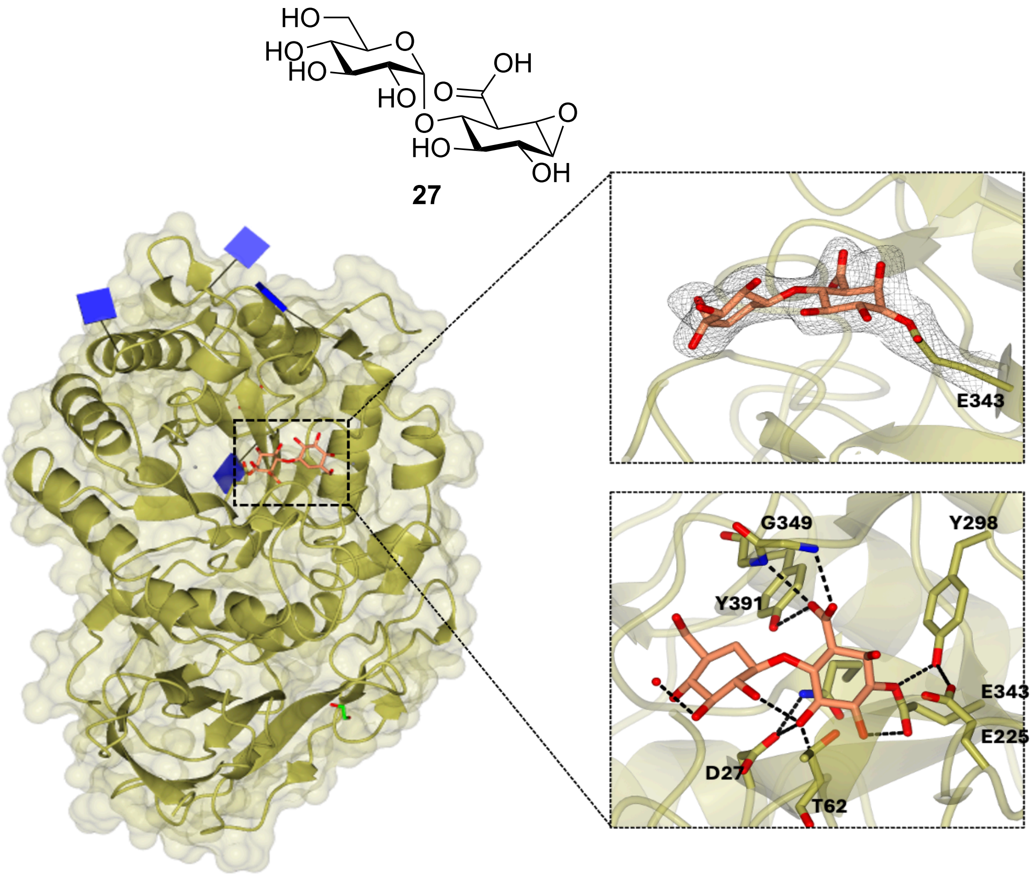


Figure 6.1- *Covalent complex of HPSE with compound **27**.*

*HPSE with compound **27** bound in the active site (right panel) and compound **27** (orange) covalently bound to the catalytic nucleophile E343 in the -1 and -2 subsites refined to 2.4 Å (right panel). Glycosylation NAG is found at Asn 162, 200, 217, 238, and 459. Images created in CCP4mg (version 2.11.0) with* F_o*-*F_c *maps (grey) have been displayed in chicken wire and contoured to 0.2* $e^-/Å^3$.

6.1.1. 2 Host-Targeted Broad-Spectrum Antivirals for Endoplasmic Reticulum α-Glucosidases

Many iminosugars and their derivatives have been assessed in clinical trials as potential host-targeted broad-spectrum antivirals but have failed due to their adverse effects from poor specificity, commonly due to off-target inhibition of intestinal α-glucosidases [120, 240, 241]. Comparatively, the 1,6-*epi*-cyclophellitol cyclosulfate discussed in **Chapter 3** is much more specific, targeting only retaining α-glucosidases. In **Chapter 3**, a panel of cyclophellitol derivatives were assessed for their inhibitory potency with ER α-GluII for reducing SARS-CoV-2 infectivity. The 1,6-*epi*-cyclophellitol cyclosulfate was identified as the lead inhibitor of the series where the inhibitor effectively reduced SARS-CoV-2 infectivity through the inhibition of ER-αGluII, halting the maturation of the viruses spike protein, causing a reduction in the infectivity of the progeny virus. Collectively, this work validates the use of cyclophellitol-based inhibitors for the inhibition of ER-αGluII, where they serve as a potential new strategy for host-targeted antivirals by reacting in a covalent and irreversible manner. Adverse side effects from **14** could still occur as a result of inhibition of lysosomal GAA, although these compounds are still more specific than iminosugars. Nevertheless, the work presented provides a foundation for developing cyclophellitols as new antiviral host-targeted agents for present and emerging viruses that rely on the ERQC.

The COVID-19 pandemic highlighted the need for host-targeted antivirals to fight the emergence of new viruses. To further validate the effectivity of **14** for reducing viral infectivity, other enveloped viruses, such as Dengue, HIV and Influenza, should be examined, as their dependency on the ERQC varies. The 1,6 and 1,2 cyclophellitol derivatives could also have future applications as anticancer compounds. The ERQC and ERAD pathways preside over cellular glycoprotein synthesis and secretion; these pathways are particularly relied upon during cancer progression [92]. Therefore, it is conceivable that ER-αGluI and ER-αGluII cyclophellitol inhibition could be used as an anticancer strategy.

6.1.2. Structural Analysis of Viral Neuraminidase with Mechanism-Based Inhibitors.

To combat the emergence of antiviral-resistant influenzas strains, many direct-acting antivirals for viral NA inhibition have been developed, many of which are analogues of sialic acid and commercialised antivirals [177-179]. In response to this, in **Chapter 4** and **Chapter 5,** I discussed the structural characterisation of viral NA N1 with N-acyl aziridine inhibitors (Synthesised by Overkleeft group and QM/MD analysis by Rovira group) based on the commercialised antiviral Oseltamivir. The series of inhibitors investigated contained mechanism-based covalent N-acyl aziridines, while the free aziridine acts as a competitive inhibitor. The competitive inhibitor was the lead candidate of the series and matched the activity profile of the commercialised antiviral Oseltamivir but with improved potency and activity with Oseltamivir-resistant strains.

The free aziridine is a potential alternative to the current competitive, commercialised antivirals, whereas the mechanism-based N-acyl aziridines serve as alternatives to the formerly mentioned DFSA. Although the mechanism-based inhibitors are not lead candidates, their covalent nature has recently increased in popularity. These types of inhibitors permanently block catalytic activity, so they have advantages in pharmacokinetics and dosing. Furthermore, the inhibitors discussed provide a framework for developing future mechanism-based inhibitors as antiviral drugs, as the aziridine offers the possibility for chemical derivatisation, such as using an electrophilic epoxide warhead.

6.2. Future challenges for glycosidase ABPP

The emergence of ABPP in the past two decades has solidified the technique as a powerful strategy for the identification and study of active enzymes both *in vivo* and *in vitro*. The first instance of ABPs was for cysteine proteases and serine hydrolases, although the probes can, in theory, be designed for any enzyme that forms a covalent enzyme-substrate adduct [27, 124, 242]. This makes retaining glucosidases a tractable target for ABP design as they employ a Koshland two-step double displacement mechanism. Previous work by Davies group and collaborators has developed multiple cyclophellitol base ABPs that have demonstrated applications in in-gel fluorescence labelling, tissue labelling, cell imaging and proteomics [33, 70, 243-246].

The field of ABPP for retaining glycosidases is now reaching maturity, as many ABPs have been designed for specific GH families with different warhead chemistries [102, 247, 248]. There is now an extensive choice of probe chemistries to choose from for retaining GHs; in general, cyclophellitol is the first choice for ABP design, as it is the most widely used. These ABPs typically adopt the ground state $^4H_3$ conformation, the conformation of many glycosidases. Still, for ones who don’t adopt this ground state, ABP design is difficult as alternatives such as fluorosugar inhibitors for ABP are rapidly turned over [33]. Additionally, retaining *exo*-glycosidases ABP has been well explored compared to the *endo*-glycosidases as the synthesis of a monosaccharide ABP is more attenable compared to the length/branching needed for *endo*-acting ABP [33]. For all the protein targets discussed in this thesis, ABPP was utilised to characterise protein-inhibitor interactions of retaining GHs [33]. However, for the inverting ER-aGluI discussed in **Chapter 3**, there is no ABP, as the synthesis of such molecules would require a more in-depth understanding of the conformational itineraries of the protein and a high level of synthetic chemistry. Opening up inventing enzymes to ABPP would have huge potential with applications in biotechnological and biomedical areas [33].

7. Appendix

7.1. Appendix: Gene constructs

7.1.1. Ct ER-αGluI sequence cloned into pPICZalphaA

MGILPSPGMPALLSLVSLLSVLLMGCVAETGVEGESILHSEIGRLNNQSLLWGPYRPNIYFGTRPRIGKSLMTGLMWGKIES
YTDFQHTVRYTCEQNEGMKGYGWDEYDPRRGGIQSIHDIQNGLDITTSFVKIPGGAHGGSWAARIKGTLNDDAPKDQK
TIVVFYVSQEGENSELEAVPSENEFGYEGDVILKGRSEALGNYKLVVTKGKGVIPQSDHDLSRLRGPGQTVVQSLTYPDEVL
WQAKPILFQQLKAGIDWLVENKYDVADPPPPWQVYLLANKPGSGNVHIVQKVFEGDFEFDILFSSESAGKEVTSKDLERE
VKQATEVFGERFARVFDLKAPFQGDNYKKFGKSMFSNLIGGIGYFYGHSLVDRSYAPEYDEENEGFWEDAAEARARHQEA
LEGPYELFTSIPSRPFFPRGFLWDEGFHLLPIADWDIDLALEIIKSWYNLMDEDGWIAREQILGAEARSKVPKEFQTQYPHY
ANPPTLFLVLDNFVERLRKNNASQPVVKDNLSLDETLSTASVDNPEVGLEYLRRLYPLLRRQFDWFRKTQAGDIKSYDREA
YSTKEAYRWRGRTVSHCLTSGLDDYPRPQPPHPGELHVDLMSWVGVMVKSLISIGSLLGATEDVEFYTKVLDAIEHNLDD
LHWSEKEGCYCDATIDEFEEHKLVCHKGYISLFPFLTGLLKPDSPKLGKLLALIGDESELWSPYGLRSLSKKDEFYGTAENYWR
SPVWININYLAIVQLYNIATQDGPYKETARDLYTRLRKNIVETVYRNWEETGFAWEQYNPETGKGQRTQHFTGWTSLVVKI
MSGHHHHHH

7.1.2. Ct ER-αGluI cloned into pFastBacTM vector

ATGAAATTTTTGGTGAACGTGGCCTTGGTGTTTATGGTGGTTTACATTAGCTACATTTATGGCGGACCCGGGAGAAAA
CTTGTACTTTCAAGGCAGCGTAGAAGGAGAATCCATCTTGCATTCTGAGATAGGTCGATTGAACAATCAAAGTCTGCTC
TGGGGCCCGTACCGGCCTAACATTTATTTCGGAACCCGGCCCCGTATTGGAAAAAGTCTTATGACTGGGCTTATGTGG
GGAAAGATTGAGTCATATACCGACTTCCAACACACCGTCCGCTACACGTGCGAACAGAATGAGGGTATGAAGGGCTAT
GGCTGGGATGAGTACGACCCCCGGAGGGGCGGGATTCAGAGCATCCATGACATCCAAAACGGCCTAGATATAACTAC
TTCGTTCGTTAAGATTCCGGGCGGGGCTCATGGTGGCTCCTGGGCAGCTCGTATTAAGGGGACGTTGAACGATGACG
CCCCTAAGGACCAAAAAACTATCGTAGTGTTTTACGTCTCTCAAGAAGGTGAAAACAGTGAATTAGAGGCTGTGCCAT
CTGAGAATGAATTTGGCTATGAGGGGGATGTAATTCTTAAAGGTAGGTCCGAAGCACTGGGTAATTATAAGCTGGTGG
TCACCAAAGGTAAAGGAGTCATACCACAATCCGATCACGATCTGTCTAGGTTGCGCGGCCCGGGACAAACAGTAGTTC
AGTCGCTTACCTATCCCGATGAAGTGCTATGGCAAGCGAAGCCGATCCTATTCCAACAACTCAAAGCTGGAATTGACTG
GTTGGTAGAAAACAAGTACGATGTTGCCGATCCGCCGCCGCCCTGGCAGGTCTACTTACTAGCAAATAAGCCTGGATC
GGGCAACGTGCACATCGTACAGAAGGTCTTTGAAGGGGATTTCGAGTTCGATATATTATTTTCAAGCGAATCTGCGGG
GAAAGAGGTCACATCTAAAGACCTTGAAAGAGAAGTGAAACAGGCTACTGAAGTTTTTGGGGAGAGATTCGCGCGA
GTATTTGATCTGAAAGCACCTTTTCAAGGGGACAATTATAAAAAGTTCGGAAAATCGATGTTCTCGAACTTAATCGGGG
GAATAGGATACTTTTACGGACACTCACTTGTTGATCGAAGTTACGCGCCAGAGTATGACGAGGAGAACGAGGGTTTTT
GGGAAGACGCCGCTGAAGCGCGTGCTAGGCACCAGGAAGCCCTCGAAGGCCCGTACGAGCTCTTCACATCCATACCA
AGTAGACCCTTTTTCCCTCGGGGATTTCTTTGGGATGAGGGTTTTCATCTCCTGCCAATTGCGGACTGGGATATAGATC
TAGCTTTGGAGATTATCAAAAGCTGGTACAATTTGATGGACGAGGACGGATGGATAGCGCGGGAACAGATACTCGGA
GCAGAAGCCAGGAGCAAGGTCCCTAAAGAATTTCAGACGCAGTATCCCCACTATGCAAACCCACCTACGCTGTTCTTA
GTGTTGGACAACTTTGTTGAACGTCTCAGGAAAAATAATGCAAGCCAGCCGGTCGTAAAGGACAATCTCAGTCTGGA
CGAAACCTTAAGCACTGCAAGTGTGGACAATCCGGAGGTTGGCCTGGAGTATCTGAGACGACTTTACCCACTATTGCG
CAGACAGTTTGATTGGTTCAGAAAAACACAAGCGGGGGACATTAAGTCATACGACCGGGAGGCCTACTCTACGAAAG
AGGCTTACCGCTGGCGCGGCCGAACTGTCTCACACTGTCTCACGTCAGGGTTAGACGACTATCCCCGACCACAACCAC
CACATCCGGGGGAGTTGCACGTAGATTTAATGTCTTGGGTCGGTGTTATGGTCAAGTCACTAATCAGCATTGGGTCCTT
GCTTGGCGCTACCGAGGATGTTGAGTTCTATACAAAAGTGCTCGACGCAATTGAACATAACCTGGATGATTTGCACTG
GTCAGAGAAGGAGGGCTGCTATTGTGACGCAACCATAGATGAGTTCGAGGAACATAAACTAGTTTGCCACAAGGGTT
ACATCTCCTTATTTCCTTTTTTGACTGGTCTATTAAAGCCGGATTCCCCCAAGCTGGGAAAGCTTCTAGCGCTCATAGGT
GACGAGAGTGAATTATGGTCTCCTTATGGTCTACGTAGCCTTTCGAAAAAAGATGAATTCTATGGGACAGCCGAAAATT
ACTGGCGTTCCCCAGTTTGGATAAACATCAATTATTTGGCCATCGTACAATTATACAACATAGCCACACAGGACGGGCC
TTATAAGGAAACGGCCCGAGATCTTTACACCAGACTGCGGAAGAATATCGTGGAGACGGTTTATCGTAACTGGGAGG
AAACTGGTTTCGCGTGGGAACAGTATAATCCCGAGACTGGCAAAGGACAGCGCACACAACACTTTACAGGTTGGACG
TCGCTTGTGGTAAAGATTATGAGCGGTCATCACCATCATCATCATTAG

7.1.3. MmER-αGluI GANAB (alpha subunit) cloned into pFastBacTM DUAL plasmid

GAATTCACCATGAAATTTTTGGTGAACGTGGCCTTGGTGTTTATGGTGGTTTACATTAGCTACATTTACTGCGGACCCG
GGCGCTGTGGATAGAAGCAACTTTAAGACCTGTGATGAGAGTTCCTTTTGCAAACGGCAGCGAAGCATTCGGCCAGG
CCTCTCTCCTTACCGTGCCTTGCTGGACACTCTGCAGCTTGGTCCTGATGCTCTTACAGTCCATCTGATCCATGAAGTCA
CCAAGGTGCTGCTTGTGCTGGAGCTCCAGGGCCTTCAGAAGAACATGACTCGGATCAGGATCGATGAGCTAGAGCCC
CGGCGGCCTCGATACCGAGTGCCAGATGTTTTAGTGGCTGACCCCCCCACAGCTAGGCTTTCAGTCTCTGGCCGTGAT
GACAACAGTGTGGAGCTAACAGTGGCTGAGGGACCCTACAAAATCATTTTGACAGCACAGCCATTCCGCCTTGACCT
GCTAGAAGATCGAAGCCTCCTGCTCAGTGTCAATGCCCGAGGACTTATGGCCTTTGAGCACCAGAGGGCCCCCAGGG
TCCCGCAAGAATCAAAAGACCCAGCTGAAGGCAATGGAGCCCAGCCTGAAGCAACACCTGGGGATGGTGACAAGCC
AGAGGAGACCCAGGAAAAGGCTGAGAAGGATGAGCCAGGAGCCTGGGAGGAGACATTCAAAACACATTCTGACAG
CAAGCCTTATGGCCCCACGTCTGTAGGTTTGGACTTTTCTCTGCCAGGAATGGAACATGTGTATGGGATCCCTGAGCAT
GCTGACAGCCTGAGACTGAAGGTCACTGAGGGCGGTGAGCCGTACCGCCTGTACAATTTGGATGTGTTCCAGTATGA
GCTGAACAACCCCATGGCTCTATATGGGTCTGTGCCTGTGCTCCTGGCACACAGCTTTCATCGAGACCTGGGCATCTTC
TGGCTTAATGCTGCTGAGACTTGGGTTGATATATCCTCCAACACGGCTGGGAAGACCCTGTTTGGGAAGATGCTTGAT
TACCTGCAGGGCTCTGGGGAGACTCCACAGACAGACATTCGTTGGATGTCAGAGAGTGGCATTATTGATGTTTTCCTA
ATGCTTGGCCCTTCGGTTTTTGATGTCTTTAGGCAGTATGCTAGTCTCACAGGGACCCAGGCATTGCCCCCACTCTTCT
CCCTCGGCTATCACCAGAGTCGCTGGAACTACCGGGATGAGGCTGATGTTTTGGAAGTGGATCAGGGTTTTGATGATC
ACAACATGCCTTGTGATGTCATTTGGTTGGACATTGAACATGCTGATGGCAAGCGGTACTTCACTTGGGACCCCACCC
GATTTCCTCAGCCCCTCAATATGCTTGAGCACTTGGCTTCCAAGAGGCGGAAGCTGGTGGCCATTGTGGACCCCCACA
TCAAGGTAGACTCTGGCTACCGAGTTCACGAAGAATTGCGAAACCATGGGCTGTATGTTAAAACTCGGGATGGCTCTG
ATTACGAGGGCTGGTGCTGGCCAGGCTCAGCTAGTTACCCTGACTTCACTAATCCAAGGATGAGGGCCTGGTGGTCTA
ACATGTTCAGCTTTGACAATTATGAGGGTTCAGCTCCTAATCTTTATGTTTGGAATGACATGAATGAACCGTCTGTGTTC
AATGGTCCTGAGGTCACCATGTTGAAGGATGCTGTGCATTATGGAGGCTGGGAGCACCGGGACATCCATAACATCTAT
GGCTTATATGTGCACATGGCGACTGCTGATGGGCTAATACAGCGCTCTGGGGGCATAGAGCGTCCCTTTGTCCTGAGT
AGGGCTTTCTTCTCAGGCTCCCAGCGCTTTGGAGCTGTGTGGACAGGGGACAACACTGCCGAATGGGATCATTTGAA
GATCTCTATCCCTATGTGTCTCAGCCTGGCACTGGTGGGGCTTTCCTTCTGTGGAGCGGATGTGGGTGGCTTCTTCAAG
AACCCAGAGCCAGAGCTGCTTGTGCGCTGGTACCAAATGGGTGCCTACCAGCCGTTCTTTCGGGCTCATGCCCACTTG
GACACTGGGCGGCGAGAGCCATGGCTGTTAGCGTCTCAATACCAAGATGCAATCCGAGATGCCTTGTTCCAGCGATAT
TCTTTGCTGCCCTTCTGGTATACCCTCTTCTATCAAGCTCACAAGGAAGGGTTTCCTGTCATGAGGCCCCTCTGGGTACA
GTATCCTGAGGATATGTCTACCTTCAGTATAGAGGATCAGTTCATGCTTGGTGATGCACTCCTTATTCACCCTGTATCGGA
TGCTGGGGCCCACGGAGTGCAGGTCTATTTGCCTGGCCAAGAAGAAGTGTGGTATGACATTCAGAGCTATCAGAAGC
ATCATGGGCCCCAGACCTTGTATCTGCCAGTAACTTTGAGCAGTATCCCCGTGTTCCAGCGTGGCGGAACCATTGTGCC
TCGATGGATGCGTGTGAGGCGCTCTTCAGACTGTATGAAGGACGATCCTATCACTCTCTTTGTTGCTCTCAGTCCCCAG
GGTACTGCCCAAGGAGAGCTCTTTCTAGATGATGGACACACCTTTAACTATCAGACTCGCCATGAGTTCCTGTTGCGGC
GGTTCTCTTTCTCTGGCAGCACACTGGTCTCTAGTTCAGCAGACCCCAAAGGCCACCTTGAGACACCTATTTGGATTG
AGCGAGTAGTCATCATGGGGGCTGGAAAGCCAGCAGCTGTGGTGCTCCAGACCAAAGGATCCCCTGAAAGTCGCCT
GTCCTTCCAGCATGACCCTGAGACCTCAGTGCTGATATTGCGTAAACCTGGTGTCAGCGTGGCATCCGACTGGAGTATT
CATCTTCGAGCTAGCTGGAGCCACCCGCAGTTCGAAAAACATCACCATCACCATCACTAAGCTT

7.1.4. MmER-αGluI PRKSCH (Beta subunit) cloned into pFastBacTM DUAL plasmid

CTCGAGCACCATGAAATTTTTGGTGAACGTGGCCTTGGTGTTTATGGTGGTTTACATTAGCTACATTTATGCGGACCCG
GGCGCTGTAGAAGTTAAGAGACCCCGGGGCGTTTCCCTCAGCAACCATCACTTCTATGAAGAATCTAAACCTTTCACC
TGTTTGGACGGCACAGCCACCATCCCATTCGATCAGGTGAACGACGACTACTGCGACTGTAAGGACGGTTCAGATGA
GCCTGGCACAGCTGCTTGTCCCAATGGCAGCTTTCACTGCACCAACACTGGGTACAAGCCCTTGTACATCCTCTCCAGC
CGGGTCAATGATGGGGTATGTGACTGCTGTGATGGCACAGACGAGTACAACAGCGGCACGGTCTGCGAGAACACCTG
CAGAGAGAAGGGTCGCAAGGAAAAAGAGTCCCTGCAGCAGCTGGCGGAAGTCACCCGTGAAGGGTTCCGCCTGAA
GAAGATTCTCATTGAGGAGTGGAAGACAGCCCGGGAAGAAAAGCAGAGTAAGCTTCTTGAGCTTCAGGCTGGAAA
GAAGTCTCTGGAAGACCAGGTAGAAACACTGCGGGCAGCGAAAGAAGAAGCAGAGAGGCCAGAGAAGGAGGCCA

AGGACCAGCACCGGAAGCTGTGGGAAGAGCAGCAAGCTGCTGCCAAGGCCCGGCGGGAACAGGAGCGGGCAGCC
AGTGCCTTCCAGGAACTTGACGACAACATGGATGGGATGGTCTCGCTGGCTGAGTTACAGACTCACCCGGAGCTGGA
CACAGATGGAGATGGAGCGCTGTCTGAGGAGGAGGCCCAGGCCCTTCTCAGTGGAGACACACAGACTGACACCACC
TCCTTCTATGACCGTGTCTGGGCTGCCATCAGGGACAAGTACCGCTCTGAGGTCCCGCCCACTGACATACCTGTTCCGG
AGGAGACTGAGCCCAAAGAGGAAAAGCCACCAGTGTTGCCACCCACAGAGGAGGAGGAAGAGGAGGAGGAGGA
GCCAGAAGAAGAGGAGGAGGAAGAGGAAGAGGAGGAGGAGGCTCCGCCCCCACTGCAGCCCCCACAGCCTCCCA
GCCCCACAGAGGATGAGAAGATGCCGCCCTATGATGAGGAGACCCAGGCCATCATCGATGCTGCACAGGAGGCCCG
GAGTAAGTTTGAGGAAGTCGAACGGTCCTTGAAAGAGATGGAAGAGTCCATCAGGAGTTTGGAACAAGAGATCTCC
TTTGATTTCGGTCCCTCTGGAGAGTTTGCATATCTCTACAGCCAATGCTACGAGCTCACCACCAATGAGTACGTCTACCG
GCTTTGCCCCTTCAAACTGGTCTCCCAGAAACCCAAACATGGGGGCTCCCCGACCAGCCTGGGCACATGGGGCTCCT
GGGCTGGCCCTGATCATGACAAGTTCAGTGCGATGAAGTACGAGCAGGGCACGGGCTGTTGGCAGGGCCCCAACCG
ATCCACCACAGTGCGCCTGCTGTGTGGCAAAGAGACTGTGGTGACCAGCACCACGGAGCCCAGTCGCTGTGAGTACC
TCATGGAGCTGATGACACCAGCAGCCTGCCCAGAGCCGCCACCAGAAGCACCCAGTGATGGGGACAAACATCACCAT
CACCATCACTAAGGTACC

7.1.5. Human NEU2 sequence cloned into pET28

CCATGGGTAGCAGCCATCACCACCACCACCACGGCACCGCGGAGAACCTGTACTTCCAAGGTAGCATGGCGAGCCTG
CCGGTGCTGCAAAAGGAAAGCGTTTTTCAGAGCGGTGCGCATGCGTACCGTATCCCGGCGCTGCTGTATCTGCCGGG
TCAGCAAAGCCTGCTGGCGTTTGCGGAGCAGCGTGCGAGCAAGAAAGACGAGCACGCGGAACTGATTGTGCTGCGT
CGTGGCGACTACGATGCGCCGACCCACCAAGTTCAGTGGCAAGCGCAGGAAGTGGTTGCGCAAGCGCGTCTGGACG
GTCACCGTAGCATGAACCCGTGCCCGCTGTATGATGCGCAGACCGGCACCCTGTTTCTGTTCTTTATCGCGATTCCGGG
TCAAGTTACCGAGCAGCAACAGCTGCAGACCCGTGCGAACGTGACCCGTCTGTGCCAGGTTACCAGCACCGACCACG
GCCGTACCTGGAGCAGCCCGCGTGATCTGACCGATGCGGCGATTGGTCCGGCGTATCGTGAATGGAGCACCTTTGCG
GTGGGTCCGGGTCACTGCCTGCAACTGAACGATCGTGCGCGTAGCCTGGTGGTTCCGGCGTATGCGTATCGTAAACTG
CACCCGATCCAGCGTCCGATTCCGAGCGCGTTCTGCTTTCTGAGCCATGACCACGGTCGTACCTGGGCGCGTGGCCAC
TTTGTGGCGCAAGATACCCTGGAGTGCCAGGTGGCGGAAGTGGAAACCGGCGAACAGCGTGTGGTTACCCTGAACG
CGCGTAGCCACCTGCGTGCGCGTGTTCAAGCGCAGAGCACCAACGACGGTCTGGATTTCCAAGAGAGCCAGCTGGT
GAAGAAACTGGTGGAGCCGCCGCCGCAGGGTTGCCAGGGCAGCGTGATCAGCTTTCCGAGCCCGCGTAGCGGTCC
GGGCAGCCCGGCGCAATGGCTGCTGTACACCCATCCGACCCACAGCTGGCAGCGTGCGGACCTGGGTGCGTATCTGA
ACCCGCGTCCGCCGGCGCCGGAGGCGTGGAGCGAACCGGTTCTGCTGGCGAAGGGTAGCTGCGCGTACAGCGACC
TGCAGAGCATGGGTACCGGTCCGGATGGTAGCCCGCTGTTTGGTTGCCTGTACGAGGCGAACGATTATGAGGAAATT
GTGTTCCTGATGTTTACCCTGAAACAAGCGTTTCCGGCGGAATATCTGCCGCAGTAAGGATCC

7.1.6. Viral neuraminidase open construct cloned into pFastBacTM

ATGAAATTTTTGGTTAATGTGGCTTTAGTGTTCATGGTCGTCTACATTTCTTATATTTACGCTGACCCAGGACACCAC
CACCACCATCATTCTAGTTCAGATTACTCGGATTTACAGAGGGTTAAACAAGAGCTGCTCGAAGAAGTCAAGAAAG
AACTTCAAAAGGTGAAGGAGGAAATCATTGAAGCCTTCGTTCAAGAGCTTCGTAAGAGAGGAGGGACGGCGGAG
AACCTATACTTCCAAGGTGTAAAATTGGCCGGTAATTCTTCGCTATGCCCCGTGAGCGGTTGGGCAATTTATTCAAA
GGATAACTCGGTACGGATCGGGTCAAAAGGTGACGTATTCGTTATTCGGGAGCCGTTTATATCGTGCAGCCCTCTC
GAATGTAGAACCTTTTTCCTTACGCAAGGTGCACTGCTGAATGACAAACACAGCAACGGTACTATCAAAGATCGTAG
CCCTTACCGCACACTAATGAGTTGCCCGATAGGCGAGGTCCCATCCCCTTACAATAGCCGATTCGAGAGTGTTGCCT
GGAGCGCCTCGGCGTGTCATGATGGAATAAATTGGCTCACAATCGGGATCTCAGGTCCAGACAATGGCGCTGTAGC
AGTTTTGAAATACAACGGTATAATAACCGATACAATCAAGAGTTGGCGAAATAACATATTACGTACTCAGGAATCCG
AGTGTGCATGCGTAAACGGCAGTTGTTTCACTGTTATGACTGACGGTCCCTCGAATGGCCAGGCATCATATAAGATC
TTCCGAATCGAGAAAGGCAAGATAGTAAAATCAGTTGAAATGAACGCGCCAAACTATCACTATGAGGAATGTAGTT
GTTACCCAGACTCATCCGAGATTACGTGTGTGTGCCGGGACAATTGGCATGGGTCTAACCGTCCTTGGGTCTCGTTT
AATCAGAACTTGGAATATCAGATAGGATACATCTGCAGCGGAATATTTGGGGATAACCCGAGGCCCAATGATAAGA
CGGGCAGTTGTGGCCCCGTGTCTTCTAACGGAGCTAATGGTGTCAAGGGCTTCAGCTTTAAGTATGGAAACGGTGT
CTGGATTGGCAGAACAAAGTCCATTTCTTCGAGGAACGGGTTTGAAATGATATGGGACCCTAACGGATGGACAGG

AACTGACAATAATTTCTCTATCAAGCAGGACATTGTCGGGATTAACGAATGGTCCGGGTATAGTGGGTCTTTTGTGC
AACATCCTGAGCTGACCGGGCTAGATTGTATTCGCCCGTGCTTCTGGGTAGAGCTCATACGCGGCAGACCGAAAGA
GAATACCATCTGGACGTCCGGTTCTTCTATCAGCTTTTGCGGTGTAAATTCAGATACCGTGGGCTGGTCCTGGCCGG
ACGGAGCGGAACTTCCTTTTACGATAGACAAATGA

7.1.7. Viral neuraminidase closed construct cloned into pFastBacTM

ATGAAATTTTTGGTTAATGTGGCTTTAGTGTTCATGGTCGTCTACATTTCTTATATTTACGCTGACCCAGGACACCAC
CACCACCATCATTCTAGTTCAGATTACTCGGATTTACAGAGGGTTAAACAAGAGCTGCTCGAAGAAGTCAAGAAAG
AACTTCAAAAGGTGAAGGAGGAAATCATTGAAGCCTTCGTTCAAGAGCTTCGTAAGAGAGGAGGGACGGCGGAG
AACCTATACTTCCAAGGTGTAAAATTGGCCGGTAATTCTTCGCTATGCCCCGTGAGCGGTTGGGCAATTTATTCAAA
GGATAACTCGGTACGGATCGGGTCAAAAGGTGACGTATTCGTTATTCGGGAGCCGTTTATATCGTGCAGCCCTCTC
GAATGTAGAACCTTTTTCCTTACGCAAGGTGCACTGCTGAATGACAAACACAGCAACGGTACTATCAAAGATCGTAG
CCCTTACCGCACACTAATGAGTTGCCCGATAGGCGAGGTCCCATCCCCTTACAATAGCCGATTCGAGAGTGTTGCCT
GGAGCGCCTCGGCGTGTCATGATGGAATAAATTGGCTCACAATCGGGATCTCAGGTCCAGACAATGGCGCTGTAGC
AGTTTTGAAATACAACGGTATAATAACCGATACAATCAAGAGTTGGCGAAATAACATATTACGTACTCAGGAATCCG
AGTGTGCATGCGTAAACGGCAGTTGTTTCACTGTTATGACTGACGGTCCCTCGAATGGCCAGGCATCATATAAGATC
TTCCGAATCGAGAAAGGCAAGATAGTAAAATCAGTTGAAATGAACGCGCCAAACTATCACTATGAGGAATGTAGTT
GTTACCCAGACTCATCCGAGATTACGTGTGTGTGCCGGGACAATTGGCATGGGTCTAACCGTCCTTGGGTCTCGTTT
AATCAGAACTTGGAATATCAGATAGGATACATCTGCAGCGGAATATTTGGGGATAACCCGAGGCCCAATGATAAGA
CGGGCAGTTGTGGCCCCGTGTCTTCTAACGGAGCTAATGGTGTCAAGGGCTTCAGCTTTAAGTATGGAAACGGTGT
CTGGATTGGCAGAACAAAGTCCATTTCTTCGAGGAACGGGTTTGAAATGATATGGGACCCTAACGGATGGACAGG
AACTGACAATAATTTCTCTATCAAGCAGGACATTGTCGGGATTAACGAATGGTCCGGGTATAGTGGGTCTTTTGTGC
AACATCCTGAGCTGACCGGGCTAGATTGTATTCGCCCGTGCTTCTGGGTAGAGCTCATACGCGGCAGACCGAAAGA
GAATACCATCTGGACGTCCGGTTCTTCTATCAGCTTTTGCGGTGTAAATTCAGATACCGTGGGCTGGTCCTGGCCGG
ACGGAGCGGAACTTCCTTTTACGATAGACAAATGA

7.2. X-ray Crystallography Data Collection and Refinement Statistics

	HPSE-**1**	AcGH79-**1**	BpHep-**1**	HPSE- **27**	NEU2
Data collection					
Space group	$P2_1$	$I2_1$	$P2_1$	$P2_1$	H 3
Cell dimensions					
a, *b*, *c* (Å)	46.8, 71.5, 79.3	82.9, 44.7, 137.02	76.5, 105.6, 114.3	46.3, 71.5, 79.3	146.22, 146.22, 64.95
β (°)	95.3	97.46		95.7	90
Resolution (Å)	46.57-2.16 (2.16-2.10)	41.15-2.05 (2.05-2.00)	77.59-1.5 (1.53-1.50)	53.0-2.49 (2.49-2.40)	38.53-1.56 (1.56-1.53)
R_{merge}	0.127 (2.82)	0.055 (0.168)	0.14 (2.83)	0.115 (1.50)	0.058 (2.36)
R_{pim}	0.06 (1.2)	0.02 (0.07)	0.04 (0.79)	0.06 (0.69)	0.02 (0.8)
I / σI	10.8 (0.8)	25.1 (7.5)	9.6 (0.9)	8.0 (1.1)	16.4 (0.9)
CC1/2	0.997 (0.42)	0.99 (0.99)	0.99 (0.48)	0.90 (0.41)	1 (0.64)
Completeness (%)	99.7 (100)	99.8 (100)	99.7 (99)	100 (99.9)	100 (100)
Redundancy	6.4 (6.6)	6.4 (6.7)	13.2(13.5)	5.4	10.2 (10.3)
Refinement					
Resolution (Å)	46.62-2.10	41.18-2.00	63.65-1.50	41.68-2.4	38.56-1.53
No. reflections	30423	33995	147634	20299	78121
R_{work} / R_{free}	0.20/0.25	0.19/0.24	0.18/0.21	0.20/0.25	0.19/0.21
No. atoms					
Protein	3600	3487	6282	3616	5669
Ligand/ion	97/3	17/0		112/2	
Water	135	441	787	10	239
B-factors					
Protein	45.35	19.14	19.86	55.04	38.2
Ligand/ion	64.22/54.57	24.78/0		89.37/99.59	
Water	43.06	28.35	29.78	47.59	-
R.m.s. deviations					
Bond lengths (Å)	0.016	0.014	0.015	0.006	-
Bond angles (°)	2.00	1.88	1.82	1.64	-
Ramachandran					
Favored	96.7	96.93	97.24	97.36	93.29
Outliers	0.00	1.45	0	0.22	1.46
Wavelength	0.98	0.98	0.98	0.98	0.98

7.3. Appendix: Publications

7.3.1. Molecular Basis for Inhibition of Heparanases and β-Glucuronidases by Siastatin B

JACS JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
pubs.acs.org/JACS
Open Access
Article

Molecular Basis for Inhibition of Heparanases and β-Glucuronidases by Siastatin B

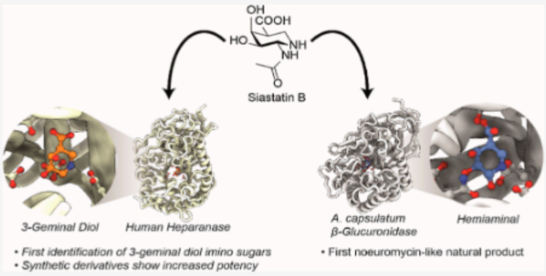
Yurong Chen,[§] Adrianus M. C. H. van den Nieuwendijk,[§] Liang Wu,[§] Elisha Moran,
Foteini Skoulikopoulou, Vera van Riet, Hermen S. Overkleeft, Gideon J. Davies,*
and Zachary Armstrong*

Cite This: *J. Am. Chem. Soc.* 2024, 146, 125−133

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Siastatin B is a potent and effective iminosugar inhibitor of three diverse glycosidase classes, namely, sialidases, β-D-glucuronidases, and *N*-acetyl-glucosaminidases. The mode of inhibition of glucuronidases, in contrast to sialidases, has long been enigmatic as siastatin B appears too bulky and incorrectly substituted to be accommodated within a β-D-glucuronidase active site pocket. Herein, we show through crystallographic analysis of protein-inhibitor complexes that siastatin B generates both a hemiaminal and a 3-geminal diol iminosugar (3-GDI) that are, rather than the parent compound, directly responsible for enzyme inhibition. The hemiaminal product is the first observation of a natural product that belongs to the noeuromycin class of inhibitors. Additionally, the 3-GDI represents a new and potent class of the iminosugar glycosidase inhibitor. To substantiate our findings, we synthesized both the *gluco*- and *galacto*-configured 3-GDIs and characterized their binding both structurally and kinetically to exo-β-D-glucuronidases and the anticancer target human heparanase. This revealed submicromolar inhibition of exo-β-D-glucuronidases and an unprecedented binding mode by this new class of inhibitor. Our results reveal the mechanism by which siastatin B acts as a broad-spectrum glycosidase inhibitor, identify a new class of glycosidase inhibitor, and suggest new functionalities that can be incorporated into future generations of glycosidase inhibitors.



■ INTRODUCTION

Iminosugars are carbohydrate mimetics that contain an endocyclic nitrogen in place of an endocyclic oxygen. This class of compounds can be both potent and selective inhibitors of carbohydrate processing enzymes.[1−3] As a result of their potency, the iminosugars migalastat and miglustat are used clinically to treat lysosomal storage disorders[4,5] while miglitol are used to treat diabetes.[6] Furthermore, iminosugars have been investigated as potential therapeutics for tumor metastasis,[7,8] cystic fibrosis,[9] and as antiviral therapeutics.[10,11]

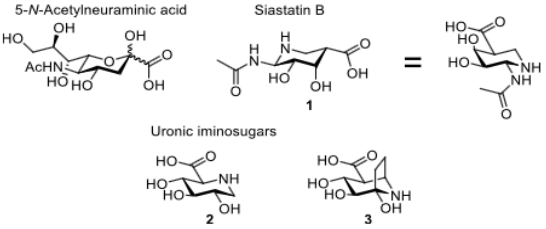


Figure 1. Structures of 5-*N*-acetylneuraminic acid, siastatin B (**1**), and the uronic acid-configured iminosugars uronic-deoxynojirimycin (**2**) and uronic-noeurostegine (**3**).

Siastatin B (**1**) is a natural product 1-*N*-iminosugar, first identified in 1974 from a *Streptomyces verticillus* culture by Umezewa and co-workers.[12] This iminosugar is a geminal diamine, in which the anomeric carbon is replaced by a nitrogen, and the 2-position has an *N*-acetyl functional group (see Figure 1). Siastatin B effectively inhibits a wide range of sialidases from viral,[13] bacterial,[14] and human origin[15] with micromolar potency as it structurally resembles 5*N*-acetylneuraminic acid. Surprisingly, siastatin B also inhibits both β-glucuronidases and human heparanase,[12,16] which cleave β-glucuronic acid residues but not sialic acid or *N*-acetylated sugars. Although siastatin B resembles uronic acid-configured iminosugars [such as the β-glucuronidase inhibitors uronic-deoxynojirimycin (**2**),[17,18] uronic-noeurostegine (**3**),[19] and uronic-isofagamine[20,21]], it bears two major modifications that

Received: April 21, 2023
Revised: December 4, 2023
Accepted: December 5, 2023
Published: December 20, 2023

https://doi.org/10.1021/jacs.3c04162
J. Am. Chem. Soc. 2024, 146, 125−133

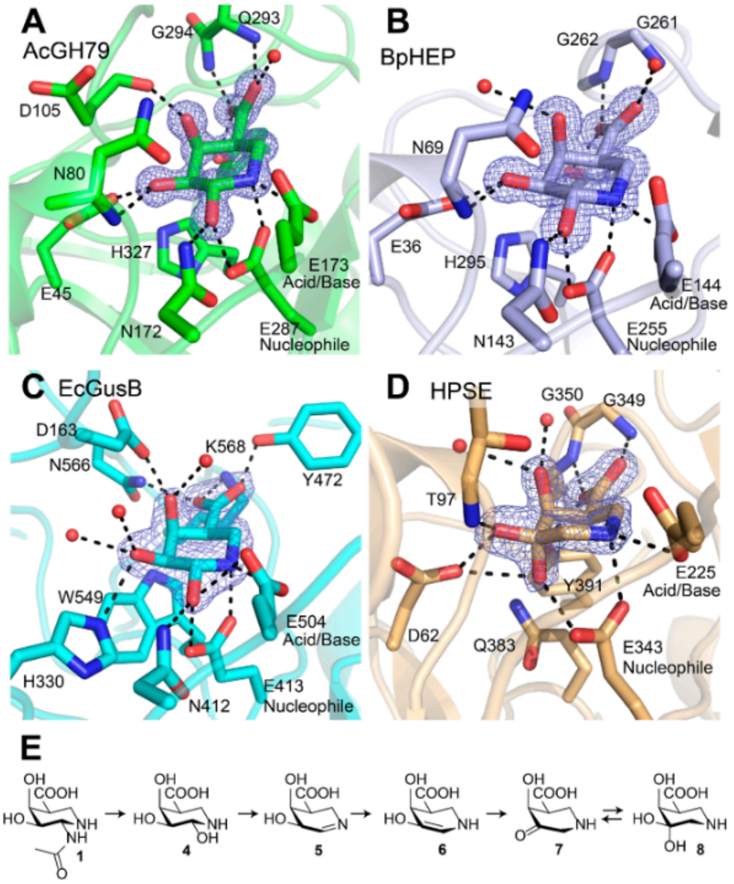


Figure 2. Structures of heparanases and β-glucuronidases soaked with siastatin B and the proposed breakdown mechanism. Siastatin breakdown products are shown to bound to (A) AcGH79, (B) BpHEP, (C) EcGusB, and (D) HPSE. Electron density ($2F_o - F_c$) is shown for the ligand as a blue mesh contoured at 2 σ (AcGH79 = 0.92 $e^-/Å^3$, BpHep = 0.77 $e^-/Å^3$, EcGusB = 0.41 $e^-/Å^3$, and HPSE = 0.61 $e^-/Å^3$). The polypeptide is shown in the cartoon form, with active site residues shown as sticks. Apparent hydrogen bonding interactions are shown as dotted black lines. (E) Proposed degradation of the siastatin B **1** into galacturonic-noeuromycin **4** and the 3-geminial diol iminosugar **8**.

one would expect to interfere with β-glucuronidase activity. The first of these is the presence of the acetamido group at the 2-position. Glycosidases are typically extremely specific for the substituents at the 2-position of their substrates[22−24] as interactions with these substituents enable the stabilization of a planar C5−O−C1−C2 oxocarbenium ion transition state.[25,26] Furthermore, investigation of the active site of human heparanase in complex with heparan oligomers[27,28] and inhibitors[29] reveals that there is little space within the active site to accommodate an *N*-acetyl group at the 2-position. The second stereochemical problem that may interfere with the binding of siastatin B to β-glucuronidases is the "*galacto*" configuration of the 4-position. In siastatin B, this hydroxyl is axial (in the $^4C_1$ conformation), while in the natural substrates of these enzymes−glucuronides−this position is "*gluco*", that is, equatorial. Together, these modifications raise the possibility that the inhibition of β-glucuronidases by siastatin B may occur in an unanticipated fashion.

One hint toward the inhibition caused by siastatin B comes from the synthesis of a series of siastatin analogues reported by Nishimura and co-authors that were modified with a trifluoroacetamido substituent at the 2-position.[30,31] These 2-trifluoroacetamido are potent inhibitors of bovine liver β-D-glucuronidase[30,31] and micromolar inhibitors of recombinant human heparanase (HPSE).[31] NMR analyses of 2-trifluoroacetamido siastatin in the media of enzyme assays have suggested that it can undergo pH-dependent decomposition and rearrangement in solution, yielding compounds that could act as the true enzyme inhibitors.[32] However, this solvent-mediated rearrangement has not been demonstrated for siastatin B, which has always been regarded as stable in aqueous solution.[33]

Our interest in heparanase inhibitors[29,34] has led us to re-examine the inhibition of heparanases and β-glucuronidases by siastatin B. Human heparanase (HPSE) is an *endo*-β-glucuronidase that catalyzes the cleavage of heparan sulfate (HS). Overexpression of HPSE strongly drives the growth of aggressive metastatic cancers[35] and leads to excessive HS degradation within the extracellular matrix, thereby facilitating cancer cell migration,[36,37] while growth factors and cytokines liberated upon HS degradation stimulate proliferation and angiogenesis.[38,39] Inhibitors of HPSE are potential anticancer therapeutics and have shown antimetastatic activity in animal models.[29,40,41] It is our hope that by understanding how glycosidases are inhibited by siastatin B, we can identify and design potent inhibitors for this enzyme. Here, we used structural biology to investigate the mode of binding of siastatin B to β-glucuronidases. This unexpectedly revealed that contaminating breakdown products are instead responsible for the inhibition of β-glucuronidases. In particular, we identify both the first natural product noeuromycin-type inhibitor and a new class of potent glycosidase inhibitor—the 3-geminal-diol iminosugars (3-GDIs). To substantiate our findings, we have synthesized both the *gluco-* and *galacto-*configured 3-GDIs and characterized their binding to HPSE and β-glucuronidases through detailed enzyme kinetics, competitive activity-based protein profiling, and structural biology.

■ RESULTS AND DISCUSSION

To determine how siastatin B inhibits β-glucuronidases, we first sought to determine whether inhibition was due to the presence of contaminants from either the commercial material or degradation products resulting from prolonged incubation in an aqueous solution or in the presence of a β-glucuronidase. NMR analysis of commercial siastatin B showed no contaminants present, and no contaminating peaks appeared after incubation at pH 5.0 for 18 h, see Figure S1. We next determined whether siastatin B degradation could be caused by prolonged incubation with a β-glucuronidase. ^{1}H NMR analysis of siastatin B incubation with recombinant human heparanase (present at concentration of 1.2 μM) showed very minor additional peaks in the 3.2−3.0 range, with no loss in the starting peaks observed.

We next performed high-resolution mass spectrometry to determine whether there were any minor contaminants present in our siastatin B sample (Figure S2). In addition to an m/z corresponding to siastatin B (219 in the positive-ion mode, 217 in the negative-ion mode), we also observed a peak at m/z = 160 (in the positive-ion mode, and 158 in the negative-ion mode) that could belong to the siastatin break down products **5**, **6**, or **7** (Figure 2). Additionally, we observed a peak at m/z = 178 in the positive-ion mode that could belong to a noeuromycin like breakdown product (**4**) or a 3-geminal-diol (**8**). These minor species could be separated and identified by LC−MS (Figure S2). We next purified the major peak using preparative HPLC−MS, although we were unable to completely remove all contaminants (see Figure S3), the resulting material had substantially reduced concentrations of the m/z = 158 and m/z = 176 peaks. This HPLC-purified material was used in inhibition assays of *E. coli* β-glucuronidase (EcGusB) and *Acidobacterium capsulatum* β-glucuronidase (AcGH79), see Figure S3. Both enzymes were inhibited less by the purified material, indicating that siastatin B contaminants are indeed inhibitors of β-glucuronidases. As the contaminating substances could be one of several isomers, uncertainty in the active species remained.

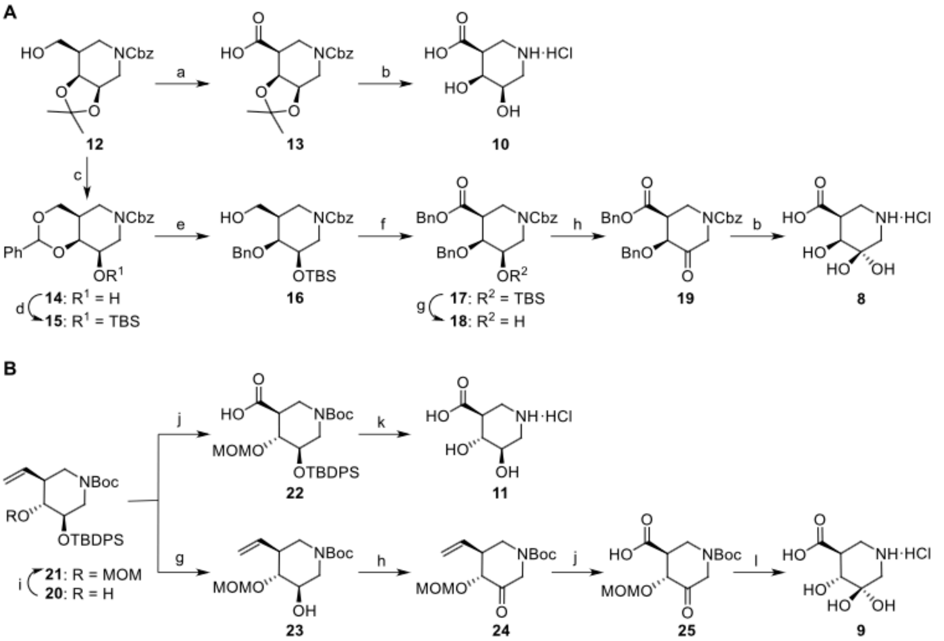
Structural Basis for Siastatin B Inhibition of Heparanases and β-Glucuronidases. To determine how siastatin B inhibits both heparanases and β-glucuronidases, we determined crystal structures of co-complexes between siastatin B and each of AcGH79, EcGusB, human heparanase (HPSE), and *Burkholderia pseudomallei* heparanase (BpHEP), see Table S1 for data collection and refinement statistics. As we showed by NMR that siastatin B is stable under enzymatic assay conditions, we expected to observe it present in the active site of the enzymes. However, we were surprised to find that although each of the structures contains a ligand present with full occupancy in the active site, none of the observed ligands are siastatin B (Figure 2). Instead, siastatin B breakdown products are present in all structures.

Remarkably, not all active sites contained the same inhibitor. Three of the four enzymes (AcGH79, BpHEP, and EcGusB) contained "galacturonic-noeuromycin" (**4**) bound in the −1 position of the active site, as opposed to siastatin B (Figure 2). We confirmed the presence of an alcohol at the 2-position—as opposed to an amine, which would result from deacetylation—through examination of the *B*-factors within the high-resolution crystal structures of AcGH79 and BpHEP, see Figure S4. When the 2-position is modeled as an alcohol, the *B*-factor of the 2-substituent is consistent with the 2-carbon. However, when the 2-position is modeled as an amine, the *B*-factor of this functional group is depressed, which is anticipated for a nitrogen being accommodated within the electron density of an oxygen.

We expect that the *galacturonic*-noeuromycin (**4**), observed in these active sites is formed through the elimination of the *N*-acetyl group of siastatin B, resulting in imine **5** that is subsequently hydrated to form hemiaminal **4**, see Figure 2E. The smaller functional group at the 2-position is both well accommodated within the active sites of these enzymes and the correct functional group (OH instead of NH-acetyl). The 2-hydroxyl is within hydrogen bonding distance of the catalytic nucleophile and an active site asparagine in all three of these structures, thereby mimicking the interactions seen for natural substrates[28] and other inhibitors.[20,29] The 4-hydroxyl of **4** has *galacto*-stereochemistry in contrast to the *gluco*-configuration of the natural enzyme substrate; however, this position also appears to have productive binding to the active site. The 4-hydroxyl of **4** forms hydrogen bond interactions with the enzyme active site, for both AcGH79 and EcGusB, see Figure 2AC. The 4-hydroxyl of **4** complexed to BpHEP does not form a hydrogen bond with the amino acids in the enzyme active site; however it forms a hydrogen bond with an active site water, see Figure 2B.

Unlike the three other co-complex structures, the structure of HPSE crystals soaked with siastatin B contained a 3-geminial diol iminosugar (**8**). Compound **8** is a further breakdown product of **4**, formed by the elimination of the 2-hydroxyl to form imine **5** which rearranges to enamine **6**, and the enol can then tautomerize to give ketoamine **7**. This ketoamine can then become hydrated at the 3-position to give the observed geminal diol **8**. A similar breakdown of noeuromycin and fuco-noeuromycin, which are unstable at neutral pH, has been previously reported and also resulted in iminosugars bearing a hydrated ketone.[42] Despite the additional hydroxyl present at the 3-position of **8**, when compared to glucuronic acid, this compound appears to be well accommodated in the active site of HPSE. Indeed, both hydroxyls present at the 3-position form hydrogen bond interactions with the active site. As well, the hydroxyl on the β-face forms a hydrogen bond with Asp62 and the backbone nitrogen of Thr97, while the hydroxyl on the α-face is positioned to form hydrogen bonds with the catalytic nucleophile Glu343 and Asp62. The hydroxyl at the 4-position of inhibitor **3** lacks the hydrogen bond with Trp391 which is seen for *gluco*-configured inhibitors.[20,29] However, as in the costructure of **4** and BpHEP, the axial 4-hydroxyl forms a hydrogen bond with water present in the crystal structure.

Scheme 1. Synthesis of D-Galacturonic Acid-Type 1-*N*-Iminosugars 8 and 10 (A) and D-Glucuronic Acid-Type 1-*N*-Iminosugars 9 and 11 (B). Reagents and Conditions: (a) (i) Dess−Martin Periodinane, DCM, rt; (ii) $NaClO_2$, NaH_2PO_4, 30% H_2O_2, CH_3CN, H_2O, 0 °C to rt, 65% over Two Steps; (b) H_2, 10% Pd/C, H_3O^+, THF, rt, 10 Quant, 8 Quant; (c) (i) 8 M HCl, MeOH, rt; (ii) $PhCH(OMe)_2$, CSA, DMF, 60 °C, 72% Over Two Steps; (d) TBSCl, Imidazole, DCM, rt, 94%; (e) $THF{\cdot}BH_3$, TMSOTf, DCM, 0 °C to rt, 97%; (f) (i) Jones Reagent, Acetone, rt; (ii) Benzyl Alcohol, DIC, DMAP, DCM, rt, 46% Over Two Steps; (g) TBAF, THF, 18 0 °C, 94%, 23 rt, 94%; (h) Dess−Martin Periodinane, DCM, rt, 19 71%, 24 93%; (i) MOMCl, DIPEA, DCM, Reflux, 87%; (j) $RuCl_3{\cdot}3H_2O$, $NaIO_4$, $CCl_4/CH_3CN/H_2O$, rt, 22 64%, 25 83%; (k) 3 M HCl in H_2O, dioxane, 100 °C, 95%; and (l) 4 M HCl in dioxane/H_2O, rt, 70%



Synthesis of Siastatin B Breakdown Products. As 3-GDI inhibitors represent a new class of glycosidase inhibitors, we sought to chemically synthesize them to better understand their binding and potency. We synthesized the geminal diol siastatin B breakdown product observed in cocrystal structure with HPSE (**8**) and its *gluco*-configured 4-epimer **9**, to probe whether this compound would be more potent. We also synthesized both galacturonic acid isofagomine **10** and glucuronic acid isofagomine **11** to gain further insights into the influence of the geminal diol at the 3-position on inhibition.

Primary alcohol **12** (see Supporting Information and Scheme S1 for its synthesis) was oxidized to carboxylic acid **13** via a two-step oxidation process (Scheme 1A). Deprotection of the isopropylidene acetal and the carboxybenzyl (Cbz) group in **13** was achieved by palladium-catalyzed hydrogenolysis in the presence of acid, affording galacturonic acid isofagomine **10** in the quantitative yield. The synthesis of geminal diol **8** proceeded through the same intermediate (**12**). The 3,4-isopropylidene acetal in **12** was removed under acidic conditions followed by a thermodynamically controlled installation of the benzylidene acetal and subsequent silylation of the remaining 3-hydroxyl group, affording fully protected **15**. Regioselective cleavage of the benzylidene acetal yielded compound **16**, the primary alcohol of which was then oxidized to carboxylic acid followed by benzyl protection to afford benzyl ester **17**. Desilylation of **17** and subsequent oxidation of the resulting secondary alcohol gave ketone **19**, which after global deprotection by catalytic hydrogenation under acidic conditions finally afforded target compound **8** (see the Supporting Information and Scheme S2 for an alternative synthetic route toward **8** via a *tert*-butyl ester intermediate).

The synthesis of *gluco*-configured iminosugars **9** and **11** commenced from common intermediate **20** (see Supporting Information and Scheme S3 for its synthesis), which was treated with methoxymethyl chloride in the presence of excessive base under heating to afford compound **21** (Scheme 1B). Oxidative cleavage of the terminal alkene using ruthenium tetroxide (generated in situ from ruthenium chloride and sodium periodate) gave carboxylic acid **22**. The acid labile protecting groups in **22** were removed by treatment with aqueous 3 M HCl to afford glucuronic acid isofagomine **11** in an excellent yield. All spectroscopic data obtained for **11** proved to be in agreement with those reported in the literature.[43−45] On the other hand, the silyl group in **21** could be readily removed by tetrabutylammonium fluoride (TBAF) to afford alcohol **23**, which was then oxidized to ketone **24** with Dess−Martin periodinane. Oxidative cleavage of the terminal alkene in **24** followed by global deprotection under acidic conditions finally yielded the target geminal diol **9**.

3-GDIs Inhibit Glucuronidases. To assess the potency of the synthesized 1-*N*-iminosugars, we first turned to the exo-acting *β*-glucuronidases AcGH79 and EcGusB, whose activity can be readily monitored with the use of the fluorogenic substrate 4-methylumbelliferyl glucuronide (MU-GlcUA). As these β-glucuronidases belong to different families—AcGH79 to the GH79 family and EcGusB to the GH2 family[46]—assessing inhibition with these enzymes also gives some insights into the applicability of these molecules. All four synthesized compounds (**8**−**11**) are able to inhibit both enzymes (see Table 1, Figures S5 and S6). The *galacto*-configured 3-GDI **8**—the breakdown product observed in the active site of HPSE—inhibited AcGH79 (K_i = 5.8 ± 0.5 μM) more potently than did EcGusB (K_i = 137 ± 3 μM). Comparison of these inhibition constants to those obtained with the *galacturonic*-isofagomine derivative (**10**) reveals that both enzymes are more strongly inhibited by *galacturonic*-isofagomine **10**, with $\Delta\Delta G$ values for K_d of more than −3 kJ·mol^{-1} for both enzymes.

Table 1. Inhibition Constants for the Synthesized Inhibitors

inhibitor	AcGH79 K_i (μM)	EcGusB K_i (μM)
8	5.8 ± 0.5	137 ± 3
9	0.520 ± 0.030	28 ± 1
10	1.5 ± 0.4	23 ± 3
11	0.022 ± 0.003	1.5 ± 0.3

We also synthesized the *gluco*-configured 1-*N*-iminosugars **9** and **11** in the hopes that by changing the 4-position stereochemistry to that of the natural substrate, we could enhance inhibition. This, indeed, turned out to be the case. The *gluco*-configured 3-GDI (**9**) was significantly more potent than the *galacto*-configured (**8**), with inhibition constants of 520 ± 30 nM and 28 ± 1 μM for AcGH79 and EcGusB, respectively. The *gluco*-configured isofagomine (**11**) proved even more potent, mirroring the trend seen with **10**, resulting in the potent inhibition of AcGH79 and EcGusB that was more than an order of magnitude better than the geminal-diol, see Tables 1 and 2.

Table 2. Free Energy Changes Resulting from 4-Epimerization and Conversion of a *gem*-Diol to an Alcohol

inhibitor	AcGH79 ΔG, K_d (kJ mol^{-1})	EcGusB ΔG, K_d (kJ mol^{-1})
	4_{ax}-OH → 4_{eq}-OH	
3-OH	−10.3	−6.7
3-*gem*-diol	−5.9	−3.9
	3-*gem*-diol → 3-OH	
4_{ax}-OH	−3.3	−4.3
4_{eq}-OH	−7.7	−7.1

We next examined whether the synthetic 1-*N*-iminosugars could inhibit HPSE and human β-glucuronidase. To assess the potency of these inhibitors, we used a competitive activity-based protein profiling (cABPP) assay, that measures the ability of the inhibitor to reduce fluorescent labeling of an enzyme by a fluorescent pan-β-glucuronidase activity-based probe (ABP).[16,34] Using this cABPP assay, we can simultaneously monitor the inhibition of HPSE and GUSB that are both present in platelet lysates.[6,29] As for the bacterial β-glucuronidases, human GUSB was also effectively inhibited by all the synthesized inhibitors (**8**−**11**), with low or sub-micromolar IC$_{50}$'s observed, see Table 3 and Figures 3, S7−S9. As observed with the bacterial β-glucuronidases, the *gluco*-configured inhibitors **9** and **11** were more potent inhibitors of human GUSB than their 4-axial *galacto*-analogues. Human heparanase was also inhibited by all four synthetic inhibitors.

Table 3. Inhibitory Potency of the Synthesized Inhibitors in the Platelet Lysate

inhibitor	HsGUSB IC$_{50}$ (μM)	HPSE IC$_{50}$ (μM)
8	4 ± 2	27 ± 3
9	1.7 ± 0.2	200 ± 80
10	0.70 ± 0.04	8 ± 1
11	0.60 ± 0.06	80 ± 20

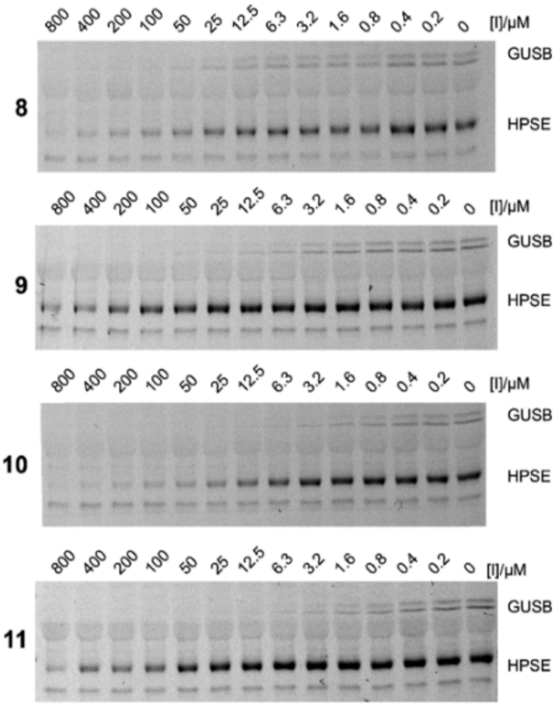


Figure 3. Competitive ABPP of platelet lysate with synthetic iminosugars. Iminosugars **8**−**11** inhibit both GusB and HPSE present in human platelet lysate. Quantitated fluorescence plots used to determine IC$_{50}$ values are available in Figure S7. Full length gel images and replicates are available in Figure S8, and Coomassie stained gels are available in Figure S9.

The IC$_{50}$'s determined for HPSE were greater than that observed for GUSB, which is anticipated for monomer-like inhibitors acting on multisubsite *endo*-acting enzymes. Surprisingly, for HPSE both the *galacto*-configured inhibitors (**8** and **10**) were more potent inhibitors than their *gluco*-configured analogues (**9** and **11**), reversing the trend seen with all of the other β-glucuronidases. However, the trend that inhibition was improved with the removal of the 3-geminal diol was maintained for both GUSB and HPSE. Furthermore, all the synthetic inhibitors were more potent than previously reported for the inhibition of GUSB or HPSE by siastatin B.[47]

Structural Basis for Inhibition by Synthetic Iminosugars. To examine the structural basis for inhibition by our panel of synthetic *gluco*- and *galacto*-configured iminosugars, we determined cocrystal structures of the 3-GDIs (**8**, **9**) and either HPSE or AcGH79 (Figure 4). Crystal structures with **8** and **9** bound to AcGH79 show that both of these inhibitors are situated in the same binding pocket as that of the *galacturonic*-noeuromycin **4**. Within the iminosugar ring, there is very little

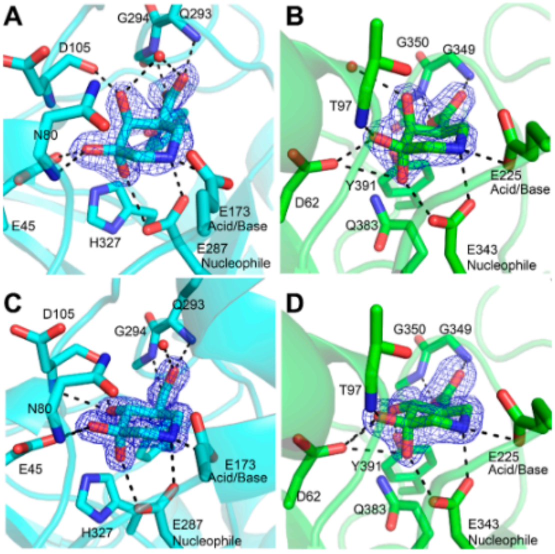


Figure 4. Structures of synthetic iminosugars bound to HPSE and AcGH79. (A) Complex between **8** and AcGH79. (B) Complex between **8** and HPSE. (C) Complex between **9** and AcGH79. (D) Complex between **9** and AcGH79. Electron density ($2F_o - F_c$) is shown for the ligand as a blue mesh contoured at 1.5 σ for *A* and *D* (*A* = 0.51 $e^-/Å^3$, *D* = 0.30 $e^-/Å^3$) and 2 σ for *B* and *C* (*B* = 0.45 $e^-/Å^3$, *C* = 0.69 $e^-/Å^3$). The polypeptide is shown in the cartoon form with active site residues shown as sticks. Apparent hydrogen bonding interactions are shown as dotted black lines. Water molecules are shown as red spheres.

conformational change between **8** and **9**, and both maintain the $^4C_1$ ring conformation seen for **4**. Furthermore, the axial 4-OH present in **8** forms the same H-bond interactions with the protein backbone (with the carbonyl of Asp105) as seen for **4**; however, it also has an additional H-bond to an active site water, see Figure 4. In the structure between the *gluco*-configured geminal diol **9** and AcGH79, the H-bond between Asp105 at the 4-position is still present—but at a different angle—and additional H-bonds between 4-OH and the backbone nitrogen of Asp105 and the carboxylate of Glu45 are present. These new H-bond interactions are the same as those observed for natural substrates[48] and other *gluco*-configured inhibitors[16,48] and are expected to be the cause of the increased inhibition of **9** over **8** for AcGH79. The axial 3-OH of geminal diol inhibitors **8** and **9** is found within hydrogen bonding distance of the catalytic nucleophile in both structures. The axial 3-OH is also within 3.0 Å of the εN of His327 in both structures; however, it is unlikely that there is a genuine 3_{ax}-HO···H-εN H-bond interaction as the 3_{ax}-OH is oriented approximately 80° to the plane of the histidine—rather than being in line with the plane, as would be expected. Furthermore, the εN of His327 forms an H-bond, at the appropriate angle, with Glu45.

As a point of comparison, we also determined a structure between AcGH79 and *gluco*-configured isofagomine **11** at 1.25 Å resolution, see Figure S10. This inhibitor was positioned nearly identically to geminal diol **9** in the active site of AcGH79. **11** also maintained all of the same H-bond interactions with the protein, with the exception of the H-bond between the axial-3-OH and the nucleophile seen for **9**. The binding of hydroxyl-containing inhibitors to enzyme active sites is associated with an energetic penalty due to desolvation that can be up to 26 kJ mol^{-1}, by some estimates.[49] Although the additional H-bond seen for **9** compensates for the desolvation of the axial hydroxyl, estimates of 18−21 kJ mol^{-1} have been made for the maximal interaction energy of a single hydroxyl H-bond.[50−52] Therefore, this single interaction is likely not enough to compensate for the additional desolvation energy, and therefore, the removal of the geminal diol is energetically favorable.

Structures of human heparanase soaked with **8** and **9** revealed interactions similar to those seen in the active site of AcGH79. The costructure of **8** soaked into crystals of HPSE showed the inhibitor in the exact same conformation and forming the same protein-inhibitor interactions as the degradation product observed when crystals were soaked with siastatin B. The costructure of **9** and HPSE has hydrogen bonds with the 4_{eq}-OH and Tyr391 and Asp62, as opposed to the hydrogen bond with water seen for **8**. The geminal diol **9** has the same hydrogen bonding network as the structure containing **11**, reported by Doherty et al.,[20] and additional H-bond between the 3_{ax}-OH with the nucleophile (Glu343), similar to what was observed in the active site of AcGH79. The 3_{ax}-OH in both structures is also within H-bonding distance of Asp62, an additional interaction that is not seen in the AcGH79-geminal diol structures. The 3_{ax}-OH in both the **8** and **9** structures is also within 3 Å of the εN of Gln383, however, as for His327 in the active site of AcGH79, this atom is positioned incorrectly to form a hydrogen bond with the axial-hydroxyl. The plane of the amide bond of Gln383 is oriented perpendicular to the axial 3-OH, and the εN of Gln383 already forms two hydrogen bonds with Tyr391 and Asp62. The presence of two additional hydrogen bonds with the active site of HPSE may be able to compensate more fully for the desolvation of an additional hydroxyl in the geminal-diol containing inhibitors. This in turn may be the reason why geminal diol inhibitor **8** is observed in HPSE, when soaked with siastatin B, instead of noeuromycin derivative **4** that is observed in the other β-glucuronidases.

■ DISCUSSION

The breakdown products of a synthetic 2-trifluoroacetamido derivative of siastatin B observed by Kondo et al. in the media of enzyme assays are the same hemiaminal **4** and geminal-diol **8** that we observed as breakdown products of siastatin B.[32] However, the 2-trifluoroacetamido derivative is a more potent inhibitor of β-glucuronidases than the unmodified parent compound.[30,31] We attribute the enhanced inhibition of 2-trifluoroacetamido-derived siastatin B to the increased rate of elimination that occurs for 2-trifluoroacetamido when compared to the *N*-acetyl group present at the 2-position, this thereby generates the active compounds at higher concentrations than that observed for siastatin B. This "pro-inhibitor" molecular design opens interesting avenues for the future design of glycosidase inhibitors. First, there is potential to optimize the leaving group at the 2-position to generate a molecule that decays into noeuromycin-like and 3-GDIs in hours or days rather than the rapid decay observed for the 2-trifluoroacetamido derivatives of siastatin B, thereby producing a prodrug molecule with a delayed release. Second, there is potential for the generation of prodrug-like molecules that only decay into their active iminosugar form when acted upon by another physiological enzyme—a strategy that could be useful for the development of chaperones for lysosomal storage diseases—or by light-activated decomposition, thereby enabling targeted delivery of inhibitors.

The class of compounds known as noeuromycins were first synthesized and reported by Bols et al. in 2001.[53] These potent glycosidase inhibitors were inspired by the natural product 1-deoxynojirimycin but were made purely by synthetic means. To the best of our knowledge, there has been no report of a natural product glycosidase inhibitor with a noeuromycin like structure. We however show here that three of the four enzymes that were soaked with siastatin B have a galacturonic acid-configured noeuromycin bound in their active site. This surprising "re-discovery" of noeuromycin-type inhibitors more than 20 years after the first synthesis of this class of molecule is a striking example of the convergence of molecular designs by both man and nature.

■ CONCLUSIONS

In conclusion, we have shown here through X-ray crystallography that two different types of 1-*N*-iminosugar inhibitors are generated from siastatin B: hemiaminal **4** and 3-GDI **8**. Notably, 3-GDIs are a new class of iminosugar glycoside hydrolase inhibitor, that have been shown to inhibit both GH2 and GH79 β-glucuronidases. Synthesis of 3-GDIs and their 3-hydroxy counterparts allowed us to dissect the mechanism of their inhibition both kinetically and structurally. These 3-GDIs are well accommodated within the active site pocket and form an additional H-bond interaction with the catalytic nucleophile. We also improved the inhibition of these 3-geminal diol inhibitors through epimerization of the 4-hydroxyl. Together these results give new insights into how the broad-spectrum inhibitor siastatin B is able to act on several families of glycoside hydrolase with very different substrate specificities. This information, in turn, should enable the future design of broad-spectrum inhibitors that can target other classes of glycosidase, beyond those targeted by siastatin B.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/jacs.3c04162.

Detailed methods for gene expression, protein purification, biochemical assays, structural biology, and complete synthetic protocols; atomic coordinates and structure factors deposited in the Protein Data Bank (PDB ID codes: 8CQI, 8OGX, 8OHQ, 8OHR, 8OHT, 8OHU, 8OHV, 8OHW, and 8OHX); NMR experiments, LC–MS experiments, enzyme kinetics experiments, full-length SDS-page gels, and additional structures; and details on crystal structure data collection and refinement (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Gideon J. Davies – *York Structural Biology Laboratory, Department of Chemistry, The University of York, YO10 5DD York, U.K.*; orcid.org/0000-0002-7343-776X; Email: gideon.davies@york.ac.uk

Zachary Armstrong – *York Structural Biology Laboratory, Department of Chemistry, The University of York, YO10 5DD York, U.K.; Leiden Institute of Chemistry, Leiden University, 2300 RA Leiden, The Netherlands*; orcid.org/0000-0002-4086-2946; Email: z.w.b.armstrong@lic.leidenuniv.nl

Authors

Yurong Chen – *Leiden Institute of Chemistry, Leiden University, 2300 RA Leiden, The Netherlands*

Adrianus M. C. H. van den Nieuwendijk – *Leiden Institute of Chemistry, Leiden University, 2300 RA Leiden, The Netherlands*

Liang Wu – *York Structural Biology Laboratory, Department of Chemistry, The University of York, YO10 5DD York, U.K.*; Present Address: Structural Biology, The Rosalind Franklin Institute, Harwell Science & Innovation Campus, Didcot OX11 0QX, U.K; orcid.org/0000-0003-0294-7065

Elisha Moran – *York Structural Biology Laboratory, Department of Chemistry, The University of York, YO10 5DD York, U.K.*; orcid.org/0000-0002-3463-6862

Foteini Skoulikopoulou – *Leiden Institute of Chemistry, Leiden University, 2300 RA Leiden, The Netherlands*

Vera van Riet – *Leiden Institute of Chemistry, Leiden University, 2300 RA Leiden, The Netherlands*

Hermen S. Overkleeft – *Leiden Institute of Chemistry, Leiden University, 2300 RA Leiden, The Netherlands*; orcid.org/0000-0001-6976-7005

Complete contact information is available at: https://pubs.acs.org/10.1021/jacs.3c04162

Author Contributions

[§]Y.C., M.C.H.v.d.N., and L.W. contributed equally.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank The Netherlands Organization for Scientific Research (NWO; TOP grant to H.S.O., Veni grant: VI.Veni.212.173 to Z.A.), the European Research Council (ERC-2011-AdG-290836 "Chembiosphing" to H.S.O. and ERC-2012-AdG-32294 "Glycopoise" to G.J.D. and ERC-2020-SyG-951231 "Carbocentre", to G.J.D. and H.S.O.), and the Biotechnology and Biological Sciences Research Council (BBSRC; BB/R001162/1 and BB/M011151/1 grants to G.J.D.). G.J.D. thanks the Royal Society for the Ken Murray Research Professorship. L.W. acknowledges support from the Wellcome Trust for a Sir Henry Dale Fellowship (218579/Z/19/Z). We thank the China Scholarship Council (CSC, Ph.D. Grant to Y.C.). Hans van den Elst and Maria Guimaraes Da Lomba Ferraz are acknowledged for technical support with LC-MS measurements. Johan Turkenburg and Sam Hart are thanked for coordinating X-ray data collection. We also thank the Diamond Light Source for access to beamlines i02, i03, i04, and i04-1 (proposal numbers mx9948, mx13587, and mx24948) that contributed to the results presented here.

■ REFERENCES

(1) Horne, G.; Wilson, F. X.; Tinsley, J.; Williams, D. H.; Storer, R. Iminosugars past, present and future: medicines for tomorrow. *Drug Discovery Today* **2011**, *16* (3–4), 107–118.

(2) Winchester, B. G. Iminosugars: from botanical curiosities to licensed drugs. *Tetrahedron: Asymmetry* **2009**, *20* (6–8), 645–651.

(3) Butters, T. D.; Dwek, R. A.; Platt, F. M. Imino sugar inhibitors for treating the lysosomal glycosphingolipidoses. *Glycobiology* **2005**, *15* (10), 43R–52R.

(4) Weidemann, F.; Jovanovic, A.; Herrmann, K.; Vardarli, I. Chaperone Therapy in Fabry Disease. *Int. J. Mol. Sci.* **2022**, *23* (3), 1887.

(5) Pineda, M.; Walterfang, M.; Patterson, M. C. Miglustat in Niemann-Pick disease type C patients: a review. *Orphanet J. Rare Dis.* **2018**, *13* (1), 140.

(6) Scott, L. J.; Spencer, C. M. Miglitol. *Drugs* **2000**, *59* (3), 521–549.

(7) Armstrong, Z.; Kuo, C.-L.; Lahav, D.; Liu, B.; Johnson, R.; Beenakker, T. J. M.; de Boer, C.; Wong, C.-S.; van Rijssel, E. R.; Debets, M. F.; Florea, B. I.; Hissink, C.; Boot, R. G.; Geurink, P. P.; Ovaa, H.; van der Stelt, M.; van der Marel, G. M.; Codée, J. D. C.; Aerts, J. M. F. G.; Wu, L.; Overkleeft, H. S.; Davies, G. J. Manno-*epi*-cyclophellitols Enable Activity-Based Protein Profiling of Human α-Mannosidases and Discovery of New Golgi Mannosidase II Inhibitors. *J. Am. Chem. Soc.* **2020**, *142* (30), 13021–13029.

(8) Chen, W.-A.; Chen, Y.-H.; Hsieh, C.-Y.; Hung, P.-F.; Chen, C.-W.; Chen, C.-H.; Lin, J.-L.; Cheng, T.-J. R.; Hsu, T.-L.; Wu, Y.-T.; Shen, C.-N.; Cheng, W.-C. Harnessing natural-product-inspired combinatorial chemistry and computation-guided synthesis to develop N-glycan modulators as anticancer agents. *Chem. Sci.* **2022**, *13* (21), 6233–6243.

(9) Esposito, A.; D'Alonzo, D.; De Fenza, M.; De Gregorio, E.; Tamanini, A.; Lippi, G.; Dechecchi, M. C.; Guaragna, A. Synthesis and Therapeutic Applications of Iminosugars in Cystic Fibrosis. *Int. J. Mol. Sci.* **2020**, *21* (9), 3353.

(10) Warfield, K. L.; Barnard, D. L.; Enterlein, S. G.; Smee, D. F.; Khaliq, M.; Sampath, A.; Callahan, M. V.; Ramstedt, U.; Day, C. W. The Iminosugar UV-4 is a Broad Inhibitor of Influenza A and B Viruses *ex Vivo* and in Mice. *Viruses* **2016**, *8* (3), 71.

(11) Sobala, Ł. F.; Fernandes, P. Z.; Hakki, Z.; Thompson, A. J.; Howe, J. D.; Hill, M.; Zitzmann, N.; Davies, S.; Stamataki, Z.; Butters, T. D.; Alonzi, D. S.; Williams, S. J.; Davies, G. J. Structure of human endo-α-1,2-mannosidase (MANEA), an antiviral host-glycosylation target. *Proc. Natl. Acad. Sci. U.S.A.* **2020**, *117* (47), 29595–29601.

(12) Umezawa, H.; Aoyagi, T.; Komiyama, T.; Morishima, H.; Hamada, M.; Takeuchi, T. Purification and characterization of a sialidase inhibitor, siastatin, produced by Streptomyces. *J. Antibiot.* **1974**, *27* (12), 963–969.

(13) Nishimura, Y. Stereoselective synthesis and transformation of siastatin B, a novel glycosidase inhibitor, directed toward new drugs for viral infection and tumor metastasis. In *Studies in Natural Products Chemistry*; Atta-ur-Rahman, Ed.; Elsevier, 1995; Vol. *16*, pp 75–121.

(14) Tailford, L. E.; Owen, C. D.; Walshaw, J.; Crost, E. H.; Hardy-Goddard, J.; Le Gall, G.; de Vos, W. M.; Taylor, G. L.; Juge, N. Discovery of intramolecular trans-sialidases in human gut microbiota suggests novel mechanisms of mucosal adaptation. *Nat. Commun.* **2015**, *6* (1), 7624.

(15) Rahman, M. M.; Hirokawa, T.; Tsuji, D.; Tsukimoto, J.; Hitaoka, S.; Yoshida, T.; Chuman, H.; Itoh, K. Novel pH-dependent regulation of human cytosolic sialidase 2 (NEU2) activities by siastatin B and structural prediction of NEU2/siastatin B complex. *Biochem. Biophys. Rep.* **2015**, *4*, 234–242.

(16) Wu, L.; Jiang, J.; Jin, Y.; Kallemeijn, W. W.; Kuo, C.-L.; Artola, M.; Dai, W.; van Elk, C.; van Eijk, M.; van der Marel, G. A.; Codée, J. D. C.; Florea, B. I.; Aerts, J. M. F. G.; Overkleeft, H. S.; Davies, G. J. Activity-based probes for functional interrogation of retaining β-glucuronidases. *Nat. Chem. Biol.* **2017**, *13* (8), 867–873.

(17) di Bello, I. C.; Dorling, P.; Fellows, L.; Winchester, B. Specific inhibition of human β-D-glucuronidase and α-L-iduronidase by a trihydroxy pipecolic acid of plant origin. *FEBS Lett.* **1984**, *176* (1), 61–64.

(18) Dashnyam, P.; Lin, H.-Y.; Chen, C.-Y.; Gao, S.; Yeh, L.-F.; Hsieh, W.-C.; Tu, Z.; Lin, C.-H. Substituent Position of Iminocyclitols Determines the Potency and Selectivity for Gut Microbial Xenobiotic-Reactivating Enzymes. *J. Med. Chem.* **2020**, *63* (9), 4617–4627.

(19) Rasmussen, T. S.; Koldsø, H.; Nakagawa, S.; Kato, A.; Schiøtt, B.; Jensen, H. H. Synthesis of uronic-Noeurostegine - a potent bacterial β-glucuronidase inhibitor. *Org. Biomol. Chem.* **2011**, *9* (22), 7807–7813.

(20) Doherty, G.; Ler, G.; Wimmer, N.; Bernhardt, P.; Ashmus, R.; Vocadlo, D.; Armstrong, Z.; Davies, G.; Maccarana, M.; Li, J.-p.; Kayal, Y.; Ferro, V. Synthesis of Uronic Acid 1-Azasugars as Putative Inhibitors of α-Iduronidase, β-Glucuronidase and Heparanase. *ChemBioChem* **2023**, *24* (4), No. e202200619.

(21) Lin, H.-Y.; Chen, C.-Y.; Lin, T.-C.; Yeh, L.-F.; Hsieh, W.-C.; Gao, S.; Burnouf, P.-A.; Chen, B.-M.; Hsieh, T.-J.; Dashnyam, P.; Kuo, Y.-H.; Tu, Z.; Roffler, S. R.; Lin, C.-H. Entropy-driven binding of gut bacterial β-glucuronidase inhibitors ameliorates irinotecan-induced toxicity. *Commun. Biol.* **2021**, *4* (1), 280.

(22) Roeser, K.-R.; Legler, G. Role of sugar hydroxyl groups in glycoside hydrolysis. Cleavage mechanism of deoxyglucosides and related substrates by β-glucosidase A3 from *Aspergillus wentii*. *Biochim. Biophys. Acta, Enzymol.* **1981**, *657* (2), 321–333.

(23) Mega, T.; Matsushima, Y. Energy of Binding of *Aspergillus oryzae* β-Glucosidase with the Substrate, and the Mechanism of Its Enzymic Action. *J. Biochem.* **1983**, *94* (5), 1637–1647.

(24) Namchuk, M. N.; Withers, S. G. Mechanism of Agrobacterium .beta.-glucosidase: kinetic analysis of the role of noncovalent enzyme/substrate interactions. *Biochemistry* **1995**, *34* (49), 16194–16202.

(25) Davies, G. J.; Planas, A.; Rovira, C. Conformational Analyses of the Reaction Coordinate of Glycosidases. *Acc. Chem. Res.* **2012**, *45* (2), 308–316.

(26) Speciale, G.; Thompson, A. J.; Davies, G. J.; Williams, S. J. Dissecting conformational contributions to glycosidase catalysis and inhibition. *Curr. Opin. Struct. Biol.* **2014**, *28*, 1–13.

(27) Wu, L.; Davies, G. J. An Overview of the Structure, Mechanism and Specificity of Human Heparanase. In *Heparanase*; Vlodavsky, I., Sanderson, R., Ilan, N., Eds.; Springer: Cham, 2020; Vol. *1221*, pp 139–167.

(28) 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* (12), 1016–1022.

(29) 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* (31), No. e2203167119.

(30) Nishimura, Y.; Kudo, T.; Kondo, S.; Takeuchi, T.; Tsuruoka, T.; Fukuyasu, H.; Shibahara, S. Totally synthetic analogues of siastatin B. III. Trifluoroacetamide analogues having inhibitory activity for tumor metastasis. *J. Antibiot.* **1994**, *47* (1), 101–107.

(31) Nishimura, Y.; Shitara, E.; Adachi, H.; Toyoshima, M.; Nakajima, M.; Okami, Y.; Takeuchi, T. Flexible Synthesis and Biological Activity of Uronic Acid-Type gem-Diamine 1-N-Iminosugars: A New Family of Glycosidase Inhibitors. *J. Org. Chem.* **2000**, *65* (1), 2–11.

(32) Kondo, K. Glycosidase inhibitors of gem-diamine 1-*N*-iminosugars structures in media of enzyme assays. *Bioorg. Med. Chem.* **2001**, *9* (5), 1091–1095.

(33) Nishimura, Y. Gem-diamine 1-N-iminosugars as versatile glycomimetics: synthesis, biological activity and therapeutic potential. *J. Antibiot.* **2009**, *62* (8), 407–423.

(34) Borlandelli, V.; Armstrong, Z.; Nin-Hill, A.; Codée, J. D. C.; Raich, L.; Artola, M.; Rovira, C.; Davies, G. J.; Overkleeft, H. S. 4-O-Substituted Glucuronic Cyclophellitols are Selective Mechanism-Based Heparanase Inhibitors. *ChemMedChem* **2023**, *18* (4), No. e202200580.

(35) Jayatilleke, K. M.; Hulett, M. D. Heparanase and the hallmarks of cancer. *J. Transl. Med.* **2020**, *18* (1), 453.

(36) 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* **1983**, *220* (4597), 611–613.

(37) Vlodavsky, I.; Fuks, Z.; Bar-Ner, M.; Ariav, Y.; Schirrmacher, V. Lymphoma cell-mediated degradation of sulfated proteoglycans in the

subendothelial extracellular matrix: relationship to tumor cell metastasis. *Cancer Res.* **1983**, *43* (6), 2704.

(38) Vlodavsky, I.; Friedmann, Y. Molecular properties and involvement of heparanase in cancer metastasis and angiogenesis. *J. Clin. Invest.* **2001**, *108* (3), 341−347.

(39) 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* (9), 1661−1663.

(40) Zubkova, O. V.; Ahmed, Y. A.; Guimond, S. E.; Noble, S.-L.; Miller, J. H.; Alfred Smith, R. A.; Nurcombe, V.; Tyler, P. C.; Weissmann, M.; Vlodavsky, I.; Turnbull, J. E. Dendrimer Heparan Sulfate Glycomimetics: Potent Heparanase Inhibitors for Anticancer Therapy. *ACS Chem. Biol.* **2018**, *13* (12), 3236−3242.

(41) Loka, R. S.; Sletten, E. T.; Barash, U.; Vlodavsky, I.; Nguyen, H. M. Specific Inhibition of Heparanase by a Glycopolymer with Well-Defined Sulfation Pattern Prevents Breast Cancer Metastasis in Mice. *ACS Appl. Mater. Interfaces* **2019**, *11* (1), 244−254.

(42) Liu, H.; Lillelund, V. H.; Andersch, J.; Liang, X.; Bols, M. Synthesis and Chemistry of Noeuromycin and Isofagomine Analogues. *J. Carbohydr. Chem.* **2004**, *23* (4), 223−238.

(43) Igarashi, Y.; Ichikawa, M.; Ichikawa, Y. Synthesis of a potent inhibitor of β-glucuronidase. *Tetrahedron Lett.* **1996**, *37*, 2707−2708.

(44) Ichikawa, Y.; Igarashi, Y.; Ichikawa, M.; Suhara, Y. 1-*N*-Iminosugars: potent and selective inhibitors of β-glycosidases. *J. Am. Chem. Soc.* **1998**, *120*, 3007−3018.

(45) Kim, Y. J.; Ichikawa, M.; Ichikawa, Y. Highly Selective Synthesis of 1-*N*-Iminosugars of the D-Glucose and -Glucuronic Acid Types. *Org. Lett.* **2000**, *65*, 2599−2602.

(46) Drula, E.; Garron, M.-L.; Dogan, S.; Lombard, V.; Henrissat, B.; Terrapon, N. The carbohydrate-active enzyme database: functions and literature. *Nucleic Acids Res.* **2022**, *50* (D1), D571−D577.

(47) Kawase, Y.; Takahashi, M.; Takatsu, T.; Arai, M.; Nakajima, M.; Tanzawa, K. A-72363 A-1, A-2, and C, Novel Heparanase Inhibitors from Streptomyces nobilis SANK 60192. II. Biological Activities. *J. Antibiot.* **1996**, *49* (1), 61−64.

(48) Michikawa, M.; Ichinose, H.; Momma, M.; Biely, P.; Jongkees, S.; Yoshida, M.; Kotake, T.; Tsumuraya, Y.; Withers, S. G.; Fujimoto, Z.; Kaneko, S. Structural and Biochemical Characterization of Glycoside Hydrolase Family 79 β-Glucuronidase from Acidobacterium capsulatum. *J. Biol. Chem.* **2012**, *287* (17), 14069−14077.

(49) Cramer, J.; Sager, C. P.; Ernst, B. Hydroxyl Groups in Synthetic and Natural-Product-Derived Therapeutics: A Perspective on a Common Functional Group. *J. Med. Chem.* **2019**, *62* (20), 8915−8930.

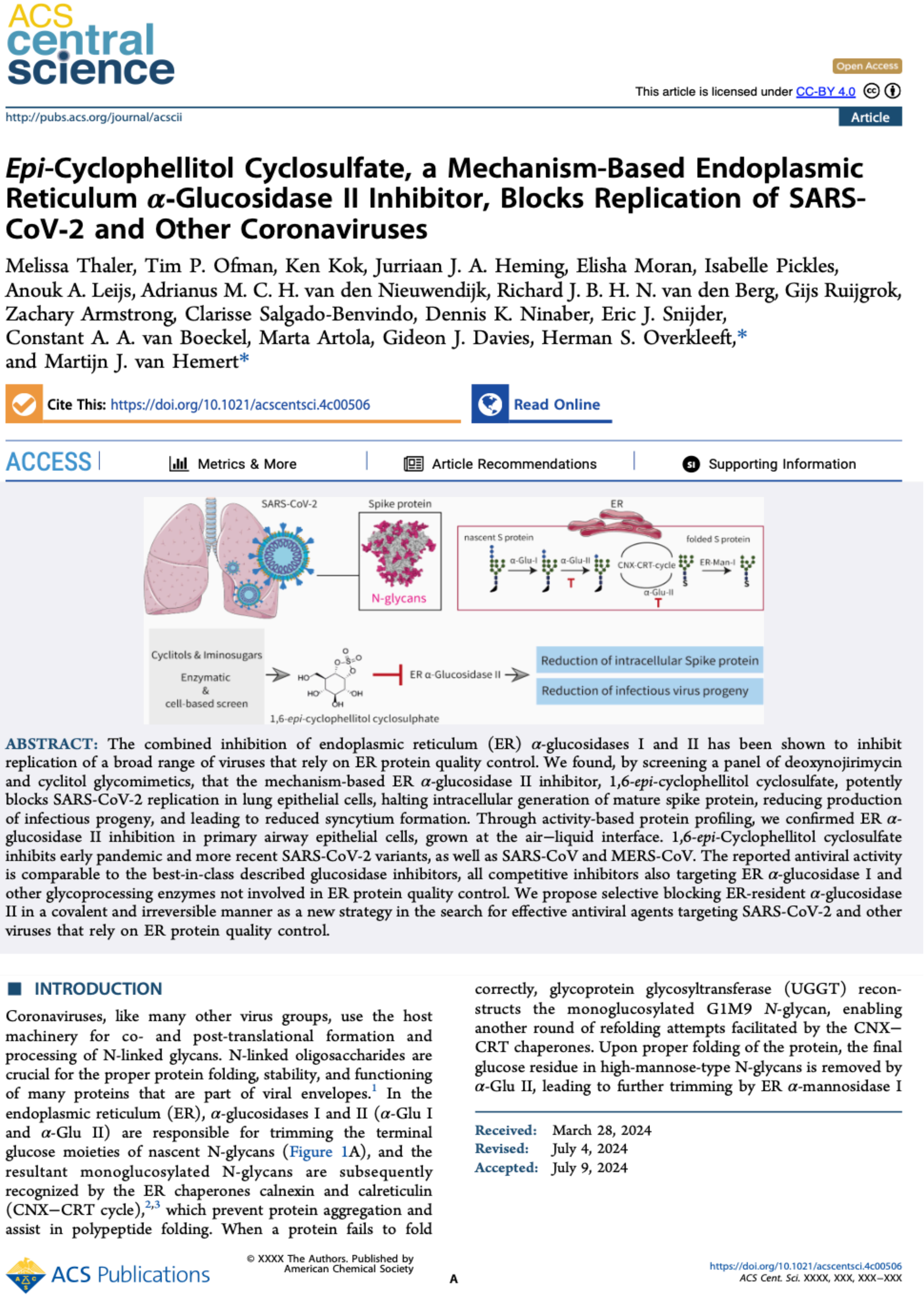
(50) Sager, C. P.; Eriş, D.; Smieško, M.; Hevey, R.; Ernst, B. What contributes to an effective mannose recognition domain? *Beilstein J. Org. Chem.* **2017**, *13*, 2584−2595.

(51) Vedani, A.; Huhta, D. W. A new force field for modeling metalloproteins. *J. Am. Chem. Soc.* **1990**, *112* (12), 4759−4767.

(52) Steiner, T. The Hydrogen Bond in the Solid State. *Angew. Chem., Int. Ed.* **2002**, *41* (1), 48−76.

(53) Liu, H.; Liang, X.; Søhoel, H.; Bülow, A.; Bols, M. Noeuromycin,[1] A Glycosyl Cation Mimic that Strongly Inhibits Glycosidases. *J. Am. Chem. Soc.* **2001**, *123* (21), 5116−5117.

7.3.2. Epi-Cyclophellitol Cyclosulfate, a Mechanism-Based Endoplasmic Reticulum α-Glucosidase II Inhibitor, Blocks Replication of SARS-CoV-2 and Other Coronavirus



ACS central science
http://pubs.acs.org/journal/acscii
Article

Epi-Cyclophellitol Cyclosulfate, a Mechanism-Based Endoplasmic Reticulum *α*-Glucosidase II Inhibitor, Blocks Replication of SARS-CoV-2 and Other Coronaviruses

Melissa Thaler, Tim P. Ofman, Ken Kok, Jurriaan J. A. Heming, Elisha Moran, Isabelle Pickles, Anouk A. Leijs, Adrianus M. C. H. van den Nieuwendijk, Richard J. B. H. N. van den Berg, Gijs Ruijgrok, Zachary Armstrong, Clarisse Salgado-Benvindo, Dennis K. Ninaber, Eric J. Snijder, Constant A. A. van Boeckel, Marta Artola, Gideon J. Davies, Herman S. Overkleeft,* and Martijn J. van Hemert*

Cite This: https://doi.org/10.1021/acscentsci.4c00506

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: The combined inhibition of endoplasmic reticulum (ER) *α*-glucosidases I and II has been shown to inhibit replication of a broad range of viruses that rely on ER protein quality control. We found, by screening a panel of deoxynojirimycin and cyclitol glycomimetics, that the mechanism-based ER *α*-glucosidase II inhibitor, 1,6-*epi*-cyclophellitol cyclosulfate, potently blocks SARS-CoV-2 replication in lung epithelial cells, halting intracellular generation of mature spike protein, reducing production of infectious progeny, and leading to reduced syncytium formation. Through activity-based protein profiling, we confirmed ER *α*-glucosidase II inhibition in primary airway epithelial cells, grown at the air−liquid interface. 1,6-*epi*-Cyclophellitol cyclosulfate inhibits early pandemic and more recent SARS-CoV-2 variants, as well as SARS-CoV and MERS-CoV. The reported antiviral activity is comparable to the best-in-class described glucosidase inhibitors, all competitive inhibitors also targeting ER *α*-glucosidase I and other glycoprocessing enzymes not involved in ER protein quality control. We propose selective blocking ER-resident *α*-glucosidase II in a covalent and irreversible manner as a new strategy in the search for effective antiviral agents targeting SARS-CoV-2 and other viruses that rely on ER protein quality control.

■ INTRODUCTION

Coronaviruses, like many other virus groups, use the host machinery for co- and post-translational formation and processing of N-linked glycans. N-linked oligosaccharides are crucial for the proper protein folding, stability, and functioning of many proteins that are part of viral envelopes.[1] In the endoplasmic reticulum (ER), *α*-glucosidases I and II (*α*-Glu I and *α*-Glu II) are responsible for trimming the terminal glucose moieties of nascent N-glycans (Figure 1A), and the resultant monoglucosylated N-glycans are subsequently recognized by the ER chaperones calnexin and calreticulin (CNX−CRT cycle),[2,3] which prevent protein aggregation and assist in polypeptide folding. When a protein fails to fold correctly, glycoprotein glycosyltransferase (UGGT) reconstructs the monoglucosylated G1M9 *N*-glycan, enabling another round of refolding attempts facilitated by the CNX−CRT chaperones. Upon proper folding of the protein, the final glucose residue in high-mannose-type N-glycans is removed by *α*-Glu II, leading to further trimming by ER *α*-mannosidase I

Received: March 28, 2024
Revised: July 4, 2024
Accepted: July 9, 2024

ACS Publications

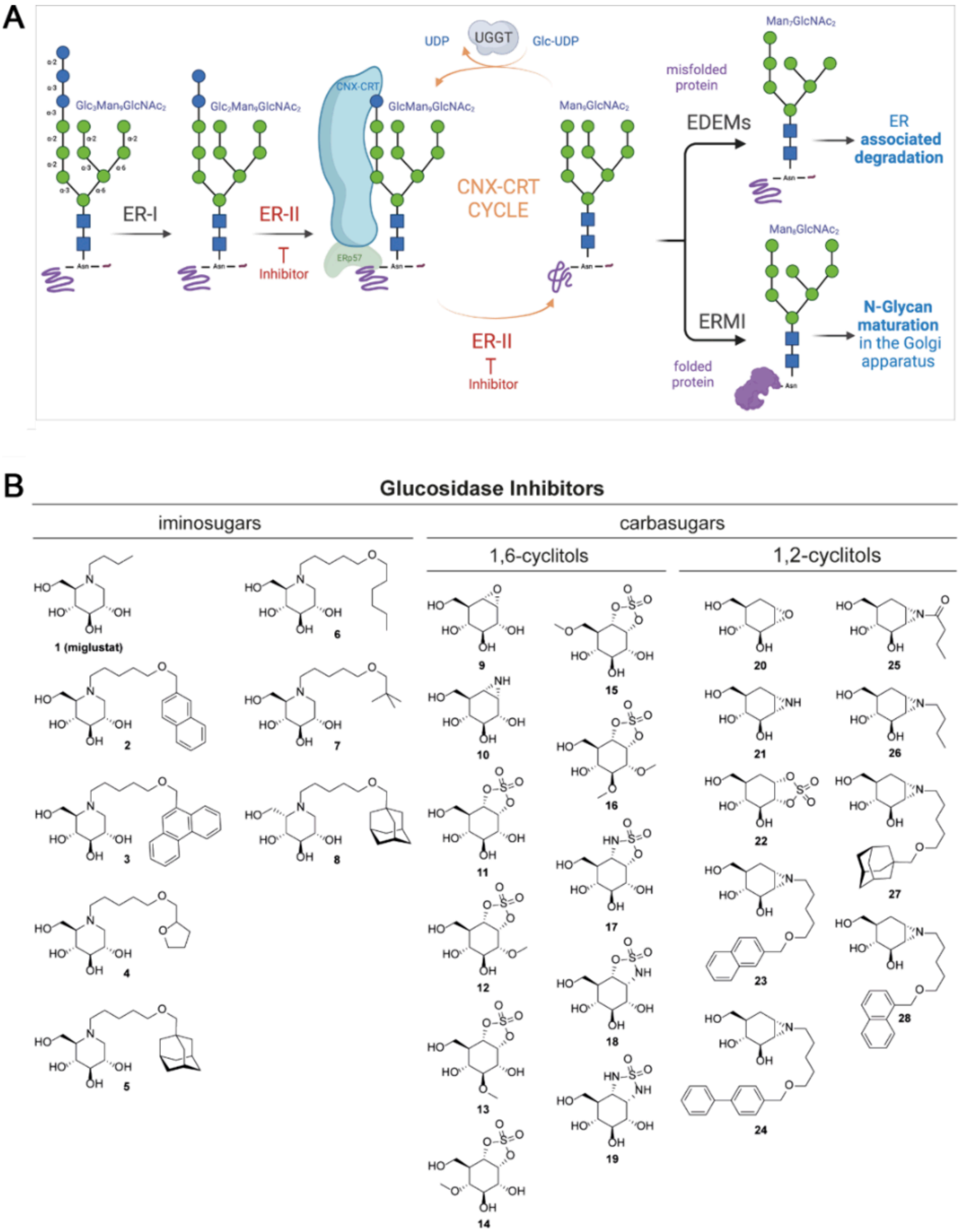


Figure 1. (A) Schematic of N-glycan processing of newly synthesized proteins in the ER lumen. Folding of nascent proteins in the ER is promoted by the calnexin−calreticulin cycle (CNX−CRT cycle), which relies on glycan trimming by ER α-Glu II (ER-II). (B) Focused library of 28 iminosugars and cyclitol subjects of the here-presented studies.

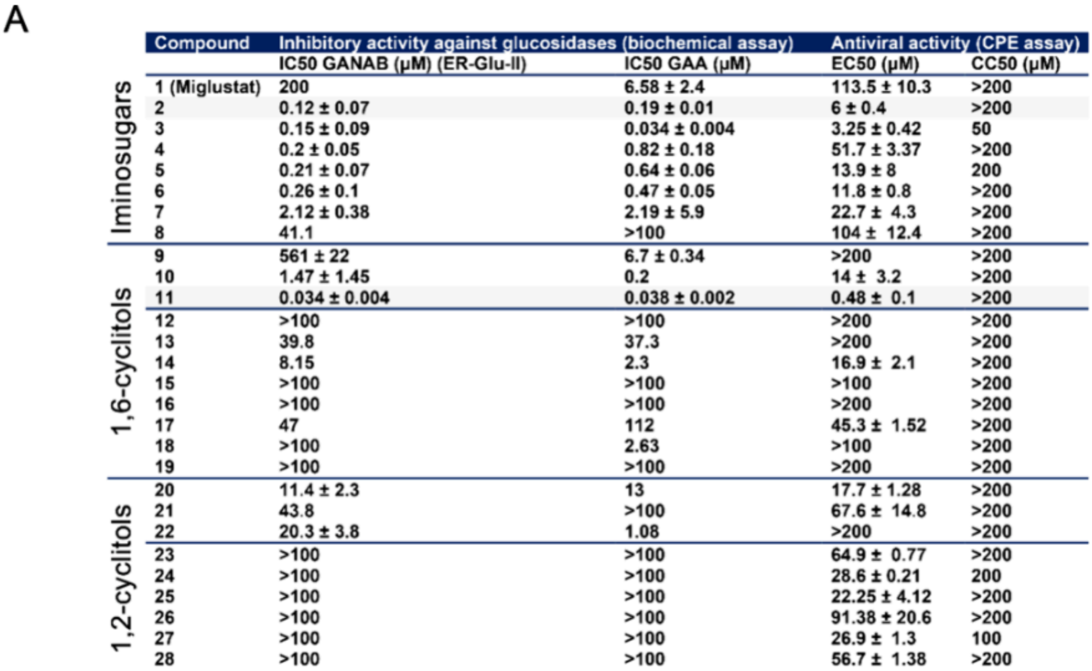
(ERMI), after which the *N*-glycoproteins are routed to the Golgi apparatus for N-glycan maturation and further post-translational modification events en route to their final destination. Glycoproteins that fail to attain their proper conformation undergo mannose trimming orchestrated by the ER degradation-enhancing mannosidase-like proteins (EDEMs) and ultimately are routed toward the ER-associated degradation (ERAD) machinery. Inhibition of ER α-Glu I and II has been shown to interfere with proper processing of nascent proteins through the CNX−CRT cycle, leading to their inappropriate folding, eventual dislocation from the ER, and proteasomal degradation.[4] This holds true for host and viral *N*-glycoproteins alike, and ER α-Glu I/II inhibition has therefore been considered as a viable strategy for antiviral therapeutics development for several decades.[5,6]

Many studies have reported the ability of iminosugars to inhibit replication of various viruses, through the blocking of ER protein quality control via ER α-Glu I/II inhibition.[7] Iminosugars are polyhydroxylated glycomimetic alkaloids featuring a basic amine, replacing the sugar ring oxygen, that is thought to interact with glycosidase active site residues that partake in enzymatic glycosidic bond hydrolysis.[8,9] The potential of iminosugars as antivirals was first reported in 1987[5,10,11] in the context of human immunodeficiency virus (HIV), which relies on the host ER machinery for glycoprotein processing.[12] These studies revealed that the two iminosugar compounds, deoxynojirimycin and castanospermine, as well as some structural analogues thereof, inhibit ER α-Glu I and II and block the production of HIV infectious progeny *in vitro*. Later studies using a host of structurally diverse iminosugars described blocking replication of a broad range of viruses *in vitro* and *in vivo*, including influenza viruses,[13−15] severe acute respiratory syndrome coronavirus (SARS-CoV),[16] dengue virus, and the hemorrhagic fever viruses Marburg and Ebola.[17,18] One of the studied iminosugars, UV-4B, showed promising results in mice, as a single high dose, which caused hallmarks of ER α-Glu I inhibition *in vivo*, protected the animals from a lethal dose of DENV or influenza virus.[15] Interestingly, patients that have N-glycosylation defects (defects in ER α-Glu I) due to a congenital disorder have also reduced susceptibility to infection with enveloped viruses that depend on host glycan processing for their replication.[19] Despite promising *in vitro* studies, phase II clinical trials with the iminosugar Celgosivir (a prodrug form of castanospermine) showed no beneficial outcomes when it was used as monotherapy for dengue and hepatitis C viral infections.[20,21] Most recently, a range of competitive α-glucosidase inhibitors have been studied during the search for antivirals against SARS-CoV-2.[22−24] The spike (S) protein of SARS-CoV-2, one of the envelope proteins on the virus surface, is heavily glycosylated with 23 reported N-glycan sites.[25] Besides shielding of antibody epitopes,[26] and modulating protein structure, N-glycosylation of S protein and its receptor binding domain (RBD) is crucial for virus infectivity, as the S protein drives virus entry by binding to the host receptor ACE2 and mediates fusion between the virus and host cell membrane.[27] N-glycans and their modulation through deletion of specific sites on the RBD were reported to be important for conformational stability and accessibility of the RBD for ACE2 binding.[28−31] Therefore, the incorporation of nonfunctional immaturely glycosylated S proteins can reduce the specific infectivity of progeny virions.[16,32] Disruption of the CNX−CRT-mediated glycoprotein processing, by iminosugars specifically, was reported to reduce the incorporation of S protein into SARS-CoV pseudovirus particles.[16] In this study, it was suggested that ER α-Glu I/II inhibition could lead to both the degradation of improperly processed S proteins in the ER and the incorporation of incompletely glycosylated S proteins into virus particles, thus having a two-pronged mode of action.

Despite the decades of research on iminosugars, no small molecules inhibiting ER α-Glu have proceeded beyond phase II clinical trials[33,34] as antivirals. With the aim of uncovering alternative inhibitor designs for antiviral drug discovery, and building upon our recent studies on mechanism-based, covalent and irreversible glycosidase inhibition,[35−40] we decided to assess a panel of mechanism-based inhibitors, side by side with a set of classical *N*-alkyl iminosugars, for their ability to inhibit SARS-CoV-2 replication through inhibition of ER α-Glu I and II. While performing the same net transformation (hydrolysis of α-glucosidic linkages), ER α-Glu I and II do so with distinct mechanisms. Both enzymes feature a carboxylic acid and a carboxylate containing amino acid in their active sites and process their substrate by acid catalysis.[8,9] Both enzymes are therefore amenable to inhibition by a basic, glucose-mimetic iminosugar. In contrast to ER α-Glu I, ER α-Glu II forms a covalent intermediate with its substrate during processing by utilizing one of the carboxylates as a nucleophile. This nucleophile can be trapped by glucomimetic cyclitols endowed with an electrophile (epoxide, aziridine, or cyclic sulfate). We have shown in the past that 1,6-*epi*-cyclophellitol (**9**, Figure 1) as well as its aziridine (**10**) and cyclic sulfate (**11**) analogues potently and selectively blocks ER α-Glu II.[35] In this study, we screened members of both compound classes, cyclitols and iminosugars, for their inhibition of ER α-Glu II and antiviral activity against SARS-CoV-2. We demonstrate that 1,6-*epi*-cyclophellitol cyclosulfate (**11**) most potently reduces the enzyme activity of α-Glu II and exerts the best antiviral efficacy against SARS-CoV-2. We also show that this compound blocks replication of all SARS-CoV-2 variants tested, as well as the pathogenic SARS-CoV and MERS-CoV, making it an interesting lead for further exploration toward a new class of antiviral drugs.

RESULTS

Efficacy of Glucosidase Inhibitors against SARS-CoV-2 Correlates with Their Activity against ER α-Glucosidase II. The panel of iminosugars and cyclitols, subject of the here-presented studies, is depicted in Figure 1B. With respect to the iminosugars, and to keep in line with literature precedents, we selected *N*-alkyl deoxynojirimycins **1**−**8**. Deoxynojirimycin (DNJ) features the glucopyranose configuration, and *N*-alkyl derivatives have been shown to be more effective glucosidase inhibitors compared to non-substituted DNJ.[41−43] This includes the benchmark analogue, *N*-butyl-DNJ **1** (Miglustat, Zavesca) which is part of almost all antiviral studies on iminosugars targeting α-Glu I/II. In fact, Miglustat is a clinical drug for the treatment of Gaucher disease and acts as a glucosylceramidase (GCS) inhibitor.[44] It also inhibits the human retaining β-glucosidases, GBA1, GBA2, and GBA3, displaying a rather broad activity profile across various glycoprocessing enzymes not involved in ER protein quality control. Besides Miglustat **1**, we included DNJ derivatives **2**−**8** to assess the influence of the hydrophobic *N*-alkyl substituent on antiviral activity. Compound **8** has the L-*ido* configuration and comprises the C6 epimer (glucopyranose numbering) of



	Compound	Inhibitory activity against glucosidases (biochemical assay)		Antiviral activity (CPE assay)	
		IC50 GANAB (μM) (ER-Glu-II)	IC50 GAA (μM)	EC50 (μM)	CC50 (μM)
Iminosugars	1 (Miglustat)	200	6.58 ± 2.4	113.5 ± 10.3	>200
	2	0.12 ± 0.07	0.19 ± 0.01	6 ± 0.4	>200
	3	0.15 ± 0.09	0.034 ± 0.004	3.25 ± 0.42	50
	4	0.2 ± 0.05	0.82 ± 0.18	51.7 ± 3.37	>200
	5	0.21 ± 0.07	0.64 ± 0.06	13.9 ± 8	200
	6	0.26 ± 0.1	0.47 ± 0.05	11.8 ± 0.8	>200
	7	2.12 ± 0.38	2.19 ± 5.9	22.7 ± 4.3	>200
	8	41.1	>100	104 ± 12.4	>200
1,6-cyclitols	9	561 ± 22	6.7 ± 0.34	>200	>200
	10	1.47 ± 1.45	0.2	14 ± 3.2	>200
	11	0.034 ± 0.004	0.038 ± 0.002	0.48 ± 0.1	>200
	12	>100	>100	>200	>200
	13	39.8	37.3	>200	>200
	14	8.15	2.3	16.9 ± 2.1	>200
	15	>100	>100	>100	>200
	16	>100	>100	>200	>200
	17	47	112	45.3 ± 1.52	>200
	18	>100	2.63	>100	>200
	19	>100	>100	>200	>200
1,2-cyclitols	20	11.4 ± 2.3	13	17.7 ± 1.28	>200
	21	43.8	>100	67.6 ± 14.8	>200
	22	20.3 ± 3.8	1.08	>200	>200
	23	>100	>100	64.9 ± 0.77	>200
	24	>100	>100	28.6 ± 0.21	200
	25	>100	>100	22.25 ± 4.12	>200
	26	>100	>100	91.38 ± 20.6	>200
	27	>100	>100	26.9 ± 1.3	100
	28	>100	>100	56.7 ± 1.38	>200

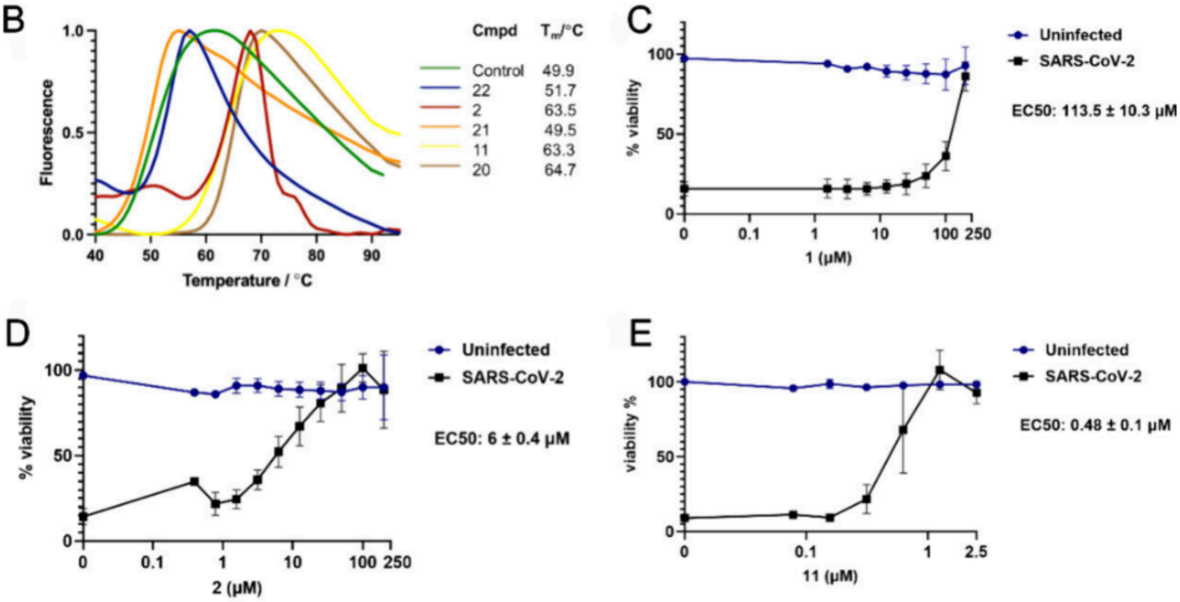


Figure 2. ER α-Glu II inhibitory potency correlates with reduction of SARS-CoV-2 mediated cytopathic effect in cell culture. (A) IC_{50} values of compounds in *in vitro* enzyme activity assays with ER α-Glu II and GAA, and EC_{50} and CC_{50} values of compounds determined by CPE reduction assays with SARS-CoV-2. (B) Thermal shift profile of preincubated ER α-Glu II with inhibitors. (C–E) SARS-CoV-2 CPE reduction assay dose–response curves of (C) Miglustat **1**, (D) naphthyl-deoxynojirimycin **2**, and (E) cyclosulfate **11**. $n = 3$ independent experiments. The viability of uninfected compound-treated cells was established by MTS assay in parallel. Mean ± SEM values are shown. The 50% inhibitory concentration (EC_{50}) values were determined by nonlinear regression with GraphPad Prism 6.

DNJ derivative **5**. Compared to **5**, L-*ido*-DNJ **8** is a much weaker ER α-Glu inhibitor, which should be reflected in its antiviral potency. With respect to the cyclitols, we previously published 1,6-*epi*-cyclophellitol **9**, 1,6-*epi*-cyclophellitol aziridine **10**, and 1,6-*epi*-cyclophellitol cyclosulfate **11** as potent and selective, mechanism-based, covalent and irreversible retaining α-glucosidase inhibitors.[35,45] Besides inhibiting ER α-Glu II, the single detected off-target (in the context of pharmacological ER protein quality control interference) is the lysosomal α-glucosidase, human acid α-glucosidase GAA. These 1,6-*epi*-cyclophellitol analogues were designed to inhibit retaining α-glucosidases exclusively (so, not inverting ones like α-Glu I), and while epoxide **9** and aziridine **10** partially inhibit the retaining β-glucosidases, GBA1 and GBA2, cyclosulfate **11** is completely inactive toward these enzymes. We also found that tempering the electrophilicity, as in cyclosulfamidates **17** and **18** and cyclosulfamide **19**, yields competitive retaining α-glucosidase inhibitors, and to investigate the effect of going from covalent to competitive inhihition within the same compound class, we included these compounds in our assays. In addition, we tested a number of structural cyclitol variations. These include 1,2-*epi*-cyclophellitols (**20**−**22**), which may block α-Glu II in a covalent, irreversible manner similar to the 1,6-*epi*-cyclophellitols.[46] A number of partially O-methylated cyclosulfates (**12**−**16**) were included to assess the effect of polarity, while compounds **23**−**28** were designed to contain alkyl substituents also present in the iminosugar series tested. The synthesis of the iminosugar and cyclitol inhibitors **1**−**11**, **17**−**22**, **25**, and **26** have been published previously.[35,41−43,45,47] The synthesis of methylated sulfates **12**−**16** and alkyl aziridines **23**, **24**, **27**, and **28** can be found in the Supporting Information (Schemes S1−S5).

The inhibitory effect of all synthesized molecules on the activity of GAA and endoplasmic reticulum α-glucosidase II (ER α-Glu II, GANAB) was determined following *in vitro* enzyme activity methods reported previously,[35] using 4-methylumbelliferyl-α-D-glucopyranoside (4-MU-α-Glc) as a fluorogenic substrate and measuring the amount of 4-MU-mediated fluorescence (Figure 2A, left panel). *N*-Alkyldeoxynojirimycins **1**−**8** all inhibited both ER α-Glu II and GAA, but with potencies varying from the nanomolar to the micromolar range. *N*-Alkyl-iminosugars **2**−**7**, featuring an extended lipophilic *N*-alkyl moiety relative to *N*-butyl-DNJ **1**, inhibited both enzymes rather more potently than this benchmark iminosugar, with **2** showing high potencies for both ER α-Glu II (IC_{50} = 0.3 ± 0.07 μM) and GAA (IC_{50} = 1.1 ± 0.09 μM). L-*ido*-Deoxynojirimycin **8** is a much weaker ER α-Glu II inhibitor than its D-*gluco* isoster **5** (both compounds containing the same adamantane-modified *N*-alkyl chain), and it showed no activity against GAA at the measured concentrations. These results match the literature trend indicating that large, hydrophobic *N*-alkyl appendages positively influence glucosidase inhibitory potency in this class of compound.[41−43,47]

With respect to the cyclitol class of compounds, 1,6-*epi*-cyclophellitol cyclosulfate **11** proved to be the most potent ER α-Glu II inhibitor of all compounds tested, with an IC_{50} value of 0.03 ± 0.007 μM. Cyclosulfate **11** was also and, together with naphthyl-iminosugar **2**, the most potent of the GAA inhibitors. Methylation of either of the four hydroxyls (or combinations thereof) in **11**, as in 1,6-*epi*-cyclophellitol cyclosulfates **12**−**16**, proved detrimental to inhibitory potency, though 4-*O*-methyl derivative **14** with IC_{50} values of 8.2 ± 0.1 μM for ER α-Glu II and 2.2 ± 0.09 μM for GAA still outperformed Miglustat (**1**) as an inhibitor of both of these enzymes. Moving from covalent (cyclosulfate, **11**) to competitive (**17**−**19**) cyclitol designs proved detrimental for ER α-Glu II inhibition, although compound **18** retains remarkable (IC_{50} = 6.1 ± 1.3 μM) inhibitory activity against GAA. 1,2-*epi*-Cyclitols **20**−**22** turned out to be only moderately active ER α-Glu II inhibitors. In contrast to the 1,6-analogues (**9**−**11**), where the cyclosulfate was more potent compared to the aziridine and epoxide, epoxide **20** was the most potent of this series.[46] Interestingly, 1,2-cyclosulfate **22** proved to be a rather potent GAA inhibitor, much more so than epoxide **20** and aziridine **21**, suggesting that conformational aspects (the epoxide and aziridine likely enforcing a half-chair conformation with respect to the cyclitol ring where the cyclosulfate will allow a chairlike conformation) are in play for this enzyme. Finally, and in contrast to what was observed for the competitive inhibitor series **1**−**8**, 1,2-cyclophellitol aziridines **23**−**28** bearing an *N*-alkyl chain (and in the case of **25** an *N*-acyl one) are much worse inhibitors for both enzymes tested (no significant inhibition up to 100 μM) when compared to the nonsubstituted aziridine **21**. In all, 1,6-*epi*-cyclophellitol cyclosulfate **11** is the most potent ER α-Glu II inhibitor, with naphthylated deoxynojirimycin **2** as the most effective of the competitive inhibitors almost on a par with **11**.

To confirm the stabilizing effect of these two compounds on the enzyme, we performed a thermal stability assay with these, as well as with the less potent inhibitors **20**−**22**, on recombinant *Mus musculus* α-Glu II, a mouse enzyme with high sequence homology to the human enzyme (Figure 2B). ER α-Glu II denaturation as a consequence of heat exposure, as well as the effect of active site-binding inhibitors on the denaturation temperature, can be monitored by a naturally quenched SYPRO orange dye. Upon denaturation of a protein, hydrophobic regions are exposed to which the dye binds, demonstrating a distinct difference in melting temperature (T_m) for each inhibitor compared to the unliganded ER α-Glu II control. *Mm*α-Glu II preincubated with compound **11** or **2** displayed melting temperatures (T_m) of 63.3 and 63.5 °C, respectively, whereas the unliganded enzyme denatured at approximately 15 °C lower (T_m = 49.9 °C). In comparison, compounds **21** and **22** gave no (49.5 °C) to marginal (51.7 °C) T_m increases, while epoxide **20**, which had the best efficacy of all 1,2-*epi*-cyclophellitols in the enzyme activity assay, gave a remarkably high T_m of 64.7 °C.

To elucidate the structure−activity relationship and predict the binding mode of the compounds before and after the covalent reaction with the nucleophilic aspartate, docking into ER α-Glu II was performed for compounds **11**, **10** and **9**. The top scoring pose of **11**, **10**, and **9** after noncovalent docking using Glide (in the Schrödinger Maestro GUI) was overlaid with the bound D-glucose molecule from the original PDB file (PDB: 5H9O) as a measure of the accuracy of the pose. The compound adopted a near-identical conformation in the binding site (Figure S1A). The ligand was also subjected to covalent docking to mimic a postreaction conformation. The outputted poses made the same hydrogen bonding interactions as the noncovalently docked pose. The top poses were overlaid with a PDB file containing a 5-fluoro-α-D-glucopyranosyl (PDB: 5HJR); the poses overlaid well in a skewed boat confirmation (Figure S1B−D), suggesting confidence in the docking results. These binding pose predictions suggest compounds **11**, **10**, and **9** are orientated correctly in the binding site of ER α-Glu II to facilitate a covalent reaction with Asp564.

All compounds were then analyzed for their antiviral activities against SARS-CoV-2, in cytopathic effect (CPE) reduction assays, in which Vero E6 cells were pretreated and infected with SARS-CoV-2 in the presence of various concentrations of a compound. Three days postinfection cell viability was measured and EC_{50} values (compound concen-

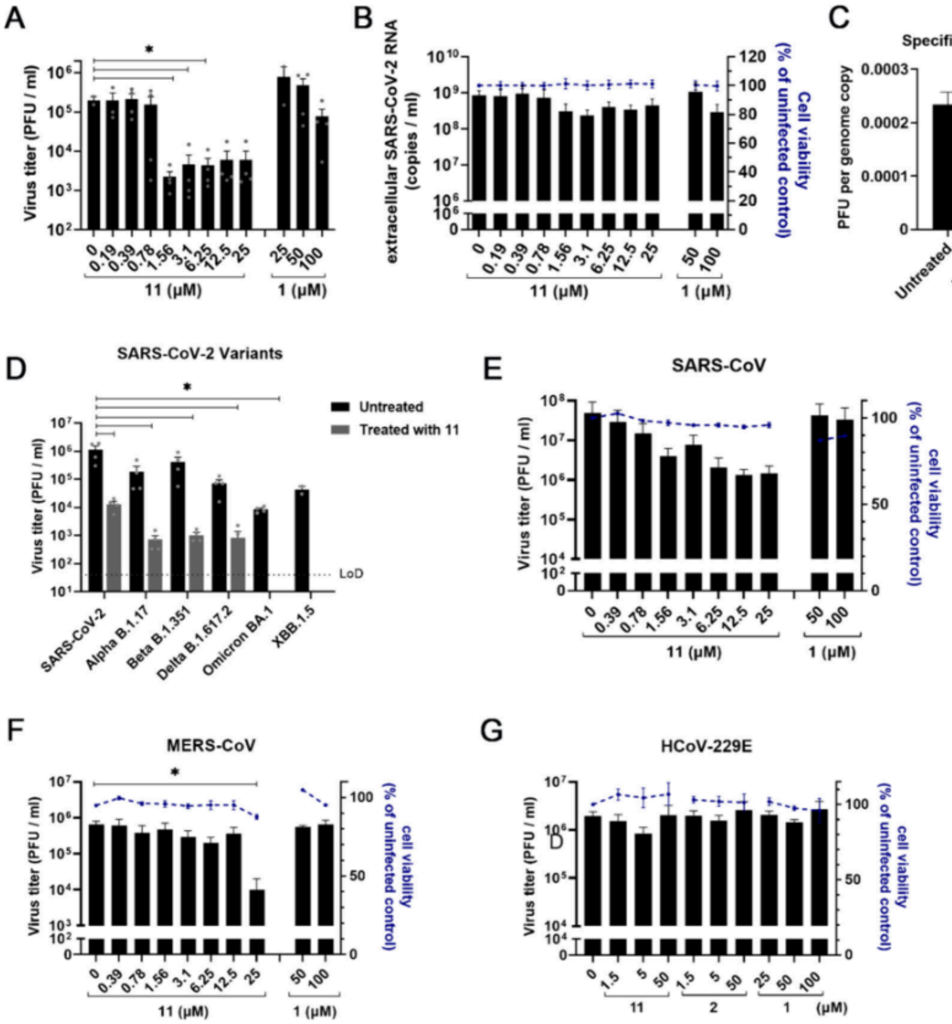


Figure 3. Spectrum of activity of 1,6-*epi*-cyclophellitol cyclosulfate **11** and iminosugars **1** and **2** against various coronaviruses. (A, B) Viral load reduction assay on H1299/ACE2 cells with SARS-CoV-2 (MOI 1) in the presence of compounds **1** or **11**. (A) Infectious virus titer and (B) extracellular viral RNA copy numbers were quantified by plaque assay and RT-qPCR, respectively. Uninfected compound-treated cells were assessed by MTS assay in parallel to measure cytotoxicity of the compounds. $n = 3$ independent experiments. Mean ± SEM values are shown. (C) The specific infectivity of treated (using 1.5 μM of compound **11**) and untreated samples was calculated by dividing the infectious virus titer (PFU/mL) by the viral RNA copy number (copies/mL). Viral load reduction assays with (D) SARS-CoV-2 variants in H1299/ACE2 cells, (E) SARS-CoV in Vero E6 cells, (F) MERS-CoV in HuH-7 cells, and (G) HCoV-229E in H1299/ACE2 cells (all with MOI 1), and treatment with **1**, **2**, or **11**. Supernatant was harvested at 16 hpi to quantify infectious progeny by plaque assay. $n = 3$ independent experiments. Uninfected compound-treated cells were measured by MTS assay in parallel to assess the cytotoxicity of the compounds. Mean ± SEM values are shown. Statistical analysis was conducted using one-way ANOVA, and significant differences are indicated by $*$, $p < 0.05$.

tration at which 50% of cell viability is reached as compared to the nontreated, infected cells) were determined (Figure 2A, right panel). Simultaneously, uninfected cells were treated with the same concentrations of compound to determine the CC_{50} (compound concentration at which cell viability is 50% of that of untreated cells due to cytotoxicity). All iminosugars **1**−**8** protected cells from SARS-CoV-2 infection in this assay, and naphthyl deoxynojirimycin **2**, being the most potent competitive ER α-Glu II inhibitor from the enzyme activity assay, also displayed the highest efficacy of the eight iminosugars assessed in blocking SARS-CoV-2 replication, with an EC_{50} value of 6 ± 0.4 μM (Figure 2D). Similar deoxynojirimycin derivatives were previously reported to have activity against SARS-CoV-2.[24,48] UV-4, an iminosugar that was previously described to be efficacious in a mouse model,[13] was tested in parallel, and its activity was compared to those of compounds **11** and **2**. The antiviral efficacy of UV-4 was similar to that of our iminosugar compound **2** (Figure S2A). In contrast, the EC_{50} value in the CPE assay for Miglustat **1** was above 100 μM (Figure 2C), which correlates to other studies which found limited antiviral activity for this compound against SARS-CoV-2.[24,49] 1,6-*epi*-Cyclophellitol cyclosulfate

11, our most potent ER α-Glu II inhibitor, also proved to be the most potent SARS-CoV-2 replication inhibitor of all compounds tested, with an EC_{50} value of 0.48 ± 0.1 μM (Figure 2E). This matches our general finding that ER α-Glu II inhibitory potency correlates with anti-SARS-CoV-2 replication efficacy (Figure 2A). Selective ER α-Glu II inhibition thus appears a promising strategy in the discovery of new antiviral agents. To validate the results obtained in the Vero E6 cell based assays, CPE reduction assays on H1299/ACE2 lung epithelial cells were performed with compounds **11**, **2**, and UV-4. With these human lung cells, comparable EC_{50} values were obtained (Figure S2B).

Given that 1,6-*epi*-cyclophellitol cyclosulfate **11** came out as the most potent compound in both the enzyme inhibition and SARS-CoV-2 CPE assays, and that this compound class, in contrast to that of iminosugars, comprises a new design class, we decided to further profile this inhibitor in more advanced virological assays to study its efficacy and mechanism of action.

1,6-*epi*-Cyclophellitol Cyclosulfate Reduces SARS-CoV-2 Infectious Progeny in Cell Culture. To investigate further the results from the CPE reduction assays, the effect of the most potent glucosidase inhibitor, 1,6-*epi*-cyclophellitol cyclosulfate **11**, was assessed in viral load reduction assays on infected H1299/ACE2 lung epithelial cells. Cells were pretreated with **11** and infected with SARS-CoV-2 at a multiplicity of infection (MOI) of 1. At 16 h postinfection (hpi) supernatant was harvested to quantify the infectious virus titer by plaque assay and extracellular viral RNA copies by RT-qPCR. Treatment of infected H1299/ACE2 lung epithelial cells with **11** resulted in a 100-fold reduction of the infectious progeny virus titer (Figure 3A). The inhibitory effect reached a plateau at 1.6 μM, and higher concentrations of **11** did not lead to more inhibition of virus replication. In contrast, Miglustat **1** reduced infectious progeny production only minimally, even at a concentration as high as 100 μM. Cyclosulfate **11** only slightly reduced extracellular viral RNA copy numbers (Figure 3B), indicating no effect on viral RNA production. This is in line with the expected mechanism of action of the compound that involves viral (structural) protein maturation, likely resulting in reduced infectivity of progeny virus. We then calculated the specific infectivity (defined as the number of infectious particles per viral RNA copy) of treated and untreated samples for the data in Figure 3A,B (Figure 3C). Treatment with compound **11** caused a decrease in specific infectivity, suggesting that the infectivity of released particles is affected. None of the treatments caused noticeable cytotoxicity in uninfected treated cells (Figure 3B). Similarly, treatment of infected Calu-3 lung epithelial cells with **11** reduced infectious progeny virus titers by ~10-fold, while no reduction in extracellular viral RNA copies was observed (Figure S3).

1,6-*epi*-Cyclophellitol Cyclosulfate Inhibits Infectious Progeny of SARS-CoV-2 Variants, SARS-CoV, and MERS-CoV, but Not HCoV-229E. To investigate the spectrum of activity against coronaviruses of 1,6-*epi*-cyclophellitol cyclosulfate **11**, its effect on the replication of SARS-CoV-2 variants Alpha, Beta, Delta, Omicron BA.1, and XBB.1.5 was tested (Figure 3D). As in the above experiments (Figure 3A), viral load reduction assays were performed, during which different cell lines were infected with the respective virus in the presence of compound, and at 16 hpi supernatant was harvested to quantify the infectious virus titer by plaque assay. Similar to the antiviral effect on the early pandemic SARS-CoV-2 isolate, treatment of H1299/ACE2 cells that were infected with other variants showed an ~100-fold reduction in infectious virus titer (Figure 3D). Viral load reduction assays with SARS-CoV on Vero E6 cells and MERS-CoV on HuH-7 cells showed a significant reduction of infectious progeny upon treatment with increasing concentrations of compound **11** (Figure 3E,F), although the efficacy of the compound was slightly lower against SARS-CoV and clearly lower against MERS-CoV. Interestingly, the viral load reduction assay with HCoV-229E on H1299/ACE2 cells did not show any reduction in virus infectivity, upon treatment with either compound **11** or **2** (Figure 3G).

1,6-*epi*-Cyclophellitol Cyclosulfate Strongly Reduces α-Glucosidase Activity and Inhibits SARS-CoV-2 in Primary Human Bronchial Epithelial Cells Cultured at the Air−Liquid Interface. We next evaluated the efficacy of 1,6-*epi*-cyclophellitol cyclosulfate **11**, in comparison to our most potent iminosugar, naphthyl-deoxynojirimycin **2**, as well as Miglustat **1** in a more advanced model of primary human bronchial epithelial cells that were cultured at the air−liquid interface (ALI-PBEC), as we described previously.[50,51] Thus, ALI-PBEC cells were infected with SARS-CoV-2 (10^5 PFU per insert; estimated MOI ~0.1) and treated with compounds on the apical side of the cells for 2 h. For uninfected controls, PBS was used instead of virus. The compounds were also present in the basal medium during the whole experiment until 48 hpi, when samples were harvested. Treatment with 0.5 μM compound **11** reduced the viral load significantly by up to 100-fold compared to the untreated control (Figure 4A). Deoxynojiriomycin derivative **2** reduced SARS-CoV-2 to similar titers, but at higher compound concentrations (10 and 100 μM), while Miglustat **1** had only a slight effect at the highest concentration measured (100 μM) (Figure 4A). Measurement of cell death (by LDH release in the supernatant) revealed that none of the compounds tested caused significant cytotoxicity at the highest concentrations (Figure 4B). We also evaluated the reduction of retaining α-glucosidases in the treated ALI-PBEC cell cultures by treatment of the cell lysate at 48 hpi with retaining α-glucosidase activity-based probe **29**, which labels GAA (isoforms at 70 and 76 kDa) and both isoforms of GANAB (~100 kDa) at pH 7[52] (Figure 4D). In line with the *in vitro* enzyme activity assay results (Figure 2A), compound **11** was most efficient in inhibiting ER α-Glu II and GAA at low concentrations (Figure 4C and Figure S4), suggesting that *in cellulo* ER α-Glu II inhibition potency correlated well with the efficacy to block SARS-CoV-2 replication.

1,6-*epi*-Cyclophellitol Cyclosulfate Inhibits SARS-CoV-2 Replication at a Postentry Step of the Viral Replication Cycle. We then investigated the mode of action of 1,6-*epi*-cyclophellitol cyclosulfate **11** by assessing which step in the viral replication cycle is inhibited. First, we assessed whether the compound affects the infectivity of virus particles, that is, has virucidal or neutralizing activity. Therefore, SARS-CoV-2 was incubated with a high concentration of compound **11** (50 μM) for 1 h at 37 °C, and subsequently the infectious virus titer was quantified by plaque assay. Control treatment with 70% ethanol led to full inactivation of the virus, while compound **11** had no effect on the infectious titer (Figure 5A). Next, we assessed if treatment early during infection had an effect on virus replication. We infected H1299/ACE2 cells with SARS-CoV-2 at an MOI 3 and started treatment with compound **11** at 1 hpi. At 2, 3, and 5 hpi, cells were harvested and RT-qPCR was performed to quantify the intracellular viral

https://doi.org/10.1021/acscentsci.4c00506
ACS Cent. Sci. XXXX, XXX, XXX−XXX

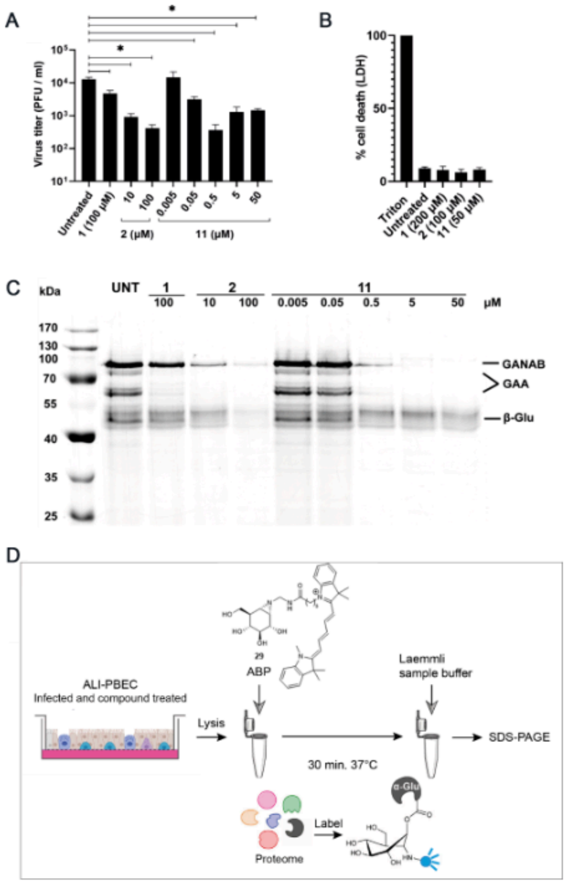


Figure 4. Reduction of SARS-CoV-2 infection in primary bronchial epithelial cells is consistent with inhibition of active ER α-glucosidase II. (A) Viral load reduction assay in ALI-PBEC. Supernatant was harvested at 48 hpi to quantify infectious progeny by plaque assay. n = 3 independent experiments. Mean ± SEM values are shown. Statistical analysis was conducted using one-way ANOVA, and significant differences are indicated by $*$, $p < 0.05$. (B) The viability of uninfected compound-treated cells was measured by LDH release assay in parallel, to assess the cytotoxicity of the compounds. Mean ± SEM values are shown. (C) Following compound treatment, cells were lysed and the lysate at pH 7.0 was treated with activity-based probe (ABP) **29** to assess cellular retaining α-glucosidase activities in a competitive activity-based protein profiling experiment. A representative gel of three independent experiments (with two biological replicates/ALI-PBEC inserts each) is shown. (D) Schematic representation of ABP labeling. Part of the figure in (D) was adapted from ref 52, and part was generated using Servier Medical Art, provided by Servier, licensed under a Creative Commons Attribution 3.0 unported license. Figure S4 shows the Gelcode Blue stained gel of (C), which demonstrated that equal amounts of protein were loaded.

genome copies. The kinetics of intracellular viral RNA accumulation were similar in untreated and compound **11** treated cells, suggesting the compound had no effect on (early) RNA replication (Figure 5B).

To evaluate whether compound **11** has an effect on host proteins (for instance, ACE2) involved in viral entry, we treated monolayers of H1299/ACE2 cells with compound **11** either 48 or 2 h before infection, during infection (0−1 h), or

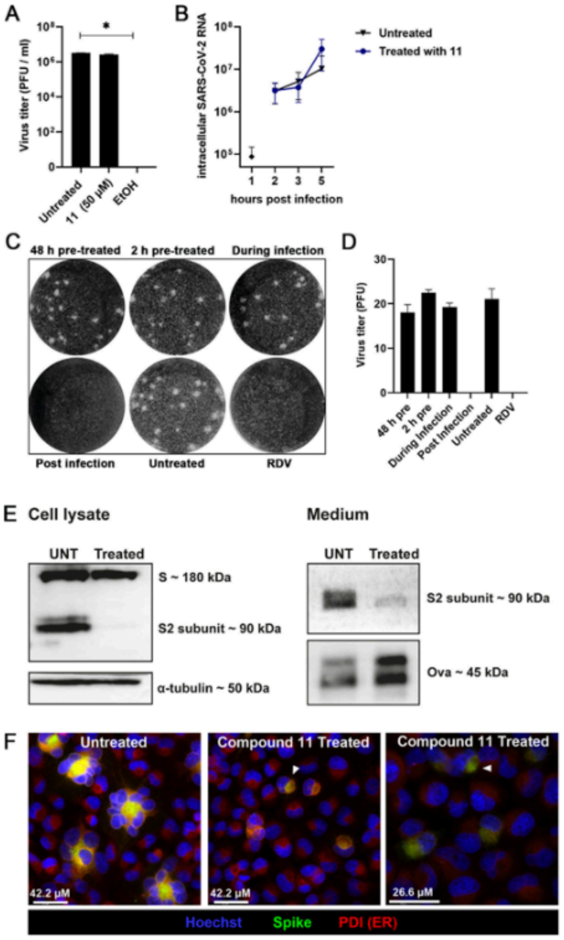


Figure 5. 1,6-*epi*-Cyclophellitol cyclosulfate **11** inhibits SARS-CoV-2 replication and syncytium formation by reducing intracellular spike protein levels and processing. (A) Virucidal activity assay in which SARS-CoV-2 was incubated with compound **11** or 70% ethanol (as control) for 1 h at RT, and (remaining) infectious progeny was quantified by plaque assay. n = 2 independent experiments. Mean ± SEM values are shown. Statistical analysis was conducted using one-way ANOVA, and significant differences are indicated by $*$, $p < 0.05$. (B) H1299/ACE2 cells were infected with SARS-CoV-2 (MOI 3) and treated with **11** from 1 hpi until harvesting at the indicated time points. Intracellular viral RNA copies were quantified by RT-qPCR. n = 3 independent experiments. (C, D) Plaque reduction assay was performed with 1 h infection and incubation for 3 days until cells were fixed and stained with crystal violet. Cells were treated with 5 μM compound **11**, either before infection (pretreatment), during infection, or after infection (postinfection) in the overlay. Treatment with RDV in the overlay was used as a control. n = 2 independent experiments. Mean ± SEM values are shown. (E) Western blot analysis of viral S protein in the medium and cell lysates of untreated (UNT) or compound **11** treated (2 μM) H1299/ACE2 cells that were infected with SARS-CoV-2 (MOI 2) and analyzed at 10 hpi using an S2-specific antibody. The medium was spiked with ovalbumin (Ova) as a recovery control and was concentrated before a sample corresponding to ~250 μL of the original medium volume was analyzed. α-Tubulin was used as a loading control for cell lysates. (F) H1299/ACE2 cells were infected with SARS-CoV-2 (MOI 0.1), fixed at 10 hpi, and the viral S protein and ER marker PDI were visualized by immunofluorescence microscopy. Cells were stained

Figure 5. continued

with human anti-SARS-CoV-2 S protein antibody (green), mouse anti-PDI antibody for ER staining (red), and Hoechst for visualizing nuclei (blue). White arrows indicate colocalization of S with PDI. Images are representative of $n = 2$ independent experiments.

starting from 1 h postinfection (hpi). The cell monolayers were infected with ∼20 PFU of SARS-CoV-2, and after 1 h the inoculum was replaced with an overlay. In one well (Post Infection) the overlay contained compound **11**. Remdesivir, a viral RNA synthesis inhibitor, was added to the overlay of another well, as a positive control for blocking virus replication in the cell. At 3 days postinfection cells were fixed and stained with crystal violet. Pretreatment of the cells with compound **11**, or treatment only during infection, had no effect on the number of plaques that developed or their morphology. Only the presence of compound **11** after infection prevented the formation of plaques, similar to treatment with remdesivir (Figure 5C,D). This result suggests that the antiviral effect of **11** is not through modulating expression or functioning of host proteins (such as the ACE2 receptor) that are essential for viral attachment to, or entry into, the host cell.

1,6-*epi*-Cyclophellitol Cyclosulfate Inhibits SARS-CoV-2 Replication through Effects on Intracellular S Protein Maturation and Infectivity of Viral Progeny. From the above-described experiments it became evident that treatment with 1,6-*epi*-cyclophellitol cyclosulfate **11** led to a reduction in virus infectivity, but not to a reduction in the number of viral genome copies (Figure 3), and that inhibition was not through an effect on the receptor or virus binding and entry, but at a postentry step other than RNA replication (Figure 5A–D). Therefore, we suspected an effect on the S protein. As shown in Figure 4, compound **11** effciently inhibited ER α-Glu II, which is crucial for the processing of N-glycosylated viral proteins such as S. To assess the effect of α-Glu II inhibition on S protein production/maturation, we performed viral load reduction assays on H1299/ACE2 cells. Cells were infected with SARS-CoV-2 (MOI 2) and treated with 2 μM compound **11** or cell culture medium. At 10 hpi medium and cell lysate were harvested to analyze S protein levels by Western blotting with an S2-specific antibody. Treatment with compound **11** led to a minor reduction in the amount of full-length S protein in the cell lysate and to the almost complete disappearance of the ∼90 kDa S2 fragment, a product of proteolytic (furin) cleavage of mature S protein in the Golgi apparatus. This indicated that treatment with **11** impaired maturation of the S protein in the ER, leading to reduced trafficking to the Golgi (Figure 5E). The amount of (processed) S2 was also strongly reduced in the medium of compound-treated cells, suggesting the compound impaired biogenesis of particles or their S protein content (Figure 5E).

Next, we set out to analyze the effect of compound **11** treatment on the level and localization of the S protein in infected cells and on the formation of syncytia, which are large multinucleated cells resulting from the interaction of S protein on the surface of infected cells with ACE2 receptors on neighboring cells, which triggers cell fusion. To this end, SARS-CoV-2-infected H1299/ACE2 cells (MOI 0.1) were treated with 5 μM compound **11** or cell culture medium as control, and at 10 hpi cells were fixed and analyzed by immunofluorescence staining for the viral S protein and the ER marker protein disulfide isomerase (Figure 5F). We observed a reduction in the amount of S protein in infected cells that were treated with compound **11** and the colocalization of S protein with the ER marker, which suggests (partial) retention of S proteins in the ER. Treatment also led to reduced syncytium formation compared to untreated infected cells, likely due to impaired maturation, and subsequent impaired trafficking of S protein to the plasma membrane.

DISCUSSION

In this study we have assessed the ER α-Glu II inhibitory potency and anti-SARS-CoV-2 activity of selected members (28 compounds in total) of two classes of glycomimetics—iminosugars and cyclitol analogues—and to what extent these two effects correlate. Deoxynojirimycin-type iminosugars as competitive inhibitors have been studied for almost four decades as candidate antivirals for pathogenic viruses that rely on ER-protein quality control and in recent years have also been explored as anti-SARS-CoV-2 agents.[14,15,18,24,53,54] In contrast, cyclophellitol-type mechanism-based inhibitors have not been considered for this purpose. The results described here support the hypothesis that mechanism-based inactivation of ER α-Glu II may lead to effective new antiviral agents to treat infections with the numerous viruses that rely on host protein glycosylation for replication. In particular, 1,6-*epi*-cyclocyclosulfate **11**, the most potent ER α-Glu II inhibitor of the tested compounds, also blocked viral replication most effectively. Although 0.5−1.6 μM doses of compound **11** reduced infectious virus titers up to 2 logs in Calu-3 cells and ALI-PBEC, higher concentrations did not lead to a further reduction and complete inhibition of virus replication was not observed at high doses. In ALI-PBEC the maximum antiviral effect was already reached at 0.5 μM, a concentration at which an almost full inhibition of ER α-Glu II was observed, suggesting that the remaining virus replication was not due to incomplete inhibition of this enzyme. Further investigations revealed that the antiviral effect is not due to effects on (glycosylation or quantity of) host cell factors that play a role in virus binding and entry into the host cell, or replication of the viral genome, suggesting it does not (noticeably) target the SARS-CoV-2 nonstructural proteins. The antiviral effect is on blocking N-glycosylation of the S protein, the most heavily N-glycosylated SARS-CoV-2 protein, which plays crucial roles in virus binding and entry. The absence of cleaved S2 fragment in compound treated cells indicates that impairing processing of S protein at the ER led to reduced trafficking of S to the Golgi and prevention of (furin) cleavage of the S1/S2 site, ultimately leading to less mature S protein for incorporation into infectious virus particles. Thus, cyclosulfate **11** acts on protein N-glycosylation/ER protein quality control, just as the *N*-alkyl deoxynojirimycin derivatives tested by us and others, but in addition compound **11** is much more selective compared to the iminosugars.[35] Considering the mechanistic mode of action of inverting and retaining glucosidases, compound **11** inhibits retaining α-glucosidases exclusively over inverting α-glucosidases, within the context of this work the lysosomal retaining α-glucosidase, GAA, as the single off-target.

Deoxynojirimycin-type iminosugars in contrast also block inverting α-glucosidases including ER α-Glu I. The finding that blocking ER α-Glu II alone is sufficient (at least in the assays reported here) for halting SARS-CoV-2 replication may therefore be beneficial for situations in which ER inhibiting α-Glu I has adverse effects. Iminosugars have often also other human glycoprocessing enzymes as off-target. *N*-Butyldeoxynojirimycin **1** (Miglustat) is applied in the clinic for the treatment of Gaucher disease, where it acts as a glucosylceramide synthase inhibitor.[44,55] It also inhibits the three human retaining *β*-glucosidases, GBA1, GBA2, and GBA3.[56] None of these enzymes plays a role in SARS-CoV-2 infections, and their inhibition may lead to adverse effects as well. Such adverse effects in contrast are not to be expected from 1,6-*epi*-cyclophellitol cyclosulfate **11**, which does not inhibit any of these enzymes (GCS, GBA1, GBA2, GBA3) as we have shown before.[35] Arguably, adverse effects such as elicited by **11** may be the result of inhibition of the lysosomal *α*-glucosidase, GAA; however, this enzyme is also inhibited by the iminosugars.[57] We therefore conclude that compound **11**, which in contrast to the iminosugars is nonbasic, thus not charged at physiological conditions, may be a good starting point for the development of new antiviral agents for the treatment of infections by SARS-CoV-2 and other (emerging) viruses that require ER-protein quality control for replication.

METHODS

Compounds and Cell Lines. Inhibitors were synthesized at the bio-organic synthesis group at the Leiden Institute of Chemistry. The synthesis of the cyclitol and iminosugar inhibitors **9–11**, **17**, **18–22**, **25**, **26**, and **1–8** have been published previously.[35,41−43,45,47] The syntheses of methylated sulfates **12–16** and alkyl aziridines **23**, **24**, **27** and **28** can be found in the Supporting Information (Schemes S1–S5). Lyophilized compounds were diluted in DMSO prior to use. Remdesivir, which was used as the compound control in different assays, was purchased from Sigma-Aldrich and dissolved in DMSO. UV-4 (SP187) was purchased from MedChemExpress and dissolved in DMSO.

Vero E6 cells and HuH-7 cells were cultured as previously described.[58] Human lung cell line H1299/ACE2 is described elsewhere.[59] These cells were cultured in Dulbecco's modified Eagle's medium with 4.5 g/L glucose with L-glutamine (DMEM; Lonza, Basel, Switzerland) supplemented with 10% fetal calf serum (FCS; CapriCorn Scientific, Ebsdorfergrund, Germany), 100 U/mL penicillin/streptomycin (P/S; Sigma-Aldrich, St. Louis, MO, USA), and 1200 μg/mL G418 for selection (InvivoGen, San Diego, CA, USA). Infections of Vero E6 cells, HuH-7 cells, and H1299/ACE2 cells were performed in Eagle's minimal essential medium with 25 mM HEPES (EMEM; Lonza) supplemented with 2% FCS, 2 mM L-glutamine (Sigma-Aldrich), and 100 U/mL P/S. Primary human bronchial epithelial cells (PBEC) were isolated and cultured as previously described.[60] All cell cultures were maintained at 37 °C in an atmosphere of 5% CO_2.

Virus Stocks. All experiments with infectious SARS-CoV, SARS-CoV-2, or MERS-CoV were performed at the LUMC biosafety level 3 facilities. The clinical isolate SARS-CoV-2/Leiden-0008 (isolated at LUMC during the first wave of the Corona pandemic in March 2020 (GenBank: MT705206.1) was used for H1299/ACE2 and ALI-PBEC infections. This virus stock was not adapted to Vero E6 cells with regard to the spike S1/S2 cleavage site (confirmed by NGS). For CPE assays in Vero E6 cells SARS-CoV-2/Leiden0002 was used (GenBank: MT510999.1). SARS-CoV-2 variant B.1.1.7 (Alpha), variant B.1.351 (Beta), and variant B.1.617 (Delta) were obtained from the University of Leuven. SARS-CoV-2 variant BA.1 (Omicron) was obtained from RIVM (strain hCoV-19/Netherlands/NH-RIVM-72291/2021, lineage B.1.1.529, GenBank: OR427989.1), and variant XBB.1.5 was isolated from a patient sample at LUMC. SARS-CoV-2/Leiden-0008 (Passage 2), SARS-CoV-2/Leiden0002, and SARS-CoV isolate Frankfurt 1[61] (Passage 4) were grown on Vero E6 cells. Alpha (Passage 4), Beta (Passage 4), Delta (Passage 4), Omicron BA.1, and XBB.1.5 (P3) variants were grown on Calu-3 cells. MERS-CoV (N3/Jordan) (GenBank: KJ614529.1) (Passage 3) and HCoV-229E were grown on HuH-7 cells. Virus titers were determined by plaque assay on Vero E6 cells, and for MERS-CoV and HCoV-229E on HuH-7 cells, as described before.[62]

***In Vitro* GAA and GANAB Enzyme Activity Assay.** Inhibition of the enzymes GAA and GANAB by the compounds was tested *in vitro* as described previously.[35] Briefly, enzymes were preincubated with a range of inhibitor concentrations for 30 min at 37 °C. The residual activities of the enzymes were then measured by adding the 4-MU−Glc substrate mixture at their corresponding optimal pHs. Reactions were quenched with 1 M NaOH−glycine (pH 10.3) upon completion, and 4-MU fluorescence was measured with an LS55 fluorescence spectrophotometer (PerkinElmer) (λ_{EX} 366 nm; λ_{EM} 445 nm). IC_{50} values reported are the mean values from three technical replicates.

Cytopathic Effect (CPE) Reduction Assay. CPE reduction assays were performed as previously described.[58] Briefly, Vero E6 cells were seeded in 96-well plates at a density of 5×10^3 cells/well. The next day, cells were infected with SARS-CoV-2/Leiden0002 in the presence of 2-fold serial dilutions of compound. Four days postinfection the CellTiter 96 aqueous nonradioactive cell proliferation kit (Promega) was used to measure the cell viability of infected (protection) and noninfected cells (assessment of cytotoxicity). EC_{50} values reported are the mean values from three independent experiments and were calculated using GraphPad Prism 6.

Expression of *Mmα*-Glu-II. The two subunits of *M. musculus* *α*-glucosidase II ganab and prkcsh were subcloned into separate vectors (pOPING and pOPINGS for ganab and prkcsh respectively) and codon optimized for mammalian expression by Genscript. Each vector was transformed into DH5*α* (Thermo Fisher) cells by heat shock. Cultures of each subunit were grown at 37 °C in LB, and the amplified DNA was purified using the PureLink HiPure plasmid filter Maxiprep kit (Invitrogen) obtaining 750 μg of DNA for both constructs. The isolated DNA was cotransfected into a 600 mL suspension of 293-F cells following the Freestyle 293 Expression system protocol (Thermo Fisher) and harvested after 4 days at 37 °C, 8% CO_2, at 135 rpm.

Purification of ER *α*-Glu-II. Cells were pelleted at 200*g*, for 20 min at 4 °C, and the clarified media was then further centrifuged for 20 min, at 5000*g* at 4 °C. The clarified media was loaded onto a pre-equilibrated 5 mL HisTrap excel column (Cytiva) with binding buffer (1× PBS, 20 mM imidazole, 5% glycerol w/v) and eluted using a buffer gradient 0–100% of elution buffer (1× PBS, 500 mM imidazole, 5% glycerol w/v) over 20 CVs. Fractions containing *Mmα*-Glu-II were concentrated and loaded onto a size exclusion S200 column (Cytiva), which was pre-equilibrated with HEPES buffer (20 mM HEPES pH 7.5 and 150 mM NaCl). The *Mmα*-Glu-II containing fractions were pooled, and a trypsinolysis was performed using sequencing grade modified trypsin (Promega), supplemented with 2 mM $CaCl_2$ for 4 h at a ratio of 1:100 (trypsin:*Mmα*-Glu-II). The size exclusion was repeated, and the resulting *Mmα*-Glu-II was pooled and concentrated to 8 mg/mL.

Thermal Shift Assays. Triplicate reactions of 10 μM *Mma*-Glu-II unliganded control and 10 μM *Mma*-Glu-II with 50 μM inhibitor were prepared to a final volume of 30 μL with buffer (20 mM HEPES pH 7.5 and 150 mM NaCl). Before the assay, 20× SYPRO orange dye was added to each reaction mixture. The assay was performed using the Stratagene Mx3005P qPCR machine, where the SYPRO orange dye was excited at λ_{EX} 517 nm and monitored at 585 nm with 2 °C min^{-1} increases from 25 to 95 °C. Readings were averaged to produce a thermal stability curve with fluorescence plotted against temperature and the T_m estimated from the midpoint.

Viral Load Reduction Assays. For SARS-CoV-2 (variants) and HCoV-229E infections, H1299/ACE2 cells were seeded in 96-well plates at a density of 10^4 cells/well and the next day infected at MOI 1. Infections with SARS-CoV-2 were incubated at 37 °C, and infections with HCoV-229E were incubated at 33 °C. For SARS-CoV or MERS-CoV infections (MOI 1), Vero E6 or HuH-7 cells were seeded in 96-well plates at a density of 2×10^4 cells/well. Cells were incubated at 37 °C. After removal of the inoculum at 1 hpi, cells were washed three times with warm PBS or medium, after which they were incubated in infection medium (EMEM). Supernatant samples were harvested at 16 hpi, and infectious virus titers were determined by plaque assay and viral RNA copy numbers by RT-qPCR. In parallel, the cytotoxicity of compound treatment was measured on uninfected cells by the CellTiter 96 aqueous nonradioactive cell proliferation kit.

Immunofluorescence Staining. For immunofluorescence imaging of viral spike protein, H1299/ACE2 cells were seeded onto glass coverslips in 24-well plates at a density of 1.6×10^5 cells/well. The next day they were infected with SARS-CoV-2/Leiden0008 (MOI 0.1) in Opti-MEM reduced serum medium (Thermo Fisher Scientific). At 16 hpi, cells were fixed with 3% warm paraformaldehyde. Immunofluorescent staining of viral spike protein was done using human antispike antibody P52 (gift from King's College) and goat-α-human IgG Alexa 488 antibody (Thermo Fisher Scientific). Staining of endoplasmic reticulum was done using mouse anti-PDI antibody (Fuller)[63] and donkey-α-mouse Cy3 antibody (Jackson).

Western Blot. For Western blot analysis, H1299/ACE2 cells were seeded in 6-well plates at a density of 6.5×10^5 cells/well and the next day infected with SARS-CoV-2/Leiden0008 at an MOI of 2. At 10 hpi supernatant was harvested and 4000 μL of medium was spiked with ovalbumin (internal recovery control) and concentrated to 150 μL using Amicon Ultra-0.5 centrifugal filter units (Merck), according to the manufacturer's instruction. An equal amount of Laemmli buffer was added, and samples were heated at 95 °C for 5 min. Samples were analyzed by SDS-PAGE (10% gel, 30 min at 90 V, then 50 min at 120 V) and subsequently blotted for 30 min in a semidry blotting system (Bio-Rad). The membrane was blocked with 1% casein in PBST for 1 h at RT, before incubation with primary antibodies overnight at 4 °C. Spike proteins were detected using SARS/SARS-CoV-2 spike protein S2-specific mAb 1A9 (Invitrogen) as the primary antibody. The loading control tubulin was detected with mouse-anti-α-tubulin antibody B-5-1-2 (abcam), and spiked ovalbumin was detected with mouse ovalbumin mAb 1D3D5 (Thermo Fisher). The next day the membrane was washed three times for 5 min with PBST and then incubated in 0.5% casein in PBST with a secondary goat-α-mouse-HRP antibody (P0447, Dako) for 1 h at RT. After washing again three times, the membrane was incubated in Clarity Western ECL Substrate (Bio-Rad) for 2 min and imaged with the Uvitec Alliance Q9 advanced imager.

RNA Isolation and RT-qPCR. RNA was isolated by magnetic bead isolation, as described in ref 51. Equine arteritis virus (EAV) in AVL lysis buffer (Qiagen) was spiked into the isolation reagent as an internal control for extracellular RNA samples. RT-qPCR was performed using TaqMan Fast Virus 1-step master mix (Thermo Fisher Scientific) and as previously described.[64] The cellular reference gene PGK1 served as a control for intracellular RNA. Primers and probes for EAV and PGK1 and the normalization procedure were described before.[62] Primers and probes for SARS-CoV-2, as well as a standard curve, were used as described previously.[64,65]

Plaque Assay. To quantify infectious virus titers, plaque assays were done on Vero E6 cells (SARS-CoV-2 and variants, SARS-CoV), H1299/ACE2 (HCoV-229E), or HuH-7 (MERS-CoV). For SARS-CoV-2 and variants, 2×10^4 cells/well were seeded in a 96-well plate, and serial dilutions of samples were inoculated for 1 h at 37 °C on a rocking platform. Inoculums were removed and 100 μL of methylcellulose overlay was added. Cells were incubated for 4 days until fixation and crystal violet staining. Alternatively, plaque assays for SARS-CoV-2 (variants) were done in 6-well plates, with Avicel overlay and 3 days of incubation. HCoV-229E samples were quantified in 12-well plates, using Avicel overlay and incubating for 4 days. MERS-CoV samples were quantified in 12-well plates with Avicel overlay or 96-well plates with methylcellulose overlay for 3 days.

Infection of ALI-PBEC and Activity-Based Probe Labeling. ALI-PBEC were pretreated with compound in the basal medium for 3 h. Cells were infected with 100 000 PFU of SARS-CoV-2/Leiden0008 per insert (estimated MOI 0.1) with compounds present in the inoculum. After 2 h at 37 °C on a rocking platform, the inoculum was removed and cells were washed three times with warm PBS. Compounds stayed present in the basal medium until 48 h postinfection. At 48 hpi the viral load was determined by plaque assay on a 200 μL apical wash (PBS incubated on the apical side of the insets for 10 min at 37 °C). For assessing cytotoxicity with the CyQuant LDH cytotoxicity assay (Thermo Fisher Scientific), 10 μL of apical wash was diluted 5 times with 40 μL of PBS. A 25 μL volume of this dilution was added to 25 μL of assay reagent and incubated for 30 min at RT in the dark. The plate was fixed and measured at a wavelength of 490 nM (Envision reader, PerkinElmer). For the activity-based probe labeling, the insets were washed one more time with PBS and processed as described previously.[52] Briefly, cells were lysed with 60 μL of potassium phosphate buffer per insert. A fluorescently labeled probe (JJB383) was diluted in McIlvaine buffer (pH 7) to a 10 μM stock and incubated for 5 min on ice. For labeling of the cell lysate, 10 μL of lysate was added to 10 μL of McIlvaine buffer and 5 μL of probe. The lysate was incubated for 30 min at 37 °C before the addition of 10 μL of Laemmli sample buffer (4×). Samples were heated at 95 °C for 5 min and separated in a 10% SDS-PAGE gel. Fluorescence was measured at a wavelength of 625 nm (Cy5) with a Uvitec Alliance Q9 imager (BioSPX). After imaging, the gels were stained with GelCode Blue stain reagent (Thermo Fisher Scientific) and visualized using a Uvitec Essential V6 system to check for equal loading.

Plaque Reduction Assay. H1299/ACE2 cells were seeded in a 6-well plate at a density of 1.3×10^5 cells/well (20% confluency), 96 h prior to infection. Cells were treated with 5 μM compound **11** either 48 or 2 h before infection, or during the 1 h infection in the inoculum. The monolayers were infected with ~20 PFU of SARS-CoV-2/Leiden0008. In the postinfection treatment, the compound (or RDV) was added to the Avicel overlay. Cells were incubated for 4 days at 37 °C before fixation and crystal violet staining.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acscentsci.4c00506.

> Additional experimental details including synthesis of the methylated sulfamidates **12**−**16** and the alkyl aziridines **23**, **24**, **27**, and **28**; additional results including SARS-CoV-2 viral load reduction assay on Calu-3 lung epithelial cells, SARS-CoV viral load reduction assay on Vero E6 cells, GelCode Blue staining of SDS-PAGE gel of activity-based protein profiling experiment, SDS-PAGE gel of activity-based protein profiling at pH 4 (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Martijn J. van Hemert − *Leiden University Center for Infectious Diseases (LUCID), Leiden University Medical Center, 2333 ZA Leiden, The Netherlands;* orcid.org/0000-0002-2617-9243; Email: m.j.van_hemert@lumc.nl

Herman S. Overkleeft − *Leiden Institute of Chemistry, Leiden University, 2311 EZ Leiden, The Netherlands;* orcid.org/0000-0001-6976-7005; Email: h.s.overkleeft@chem.leidenuniv.nl

Authors

Melissa Thaler − *Leiden University Center for Infectious Diseases (LUCID), Leiden University Medical Center, 2333 ZA Leiden, The Netherlands*

Tim P. Ofman − *Leiden Institute of Chemistry, Leiden University, 2311 EZ Leiden, The Netherlands*

Ken Kok − *Leiden Institute of Chemistry, Leiden University, 2311 EZ Leiden, The Netherlands*

Jurriaan J. A. Heming − *Leiden Institute of Chemistry, Leiden University, 2311 EZ Leiden, The Netherlands*

Elisha Moran − *Department of Chemistry, University of York, York YO10 5DD, United Kingdom;* orcid.org/0000-0002-3463-6862

Isabelle Pickles − *Department of Chemistry, University of York, York YO10 5DD, United Kingdom;* orcid.org/0000-0002-8495-5570

Anouk A. Leijs − *Leiden University Center for Infectious Diseases (LUCID), Leiden University Medical Center, 2333 ZA Leiden, The Netherlands*

Adrianus M. C. H. van den Nieuwendijk − *Leiden Institute of Chemistry, Leiden University, 2311 EZ Leiden, The Netherlands*

Richard J. B. H. N. van den Berg − *Leiden Institute of Chemistry, Leiden University, 2311 EZ Leiden, The Netherlands*

Gijs Ruijgrok − *Leiden Institute of Chemistry, Leiden University, 2311 EZ Leiden, The Netherlands*

Zachary Armstrong − *Leiden Institute of Chemistry, Leiden University, 2311 EZ Leiden, The Netherlands;* orcid.org/0000-0002-4086-2946

Clarisse Salgado-Benvindo − *Leiden University Center for Infectious Diseases (LUCID), Leiden University Medical Center, 2333 ZA Leiden, The Netherlands*

Dennis K. Ninaber − *Department of Pulmonology, Leiden University Medical Center, 2333 ZA Leiden, The Netherlands*

Eric J. Snijder − *Leiden University Center for Infectious Diseases (LUCID), Leiden University Medical Center, 2333 ZA Leiden, The Netherlands*

Constant A. A. van Boeckel − *Leiden Institute of Chemistry, Leiden University, 2311 EZ Leiden, The Netherlands*

Marta Artola − *Leiden Institute of Chemistry, Leiden University, 2311 EZ Leiden, The Netherlands;* orcid.org/0000-0002-3051-3902

Gideon J. Davies − *Department of Chemistry, University of York, York YO10 5DD, United Kingdom;* orcid.org/0000-0002-7343-776X

Complete contact information is available at: https://pubs.acs.org/10.1021/acscentsci.4c00506

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors are grateful for funding from the European Research Council (ERC-2020-SyG-951231 Carbocentre, to G.J.D. and H.S.O.). G.J.D. is funded by the Royal Society Ken Murray Research Professorship. Z.A. thanks the NWO for support through the Veni grant: VI.Veni.212.173. C.S.-B. was supported by the Coordination for the Improvement of Higher Education Personnel (CAPES) (Process No. 88881.171440/2018-01), Ministry of Education, Brazil. Figure [1] was created with BioRender.com. The graphic for the table of contents was partly generated using Servier Medical Art, provided by Servier, licensed under a Creative 995 Commons Attribution 3.0 unported license.

■ REFERENCES

(1) Feng, T.; Zhang, J.; Chen, Z.; Pan, W.; Chen, Z.; Yan, Y.; Dai, J. Glycosylation of viral proteins: Implication in virus-host interaction and virulence. *Virulence* **2022**, *13*, 670−683.

(2) O'Keefe, S.; Roebuck, Q. P.; Nakagome, I.; Hirono, S.; Kato, A.; Nash, R.; High, S. Characterizing the selectivity of ER α-glucosidase inhibitors. *Glycobiology* **2019**, *29*, 530−542.

(3) Hebert, D. N.; Foellmer, B.; Helenius, A. Glucose trimming and reglucosylation determine glycoprotein association with calnexin in the endoplasmic reticulum. *Cell* **1995**, *81*, 425−33.

(4) Kostova, Z.; Wolf, D. H. For whom the bell tolls: protein quality control of the endoplasmic reticulum and the ubiquitin-proteasome connection. *EMBO J.* **2003**, *22*, 2309−2317.

(5) Gruters, R. A.; Neefjes, J. J.; Tersmette, M.; de Goede, R. E.; Tulp, A.; Huisman, H. G.; Miedema, F.; Ploegh, H. L. Interference with HIV-induced syncytium formation and viral infectivity by inhibitors of trimming glucosidase. *Nature* **1987**, *330*, 74−7.

(6) Fuhrmann, U.; Bause, E.; Ploegh, H. Inhibitors of oligosaccharide processing. *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression* **1985**, *825*, 95−110.

(7) Chang, J.; Block, T. M.; Guo, J.-T. Antiviral therapies targeting host ER alpha-glucosidases: Current status and future directions. *Antiviral Res.* **2013**, *99*, 251−260.

(8) Koshland, D. E. Stereochemistry and the mechanism of enzymatic reactions. *Biol. Rev.* **1953**, *28*, 416−436.

(9) Koshland, D. E. Application of a Theory of Enzyme Specificity to Protein Synthesis. *Proc. Natl. Acad. Sci. U. S. A.* **1958**, *44*, 98−104.

(10) Walker, B. D.; Kowalski, M.; Goh, W. C.; Kozarsky, K.; Krieger, M.; Rosen, C.; Rohrschneider, L.; Haseltine, W. A.; Sodroski, J. Inhibition of human immunodeficiency virus syncytium formation and virus replication by castanospermine. *Proc. Natl. Acad. Sci. U. S. A.* **1987**, *84*, 8120−4.

(11) Fleet, G. W.; Karpas, A.; Dwek, R. A.; Fellows, L. E.; Tyms, A. S.; Petursson, S.; Namgoong, S. K.; Ramsden, N. G.; Smith, P. W.; Son, J. C.; et al. Inhibition of HIV replication by amino-sugar derivatives. *FEBS Lett.* **1988**, *237*, 128−32.

(12) Montagnier, L.; Clavel, F.; Krust, B.; Chamaret, S.; Rey, F.; Barré-Sinoussi, F.; Chermann, J. C. Identification and antigenicity of the major envelope glycoprotein of lymphadenopathy-associated virus. *Virology* **1985**, *144*, 283−9.

(13) Warfield, K. L.; Barnard, D. L.; Enterlein, S. G.; Smee, D. F.; Khaliq, M.; Sampath, A.; Callahan, M. V.; Ramstedt, U.; Day, C. W. The Iminosugar UV-4 is a Broad Inhibitor of Influenza A and B Viruses ex Vivo and in Mice. *Viruses* **2016**, *8*, 71.

(14) Stavale, E. J.; Vu, H.; Sampath, A.; Ramstedt, U.; Warfield, K. L. In vivo therapeutic protection against influenza A (H1N1) oseltamivir-sensitive and resistant viruses by the iminosugar UV-4. *PLoS One* **2015**, *10*, No. e0121662.

(15) Warfield, K. L.; Alonzi, D. S.; Hill, J. C.; Caputo, A. T.; Roversi, P.; Kiappes, J. L.; Sheets, N.; Duchars, M.; Dwek, R. A.; Biggins, J.; Barnard, D.; Shresta, S.; Treston, A. M.; Zitzmann, N. Targeting Endoplasmic Reticulum α-Glucosidase I with a Single-Dose Iminosugar Treatment Protects against Lethal Influenza and Dengue Virus Infections. *J. Med. Chem.* **2020**, *63*, 4205−4214.

(16) Fukushi, M.; Yoshinaka, Y.; Matsuoka, Y.; Hatakeyama, S.; Ishizaka, Y.; Kirikae, T.; Sasazuki, T.; Miyoshi-Akiyama, T. Monitoring of S Protein Maturation in the Endoplasmic Reticulum by Calnexin Is Important for the Infectivity of Severe Acute Respiratory Syndrome Coronavirus. *J. Virol.* **2012**, *86*, 11745−11753.

(17) Chang, J.; Warren, T. K.; Zhao, X.; Gill, T.; Guo, F.; Wang, L.; Comunale, M. A.; Du, Y.; Alonzi, D. S.; Yu, W.; Ye, H.; Liu, F.; Guo, J.-T.; Mehta, A.; Cuconati, A.; Butters, T. D.; Bavari, S.; Xu, X.; Block, T. M. Small molecule inhibitors of ER α-glucosidases are active against multiple hemorrhagic fever viruses. *Antiviral Res.* **2013**, *98*, 432−440.

(18) Perry, S. T.; Buck, M. D.; Plummer, E. M.; Penmasta, R. A.; Batra, H.; Stavale, E. J.; Warfield, K. L.; Dwek, R. A.; Butters, T. D.; Alonzi, D. S.; Lada, S. M.; King, K.; Klose, B.; Ramstedt, U.; Shresta, S. An iminosugar with potent inhibition of dengue virus infection in vivo. *Antiviral Res.* **2013**, *98*, 35−43.

(19) Sadat, M. A.; Moir, S.; Chun, T.-W.; Lusso, P.; Kaplan, G.; Wolfe, L.; Memoli, M. J.; He, M.; Vega, H.; Kim, L. J. Y.; Huang, Y.; Hussein, N.; Nievas, E.; Mitchell, R.; Garofalo, M.; Louie, A.; Ireland, D. C.; Grunes, C.; Cimbro, R.; Patel, V.; Holzapfel, G.; Salahuddin, D.; Bristol, T.; Adams, D.; Marciano, B. E.; Hegde, M.; Li, Y.; Calvo, K. R.; Stoddard, J.; Justement, J. S.; Jacques, J.; Long Priel, D. A.; Murray, D.; Sun, P.; Kuhns, D. B.; Boerkoel, C. F.; Chiorini, J. A.; Di Pasquale, G.; Verthelyi, D.; Rosenzweig, S. D. Glycosylation, Hypogammaglobulinemia, and Resistance to Viral Infections. *N. Engl. J. Med.* **2014**, *370*, 1615−1625.

(20) Sung, C.; Wei, Y.; Watanabe, S.; Lee, H. S.; Khoo, Y. M.; Fan, L.; Rathore, A. P.; Chan, K. W.; Choy, M. M.; Kamaraj, U. S.; Sessions, O. M.; Aw, P.; de Sessions, P. F.; Lee, B.; Connolly, J. E.; Hibberd, M. L.; Vijaykrishna, D.; Wijaya, L.; Ooi, E. E.; Low, J. G.; Vasudevan, S. G. Extended Evaluation of Virological, Immunological and Pharmacokinetic Endpoints of CELADEN: A Randomized, Placebo-Controlled Trial of Celgosivir in Dengue Fever Patients. *PLoS Negl Trop Dis* **2016**, *10*, No. e0004851.

(21) Durantel, D. Celgosivir, an alpha-glucosidase I inhibitor for the potential treatment of HCV infection. *Curr. Opin. Invest. Drugs* **2009**, *10*, 860−870.

(22) Clarke, E. C.; Nofchissey, R. A.; Ye, C.; Bradfute, S. B. The iminosugars celgosivir, castanospermine and UV-4 inhibit SARS-CoV-2 replication. *Glycobiology* **2021**, *31*, 378−384.

(23) Karade, S. S.; Hill, M. L.; Kiappes, J. L.; Manne, R.; Aakula, B.; Zitzmann, N.; Warfield, K. L.; Treston, A. M.; Mariuzza, R. A. N-Substituted Valiolamine Derivatives as Potent Inhibitors of Endoplasmic Reticulum α-Glucosidases I and II with Antiviral Activity. *J. Med. Chem.* **2021**, *64*, 18010−18024.

(24) Ferjancic, Z.; Bihelovic, F.; Vulovic, B.; Matovic, R.; Trmcic, M.; Jankovic, A.; Pavlovic, M.; Djurkovic, F.; Prodanovic, R.; Djurdjevic Djelmas, A.; Kalicanin, N.; Zlatovic, M.; Sladic, D.; Vallet, T.; Vignuzzi, M.; Saicic, R. N. Development of iminosugar-based glycosidase inhibitors as drug candidates for SARS-CoV-2 virus via molecular modelling and in vitro studies. *J. Enzyme Inhib. Med. Chem.* **2024**, *39*, 2289007.

(25) Gong, Y.; Qin, S.; Dai, L.; Tian, Z. The glycosylation in SARS-CoV-2 and its receptor ACE2. *Signal Transduction and Targeted Therapy* **2021**, *6*, 396.

(26) Grant, O. C.; Montgomery, D.; Ito, K.; Woods, R. J. Analysis of the SARS-CoV-2 spike protein glycan shield reveals implications for immune recognition. *Sci. Rep.* **2020**, *10*, 14991.

(27) Jackson, C. B.; Farzan, M.; Chen, B.; Choe, H. Mechanisms of SARS-CoV-2 entry into cells. *Nat. Rev. Mol. Cell Biol.* **2022**, *23*, 3−20.

(28) Bouwman, K. M.; Tomris, I.; Turner, H. L.; van der Woude, R.; Shamorkina, T. M.; Bosman, G. P.; Rockx, B.; Herfst, S.; Snijder, J.; Haagmans, B. L.; Ward, A. B.; Boons, G.-J.; de Vries, R. P. Multimerization- and glycosylation-dependent receptor binding of SARS-CoV-2 spike proteins. *PLOS Pathogens* **2021**, *17*, No. e1009282.

(29) Henderson, R.; Edwards, R. J.; Mansouri, K.; Janowska, K.; Stalls, V.; Kopp, M.; Haynes, B. F.; Acharya, P. Glycans on the SARS-CoV-2 Spike Control the Receptor Binding Domain Conformation. *bioRxiv*, June 30, 2020. DOI: 10.1101/2020.06.26.173765.

(30) Huang, H.-C.; Lai, Y.-J.; Liao, C.-C.; Wang, F.-Y.; Huang, K.-B.; Lee, I. J.; Chou, W.-C.; Wang, S.-H.; Wang, L.-H.; Hsu, J.-M.; Sun, C.-P.; Kuo, C.-T.; Wang, J.; Hsiao, T.-C.; Yang, P.-J.; Lee, T.-A.; Huang, W.; Li, F.-A.; Shen, C.-Y.; Lin, Y.-L.; Tao, M.-H.; Li, C.-W. Targeting conserved N-glycosylation blocks SARS-CoV-2 variant infection in vitro. *eBioMedicine* **2021**, *74*, 103712.

(31) Casalino, L.; Gaieb, Z.; Goldsmith, J. A.; Hjorth, C. K.; Dommer, A. C.; Harbison, A. M.; Fogarty, C. A.; Barros, E. P.; Taylor, B. C.; McLellan, J. S.; Fadda, E.; Amaro, R. E. Beyond Shielding: The Roles of Glycans in the SARS-CoV-2 Spike Protein. *ACS Cent Sci.* **2020**, *6*, 1722−1734.

(32) Casas-Sanchez, A.; Romero-Ramirez, A.; Hargreaves, E.; Ellis, C. C.; Grajeda, B. I.; Estevao, I. L.; Patterson, E. I.; Hughes, G. L.; Almeida, I. C.; Zech, T.; Acosta-Serrano, Á. Inhibition of Protein N-Glycosylation Blocks SARS-CoV-2 Infection. *mBio* **2022**, *13*, No. e03718-21.

(33) Watanabe, S.; Chan, KW-K; Dow, G.; Ooi, E. E.; Low, J. G.; Vasudevan, S. G. Optimizing celgosivir therapy in mouse models of dengue virus infection of serotypes 1 and 2: The search for a window for potential therapeutic efficacy. *Antiviral Res.* **2016**, *127*, 10−19.

(34) Hoechst Marion Roussel. A randomized, double-blind active-controlled, dose-ranging study of safety and efficacy of chronically administered MDL 28,574A in the treatment of HIV-infected patients. *ClinicalTrials.gov*, June 24, 2005. NLM identifier: NCT00002151. https://clinicaltrials.gov/study/NCT00002151.

(35) Artola, M.; Wu, L.; Ferraz, M. J.; Kuo, C.-L.; Raich, L.; Breen, I. Z.; Offen, W. A.; Codée, J. D. C.; van der Marel, G. A.; Rovira, C.; Aerts, J. M. F. G.; Davies, G. J.; Overkleeft, H. S. 1,6-Cyclophellitol Cyclosulfates: A New Class of Irreversible Glycosidase Inhibitor. *ACS Central Science* **2017**, *3*, 784−793.

(36) Schröder, S. P.; Petracca, R.; Minnee, H.; Artola, M.; Aerts, J. M. F. G.; Codée, J. D. C.; van der Marel, G. A.; Overkleeft, H. S. A Divergent Synthesis of l-arabino- and d-xylo-Configured Cyclophellitol Epoxides and Aziridines. *Eur. J. Org. Chem.* **2016**, *2016*, 4787−4794.

(37) Jiang, J.; Artola, M.; Beenakker, T. J. M.; Schröder, S. P.; Petracca, R.; de Boer, C.; Aerts, J. M. F. G.; van der Marel, G. A.; Codée, J. D. C.; Overkleeft, H. S. The Synthesis of Cyclophellitol-Aziridine and Its Configurational and Functional Isomers. *Eur. J. Org. Chem.* **2016**, *2016*, 3671−3678.

(38) Willems, L. I.; Jiang, J.; Li, K.-Y.; Witte, M. D.; Kallemeijn, W. W.; Beenakker, T. J. N.; Schröder, S. P.; Aerts, J. M. F. G.; van der Marel, G. A.; Codée, J. D. C.; Overkleeft, H. S. From Covalent Glycosidase Inhibitors to Activity-Based Glycosidase Probes. *Chem. Eur. J.* **2014**, *20*, 10864−10872.

(39) de Boer, C.; McGregor, N. G. S.; Peterse, E.; Schröder, S. P.; Florea, B. I.; Jiang, J.; Reijngoud, J.; Ram, A. F. J.; van Wezel, G. P.; van der Marel, G. A.; Codée, J. D. C.; Overkleeft, H. S.; Davies, G. J. Glycosylated cyclophellitol-derived activity-based probes and inhibitors for cellulases. *RSC Chemical Biology* **2020**, *1*, 148−155.

(40) Schröder, S. P.; van de Sande, J. W.; Kallemeijn, W. W.; Kuo, C.-L.; Artola, M.; van Rooden, E. J.; Jiang, J.; Beenakker, T. J. M.; Florea, B. I.; Offen, W. A.; Davies, G. J.; Minnaard, A. J.; Aerts, J. M. F. G.; Codée, J. D. C.; van der Marel, G. A.; Overkleeft, H. S. Towards broad spectrum activity-based glycosidase probes: synthesis and evaluation of deoxygenated cyclophellitol aziridines. *Chem. Commun.* **2017**, *53*, 12528−12531.

(41) Wennekes, T.; Meijer, A. J.; Groen, A. K.; Boot, R. G.; Groener, J. E.; van Eijk, M.; Ottenhoff, R.; Bijl, N.; Ghauharali, K.; Song, H.; O'Shea, T. J.; Liu, H.; Yew, N.; Copeland, D.; van den Berg, R. J.; van der Marel, G. A.; Overkleeft, H. S.; Aerts, J. M. Dual-Action Lipophilic Iminosugar Improves Glycemic Control in Obese Rodents by Reduction of Visceral Glycosphingolipids and Buffering of Carbohydrate Assimilation. *J. Med. Chem.* **2010**, *53*, 689−698.

(42) Ghisaidoobe, A.; Bikker, P.; de Bruijn, A. C. J.; Godschalk, F. D.; Rogaar, E.; Guijt, M. C.; Hagens, P.; Halma, J. M.; van't Hart, S. M.; Luitjens, S. B.; van Rixel, V. H. S.; Wijzenbroek, M.; Zweegers, T.; Donker-Koopman, W. E.; Strijland, A.; Boot, R.; van der Marel, G.; Overkleeft, H. S.; Aerts, J. M. F. G.; van den Berg, R. J. B. H. N. Identification of Potent and Selective Glucosylceramide Synthase Inhibitors from a Library of N-Alkylated Iminosugars. *ACS Med. Chem. Lett.* **2011**, *2*, 119−123.

(43) Ghisaidoobe, A. T.; van den Berg, R. J. B. H. N; Butt, S. S.; Strijland, A.; Donker-Koopman, W. E.; Scheij, S.; van den Nieuwendijk, A. M. C. H.; Koomen, G.-J.; van Loevezijn, A.; Leemhuis, M.; Wennekes, T.; van der Stelt, M.; van der Marel, G. A.; van Boeckel, C. A. A.; Aerts, J. M. F. G.; Overkleeft, H. S. Identification and Development of Biphenyl Substituted Iminosugars as Improved Dual Glucosylceramide Synthase/Neutral Glucosylceramidase Inhibitors. *J. Med. Chem.* **2014**, *57*, 9096−9104.

(44) Ficicioglu, C. Review of miglustat for clinical management in Gaucher disease type 1. *Ther Clin Risk Manag* **2008**, *4*, 425−31.

(45) Kok, K.; Kuo, C.-L.; Katzy, R. E.; Lelieveld, L. T.; Wu, L.; Roig-Zamboni, V.; van der Marel, G. A.; Codée, J. D. C.; Sulzenbacher, G.; Davies, G. J.; Overkleeft, H. S.; Aerts, J. M. F. G.; Artola, M. 1,6-epi-Cyclophellitol Cyclosulfamidate Is a Bona Fide Lysosomal α-Glucosidase Stabilizer for the Treatment of Pompe Disease. *J. Am. Chem. Soc.* **2022**, *144*, 14819−14827.

(46) Ofman, T. P.; Heming, J. J. A.; Nin-Hill, A.; Küllmer, F.; Steneker, R.; Klein, A.; Moran, E.; Bennett, M.; Ruijgrok, G.; Kok, K.; Armstrong, Z. W. B.; Aerts, J. M. F. G.; van der Marel, G. A.; Rovira, C.; Davies, G. J.; Artola, M.; Codée, J. D. C.; Overkleeft, H. S. Conformational and Electronic Variations in 1,2- and 1,6-Cyclophellitols and Their Impact on Retaining α-Glucosidase Inhibition. *Chem. Eur. J.* **2024**, *30*, e202400723.

(47) Lahav, D.; Liu, B.; van den Berg, R. J. B. H. N.; van den Nieuwendijk, A. M. C. H.; Wennekes, T.; Ghisaidoobe, A. T.; Breen, I.; Ferraz, M. J.; Kuo, C.-L.; Wu, L.; Geurink, P. P.; Ovaa, H.; van der Marel, G. A.; van der Stelt, M.; Boot, R. G.; Davies, G. J.; Aerts, J. M. F. G.; Overkleeft, H. S. A Fluorescence Polarization Activity-Based Protein Profiling Assay in the Discovery of Potent, Selective Inhibitors for Human Nonlysosomal Glucosylceramidase. *J. Am. Chem. Soc.* **2017**, *139*, 14192−14197.

(48) Karade, S. S.; Franco, E. J.; Rojas, A. C.; Hanrahan, K. C.; Kolesnikov, A.; Yu, W.; MacKerell, A. D., Jr; Hill, D. C.; Weber, D. J.; Brown, A. N.; Treston, A. M.; Mariuzza, R. A. Structure-Based Design of Potent Iminosugar Inhibitors of Endoplasmic Reticulum α-Glucosidase I with Anti-SARS-CoV-2 Activity. *J. Med. Chem.* **2023**, *66*, 2744−2760.

(49) Rajasekharan, S.; Milan Bonotto, R.; Nascimento Alves, L.; Kazungu, Y.; Poggianella, M.; Martinez-Orellana, P.; Skoko, N.; Polez, S.; Marcello, A. Inhibitors of Protein Glycosylation Are Active against the Coronavirus Severe Acute Respiratory Syndrome Coronavirus SARS-CoV-2. *Viruses* **2021**, *13*, 808.

(50) Thaler, M.; Wang, Y.; van der Does, A. M.; Faiz, A.; Ninaber, D. K.; Ogando, N. S.; Beckert, H.; Taube, C.; Salgado-Benvindo, C.; Snijder, E. J.; Bredenbeek, P. J.; Hiemstra, P. S.; van Hemert, M. J. Impact of Changes in Human Airway Epithelial Cellular Composition and Differentiation on SARS-CoV-2 Infection Biology. *Journal of Innate Immunity* **2023**, *15*, 562−580.

(51) Thaler, M.; Salgado-Benvindo, C.; Leijs, A.; Tas, A.; Ninaber, D. K.; Arbiser, J. L.; Snijder, E. J.; van Hemert, M. J. R-Propranolol Has Broad-Spectrum Anti-Coronavirus Activity and Suppresses Factors Involved in Pathogenic Angiogenesis. *Int. J. Mol. Sci.* **2023**, *24*, 4588.

(52) Jiang, J.; Kuo, C.-L.; Wu, L.; Franke, C.; Kallemeijn, W. W.; Florea, B. I.; van Meel, E.; van der Marel, G. A.; Codée, J. D. C.; Boot, R. G.; Davies, G. J.; Overkleeft, H. S.; Aerts, J. M. F. G. Detection of Active Mammalian GH31 α-Glucosidases in Health and Disease Using In-Class, Broad-Spectrum Activity-Based Probes. *ACS Central Science* **2016**, *2*, 351−358.

(53) Warfield, K. L.; Plummer, E. M.; Sayce, A. C.; Alonzi, D. S.; Tang, W.; Tyrrell, B. E.; Hill, M. L.; Caputo, A. T.; Killingbeck, S. S.; Beatty, P. R.; Harris, E.; Iwaki, R.; Kinami, K.; Ide, D.; Kiappes, J. L.; Kato, A.; Buck, M. D.; King, K.; Eddy, W.; Khaliq, M.; Sampath, A.; Treston, A. M.; Dwek, R. A.; Enterlein, S. G.; Miller, J. L.; Zitzmann, N.; Ramstedt, U.; Shresta, S. Inhibition of endoplasmic reticulum glucosidases is required for in vitro and in vivo dengue antiviral activity by the iminosugar UV-4. *Antiviral Res.* **2016**, *129*, 93−98.

(54) Fukushi, M.; Yoshinaka, Y.; Matsuoka, Y.; Hatakeyama, S.; Ishizaka, Y.; Kirikae, T.; Sasazuki, T.; Miyoshi-Akiyama, T. Monitoring of S protein maturation in the endoplasmic reticulum by calnexin is important for the infectivity of severe acute respiratory syndrome coronavirus. *J. Virol* **2012**, *86*, 11745−53.

(55) Elstein, D.; Hollak, C.; Aerts, J. M. F. G.; van Weely, S.; Maas, M.; Cox, T. M.; Lachmann, R. H.; Hrebicek, M.; Platt, F. M.; Butters, T. D.; Dwek, R. A.; Zimran, A. Sustained therapeutic effects of oral miglustat (Zavesca, N-butyldeoxynojirimycin, OGT 918) in type I Gaucher disease. *J. Inherited Metab. Dis.* **2004**, *27*, 757−766.

(56) Kiappes, J. L.; Hill, M. L.; Alonzi, D. S.; Miller, J. L.; Iwaki, R.; Sayce, A. C.; Caputo, A. T.; Kato, A.; Zitzmann, N. ToP-DNJ, a Selective Inhibitor of Endoplasmic Reticulum α-Glucosidase II Exhibiting Antiflaviviral Activity. *ACS Chem. Biol.* **2018**, *13*, 60−65.

(57) Fiege, L.; Duran, I.; Marquardt, T. Improved Enzyme Replacement Therapy with Cipaglucosidase Alfa/Miglustat in Infantile Pompe Disease. *Pharmaceuticals (Basel)* **2023**, *16*, 1199.

(58) Salgado-Benvindo, C.; Leijs, A. A.; Thaler, M.; Tas, A.; Arbiser, J. L.; Snijder, E. J.; van Hemert, M. J. Honokiol Inhibits SARS-CoV-2 Replication in Cell Culture at a Post-Entry Step. *Microbiol. Spectrum* **2023**, *11*, No. e03273-22.

(59) Salgado-Benvindo, C.; Tas, A.; Zevenhoven-Dobbe, J. C.; van der Meer, Y.; Sidorov, I. A.; Leijs, A. A.; Gelderloos, A. T.; van Kasteren, P. B.; Snijder, E. J.; van Hemert, M. J.; Wanningen, P. Characterization of SARS-CoV-2 replication in human H1299/ACE2 cells: a versatile and practical infection model for antiviral research and beyond. *Antiviral Research* **2024**, *227*, 105903.

(60) Ninaber, D. K.; van der Does, A. M.; Hiemstra, P. S. Isolating Bronchial Epithelial Cells from Resected Lung Tissue for Biobanking and Establishing Well-Differentiated Air-Liquid Interface Cultures. *J. Vis. Exp.* **2023**, No. 195, e65102.

(61) Drosten, C.; Günther, S.; Preiser, W.; Van Der Werf, S.; Brodt, H.-R.; Becker, S.; Rabenau, H.; Panning, M.; Kolesnikova, L.; Fouchier, R. A.; et al. Identification of a novel coronavirus in patients with severe acute respiratory syndrome. *N. Engl. J. Med.* **2003**, *348*, 1967−1976.

(62) Kovacikova, K.; Morren, B. M.; Tas, A.; Albulescu, I. C.; van Rijswijk, R.; Jarhad, D. B.; Shin, Y. S.; Jang, M. H.; Kim, G.; Lee, H. W.; Jeong, L. S.; Snijder, E. J.; van Hemert, M. J. 6'-β-Fluoro-

Homoaristeromycin and 6′-Fluoro-Homoneplanocin A Are Potent Inhibitors of Chikungunya Virus Replication through Their Direct Effect on Viral Nonstructural Protein 1. *Antimicrob. Agents Chemother.* **2020**, *64*, e02532-19.

(63) Vaux, D.; Tooze, J.; Fuller, S. Identification by anti-idiotype antibodies of an intracellular membrane protein that recognizes a mammalian endoplasmic reticulum retention signal. *Nature* **1990**, *345*, 495−502.

(64) Salgado-Benvindo, C.; Thaler, M.; Tas, A.; Ogando, N. S.; Bredenbeek, P. J.; Ninaber, D. K.; Wang, Y.; Hiemstra, P. S.; Snijder, E. J.; van Hemert, M. J. Suramin Inhibits SARS-CoV-2 Infection in Cell Culture by Interfering with Early Steps of the Replication Cycle. *Antimicrob. Agents Chemother.* **2020**, *64*, e00900-20.

(65) Corman, V. M.; Landt, O.; Kaiser, M.; Molenkamp, R.; Meijer, A.; Chu, D. K.; Bleicker, T.; Brunink, S.; Schneider, J.; Schmidt, M. L.; Mulders, D. G.; Haagmans, B. L.; van der Veer, B.; van den Brink, S.; Wijsman, L.; Goderski, G.; Romette, J. L.; Ellis, J.; Zambon, M.; Peiris, M.; Goossens, H.; Reusken, C.; Koopmans, M. P.; Drosten, C. Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. *Eurosurveillance* **2020**, *25*, 25.

7.3.3. Conformational and Electronic Variations in 1,2- and 1,5 α-Cyclophellitols and their Impact on Retaining α-Glucosidase Inhibition

Chemistry–A European Journal
Research Article
doi.org/10.1002/chem.202400723
Chemistry Europe
European Chemical Societies Publishing
www.chemeurj.org

Conformational and Electronic Variations in 1,2- and 1,5a-Cyclophellitols and their Impact on Retaining α-Glucosidase Inhibition

Tim P. Ofman,[a] Jurriaan J. A. Heming,[a] Alba Nin-Hill,[b] Florian Küllmer,[a] Elisha Moran,[c] Megan Bennett,[c] Roy Steneker,[a] Anne-Mei Klein,[a] Gijs Ruijgrok,[a] Ken Kok,[a] Zach W. B. Armstrong,[c] Johannes M. F. G. Aerts,[a] Gijsbert A. van der Marel,[a] Carme Rovira,[b, d] Gideon J. Davies,[c] Marta Artola,[a] Jeroen D. C. Codée,[a] and Herman S. Overkleeft*[a]

Glycoside hydrolases (glycosidases) take part in myriad biological processes and are important therapeutic targets. Competitive and mechanism-based inhibitors are useful tools to dissect their biological role and comprise a good starting point for drug discovery. The natural product, cyclophellitol, a mechanism-based, covalent and irreversible retaining β-glucosidase inhibitor has inspired the design of diverse α- and β-glycosidase inhibitor and activity-based probe scaffolds. Here, we sought to deepen our understanding of the structural and functional requirements of cyclophellitol-type compounds for effective human α-glucosidase inhibition. We synthesized a comprehensive set of α-configured 1,2- and 1,5a-cyclophellitol analogues bearing a variety of electrophilic traps. The inhibitory potency of these compounds was assessed towards both lysosomal and ER retaining α-glucosidases. These studies revealed the 1,5a-cyclophellitols to be the most potent retaining α-glucosidase inhibitors, with the nature of the electrophile determining inhibitory mode of action (covalent or non-covalent). DFT calculations support the ability of the 1,5a-cyclophellitols, but not the 1,2-congeners, to adopt conformations that mimic either the Michaelis complex or transition state of α-glucosidases.

Introduction

Carbohydrates are found abundantly in nature and are essential in numerous biological processes.[1–4] The vast structural variation found in glycans is evident in the large variety of hydrolytic enzymes that have emerged and that are responsible for their processing and degradation. This large family of glycoside hydrolases is categorized in over 180 subfamilies, based on their primary sequence, tertiary structure and function (www.cazy.org).[5] Understanding their mode of action is an important step in the rational design of compounds that can selectively and efficiently inhibit specific glycoside hydrolases. Retaining glycoside hydrolases, which comprise a large part of the known glycosidases, encompass the human GH31 retaining α-glucosidases subject of the here-presented studies. Retaining α-glucosidases classically employ a Koshland double displacement mechanism (Figure 1A).[6] Typically, two carboxylic acid residues residing in the enzyme active site are positioned in such a way that one residue can act as a nucleophile and the other as a catalytic acid/base. Upon binding of the substrate in the active site, a Michaelis complex is formed with the substrate (in general the case of retaining α-glucosidases) adopting a $^4C_1$-conformation. In this way, the leaving group is positioned axially, allowing protonation by the catalytic acid-base and subsequent nucleophilic displacement of the aglycon by the nucleophilic acid residue. This process proceeds through a glucosyl $^4H_3$ oxocarbenium ion-like transition state and results in the formation of a covalent intermediate, with the bound glucose adopting a $^1S_3$-conformation. Next and following the expulsion of the aglycon, water enters the active site. Following a reversed conformational itinerary ($^1S_3 \rightarrow {}^4H_3 \rightarrow {}^4C_1$), α-glucose is released and the enzyme returned to its resting phase ready for another catalytic cycle.[7–11] Knowledge of these conformational itineraries allows both the interpretation and design of potent

---

[a] *Dr. T. P. Ofman, J. J. A. Heming, Dr. F. Küllmer, R. Steneker, A.-M. Klein, G. Ruijgrok, K. Kok, Prof. Dr. J. M. F. G. Aerts, Prof. Dr. G. A. van der Marel, Dr. M. Artola, Prof. Dr. J. D. C. Codée, Prof. Dr. H. S. Overkleeft*
Leiden Institute of Chemistry, Leiden University
Einsteinweg 55, 2333 CC Leiden, The Netherlands
E-mail: h.s.overkleeft@chem.leidenuniv.nl

[b] *Dr. A. Nin-Hill, Prof. Dr. C. Rovira*
Departament de Química Inorgànica i Orgànica (Secció de Química Orgànica) and Institut de Química Teòrica i Computacional (IQTCUB)
Universitat de Barcelona
Martí i Franques 1–11, E-08028 Barcelona, Spain

[c] *E. Moran, M. Bennett, Dr. Z. W. B. Armstrong, Prof. Dr. G. J. Davies*
York Structural Biology Laboratory, Department of Chemistry
University of York
Heslington YO10 5DD, United Kingdom

[d] *Prof. Dr. C. Rovira*
Institució Catalana de Recerca i Estudis Avançats (ICREA)
08020 Barcelona (Spain)

Supporting information for this article is available on the WWW under https://doi.org/10.1002/chem.202400723

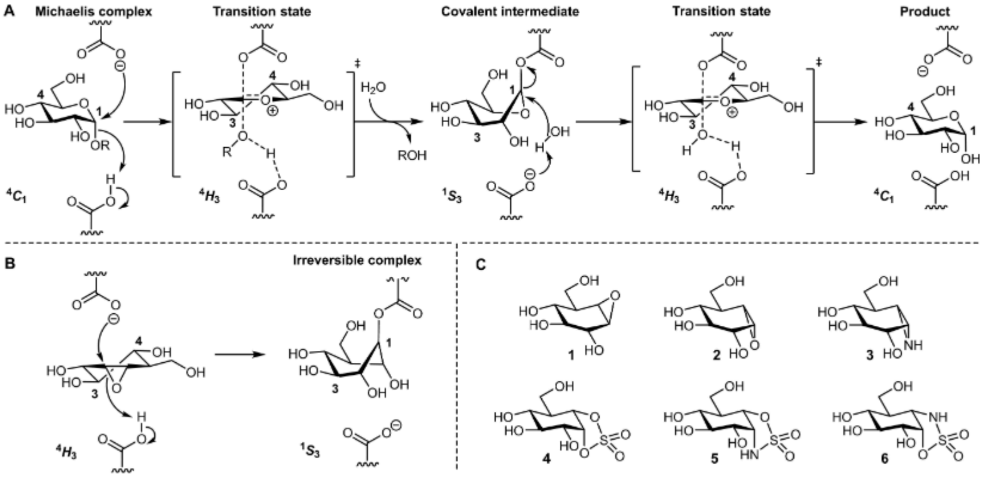


Figure 1. (A) Mechanistic itinerary of retaining α-glucosidases, (B) mechanism of inhibition by 1,5a-*epi*-cyclophellitol **2** and (C) a selection of established covalent and non-covalent inhibitors of retaining β-glucosidases (**1**) and α-glucosidases (**2–6**).

inhibitors and activity-based probes, such as those based on cyclophellitol, as described here.

Cyclophellitol (**1**), a natural product found in the *Phellinus sp.* mushroom, is a potent mechanism-based, covalent and irreversible retaining β-glucosidase inhibitor.[12] Cyclophellitol is the carbocyclic analogue of natural retaining β-glucosidase substrates (β-glucopyranosides), bearing an epoxide bridging the C1 and C5a position (IUPAC numbering).[13–15] This epoxide also effectively constrains the cyclohexane into a ${}^4H_3$ conformation thereby mimicking the transition state (TS) during hydrolysis of β-glucopyranoses by retaining β-glucosidases. Not long after the discovery of cyclophellitol and fuelled by its rare mode of action, the 1,5a-epimer of cyclophellitol was synthesized and shown to be a mechanism-based inhibitor of retaining α-glucosidases.[16]

Upon binding of 1,5a-*epi*-cyclophellitol (**2**) to the active site of a retaining α-glucosidase, the nucleophilic carboxylate opens the (protonated) epoxide in a *trans*-diaxial fashion forming an irreversible, covalent ester linkage with the inhibitor, thereby incapacitating the enzyme (Figure 1B). This *modus operandi* has been well-appreciated in the design of activity-based protein profiling (ABPP) as tools in chemical glycobiology.[17–21] Previously, we have shown 1,5a-*epi*-cyclophellitol (**2**) and its nitrogen congener (**3**) to be effective irreversible inhibitors with micromolar to nanomolar affinities for human acid α-glucosidase (GAA) and endoplasmic reticulum glucosidase II (ER-II, Figure 1C).[20,22] Notably, given the occasional off-target effects of 1,5a-*epi*-cyclophellitol, replacing the epoxide by a cyclic sulfate (as in **4**) resulted in an inhibitor with excellent potency and selectivity for α-glucosidases.[23] Conformationally, compound **4** does not exhibit ${}^4H_3$ character since the ring is not distorted by a strained three-membered ring. Rather, a ${}^4C_1$ conformation is adopted mimicking the structure of the α-glucosyl substrate in the Michaelis complex instead. As a follow up study and with the aim to reduce the electrophilicity of the cyclic sulfate, the corresponding cyclic sulfamidates (**5** and **6**) were developed as retaining α-glucosidase inhibitors.[24] Due to the non-covalent binding mode of compound **5**, this ligand was further evaluated as enzyme stabilizer and pharmacological chaperone for the possible treatment of Pompe disease, in which the lysosomal α-glucosidase GAA is genetically impaired.[25–27] Altogether, these results invite for a more in-depth study of modified cyclophellitol analogues as mechanism-based inhibitors.

Here, the synthesis and inhibitory potential of α-1,2-cyclophellitol (**12–21**, Figure 2) in comparison with α-1,5a-*epi*-cyclophellitols (**2–11**) is described. The inhibitory potencies and mode of action of the focused library of cyclophellitols on GAA and ER-II was studied in comparison to their parent compound, α-1,5a-*epi*-cyclophellitol (**2**). Some compounds proved to be more potent than lead structure (**2**), with low micromolar inhibition constants (IC_{50}) observed for these. In all, the work presented here comprises expansion of the cyclophellitol scaffold in the design of both covalent and competitive retaining glycosidase inhibitors, including chemistries that can be readily adapted to differently configured carbohydrate mimetics and that would target glycosidases other than the α-glucosidase ones studied here.

Results and Discussion

Compound synthesis. Compound **22** (Scheme 1), the key intermediate from which all 1,2-cyclophellitols were derived, was synthesized in eight steps according to procedures of Crotti *et al.* and Nagarajan *et al.* (See SI, Scheme S1).[28,29] Epoxidation of the double bond in **22** under the *aegis* of *m*-CPBA yielded a separable mixture of epoxides **23** and **24** in 53 % and 26 % yield, respectively.

Reductive deprotection (Pd/C, H_2) of α-epoxide **23** yielded **12** (83 %) as the first target compound, of which all spectro-

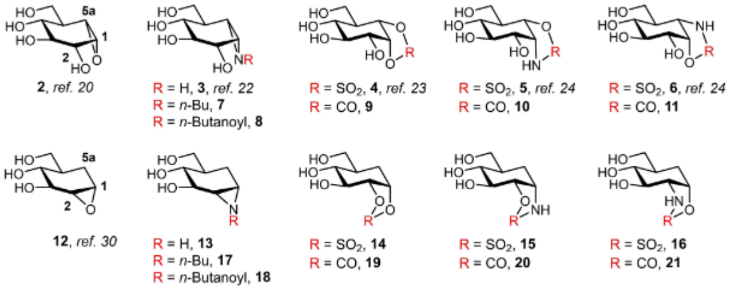
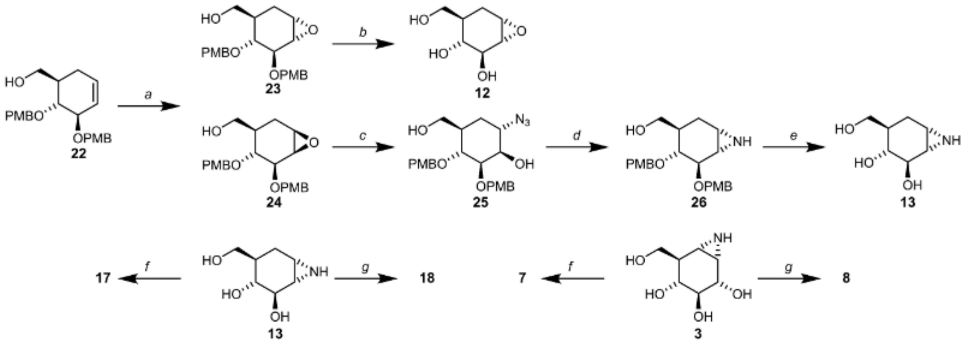
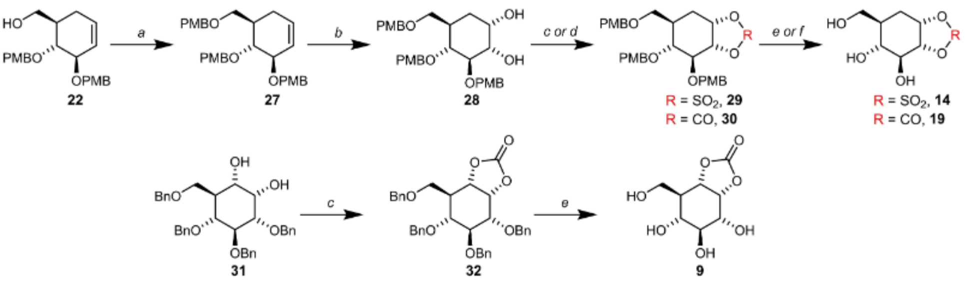


Figure 2. Twenty cyclophellitol analogues studied in this paper for their inhibitory activity against human retaining α-glucosidases.



Scheme 1. Synthesis of compounds **7**, **8**, **12**, **13**, **17** and **18**. Reagents and conditions: (a) *m*-CPBA, DCM, 18 h, rt, 53 % (**23**), 26 % (**24**); (b) Pd/C, H_2, MeOH, 1 h, rt (83 %); (c) NaN_3, DMF, 18 h, 130 °C (60 %); (d) TPP, CH_3CN, 18 h, 60 °C (43 %); (e) Na, NH_3, 1 h, −60 °C (92 %); (f) butyl iodide, K_2CO_3, DMF, 18 h, 80 °C, 84 % (**17**), 24 % (**7**); (g) butanoyl chloride, Et_3N, MeOH, 30 min, 0 °C, 50 % (**18**), 64 % (**8**).

scopic data were in full agreement with those reported in literature.[30] Following established procedures, [31] opening of the β-epoxide with sodium azide under elevated temperatures yielded **25** (60 %) which was subsequently treated with triphenylphosphine to undergo an intramolecular Staudinger cyclisation to yield α-aziridine **26** as the single regioisomer (43 %). Removal of the 4-methoxybenzyl protecting groups in **26** was accomplished using Birch conditions (Na, NH_3) yielding target compound **13** (92 %). Treatment of the α-aziridine with either butyl iodide or butanoyl chloride yielded the butylated aziridine **17** and butanoylated aziridine **18** in 84 % and 50 % respectively. The butylated and butanoylated 1,5a-α-aziridines (**7** and **8**) were obtained *via* identical conditions starting from 1,5a-α-aziridine **3**, which was in turn synthesized according to procedures optimized as published previously.[22,32] Cyclic sulfate **14** and carbonate **19** were constructed starting with protection of the primary hydroxyl in **22** as the 4-methoxybenzyl ether under Williamson etherification conditions (NaH, PMBCl) to yield compound **27** in 85 % (Scheme 2). Subsequent dihydroxylation of the alkene ($RuCl_3$, $NaIO_4$) yielded solely α-*cis*-diol **28** (85 %)



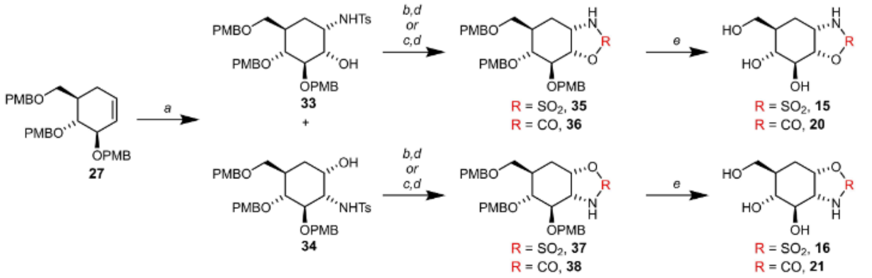
Scheme 2. Synthesis of **9**, **14** and **19**. Reagents and conditions: (a) PMBCl, NaH, DMF, 16 h, rt (85 %); (b) $RuCl_3$, $NaIO_4$, 1:4:4 H_2O:CH_3CN:EtOAc, 1 h, 0 °C (85 %); (c) triphosgene, pyridine, DCM, 1.5 h, rt, 86 % (**30**), 90 % (**32**); (d) (i) $SOCl_2$, Et_3N, DCM, 1 h, rt; (ii) $RuCl_3$, $NaIO_4$, 1:1 H_2O:CH_3CN, 15 min, 0 °C, 46 % (**29**); (e) $Pd(OH)_2$/C, H_2, MeOH, 18 h, rt, 68 % (**19**), quant. (**9**); (f) Pd/C, H_2, MeOH, 18 h, rt, 83 % (**14**).

which was either sulfurylated ($SOCl_2$, Et_3N, *then* $RuCl_3$, $NaIO_4$) or carbonylated (triphosgene, Et_3N) to yield cyclic sulfate **29** (46 %) and cyclic carbonate **30** (86 %) respectively. Hydrogenation yielded the final compounds **14** and **19** in 83 % and 68 % respectively. In addition, 1,5a-α-carbonate **9** was obtained *via* identical conditions starting from 1,5a-α-*cis*-diol **31**, which was in turn synthesized according to procedures published previously.[23] Cyclic sulfamidates **15** and **16** and carbamates **20** and **21** were constructed *via* a stereoselective Sharpless aminohydroxylation on alkene **27** ($K_2[OsO_2(OH)_4]$, chloramine-T, TEBACl) to give a separable, regioisomeric mixture of α-*cis*-amino alcohols **33** and **34** in 54 % and 31 % respectively (Scheme 3).[33] Both α-*cis*-amino alcohols could be transformed into their corresponding cyclic sulfamidates by treatment with sulfuryl chloride and Et_3N at low temperatures (−78 °C) in quantitative yields. Subsequent removal of the *N*-tosyl functionality under reductive conditions (Na, naphthalene) gave rise to cyclic sulfamidates **35** and **37** in 75 % and 85 % yield respectively. Alternatively, treatment of the individual amino alcohols **33** and **34** with triphosgene and pyridine followed by subsequent reductive detosylation (Na, naphthalene) yielded cyclic carbamates **36** and **38** (71 % and 79 % respectively, yield over two steps). Global deprotection with TFA and triethylsilane as cation scavenger afforded target sulfamidates **15** and **16** and carbamates **20** and **21** in 94 %, 81 %, 67 % and 32 % respectively.

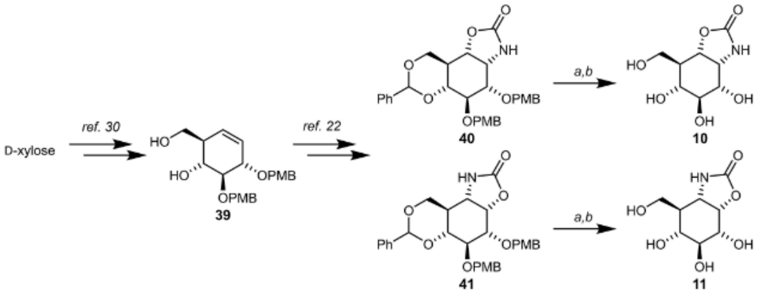
Cyclic carbamates **10** and **11** were constructed *via* global deprotection of intermediates **40** and **41** (Scheme 4). Compounds **40** and **41** were synthesized from parent structure **39**,[34–36] according to modified literature procedures (See SI, Scheme S2).[24] A two-step deprotection sequence then proceeded smoothly by Apparent IC_{50} values were determined by measuring hydrolysis of the fluorogenic substrate, 4-methylulbelliferyl-α-D-glucose, where release of fluorescent product (4-methylumbelliferonate) is determined in terms of relative absorption (see SI). Epoxide **12** proved to be a micromolar inhibitor of rhGAA treating compounds **40** and **41** with a Brønsted acid (*p*-TsOH, MeOH) followed by treating the purified intermediates (S18 and S19) with TFA and triethylsilane. This afforded the target structures **10** and **11** in moderate to good yield (74 % and 86 % respectively).

***In vitro* inhibition of human acid α-glucosidase and ER α-glucosidase II**. With inhibitors **2–21** in hand, attention was turned to evaluating their inhibitory potencies as inhibitors against recombinant human acid α-glucosidase (rhGAA) and ER α-glucosidase II (ER-II, Table 1). Although inhibition by covalent inhibitors reflects both initial binding and subsequent covalent rate (see below), IC_{50} measurements provide a facile initial way to compare inhibitory potency.

($IC_{50} = 32.4 \pm 4.6\ \mu M$) making it slightly less potent than its 1,5a-counterpart **2** ($IC_{50} = 6.7 \pm 0.34\ \mu M$).[23] In contrast, superior inhibitory potency was observed for **12** ($IC_{50} = 11.3 \pm 0.5\ \mu M$) when screened against ER-II. Aziridine **13** showed to be inactive on rhGAA, while inhibiting ER-II in the micromolar range ($IC_{50} = 46.2 \pm 2.6\ \mu M$), whilst its 1,5a-counterpart **3** tested to be a submicromolar inhibitor of both rhGAA ($IC_{50} = 0.34 \pm 0.091\ \mu M$) and ER-II ($IC_{50} = 1.5 \pm 0.08\ \mu M$).[23] *n*-Butyl- and *n*-butanoyl 1,2-azir-



Scheme 3. Synthesis of cyclic sulfamidates **15** and **16**, and carbamates **20** and **21**. Reagents and conditions: (a) Chloramine-T, TEBACl, $K_2[OsO_2(OH)_4]$, 1 : 1 $CHCl_3$:H_2O, 18 h, 60 °C, 54 % (**33**), 31 % (**34**); (b) SO_2Cl_2, Et_3N, DCM, 2 h, −78 °C, quant. (**S4**), quant. (**S5**); (c) triphosgene, pyridine, DCM, 3 h, rt; (d) naphthalene, Na, THF, 30 min, −78 °C, 75 % (**35**), 71 % (**36**), 85 % (**37**), 79 % (**38**); (e) TFA, TES, DCM, 1 h, 0 °C, 94 % (**15**), 67 % (**20**), 81 % (**16**), 32 % (**21**).



Scheme 4. Synthesis of cyclic carbamates **10** and **11**. Reagents and conditions: (a) *p*-TsOH, MeOH, 1 h, 40 °C (b) TFA, TES, DCM, 1 h, 0 °C, 74 % (**10**), 86 % (**11**).

Table 1. Apparent IC$_{50}$ values for *in vitro* inhibition of rhGAA and ER-II.[a]

Compound	*In vitro* rhGAA IC$_{50}$ (μM)	*In vitro* ER-II IC$_{50}$ (μM)	Compound	*In vitro* rhGAA IC$_{50}$ (μM)	*In vitro* ER-II IC$_{50}$ (μM)
2	6.7 ± 0.34 [15 ± 2][b]	> 100 [> 100][b]	**12**	32.4 ± 4.6	11.3 ± 0.5
3	0.34 ± 0.091 [0.038 ± 0.003][b]	1.5 ± 0.08 [1.4 ± 0.1][b]	**13**	> 100	46.2 ± 2.6
4	0.060 ± 0.006 [0.082 ± 0.001][b]	0.029 ± 0.007 [0.029 ± 0.002][b]	**14**	2.63 ± 0.07	26.1 ± 3.2
5	6.1 ± 1.25 [5.17 ± 0.195][c]	> 100 [> 100][c]	**15**	> 100	> 100
6	20.2 ± 5.9 [112 ± 2.54][c]	1.03 ± 0.15 [47 ± 1.75][c]	**16**	> 100	> 100
7	1.96 ± 0.51	0.45 ± 0.03	**17**	> 100	> 100
8	0.84 ± 0.043	5.47 ± 0.91	**18**	> 100	> 100
9	> 100	> 100	**19**	> 100	> 100
10	12.5 ± 3.1	> 100	**20**	> 100	> 100
11	> 100	> 100	**21**	> 100	> 100

[a] All apparent IC$_{50}$ values are determined from three technical triplicates. [b] Value in brackets contains literature reference data.[23] [c] Value in brackets contains literature reference data.[24]

idines **17** and **18** proved inactive as inhibitors of rhGAA and ER-II. Their 1,5a-analogues, compounds **7** and **8**, however, yielded micromolar rhGAA inhibitors. Slightly reduced inhibitory potency was observed for compound **7** and **8** (IC$_{50}$ = 1.96 ± 0.51 μM and 0.84 ± 0.043 μM, respectively) in comparison to unfunctionalized aziridine **3**.

Turning to ER-II, compound **7** proved to be a 3-fold more potent inhibitor (IC$_{50}$ = 0.45 ± 0.03 μM) whilst compound **8** showed a 3-fold reduction in inhibitory potency (IC$_{50}$ = 5.47 ± 0.91 μM) compared to aziridine **3**. Cyclic sulfate **14** appeared to be a low-micromolar inhibitor of rhGAA (IC$_{50}$ = 2.63 ± 0.07 μM), in contrast to its somewhat weaker inhibition of ER-II (IC$_{50}$ = 26.1 ± 3.2 μM), giving compound **14** a roughly tenfold selectivity against rhGAA over ER-II. Its 1,5a-counterpart **4** proved to be a 100-fold more potent inhibitor of both rhGAA and ER-II (IC$_{50}$ = 0.060 ± 0.006 μM and 0.029 ± 0.007 μM respectively).[23] Both the 1,2-cyclic sulfamidates **15** and **16** lacked the ability to reduce enzyme activities of both rhGAA as ER-II up to concentrations of 100 μM, demonstrating that migration of the sulfamidate from 1,5a to 1,2 does not lead to effective inhibitors. With regard to cyclic carbamates **10**, **11**, **20** and **21**, only compound **10** appeared to be an active inhibitor of rhGAA (IC$_{50}$ = 12.5 ± 3.1 μM), which is in line with the structural relationship observed for the cyclic sulfamidates. In addition, inferior apparent IC$_{50}$ values are observed for the 1,5a-(*N*,*O*)-regioisomers (**5** and **10**), which are roughly an order of magnitude more potent inhibitors when compared to the 1,5a-(*O*,*N*)-regioisomers (**6** and **11**). Finally, neither cyclic carbonates **9** nor **19** proved able to block rhGAA or ER-II activity.

Having identified the inhibitory potencies of compounds **2**–**21**, focus was shifted to determining the kinetic parameters and the mode of binding of some of the most active inhibitors on rhGAA. For this, rhGAA was incubated with a fixed substrate concentration and various inhibitor concentrations. Subsequently, apparent IC$_{50}$ values were measured under varying incubation times. Compounds **2**, **3**, **4**, **6**, **7**, **8**, **12** and **14** showed a gradual decrease in enzyme activity indicating these compounds to be covalent and irreversible binders (Table 2). In

Table 2. Inhibitor kinetic constants for recombinant human α-glucosidase (rhGAA).[a]

Compound	Kinetic parameters in rhGAA k_{inact}/K_I (min^{-1} mM^{-1})	Mode of binding	Compound	Kinetic parameters in rhGAA k_{inact}/K_I (min^{-1} mM^{-1})	Mode of binding
2	0.1511 ± 0.0101 [0.37][b]	Covalent[b]	**12**	0.1526 ± 0.0102	Covalent
3	N.D. [58.0][b]	Covalent[b]	**13**	N.D.	N.D.
4	62.41 ± 3.82 [64.3][b]	Covalent[b]	**14**	1.389 ± 0.0592	Covalent
5	N.D.	Non-covalent[c]	**15**	N.D.	N.D.
6	0.06169 ± 0.0063 [0.7675][c]	Covalent[c]	**16**	N.D.	N.D.
7	N.D.	Covalent	**17**	N.D.	N.D.
8	N.D.	Covalent	**18**	N.D.	N.D.
9	N.D.	N.D.	**19**	N.D.	N.D.
10	N.D.	Non-covalent	**20**	N.D.	N.D.
11	N.D.	N.D.	**21**	N.D.	N.D.

[a] All apparent IC$_{50}$ values are determined from two technical triplicates. [b] Value in brackets contains literature reference data.[23] [c] Value in brackets contains literature reference data.[24] N.D.: not determined.

contrast, both compound **5** and **10** appeared to be competitive inhibitors of rhGAA, as indicated by the observed lack of time dependency of the enzyme activity. Compounds **12** and **14** display pseudo first order kinetics due to fast inhibition against rhGAA, limiting measurement of a combined $k_{\text{inact}}/K_{\text{I}}$ ratio.

Conformational free energy landscapes and target engagement. With the aim to obtain some more insight in the binding mode of the new 1,2-cyclophellitols, the gas-phase conformational free energy landscapes of the most potent 1,2-cyclophellitol inhibitors (**12**–**14**) were mapped next. Free energy landscapes (FELs) were computed by means of metadynamics simulations based on density functional theory (DFT, see Supporting Information).[37–47] For compounds **12** and **13**, the lowest energy conformation calculated is centred at $^4H_5$, with relative energies quickly inclining around this energy minimum (Figure 3B and D, respectively). This suggests compounds **12** and **13** to be relatively rigid and to mainly occupy this $^4H_5$ conformation. In contrast, their 1,5a-counterparts, **2** and **3**, exhibit local energy minima around the $^4H_3$ conformation, related to a 60° displacement along the ϕ-axis (Figure 3A, C, respectively). The lack of flexibility exhibited by compounds **12** and **13** prevents adoption of the $^4H_3$ conformation, required for suitable mimicry of the transition state.

The energy minimum of cyclic sulfate **14** is located around $^4C_1$, with a relatively wide minimum expanding toward the $^4H_3$-^4E-$^4H_5$ region, with even an additional energy minimum extending toward the $B_{3,0}$-$^1S_3$ region (Figure 3F). Again a 60° shift along the ϕ-axis is observed when compared to **4** (Figure 3E).[23] Interestingly, here an additional energy minimum is observed around the $B_{3,0}$-$^1S_3$ region not observed in parent structure **4**. These many low-energy conformations suggest **14** not only to adopt a $^4C_1$ conformation (mimicking the structure of the substrate at the Michaelis complex), but to be flexible enough to reach the $^4H_3$ transition state conformation. Thus, based on its conformational preferences, compound **14** should, as observed during *in vitro* assays, inhibit α-glucosidases with higher potencies in comparison to compounds **12** and **13**.

In order to ascertain whether the 1,2-cyclophellitol inhibitors binds to α-glucosidases of the CAZY GH31 family as predicted by the conformational analyses, the structure of **13** in complex with the model GH31 system Agd31B from *Cellvibrio japonicus* (a bacterial homologue of GAA) was determined at 1.9 Å resolution. **13** binds in the −1 subsite as expected (subsite nomenclature according to literature).[23,24,48] Serendipitously, as occasionally occurs with crystal soaks, the aziridine bound unopened in approximate $^1S_3/^4H_3$ conformation (Figure 4A). Asp412, the nucleophile, sits 3 Å "above" the atom equivalent to C1 of a glycoside with geometry poised for nucleophilic attack, and with Asp480, the acid/base, just 2.3 Å from the aziridine nitrogen. The structure confirms, "on enzyme", the conformational design of the 1,2-cyclophellitol as it lies in the low energy, favoured, region of the free energy landscape (Figure 3D).

Furthermore, we were struck by the shape of the 1,2-cyclophellitol design, in general, and the manner in which the shape and stereo/regiochemistry allows for substituents to lay astride the 1 and 2 positions tucked under below the sugar ring. Building on proposals that 1,2-cyclic phosphates could act as weak glycogen phosphorylase inhibitors,[49,50] we also wondered if 1,2-linked compounds would bind to *E. coli* maltodextrin phosphorylase, MalP, (a glycogen phosphorylase homolog).[51,52] Whilst we were unable to measure the binding constants for the compounds, soaking of crystals of MalP with **21** yielded a 1.9 Å structure which clearly showed binding in the catalytic centre (Figure 4B). suggesting that this scaffold may offer inspiration for future phosphorylase and glycosyltransferase inhibitor designs.

Discussion

This study reports on the preparation of a focused series of 1,2-*epi*-cyclophellitol analogues in comparison with their 1,5a-counterparts and their evaluation as inhibitors of recombinant human acid α-glucosidase (rhGAA) and ER α-glucosidase II (ER-II) and related enzymes. All 1,2-analogues (**12**–**21**) revealed reduced inhibitory potencies in comparison to their 1,5a-counterpart, with only compounds **12**–**14** exhibiting inhibitory potencies below 100 μM. The work presented here includes some 1,5a-*epi*-cyclophellitol derivatives with improved inhibition properties compared to what we reported previously. Of these, *N*-alkyl-1,5a-aziridine **7** inhibits ER-II more potently compared to its non-alkylated counterpart **3**, resulting in a nanomolar IC_{50} against ER-II with a 5-fold selectivity over rhGAA. Additionally, 1,5a-(*N*,*O*)-carbamate **10** revealed to be a non-covalent, low-micromolar inhibitor of rhGAA. Here, a strong structural relationship can be drawn with 1,5a-(*N*,*O*)-sulfamidate **5**, suggesting identical enzyme interactions are at play. Therefore, compound **10** may be an interesting candidate for further study as enzyme stabilizer that can be potentially used in treatment of Pompe disease. Conformational free energy landscapes revealed the ground state of compounds **12**–**14** to have undergone a shift in lowest energy conformation in comparison to their parent structures (**2**–**4**). As a result of this shift, the lowest energy conformation (approximated as $^4H_5$) neither resembles the conformation of the Michaelis complex nor the transition state during hydrolysis reflected by an energetic penalty for distortion to the conformation observed in crystal structures. This conformational shift may explain the overall reduction in observed inhibitory potencies of the 1,2-cyclophellitols in contrast to their 1,5a-counterparts. Although given the observed binding on maltodextrin phosphorylase, it is possible that the geometry of the compounds may inspire new inhibitor designs for transferases. In all, this study into structure-activity relationships of cyclophellitol analogues as human α-glucosidase inhibitors may fuel future design of constructs to effectively act on glycoside hydrolases of various sources and acting on various substrate glycosides.

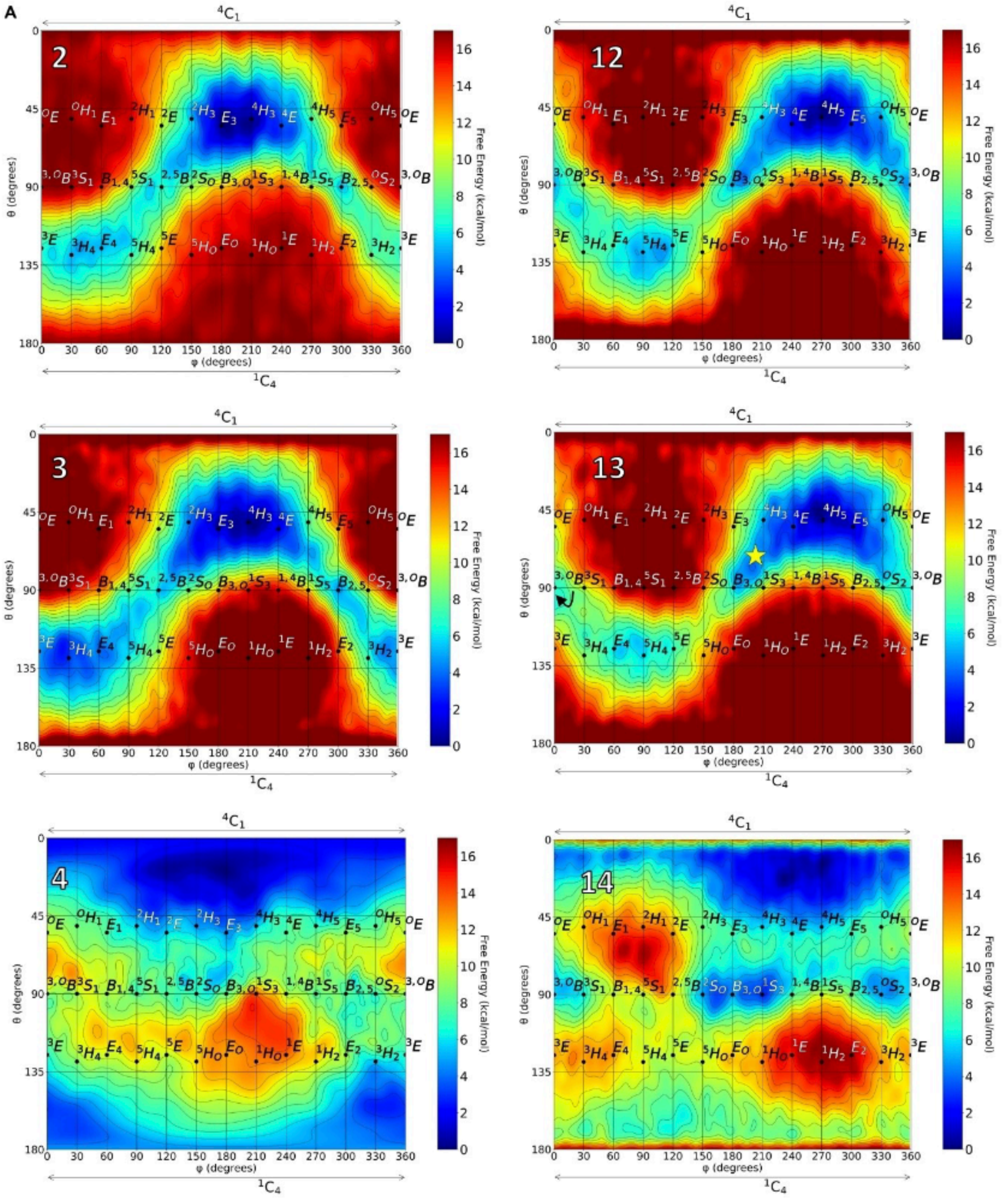


Figure 3. Gas-phase conformational free energy landscapes of (A) 1,5a-*epi*-cyclophellitol **2**, (B) 1,2-epoxide **12**, (C) 1,5a-aziridine **3**, (D) 1,2-aziridine **13** (E) cyclic 1,5a-sulfate **4** (adapted from Artola *et al.*[23]), and (F) cyclic 1,2-sulfate **14**. Isolines are 1 kcal/mol with the x and y axis representing the angle (in degrees) in Cremer-Pople puckering coordinates (ϕ and θ respectively). (D) The yellow star indicates the observed conformation of **13** in the −1 subsite of GH31 Agd31B from *Cellvibrio japonicus*. The arrow clarifies that label $^{3,O}B$ corresponds to the 0° point on the x axis and that all conformation labels appear on the right side of their corresponding points.

Experimental Section

Please find details on computational modelling, IC_{50} and kinetics determination, synthesis procedures, and experimental details in the Supporting Information.

Acknowledgements

This work was supported by the European Research Council (ERC-CoG-726072 "GLYCONTROL" to J.D.C.C. and ERC-2020-SyG-951231 "CARBOCENTRE" to H.S.O., C.R. and G.J.D.). GJD was

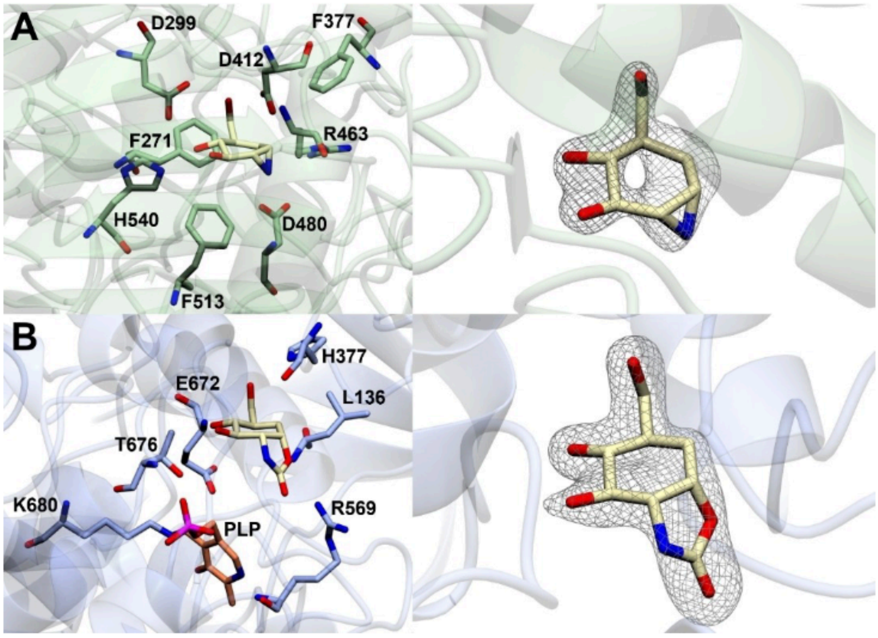


Figure 4. Comparison of the binding modes of 1,2-cyclophellitol based inhibitors towards glycoside hydrolases. All images were created and visualised in CCP4 mg (v. 2.10.11). All F_o-F_c maps (grey) have been displayed in chicken wire and contoured to 3.0 σ. A) Left hand panel: The active site of Agd31B, from *Cellvibrio japonicus*, in complex with one molecule of **13** (yellow). Active site residues (green) are labelled accordingly. Right hand panel: Omit density map for **13**. B) Left hand panel: The active site of MalP from *Escherichia coli*, in complex with one molecule of **21** (yellow). Active site residues (blue) and the PLP cofactor (orange) are labelled accordingly. Right hand panel: Omit density map for **21**.

supported by the Royal Society "Ken Murray" Research Professorship.We acknowledge Dr. Johan P. Turkenburg and Sam Hart for assistance with X-ray data collection; the staff of the Diamond Light Source (U.K.) for provision of i03 and i04 beamline facilities (proposal numbers mx32736 and mx24948). We thank Dr. Lluís Raich for technical assistance. The Viking cluster was used during this project, which is a high performance compute facility provided by the University of York. We are grateful for computational support from the University of York, IT Services and the Research IT team.

Conflict of Interests

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available in the supplementary material of this article.

Keywords: Cyclophellitol · Glucosidase · Inhibitor · Conformational Analysis · Carbasugar

[1] L. L. Kiessling, R. A. Splain, *Annu. Rev. Biochem.* **2010**, *79*, 619–653.
[2] C. R. Bertozzi, L. L. Kiessling, *Science* **2001**, *291*, 2357–2364.
[3] R. A. Dwek, *Chem. Rev.* **1996**, *96*, 683–720.
[4] H. Ghazarian, B. Idoni, S. B. A. Oppenheimer, *Acta Histochem.* **2011**, *113*, 236–247.
[5] E. Drula, M. L. Garron, S. Dogan, V. Lombard, B. Henrissat, N. Terrapon, *Nucleic Acids Res.* **2022**, *50*, D571–D577.
[6] D. E. Koshland, *Proc. Natl. Acad. Sci. USA* **1958**, *44*, 98–104.
[7] A. Vasella, G. J. Davies, M. Böhm, *Curr. Opin. Chem. Biol.* **2002**, *6*, 619–629.
[8] D. L. Zechel, S. G. Withers, *Curr. Opin. Chem. Biol.* **2001**, *5*, 643–649.
[9] C. S. Rye, S. G. Withers, *Curr. Opin. Chem. Biol.* **2000**, *4*, 573–580.
[10] J. D. McCarter, S. G. Withers, *Curr. Opin. Struct. Biol.* **1994**, *4*, 885–892.
[11] G. Speciale, A. J. Thompson, G. J. Davies, S. J. Williams, *Curr. Opin. Struct. Biol.* **2014**, *28*, 1–13.
[12] S. Atsumi, H. Iinuma, H. Naganawa, H. Nakamura, T. Takeuchi, K. Umezawa, Y. Iitaka, *J. Antibiot.* **1990**, *43*, 49–53.
[13] T. K. M. Shing, Y. Cui, Y. Tang, *J. Chem. Soc. Chem. Commun.* **1991**, *11*, 756–757.
[14] K. Tatsuta, Y. Niwata, K. Umezawa, K. Toshima, M. Nakata, *J. Antibiot.* **1991**, *44*, 912–914.
[15] A. D. Mcnaught, H. B. F. Dixon, F. Cornish-Bowden, M. A. Chester, A. J. Barrett, J. C. Rigg, D. Horton, L. Anderson, D. C. Baker, H. H. Baer, J. N. Bemiller, B. Bossenbroek, R. W. Jeanloz, K. L. Loening, W. A. Szarek, R. S. Tipson, W. J. Whelan, R. L. Whistler, *Pure Appl. Chem.* **1996**, *68*, 1919–2008.
[16] R. E. McDevitt, B. A. Fraser-Reid, *J. Org. Chem.* **2002**, *59*, 3250–3252.
[17] S. P. Schröder, R. Petracca, H. Minnee, M. Artola, J. M. F. G. Aerts, J. D. C. Codée, G. A. van der Marel, H. S. Overkleeft, *Eur. J. Org. Chem.* **2016**, *28*, 4787–4794.
[18] C. De Boer, N. G. S. McGregor, E. Peterse, S. P. Schröder, B. I. Florea, J. Jiang, J. Reijngoud, A. F. J. Ram, G. P. Van Wezel, G. A. van der Marel, J. D. C. Codée, H. S. Overkleeft, G. J. Davies, *RSC Chem. Biol.* **2020**, *1*, 148–155.

[19] B. F. Cravatt, A. T. Wright, J. W. Kozarich, *Annu. Rev. Biochem.* **2008**, *77*, 383–414.
[20] S. P. Schröder, J. W. van de Sande, W. W. Kallemeijn, C.-L. Kuo, M. Artola, E. J. van Rooden, J. Jiang, T. J. M. Beenakker, B. I. Florea, W. A. Offen, G. J. Davies, A. J. Minnaard, J. M. F. G. Aerts, J. D. C. Codée, G. A. van der Marel, H. S. Overkleeft, *Chem. Commun.* **2017**, *53*, 12528–12531.
[21] L. I. Willems, H. S. Overkleeft, S. I. Van Kasteren, *Bioconjugate Chem.* **2014**, *25*, 1181–1191.
[22] J. Jiang, M. Artola, T. J. M. Beenakker, S. P. Schröder, R. Petracca, C. de Boer, J. M. F. G. Aerts, G. A. van der Marel, J. D. C. Codée, H. S. Overkleeft, *Eur. J. Org. Chem.* **2016**, *22*, 3671–3678.
[23] M. Artola, L. Wu, M. J. Ferraz, C.-L. Kuo, L. Raich, I. Z. Breen, W. A. Offen, J. D. C. C. Codée, G. A. van der Marel, C. Rovira, J. M. F. G. Aerts, G. J. Davies, H. S. Overkleeft, *ACS Cent. Sci.* **2017**, *3*, 784–793.
[24] K. Kok, C. L. Kuo, R. E. Katzy, L. T. Lelieveld, L. Wu, V. Roig-Zamboni, G. A. van der Marel, J. D. C. Codée, G. Sulzenbacher, G. J. Davies, H. S. Overkleeft, J. M. F. G. Aerts, M. Artola, *J. Am. Chem. Soc.* **2022**, *144*, 14819–14827.
[25] M. Alomari, M. Taha, F. Rahim, M. Selvaraj, N. Iqbal, S. Chigurupati, S. Hussain, N. Uddin, N. B. Almandil, M. Nawaz, R. Khalid Farooq, K. M. Khan, *Bioorg. Chem.* **2021**, *108*, 104638.
[26] F. Azimi, H. Azizian, M. Najafi, F. Hassanzadeh, H. Sadeghi-aliabadi, J. B. Ghasemi, M. Ali Faramarzi, S. Mojtabavi, Larijani, B. L. Saghaei, M. Mahdavi, *Bioorg. Chem.* **2021**, *114*, 105127.
[27] S. P. Schröder, L. Wu, M. Artola, T. Hansen, W. A. Offen, M. J. Ferraz, K.-Y. Li, J. M. F. G. Aerts, G. A. van der Marel, J. D. C. Codée, G. J. Davies, H. S. Overkleeft, *J. Am. Chem. Soc.* **2018**, *140*, 5045–5048.
[28] I. Frau, V. Di Bussolo, L. Favero, M. Pineschi, P. Crotti, *Chirality* **2011**, *23*, 820–826.
[29] A. V. R. L. Sudha, M. Nagarajan, *Chem. Commun.* **1998**, *3*, 31.
[30] V. W.-F. Tai, P. H. Fung, Y. S. Wong, T. K. M. Shing, *Tetrahedron: Asymmetry* **1994**, *5*, 1353–1362.
[31] D. Lu, S. Zhu, L. F. Sobala, G. Bernardo-Seisdedos, O. Millet, Y. Zhang, J. Jiménez-Barbero, G. J. Davies, M. Sollogoub, *Org. Lett.* **2018**, *20*, 7488–7492.
[32] M. Artola, S. Wouters, S. P. Schröder, C. de Boer, Y. Chen, R. Petracca, A. M. C. H. van den Nieuwendijk, J. M. F. G. Aerts, G. A. van der Marel, J. D. C. Codée, H. S. Overkleeft, *Eur. J. Org. Chem.* **2019**, *6*, 1397–1404.
[33] H. Tsunoda, S. Ogawa, *Liebigs Ann.* **1995**, *2*, 267–277.
[34] F. G. Hansen, E. Bundgaard, R. Madsen, *J. Org. Chem.* **2005**, *70*, 10139–10142.
[35] G. Luchetti, K. Ding, M. D'Alarcao, A. Kornienko, *Synthesis* **2008**, *19*, 3142–3147.
[36] B. Lacourt-Gadras, M. Grignon-Dubois, B. Rezzonico, *Carbohydr. Res.* **1992**, *235*, 281–288.
[37] R. Car, M. Parrinello, *Phys. Rev. Lett.* **1985**, *55*, 2471–2474.
[38] P. Hohenberg, W. Kohn, *Phys. Rev. Lett.* **1964**, *136*, B864.
[39] W. Kohn, *Phys. Rev. A* **1965**, *1133*, 140–148.
[40] CPMD V3.9 Copyright IBM Corp 1990–2004, Copyright MPI für Festkörperforschung Stuttgart 1997–2001.
[41] J. P. Perdew, K. Burke, M. Ernzerhof, *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.
[42] A. Ardèvol, C. Rovira, *J. Am. Chem. Soc.* **2015**, *137*, 7528–7547.
[43] M. Marianski, A. Supady, T. Ingram, M. Schneider, C. Baldauf, *J. Chem. Theory Comput.* **2016**, *12*, 6157–6168.
[44] N. Troullier, J. L. Martins, *Phys. Rev. B* **1991**, *43*, 1993–2006.
[45] A. Laio, M. Parrinello, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 12562–12566.
[46] G. A. Tribello, M. Bonomi, D. Branduardi, C. Camilloni, G. Bussi, *Comput. Phys. Commun.* **2014**, *185*, 604–613.
[47] P. Tiwary, M. Parrinello, *J. Phys. Chem. B* **2015**, *119*, 736–742.
[48] J. Larsbrink, A. Izumi, G. R. Hemsworth, G. J. Davies, H. Brumer, *J. Biochem.* **2012**, *287*, 43288–43299.
[49] F. C. Kokesh, R. K. Stephenson, Y. Kakuda, *Biochim. Biophys. Acta* **1977**, *483*, 258–262.
[50] H. Y. Hu, M. Allen, *Biochim. Biophys. Acta* **1978**, *525*, 55–60.
[51] M. O'Reilly, K. A. Watson, R. Schinzel, D. Palm, L. N. Johnson, *Nature Struc. Bio.* **1997**, *4*, 405–412.
[52] S. Geremia, M. Campagnolo, R. Schinzel, L. N. Johnson, *J. Mol. Biol.* **2002**, *322*, 413–423.

Manuscript received: February 22, 2024
Accepted manuscript online: April 16, 2024
Version of record online: April 24, 2024

8. References

1. Davies, G. and B. Henrissat, *Structures and mechanisms of glycosyl hydrolases.* Structure, 1995. **3**(9): p. 853-9.
2. Gabius, H.-J., H.-C. Siebert, S. André, J. Jiménez-Barbero, and H. Rüdiger, *Chemical Biology of the Sugar Code.* ChemBioChem, 2004. **5**(6): p. 740-764.
3. Schjoldager, K.T., Y. Narimatsu, H.J. Joshi, and H. Clausen, *Global view of human protein glycosylation pathways and functions.* Nat. Rev. Mol. Cell Biol, 2020. **21**(12): p. 729-749.
4. Spiro, R.G., *Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds.* Glycobiology, 2002. **12**(4): p. 43r-56r.
5. Ajit Varki, R.D.C., Jeffrey D. Esko, Pamela Stanley, Gerald W. Hart, Markus Aebi, Debra Mohnen, Taroh Kinoshita, Nicolle H. Packer, James H. Prestegard, Ronald L. Schnaar, and Peter H. Seeberger, *Monosaccharide Diversity*. 4th ed. Essentials of Glycobiology [Internet]. 2022: Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press.
6. Lowe, J.B. and J.D. Marth, *A genetic approach to Mammalian glycan function.* Annu Rev Biochem, 2003. **72**: p. 643-91.
7. Varki, A., *Biological roles of glycans.* Glycobiology, 2017. **27**(1): p. 3-49.
8. Reily, C., T.J. Stewart, M.B. Renfrow, and J. Novak, *Glycosylation in health and disease.* Nat. Rev. Nephrol., 2019. **15**(6): p. 346-366.
9. Ajit Varki, R.D.C., Jeffrey D. Esko, Pamela Stanley, Gerald W. Hart, Markus Aebi, Debra Mohnen, Taroh Kinoshita, Nicolle H. Packer, James H. Prestegard, Ronald L. Schnaar, and Peter H. Seeberger, *Glycosyltransferases and Glycan-Processing Enzymes*. 4th ed. Essentials of Glycobiology 2022.
10. Cantarel, B.L., P.M. Coutinho, C. Rancurel, T. Bernard, V. Lombard, and B. Henrissat, *The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics.* Nucleic Acids Res, 2009. **37**(Database issue): p. D233-8.
11. Drula, E., M.-L. Garron, S. Dogan, V. Lombard, B. Henrissat, and N. Terrapon, *The carbohydrate-active enzyme database: functions and literature.* Nucleic Acids Research, 2021. **50**(D1): p. D571-D577.
12. Davies, G.J. and S.J. Williams, *Carbohydrate-active enzymes: sequences, shapes, contortions and cells.* Biochem Soc Trans, 2016. **44**(1): p. 79-87.
13. Vocadlo, D.J. and G.J. Davies, *Mechanistic insights into glycosidase chemistry.* Curr Opin Chem Biol, 2008. **12**(5): p. 539-55.
14. KOSHLAND Jr., D.E., *STEREOCHEMISTRY AND THE MECHANISM OF ENZYMATIC REACTIONS.* Biological Reviews, 1953. **28**(4): p. 416-436.
15. Cremer, D. and J.A. Pople, *General definition of ring puckering coordinates.* Journal of the American Chemical Society, 1975. **97**(6): p. 1354-1358.
16. Davies, G.J., A. Planas, and C. Rovira, *Conformational Analyses of the Reaction Coordinate of Glycosidases.* ACS, 2012. **45**(2): p. 308-316.
17. Morais, M.A.B., A. Nin-Hill, and C. Rovira, *Glycosidase mechanisms: Sugar conformations and reactivity in endo- and exo-acting enzymes.* Current Opinion in Chemical Biology, 2023. **74**: p. 102282.
18. Kim, Y., H. Li, J. Choi, J. Boo, H. Jo, J.Y. Hyun, and I. Shin, *Glycosidase-targeting small molecules for biological and therapeutic applications.* Chemical Society Reviews, 2023. **52**(20): p. 7036-7070.

19. Speciale, G., A.J. Thompson, G.J. Davies, and S.J. Williams, *Dissecting conformational contributions to glycosidase catalysis and inhibition.* Curr. Opin. Struct. Biol, 2014. **28**: p. 1-13.
20. Rempel, B.P. and S.G. Withers, *Covalent inhibitors of glycosidases and their applications in biochemistry and biology.* Glycobiology, 2008. **18**(8): p. 570-586.
21. Hermans, M.M., M.A. Kroos, J. van Beeumen, B.A. Oostra, and A.J. Reuser, *Human lysosomal alpha-glucosidase. Characterization of the catalytic site.* JBC, 1991. **266**(21): p. 13507-13512.
22. HERRCHEN, M. and G. LEGLER, *Identification of an essential carboxylate group at the active site of lacZβ-galactosidase from Escherichia coli.* Eur J Biochem 1984. **138**(3): p. 527-531.
23. Iwanami, S., H. Matsui, A. Kimura, H. Ito, H. Mori, M. Honma, and S. Chiba, *Chemical Modification and Amino Acid Sequence of Active Site in Sugar Beet α-Glucosidase.* Biosci. Biotechnol. Biochem., 1995. **59**(3): p. 459-463.
24. Kimura, A., M. Takata, Y. Fukushi, H. Mori, H. Matsui, and S. Chiba, *A Catalytic Amino Acid and Primary Structure of Active Site in Aspevgillus niger α-Glucosidase.* Biosci. Biotechnol. Biochem., 1997. **61**(7): p. 1091-1098.
25. Caron, G. and S.G. Withers, *Conduritol aziridine: A new mechanism-based glucosidase inactivator.* BBRC, 1989. **163**(1): p. 495-499.
26. Speers, A.E. and B.F. Cravatt, *Chemical strategies for activity-based proteomics.* Chembiochem, 2004. **5**(1): p. 41-7.
27. Cravatt, B.F., A.T. Wright, and J.W. Kozarich, *Activity-based protein profiling: from enzyme chemistry to proteomic chemistry.* Annu Rev Biochem, 2008. **77**: p. 383-414.
28. Willems, L.I., H.S. Overkleeft, and S.I. van Kasteren, *Current Developments in Activity-Based Protein Profiling.* Bioconjug Chem, 2014. **25**(7): p. 1181-1191.
29. Stubbs, K.A., A. Scaffidi, A.W. Debowski, B.L. Mark, R.V. Stick, and D.J. Vocadlo, *Synthesis and use of mechanism-based protein-profiling probes for retaining beta-D-glucosaminidases facilitate identification of Pseudomonas aeruginosa NagZ.* J Am Chem Soc, 2008. **130**(1): p. 327-35.
30. Hekmat, O., Y.W. Kim, S.J. Williams, S. He, and S.G. Withers, *Active-site peptide "fingerprinting" of glycosidases in complex mixtures by mass spectrometry. Discovery of a novel retaining beta-1,4-glycanase in Cellulomonas fimi.* J Biol Chem, 2005. **280**(42): p. 35126-35.
31. Phenix, C.P., B.P. Rempel, K. Colobong, D.J. Doudet, M.J. Adam, L.A. Clarke, and S.G. Withers, *Imaging of enzyme replacement therapy using PET.* Proc Natl Acad Sci U S A, 2010. **107**(24): p. 10842-7.
32. Braun, C., G.D. Brayer, and S.G. Withers, *Mechanism-based Inhibition of Yeast α-Glucosidase and Human Pancreatic α-Amylase by a New Class of Inhibitors: 2-DEOXY-2,2-DIFLUORO-α-GLYCOSIDES (*).* JBC, 1995. **270**(45): p. 26778-26781.
33. Wu, L., Z. Armstrong, S.P. Schröder, C. de Boer, M. Artola, J.M. Aerts, H.S. Overkleeft, and G.J. Davies, *An overview of activity-based probes for glycosidases.* Curr Opin Chem Biol, 2019. **53**: p. 25-36.
34. Wu, L., C.M. Viola, A.M. Brzozowski, and G.J. Davies, *Structural characterization of human heparanase reveals insights into substrate recognition.* Nat Struct Mol Biol, 2015. **22**(12): p. 1016-22.

35. Wu, L., N. Wimmer, G.J. Davies, and V. Ferro, *Structural insights into heparanase activity using a fluorogenic heparan sulfate disaccharide.* Chem. Comm, 2020. **56**(89): p. 13780-13783.
36. Coombe, D.R. and N.S. Gandhi, *Heparanase: A Challenging Cancer Drug Target.* Frontiers in Oncology, 2019. **9**.
37. Nishimura, Y., W. Wang, S. Kondo, T. Aoyagi, and H. Umezawa, *Siastatin B, a potent neuraminidase inhibitor: the total synthesis and absolute configuration.* ACS, 1988. **110**(21): p. 7249-7250.
38. de Boer, C., Z. Armstrong, V.A.J. Lit, U. Barash, G. Ruijgrok, I. Boyango, M.M. Weitzenberg, S.P. Schröder, A.J.C. Sarris, N.J. Meeuwenoord, P. Bule, Y. Kayal, N. Ilan, J.D.C. Codée, I. Vlodavsky, H.S. Overkleeft, G.J. Davies, and L. Wu, *Mechanism-based heparanase inhibitors reduce cancer metastasis in vivo.* PNAS, 2022. **119**(31): p. e2203167119.
39. De Pasquale, V. and L.M. Pavone, *Heparan sulfate proteoglycans: The sweet side of development turns sour in mucopolysaccharidoses.* BBA 2019. **1865**(11): p. 165539.
40. Simon Davis, D.A. and C.R. Parish, *Heparan Sulfate: A Ubiquitous Glycosaminoglycan with Multiple Roles in Immunity.* Front. immunol, 2013. **4**.
41. Iozzo, R.V., *Heparan sulfate proteoglycans: intricate molecules with intriguing functions.* J Clin Invest, 2001. **108**(2): p. 165-7.
42. Sarrazin, S., W.C. Lamanna, and J.D. Esko, *Heparan sulfate proteoglycans.* Cold Spring Harb Perspect Biol, 2011. **3**(7).
43. Turnbull, J.E., R.L. Miller, Y. Ahmed, T.M. Puvirajesinghe, and S.E. Guimond, *Glycomics profiling of heparan sulfate structure and activity.* Methods Enzymol, 2010. **480**: p. 65-85.
44. Esko, J.D. and S.B. Selleck, *Order Out of Chaos: Assembly of Ligand Binding Sites in Heparan Sulfate.* Annu. Rev. Biochem, 2002. **71**(1): p. 435-471.
45. Couchman, J.R. and C.A. Pataki, *An introduction to proteoglycans and their localization.* J Histochem Cytochem, 2012. **60**(12): p. 885-97.
46. Bishop, J.R., M. Schuksz, and J.D. Esko, *Heparan sulphate proteoglycans fine-tune mammalian physiology.* Nature, 2007. **446**(7139): p. 1030-1037.
47. Michikawa, M., H. Ichinose, M. Momma, P. Biely, S. Jongkees, M. Yoshida, T. Kotake, Y. Tsumuraya, S.G. Withers, Z. Fujimoto, and S. Kaneko, *Structural and biochemical characterization of glycoside hydrolase family 79 β-glucuronidase from Acidobacterium capsulatum.* J Biol Chem, 2012. **287**(17): p. 14069-77.
48. Mayfosh, A.J., T.K. Nguyen, and M.D. Hulett, *The Heparanase Regulatory Network in Health and Disease.* Int J Mol Sci, 2021. **22**(20).
49. Wilson, J.C., A.E. Laloo, S. Singh, and V. Ferro, *1H NMR spectroscopic studies establish that heparanase is a retaining glycosidase.* Biochem Biophys Res Commun, 2014. **443**(1): p. 185-8.
50. Peterson, S. and J. Liu, *Deciphering mode of action of heparanase using structurally defined oligosaccharides.* J Biol Chem, 2012. **287**(41): p. 34836-43.
51. Borlandelli, V., Z. Armstrong, A. Nin-Hill, J.D.C. Codée, L. Raich, M. Artola, C. Rovira, G.J. Davies, and H.S. Overkleeft, *4-O-Substituted Glucuronic Cyclophellitols are Selective Mechanism-Based Heparanase Inhibitors.* ChemMedChem, 2023. **18**(4): p. e202200580.

52. Hulett, M.D., J.R. Hornby, S.J. Ohms, J. Zuegg, C. Freeman, J.E. Gready, and C.R. Parish, *Identification of active-site residues of the pro-metastatic endoglycosidase heparanase.* Biochemistry, 2000. **39**(51): p. 15659-67.
53. Okada, Y., S. Yamada, M. Toyoshima, J. Dong, M. Nakajima, and K. Sugahara, *Structural recognition by recombinant human heparanase that plays critical roles in tumor metastasis. Hierarchical sulfate groups with different effects and the essential target disulfated trisaccharide sequence.* J Biol Chem, 2002. **277**(45): p. 42488-95.
54. Theocharis, A.D., D. Manou, and N.K. Karamanos, *The extracellular matrix as a multitasking player in disease.* The FEBS Journal, 2019. **286**(15): p. 2830-2869.
55. Ilan, N., M. Elkin, and I. Vlodavsky, *Regulation, function and clinical significance of heparanase in cancer metastasis and angiogenesis.* Int J Biochem Cell Biol, 2006. **38**(12): p. 2018-39.
56. Winkler, J., A. Abisoye-Ogunniyan, K.J. Metcalf, and Z. Werb, *Concepts of extracellular matrix remodelling in tumour progression and metastasis.* Nat Commun, 2020. **11**(1): p. 5120.
57. Koganti, R., R. Suryawanshi, and D. Shukla, *Heparanase, cell signaling, and viral infections.* Cell Mol Life Sci, 2020. **77**(24): p. 5059-5077.
58. Dalziel, M., M. Crispin, C.N. Scanlan, N. Zitzmann, and R.A. Dwek, *Emerging principles for the therapeutic exploitation of glycosylation.* Science, 2014. **343**: p. 1235681.
59. Liu, L.P., X.P. Sheng, T.K. Shuai, Y.X. Zhao, B. Li, and Y.M. Li, *Helicobacter pylori promotes invasion and metastasis of gastric cancer by enhancing heparanase expression.* World J Gastroenterol, 2018. **24**(40): p. 4565-4577.
60. Jayatilleke, K.M. and M.D. Hulett, *Heparanase and the hallmarks of cancer.* J. Transl. Med, 2020. **18**(1): p. 453.
61. Hanahan, D. and R.A. Weinberg, *Hallmarks of cancer: the next generation.* Cell, 2011. **144**(5): p. 646-74.
62. Ritchie, J.P., V.C. Ramani, Y. Ren, A. Naggi, G. Torri, B. Casu, S. Penco, C. Pisano, P. Carminati, M. Tortoreto, F. Zunino, I. Vlodavsky, R.D. Sanderson, and Y. Yang, *SST0001, a chemically modified heparin, inhibits myeloma growth and angiogenesis via disruption of the heparanase/syndecan-1 axis.* Clin Cancer Res, 2011. **17**(6): p. 1382-93.
63. Medicine, U.S.N.L.o. *A Phase III PI-88 in the Adjuvant Treatment of Subjects With Hepatitis Virus Related HCC After Surgical Resection (PATRON)*. 2020; Available from: A Phase III PI-88 in the Adjuvant Treatment of Subjects With Hepatitis Virus Related HCC After Surgical Resection (PATRON).
64. Adams, P.D., P.V. Afonine, G. Bunkóczi, V.B. Chen, I.W. Davis, N. Echols, J.J. Headd, L.W. Hung, G.J. Kapral, R.W. Grosse-Kunstleve, A.J. McCoy, N.W. Moriarty, R. Oeffner, R.J. Read, D.C. Richardson, J.S. Richardson, T.C. Terwilliger, and P.H. Zwart, *PHENIX: a comprehensive Python-based system for macromolecular structure solution.* Acta Crystallogr., Sect. D: Struct. Biol., 2010. **66**: p. 213.
65. Medicine, U.S.N.L.o. *M402 in Combination With Nab-Paclitaxel and Gemcitabine in Pancreatic Cancer*. 2012.
66. Rivara, S., F.M. Milazzo, and G. Giannini, *Heparanase: a rainbow pharmacological target associated to multiple pathologies including rare diseases.* Future Med. Chem, 2016. **8**(6): p. 647-680.
67. Medicine, U.S.N.l.o. *Pixatimod (PG545) Plus Nivolumab in PD-1 Relapsed/Refractory Metastatic Melanoma and NSCLC and With Nivolumab and Low-dose*

Cyclophosphamide in MSS Metastatic Colorectal Carcinoma (mCRC). [cited 2024; Available from: https://classic.clinicaltrials.gov/ct2/show/NCT05061017.

68. Courtney, S.M., P.A. Hay, R.T. Buck, C.S. Colville, D.W. Porter, D.I.C. Scopes, F.C. Pollard, M.J. Page, J.M. Bennett, M.L. Hircock, E.A. McKenzie, C.R. Stubberfield, and P.R. Turner, *2,3-Dihydro-1,3-dioxo-1H-isoindole-5-carboxylic acid derivatives: a novel class of small molecule heparanase inhibitors.* Bioorg Med Chem Lett, 2004. **14**(12): p. 3269-3273.
69. Li, K.-Y., J. Jiang, M.D. Witte, W.W. Kallemeijn, W.E. Donker-Koopman, R.G. Boot, J.M.F.G. Aerts, J.D.C. Codée, G.A. van der Marel, and H.S. Overkleeft, *Exploring functional cyclophellitol analogues as human retaining beta-glucosidase inhibitors.* OBC, 2014. **12**(39): p. 7786-7791.
70. Wu, L., 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, *Activity-based probes for functional interrogation of retaining β-glucuronidases.* Nat. Chem. Biol, 2017. **13**(8): p. 867-873.
71. Chen, Y., A.M.C.H. van den Nieuwendijk, L. Wu, E. Moran, F. Skoulikopoulou, V. van Riet, H.S. Overkleeft, G.J. Davies, and Z. Armstrong, *Molecular Basis for Inhibition of Heparanases and β-Glucuronidases by Siastatin B.* JACS, 2023.
72. Keil, J.M., G.R. Rafn, I.M. Turan, M.A. Aljohani, R. Sahebjam-Atabaki, and X.-L. Sun, *Sialidase Inhibitors with Different Mechanisms.* J. Med. Chem, 2022. **65**(20): p. 13574-13593.
73. Umezawa, H., T. Aoyagi, T. Komiyama, H. Morishima, and M. Hamada, *Purification and characterization of a sialidase inhibitor, siastatin, produced by Streptomyces.* J Antibiot (Tokyo), 1974. **27**(12): p. 963-9.
74. Murshudov, G.N., P. Skubák, A.A. Lebedev, N.S. Pannu, R.A. Steiner, R.A. Nicholls, M.D. Winn, F. Long, and A.A. Vagin, *REFMAC5 for the refinement of macromolecular crystal structures.* Acta Crystallogr., Sect. D: Struct. Biol., 2011. **67**: p. 355.
75. Emsley, P. and K. Cowtan, *Coot: model-building tools for molecular graphics.* Acta Crystallogr D Biol Crystallogr, 2004. **60**(Pt 12 Pt 1): p. 2126-32.
76. Nishimura, Y., T. Kudo, S. Kondo, T. Takeuchi, T. Tsuruoka, H. Fukuyasu, and S. Shibahara, *Totally synthetic analogues of siastatin B. III. Trifluoroacetamide analogues having inhibitory activity for tumor metastasis.* J Antibiot (Tokyo), 1994. **47**(1): p. 101-7.
77. Nishimura, Y., E. Shitara, H. Adachi, M. Toyoshima, M. Nakajima, Y. Okami, and T. Takeuchi, *Flexible synthesis and biological activity of uronic acid-type gem-diamine 1-N-iminosugars: a new family of glycosidase inhibitors.* J Org Chem, 2000. **65**(1): p. 2-11.
78. Kondo, K.I., H. Adachi, E. Shitara, F. Kojima, and Y. Nishimura, *Glycosidase inhibitors of gem-diamine 1-N-iminosugars. structures in media of enzyme assays.* Bioorg Med Chem, 2001. **9**(5): p. 1091-5.
79. Takatsu, T., M. Takahashi, Y. Kawase, R. Enokita, T. Okazaki, H. Matsukawa, K. Ogawa, Y. Sakaida, T. Kagasaki, T. Kinoshita, M. Nakajima, and K. Tanzawa, *A-72363 A-1, A-2 and C, novel heparanase inhibitors from Streptomyces nobilis SANK 60192. I. Taxonomy of producing organism, fermentation, isolation and structure elucidation.* J Antibiot (Tokyo), 1996. **49**(1): p. 54-60.
80. Geraghty, R.J., M.T. Aliota, and L.F. Bonnac, *Broad-Spectrum Antiviral Strategies and Nucleoside Analogues.* Viruses, 2021. **13**(4).

81. Warfield, K.L., D.S. Alonzi, J.C. Hill, A.T. Caputo, P. Roversi, J.L. Kiappes, N. Sheets, M. Duchars, R.A. Dwek, J. Biggins, D. Barnard, S. Shresta, A.M. Treston, and N. Zitzmann, *Targeting Endoplasmic Reticulum α-Glucosidase I with a Single-Dose Iminosugar Treatment Protects against Lethal Influenza and Dengue Virus Infections.* J Med Chem, 2020. **63**(8): p. 4205-4214.
82. Punekar, M., M. Kshirsagar, C. Tellapragada, and K. Patil, *Repurposing of antiviral drugs for COVID-19 and impact of repurposed drugs on the nervous system.* Microb Pathog, 2022. **168**: p. 105608.
83. Warren, T.K., R. Jordan, M.K. Lo, A.S. Ray, R.L. Mackman, V. Soloveva, D. Siegel, M. Perron, R. Bannister, H.C. Hui, N. Larson, R. Strickley, J. Wells, K.S. Stuthman, S.A. Van Tongeren, N.L. Garza, G. Donnelly, A.C. Shurtleff, C.J. Retterer, D. Gharaibeh, R. Zamani, T. Kenny, B.P. Eaton, E. Grimes, L.S. Welch, L. Gomba, C.L. Wilhelmsen, D.K. Nichols, J.E. Nuss, E.R. Nagle, J.R. Kugelman, G. Palacios, E. Doerffler, S. Neville, E. Carra, M.O. Clarke, L. Zhang, W. Lew, B. Ross, Q. Wang, K. Chun, L. Wolfe, D. Babusis, Y. Park, K.M. Stray, I. Trancheva, J.Y. Feng, O. Barauskas, Y. Xu, P. Wong, M.R. Braun, M. Flint, L.K. McMullan, S.S. Chen, R. Fearns, S. Swaminathan, D.L. Mayers, C.F. Spiropoulou, W.A. Lee, S.T. Nichol, T. Cihlar, and S. Bavari, *Therapeutic efficacy of the small molecule GS-5734 against Ebola virus in rhesus monkeys.* Nature, 2016. **531**(7594): p. 381-5.
84. de Wit, E., F. Feldmann, J. Cronin, R. Jordan, A. Okumura, T. Thomas, D. Scott, T. Cihlar, and H. Feldmann, *Prophylactic and therapeutic remdesivir (GS-5734) treatment in the rhesus macaque model of MERS-CoV infection.* Proc Natl Acad Sci U S A, 2020. **117**(12): p. 6771-6776.
85. Beigel, J.H., K.M. Tomashek, L.E. Dodd, A.K. Mehta, B.S. Zingman, A.C. Kalil, E. Hohmann, H.Y. Chu, A. Luetkemeyer, S. Kline, D. Lopez de Castilla, R.W. Finberg, K. Dierberg, V. Tapson, L. Hsieh, T.F. Patterson, R. Paredes, D.A. Sweeney, W.R. Short, G. Touloumi, D.C. Lye, N. Ohmagari, M.D. Oh, G.M. Ruiz-Palacios, T. Benfield, G. Fätkenheuer, M.G. Kortepeter, R.L. Atmar, C.B. Creech, J. Lundgren, A.G. Babiker, S. Pett, J.D. Neaton, T.H. Burgess, T. Bonnett, M. Green, M. Makowski, A. Osinusi, S. Nayak, and H.C. Lane, *Remdesivir for the Treatment of Covid-19 - Final Report.* N Engl J Med, 2020. **383**(19): p. 1813-1826.
86. Karim, M., C.W. Lo, and S. Einav, *Preparing for the next viral threat with broad-spectrum antivirals.* J Clin Invest, 2023. **133**(11).
87. Bekerman, E., G. Neveu, A. Shulla, J. Brannan, S.Y. Pu, S. Wang, F. Xiao, R. Barouch-Bentov, R.R. Bakken, R. Mateo, J. Govero, C.M. Nagamine, M.S. Diamond, S. De Jonghe, P. Herdewijn, J.M. Dye, G. Randall, and S. Einav, *Anticancer kinase inhibitors impair intracellular viral trafficking and exert broad-spectrum antiviral effects.* J Clin Invest, 2017. **127**(4): p. 1338-1352.
88. Warfield, K.L., E.M. Plummer, A.C. Sayce, D.S. Alonzi, W. Tang, B.E. Tyrrell, M.L. Hill, A.T. Caputo, S.S. Killingbeck, P.R. Beatty, E. Harris, R. Iwaki, K. Kinami, D. Ide, J.L. Kiappes, A. Kato, M.D. Buck, K. King, W. Eddy, M. Khaliq, A. Sampath, A.M. Treston, R.A. Dwek, S.G. Enterlein, J.L. Miller, N. Zitzmann, U. Ramstedt, and S. Shresta, *Inhibition of endoplasmic reticulum glucosidases is required for in vitro and in vivo dengue antiviral activity by the iminosugar UV-4.* Antiviral Res, 2016. **129**: p. 93-98.
89. DeWald, L.E., C. Starr, T. Butters, A. Treston, and K.L. Warfield, *Iminosugars: a host-targeted approach to combat Flaviviridae infections.* Antiviral Res., 2020. **184**: p. 104881.

90. Helenius, A. and M. Aebi, *Roles of N-linked glycans in the endoplasmic reticulum.* Annu Rev Biochem, 2004. **73**: p. 1019-49.
91. O'Keefe, S., Q.P. Roebuck, I. Nakagome, S. Hirono, A. Kato, R. Nash, and S. High, *Characterizing the selectivity of ER α-glucosidase inhibitors.* Glycobiology, 2019. **29**(7): p. 530-542.
92. Tax, G., A. Lia, A. Santino, and P. Roversi, *Modulation of ERQC and ERAD: A Broad-Spectrum Spanner in the Works of Cancer Cells?* J Oncol, 2019. **2019**: p. 8384913.
93. Williams, S.J. and E.D. Goddard-Borger, *α-glucosidase inhibitors as host-directed antiviral agents with potential for the treatment of COVID-19.* Biochem Soc Trans, 2020. **48**(3): p. 1287-1295.
94. Barker, M.K. and D.R. Rose, *Specificity of processing α-glucosidase I is guided by the substrate conformation: crystallographic and in silico studies.* J. Biol. Chem., 2013. **288**: p. 13563.
95. Karade, S.S., A. Kolesnikov, A.M. Treston, and R.A. Mariuzza, *Identification of Endoplasmic Reticulum α-Glucosidase I from a Thermophilic Fungus as a Platform for Structure-Guided Antiviral Drug Design.* Biochem, 2022. **61**(10): p. 822-832.
96. Warfield, K.L., D.S. Alonzi, J.C. Hill, A.T. Caputo, P. Roversi, J.L. Kiappes, N. Sheets, M. Duchars, R.A. Dwek, J. Biggins, D. Barnard, S. Shresta, A.M. Treston, and N. Zitzmann, *Targeting endoplasmic reticulum α-glucosidase I with a single-dose iminosugar treatment protects against lethal influenza and dengue virus infections.* J. Med. Chem, 2020. **63**: p. 4205.
97. Warfield, K.L., D.S. Alonzi, J.C. Hill, A.T. Caputo, P. Roversi, J.L. Kiappes, N. Sheets, M. Duchars, R.A. Dwek, J. Biggins, D. Barnard, S. Shresta, A.M. Treston, and N. Zitzmann, *Targeting Endoplasmic Reticulum α-Glucosidase I with a Single-Dose Iminosugar Treatment Protects against Lethal Influenza and Dengue Virus Infections.* J. Med. Chem, 2020. **63**(8): p. 4205-4214.
98. Caputo, A.T., D.S. Alonzi, L. Marti, I.-B. Reca, J.L. Kiappes, W.B. Struwe, A. Cross, S. Basu, E.D. Lowe, B. Darlot, A. Santino, P. Roversi, and N. Zitzmann, *Structures of mammalian ER α-glucosidase II capture the binding modes of broad-spectrum iminosugar antivirals.* PNAS, 2016. **113**(32): p. E4630.
99. De Masi, R. and S. Orlando, *GANAB and N-Glycans Substrates Are Relevant in Human Physiology, Polycystic Pathology and Multiple Sclerosis: A Review.* Int J Mol Sci, 2022. **23**(13).
100. Pelletier, M.F., A. Marcil, G. Sevigny, C.A. Jakob, D.C. Tessier, E. Chevet, R. Menard, J.J.M. Bergeron, and D.Y. Thomas, *The heterodimeric structure of glucosidase II is required for its activity, solubility, and localization in vivo.* Glycobiology, 2000. **10**(8): p. 815-827.
101. Caputo, A.T., D.S. Alonzi, L. Marti, I.B. Reca, J.L. Kiappes, W.B. Struwe, A. Cross, S. Basu, E.D. Lowe, B. Darlot, A. Santino, P. Roversi, and N. Zitzmann, *Structures of mammalian ER α-glucosidase II capture the binding modes of broad-spectrum iminosugar antivirals.* Proc. Natl. Acad. Sci. U.S.A., 2016. **113**: p. E4630.
102. Artola, M., L. Wu, M.J. Ferraz, C.-L. Kuo, L. Raich, I.Z. Breen, W.A. Offen, J.D.C. Codée, G.A. van der Marel, C. Rovira, J.M.F.G. Aerts, G.J. Davies, and H.S. Overkleeft, *1,6-Cyclophellitol Cyclosulfates: A New Class of Irreversible Glycosidase Inhibitor.* ACS Cent. Sci, 2017. **3**(7): p. 784-793.
103. Shimada, S., B.G. Ng, A.L. White, K.K. Nickander, C. Turgeon, K.L. Liedtke, C.T. Lam, E. Font-Montgomery, C.M. Lourenco, M. He, D.S. Peck, L.A. Umana, C.L. Uhles, D.

Haynes, P.G. Wheeler, M.J. Bamshad, D.A. Nickerson, T. Cushing, R. Gates, N. Gomez-Ospina, H.M. Byers, F.B. Scalco, N.N. Martinez, R. Sachdev, L. Smith, A. Poduri, S. Malone, R.V. Harris, I.E. Scheffer, S.D. Rosenzweig, D.R. Adams, W.A. Gahl, M.C.V. Malicdan, K.M. Raymond, H.H. Freeze, and L.A. Wolfe, *Clinical, biochemical and genetic characteristics of MOGS-CDG: a rare congenital disorder of glycosylation.* J Med Genet, 2022.

104. Sadat, M.A., S. Moir, T.W. Chun, P. Lusso, G. Kaplan, L. Wolfe, M.J. Memoli, M. He, H. Vega, L.J.Y. Kim, Y. Huang, N. Hussein, E. Nievas, R. Mitchell, M. Garofalo, A. Louie, D.C. Ireland, C. Grunes, R. Cimbro, V. Patel, G. Holzapfel, D. Salahuddin, T. Bristol, D. Adams, B.E. Marciano, M. Hegde, Y. Li, K.R. Calvo, J. Stoddard, J.S. Justement, J. Jacques, D.A.L. Priel, D. Murray, P. Sun, D.B. Kuhns, C.F. Boerkoel, J.A. Chiorini, G. Di Pasquale, D. Verthelyi, and S.D. Rosenzweig, *Glycosylation, hypogammaglobulinemia, and resistance to viral infections.* N Engl J Med, 2014. **370**(17): p. 1615-1625.
105. Watanabe, Y., T.A. Bowden, I.A. Wilson, and M. Crispin, *Exploitation of glycosylation in enveloped virus pathobiology.* Biochim Biophys Acta Gen Subj, 2019. **1863**(10): p. 1480-1497.
106. Lin, Q., L. Pei, Z. Zhao, X. Zhuang, and H. Qin, *Glycoprotein α-Subunit of Glucosidase II (GIIα) is a novel prognostic biomarker correlated with unfavorable outcome of urothelial carcinoma.* BMC Cancer, 2022. **22**(1): p. 817.
107. Karade, S.S., E.J. Franco, A.C. Rojas, K.C. Hanrahan, A. Kolesnikov, W. Yu, A.D. MacKerell, Jr., D.C. Hill, D.J. Weber, A.N. Brown, A.M. Treston, and R.A. Mariuzza, *Structure-Based Design of Potent Iminosugar Inhibitors of Endoplasmic Reticulum α-Glucosidase I with Anti-SARS-CoV-2 Activity.* J. Med. Chem, 2023. **66**(4): p. 2744-2760.
108. Chang, J., L. Wang, D. Ma, X. Qu, H. Guo, X. Xu, M. Mason Peter, N. Bourne, R. Moriarty, B. Gu, J.-T. Guo, and M. Block Timothy, *Novel Imino Sugar Derivatives Demonstrate Potent Antiviral Activity against Flaviviruses.* AAC, 2009. **53**(4): p. 1501-1508.
109. Fukushi, M., Y. Yoshinaka, Y. Matsuoka, S. Hatakeyama, Y. Ishizaka, T. Kirikae, T. Sasazuki, and T. Miyoshi-Akiyama, *Monitoring of S Protein Maturation in the Endoplasmic Reticulum by Calnexin Is Important for the Infectivity of Severe Acute Respiratory Syndrome Coronavirus.* J Virol, 2012. **86**(21): p. 11745-11753.
110. Qu, X., X. Pan, J. Weidner, W. Yu, D. Alonzi, X. Xu, T. Butters, T. Block, J.-T. Guo, and J. Chang, *Inhibitors of Endoplasmic Reticulum α-Glucosidases Potently Suppress Hepatitis C Virus Virion Assembly and Release.* Antimicrob. Agents Chemother, 2011. **55**(3): p. 1036-1044.
111. Burke, B., K. Matlin, E. Bause, G. Legler, N. Peyrieras, and H. Ploegh, *Inhibition of N-linked oligosaccharide trimming does not interfere with surface expression of certain integral membrane proteins.* EMBO, 1984. **3**(3): p. 551-556-556.
112. Walker, B.D., M. Kowalski, W.C. Goh, K. Kozarsky, M. Krieger, C. Rosen, L. Rohrschneider, W.A. Haseltine, and J. Sodroski, *Inhibition of human immunodeficiency virus syncytium formation and virus replication by castanospermine.* Proc Natl Acad Sci U S A, 1987. **84**(22): p. 8120-4.
113. Gruters, R.A., J.J. Neefjes, M. Tersmette, R.E.Y. de Goede, A. Tulp, H.G. Huisman, F. Miedema, and H.L. Ploegh, *Interference with HIV-induced syncytium formation and viral infectivity by inhibitors of trimming glucosidase.* Nature, 1987. **330**(6143): p. 74-77.

114. Chang, J., T.M. Block, and J.T. Guo, *Antiviral therapies targeting host ER α-glucosidases: current status and future directions.* Antiviral Res., 2013. **99**: p. 251.

115. Low, J.G., C. Sung, L. Wijaya, Y. Wei, A.P.S. Rathore, S. Watanabe, B.H. Tan, L. Toh, L.T. Chua, Y.a. Hou, A. Chow, S. Howe, W.K. Chan, K.H. Tan, J.S. Chung, B.P. Cherng, D.C. Lye, P.A. Tambayah, L.C. Ng, J. Connolly, M.L. Hibberd, Y.S. Leo, Y.B. Cheung, E.E. Ooi, and S.G. Vasudevan, *Efficacy and safety of celgosivir in patients with dengue fever (CELADEN): a phase 1b, randomised, double-blind, placebo-controlled, proof-of-concept trial.* Lancet Infect Dis, 2014. **14**(8): p. 706-715.

116. Gu, B., P. Mason, L. Wang, P. Norton, N. Bourne, R. Moriarty, A. Mehta, M. Despande, R. Shah, and T. Block, *Antiviral profiles of novel iminocyclitol compounds against bovine viral diarrhea virus, West Nile virus, dengue virus and hepatitis B virus.* Antivir Chem Chemother, 2007. **18**(1): p. 49-59.

117. Warfield, K.L., K.R. Schaaf, L.E. DeWald, K.B. Spurgers, W. Wang, E. Stavale, M. Mendenhall, M.H. Shilts, T.B. Stockwell, D.L. Barnard, U. Ramstedt, and S.R. Das, *Lack of selective resistance of influenza A virus in presence of host-targeted antiviral, UV-4B.* Sci. Rep., 2019. **9**: p. 7484.

118. Plummer, E., M.D. Buck, M. Sanchez, J.A. Greenbaum, J. Turner, R. Grewal, B. Klose, A. Sampath, K.L. Warfield, B. Peters, U. Ramstedt, and S. Shresta, *Dengue virus evolution under a host-targeted antiviral.* J. Virol., 2015. **89**: p. 5592.

119. Perry, S.T., M.D. Buck, E.M. Plummer, R.A. Penmasta, H. Batra, E.J. Stavale, K.L. Warfield, R.A. Dwek, T.D. Butters, D.S. Alonzi, S.M. Lada, K. King, B. Klose, U. Ramstedt, and S. Shresta, *An iminosugar with potent inhibition of dengue virus infection in vivo.* Antiviral Res., 2013. **98**: p. 35.

120. Perry, S.T., M.D. Buck, E.M. Plummer, R.A. Penmasta, H. Batra, E.J. Stavale, K.L. Warfield, R.A. Dwek, T.D. Butters, D.S. Alonzi, S.M. Lada, K. King, B. Klose, U. Ramstedt, and S. Shresta, *An iminosugar with potent inhibition of dengue virus infection in vivo.* Antiviral Res, 2013. **98**(1): p. 35-43.

121. Tierney, M., J. Pottage, H. Kessler, M. Fischl, D. Richman, T. Merigan, W. Powderly, S. Smith, A. Karim, J. Sherman, and et al., *The tolerability and pharmacokinetics of N-butyl-deoxynojirimycin in patients with advanced HIV disease (ACTG 100). The AIDS Clinical Trials Group (ACTG) of the National Institute of Allergy and Infectious Diseases.* J Acquir Immune Defic Syndr Hum Retrovirol, 1995. **10**(5): p. 549-53.

122. Ficicioglu, C., *Review of miglustat for clinical management in Gaucher disease type 1.* Ther Clin Risk Manag, 2008. **4**(2): p. 425-31.

123. Karade, S.S., M.L. Hill, J.L. Kiappes, R. Manne, B. Aakula, N. Zitzmann, K.L. Warfield, A.M. Treston, and R.A. Mariuzza, *N-substituted valiolamine derivatives as potent inhibitors of endoplasmic reticulum α-glucosidases I and II with antiviral activity.* J. Med. Chem., 2021. **64**: p. 18010.

124. Jiang, J., C.-L. Kuo, L. Wu, C. Franke, W.W. Kallemeijn, B.I. Florea, E. van Meel, G.A. van der Marel, J.D.C. Codée, R.G. Boot, G.J. Davies, H.S. Overkleeft, and J.M.F.G. Aerts, *Detection of Active Mammalian GH31 α-Glucosidases in Health and Disease Using In-Class, Broad-Spectrum Activity-Based Probes.* ACS Cent. Sci, 2016. **2**(5): p. 351-358.

125. Sastry, G.M., M. Adzhigirey, T. Day, R. Annabhimoju, and W. Sherman, *Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments.* J Comput Aided Mol Des, 2013. **27**(3): p. 221-34.

126. Yang, Y., K. Yao, M.P. Repasky, K. Leswing, R. Abel, B.K. Shoichet, and S.V. Jerome, *Efficient Exploration of Chemical Space with Docking and Deep Learning.* J Chem Theory Comput, 2021. **17**(11): p. 7106-7119.
127. Friesner, R.A., R.B. Murphy, M.P. Repasky, L.L. Frye, J.R. Greenwood, T.A. Halgren, P.C. Sanschagrin, and D.T. Mainz, *Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes.* J Med Chem, 2006. **49**(21): p. 6177-96.
128. Halgren, T.A., R.B. Murphy, R.A. Friesner, H.S. Beard, L.L. Frye, W.T. Pollard, and J.L. Banks, *Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening.* J Med Chem, 2004. **47**(7): p. 1750-9.
129. Friesner, R.A., J.L. Banks, R.B. Murphy, T.A. Halgren, J.J. Klicic, D.T. Mainz, M.P. Repasky, E.H. Knoll, M. Shelley, J.K. Perry, D.E. Shaw, P. Francis, and P.S. Shenkin, *Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy.* J Med Chem, 2004. **47**(7): p. 1739-49.
130. Lu, C., C. Wu, D. Ghoreishi, W. Chen, L. Wang, W. Damm, G.A. Ross, M.K. Dahlgren, E. Russell, C.D. Von Bargen, R. Abel, R.A. Friesner, and E.D. Harder, *OPLS4: Improving Force Field Accuracy on Challenging Regimes of Chemical Space.* J Chem Theory Comput, 2021. **17**(7): p. 4291-4300.
131. Johnston, R.C., K. Yao, Z. Kaplan, M. Chelliah, K. Leswing, S. Seekins, S. Watts, D. Calkins, J. Chief Elk, S.V. Jerome, M.P. Repasky, and J.C. Shelley, *Epik: pK(a) and Protonation State Prediction through Machine Learning.* J Chem Theory Comput, 2023. **19**(8): p. 2380-2388.
132. Zhu, K., K.W. Borrelli, J.R. Greenwood, T. Day, R. Abel, R.S. Farid, and E. Harder, *Docking covalent inhibitors: a parameter free approach to pose prediction and scoring.* J Chem Inf Model, 2014. **54**(7): p. 1932-40.
133. Pettersen, E.F., T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, and T.E. Ferrin, *UCSF Chimera—A visualization system for exploratory research and analysis.* J. Comput. Chem, 2004. **25**(13): p. 1605-1612.
134. Kurasawa, J.H., A. Park, C.R. Sowers, R.A. Halpin, A. Tovchigrechko, C.L. Dobson, A.E. Schmelzer, C. Gao, S.D. Wilson, and Y. Ikeda, *Chemically Defined, High-Density Insect Cell-Based Expression System for Scalable AAV Vector Production.* Mol Ther Methods Clin Dev, 2020. **19**: p. 330-340.
135. Karbalaei, M., S.A. Rezaee, and H. Farsiani, *Pichia pastoris: A highly successful expression system for optimal synthesis of heterologous proteins.* J Cell Physiol, 2020. **235**(9): p. 5867-5881.
136. Weidner, M., M. Taupp, and S.J. Hallam, *Expression of recombinant proteins in the methylotrophic yeast Pichia pastoris.* J Vis Exp, 2010(36).
137. Ingram, Z., A. Patkar, D. Oh, K.K. Zhang, C. Chung, J. Lin-Cereghino, and G.P. Lin-Cereghino, *Overcoming Obstacles in Protein Expression in the Yeast Pichia pastoris: Interviews of Leaders in the Pichia Field.* Pac J Health (Stockt), 2021. **4**(1).
138. Gorrec, F., *The MORPHEUS protein crystallization screen.* J Appl Crystallogr, 2009. **42**(Pt 6): p. 1035-1042.
139. Ofman, T.P., J.J.A. Heming, A. Nin-Hill, F. Küllmer, E. Moran, M. Bennett, R. Steneker, A.-M. Klein, G. Ruijgrok, K. Kok, Z.W.B. Armstrong, J.M.F.G. Aerts, G.A. van der Marel, C. Rovira, G.J. Davies, M. Artola, J.D.C. Codée, and H.S. Overkleeft, *Conformational and Electronic Variations in 1,2- and 1,5a-Cyclophellitols and their Impact on Retaining α-Glucosidase Inhibition.* Chem. Eur. J, 2024. **30**(31): p. e202400723.

140. Larsbrink, J., A. Izumi, G.R. Hemsworth, G.J. Davies, and H. Brumer, *Structural Enzymology of Cellvibrio japonicus Agd31B Protein Reveals α-Transglucosylase Activity in Glycoside Hydrolase Family 31*.* JBC, 2012. **287**(52): p. 43288-43299.
141. Cavasotto, C.N. and M.G. Aucar, *High-Throughput Docking Using Quantum Mechanical Scoring.* Front Chem, 2020. **8**: p. 246.
142. Thaler, M., T.P. Ofman, K. Kok, J.J.A. Heming, E. Moran, I. Pickles, A.A. Leijs, A.M.C.H. van den Nieuwendijk, R.J.B.H.N. van den Berg, G. Ruijgrok, Z. Armstrong, C. Salgado-Benvindo, D.K. Ninaber, E.J. Snijder, C.A.A. van Boeckel, M. Artola, G.J. Davies, H.S. Overkleeft, and M.J. van Hemert, *Epi-Cyclophellitol Cyclosulfate, a Mechanism-Based Endoplasmic Reticulum α-Glucosidase II Inhibitor, Blocks Replication of SARS-CoV-2 and Other Coronaviruses.* ACS, 2024.
143. Ghosh, S., *Sialic acid and biology of life: An introduction*, in *Sialic Acids and Sialoglycoconjugates in the Biology of Life, Health and Disease*. 2020. p. 1 - 61.
144. Nicholls, J.M., R.B. Moss, and S.M. Haslam, *The use of sialidase therapy for respiratory viral infections.* Antiviral Res, 2013. **98**(3): p. 401-9.
145. Visser, E.A., S.J. Moons, S.B.P.E. Timmermans, H. de Jong, T.J. Boltje, and C. Büll, *Sialic acid O-acetylation: From biosynthesis to roles in health and disease.* JBC, 2021. **297**(2): p. 100906.
146. Schauer, R., *Achievements and challenges of sialic acid research.* Glycoconj J, 2000. **17**(7-9): p. 485-99.
147. Tringali, C., N. Papini, P. Fusi, G. Croci, G. Borsani, A. Preti, P. Tortora, G. Tettamanti, B. Venerando, and E. Monti, *Properties of Recombinant Human Cytosolic Sialidase HsNEU2: THE ENZYME HYDROLYZES MONOMERICALLY DISPERSED GM1 GANGLIOSIDE MOLECULES *.* JBC, 2004. **279**(5): p. 3169-3179.
148. Chavas, L.M.G., C. Tringali, P. Fusi, B. Venerando, G. Tettamanti, R. Kato, E. Monti, and S. Wakatsuki, *Crystal Structure of the Human Cytosolic Sialidase Neu2: EVIDENCE FOR THE DYNAMIC NATURE OF SUBSTRATE RECOGNITION*.* JBC, 2005. **280**(1): p. 469-475.
149. Amaya, M.a.F., A.G. Watts, I. Damager, A. Wehenkel, T. Nguyen, A. Buschiazzo, G. Paris, A.C. Frasch, S.G. Withers, and P.M. Alzari, *Structural Insights into the Catalytic Mechanism of Trypanosoma cruzi trans-Sialidase.* Structure, 2004. **12**(5): p. 775-784.
150. Watts, A.G., I. Damager, M.L. Amaya, A. Buschiazzo, P. Alzari, A.C. Frasch, and S.G. Withers, *Trypanosoma cruzi trans-sialidase operates through a covalent sialyl-enzyme intermediate: tyrosine is the catalytic nucleophile.* J Am Chem Soc, 2003. **125**(25): p. 7532-3.
151. Wicki, J., D.R. Rose, and S.G. Withers, *5 - Trapping Covalent Intermediates on β-Glycosidases*, in *Methods Enzymol.*, D.L. Purich, Editor. 2002, Academic Press. p. 84-105.
152. Zechel, D.L. and S.G. Withers, *Glycosidase Mechanisms: Anatomy of a Finely Tuned Catalyst.* ACR, 2000. **33**(1): p. 11-18.
153. Amaya, M.F., A. Buschiazzo, T. Nguyen, and P.M. Alzari, *The High Resolution Structures of Free and Inhibitor-bound Trypanosoma rangeli Sialidase and its Comparison with T.cruziTrans-sialidase.* JMB, 2003. **325**(4): p. 773-784.
154. Li, Y., H. Cao, H. Yu, Y. Chen, K. Lau, J. Qu, V. Thon, G. Sugiarto, and X. Chen, *Identifying selective inhibitors against the human cytosolic sialidase NEU2 by substrate specificity studies.* Molecular bioSystems, 2011. **7**(4): p. 1060-1072.

155. Heimerl, M., T. Gausepohl, J.H. Mueller, and M. Ricke-Hoch, *Neuraminidases-Key Players in the Inflammatory Response after Pathophysiological Cardiac Stress and Potential New Therapeutic Targets in Cardiac Disease.* Biol, 2022. **11**(8).
156. Zhang, Y., R. Wang, Y. Feng, and F. Ma, *The role of sialyltransferases in gynecological malignant tumors.* Life Sci., 2020. **263**: p. 118670.
157. Ulloa, F. and F.X. Real, *Differential distribution of sialic acid in alpha2,3 and alpha2,6 linkages in the apical membrane of cultured epithelial cells and tissues.* J Histochem Cytochem, 2001. **49**(4): p. 501-10.
158. Bassagañas, S., M. Pérez-Garay, and R. Peracaula, *Cell surface sialic acid modulates extracellular matrix adhesion and migration in pancreatic adenocarcinoma cells.* Pancreas, 2014. **43**(1): p. 109-17.
159. Tokuyama, S., S. Moriya, S.i. Taniguchi, A. Yasui, J.-i. Miyazaki, S. Orikasa, and T. Miyagi, *Suppression of pulmonary metastasis in murine B16 melanoma cells by transfection of a sialidase cDNA.* Int. J. Cancer, 1997. **73**(3): p. 410-415.
160. Bourguet, E., S. Figurska, and M.M. Frączek, *Human Neuraminidases: Structures and Stereoselective Inhibitors.* Journal of Medicinal Chemistry, 2022. **65**(4): p. 3002-3025.
161. Moghadami, M., *A Narrative Review of Influenza: A Seasonal and Pandemic Disease.* Iran J Med Sci, 2017. **42**(1): p. 2-13.
162. Shtyrya, Y.A., L.V. Mochalova, and N.V. Bovin, *Influenza virus neuraminidase: structure and function.* Acta Naturae, 2009. **1**(2): p. 26-32.
163. Luo, M., *Influenza virus entry.* Adv Exp Med Biol, 2012. **726**: p. 201-21.
164. Gamblin, S.J. and J.J. Skehel, *Influenza hemagglutinin and neuraminidase membrane glycoproteins.* JBC, 2010. **285**(37): p. 28403-28409.
165. Gao, Z., K. Robinson, D.M. Skowronski, G. De Serres, and S.G. Withers, *Quantification of the total neuraminidase content of recent commercially-available influenza vaccines: Introducing a neuraminidase titration reagent.* Vaccine, 2020. **38**(4): p. 715-718.
166. Tripathi, S., J. Batra, and S.K. Lal, *Interplay between influenza A virus and host factors: targets for antiviral intervention.* Archives of Virology, 2015. **160**(8): p. 1877-1891.
167. Ditmar, M.F., *CHAPTER 11 - Infectious Diseases*, in *Pediatric Secrets*, R.A. Polin and M.F. Ditmar, Editors. 2011, Mosby: Philadelphia. p. 354-422.
168. Ellis, D., J. Lederhofer, O.J. Acton, Y. Tsybovsky, S. Kephart, C. Yap, R.A. Gillespie, A. Creanga, A. Olshefsky, T. Stephens, D. Pettie, M. Murphy, C. Sydeman, M. Ahlrichs, S. Chan, A.J. Borst, Y.-J. Park, K.K. Lee, B.S. Graham, D. Veesler, N.P. King, and M. Kanekiyo, *Structure-based design of stabilized recombinant influenza neuraminidase tetramers.* Nat. Commun, 2022. **13**(1): p. 1825.
169. Mitnaul, L.J., M.R. Castrucci, K.G. Murti, and Y. Kawaoka, *The cytoplasmic tail of influenza A virus neuraminidase (NA) affects NA incorporation into virions, virion morphology, and virulence in mice but is not essential for virus replication.* J Virol, 1996. **70**(2): p. 873-9.
170. McKimm-Breschkin, J.L., J. Williams, S. Barrett, K. Jachno, M. McDonald, P.G. Mohr, T. Saito, and M. Tashiro, *Reduced susceptibility to all neuraminidase inhibitors of influenza H1N1 viruses with haemagglutinin mutations and mutations in non-conserved residues of the neuraminidase.* J Antimicrob Chemother, 2013. **68**(10): p. 2210-21.
171. Harris, A.K., J.R. Meyerson, Y. Matsuoka, O. Kuybeda, A. Moran, D. Bliss, S.R. Das, J.W. Yewdell, G. Sapiro, K. Subbarao, and S. Subramaniam, *Structure and accessibility of*

HA trimers on intact 2009 H1N1 pandemic influenza virus to stem region-specific neutralizing antibodies. Proc Natl Acad Sci U S A, 2013. **110**(12): p. 4592-7.

172. Rosu Miruna, E., A. Kok, M. Bestebroer Theo, D. de Meulder, P. Verveer Elwin, R. Pronk Mark, J.M. Dekker Lennard, M. Luider Theo, M. Richard, M.A. van den Brand Judith, A.M. Fouchier Ron, and S. Herfst, *Contribution of Neuraminidase to the Efficacy of Seasonal Split Influenza Vaccines in the Ferret Model.* Virol. J, 2022. **96**(6): p. e01959-21.
173. Smith, B.J., P.M. Colman, M. Von Itzstein, B. Danylec, and J.N. Varghese, *Analysis of inhibitor binding in influenza virus neuraminidase.* Prot. Sci, 2001. **10**(4): p. 689-696.
174. Laborda, P., S.Y. Wang, and J. Voglmeir, *Influenza Neuraminidase Inhibitors: Synthetic Approaches, Derivatives and Biological Activity.* Molecules, 2016. **21**(11).
175. Gubareva, L. and T. Mohan, *Antivirals Targeting the Neuraminidase.* Cold Spring Harb Perspect Med, 2022. **12**(1).
176. Hurt, A.C., K. Hardie, N.J. Wilson, Y.-M. Deng, M.O.B.N.G.D.L. Studies, N. Gehrig, and A. Kelso, *Community Transmission of Oseltamivir-Resistant A(H1N1)pdm09 Influenza.* N Engl J Med, 2011. **365**(26): p. 2541-2542.
177. Mooney, C.A., S.A. Johnson, P. ’t Hart, L. Quarles van Ufford, C.A.M. de Haan, E.E. Moret, and N.I. Martin, *Oseltamivir Analogues Bearing N-Substituted Guanidines as Potent Neuraminidase Inhibitors.* J. Med. Chem, 2014. **57**(7): p. 3154-3160.
178. Xie, Y., D. Xu, B. Huang, X. Ma, W. Qi, F. Shi, X. Liu, Y. Zhang, and W. Xu, *Discovery of N-Substituted Oseltamivir Derivatives as Potent and Selective Inhibitors of H5N1 Influenza Neuraminidase.* J. Med. Chem, 2014. **57**(20): p. 8445-8458.
179. Schade, D., J. Kotthaus, L. Riebling, J. Kotthaus, H. Müller-Fielitz, W. Raasch, O. Koch, N. Seidel, M. Schmidtke, and B. Clement, *Development of Novel Potent Orally Bioavailable Oseltamivir Derivatives Active against Resistant Influenza A.* J. Med. Chem, 2014. **57**(3): p. 759-769.
180. Fraser, B.H., S. Hamilton, A.M. Krause-Heuer, P.J. Wright, I. Greguric, S.P. Tucker, A.G. Draffan, V.V. Fokin, and K.B. Sharpless, *Synthesis of 1,4-triazole linked zanamivir dimers as highly potent inhibitors of influenza A and B.* MedChemComm, 2013. **4**(2): p. 383-386.
181. Yan, Z.-L., A.-Y. Liu, X.-X. Wei, Z. Zhang, L. Qin, Q. Yu, P. Yu, K. Lu, and Y. Yang, *Divalent oseltamivir analogues as potent influenza neuraminidase inhibitors.* Carbohydr. Res, 2019. **477**: p. 32-38.
182. Fu, L., Y. Bi, Y. Wu, S. Zhang, J. Qi, Y. Li, X. Lu, Z. Zhang, X. Lv, J. Yan, G.F. Gao, and X. Li, *Structure-Based Tetravalent Zanamivir with Potent Inhibitory Activity against Drug-Resistant Influenza Viruses.* J. Med. Chem, 2016. **59**(13): p. 6303-6312.
183. Kim, J.-H., R. Resende, T. Wennekes, H.-M. Chen, N. Bance, S. Buchini, A.G. Watts, P. Pilling, V.A. Streltsov, M. Petric, R. Liggins, S. Barrett, J.L. McKimm-Breschkin, M. Niikura, and S.G. Withers, *Mechanism-Based Covalent Neuraminidase Inhibitors with Broad-Spectrum Influenza Antiviral Activity.* Science, 2013. **340**(6128): p. 71-75.
184. Shelef, O., S. Gutkin, D. Feder, A. Ben-Bassat, M. Mandelboim, Y. Haitin, N. Ben-Tal, E. Bacharach, and D. Shabat, *Ultrasensitive chemiluminescent neuraminidase probe for rapid screening and identification of small-molecules with antiviral activity against influenza A virus in mammalian cells.* Chem. Sci, 2022. **13**(42): p. 12348-12357.
185. Lu, C.P., C.T. Ren, Y.N. Lai, S.H. Wu, W.M. Wang, J.Y. Chen, and L.C. Lo, *Design of a mechanism-based probe for neuraminidase to capture influenza viruses.* Angew Chem Int Ed Engl, 2005. **44**(42): p. 6888-92.

186. Minami, A., Y. Kurebayashi, T. Takahashi, T. Otsubo, K. Ikeda, and T. Suzuki, *The Function of Sialidase Revealed by Sialidase Activity Imaging Probe.* Int J Mol Sci, 2021. **22**(6).
187. Kurebayashi, Y., T. Takahashi, T. Otsubo, K. Ikeda, S. Takahashi, M. Takano, T. Agarikuchi, T. Sato, Y. Matsuda, A. Minami, H. Kanazawa, Y. Uchida, T. Saito, Y. Kawaoka, T. Yamada, F. Kawamori, R. Thomson, M. von Itzstein, and T. Suzuki, *Imaging of influenza virus sialidase activity in living cells.* Sci. Rep, 2014. **4**(1): p. 4877.
188. Agirre, J., M. Atanasova, H. Bagdonas, C.B. Ballard, A. Baslé, J. Beilsten-Edmands, R.J. Borges, D.G. Brown, J.J. Burgos-Mármol, J.M. Berrisford, P.S. Bond, I. Caballero, L. Catapano, G. Chojnowski, A.G. Cook, K.D. Cowtan, T.I. Croll, J. Debreczeni, N.E. Devenish, E.J. Dodson, T.R. Drevon, P. Emsley, G. Evans, P.R. Evans, M. Fando, J. Foadi, L. Fuentes-Montero, E.F. Garman, M. Gerstel, R.J. Gildea, K. Hatti, M.L. Hekkelman, P. Heuser, S.W. Hoh, M.A. Hough, H.T. Jenkins, E. Jiménez, R.P. Joosten, R.M. Keegan, N. Keep, E.B. Krissinel, P. Kolenko, O. Kovalevskiy, V.S. Lamzin, D.M. Lawson, A.A. Lebedev, A.G.W. Leslie, B. Lohkamp, F. Long, M. Malý, A.J. McCoy, S.J. McNicholas, A. Medina, C. Millán, J.W. Murray, G.N. Murshudov, R.A. Nicholls, M.E.M. Noble, R. Oeffner, N.S. Pannu, J.M. Parkhurst, N. Pearce, J. Pereira, A. Perrakis, H.R. Powell, R.J. Read, D.J. Rigden, W. Rochira, M. Sammito, F. Sánchez Rodríguez, G.M. Sheldrick, K.L. Shelley, F. Simkovic, A.J. Simpkin, P. Skubak, E. Sobolev, R.A. Steiner, K. Stevenson, I. Tews, J.M.H. Thomas, A. Thorn, J.T. Valls, V. Uski, I. Usón, A. Vagin, S. Velankar, M. Vollmar, H. Walden, D. Waterman, K.S. Wilson, M.D. Winn, G. Winter, M. Wojdyr, and K. Yamashita, *The CCP4 suite: integrative software for macromolecular crystallography.* Acta Crystallogr D Struct Biol, 2023. **79**(Pt 6): p. 449-461.
189. Matrosovich, M., T. Matrosovich, W. Garten, and H.D. Klenk, *New low-viscosity overlay medium for viral plaque assays.* Virol J, 2006. **3**: p. 63.
190. Tahara, N., I. Tachibana, K. Takeo, S. Yamashita, A. Shimada, M. Hashimoto, S. Ohno, T. Yokogawa, T. Nakagawa, F. Suzuki, and A. Ebihara, *Boosting Auto-Induction of Recombinant Proteins in Escherichia coli with Glucose and Lactose Additives.* Protein Pept Lett, 2021. **28**(10): p. 1180-1190.
191. Inc, T.F.S., *An efficient site-specific transposition system to generate baculovirus for high-level expression of recombinant proteins.* 2018.
192. Wilde, M., M. Klausberger, D. Palmberger, W. Ernst, and R. Grabherr, *Tnao38, high five and Sf9--evaluation of host-virus interactions in three different insect cell lines: baculovirus production and recombinant protein expression.* Biotechnol Lett, 2014. **36**(4): p. 743-9.
193. van der Vries, E., P.J. Collins, S.G. Vachieri, X. Xiong, J. Liu, P.A. Walker, L.F. Haire, A.J. Hay, M. Schutten, A.D. Osterhaus, S.R. Martin, C.A. Boucher, J.J. Skehel, and S.J. Gamblin, *H1N1 2009 pandemic influenza virus: resistance of the I223R neuraminidase mutant explained by kinetic and structural analysis.* PLoS Pathog, 2012. **8**(9): p. e1002914.
194. Air, G.M., *Influenza neuraminidase*. Influenza and Other Respiratory Viruses. Vol. 6. 2012: John Wiley & Sons, Ltd. 245-256.
195. Smith, B.J., T. Huyton, R.P. Joosten, J.L. McKimm-Breschkin, J.G. Zhang, C.S. Luo, M.Z. Lou, N.E. Labrou, and T.P. Garrett, *Structure of a calcium-deficient form of influenza virus neuraminidase: implications for substrate binding.* Acta Crystallogr D Biol Crystallogr, 2006. **62**(Pt 9): p. 947-52.

196. Han, N., J.T.Y. Ng, Y. Li, Y. Mu, and Z. Huang, *Plasticity of the 340-Loop in Influenza Neuraminidase Offers New Insight for Antiviral Drug Development.* Int J Mol Sci, 2020. **21**(16).
197. Checkmahomed, L., B. Padey, A. Pizzorno, O. Terrier, M. Rosa-Calatrava, Y. Abed, M. Baz, and G. Boivin, *In Vitro Combinations of Baloxavir Acid and Other Inhibitors against Seasonal Influenza A Viruses.* Viruses, 2020. **12**(10).
198. Hansen, T., A. Nin-Hill, J.D.C. Codée, T.A. Hamlin, and C. Rovira, *Rational Tuning of the Reactivity of Three-Membered Heterocycle Ring Openings via SN2 Reactions.* Chem. Eur. J. , 2022. **28**(60): p. e202201649.
199. Carugo, O. and K. Djinović-Carugo, *Structural biology: A golden era.* PLoS Biol, 2023. **21**(6): p. e3002187.
200. Kendrew, J.C., G. Bodo, H.M. Dintzis, R.G. Parrish, H. Wyckoff, and D.C. Phillips, *A Three-Dimensional Model of the Myoglobin Molecule Obtained by X-Ray Analysis.* Nature, 1958. **181**(4610): p. 662-666.
201. Frank, J., *Advances in the field of single-particle cryo-electron microscopy over the last decade.* Nat Protoc, 2017. **12**(2): p. 209-212.
202. Weissenberger, G., R.J.M. Henderikx, and P.J. Peters, *Understanding the invisible hands of sample preparation for cryo-EM.* Nature Methods, 2021. **18**(5): p. 463-471.
203. Doerr, A., *Single-particle cryo-electron microscopy.* Nature Methods, 2016. **13**(1): p. 23-23.
204. Lee, Y., K.Y. Oang, D. Kim, and H. Ihee, *A comparative review of time-resolved x-ray and electron scattering to probe structural dynamics.* Struct Dyn, 2024. **11**(3): p. 031301.
205. Orlova, E.V. and H.R. Saibil, *Structural analysis of macromolecular assemblies by electron microscopy.* Chem Rev, 2011. **111**(12): p. 7710-48.
206. Fernandez-Leiro, R. and S.H.W. Scheres, *Unravelling biological macromolecules with cryo-electron microscopy.* Nature, 2016. **537**(7620): p. 339-346.
207. Dubochet, J., J. Lepault, R. Freeman, J.A. Berriman, and J.C. Homo, *Electron microscopy of frozen water and aqueous solutions.* J. Microsc, 1982. **128**(3): p. 219-237.
208. Taylor, K.A. and R.M. Glaeser, *Electron diffraction of frozen, hydrated protein crystals.* Science, 1974. **186**(4168): p. 1036-7.
209. Shi, D. and R. Huang, *Analysis and comparison of electron radiation damage assessments in Cryo-EM by single particle analysis and micro-crystal electron diffraction.* Front Mol Biosci, 2022. **9**: p. 988928.
210. Cheng, Y., N. Grigorieff, P.A. Penczek, and T. Walz, *A primer to single-particle cryo-electron microscopy.* Cell, 2015. **161**(3): p. 438-449.
211. Thompson, R.F., M. Walker, C.A. Siebert, S.P. Muench, and N.A. Ranson, *An introduction to sample preparation and imaging by cryo-electron microscopy for structural biology.* Methods, 2016. **100**: p. 3-15.
212. Cheng, Y., *Single-Particle Cryo-EM at Crystallographic Resolution.* Cell, 2015. **161**(3): p. 450-457.
213. Dubochet, J., J.J. Chang, R. Freeman, J. Lepault, and A.W. McDowall, *Frozen aqueous suspensions.* Ultramicroscopy, 1982. **10**(1): p. 55-61.
214. Russo, C.J. and L.A. Passmore, *Electron microscopy: Ultrastable gold substrates for electron cryomicroscopy.* Science, 2014. **346**(6215): p. 1377-80.

215. Yip, K.M., N. Fischer, E. Paknia, A. Chari, and H. Stark, *Atomic-resolution protein structure determination by cryo-EM.* Nature, 2020. **587**(7832): p. 157-161.
216. Shen, P.I., J.; Brasch, J. *CryoEM 101*. 2024.
217. Carroni, M. and H.R. Saibil, *Cryo electron microscopy to determine the structure of macromolecular complexes.* Methods, 2016. **95**: p. 78-85.
218. Cong, Y. and S.J. Ludtke, *Single particle analysis at high resolution.* Methods Enzymol, 2010. **482**: p. 211-35.
219. Danev, R. and W. Baumeister, *Expanding the boundaries of cryo-EM with phase plates.* Curr Opin Struct Biol, 2017. **46**: p. 87-94.
220. Wade, R.H., *A brief look at imaging and contrast transfer.* Ultramicroscopy, 1992. **46**(1): p. 145-156.
221. van Heel, M., B. Gowen, R. Matadeen, E.V. Orlova, R. Finn, T. Pape, D. Cohen, H. Stark, R. Schmidt, M. Schatz, and A. Patwardhan, *Single-particle electron cryo-microscopy: towards atomic resolution.* QRB, 2000. **33**(4): p. 307-369.
222. Rohou, A. and N. Grigorieff, *CTFFIND4: Fast and accurate defocus estimation from electron micrographs.* J Struct Biol, 2015. **192**(2): p. 216-21.
223. Sigworth, F.J., *Principles of cryo-EM single-particle image processing.* Microscopy, 2016. **65**(1): p. 57-67.
224. Zivanov, J., T. Nakane, B.O. Forsberg, D. Kimanius, W.J. Hagen, E. Lindahl, and S.H. Scheres, *New tools for automated high-resolution cryo-EM structure determination in RELION-3.* Elife, 2018. **7**.
225. Nogales, E. and S.H. Scheres, *Cryo-EM: A Unique Tool for the Visualization of Macromolecular Complexity.* Mol Cell, 2015. **58**(4): p. 677-89.
226. Scheres, S.H.W., *Chapter Six - Processing of Structurally Heterogeneous Cryo-EM Data in RELION*, in *Methods in Enzymology*, R.A. Crowther, Editor. 2016, Academic Press. p. 125-157.
227. Toader, B.J., K.; Schwab, J.; Kimanius, D.; Pyle, E.; Nakane, T.; Scheres, S. . *Single particle tutorial*. RELION Documentation 2024 5th [Available from: https://relion.readthedocs.io/en/release-5.0/.
228. He, J., P. Lin, J. Chen, H. Cao, and S.-Y. Huang, *Model building of protein complexes from intermediate-resolution cryo-EM maps with deep learning-guided automatic assembly.* Nature Comms. , 2022. **13**(1): p. 4066.
229. Jumper, J., R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S.A.A. Kohl, A.J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A.W. Senior, K. Kavukcuoglu, P. Kohli, and D. Hassabis, *Highly accurate protein structure prediction with AlphaFold.* Nature, 2021. **596**(7873): p. 583-589.
230. Zivanov, J., T. Nakane, B.O. Forsberg, D. Kimanius, W.J.H. Hagen, E. Lindahl, and S.H.W. Scheres, *New tools for automated high-resolution cryo-EM structure determination in RELION-3.* eLife, 2018. **7**: p. e42166.
231. Burt, A., B. Toader, R. Warshamanage, A. von Kügelgen, E. Pyle, J. Zivanov, D. Kimanius, T.A.M. Bharat, and S.H.W. Scheres, *An image processing pipeline for electron cryo-tomography in RELION-5.* bioRxiv, 2024: p. 2024.04.26.591129.
232. Zivanov, J., T. Nakane, and S.H.W. Scheres, *A Bayesian approach to beam-induced motion correction in cryo-EM single-particle analysis.* IUCrJ, 2019. **6**(Pt 1): p. 5-17.

233. Emsley, P., B. Lohkamp, W.G. Scott, and K. Cowtan, *Features and development of Coot.* Acta Crystallogr., Sect. D: Struct. Biol., 2010. **66**: p. 486.

234. Liebschner, D., P.V. Afonine, M.L. Baker, G. Bunkóczi, V.B. Chen, T.I. Croll, B. Hintze, L.W. Hung, S. Jain, A.J. McCoy, N.W. Moriarty, R.D. Oeffner, B.K. Poon, M.G. Prisant, R.J. Read, J.S. Richardson, D.C. Richardson, M.D. Sammito, O.V. Sobolev, D.H. Stockwell, T.C. Terwilliger, A.G. Urzhumtsev, L.L. Videau, C.J. Williams, and P.D. Adams, *Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix.* Acta Crystallogr., Sect. D: Struct. Biol., 2019. **75**: p. 861.

235. Burmeister, W.P., B. Henrissat, C. Bosso, S. Cusack, and R.W.H. Ruigrok, *Influenza B virus neuraminidase can synthesize its own inhibitor.* Structure, 1993. **1**(1): p. 19-26.

236. Córdula, C.R., M.A. Lima, S.K. Shinjo, T.F. Gesteira, L. Pol-Fachin, V.J. Coulson-Thomas, H. Verli, E.A. Yates, T.R. Rudd, M.A. Pinhal, L. Toma, C.P. Dietrich, H.B. Nader, and I.L. Tersariol, *On the catalytic mechanism of polysaccharide lyases: evidence of His and Tyr involvement in heparin lysis by heparinase I and the role of Ca2+.* Mol Biosyst, 2014. **10**(1): p. 54-64.

237. Kato, K., N. Miyazaki, T. Hamaguchi, Y. Nakajima, F. Akita, K. Yonekura, and J.-R. Shen, *High-resolution cryo-EM structure of photosystem II reveals damage from high-dose electron beams.* Communications Biology, 2021. **4**(1): p. 382.

238. Ravelli, R.B. and S.M. McSweeney, *The 'fingerprint' that X-rays can leave on structures.* Structure, 2000. **8**(3): p. 315-28.

239. Yano, J., J. Kern, K.D. Irrgang, M.J. Latimer, U. Bergmann, P. Glatzel, Y. Pushkar, J. Biesiadka, B. Loll, K. Sauer, J. Messinger, A. Zouni, and V.K. Yachandra, *X-ray damage to the Mn4Ca complex in single crystals of photosystem II: a case study for metalloprotein crystallography.* Proc Natl Acad Sci U S A, 2005. **102**(34): p. 12047-52.

240. Warfield, K.L., E.M. Plummer, A.C. Sayce, D.S. Alonzi, W. Tang, B.E. Tyrrell, M.L. Hill, A.T. Caputo, S.S. Killingbeck, P.R. Beatty, E. Harris, R. Iwaki, K. Kinami, D. Ide, J.L. Kiappes, A. Kato, M.D. Buck, K. King, W. Eddy, M. Khaliq, A. Sampath, A.M. Treston, R.A. Dwek, S.G. Enterlein, J.L. Miller, N. Zitzmann, U. Ramstedt, and S. Shresta, *Inhibition of endoplasmic reticulum glucosidases is required for in vitro and in vivo dengue antiviral activity by the iminosugar UV-4.* Antiviral Res., 2016. **129**: p. 93.

241. Stavale, E.J., H. Vu, A. Sampath, U. Ramstedt, and K.L. Warfield, *In vivo therapeutic protection against influenza A (H1N1) oseltamivir-sensitive and resistant viruses by the iminosugar UV-4.* PLoS One, 2015. **10**(3): p. e0121662.

242. Liu, Y., M.P. Patricelli, and B.F. Cravatt, *Activity-based protein profiling: the serine hydrolases.* Proc Natl Acad Sci U S A, 1999. **96**(26): p. 14694-9.

243. Jiang, J., W.W. Kallemeijn, D.W. Wright, A. van den Nieuwendijk, V.C. Rohde, E.C. Folch, H. van den Elst, B.I. Florea, S. Scheij, W.E. Donker-Koopman, M. Verhoek, N. Li, M. Schürmann, D. Mink, R.G. Boot, J.D.C. Codée, G.A. van der Marel, G.J. Davies, J. Aerts, and H.S. Overkleeft, *In vitro and in vivo comparative and competitive activity-based protein profiling of GH29 α-l-fucosidases.* Chem Sci, 2015. **6**(5): p. 2782-false.

244. Jiang, J., C.L. Kuo, L. Wu, C. Franke, W.W. Kallemeijn, B.I. Florea, E. van Meel, G.A. van der Marel, J.D. Codée, R.G. Boot, G.J. Davies, H.S. Overkleeft, and J.M. Aerts, *Detection of Active Mammalian GH31 α-Glucosidases in Health and Disease Using In-Class, Broad-Spectrum Activity-Based Probes.* ACS Cent Sci, 2016. **2**(5): p. 351-8.

245. Jiang, J., T.J.M. Beenakker, W.W. Kallemeijn, G.A. van der Marel, H. van den Elst, J.D.C. Codée, J.M.F.G. Aerts, and H.S. Overkleeft, *Comparing Cyclophellitol N-Alkyl and N-*

Acyl Cyclophellitol Aziridines as Activity-Based Glycosidase Probes. Chem. Eur. J, 2015. **21**(30): p. 10861-10869.

246. Artola, M., C.-L. Kuo, L.T. Lelieveld, R.J. Rowland, G.A. van der Marel, J.D.C. Codée, R.G. Boot, G.J. Davies, J.M.F.G. Aerts, and H.S. Overkleeft, *Functionalized Cyclophellitols Are Selective Glucocerebrosidase Inhibitors and Induce a Bona Fide Neuropathic Gaucher Model in Zebrafish.* Journal of the American Chemical Society, 2019. **141**(10): p. 4214-4218.
247. Witte, M.D., W.W. Kallemeijn, J. Aten, K.Y. Li, A. Strijland, W.E. Donker-Koopman, A.M. van den Nieuwendijk, B. Bleijlevens, G. Kramer, B.I. Florea, B. Hooibrink, C.E. Hollak, R. Ottenhoff, R.G. Boot, G.A. van der Marel, H.S. Overkleeft, and J.M. Aerts, *Ultrasensitive in situ visualization of active glucocerebrosidase molecules.* Nat Chem Biol, 2010. **6**(12): p. 907-13.
248. Howard, S. and S.G. Withers, *Bromoketone C-Glycosides, a New Class of β-Glucanase Inactivators.* Journal of the American Chemical Society, 1998. **120**(40): p. 10326-10331.