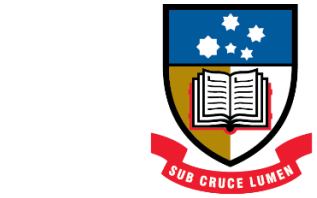
Probing a Promiscuous Binding Pocket of the Proteasome

Submitted in partial fulfilment of the degree

Master of Philosophy (Chemical Sciences)



THE UNIVERSITY *of* ADELAIDE

School of Physical Sciences

Department of Chemistry

Dion Joel Lammas Turner

B.Sc. (Adv.)

Supervisors:

Professor Andrew Abell

& Dr. Thomas Avery

November 2019

Table of Contents

Abstract

The proteasome is a multi-catalytic protease complex responsible for the degradation of unneeded, damaged or misfolded proteins. The proteasome is a validated target for the treatment of haematological malignancies such as multiple myeloma and mantle cell lymphoma, as demonstrated by the FDA approved proteasome inhibitors bortezomib, carfilzomib and ixazomib. These inhibitors, especially bortezomib, suffer from poor specificity and relatively high prevalence of resistance, therefore new inhibitors should be designed such that these characteristics improve. This thesis details probing of the relatively unexplored primed site binding channel of the β5 subunit of the proteasome with P_2 extended proteasome inhibitors. This work indicates the primed site binding channel as a promiscuous target for interaction which may aid in increasing the specificity of proteasome inhibitors

Chapter 1 introduces the structure and activity of the proteasome and its implications and relevance to human diseases. Inhibition of the proteasome by small molecule inhibitors is discussed, including the main classes, exemplary inhibitors, their mechanisms and applications. The primed site binding channel is then identified via examination of the crystal structure of the proteasome as a pocket which provides potential for new inhibitor-enzyme interactions.

Chapter 2 details the design, synthesis and evaluation of inhibitors **2.01**-**2.04** which probe the extent of the promiscuity of the primed site binding channel. The collection of published inhibitors which are known to, or are likely to, occupy the primed binding sites demonstrate the primed site binding channel as promiscuous regarding the substituents it accepts. The P_2 residue of bortezomib was identified as providing an access point to the primed binding sites. Imidazolyl and phenyl substituents were demonstrated to be accommodated by the primed site binding channel, with greater potency found for longer extensions into the pocket, or inhibitors with a phenyl substituent within the pocket.

Chapter 3 describes further probing of the primed site binding channel with the azobenzene-containing proteasome inhibitor **3.01**, which can be converted between *cis*- and *trans*-enriched isomeric states using light. The azobenzene substituent was placed at the P_2 position of a bortezomib-inspired inhibitor and allowed probing of the primed binding sites with greater conformation predictability. Remarkably, despite significant change in conformation between the *cis* and *trans* isomers, there is little difference between the low nanomolar range potencies of the isomeric states. This further indicates the significant promiscuity of the primed site binding channel.

Chapter 4 presents the evaluation of inhibitors **2.01**-**2.04** and the thermally adapted state of **3.01** alongside bortezomib against bovine α-chymotrypsin to examine the specificity of such inhibitors. Primed site-occupying inhibitors **2.01** and **2.04** demonstrate more than 2.5 times greater specificity towards the β5 subunit of the proteasome over α-chymotrypsin. This result indicates occupying the primed site binding channel as an effective strategy of improving proteasome inhibitor specificity, which may be critical in improving upon the currently available proteasome inhibitors.

Declaration

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint award of this degree.

I give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

I acknowledge the support I have received for my research through the provision of an Australian Government Research Training Program Scholarship.

Dion Turner

29/11/2019

Acknowledgements

To my supervisors, Professor Andrew Abell and Dr. Thomas Avery, I extend a most sincere thank you for your support and expert guidance over the past two years. Allowing me to communicate my ideas, suggesting new ones once I had run out and steering me back on track helped me complete the ever-changing goals of my project and degree. I also thank you for the guidance and assistance in writing this thesis, which would be non-existent without you.

Countless thanks to the members of the Abell Group for their friendship as well as scientific and non-scientific support. I greatly appreciate your assistance in becoming familiar with the environment of research and showing me the ropes in and around the lab. Special thanks to Kwang Jun Lee for his assistance with molecular docking, Kathryn Palasis and Dr. Michelle Zhang for their help with biological assays and Aimee Horsfall, Dr. Thomas Avery and Dr. John Horsley for running HRMS samples. I also thank all technical and administrative staff of The University of Adelaide, School of Physical Sciences and the chemistry department who have given me assistance throughout my degree.

To the Institute for Photonics and Advanced Sensing (IPAS), Centre for Nanoscale BioPhotonics (CNBP) and their members, thank you for the great opportunities, knowledge and friendships you have allowed me to gain over the past two years. I also acknowledge CNBP for their funding support, IPAS for their facilities and the Australian National Fabrication Facility for providing essential instruments.

To my family and friends, thank you for all the support and necessary distraction throughout my entire university experience which has brought me to this point. Thank you for listening to me express my passion for science, especially when you had no clue what I was talking about. A special “cheers” to beer club, for providing a debrief/therapy/chill session which I would always look forward to at the end of the week.

Specifically, to my parents: my appreciation for you goes beyond words. Your unwavering support and desire for me to achieve my goals has been a huge motivation. Thank you for everything you have done to help me reach this point in my life.

Finally, to my partner Caitlin: I owe a great amount of my success to your love and support. You have provided me with motivation even at the toughest points and helped me manage all the new experiences throughout out time together. Completing this thesis would have been a much longer and more tortuous experience if it were not for the support and motivation you provided.

Abbreviations

AMC: 7-amino-4-methylcoumarin

Boc: *tert*-butyloxycarbonyl

DCM: dichloromethane

DIPEA: *N*,*N*-diisopropylethylamine

DMF: *N*,*N*-dimethylformamide

DMSO; Dimethyl sulfoxide

DPPP: 1,3-bis(diphenylphosphino)propane

EDC: 3-(ethyliminomethyleneamino)-*N*,*N*-dimethylpropan-1-amine

FDA: Food and Drug Administration

HATU: 1-[bis(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-b]pyridinium 3-oxide hexafluorophosphate

HBTU: 3-[bis(dimethylamino)methyliumyl]-3*H*-benzotriazol-1-oxide hexafluorophosphate

HPLC: High-Performance Liquid Chromatography

IC_{50}: half maximal inhibitory concentration

MES: 2-(*N*-morpholino)ethanesulfonic acid

MHC-1: Major Histocompatability Complex class I

MS: Mass Spectrometry

NF-κB: Nuclear Factor kappa-light-chain-enhancer of B cells

NMR: Nuclear Magnetic Resonance

PDB: Protein Data Bank

PyBOP: (benzotriazol-1-yloxy)tripyrrolidinophosphonium hexafluorophosphate

RP-HPLC: Reverse Phase High-Performance Liquid Chromatography

TEV: Tobacco Etch Virus

THF: tetrahydrofuran

TLC: Thin-Layer Chromatography

UV: Ultraviolet

Chapter 1: Introduction

1.1 Proteases and Protease Inhibitors

Proteases catalyse the hydrolysis of amide bonds in proteins and peptides and they are essential for cellular processes such as signalling,[1-3] cell proliferation,[4,5] cell differentiation[6,7] and apoptosis.[8,9] They occupy 2% of the genomes in all living species[10] and are involved in such processes via degradation, modification, activation and deactivation of proteins.

A protease accommodates its substrate in an active site cleft, with the interaction defined by Schechter and Berger nomenclature[11] as depicted in Figure 1.01. The amino acid residues towards the N-terminus from the cleavage site are denoted as P_1-P_n, and the residues towards the C-terminus are denoted at P_1'-P_n'. The amide bond to be hydrolysed, the scissile bond, joins the P_1 and P_1' residues. The subsites of the protease are assigned as S_1-S_n and S_1'-S_n' to correspond with the residues which reside within each subsite.

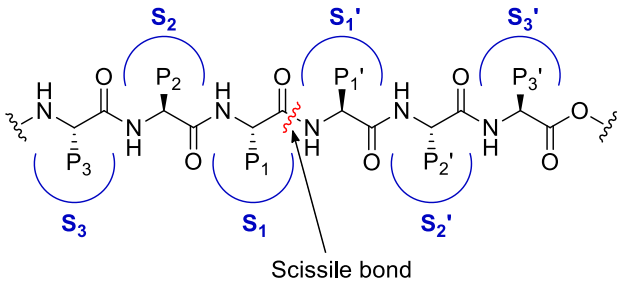


Figure 1.01 A schematic of a typical peptide bound to the active site of a protease, labelled with the Schechter and Berger[11] nomenclature. The peptide is represented in black with the amino acid side chains denoted as P_n-P_n', the enzyme subsites are represented in blue and denoted as S_n-S_n'.

Proteases recognise specific sequences of amino acids with specificity mainly dictated by the make-up and composition of the S_1 pocket.[12] This preference generally limits an protease to cleave a substrate on the C-terminal side of a specific type of amino acid residue. For example, trypsin cleaves on the C-terminal side of positively charged lysine or arginine residues,[13] while chymotrypsin cleaves on the C-terminal side of large, hydrophobic tyrosine, phenylalanine and

tryptophan residues.[14] However, some proteases exhibit striking specificity, for example nuclear-inclusion-a-endopeptidase, also known as TEV protease, cleaves on the C-terminal side of the glutamine (Gln) residue in the sequence Glu-Xaa-Xaa-Tyr-Xaa-Gln-Ser or Gly.[15]

The modulation of protease activity can result in significant changes in cellular functions given their key role in cellular processes. One such way of modulating protease activity is by directly inhibiting its ability to catalyse the amide bond hydrolysis reaction. Competitive reversible inhibitors function by competing for, and blocking, the active site of the protease. Synthetic protease inhibitors are commonly small, peptide-based or -inspired molecules which replicate or mimic the structure of natural substrates.[16] Many protease inhibitors contain a C-terminal electrophile that reacts covalently with the catalytic residue of the protease, with the resulting structure often mimicking the intermediate species in the amide bond hydrolysis.[17]

1.2 The Proteasome Structure and Activity

The protease responsible for the majority of cytosolic protein degradation in eukaryotes is the proteasome.[18] The proteasome is a protease complex which degrades unneeded, damaged or misfolded proteins, and is thus of great importance in maintaining intracellular protein homeostasis.[19] The finely tuned function of the proteasome also plays a critical role in immune responses,[20,21] cell cycle progression,[22,23] regulation of transcription,[24] and apoptosis.[25,26] Such is the proteasome's importance, it constitutes nearly 2% of all protein within cells.[27]

The 26S form of the proteasome is a 2.4 MDa complex consisting of a 20S core particle with 19S regulatory caps on one end (Figure 1.02). The 20S complex found in eukaryotes is arranged in a barrel-like structure, with two outer and two inner heptameric rings. The two outer rings, α1-α7, control the entry and release of substrates into the proteolytic barrel via its interactions with the 19S regulatory particle. The inner rings, β1-β7, form a chamber where the proteolytic activity occurs. Three of the β subunits are N-terminal threonine proteases and have well-

characterised protease activity.[28] The β1, β2 and β5 subunits are also known as the caspase-like, trypsin-like and chymotrypsin-like subunits, respectively. These subunits are named for their substrate specificity at the P_1 position. β1 cleaves preferentially at the C-terminal side of acidic residues, β2 cleaves preferentially at the C-terminal side of basic residues, and β5 cleaves preferentially at the C-terminal side of hydrophobic residues.[27,29] The substrate binding pockets responsible for proteolysis are defined by interactions between the β subunits.[27,29] As such, it is more accurate to label the proteasome not as a complex of individual proteases, but as a multi-catalytic proteolysis machine which only functions when the structure is complete. The advantage of containing multiple proteolytic subunits with varying specificities is that cellular proteins can be degraded with high processivity.

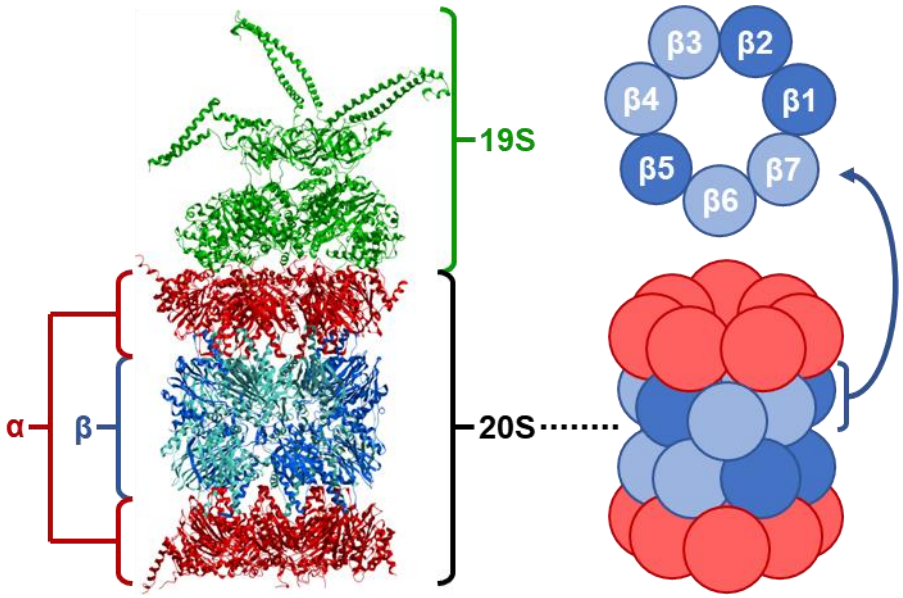
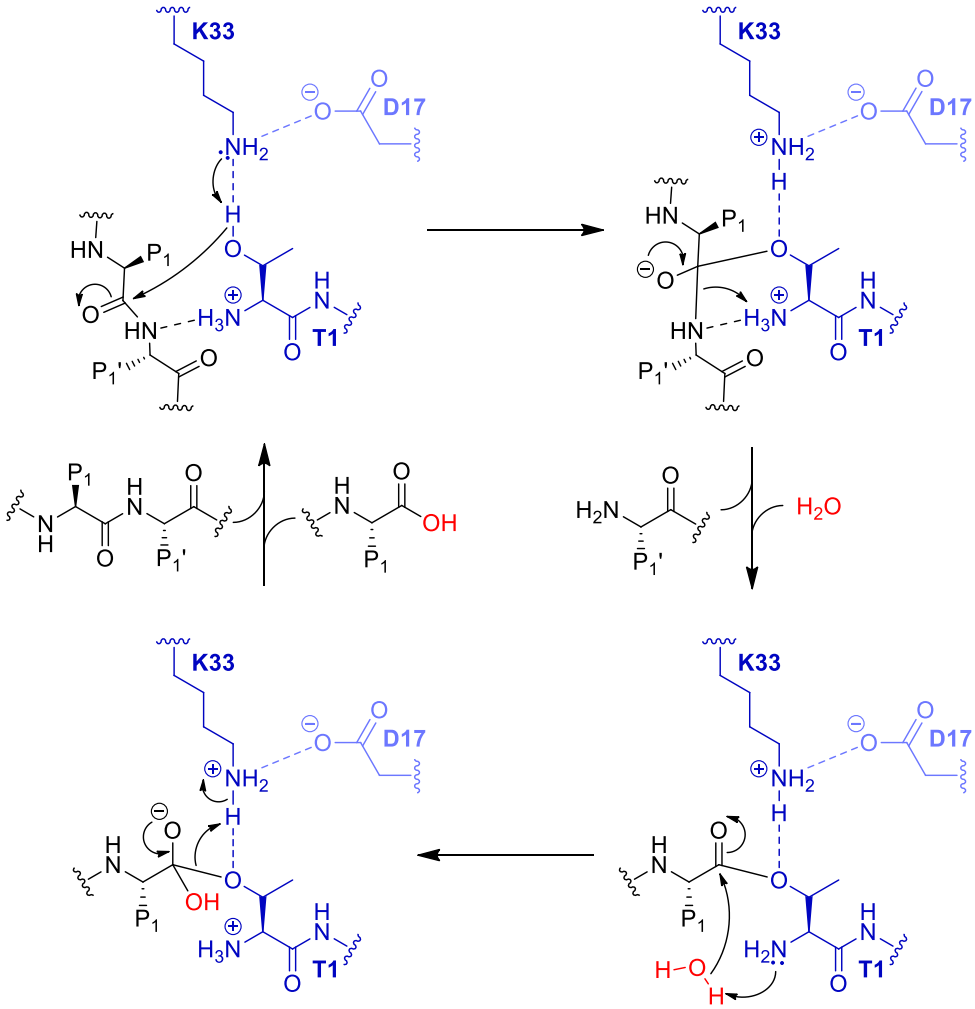


Figure 1.02 Forms of proteasome structure. Structure of the human 26S proteasome (PDB code: 5L4G)[30] consisting of the 20S core particle and the 19S regulatory cap with a cartoon representation showing the arrangement of proteasomal catalytic and non-catalytic β subunits.

Knowledge of the catalytic mechanism of the β1, β2 and β5 subunits is mainly informed by site-directed mutagenesis,[31] X-ray crystalography,[32] and enzyme inhibition studies.[33] A recently updated theory proposed by Huber *et al.*[32] suggests that the active sites of the various eukaryotic proteasomes contain a catalytic triad comprised of Thr1, Lys33 and Asp/Glu17 (Scheme 1.01).

Lys33-NH_2 is proposed to accept a proton from Thr1-OH to enable the nucleophilic attack by Thr1 upon the carbonyl carbon of the amide to be hydrolysed. Asp/Glu17 functions by orientating Lys33-NH_2 and stabilising the protonated form via hydrogen bonding. The positively charged N-terminal Thr1-NH_3^+ likely forms a hydrogen bond with the amide nitrogen, with donation of a proton to the resulting free amine to promote proteolysis. The resulting ester intermediate is then hydrolysed by a water molecule hydrogen bonded to the N-terminal free amine. A proton is donated by this water molecule to restore the N-terminus to its protonated form, and the Thr1 hydroxyl proton is returned from the protonated Lys33-NH_3^+ to restore the enzyme for the next cycle of proteolysis.



Scheme 1.01 Proposed catalytic mechanism of amide bond hydrolysis for the catalytic subunits of the proteasome.[32] The catalytic triad consisting of Thr1, Lys33 and Asp/Glu17 catalyses the hydrolysis of an amide bond with the aid of a water molecule. Thr1 is the active site nucleophile

whilst Lys33 removes a proton from the Thr1 hydroxyl and Asp/Glu17 stabilises the resulting positive charge on the Lys33 side-chain nitrogen.

Proteins to be degraded by the proteasome are tagged with the protein ubiquitin.[34] The subsequent steps towards degradation involve further ubiquitin conjugation to give a protein with a poly-ubiquitin chain attached. The 19S regulatory cap of the 26S proteasome recognises the poly-ubiquitin chain and allows the degradation of tagged proteins into short peptides, generally 6-9 amino acid residues in length, but in actual fact ranging from 3 to 22 amino acid residues.[35] These peptide fragments are further degraded by other proteases into individual amino acids or displayed to immune cells by major histocompatibility complex class I (MHC-I) receptors on the cell surface.[36] The 2004 Nobel Prize in Chemistry was awarded to Avram Hershko, Aaron Ciechanover and Irwin Rose for the discovery and characterisation of the ubiquitin-mediated protein degradation pathway.

As noted above, many cellular functions depend on the proteasome[27,28] (Figure 1.03) and abnormal proteasome function is thus linked to a number of diseases in humans. Aggregation of abnormal proteins occurs when the proteasome is dysfunctional, a key causative factor for Alzheimer's, Parkinson's and Huntington's disease, as well as other neurodegenerative disorders and cataract.[37] Proteasome dysfunction in cardiac tissue can cause congestive cardiac failure, ventricular hypertrophy, and ischaemia.[38] The proteasome plays a role in the activation of the cytokine expression regulator and anti-apoptotic protein complex, Nuclear Factor kappa-light-chain-enhancer of B cells (NF-κB).[39] Increased proteasome activity thus increases the activation of NF-κB, which is implicated in causing inflammatory[40] and autoimmune diseases.[41] Additionally, production and activity of the proteasome in cancer cells is increased, thus aiding in malignant transformation and the degradation of pro-apoptotic factors.[42]

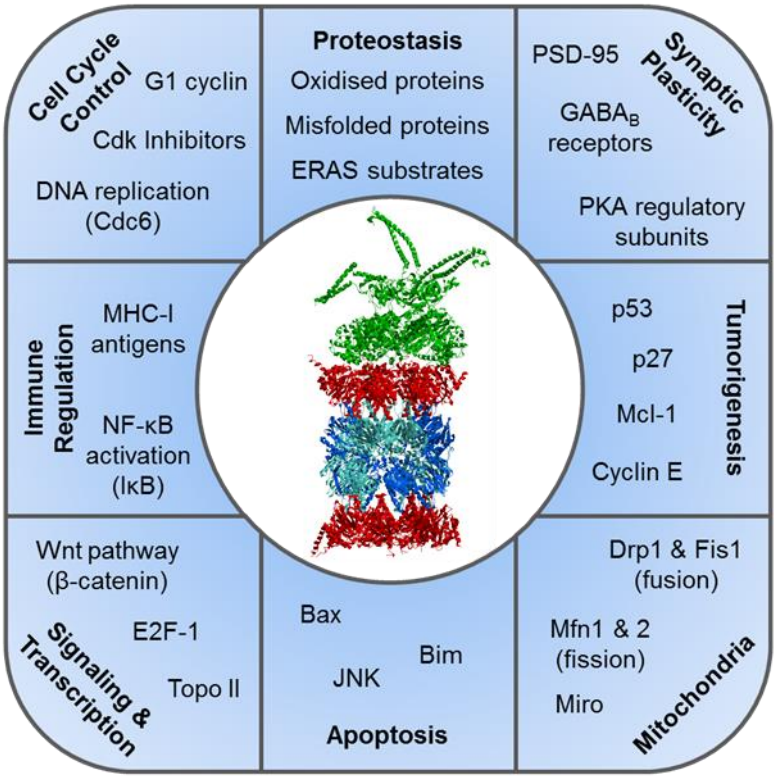


Figure 1.03 Areas of cellular functions where the proteasome plays a crucial role. Bax, bcl-2-like protein 4; Bim, bcl-2-like protein 11; Cdk, cyclin-dependent kinase; Drp1, dynamin-1-like protein; ERAD, endoplasmic-reticulum-associated protein degradation; E2F-1, target of retinoblastoma protein; Fis1, mitochondrial fission 1 protein; GABA, gamma-aminobutyric acid; JNK, C-Jun-amino-terminal kinase; Mfn, mitofusin; MHC-I, major histocompatibility complex-I; Miro, mitochondrial Rho GTPase; NF-kB, nuclear factor–kB; PKA, protein kinase A; PSD-95, postsynaptic density protein 95; Topo II, type II topoisomerase; Wnt, wingless-type. Adapted from Thibaudeau and Smith.[27]

1.3 Inhibition of the Proteasome

Identification of the first proteasome inhibitors allowed interrogation of proteasome function in sophisticated biological systems, to significantly advance knowledge of pathogenesis of diseases, cell cycle regulation, immune surveillance and protein aggregation.[27,43] Early inhibitors were short peptide aldehydes, with the exemplary MG-132 (Figure 1.04) still commonly used in research today as a cheap, potent and reversible inhibitor of the proteasome.

Proteasome inhibitors were hypothesised to be susceptible to cause apoptosis of neoplastic and malignant cancer cells, as these cells have fewer checkpoint mechanisms that prevent apoptosis.[44] Such inhibitors were subsequently shown to be cytotoxic towards transformed malignant cell cultures with significantly less effect on healthy and non-transformed cultures.[45,46] With MG-132 as a lead compound, the company MyoGenics developed a dipeptide boronic acid, PS-341 (Figure 1.04), which was shown to inhibit the β5 active site (as well as the β2 subunit to a lesser extent) in a slowly reversible manner.[47,48]

In phase I clinical trials against multiple myeloma, PS-341 extraordinarily caused a complete response in one patient.[49] Multiple myeloma is a malignancy of plasma cells which at the time lacked viable treatment options and is currently considered incurable. After displaying further remarkable results in phase II clinical trials against multiple myeloma,[50] The FDA approved PS-341 in 2003 as a third-line treatment for multiple myeloma.[51] PS-341 was renamed bortezomib (by which it shall be referred to henceforth) and marketed as VELCADE® by Millenium Pharmaceuticals. The development of bortezomib revolutionised the treatment of multiple myeloma and it is now used in the first-line treatment of multiple myeloma and mantle-cell lymphoma, as well as retreatment of relapsed adult multiple myeloma patients who have previously responded well to bortezomib.[52]

The clinical success of bortezomib is limited by multiple factors including resistance,[53] dose-limiting toxicities[54] and a poor pharmacokinetic profile.[55] Development of the inhibitor chemistry and the understanding of the catalytic mechanism unique to the proteasome has enabled the design and discovery of second-generation proteasome inhibitors which possess improved characteristics. The following section discusses the major classes of available proteasome inhibitors and their chemistry, pharmacology and utility in both research and clinical settings.

1.3.1 Classes of Proteasome Inhibitors

Despite containing three unique active sites, specific inhibition of the β5 subunit of the proteasome is sufficient to disrupt protein degradation.[56] Conversely, specific inhibition of either the β1 or β2 subunit does not result in a substantial reduction of protein degradation.[57] As a result, the majority of proteasome inhibitors preferentially inhibit the β5 subunit, with some activity against β1 and even less activity against β2.

The vast majority of proteasome inhibitors react with the catalytic N-terminal threonine residue of the proteasome to form reversible or irreversible covalent adducts. Thus, the reactivity of the inhibitor warhead is a large factor in the inhibitor's activity.[47] The following sections discuss the four main classes of proteasome inhibitors differentiated by the C-terminal warhead: Peptide aldehydes, peptide boronates, epoxyketones and β-lactones, see Figure 1.04.

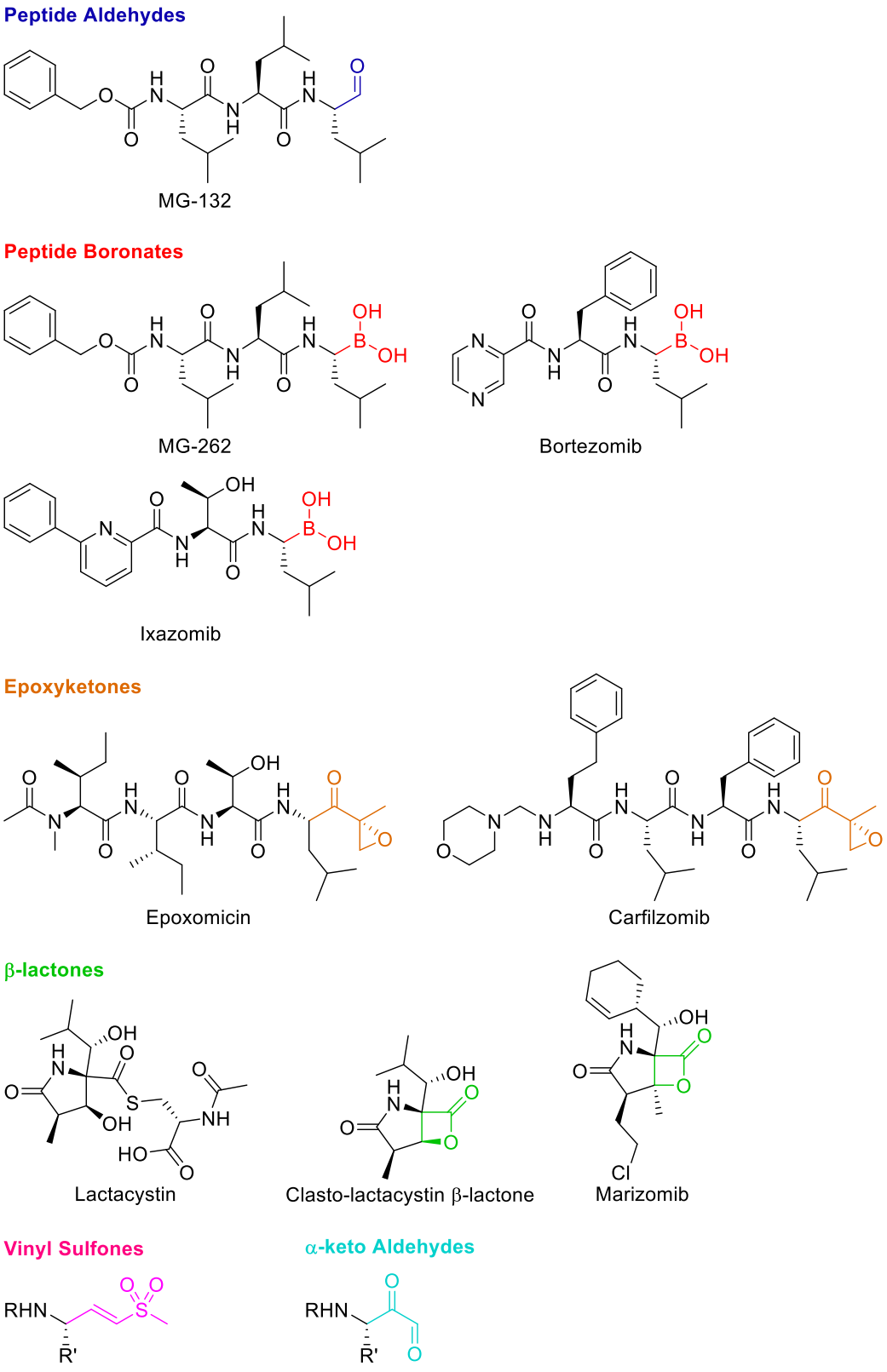
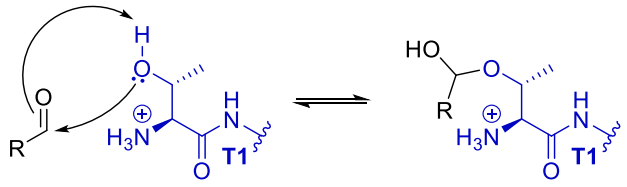


Figure 1.04 Structures of proteasome inhibitors discriminated by their warhead.

A. Peptide Aldehydes

MG-132 exemplifies the peptide aldehyde class of inhibitors and its common limitations, such as rapid metabolism or oxidation and low specificity towards the proteasome over other proteases such as cathepsins and calpains.[27] As such, MG-132 and other peptide aldehydes are not suitable for the treatment of disease. However, MG-132 remains as a useful tool in research as it is cell permeable, slow binding and fast dissociating, which allows the quick reversal of inhibition by changing to media without inhibitor.[58] The active site Thr1 hydroxyl reacts with the C-terminal aldehyde of MG-132 to give a hemiacetal (Scheme 1.02).[59] This reaction is reversible and results in a species that mimics a tetrahedral intermediate in the hydrolysis of an amide bond.



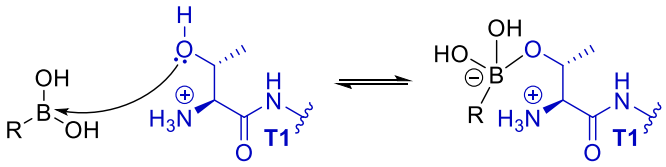
Scheme 1.02 Mechanism of covalent proteasome inhibition by peptide aldehydes.[59] The peptide aldehyde (black) is attacked by the Thr1 (blue) hydroxyl to form a hemiacetal.

B. Peptide Boronates

The peptide boronate class of proteasome inhibitors was developed to improve on the poor specificity and *in vivo* stability of peptide aldehydes.[47,48] The boronic acid warhead also imparts greater potency towards the proteasome relative to peptide aldehydes.[47] For example, the boronic acid analogue of MG-132, MG-262 (Figure 1.04), has 100-fold greater potency in comparison to MG-132.[47] Furthermore, peptide boronates have greater metabolic stability and do not inhibit cysteine proteases.

Similar to the covalent binding of peptide aldehydes to the proteasome, peptide boronates form a boron-oxygen bond between the boronic acid and the Thr1 hydroxyl (Scheme 1.03).[60] Again, the four-coordinate boron atom mimics the tetrahedral intermediate in amide bond hydrolysis,

but this species usually has dissociation half-lives on the order of hours and is thus considered slowly reversible.



Scheme 1.03 Mechanism of covalent proteasome inhibition by peptide boronates.[60] The peptide boronate (black) is attacked by the Thr1 (blue) hydroxyl to form a four-coordinate boron species.

Bortezomib has a small difference between therapeutic and fatal doses, relatively high prevalence of resistance and the dose limiting side-effect of peripheral neuropathy.[61] Peripheral neuropathy is a condition where the peripheral nerves are damaged in a likely irreversible fashion. This can cause tingling, numbness, severe pain and decreased mobility. Up to 55% of patients experience peripheral neuropathy as a result of bortezomib treatment which, depending on the severity, can require a reduction or cessation of treatment.[62] Furthermore, anaemia, neutropenia, leukopenia, thrombocytopenia, fatigue and nausea/vomiting are side effects experienced by greater than 20% of patients receiving bortezomib.[63] The cause of bortezomib's side effects are likely not just proteasome inhibition in healthy tissue, but also inhibition of other proteases. *In vitro* studies demonstrated that bortezomib causes peripheral neuropathy in a proteasome-independent manner. A possible mechanism for bortezomib-induced peripheral neuropathy involves the inhibition of HtrA2, a serine protease known to play a role in neuronal survival. Bortezomib also inhibits many other serine proteases including chymotrypsin, chymase, cathepsin A, cathepsin G and dipeptidyl-peptidase II.

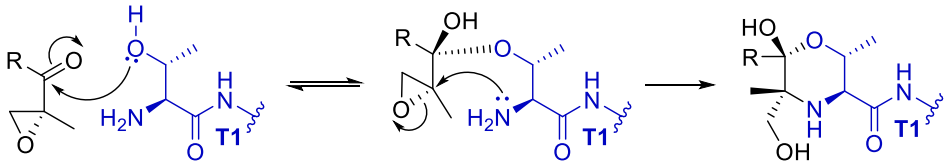
Resistance of multiple myeloma cells against bortezomib can significantly reduce the efficacy of bortezomib against multiple myeloma. The mechanisms currently associated with bortezomib resistance are mutations within the binding pocket of the β5 subunit as well as changes in the expression and subunit composition of the proteasome.[64,65] Resistance can also

arise when levels of antiapoptotic proteins are increased and proapoptotic proteins are decreased in response to the proteasome no longer degrading the proapoptotic factors.[66]

The peptide boronate ixazomib (NINLARO®) (Figure 1.04), identified as alternative to bortezomib, is an orally available proteasome inhibitor FDA approval for the treatment of multiple myeloma in adult patients who have received at least one prior treatment. Ixazomib is administered as a citrate boronic ester, which hydrolyses to the boronic acid in stomach acid or plasma, the demonstrated drug distribution of ixazomib is five times greater than that of bortezomib.[67] Ixazomib also reduces the incidence of peripheral neuropathy in patients to 28%,[68] which overall results in more patients receiving the best dose to treat multiple myeloma. Patients who have multiple myeloma resistant against bortezomib can be effectively treated by ixazomib, which does not have a frequency of resistance as high as bortezomib.[69] This suggests the characteristics imparted upon the inhibitor by the P_2 residue has great effect on the pharmacokinetic and pharmacodynamic profile of the drug.

C. Epoxyketones

Epoxyketones demonstrate significant specificity towards the proteasome. Epoxomicin (Figure 1.04), a metabolite from a strain of *Antinomyces*, is a tetrapeptide containing a C-terminal α',β'-epoxyketone moiety which displayed antitumour activity[70] later discovered to arise from proteasome inhibition.[71] The α',β'-epoxyketone moiety imparts great affinity and specificity towards the proteasome. The unusual binding mechanism (Scheme 1.04) of epoxyketone proteasome inhibitors was identified by examining the crystal structure of the 20S proteasome in complex with Epoxomicin.[72] The Thr1-OH attacks the carbonyl of the α',β'-epoxyketone followed by attack of the α'-carbon to irreversibly form a morpholino adduct.

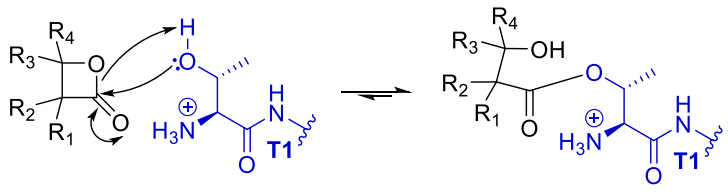


Scheme 1.04 Mechanism of covalent proteasome inhibition by epoxyketones.[72] The α',β'-epoxyketone (black) forms a morpholino adduct with the catalytic N-terminal threonine of the proteasome (blue).

The epoxyketone carfilzomib (KYPROLIS®) (Figure 1.04) is a second-generation inhibitor of the proteasome approved by the FDA to treat patients with relapsed or refractory multiple myeloma who have also received one to three previous treatments. Carfilzomib does not inhibit cysteine proteases and most serine proteases, and thus has fewer off target effects.[73] Peripheral neuropathy is only experienced by up to 12% of patients receiving carfilzomib,[74] allowing more patients to receive the optimum dose to treat their disease. This indicates that inhibitors with greater specificity towards the proteasome over serine proteases are less likely to give rise to peripheral neuropathy in patients.

D. β-lactones

The first non-peptidic compound found to inhibit the proteasome was lactacystin (Figure 1.04), a metabolite produced by *Streptomyces* bacteria.[75] Lactacystin itself does not inhibit the proteasome and is not a β-lactone, but is converted to the active inhibitor, clasto-lactacystin β-lactone (Figure 1.04), in aqueous environments at neutral pH.[76,77] The β-lactone is ring-opened via nucleophilic attack by the Thr1 hydroxyl, thus acylating the proteasome and mimicking the acyl intermediate of amide hydrolysis (Scheme 1.05).[78]

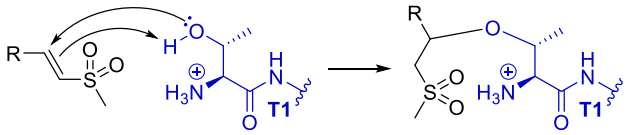


Scheme 1.05 Mechanism of covalent proteasome inhibition by β-lactones.[78]

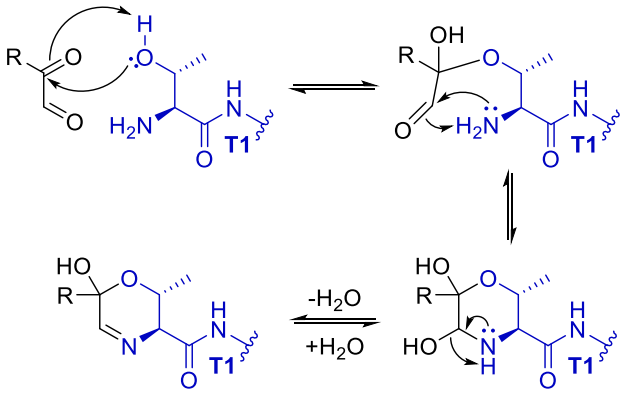
E. Other Classes

A range of peptidic and non-peptidic compounds have been demonstrated to inhibit the proteasome in a non-covalent manner.[33,79-81] Unlike bortezomib, these non-covalent inhibitors are generally unable to accomplish total proteasome inhibition in cancerous cells.[33,82] However, the partial inhibition is still able to perturb the proliferative nature of cancer cells, whilst reducing cytotoxicity towards healthy cells relative to covalent inhibitors.[33,82] Despite these promising characteristics, there has been no evaluation of non-covalent proteasome inhibitors in clinical phase trials, suggesting poor translation of *in vitro* results into *in vivo* studies.

Vinyl sulfones act as Michael acceptors and covalently attach to the catalytic Thr1 hydroxyl of the proteasome via conjugate addition (Scheme 1.07).[83] The use of vinyl sulfones as proteasome inhibitors is generally limited to proteasome activity probes. The binding mechanism (Scheme 1.08) of the reversible α-Keto aldehydes class of inhibitors is related to that of epoxyketones.[84] The α-Keto aldehyde warhead first reacts with the Thr1 hydroxyl and then the N-terminal amine to form a 5,6-dihydro-2*H*-1,4-oxazine ring.



Scheme 1.07 Mechanism of proteasome inhibition by vinyl sulfones occurs via Conjugate addition of the hydroxyl from the catalytic threonine residue (blue) to the vinyl sulfone (black).[83]



Scheme 1.08 Mechanism of proteasome inhibition by α-keto aldehydes.[84] The catalytic threonine residue of the proteasome (blue) forms a 5,6-dihydro-2*H*-1,4-oxazine ring with the α-ketoaldehyde (black).

1.3.2 Primed Site Binding Channel of the β5 Subunit of the Proteasome

Across the nearly three decades of proteasome inhibitor research, the evaluation of the C-terminal warheads discussed above and the characteristics of inhibitors resulting from such warheads has been extensive. Furthermore, as most inhibitors occupy the S_1-S_3 pockets of the β5 subunit of the proteasome, a great deal is known about the effect P_1-P_3 residues of proteasome inhibitors have on the affinity and specificity towards subunits of the proteasome. The P_1 and P_3 residues have the greatest effect on inhibitor affinity and specificity due to the well-defined β5 S_1 and S_3 pockets.[27] The β5 S_2 pocket is partially solvent exposed and can accommodate structures with varying sizes and functional groups.[27,85]

Considerably less is known about the primed site binding pockets of the β5 subunit of the proteasome. Inhibitors which occupy the primed sites mimic the residues of substrates which are downstream from the scissile bond. Generally, proteasome inhibitor which occupy the primed sites are not only more specific towards the proteasome over other proteases, but they have relatively high specificity for one proteasomal subunit.[86] In the case of such primed site-occupying inhibitors, S_1 is usually occupied by a small hydrophobic group, with the remaining structure extending away from the other non-primed sites and towards the primed sites. This

demonstrates that specificity and affinity towards the proteasome can be achieved without occupying all of S_1-S_3.

The structure of the β5 subunit primed site binding channel is well-understood due to X-ray crystallography.[87] Figure 1.05 shows the primed site binding channel splits into two pockets not very far downstream from the catalytic Thr1 residue. These two pockets either side of the channel have very different compositions, one being a large hydrophobic pocket comprised of tyrosine, phenylalanine and leucine. The other pocket contains the hydrophilic residues tyrosine, aspartic acid, serine and glutamic acid, which can form strong secondary interactions such as hydrogen bonding and salt bridges. The proximity of the primed pockets to this residue allows access to these primed binding sites with a small-molecule inhibitor. Accessing and interacting with these unique pockets is a strategy which could improve both the affinity and specificity for the β5 subunit of the proteasome.[86]

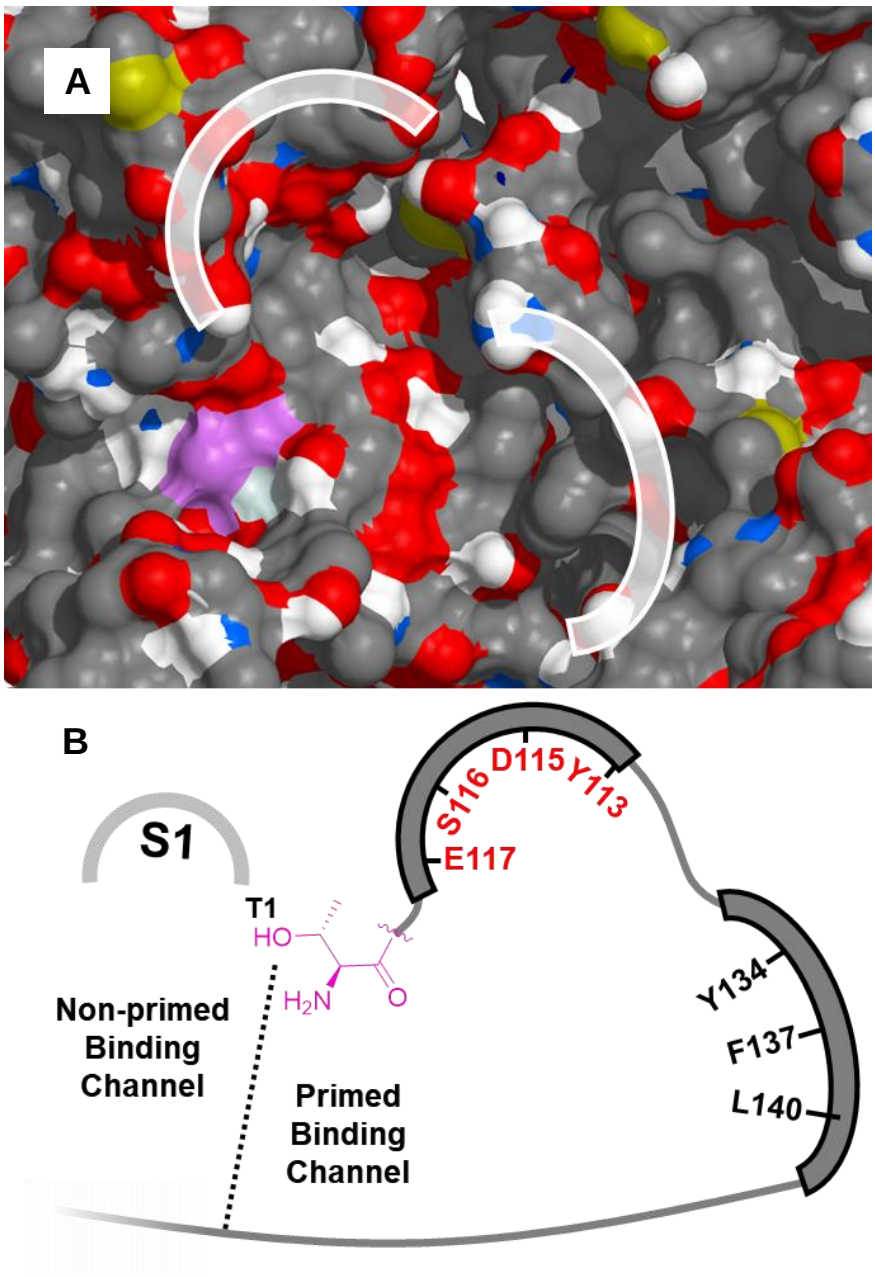


Figure 1.05 (A) X-ray structure (PDB code: 5L4G) and (B) cartoon representation of the primed site binding channel of the β5 subunit of the human proteasome. The catalytic threonine residue (Thr1) is highlighted in pink. The primed site binding channel splits into two pockets, a large hydrophobic pocket containing a tyrosine, phenylalanine and leucine, the other a smaller hydrophilic pocket containing a tyrosine, aspartic acid, serine and glutamic acid.

1.5 Thesis Aims and Overview

The primed binding sites of the β5 subunit of the proteasome are relatively unexplored by inhibitors and provide an opportunity for favourable enzyme-inhibitor interactions. Examining the affinity of inhibitors with different groups occupying the primed binding sites can inform

the future design of proteasome inhibitors such that they have high affinity and specificity towards the proteasome. As a result, new therapeutic treatments for haematological malignancies with improved pharmacokinetic and pharmacodynamic properties may emerge, especially those with reduced off-target effects.

This thesis describes the design, synthesis and evaluation of proteasome inhibitors intended to, as a collective, probe the promiscuity of the primed site pockets. The development of bortezomib-inspired proteasome inhibitors **2.01**-**2.04** with extended P_2 residues to access the primed binding sites is describes in Chapter 2. The primed site binding channel is relatively accommodating to imidazolyl substituents, suggesting it can be utilised in the design of proteasome inhibitors with greater specificity. Chapter 3 includes the design, synthesis and evaluation of **3.01**, a photoswitchable inhibitor of the proteasome, in order to further define the promiscuity of the primed site binding channel. Despite the significant change in geometry and dipole moment associated with the *cis* and *trans* isomers, the *cis*- and *trans*-enriched states show a remarkable similarity in potency. This further reinforces the promiscuity of this binding pocket and suggests there is significant scope for further modifications within the primed site binding channel. An investigation of specificity towards the proteasome over the serine protease, α-chymotrypsin, is detailed in chapter 4. Compounds **2.01**, **2.03** and **2.04** are greater than 2.5-fold more specific than bortezomib for the β5 subunit of the proteasome over α-chymotrypsin. This confirms occupying the promiscuous primed site binding channel as an effective strategy for improving the specificity of proteasome inhibitors towards the proteasome over other proteases.

Chapter 2: Probing the Primed Sites of the Proteasome

2.1 Introduction

The earliest evidence towards the primed binding pockets being a viable target for inhibitor binding came from a co-crystal structure of the 20S proteasome with epoxomicin.[72] Although epoxomicin does not itself occupy any of the primed site binding pockets, a component of the crystallisation buffer, 2-(*N*-morpholino)ethanesulfonic acid (MES), occupied the aforementioned smaller, hydrophilic primed binding site whilst hydrogen bonding with epoxomicin (Figure 2.01). Such a result indicates that this binding site should be further examined to develop inhibitors targeting the primed site binding channel.

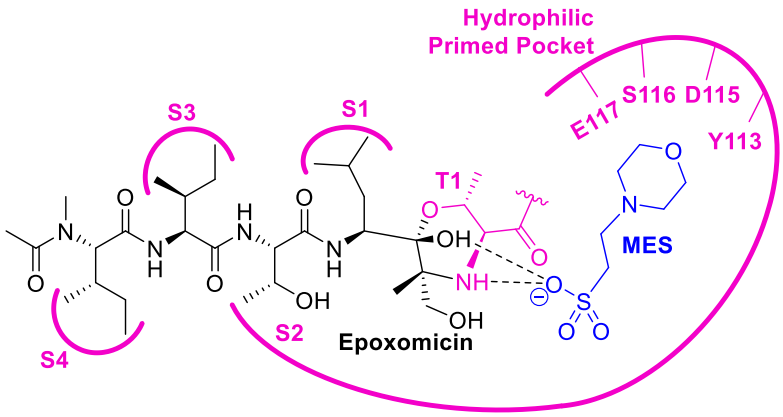
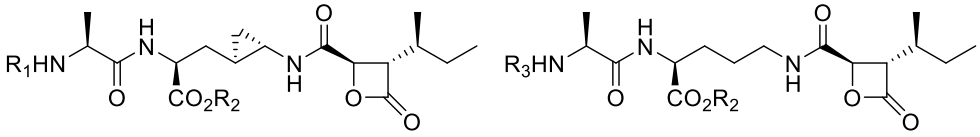


Figure 2.01 Schematic representation of the electron density map from Groll *et al.*[33] demonstrating epoxomicin complexed by the β5 subunit of the yeast 20S proteasome and MES occupying the small, hydrophilic primed binding pocket. The occupation of the pocket by MES indicates this pocket may have the ability to form favourable interactions with an inhibitor.

The first group of proteasome inhibitors shown by X-ray crystallography to occupy the primed binding sites of the β5 subunit of the proteasome are derivatives of the natural product belactosin A. The *Streptomyces sp*. β-lactone metabolites belactosin A and C (Figure 2.02) showed promising antitumour activity[88] later elucidated to arise from inhibition of the β5 subunit of the proteasome.[89] Belactosin A has reasonable affinity (IC_{50} = 0.21 μM) towards the β5 subunit of the proteasome; however, poor cell permeability gave rise to poor growth-

inhibition of HeLa S3 cells (IC_{50} = 51 μM). Modification of the free carboxylic acid group of belactosin A to a benzyl ester created KF33955 (Figure 1.07), a cell-permeable inhibitor with greater proteasome β5 subunit affinity (IC_{50} = 0.048 μM) and potent growth-inhibitory activity of HeLa S3 cells (IC_{50} = 0.46 μM).[89] The impressive results of this primed site-occupying inhibitor confirms the primed site binding channel as a target for beneficial interactions.



Belactosin A (R_1 = R_2 = H)
KF33955 (R_1 = Cbz, R_2 = Bn)

Belactosin C (R_3 = R_4 = H)
Belactosin C derivative 1
(R_3 = 2-naphthylmethoxycarbonyl, R_4 = Bn)

Figure 2.02 Structures of primed site-occupying inhibitors belactosin A and C and their respective derivatives KF33955 and belactosin derivative 1.

The binding mode of the di-protected belactosin C derivative 1 (Figure 2.02) to the β5 subunit of the proteasome was revealed by X-ray crystallography and demonstrated that it occupies the primed site binding channel (Figure 2.03).[90] The *sec*-butyl group attached to the α-carbon of the β-lactone occupies the S1 pocket of the β5 subunit, whilst the primed site binding channel is occupied by the bulky aryl substituents. A series of subsequent structure-activity relationship studies[89,91] elucidated that unnatural *cis*-cyclopropane derivatives improve upon the affinity of the natural *trans* configuration of the cyclopropane found in belactosin A and its previous derivatives. This series of belactosin A *cis*-cyclopropane derivatives occupied the smaller, hydrophilic pocket with hydrophobic groups such as phenyl (belactosin A derivative 1, Figure 2.05) or naphthoyl instead of structures which can form stronger secondary interactions with the hydrophilic residues in the pocket (Figure 2.04).[91] The significant difference in binding mode between the inhibitors shown in Figure 2.04 despite their similar potency demonstrates the promiscuity of the primed site binding channel.

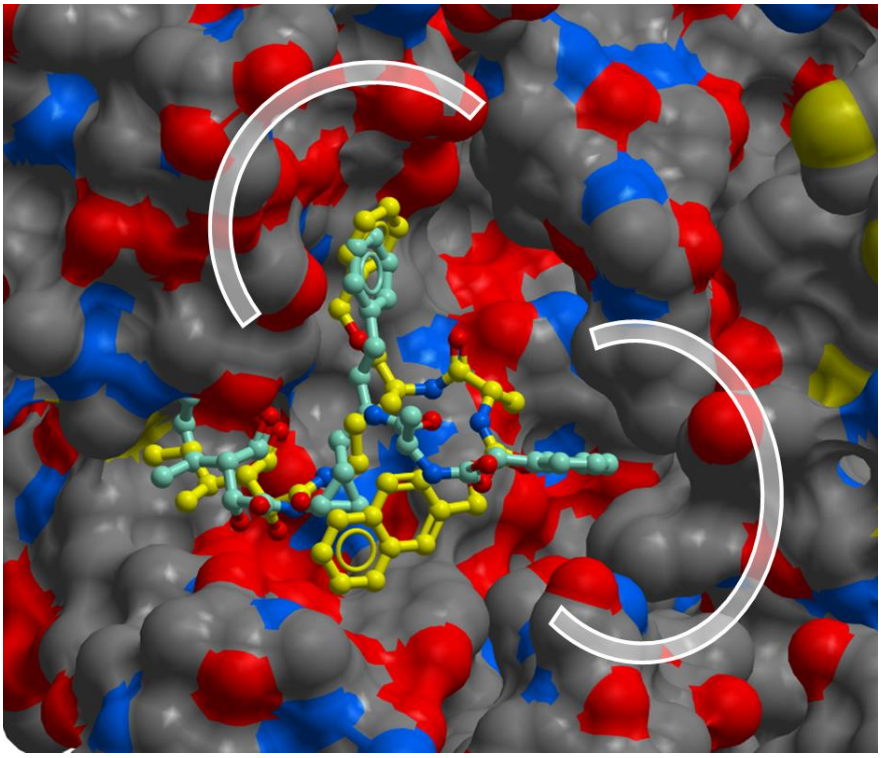


Figure 2.04 Co-crystal structure of the β5 subunit of the yeast 20S proteasome in complex with the di-protected belactosin C derivative 1 (yellow) (PDB code: 3TDD) superimposed with belactosin A derivative 1 (teal) (PDB code: 4J70), an un-natural *cis*-cyclopropane derivative of belactosin A. belactosin C derivative 1 occupies the large hydrophobic pocket with a bulky naphthoyl protecting group whilst occupying the small hydrophilic pocket with a phenyl group. Belactosin A derivative 1 occupies the small hydrophilic and large hydrophobic primed pocket with hydrophobic phenyl groups.

The next iteration of this series of primed site-occupying inhibitors of the proteasome was belactosin A derivative 2, which contained a simplified primed site occupying inhibitor and had reduced proteasome affinity and cell growth inhibitory activity (Figure 2.05).[92] Potency was restored upon attachment of the primed site-occupying substituent of belactosin A derivative 2 with a peptide boronate backbone similar to that of bortezomib, creating the first primed site-occupying peptide boronate inhibitor of the proteasome (boronate inhibitor 1, Figure 2.05).[93] This inhibitor occupies both the small and large primed binding pockets with an identical benzyl ether group. Further simplification by removing one of these groups (boronate inhibitor 2, Figure 2.05) retained good affinity and specificity towards the β5 subunit of the proteasome[93]

indicating the promiscuous primed site binding channel does not require bulky substituents. However, the binding mode of the inhibitor's remaining benzyl ether group is not known, as it could reside within either of the aforementioned primed binding pockets.

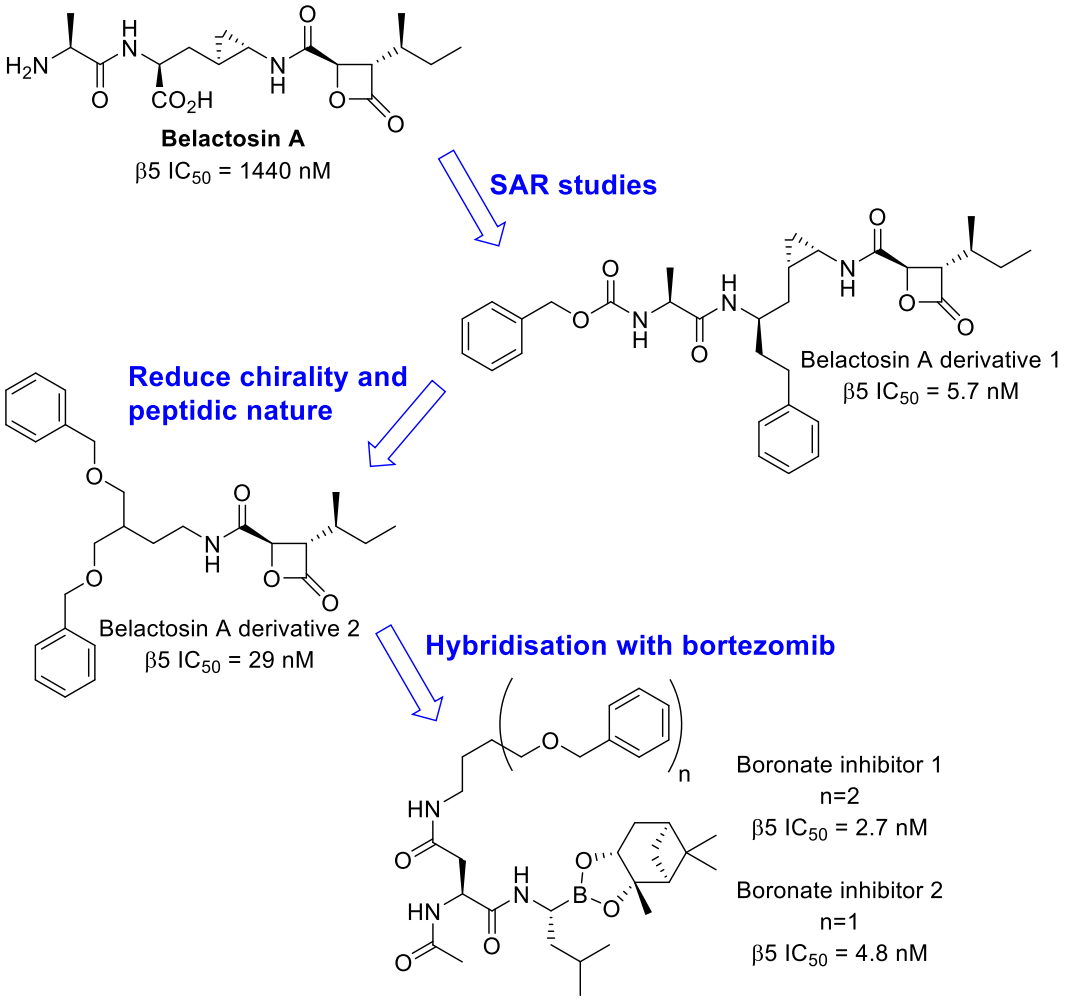


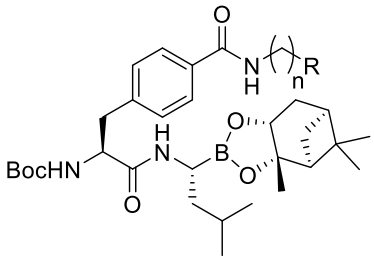
Figure 2.05 A series of proteasome inhibitors based off the natural product belactosin A and their IC_{50} values against the β5 subunit of the proteasome.

The primed site binding channel of the β5 subunit of the proteasome is promiscuous regarding the substituents it can accommodate. Thus far, primed site occupying inhibitors have placed hydrophobic aryl substituents within the pockets, neglecting to probe the limits of the promiscuity of the primed site binding channel. One of the pockets in the primed site binding channel contains the residues tyrosine, aspartic acid, serine and glutamic acid, all of which can participate in strong secondary interactions with an inhibitor, such as hydrogen bonding and

salt bridges. Despite this, published inhibitors which occupy the pocket have not attempted to take advantage of this opportunity.[91-99] As such, this chapter describes the probing of the primed site binding channel's promiscuity with P_2 extended bortezomib-inspired inhibitors.

2.2 Inhibitor Design

The set of inhibitors described in this chapter were designed to probe the promiscuous primed site binding channel of the β5 subunit of the proteasome. Bortezomib is an ideal choice for the inspiration of these inhibitors as it has been extensively studied during development, clinical trials and post-FDA approval, as well as providing access to the primed site binding channel via the phenylalanine residue. As such, the set of inhibitors (compounds **2.01**-**2.04**) shown in Figure 2.06 share the Phe-boroLeu backbone of bortezomib and have an extension from the P_2 phenylalanine for the purpose of probing the primed binding sites.



2.01 n=2, R=Ph
2.02 n=3, R=Ph
2.03 n=2, R=4-Im
2.04 n=3, R=4-Im

Figure 2.06 Structure of target inhibitors **2.01**-**2.04** designed to probe the promiscuity of the primed site binding channel of the β5 subunit of the proteasome.

Extending a structure from the P_2 phenylalanine residue of bortezomib is likely to occupy the primed site binding channel as the S_2 subsite of the β5 subunit of the proteasome is connected to the primed site binding channel. Furthermore, the 4-carbon of the P_2 phenylalanine of bortezomib is orientated such that it is pointing in the direction of the primed site binding channel. A superimposition of the cocrystal structures of bortezomib and primed site occupying

belactosin A derivative 1 with the yeast 20S proteasome reveal that the 4-carbon of bortezomib's phenylalanine phenyl ring is approximately coincident with a carbon of the *cis*-cyclopropane in belactosin A derivative 1 (Figure 2.07).

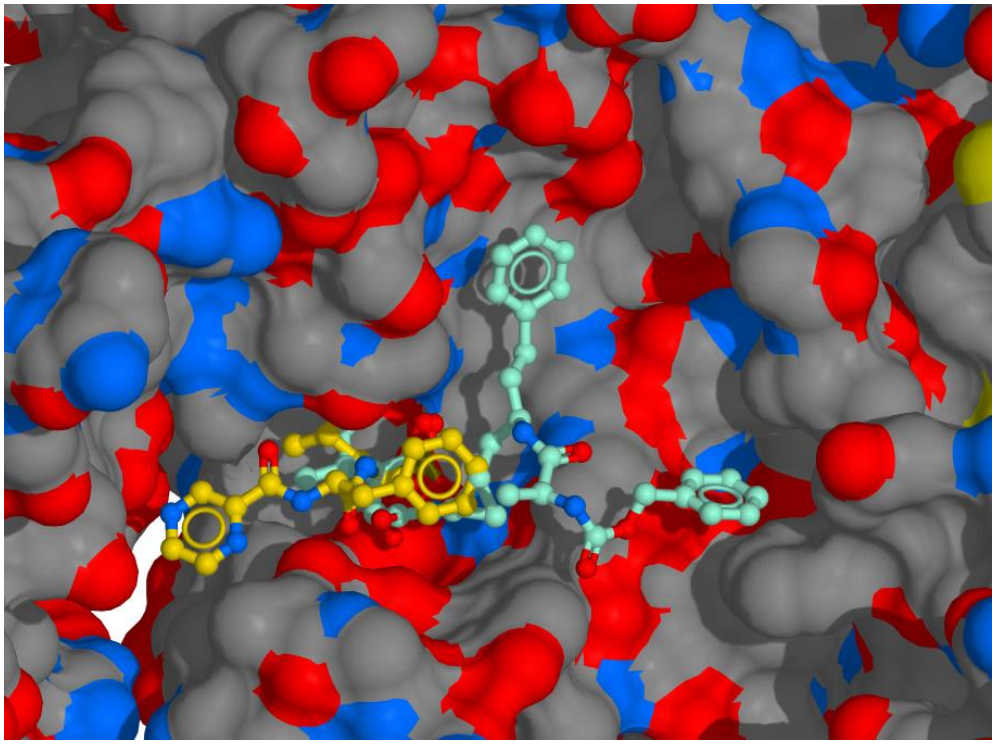


Figure 2.07 Superimposed crystal structures of bortezomib (yellow) (PDB code: 3MG0) and belactosin A derivative 1 (teal) (PDB code: 4J70) complexed by the yeast 20S proteasome. The phenyl ring 4-carbon of bortezomib approximately coincides with a carbon of the cyclopropane in belactosin A derivative 1, indicating the hydrophilic pocket can be accessed via the P_2 phenylalanine residue of bortezomib-based inhibitors.

The N-terminal pyrazinyl carbonyl protecting group of bortezomib was substituted with *tert*-butyloxycarbonyl (Boc) (Figure 2.08) in order to provide a shorter and less expensive synthetic route. Whilst this substitution will influence the binding of inhibitors, it does not exclude comparison between the inhibitors in this thesis and thus does not limit the aim of this thesis: to probe the hydrophilic primed binding pocket of the β5 subunit of the proteasome.

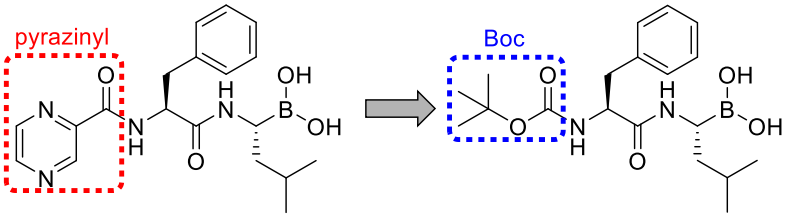


Figure 2.08 Replacement of the N-terminal pyrazinyl protecting group of bortezomib with *tert*-butyloxycarbonyl allows for a shorter and less expensive synthetic route.

Probing of the primed site binding channel of the β5 subunit of the proteasome by a bortezomib-inspired inhibitor requires a structure extending from the 4-carbon of the P_2 phenylalanine residue. The approach towards the design of such an extension was to separate the extension into three components (Figure 2.09). Firstly, the head group at the end of the extension with the purpose of probing the interaction with the primed pocket; secondly, the length of the alkyl chain extension so that the head group can reach the desired pocket; and finally, the linker attaching the extension to the phenyl 4-carbon of the P_2 residue, phenylalanine.

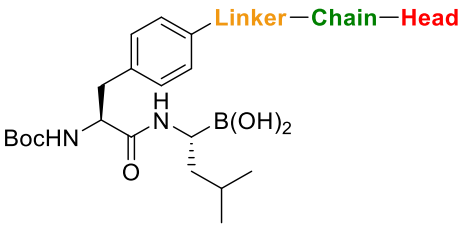


Figure 2.09 Approach towards the design of the extension from the 4-carbon of the P_2 phenylalanine of a bortezomib-inspired proteasome inhibitor. The design of the extension was split into three components: the head group at the end of the extension, the length of the alkyl chain and the linker attaching the extension to the P_2 phenylalanine.

Firstly, phenyl (Figure 2.10 A) and imidazolyl (Figure 2.10 B) substituents were chosen as the head groups as their properties, such as hydrophobicity or hydrophilicity and ability to form hydrogen bonds or salt bridges, differ significantly. The phenyl head group is unable to form hydrogen bonds or salt bridges with the hydrophilic primed pocket. Furthermore, the phenyl group provides a reference point to allow comparison of the compounds in this thesis with

belactosin C derivative 1 and belactosin A derivative as they occupy the primed binding pocket with a phenyl substituent. An imidazolyl group is an isosteric substitution of a phenyl group which has the unique ability to act as a hydrogen bond donor, hydrogen bond acceptor and as a cation in a salt bridge (Figure 2.10 B).[100] The residues in the hydrophilic binding pocket can interact with an imidazolyl group in any of those three conditions. Tyr113 and Ser116 can donate a hydrogen bond via their hydroxyls, Asp115 and Glu117 are anionic and can form salt bridges, whilst all four are able to act as hydrogen bond acceptors.

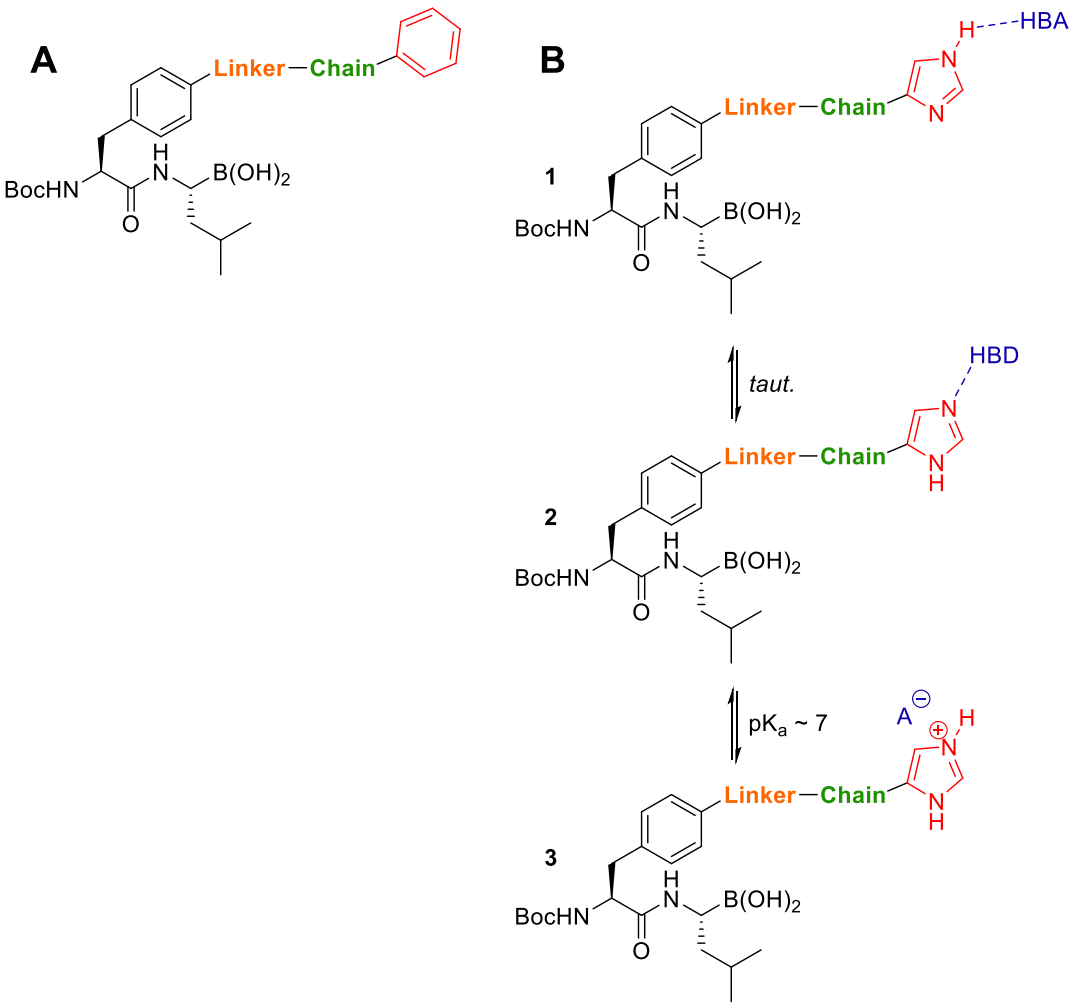


Figure 2.10 General structures of proteasome inhibitors with a (A) phenyl or (B) imidazolyl head groups. Imidazolyl groups have the ability to be a (1) hydrogen bond donor, (2) hydrogen bond acceptor and (3) the cation in a salt bridge.[100]

The group connecting the extension to the P_2 phenylalanine residue should involve well-known chemistry in its synthesis, be stable in any environments the compound will be subjected to for

evaluation and be compatible for binding to the proteasome. In order to be compatible for binding, the linker group must be relatively small and non-reactive. As such, a phenol ether linkage (Figure 2.11 A) is the smallest possible linkage with the added advantage of potential synthetic routes containing the natural amino acid tyrosine. A substitution reaction between the tyrosine phenol and an alkyl halide could afford the desired attachments, which could also be achieved via a Mitsunobu reaction between the phenol and an alcohol. The second linker group considered was an amide derivative of an aniline (Figure 2.11 B), which has the potential to be synthesised using the multitude of known amide coupling reaction conditions. Finally, the reverse configuration of the amide (Figure 2.11 C), such that it is a derivative of a benzoic acid, is also a candidate for the linker group. Synthesis of these compounds could also be achieved via amide coupling chemistry.

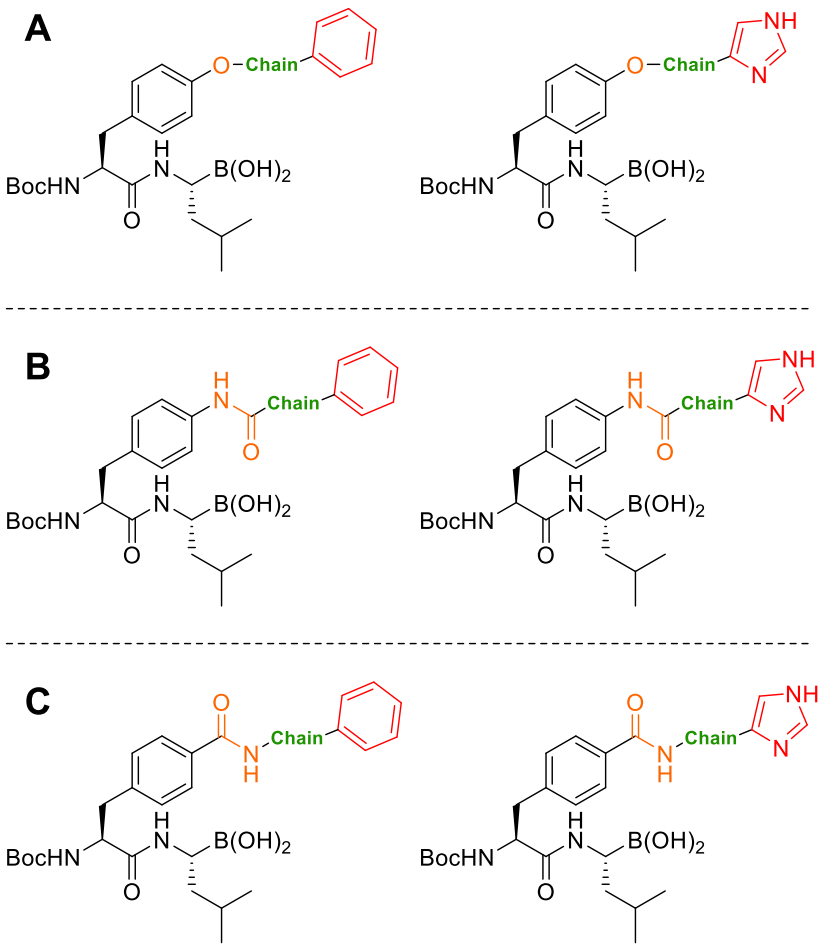


Figure 2.11 General structure of P_2 extended proteasome inhibitors with (A) phenol ether, (B) aniline amide and (C) benzoic acid amide linker.

Occupying one of the primed site binding pockets, the structure connecting the linker and head group must be an appropriate length. The simplicity, flexibility and scalable length of alkyl chains makes them a natural choice to provide length to the extension from the P_2 phenylalanine residue. Visual inspection of the binding mode of bortezomib reveals that the length of extension excluding the head group should be between 3 and 6 atoms from the phenyl 4-carbon of the P_2 phenylalanine residue. The structures of this preliminary set of proposed inhibitors are shown in Figure 2.12 below.

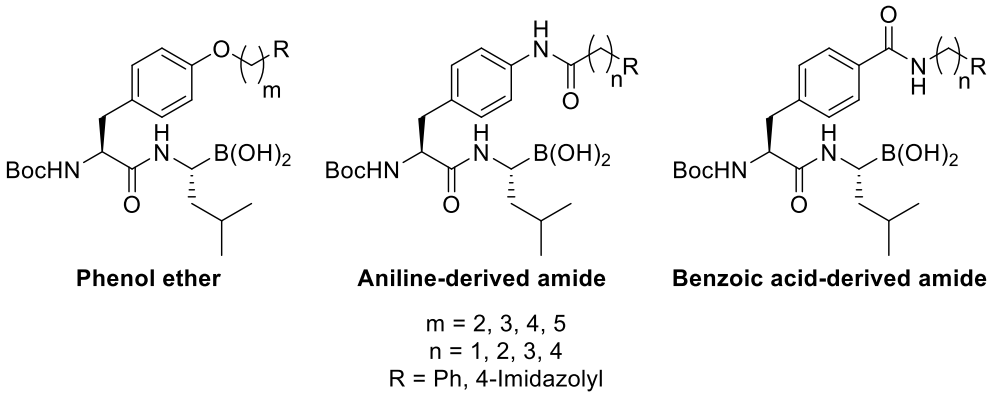


Figure 2.12 Structures of the preliminary set of proposed inhibitors which include all combinations of linker, alkyl chain length and head group.

The covalent docking function of the program 'Molsoft ICM-Pro' was used to dock the potential synthesis targets with the β5 subunit of the proteasome (benzoic acid-derived amides docked with the proteasome can be found in Figure 2.15). A 1-carbon chain length is not sufficient to place the head group of a benzoic acid-derived amide within the small hydrophilic pocket (Figure 2.15 A). On the other hand, an alkyl chain length of 4 is likely too long to be accommodated within the primed site binding channel without steric strain on the extension (Figure 2.15 D). Alkyl chain lengths of 2 or 3 are likely to be the optimum lengths to allow interactions between the head group and the primed binding sites (Figure 2.15 B and C). As such, the target compounds for synthesis are the aniline-derived or benzoic acid-derived amides with an alkyl chain length of 2 or 3 and phenol ethers with alkyl chain length of 3 or 4, due to the linker group being one atom shorter.

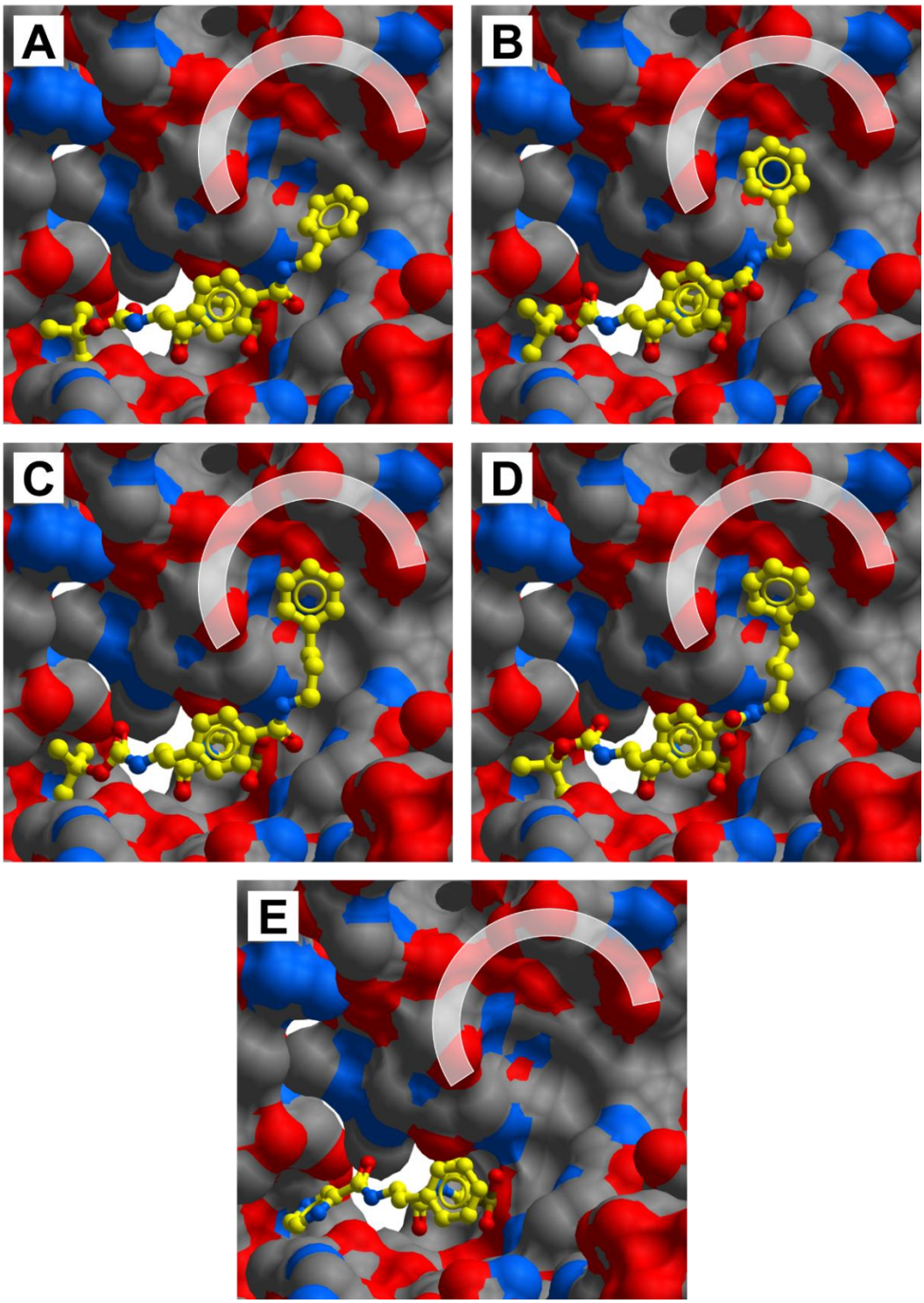


Figure 2.15 Covalent docking structures of the set of phenyl head group benzoic acid-derived amides with the β5 subunit of the proteasome (PDB code: 5LF3) used to qualitatively indicate the optimum extension lengths to interact with the hydrophilic primed binding pocket. (A) 1-carbon alkyl chain; (B) 2-carbon alkyl chain; (C) 3-carbon alkyl chain; (D) 4-carbon alkyl chain; (E) co-crystal structure of bortezomib with the 20S proteasome (PDB code: 5LF3). Produced using the covalent docking function of 'Molsoft ICM-Pro'.

2.3 Synthesis

All target compounds are dipeptides, which means the natural first disconnection is the amide bond. The second disconnection is the group attaching the extension to the phenyl ring of phenylalanine. The disconnection for the alkyl aryl ether is the bond between the alkyl carbon and the oxygen, whereas for the two amide derivatives the disconnection is again the amide bond (Figure 2.16). To avoid chemospecificity problems, the free acid of phenylalanine must be esterified during the formation of the group which attaches the extension to the P_2 phenylalanine.

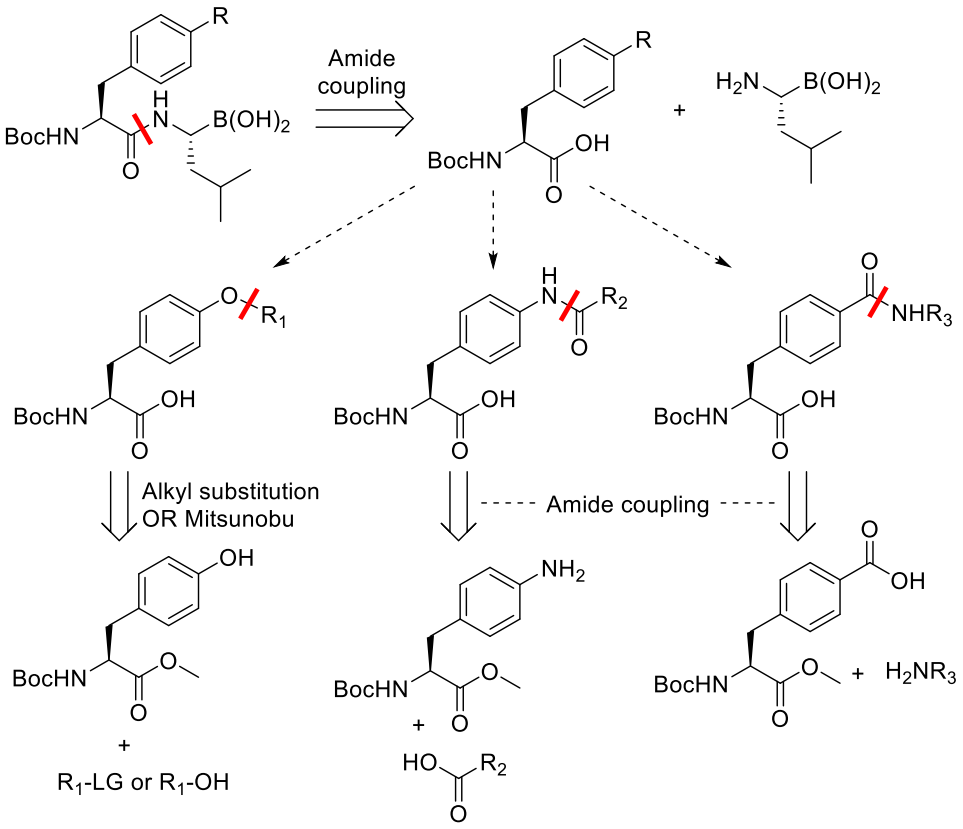
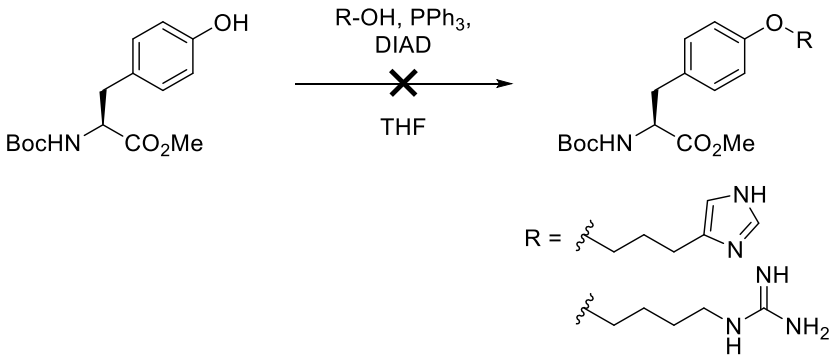


Figure 2.16 Retrosynthetic analysis of preliminary proposed inhibitors.

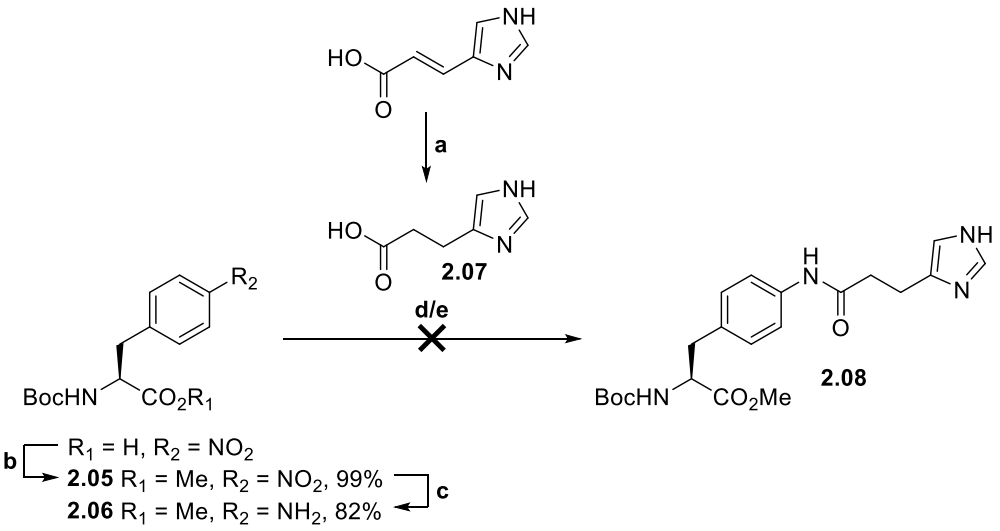
Synthesis of phenol ether compounds with an alkyl chain length of two via alkyl substitution has the problem of elimination of the leaving group to form styrene. As such, the synthesis of the alkyl aryl ether compounds was attempted via Mitsunobu chemistry.[101,102] An attempt at the formation of the phenol ether with the imidazolyl substrate via a Mitsunobu reaction was unsuccessful, indicating the imidazolyl nitrogens may interfere with the reaction (Scheme

2.01). This is consistent with a guanidinium substrate also not forming the desired ether under the same Mitsunobu conditions (Scheme 2.01). Thus, protecting group strategies were considered but were likely incompatible with the other protecting groups on the compound (Boc, boronic ester). Planned synthesis of the phenol ether compounds was therefore abandoned as the set could not be completed in its entirety.



Scheme 2.01 Attempted synthesis of phenol ethers via Mitsunobu chemistry.

As a common intermediate in the synthesis of the aniline-derived amide compounds, *N*-Boc-4-amino-L-phenylalanine methyl ester, **2.06**, was synthesised via methyl esterification of *N*-Boc-4-nitro-L-phenylalanine with methyl iodide to yield **2.05**. Subsequent reduction of the nitro group afforded the aniline **2.06** (Scheme 2.02). Amide coupling utilising HATU and PyBOP as the coupling reagents for the formation of the amide linker to the aniline did not yield any product. The low reactivity of anilines is the likely reason for the failure of the reactions. As such, synthesis of the aniline-derived amide compounds was abandoned in favour of the amides derived from benzoic acid.

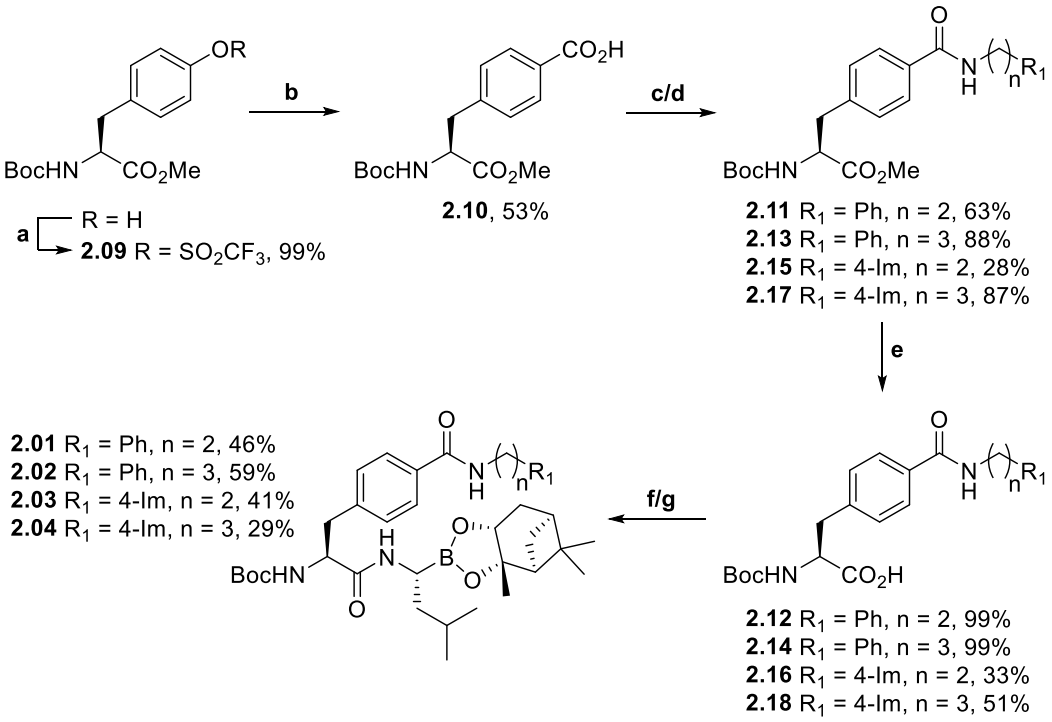


Scheme 2.02 Synthesis of *N*-boc-4-amino-L-phenylalanine methyl ester **2.06** and attempted synthesis of aniline-derived amides. Reagents and conditions: a) NH_4HCO_2, Pd/C, MeOH, reflux, 6 h; b) MeI, $NaHCO_3$, DMF, r.t., 18 h; c) zinc dust, NH_4Cl, MeOH, H_2O, 0 °C to r.t., 2.5 h; d) HATU, DIPEA, DMF, r.t., o/n; e) PyBOP, DIPEA, DMF, r.t., o/n.

Synthesis of boronic esters **2.01**-**2.04** is outlined in Scheme 2.03. The benzoic acid amide derivatives have a common intermediate, namely *N*-Boc-4-carboxy-L-phenylalanine methyl ester (Compound **2.10**). This was synthesised using conditions from Horatscheck *et al.*,[103] where *N*-Boc-L-tyrosine methyl ester is converted to its corresponding *O*-triflate, **2.09**, via reaction with *N*-phenyl-bis(trifluoromethane sulfonimide) in the presence of triethylamine. Subsequent hydroxycarbonylation via reaction with carbon monoxide catalysed by palladium(II) acetate with 1,3-bis(diphenylphosphino)propane (DPPP) yielded *N*-Boc-4-carboxyphenylalanine methyl ester **2.10**.

HATU mediated coupling of **2.10** with phenethylamine or 3-phenylpropylamine gives compounds **2.11** and **2.13**, respectively. Using the same conditions, **2.10** gave no desired product in the reaction with histamine. However, switching to the less reactive HBTU allowed the coupling of histamine with **2.10** to give the desired amide **2.15** and the same reaction with homohistamine afforded amide **2.17**. Following methyl ester hydrolysis of **2.11** and **2.13** by lithium hydroxide in 1:1 THF/H_2O, HATU mediated coupling with (*R*)-boroleucine pinanediol

gave the boronic esters **2.01** and **2.02**, respectively. Similarly, hydrolysis of **2.15** and **2.17** followed by HBTU mediated coupling with (*R*)-boroleucine pinanediol afforded the respective amides **2.03** and **2.04**.

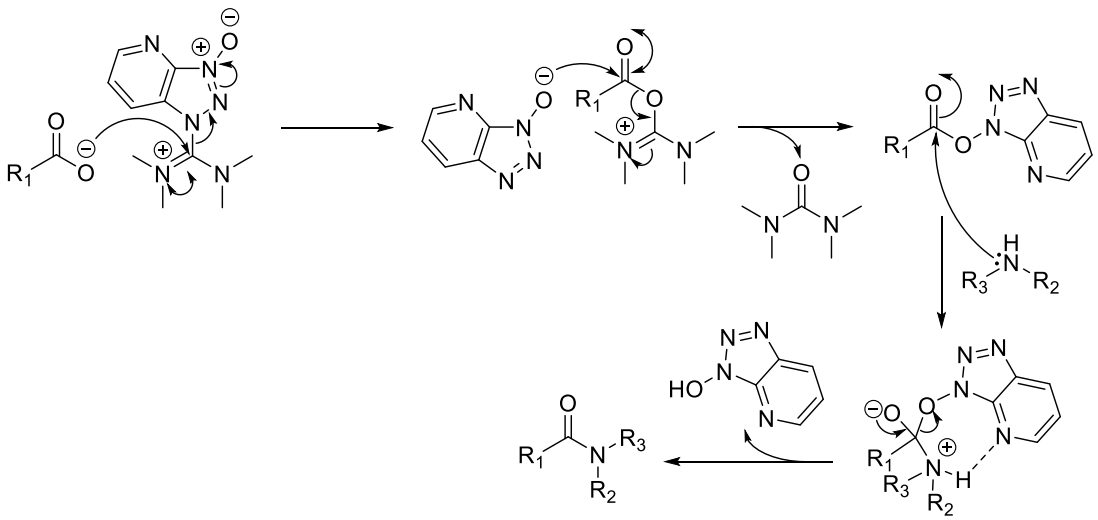


Scheme 2.03 Synthesis of proteasome inhibitors **2.01**-**2.04**. Reagents and conditions: a) *N*-phenyl-bis(trifluoromethane sulfonimide), Et_3N, DCM 0 °C, 1.5 h; b) CO, $Pd(OAc)_2$, DPPP, DIPEA, 3:1 DMF/H_2O, 70 °C to r.t., o/n; c) $R_1(CH_2)_nNH_2$, HATU, DIPEA, DMF, r.t., o/n; d) $R(CH_2)_nNH_2$, HBTU, DIPEA, DMF, r.t., o/n; e) LiOH, 1:1 H_2O/THF, 0 °C → r.t., 3 h; f) boroLeu pinanediol boronic ester trifluoroacetate, HATU, DIPEA, DMF, r.t., o/n; g) boroLeu pinanediol boronic ester trifluoroacetate, HATU, DIPEA, DMF, r.t., o/n.

The yields of the coupling reactions varied wildly (28-88%) between substrates and coupling reagents. In all cases, coupling of phenyl-containing substrates was higher yielding than coupling of their imidazolyl containing analogues. This observed trend is possibly due to a combination of the differences between coupling reagents (HATU vs HBTU) and the presence of the imidazole, which may interfere in coupling reactions. Similarly, methyl ester hydrolysis of compounds containing imidazolyl rings was lower yielding than for compounds containing

phenyl rings. However, this is likely due to a problem with extraction of the product from the aqueous layer as acidification to pH 1 will protonate the imidazolyl rings, thus making it positively charged and have greater hydrophilicity.

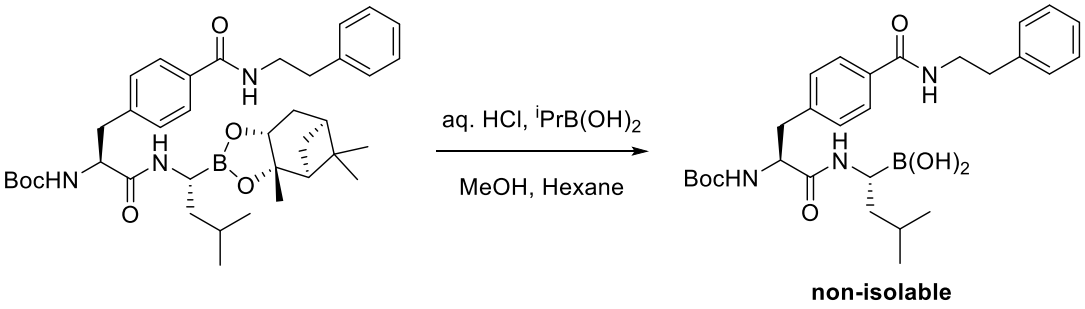
The greater reactivity of HATU relative to HBTU may result in the imidazole N-H interfering with a step in the activation of the carboxylic acid or the subsequent formation of the amide. No explanation was found after an extensive search in the literature, however the HATU mechanism (Scheme 2.04) is suggested[104] to involve a 7-membered ring transition state where the amine nucleophile is stabilised via hydrogen bonding with the pyridinyl nitrogen. The imidazole N-H may prevent this favourable transition state via hydrogen bonding to the pyridinyl nitrogen of HATU. A search of the Reaxys® database for the amide coupling of histamine with any carboxylic acid (R-COOH) gave 141 results. Out of all coupling reagents, HBTU was the most common choice followed by EDC and various BOP derivatives. HATU was the choice of coupling reagent for only 5 reactions.



Scheme 2.04 Mechanism of HATU mediated amide coupling.

Conversion of boronic ester **2.01** to the corresponding boronic acid **2.19** was attempted with a hexane/methanol biphasic reaction condition with *iso*-butyl boronic acid and 1M aqueous HCl (Scheme 2.05). The resulting product was too polar to move from the baseline of a normal phase

thin layer chromatography (TLC) plate and difficult to separate on reverse phase silica. Thus, all final compounds are the pinanediol boronic esters, **2.01**-**2.04**. The boronic ester cleaves in aqueous environments, thus not compromising the ability of the boronate to function as the covalent warhead.[105,106]



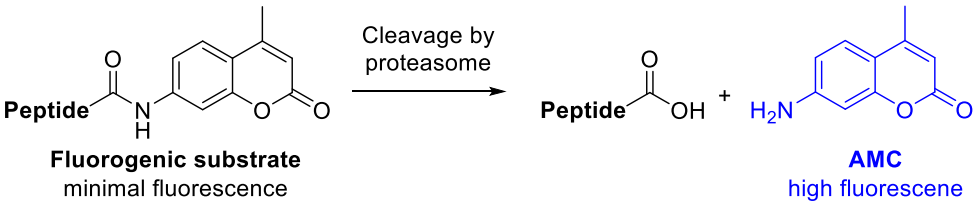
Scheme 2.05 Attempted hydrolysis of pinanediol boronic acid **2.01** to boronic acid **2.19**.

2.4 Proteasome Inhibition Assay

The aim of this project, being to probe the primed binding pocket of the β5 subunit of the proteasome, dictates that the evaluation of the inhibitors should be via an enzyme-based assay. Cell-based assays introduce variables such as cell-permeability, metabolism and off-target effects which would obscure the measurement of affinity towards the enzyme. Thus, the inhibitors described in this chapter were evaluated with an enzyme-based assay.

Due to the inherent high level of sensitivity, a fluorometric assay was chosen to evaluate the inhibitors. Fluorometric assays have been used to quantify proteasome activity since the discovery of the proteasome in 1983 by Wilk and Orlowski.[43] These assays utilise fluorogenic peptide substrates which are degraded by only one of the proteasome subunits and are thus a convenient and sensitive tool for the measurement of proteasome activity, where the fluorescence measured is proportional to the activity of a specific subunit of the proteasome. The fluorogenic substrates are generally three to four amino acid residues in length with a reporting group attached at the C-terminus. Upon cleavage by a proteasomal subunit, a highly fluorescent reporting group is released (Scheme 2.06). The most common fluorophore used as

a reporter group for measurement of proteasome activity is 7-amino-4-methylcoumarin (AMC, λ_{ex} = 380 nm, λ_{em} = 460 nm), which is attached to the C-terminus of a peptide with an amide bond via its amine (Figure 2.07). The most commonly used fluorogenic substrates for the β1, β2 and β5 subunits of the proteasome are Ac-Nle-Pro-Nle-Asp-amc, Bz-Val-Gly-Arg-amc and Suc-Lys-Lys-Val-Tyr-amc, respectively.[107]



Scheme 2.06 Cleavage of a fluorogenic peptide substrate which releases 7-amino-4-methylcoumarin (AMC, blue), a highly fluorescent reporter group for the measurement of proteasome activity.

Compounds **2.01-2.04** were evaluated alongside bortezomib in an assay against all three subunits of the proteasome. As a triage for human proteasome, 20S rabbit proteasome was used as it is considerably cheaper whilst possessing a very high level of homology to the human 20S proteasome. The assay was performed in triplicate against 7 serial dilutions of each inhibitor (25 pM to 25 μM) and were activated by addition of the enzyme and the fluorescence of AMC was measured after incubation at 37 °C for 2 h. Three controls were used for the assay: 1) a buffer blank; 2) a substrate blank in buffer; and 3) a positive control with enzyme and substrate with no inhibitor for the purpose of normalising the data with the positive control at 100% activity. The IC_{50} value was calculated GraphPad Prism 8 using a four-parameter fit of the fraction of activity relative to the positive control against the log of inhibitor concentration (Figure 2.17). An IC_{50} value is the concentration of inhibitor required to reduce the enzyme activity by 50%.

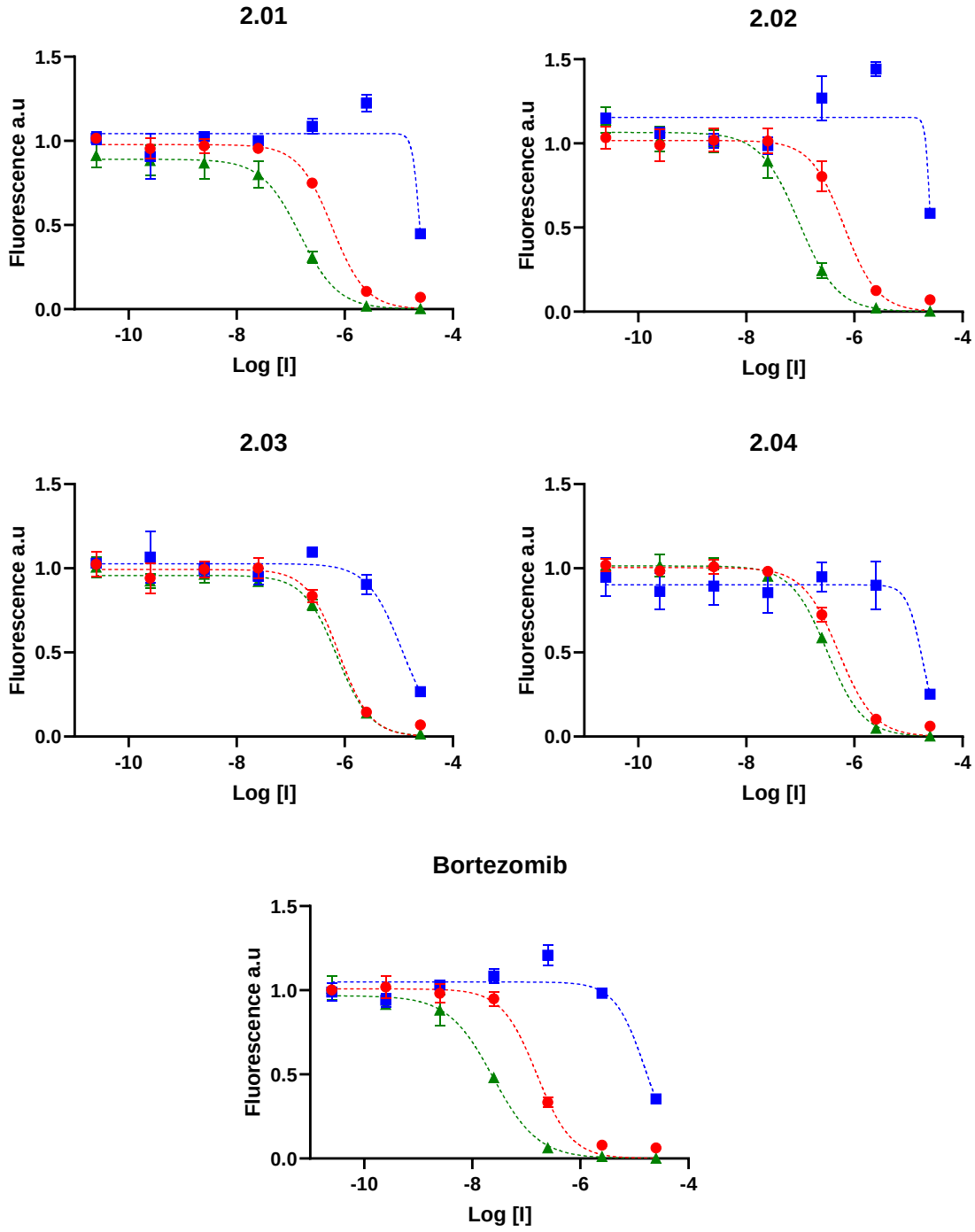


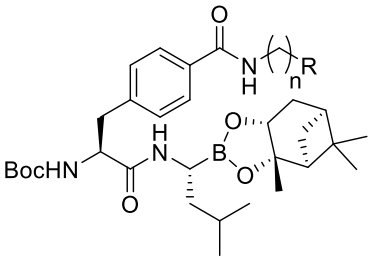
Figure 2.17 Inhibition curves of compounds **2.01**-**2.04** and bortezomib with a four-parameter inhibitor vs. response fit against β1 (red), β2 (blue), and β5 (green) subunits of the rabbit 20S proteasome.

The calculated IC_{50} values are presented in Table 2.01. Compounds **2.01**-**2.04** all inhibit the β5 subunit of the proteasome with an IC_{50} in the nanomolar range. However, all compounds are less potent that bortezomib which has nearly 4-fold greater potency against the β5 subunit that the next most potent inhibitor, **2.02**. Compounds with a phenyl head group, **2.01** and **2.02**, have

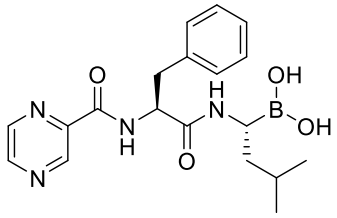
greater potency than those with an imidazolyl head group, **2.03** and **2.04**. Additionally, the three carbon alkyl chain analogues, **2.02** and **2.04**, are more potent than their two carbon alkyl chain counterparts, **2.01** and **2.03**, respectively. All compounds also have nanomolar IC_{50} values against the β1 subunit of the proteasome, however the trends found in IC_{50} values against the β5 subunit do not appear against the β1 subunit. Neither head group nor alkyl chain length has definitively greater potency against the β1 subunit, suggesting a drastically different binding pocket is present in the β1 subunit compared to the β5 subunit. All five compounds only begin to inhibit the β2 subunit at the highest measured concentration (25 μM).

Bortezomib is approximately 6.5-fold more potent against the β5 subunit over the β1 subunit, a specificity only surpassed in this assay by compound **2.02**, which is approximately 6.8-fold more potent against the β5 subunit. Compound **2.01** is the only other inhibitor with reasonable specificity for inhibition of the β5 subunit of the β1 subunit, with a specificity of approximately 4-fold. Both inhibitors with an imidazolyl head group, **2.03** and **2.04**, have relatively small differences in IC_{50} values between the β5 and β1 subunits.

Table 2.01 IC_{50} values with standard errors of compounds **2.01**-**2.04** and bortezomib against the β1, β2 and β5 subunits of the rabbit 20S proteasome.



2.01 n=2, R=Ph
2.02 n=3, R=Ph
2.03 n=2, R=4-Im
2.04 n=3, R=4-Im



Bortezomib

Compound	IC_{50} β1 ± SE (nM)	IC_{50} β2 (nM)	IC_{50} β5 ± SE (nM)
2.01	583 ± 55	>10000	149 ± 22
2.02	648 ± 93	>10000	95 ± 14
2.03	782 ± 106	>10000	720 ± 71
2.04	512 ± 39	>10000	312 ± 17
Bortezomib	155 ± 17	>10000	24 ± 3

The diminished potency of imidazolyl head group inhibitors, **2.03** and **2.04**, relative to inhibitors with phenyl head groups, **2.01** and **2.02**, indicates there are no beneficial interactions of these groups with the hydrophilic primed pocket. Whilst these results are informative regarding the promiscuity of the primed site binding channel, the binding mode of the inhibitors remains uncertain. However, the results do not exclude the possibility of the imidazolyl group from occupying the pocket, as the residues and imidazolyl group may occupy the pocket but also not be in the correct orientation to form beneficial secondary interactions. Alternatively, the reason for the greater potency of inhibitors **2.01** and **2.02** compared to **2.03** and **2.04** could be that the extension into the primed site binding channel may occupy the hydrophobic pocket. The flexibility of the extension into the primed site binding channel prevents extrapolation of the binding mode without further conclusive studies such as X-ray crystallography or cryogenic electron microscopy.

Extending into the primed site binding channel of the β5 subunit of the proteasome via an attachment to the P_2 phenylalanine residue of a bortezomib-inspired inhibitor can retain potency with IC_{50} values in the nanomolar range. Occupation of a primed site binding pocket with an imidazolyl substituent whilst retaining respectable potency indicates the primed site binding channel is not only promiscuous with accommodating hydrophobic substituents such as phenyl and naphthyl, but also to hydrophilic substituents. Such a result has implications in the design of future proteasome inhibitors, as the primed site binding channel is accommodating to substituents which may not be accommodated by other proteases inhibited by proteasome inhibitors, thereby increasing specificity towards the proteasome. Evaluation of inhibitors **2.01**-**2.04** against α-chymotrypsin, an off-target of bortezomib, is detailed in chapter 4.

2.5 Chapter Conclusions

In summary, the 4-carbon of the P_2 phenylalanine of bortezomib was identified to provide a point of attachment for an extension into the primed site binding channel of the β5 subunit of the proteasome. As such, four inhibitors, **2.01**-**2.04**, were designed for the purpose of probing the promiscuity of the primed site binding channel, which thus far has only been occupied by hydrophobic substituents such as phenyl and naphthyl. The four inhibitors were evaluated alongside bortezomib in an enzyme assay against all three subunits of the rabbit 20S proteasome. IC_{50} values of the four inhibitors against the β5 subunit were in the nanomolar range, indicating the primed site binding channel accommodates hydrophilic substituents such as imidazolyl as well as hydrophobic substituents. The demonstrated promiscuity of the primed site binding channel suggests occupation of the binding channel could be exploited in the design of proteasome inhibitors which aim to increase specificity towards the proteasome. The binding mode of the extension into the primed site binding channel could not be extrapolated from the assay results due to their flexibility. As such, reducing the flexibility of the extension could reduce the possible binding conformations, resulting in probing of the primed site binding channel with greater predictability and accuracy.

Chapter 3: Photoswitchable Proteasome Inhibitor

3.1 Introduction

The work detailed in chapter 2 demonstrated that the primed site binding channel of the β5 subunit of the proteasome can accommodate not only hydrophobic substituents such as phenyl and naphthyl, but also less hydrophobic substituents with hydrogen bonding capabilities such as imidazole. The promiscuity of the primed site binding channel resulted in the primed site occupying P_2 extended inhibitors **2.01**-**2.04** to have IC_{50} values in the nanomolar range against the β5 subunit of the proteasome. However, due to the large size of the primed site binding channel, the binding mode remains undetermined. As such, the aim of this chapter is to probe the promiscuity of the primed site binding channel with greater accuracy and predictability of binding conformation. Compound **3.01** is an azobenzene-containing inhibitor of the proteasome which has less degrees of flexibility compared to the inhibitors in Chapter 2. The inclusion of an azobenzene moiety into the structure of compound **3.01** allowed conformational control of the substituent occupying the primed site binding channel, thus allowing further probing of the promiscuity of the primed site binding channel.

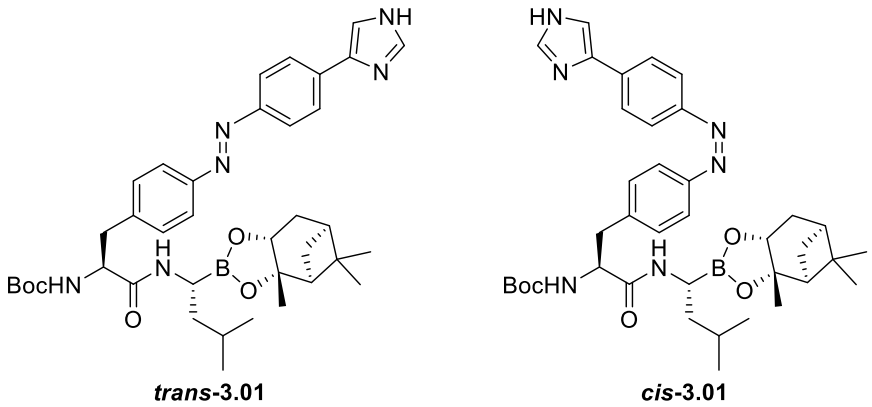


Figure 3.01 The *trans* and *cis* isomers of compound **3.01**, an azobenzene-containing inhibitor of the proteasome which, upon irradiation, can photoswitch between the two geometrically distinct isomers.

3.1.1 Azobenzene as a Molecular Photoswitch

Azobenzene was first identified by the German chemist Eilhard Mitscherlich in 1843,[108] however, the photoisomerisation capabilities of azobenzene were not described until over a century later by Hartley.[109] Azobenzenes contain a nitrogen-nitrogen double bond of which rotation is forbidden in its electronic ground state. This results in the phenyl substituents on each nitrogen to be arranged either on the same side (*Z*/*cis* isomer), or on the opposite side (*E*/*trans* isomer) of the plane of the N=N bond. The *trans* isomer is approximately 50 kJ.mol^{-1} more stable than the *cis* isomer.[110] Thus, upon irradiation with ultraviolet (UV) wavelengths between 320-350 nm, *trans*-azobenzene isomerises to *cis*-azobenzene. This isomerisation is reversible via thermal energy or irradiation with wavelengths between 400-450 nm (Figure 1.11).[111] The significant difference in geometry between isomers also gives rise to a large change in polarity, allowing azobenzenes to be used as structural molecular switches operated by light. The ability to photoisomerise between *cis* and *trans* isomers, as well as azobenzene's other chemical and spectroscopic properties, means azobenzenes are used in molecular machines,[112] light-driven liquid motion,[113] dyes,[114] metal ion chelators[115] and photoswitchable inhibitors in photopharmacology.[116,117]

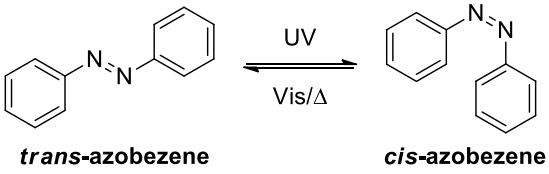


Figure 1.11 Reversible *cis*/*trans* isomerisation of azobenzene.

3.1.3 Photoswitchable Inhibitors of the Proteasome

There are two published instances of photoswitchable proteasome inhibitors. Hansen *et al.*[118] created six proteasome inhibitors inspired by bortezomib with an N-terminal azobenzene with a series of substitutions on the azobenzene (Figure 1.13 A). As such, these inhibitors do not occupy the primed site binding channel of the β5 subunit of the proteasome. The Abell group[119] designed two sets of photoswitchable azobenzene-containing proteasome inhibitors, one based

on delanzomib, the other based on bortezomib. The delanzomib-based inhibitor series (Figure 1.13 B) contained an N-terminal azobenzene moiety and does not occupy the primed site binding channel. However, the bortezomib-based inhibitor series replaced the phenyl at the P_2 position with an azobenzene moiety (Figure 1.13 C). One such inhibitor was a biphenyl azobenzene derivative, which places significant steric bulk within the primed site binding channel. A 5-fold difference in potency was observed between the two isomeric states for the biphenyl azobenzene derivative in an enzyme assay against the β5 subunit of the proteasome, with the IC_{50} of both states in the low nanomolar range. This indicates there is a strong preference for the location of a bulky hydrophobic substituent within the primed site binding channel.

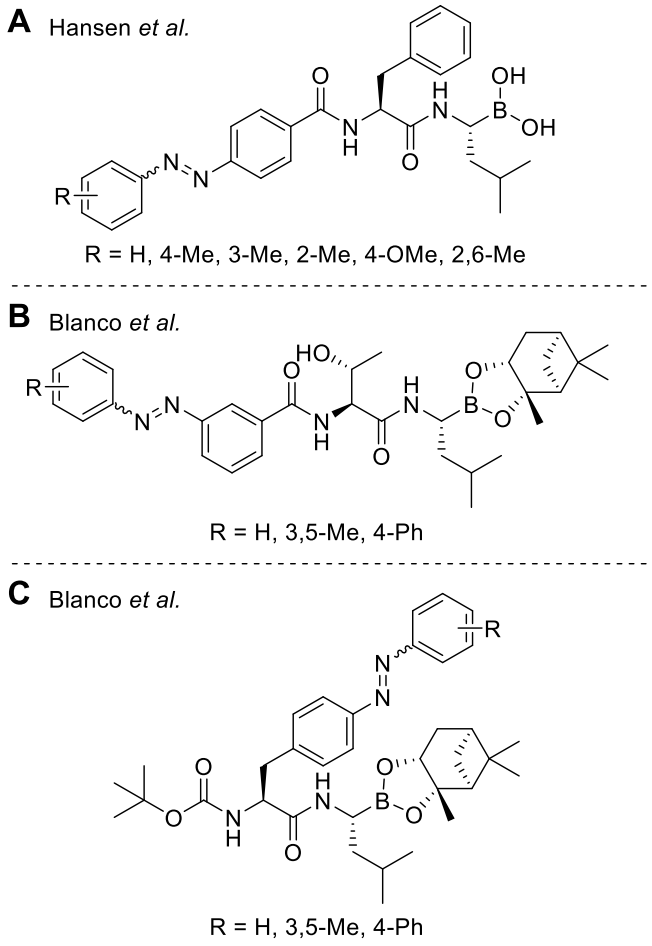


Figure 1.13 Photoswitchable inhibitors from (A) Hansen *et al.*,[118] inspired by bortezomib; (B) Blanco *et al.*,[119] inspired by delanzomib and (C) Blanco *et al.*,[119] inspired by bortezomib.

This work explored the primed site binding channel of the β5 subunit of the proteasome with a photoswitchable moiety incapable of forming secondary interactions such as hydrogen bonding and salt bridges. Hence, this chapter details the probing of the primed site binding channel with a photoswitchable substituent with the ability to form hydrogen bonds or salt bridges. As a result, the extent of the promiscuity presented by the primed site binding channel is investigated.

3.2 Photoswitchable Inhibitor Design

The purpose of the inhibitor described in this chapter was to probe the promiscuity of the primed site binding channel with an inhibitor with greater conformational control compared to the inhibitors described in Chapter 2. As such, the well-defined and distinct geometry of the two isomers of azobenzene allows control of inhibitor conformation via incorporating the azobenzene into the substituent occupying the primed site binding channel. Inspired by the photoswitchable inhibitors of the proteasome by Blanco *et al.*,[119] the azobenzene was placed at the P_2 position of a bortezomib-inspired inhibitor. An imidazolyl group was added to the end of the azobenzene moiety for the purpose of probing the pocket for secondary interactions such as hydrogen bonding or salt bridges. The resulting inhibitor, compound **3.01**, is essentially a combination of inhibitors **2.03** and **2.04** with the bortezomib-inspired photoswitchable inhibitors from Blanco *et al*. (Figure 3.02).[119]

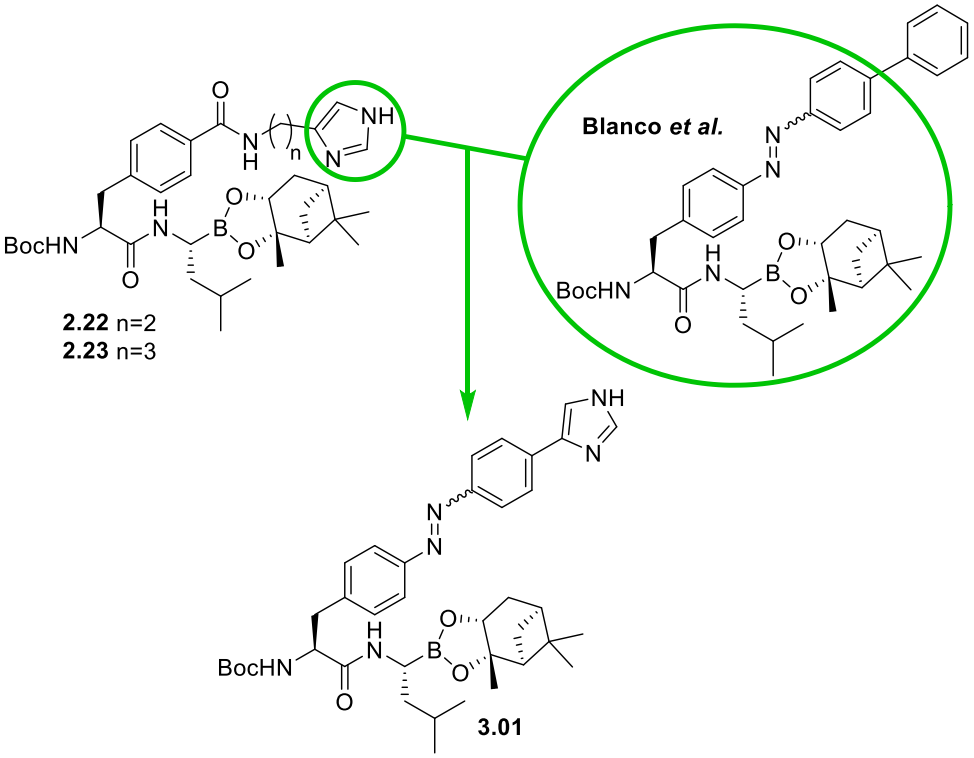


Figure 3.02 Compound **3.01** is a combination of the structures of inhibitors **2.03** and **2.04** with the bortezomib inspired photoswitchable inhibitors from Blanco *et al*.[119]

The large geometric change resulting from the isomerism of an azobenzene, means that the imidazolyl substituent of **3.01** will be situated in drastically different positions of the primed site binding channel. Molecular docking of the *cis* and *trans* isomers of **3.01** using the covalent docking function of the program 'Molsoft ICM-Pro' reveals possible binding conformations for each isomer, which can be found in Figure 3.03. The azobenzene moiety with an attached imidazolyl substituent is expected to be fully conjugated and planar in ***trans*-3.01**. As the entrance to the primed pocket from the S2 subsite is small, ***trans*-3.01** is restricted to a binding conformation which does not place the imidazolyl substituent within either of the pockets identified in Chapter 1. Conversely, the conformation of ***cis*-3.01** is non-planar due to sterics and allows the imidazolyl substituent to be placed in the hydrophilic pocket without significant strain on the molecule.

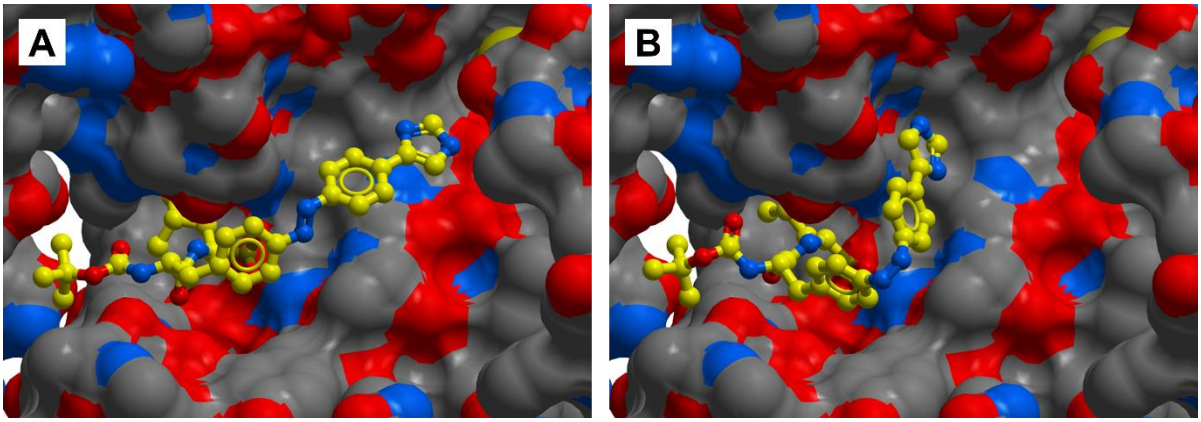


Figure 3.03 Covalent docking structures of A) ***trans*-3.01** and B) ***cis*-3.01** with the β5 subunit of the proteasome illustrating the possible binding modes of each isomer. Produced using the covalent docking function of 'Molsoft ICM-Pro'.

3.3 Synthesis

Similar to Chapter 2, the possible disconnections identified in a retrosynthetic analysis of **3.01** are the amide bond between the P_1 and P_2 residues and the azo bond of the azobenzene moiety (Figure 3.04). The same synthetic strategy was employed for the synthesis of the bortezomib-inspired photoswitchable proteasome inhibitors in Blanco *et al.*[119]

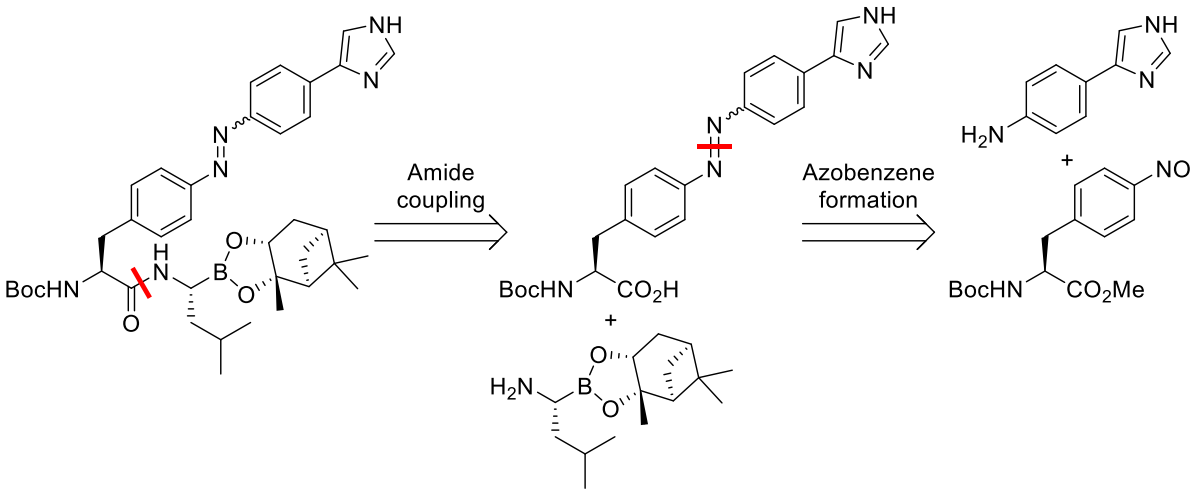
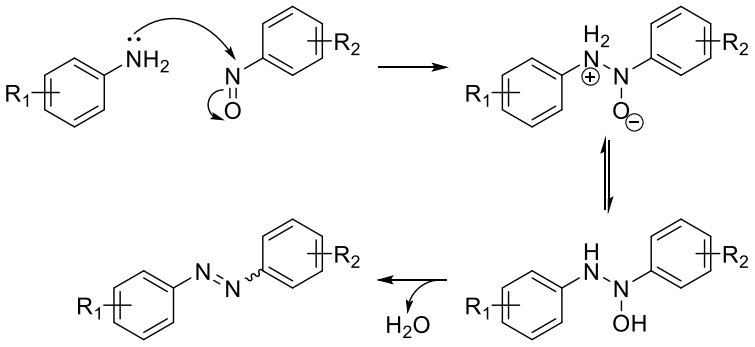


Figure 3.04 Retrosynthetic analysis of compound **3.01** displaying the disconnection of the amide bond between the P_1 and P_2 residues, as well as the disconnection of the azo bond of the azobenzene.

The Mills reaction is one of the many synthetic methods available to form an azobenzene and was the method of choice for the azobenzene formation in the synthetic route towards **3.01**. The

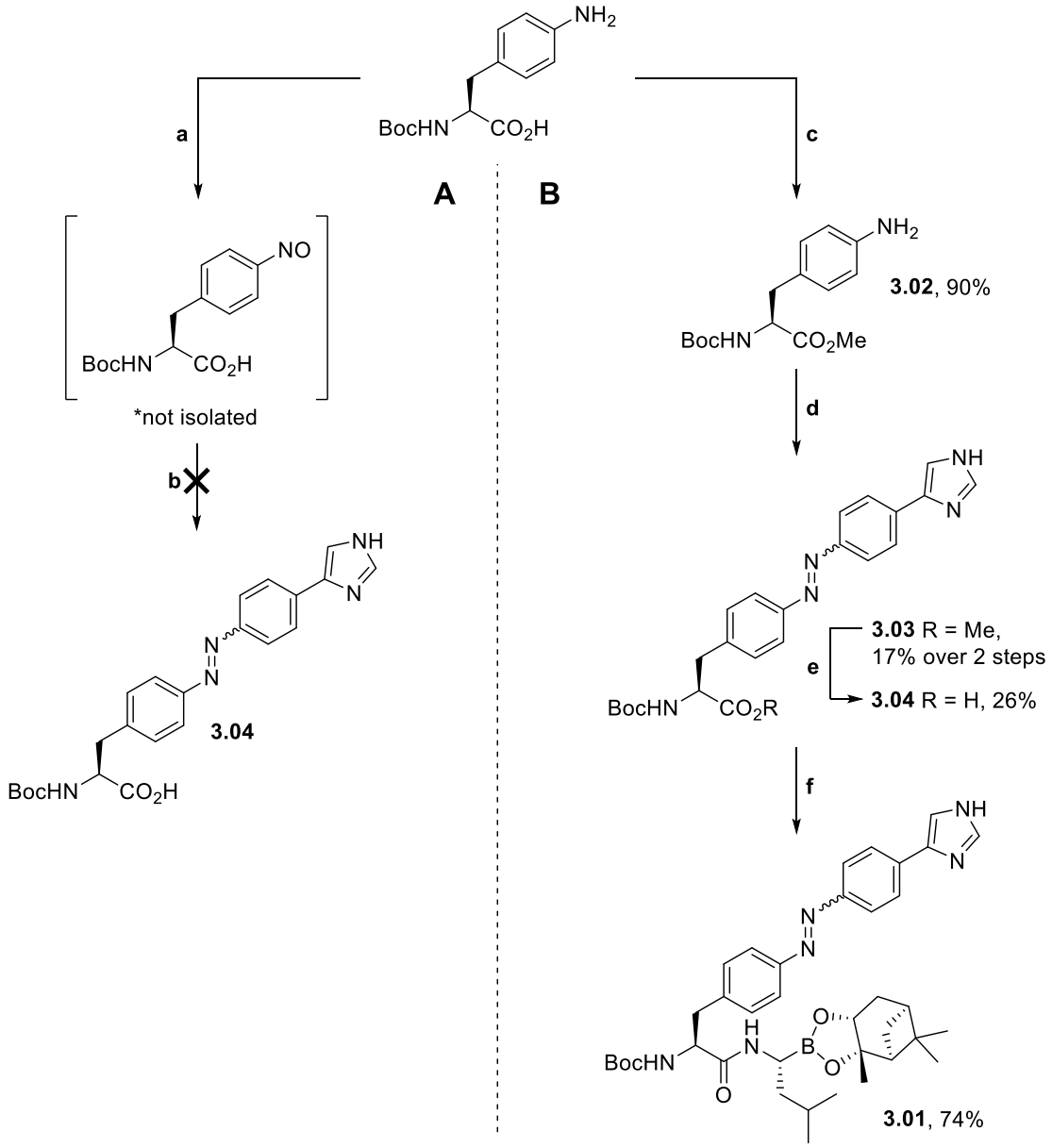
reaction of aryl nitroso derivatives with aniline derivatives in glacial acetic acid affords azobenzenes in good yield. The mechanism of the Mills reaction is analogous to the formation of an imine formation reaction, except the nitroso acts as the electrophile instead of a carbonyl (Scheme 3.01).[120] Aryl nitroso compounds can be obtained via the reaction of the corresponding aniline with Oxone® ($2KHSO_5 \cdot KHSO_4 \cdot K_2SO_4$) in a biphasic solvent system such as H_2O/DCM.



Scheme 3.01 Mechanism of the Mills reaction. Nucleophilic attack of an aniline on the nitrogen of an aryl nitroso in acidic conditions results in a hydroxyhydrazine which subsequently undergoes dehydration to form the azo bond.

An attempt at the synthesis of photoswitchable proteasome inhibitor **3.01** is outlined in Scheme 3.02. Oxidation of *N*-Boc-4-amino-L-phenylalanine to the corresponding nitroso by Oxone® in 1:1 H_2O/DCM was successful. The crude mixture obtained from evaporation of the organic layer contained approximately 85% of the nitroso compound and the pure compound was not isolated due to the reactivity and potential degradation of the nitroso during purification. As such, the crude nitroso was reacted overnight with 4-(1*H*-imidazol-4-yl)aniline in glacial acetic acid in an attempt to form the azobenzene. After work up, less than 1% (as determined by analytical HPLC) of the crude mixture was the penultimate azobenzene **3.04**. In order to prevent possible interference by the free acid, *N*-Boc-4-amino-L-phenylalanine was esterified via reaction with thionyl chloride in methanol to give the corresponding methyl ester **3.02** in 97% yield. Conversion to the nitroso and subsequent reaction with 4-(1*H*-imidazol-4-yl)aniline via

the same reaction conditions as the previous attempt afforded the azobenzene **3.03** in a 17% yield over two steps. Hydrolysis of the methyl ester by LiOH in 1:1 H_2O/THF followed by HBTU mediated coupling of (R)-boroleucine pinanediol gave the target compound **3.01**. As per the inhibitors synthesised in Chapter 2, the final compound was left as the boronic ester and not hydrolysed to the boronic acid due to the likely purification problems.



Scheme 3.02 Synthetic routes for the (A) attempted synthesis of compound **3.04** and (B) photoswitchable inhibitor **3.01**. a) Oxone®, 1:1 H_2O/DCM, r.t., 3 h; b) 4-(1H-imidazol-4-yl)-aniline, glacial acetic acid, r.t., o/n; c) $SOCl_2$, MeOH, -10 °C → r.t., 3 h; d) 1. Oxone®, 1:1 H_2O/DCM, r.t., 3 h; 2. 4-(1*H*-imidazol-4-yl)-aniline, glacial acetic acid, r.t., o/n; e) LiOH, 1:1 H_2O/THF, 4 °C → r.t., 3 h; f) boroLeu pinanediol boronic ester trifluoroacetate, HBTU, DIPEA, DMF, r.t., o/n.

3.4 Photoswitching

The *cis* and *trans* isomers of azobenzene derivatives in solution are in an equilibrium with the equilibrium position dictated by the light irradiating the solution as well as the temperature of the solution. Generally, complete switching to 100% *cis* or *trans* isomer is not observed in environments which enzyme inhibitors are subjected to. As such, the ratio of the *cis* and *trans* isomers in the thermally adapted (*trans*-enriched) and photostationary (*cis*-enriched) states needs to be determined.

Initially, the absorption maximum for ***trans*-3.01** in DMSO was determined to be 382 nm via UV/Vis absorption spectroscopy (Figure 3.05). The wavelength of the absorption maximum for ***trans*-3.01** is longer than that of azobenzene ($\lambda_{max} \approx 320$ nm)[121] due to the electron-rich imidazolyl substituent which lowers the energy of the $\pi \rightarrow \pi^*$ electronic transition. The wavelength used in all instances to irradiate a sample of **3.01** to give the photostationary state was 365 nm. Whilst the λ_{max} of ***trans*-3.01** is nearly 20 nm longer, absorption of 365 nm UV irradiation is 89% of the absorption at λ_{max}, thus making 365 nm a suitable wavelength for the photoswitching of **3.01**. In the *cis*-enriched photostationary state, a small shoulder peak was observed at 460 nm, corresponding to the $n \rightarrow \pi^*$ electronic transition (Figure 3.05).

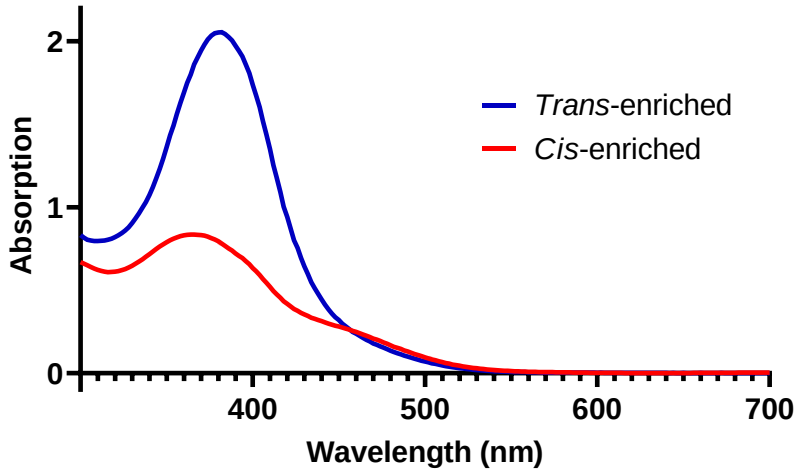


Figure 3.05 UV/vis absorption spectrum of the thermally adapted *trans*-enriched (blue) and photostationary *cis*-enriched (red) states of **3.01**.

^{1}H Nuclear Magnetic Resonance (NMR) spectroscopy was utilised to provide a good estimate of the *cis* to *trans* isomer ratio via integrating the proton signals from the azobenzene (Figure 3.06, Table 3.01). As the compound was irradiated in 100% DMSO to reach the photostationary state for the assay, d_6-DMSO was used as the solvent for the switching and ^{1}H NMR experiments. The half-life of ***cis*-3.01** was determined by running ^{1}H NMR experiments every 10 minutes for 120 minutes and integrating the azobenzene signals to give a ratio of the *cis* and *trans* isomers (Table 3.01).

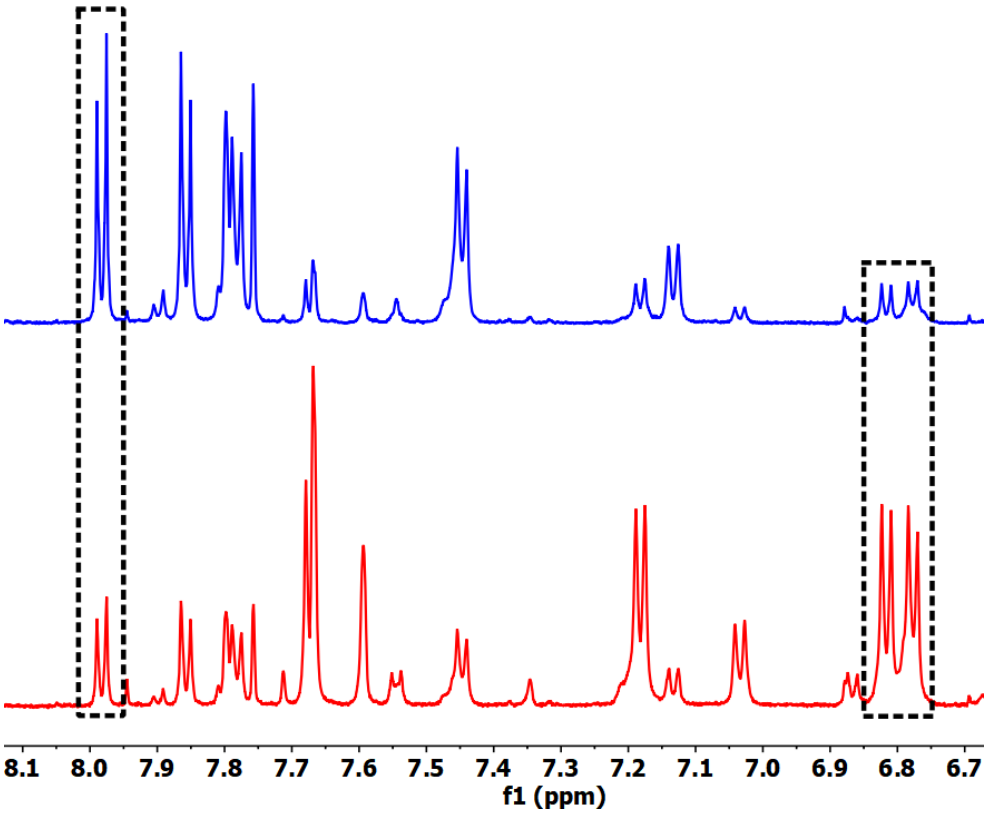


Figure 3.06 Stacked ^{1}H NMR spectra excerpts of the *trans*-enriched (blue) and *cis*-enriched (red) states of azobenzene-containing proteasome inhibitor **3.01**. The peaks used to determine the *trans*:*cis* ratio of each isomeric state are highlighted with a dashed black box.

Table 3.01 Ratio of the *cis* and *trans* isomers of **3.01** in their thermally adapted and photostationary states with the half-life of ***cis*-3.01** as determined by the integration of the azobenzene proton signals in the ^{1}H NMR spectra.

State	*Trans*:*cis* ratio	$t_{1/2}$ (min)
Thermally adapted	85:15	-
Photostationary	25:75	487.2

3.5 Proteasome Inhibition Assay

The *cis*- and *trans*-enriched isomeric states of compound **3.01** were evaluated against the β1 and β5 subunits of the proteasome under the same assay conditions used in Chapter 2 to evaluate compounds **2.01**-**2.04**. Serial dilutions of **3.01** were irradiated in 100% DMSO with an 8 W 365 nm UV lamp to reach the photostationary state before dilution with buffer to the appropriate concentrations. The fluorescence produced by AMC cleaved from the substrates was measured over time to avoid measuring inhibition by the *trans* isomer resulting from thermal *cis*-to-*trans* isomerisation from the photostationary state. The activity of the enzyme in each well was then determined by the slope of the initial linear period of the trace. Enzyme activity was calculated in the same manner for all wells in order to maintain consistency. The slope of the fluorescence trace from the positive control wells was again designated as 100% activity to normalise the activity of the enzyme with inhibitor. The activity of the enzyme at each concentration was then converted to a fraction of the positive control and plotted against the log of the inhibitor concentration (Figure 3.07). These plots were subsequently fitted with a four-parameter curve which was and used to calculate the IC_{50} values of the two isomeric states.

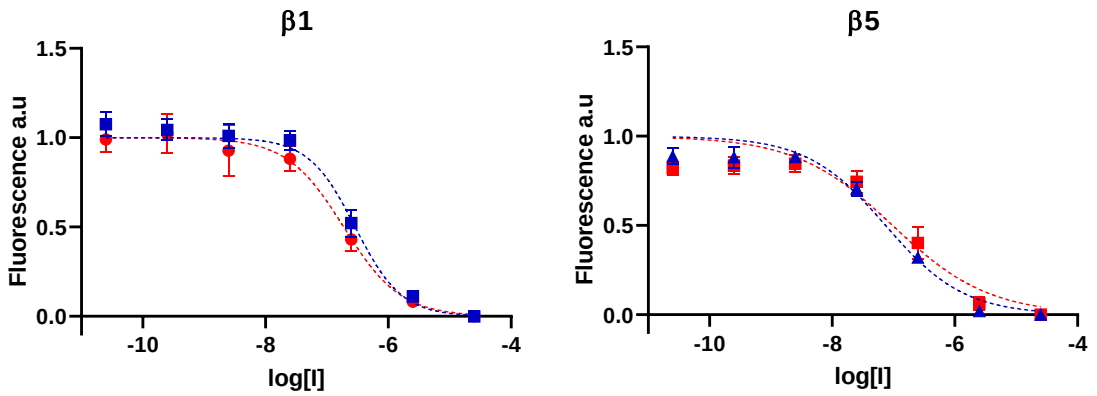
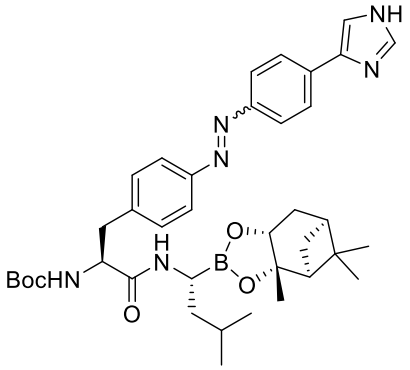


Figure 3.07 Inhibition curves of the thermally adapted (blue) and photostationary (red) states with a four-parameter inhibitor vs. response fit against β1 and β5 subunits of the rabbit 20S proteasome.

The IC_{50} values for the *trans*-enriched and *cis*-enriched isomeric states are shown in Table 3.02. Both states have low nanomolar-range IC_{50} values against the β5 subunit of the rabbit 20S proteasome. The thermally adapted state has a slightly lower IC_{50} value against the β5 subunit compared to the photostationary state, indicating that ***trans*-3.01** is a more potent inhibitor than ***cis*-3.01**. However, considering the drastic difference in geometry of the primed site occupying azobenzene substituent there is a remarkable similarity in potency. This result further confirms the promiscuity of the β5 primed site binding channel and indicates it can accommodate a diverse range of substituents with different sizes, geometries and polarities. In comparison to compounds **2.01**-**2.04**, the *trans*-enriched isomeric state of **3.01** is more potent against the β5 subunit, suggesting the primed site binding channel prefers bulky substituents. Interestingly, the *cis*-enriched isomeric state has a lower IC_{50} than the thermally adapted state against the β1 subunit of the proteasome, the opposite of the result seen against the β5 subunit. A similar result was observed by Blanco *et al.* for the biphenyl azobenzene derivative, suggesting a notable preference for the conformation of bulky azobenzene substituents at the P_2 position.

Table 3.02 IC_{50} values with standard errors of the thermally adapted and photostationary states of **3.01** against the β1, β2 and β5 subunits of the rabbit 20S proteasome.



	IC_{50} β1 ± SE (nM)	IC_{50} β5 ± SE (nM)
Thermally adapted state (*trans*-enriched)	312 ± 120	75 ± 13
Photostationary state (*cis*-enriched)	197 ± 59	98 ± 19
Fold difference	1.6	1.3

3.6 Chapter Conclusions

In summary, an azobenzene-containing photoswitchable inhibitor of the proteasome, **3.01**, was designed with the purpose of probing the promiscuity of the primed site binding channel of the proteasome with a substituent which has a relatively high degree of conformational predictability. The azobenzene moiety incorporated into the P2 position of a bortezomib-inspired inhibitor with a terminal imidazolyl group was demonstrated by molecular docking to have two distinct binding conformations within the primed site binding channel. An enzyme assay against the β1 and β5 subunits of the rabbit 20S proteasome was carried out to evaluate the potency of the thermally adapted and photostationary states of **3.01**. Remarkably, the IC_{50} value of both states against the β5 subunit were nearly identical despite the significant change in geometry and polarity associated with the isomerisation of the azobenzene substituent. The promiscuity of the primed site binding channel of the β5 subunit of the proteasome thus extends to accommodating substituents with less hydrophobicity. This result further suggests the promiscuity of the primed site binding channel also shown in Chapter 2 could be exploited by

occupying the binding channel with a substituent not accommodated by other proteases, therefore increasing the specificity towards the proteasome.

Chapter 4: Specificity of Proteasome Inhibitors

4.1 Introduction

Treatment of multiple myeloma and mantle cell lymphoma with the dipeptide boronic acid proteasome inhibitor bortezomib causes peripheral neuropathy in 35% to 52% of patients. A more recently developed and approved drug for the treatment of multiple myeloma is carfilzomib, a tetrapeptide epoxyketone proteasome inhibitor. The frequency of treatment-emergent peripheral neuropathy for patients receiving carfilzomib is up to 12%, a significant reduction in frequency compared to bortezomib. Arastu-Kapur *et al.*[73] demonstrated that *in vitro* bortezomib-induced neurodegeneration occurs via a proteasome-independent mechanism, indicating inhibition of other proteases may be the cause of peripheral neuropathy in patients receiving bortezomib. Additionally, bortezomib has potent inhibitory activity against the serine proteases cathepsin A and G, chymotrypsin, chymase and HtrA2/Omi whilst carfilzomib does not.[73] HtrA2/Omi is a protease involved with neuronal cell survival[122] and was identified as a possible off-target of bortezomib involved in the mechanism that causes bortezomib-induced peripheral neuropathy.[73]

In Chapters 2 and 3, the results indicated that the primed site binding channel of the β5 subunit of the proteasome was promiscuous regarding the accommodation of substituents with varied size, geometry and polarities. As such, the primed site binding channel was identified as a possible target for accommodating substituents which may not be accommodated by other proteases, thus creating an inhibitor with greater specificity towards the proteasome. Furthermore, carfilzomib demonstrates that an increase in specificity is correlated to reduced side effects,[73] making the specificity of proteasome inhibitors a vital characteristic in their design. Therefore, compounds **2.01-2.04** and the *trans*-enriched thermally adapted state of **3.01** were thus evaluated in an enzyme assay against bovine α-chymotrypsin alongside bortezomib.

4.2 α-Chymotrypsin Inhibition Assay

The assay performed utilised the substrate Ala-Ala-Phe-amc which is cleaved by bovine α-chymotrypsin to release 7-amino-4-methylcoumarin (AMC) as a fluorescent reporting group in order to measure the activity of the enzyme. 7 serial dilutions of each inhibitor were tested; however, the highest concentration was omitted from analysis as compounds **2.01**, **2.02** and **3.01** were insoluble at this concentration. Initiation of the assay was via addition of the enzyme and fluorescence was measured continuously over 10 minutes. The slope of each fluorescence trace is thus proportional to the activity of the enzyme, where the 100% activity positive control was the slope from the enzyme and substrate without inhibitor. Each slope was converted to a fraction of the slope of the positive control and plotted against the log of the concentration of inhibitor. These data were subsequently analysed by 'GraphPad Prism 8' to give a four-parameter fit from which the IC_{50} values could be calculated for each inhibitor (Figure 4.01).

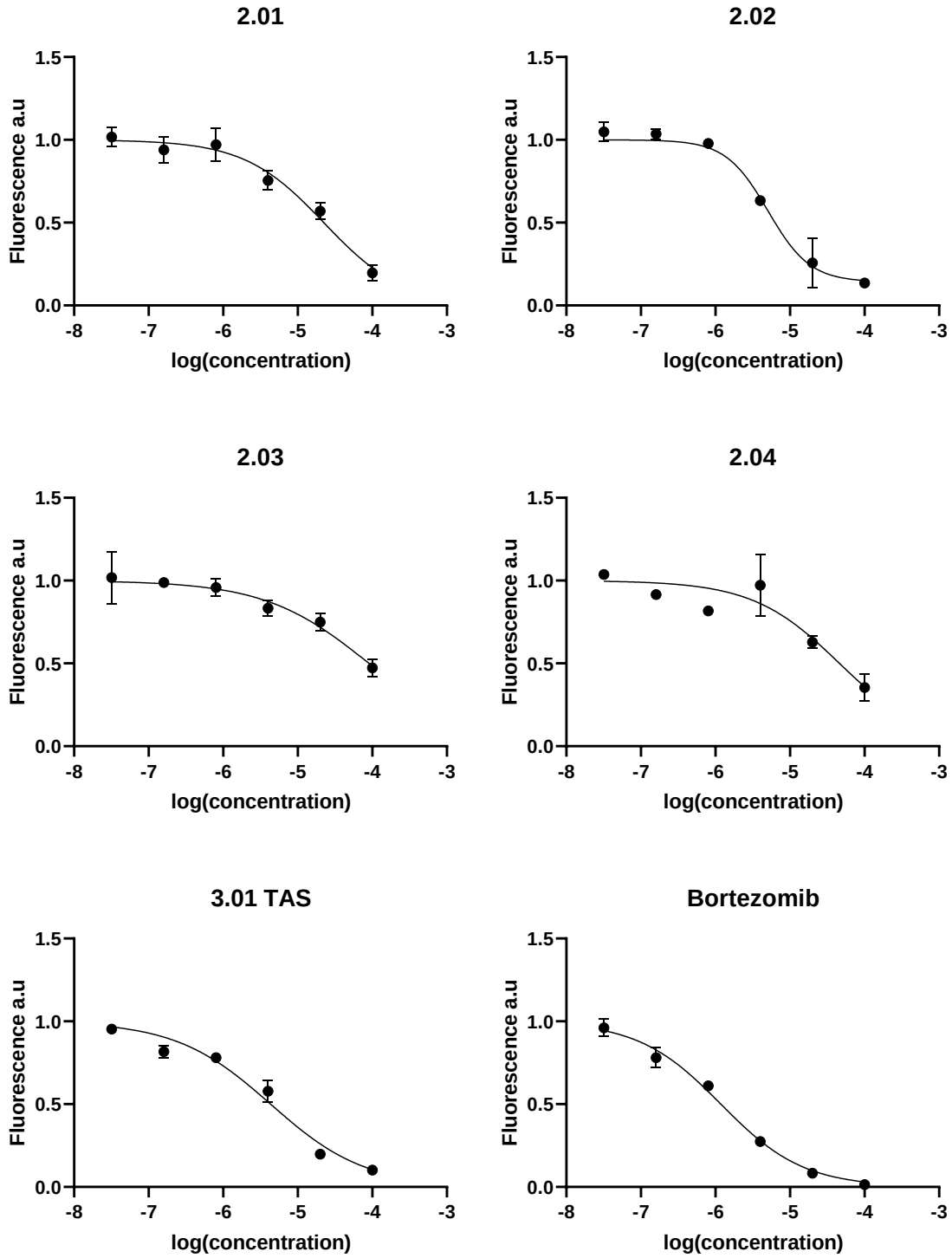


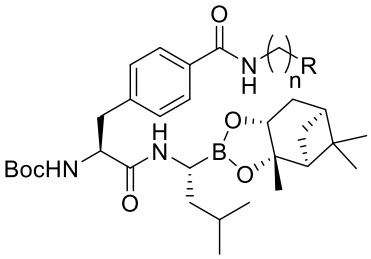
Figure 4.01 Inhibition curves of compounds **2.01-2.04**, and the *trans*-enriched thermally adapted state of **3.01** (3.01 TAS) with a four-parameter inhibitor vs. response fit against bovine α-chymotrypsin.

The IC_{50} values of compounds **2.01-2.04**, and the *trans*-enriched thermally adapted state of **3.01** are reported in Table 4.01 with the IC_{50} values of each compound against the β5 subunit of the proteasome. The specificity of each compound (cmpd) relative to bortezomib for the β5 subunit

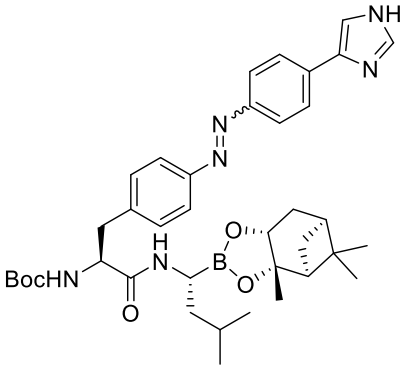
of the rabbit 20 S proteasome over bovine α-chymotrypsin (chymo.) is also reported as calculated by the formula:

$$Selectivity\ relative\ to\ bortezomib = \left(\frac{cmpd\ \mathrm{IC_{50}}(chymo.)}{cmpd\ \mathrm{IC_{50}}(proteasome)}\right)\Big/\left(\frac{bort\ \mathrm{IC_{50}}(chymo.)}{bort\ \mathrm{IC_{50}}(proteasome)}\right).$$

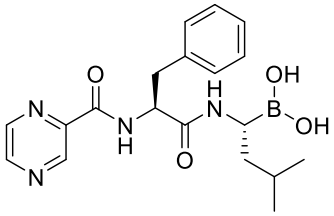
Table 4.01 IC_{50} values with standard errors of compounds **2.01**-**2.04**, the *trans*-enriched thermally adapted state of **3.01** and bortezomib against bovine α-chymotrypsin the β5 subunit of the rabbit 20S proteasome. As a measure of specificity, the ratio of the IC_{50} of each compound relative to bortezomib is reported as the specificity relative to bortezomib.



2.01 n=2, R=Ph
2.02 n=3, R=Ph
2.03 n=2, R=4-Im
2.04 n=3, R=4-Im



3.01
85:15 trans:cis



Bortezomib

Compound	Bovine α-chymotrypsin IC_{50} ± SE (nM)	β5 IC_{50} ± SE (nM)	Specificity relative to bortezomib
2.01	28 000 ± 17 000	149 ± 22	3.76
2.02	5 100 ± 800	95 ± 14	1.07
2.03	>90 000	720 ± 71	>2.5
2.04	>45 000	312 ± 17	>2.88
3.01	4 600 ± 1 600	155 ± 27	0.59
Bortezomib	1 200 ± 200	24 ± 3	-

Bortezomib is again the most potent inhibitor against bovine α-chymotrypsin as it was against the β1 and β5 subunits of the proteasome. All compounds are relatively low potency inhibitors in the micromolar range of bovine α-chymotrypsin. Similar to the β5 subunit of the proteasome,

three carbon alkyl chain analogues, **2.02** and **2.04**, are more potent than their respective two carbon alkyl chain analogues, **2.01** and **2.03**. Further similarity to the potency against the β5 subunit is present with imidazolyl head group compounds, **2.03** and **2.04**, being less potent against α-chymotrypsin than the phenyl head group compounds, **2.01** and **2.02**. The specificity as measured by the ratio of IC_{50} values against each target relative to bortezomib indicates that compounds **2.01**-**2.04** are all more specific towards the β5 subunit of the proteasome over bovine α-chymotrypsin. Relative to bortezomib, compounds **2.01**, **2.03** and **2.04** considerably more specific towards inhibition of the β5 subunit of the proteasome over inhibition of α-chymotrypsin whilst compound **2.02** has slight specificity for the β5 subunit of the proteasome over α-chymotrypsin. Interestingly, the *trans*-enriched thermally adapted state of **3.01** is less specific towards inhibition of inhibition of the β5 subunit of the proteasome over inhibition of α-chymotrypsin relative to bortezomib. These results suggest that the occupying the primed site binding channel of the β5 subunit of the proteasome is an effective strategy for increasing the specificity of proteasome inhibitors towards the β5 subunit over α-chymotrypsin.

4.3 Chapter Conclusions

Compounds **2.01**-**2.04** and the *trans*-enriched thermally adapted state of **3.01** were evaluated alongside bortezomib in a fluorescence assay against bovine α-chymotrypsin. Whilst all compounds have relatively low potency against α-chymotrypsin with IC_{50} values in the micromolar range, compounds **2.01**, **2.02** and **2.04** show a considerable improvement upon the specificity of bortezomib for the β5 subunit of the proteasome over chymotrypsin. The thermally adapted state of **3.01** has approximately half the specificity of bortezomib for the β5 subunit of the proteasome over α-chymotrypsin. The improvement in specificity shown by compounds **2.01**, **2.02** and **2.04** demonstrate evidence towards proteasome inhibitor specificity being improved by occupation of the primed side binding channel. Evaluation of these compounds against other serine proteases inhibited by bortezomib, such as cathepsin A and G, chymase and HtrA2/Omi, is necessary to fully address whether occupying the primed site

binding channel of the β5 subunit of the proteasome is an effective strategy for improving specificity towards the proteasome. If overall specificity towards the proteasome can be significantly improved via occupying the primed site binding channel, an optimised primed site occupying inhibitor may improve upon the side effect profile of bortezomib.

Chapter 5: Experimental Methods

5.1 General Methods

5.1.1 Chemical Syntheses

All starting materials and reagents were purchased from commercial sources and used without further purification. Anhydrous DMF and methanol were purchased from Sigma-Aldrich (Castle Hill, NSW). Reported yields are isolated yields of pure product according to NMR spectroscopy and/or analytical HPLC.

5.1.2 Chromatography

Thin-layer chromatography (TLC) was performed on Merck aluminium TLC plate with silica gel 60 F_{254}. Developed TLC plates were visualised under a 254 nm UV lamp and subsequently dipped in either potassium permanganate, ninhydrin or vanillin solutions. Flash chromatography was performed with Carl Roth silica gel 60 230-400 mesh with a positive pressure of nitrogen. Analytical HPLC using a Phenomenex Luna C18(2) column was used to confirm product purity for all synthesised compounds being evaluated in enzyme assays.

5.1.3 NMR Spectroscopy

NMR spectra of compounds were obtained at room temperature on a 500 MHz Agilent DD2 console using VnmrJ 4.2 software or a 600 MHz Agilent DD2 console + cryoprobe using VnmrJ 4.2 software. Reported chemical shifts are in parts per million (ppm) on a δ scale where TMS is referenced at 0.00 ppm. Solvents used for NMR spectroscopy analysis were $CDCl_3$, d_6-DMSO and CD_3OD and their reference peaks can be found in Gottlieb *et al.*[123] Peak spin multiplicities are indicated by combinations of the following symbols: singlet (s), broad singlet (br s), doublet (d), triplet (t), quartet (q), pentet (p) and multiplet (m).

5.1.4 Mass Spectrometry

Product identity was confirmed by High Resolution Mass Spectrometry (HRMS) on an Agilent ESI-TOF LCMS. All masses reported are within 5 ppm of the calculated theoretical masses.

5.1.5 Photoswitching

Compound **3.01** was dissolved in DMSO or d_6-DMSO and irradiated with a Philips TL 8W BLB 365 nm lamp for 1.5 h in 90-well plates or in a quartz NMR tube. Water was excluded from the atmosphere of the quartz NMR tube by dissolving the sample of **3.01** in a freshly cracked vial of d_6-DMSO in a glove box. Any experiments with the *cis*-enriched photostationary state of **3.01** were performed immediately after irradiation to avoid *cis*-to-*trans* isomerisation.

5.2 Synthesis

5.1.1 General Procedures

General Procedure A: HATU or HBTU Mediated Amide Coupling

To a solution of carboxylic acid (1.1 equiv), HATU or HBTU (1.1 equiv) and DIPEA (5 equiv) in anhydrous DMF (50 mL/g of acid) was added amine (1.0 equiv). The mixture was stirred under N_2 at room temperature overnight before being diluted by sat. aq. NH_4Cl (500 mL/g of acid) and extracted with EtOAc (3 x 500 mL/g of acid). The combined organic extracts were washed with water (3 x 500 mL/g of acid) and brine (500 mL/g of acid) and then dried over $MgSO_4$ and filtered. The solvent was removed *in vacuo* and the resulting crude product purified by flash chromatography to give the pure amide.

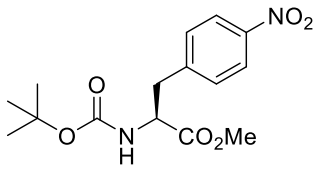
General Procedure B: LiOH Hydrolysis of Amino Acid Methyl Esters

To a solution of methyl ester (1.0 equiv) in THF (50 mL/g of methyl ester) cooled on ice was added a 1M aq. solution of LiOH (4.0 equiv). The mixture was stirred vigorously whilst being allowed to warm to room temperature for 3 h. The solution was acidified to pH 1 with 1M aq.

HCl and extracted with EtOAc (3 x 200 mL/g of methyl ester). The combined organic layers were dried over $MgSO_4$, filtered and concentrated *in vacuo* to give the pure carboxylic acid.

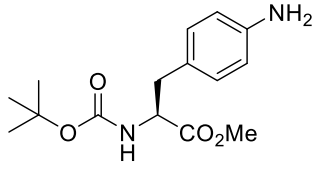
5.1.2 Synthesis for Chapter 2

(*S*)-methyl 2-[(*tert*-butoxycarbonyl)amino]-3-(4-nitrophenyl)propanoate **2.05**



To a suspension of *N*-Boc-4-nitro-L-phenylalanine (1.00 g, 3.22 mmol) and Na_2CO_3 (1.71 g, 16.1 mmol) in DMF (10 mL) was added methyl iodide (1.00 mL, 16.1 mmol) and the mixture stirred at room temperature overnight. Excess methyl iodide and DMF were mostly removed *in vacuo* and the residue was diluted with water (20 mL). The resulting solid was collected via vacuum filtration and washed with water to give *N*-Boc-4-nitro-L-phenylalanine methyl ester **2.05** as an off-white solid (1.04 g, 99%). ^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.19 – 8.13 (m, 2H), 7.35 – 7.28 (m, 2H), 5.04 (d, *J* = 6.1 Hz, 1H), 4.63 (d, *J* = 7.2 Hz, 1H), 3.74 (s, 3H), 3.27 (dd, *J* = 13.7, 5.8 Hz, 1H), 3.12 (dd, *J* = 14.0, 6.3 Hz, 1H), 1.41 (s, 9H). ^{13}C NMR (126 MHz, Chloroform-*d*) δ 130.4, 123.8, 54.3, 52.7, 38.6, 28.4.

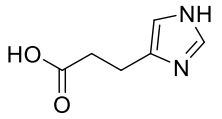
(*S*)-methyl 3-(4-aminophenyl)-2-[(*tert*-butoxycarbonyl)amino]propanoate **2.06**



N-Boc-4-nitro-L-phenylalanine methyl ester **2.05** (0.97 g, 2.99 mmol), zinc dust (1.96 g, 29.90 mmol) and ammonium chloride (2.03 g, 44.85 mmol) was stirred in MeOH (20 mL) for 30 minutes at 0 °C. Water (10 mL) was added and the reaction mixture was stirred at room temperature for 2 h. The mixture was filtered through Celite® and rinsed through with MeOH (20 mL). The filtrate was concentrated *in vacuo* and the residue taken up in EtOAc (30 mL) and washed with water (2 x 30 mL) and sat. aq. $NaHCO_3$ (30 mL). The organic layer was dried over

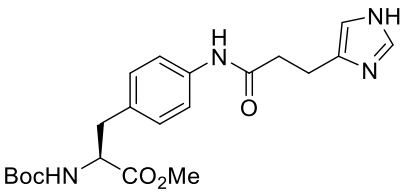
$MgSO_4$, filtered and concentrated *in vacuo* to give *N*-Boc-4-amino-L-phenylalanine methyl ester **2.06** as an orange oil (0.72 g, 82%). R_f (55% EtOAc/hexane) = 0.53. ^{1}H NMR (500 MHz, Chloroform-*d*) δ 6.93 – 6.87 (m, 2H), 6.64 – 6.58 (m, 2H), 4.93 (d, *J* = 8.2 Hz, 1H), 4.53 – 4.48 (m, 1H), 3.70 (s, 3H), 3.60 (s, 2H), 2.97 (t, *J* = 5.9 Hz, 2H), 1.42 (s, 9H). ^{13}C NMR (126 MHz, Chloroform-*d*) δ 172.7, 145.5, 130.3, 125.9, 115.4, 54.8, 52.2, 37.6, 28.5.

3-(1*H*-imidazol-4-yl)propanoic acid **2.07**



To a suspension of urocanic acid (100 mg, 0.72 mmol) and 10% Pd/C (200 mg) in MeOH (5 mL) under N_2 was added NH_4HCO_2 (410 mg, 6.5 mmol). The mixture was stirred and heated at reflux for 6h. The mixture was allowed to cool to room temperature and filtered over Celite®. The filtrate was concentrated *in vacuo* to give compound **2.07** as a white solid (100 mg, 99%). R_f (25% MeOH/DCM) = 0.03. ^{1}H NMR (500 MHz, DMSO-d_6) δ 7.40 (s, 1H), 6.59 (s, 1H), 2.64 (t, *J* = 7.4 Hz, 2H), 2.19 (t, *J* = 7.4 Hz, 2H). ^{13}C NMR (126 MHz, DMSO-d_6) δ 175.8, 133.8, 96.0, 38.7, 37.5, 12.6.

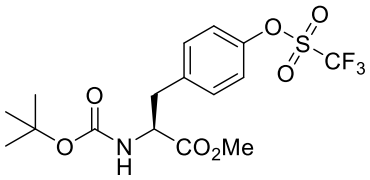
(*S*)-methyl 3-{4-[3-(1*H*-imidazol-4-yl)propanamido]phenyl}-2-[(*tert*-butoxycarbonyl)amino]propanoate **2.08**



HATU-mediated coupling method: The coupling of carboxylic acid **2.07** (52 mg, 0.37 mmol) with aniline **2.06** (100 mg, 0.34 mmol) was attempted via general procedure A using HATU. No product was formed according to TLC and MS.

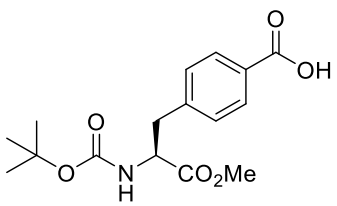
PyBOP-mediated coupling method: To a solution of carboxylic acid **2.07** (47 mg, 0.34 mmol), PyBOP (193 mg, 0.37 mmol) and DIPEA (0.18 mL, 1.02 mmol) in anhydrous DMF (5 mL) was added aniline **2.06** (100 mg, 0.34 mmol). The mixture was stirred under N_2 at room temperature overnight. No product was formed according to TLC and MS.

(*S*)-methyl 2-[(*tert*-butoxycarbonyl)amino]-3-(4-{[(trifluoromethyl)sulfonyl]oxy}phenyl)propanoate **2.09**



To a solution of *N*-Boc-L-tyrosine methyl ester (3.00 g, 10.16 mmol) and *N*-phenyl-bis(trifluoromethane sulfonimide) (3.99 g, 11.17 mmol) in dry DMC at 0 °C was added triethylamine (1.56 mL, 11.17 mmol). The reaction mixture was stirred for 1.5 h at 0 °C and then allowed to warm to room temperature whilst stirring for 1 h. The mixture was washed with water (20 mL), 1 M NaOH (3 x 20 mL), and brine (20 mL). The organic layer was dried over $MgSO_4$, filtered and concentrated *in vacuo* to give **2.09** as a clear, colourless oil (4.34 g, 99%). R_f (1:2 EtOAc/hexane) = 0.58. ^{1}H NMR (500 MHz, DMSO-d_6) δ 7.42 (t, J = 7.0 Hz, 4H), 4.28 – 4.19 (m, 1H), 3.61 (s, 3H), 3.07 (dd, J = 13.8, 4.9 Hz, 1H), 2.89 (dd, J = 13.8, 10.5 Hz, 1H), 1.29 (s, 9H). ^{13}C NMR (126 MHz, DMSO-d_6) δ 155.3, 147.9, 138.8, 131.3, 121.0, 118.3 (d, J = 317 Hz), 78.3, 54.6, 51.8, 45.7, 35.7, 28.0.

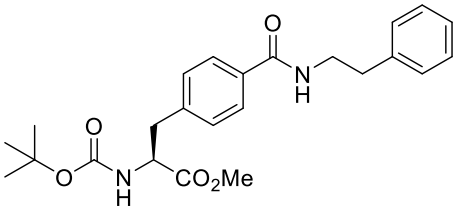
(*S*)-4-{2-[(*tert*-butoxycarbonyl)amino]-3-methoxy-3-oxopropyl}benzoic acid **2.10**



A solution of triflate **2.09** (4.32 g, 10.11 mmol), 1,3-bis(diphenylphosphino)propane (0.25 g, 0.06 mmol) and $Pd(OAc)_2$ (0.068 g, 0.03 mmol) in DMF/H_2O (15 mL, 3:1 v/v) was saturated

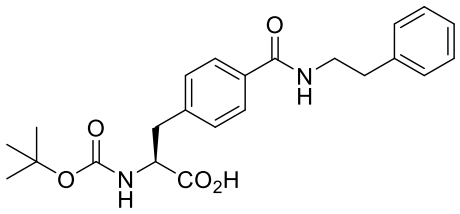
with gaseous CO for 10 min whilst heating to 70 °C. DIPEA (1.94 mL, 11.12 mmol) was added and the mixture was stirred under CO atmosphere at 70 °C for 1.5 h and then left to cool to room temperature and stir overnight. Ethyl acetate (30 mL) and sat. aq. $NaHCO_3$ (5 mL) were added and the layers separated. The organic layer was extracted with sat. aq. $NaHCO_3$ (2 x 10 mL). The aqueous layers were combined and acidified with 1 M aq. HCl until a precipitate formed. Ethyl acetate (20 mL) was added and the aqueous layer was extracted further with ethyl acetate (2 x 20 mL). The combined organic layers were dried over $MgSO_4$, filtered and concentrated *in vacuo*. Lyophilization yielded **2.10** as a white powder (1.72 g, 53%). R_f (40% EtOAc/hexane) = 0.32. ^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.02 (d, *J* = 7.9 Hz, 2H), 7.24 (d, *J* = 8.1 Hz, 2H), 5.02 (d, *J* = 7.5 Hz, 1H), 4.63 (d, *J* = 7.5 Hz, 1H), 3.73 (s, 3H), 3.21 (dd, *J* = 13.5, 5.7 Hz, 1H), 3.11 (dd, *J* = 13.6, 6.1 Hz, 1H), 1.42 (s, 9H). ^{13}C NMR (126 MHz, Chloroform-*d*) δ 130.6, 129.7, 52.5, 28.4.

(*S*)-methyl 2-[(*tert*-butoxycarbonyl)amino]-3-[4-(phenethylcarbamoyl)phenyl]propanoate **2.11**



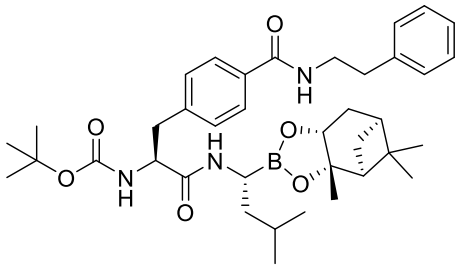
N-Boc-4-carboxy-L-phenylalanine methyl ester **2.10** (100 mg, 0.309 mmol) was coupled with phenethylamine (34 mg, 0.282 mmol) according to general procedure A using HATU, with the residue purified by flash chromatography (silica, 50% EtOAc/hexane) to give compound **2.11** as a white solid (83 mg, 63%). R_f (50% EtOAc/hexane) = 0.48. 1H NMR (500 MHz, Chloroform-*d*) δ 7.61 (d, *J* = 7.9 Hz, 2H), 7.34 (dd, *J* = 8.4, 6.5 Hz, 2H), 7.17 (d, *J* = 7.8 Hz, 2H), 6.07 (s, 1H), 4.96 (d, *J* = 8.2 Hz, 1H), 4.62 – 4.57 (m, 1H), 3.76 – 3.68 (m, 5H), 3.16 (dd, *J* = 13.9, 5.8 Hz, 1H), 3.07 (dd, *J* = 13.4, 6.0 Hz, 1H), 2.93 (t, *J* = 6.9 Hz, 2H), 1.41 (s, 9H). ^{13}C NMR (126 MHz, Chloroform-*d*) δ 129.7, 198.0, 128.9, 127.2, 65.6, 53.6, 35.9, 28.4.

(*S*)-2-[(*tert*-butoxycarbonyl)amino]-3-[4-(phenethylcarbamoyl)phenyl]propanoic acid **2.12**



Methyl ester **2.11** (65 mg, 0.152 mmol) was hydrolysed according to general procedure B to give compound **2.12** as a white solid (62 mg, 99%). R_f (10% MeOH/DCM) = 0.17. ^{1}H NMR (500 MHz, DMSO-d_6) δ 12.63 (s, 1H), 8.48 (t, *J* = 5.6 Hz, 1H), 7.73 (d, *J* = 7.9 Hz, 2H), 7.33 – 7.26 (m, 4H), 7.26 – 7.21 (m, 2H), 7.23 – 7.16 (m, 1H), 7.10 (d, *J* = 8.5 Hz, 1H), 4.11 (td, *J* = 9.5, 4.3 Hz, 1H), 3.50 – 3.43 (m, 1H), 3.06 (dd, *J* = 13.9, 4.5 Hz, 1H), 2.91 – 2.80 (m, 3H), 1.31 (s, 9H).

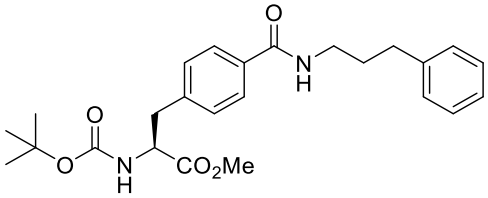
tert-butyl [(*S*)-1-({(*R*)-3-methyl-1-[(3a*S*,4*S*,6*S*,7a*R*)-3a,5,5-trimethylhexahydro-4,6-methanobenzo[d][1,3,2]dioxaborol-2-yl]butyl}amino)-1-oxo-3-[4-(phenethylcarbamoyl)phenyl]propan-2-yl]carbamate **2.01**



Carboxylic acid **2.12** (60 mg, 0.145 mmol) was coupled with (*R*)-Boroleucine (1*S*,2*S*,3*R*,5*S*)-(+)-2,3-pinanediol ester trifluoroacetate (35 mg, 0.132 mmol) according to general procedure A using HATU, with the residue purified by flash chromatography (silica, 50% EtOAc/hexane) to give compound **2.01** as a white solid (40 mg, 46%). R_f (50% EtOAc/hexane) = 0.41. Analytical HPLC: 20-100% aq. acetonitrile over 20 min R_t = 17.0 min. ^{1}H NMR (500 MHz, Methanol-d_4) δ 7.73 (d, *J* = 7.9 Hz, 2H), 7.35 (d, *J* = 7.9 Hz, 2H), 7.31 – 7.22 (m, 4H), 7.22 – 7.15 (m, 1H), 4.51 (t, *J* = 7.6 Hz, 1H), 4.23 – 4.17 (m, 1H), 3.59 (t, *J* = 7.4 Hz, 2H), 3.10 (dd, *J* = 13.7, 6.9 Hz, 1H), 2.98 (dd, *J* = 13.8, 8.3 Hz, 1H), 2.90 (t, *J* = 7.4 Hz, 2H), 2.68 (t, *J* = 7.8 Hz, 1H), 2.39 – 2.30 (m, 1H), 2.16 – 2.09 (m, 1H), 1.96 (t, *J* = 5.6 Hz, 1H), 1.90 – 1.84 (m,

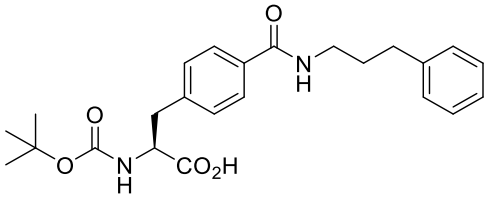
1H), 1.79 (d, *J* = 14.1 Hz, 1H), 1.57 – 1.48 (m, 1H), 1.47 – 1.22 (m, 18H), 0.88 (h, *J* = 6.7 Hz, 9H). ^{13}C NMR (126 MHz, Methanol-*d*$_4$) δ 130.7, 129.9, 129.5, 77.4, 39.2, 36.6, 29.7, 28.6, 27.8, 26.6. HRMS (ESI+) calcd for $C_{38}H_{54}BN_3O_6$: 682.4003 [$(M+Na)^+$], found: 682.4009.

(*S*)-methyl 2-[(*tert*-butoxycarbonyl)amino]-3-{4-[(3-phenylpropyl)carbamoyl]phenyl}propanoate **2.13**



N-Boc-4-carboxy-L-phenylalanine methyl ester **2.10** (100 mg, 0.309 mmol) was coupled with 3-phenylpropylamine (40 µL, 0.141 mmol) according to general procedure A using HATU, with the residue purified by flash chromatography (silica, 50% EtOAc/hexane) to give compound **2.13** as a white solid (108 mg, 88%). R_f (50% EtOAc/hexane) = 0.45. ^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.59 (d, *J* = 8.0 Hz, 2H), 7.33 – 7.26 (m, 2H), 7.24 – 7.15 (m, 5H), 6.00 (s, 1H), 4.97 (d, *J* = 8.2 Hz, 1H), 4.60 (d, *J* = 7.2 Hz, 1H), 3.71 (s, 3H), 3.50 (q, *J* = 6.6 Hz, 2H), 3.17 (dd, *J* = 13.6, 5.8 Hz, 1H), 3.07 (dd, *J* = 14.0, 6.0 Hz, 1H), 2.74 (t, *J* = 7.5 Hz, 2H), 1.97 (p, *J* = 7.3 Hz, 2H), 1.42 (s, 9H). ^{13}C NMR (126 MHz, Chloroform-*d*) δ 129.7, 128.7, 128.6, 127.2, 126.2, 77.4, 77.2, 76.9, 53.6, 40.1, 28.5, 15.6.

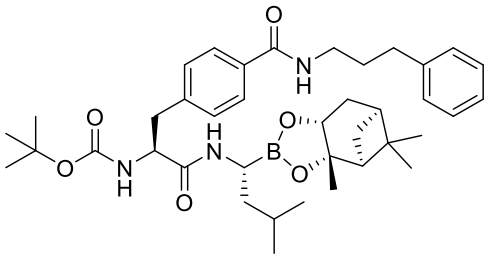
(*S*)-2-[(*tert*-butoxycarbonyl)amino]-3-{4-[(3-phenylpropyl)carbamoyl]phenyl}propanoic acid **2.14**



Methyl ester **2.13** (108 mg, 0.245 mmol) was hydrolysed according to general procedure B to give compound **2.14** as a white solid (103 mg, 99%). R_f (10% MeOH/DCM) = 0.10. ^{1}H NMR (500 MHz, DMSO-*d*$_6$) δ 12.63 (s, 1H), 8.40 (t, *J* = 5.7 Hz, 1H), 7.75 (d, *J* = 7.9 Hz, 2H), 7.34

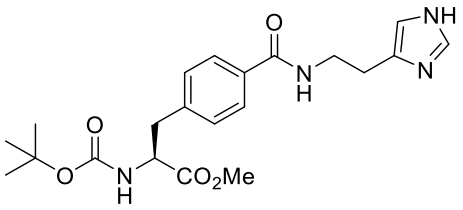
– 7.25 (m, 4H), 7.25 – 7.20 (m, 2H), 7.20 – 7.14 (m, 1H), 7.10 (d, *J* = 8.4 Hz, 1H), 4.15 – 4.07 (m, 1H), 3.30 – 3.23 (m, 2H), 3.06 (dd, *J* = 13.8, 4.6 Hz, 1H), 2.87 (dd, *J* = 13.8, 10.3 Hz, 1H), 2.62 (t, *J* = 7.7 Hz, 2H), 1.82 (p, *J* = 7.3 Hz, 2H), 1.31 (s, 9H).

tert-butyl [(*S*)-1-({(*R*)-3-methyl-1-[(3a*S*,4*S*,6*S*,7a*R*)-3a,5,5-trimethylhexahydro-4,6-methanobenzo[d][1,3,2]dioxaborol-2-yl]butyl}amino)-1-oxo-3-{4-[(3-phenylpropyl)carbamoyl]phenyl}propan-2-yl]carbamate **2.02**



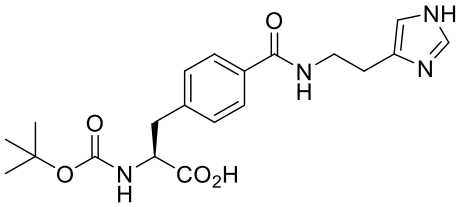
Carboxylic acid **2.14** (100 mg, 0.234 mmol) was coupled with (*R*)-Boroleucine (1*S*,2*S*,3*R*,5*S*)-(+)-2,3-pinanediol ester trifluoroacetate (56 mg, 0.213 mmol) according to general procedure A using HATU, with the residue purified by flash chromatography (silica, 50% EtOAc/hexane) to give compound **2.02** as a white solid (85 mg, 59%). R_f (50% EtOAc/hexane) = 0.48. Analytical HPLC: 20-100% aq. acetonitrile over 20 min, R_t = 17.4 min. ^{1}H NMR (500 MHz, Methanol-d_4) δ 7.77 (d, *J* = 7.9 Hz, 2H), 7.36 (d, *J* = 8.0 Hz, 2H), 7.29 – 7.23 (m, 2H), 7.23 – 7.18 (m, 2H), 7.18 – 7.12 (m, 1H), 4.51 (t, *J* = 7.6 Hz, 1H), 4.20 (dd, *J* = 8.6, 2.3 Hz, 1H), 3.45 – 3.34 (m, 2H), 3.11 (dd, *J* = 13.7, 7.0 Hz, 1H), 2.99 (dd, *J* = 13.7, 8.3 Hz, 1H), 2.72 – 2.64 (m, 3H), 2.39 – 2.31 (m, 1H), 2.13 (d, *J* = 9.6 Hz, 1H), 1.99 – 1.89 (m, 3H), 1.89 – 1.84 (m, 1H), 1.79 (d, *J* = 14.1 Hz, 1H), 1.56 – 1.47 (m, 1H), 1.42 (d, *J* = 10.3 Hz, 1H), 1.37 (d, *J* = 6.5 Hz, 9H), 1.29 (d, *J* = 1.7 Hz, 6H), 1.27 – 1.21 (m, 2H), 0.93 – 0.83 (m, 9H). ^{13}C NMR (126 MHz, Methanol-d_4) δ 143.1, 141.6, 134.4, 130.8, 129.4, 128.5, 126.9, 84.9, 76.9, 41.8, 40.7, 39.2, 38.8, 34.4, 29.8, 28.6, 27.8, 26.6. HRMS (ESI+) calcd for $C_{39}H_{56}BN_3O_6$: 696.4160 [$(M+Na)^+$], found 696.4167.

(*S*)-methyl 3-(4-{[2-(1*H*-imidazol-4-yl)ethyl]carbamoyl}phenyl)-2-[(*tert*-butoxycarbonyl)amino]propanoate **2.15**



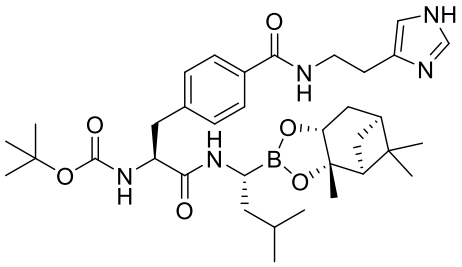
N-Boc-4-carboxy-L-phenylalanine methyl ester **2.10** (300 mg, 0.927 mmol) was coupled with histamine dihydrochloride (156 mg, 0.843 mmol) according to general procedure A using HBTU, with the residue purified by flash chromatography (silica, 10% MeOH/chloroform) to give compound **2.15** as a white solid (100 mg, 28%). R_f (10% MeOH/chloroform) = 0.16. ^{1}H NMR (500 MHz, Methanol-d_4) δ 7.72 (d, J = 7.9 Hz, 2H), 7.60 (s, 1H), 7.30 (d, J = 7.9 Hz, 2H), 6.87 (s, 1H), 4.42 – 4.36 (m, 1H), 3.70 (s, 3H), 3.61 (t, J = 7.2 Hz, 2H), 3.17 (dd, J = 13.7, 5.8 Hz, 1H), 3.01 – 2.84 (m, 3H), 1.37 (s, 9H). ^{13}C NMR (126 MHz, Chloroform-*d*) δ 129.6, 127.4, 54.4, 52.5, 40.0, 38.5, 28.4, 23.1.

(*S*)-3-(4-{[2-(1*H*-imidazol-4-yl)ethyl]carbamoyl}phenyl)-2-[(*tert*-butoxycarbonyl)amino]propanoic acid **2.16**



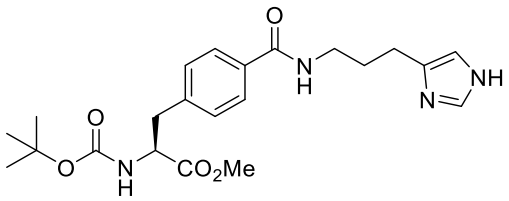
Methyl ester **2.15** (100 mg, 0.240 mmol) was hydrolysed according to general procedure B to give compound **2.16** as a white solid (32 mg, 33%). R_f (10% MeOH/DCM) = 0.00. ^{1}H NMR (500 MHz, Methanol-d_4) δ 8.77 (s, 1H), 7.71 (d, J = 7.9 Hz, 2H), 7.36 – 7.31 (m, 3H), 4.36 (dd, J = 9.2, 4.9 Hz, 1H), 3.69 (t, J = 6.7 Hz, 2H), 3.23 (dd, J = 13.8, 4.9 Hz, 1H), 3.03 (t, J = 6.7 Hz, 2H), 2.96 (dd, J = 13.8, 9.2 Hz, 1H), 1.37 (s, 9H).

tert-butyl [(*S*)-3-(4-{[2-(1*H*-imidazol-4-yl)ethyl]carbamoyl}phenyl)-1-({(*R*)-3-methyl-1-[(3a*S*,4*S*,6*S*,7a*R*)-3a,5,5-trimethylhexahydro-4,6-methanobenzo[d][1,3,2]dioxaborol-2-yl]butyl}amino)-1-oxopropan-2-yl]carbamate **2.03**



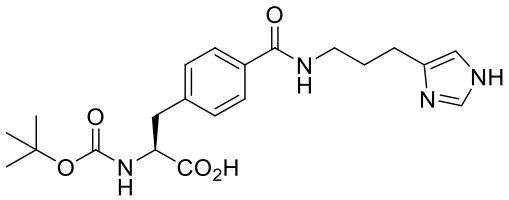
Carboxylic acid **2.16** (20 mg, 0.50 mmol) was coupled with (*R*)-Boroleucine (1*S*,2*S*,3*R*,5*S*)-(+)-2,3-pinanediol ester trifluoroacetate (12 mg, 0.0.45 mmol) according to general procedure A using HBTU, with the residue purified by flash chromatography (silica, 10% MeOH/DCM) to give compound **2.03** as a white solid (12 mg, 41%). R_f (10% MeOH/DCM) = 0.22. Analytical HPLC: 40-100% over 20 min, R_t = 8.1 min. ^{1}H NMR (500 MHz, Methanol-d_4) δ 8.00 (s, 1H), 7.75 (d, *J* = 8.0 Hz, 2H), 7.36 (d, *J* = 8.0 Hz, 2H), 7.03 (s, 1H), 4.57 – 4.46 (m, 2H), 4.20 (dd, *J* = 8.8, 2.2 Hz, 1H), 3.64 (t, *J* = 7.0 Hz, 2H), 3.11 (dd, *J* = 13.7, 6.7 Hz, 1H), 3.02 – 2.91 (m, 4H), 2.69 (t, *J* = 7.8 Hz, 1H), 2.39 – 2.31 (m, 1H), 2.23 – 2.12 (m, 2H), 2.05 – 2.02 (m, 1H), 1.96 (t, *J* = 5.7 Hz, 1H), 1.88 – 1.85 (m, 1H), 1.83 – 1.76 (m, 1H), 1.61 (s, 1H), 1.52 (p, *J* = 7.6, 7.1 Hz, 1H), 1.36 (s, 9H), 1.31 – 1.27 (m, 7H), 0.92 – 0.82 (m, 9H). ^{13}C NMR (126 MHz, Methanol-d_4) δ 211.5, 170.2, 157.8, 143.2, 134.9, 133.7, 133.0, 130.6, 128.3, 117.6, 108.8, 80.6, 70.6, 68.3, 56.1, 39.6, 38.5, 33.1, 32.1, 30.9, 30.7, 30.5, 29.5, 28.7, 25.9, 24.8, 23.7, 14.4. HRMS (ESI+) calcd for $C_{35}H_{52}BN_5O_6$: 650.4089 [$(M+H)^+$], found 650.4100.

(*S*)-methyl 3-(4-{[3-(1*H*-imidazol-4-yl)propyl]carbamoyl}phenyl)-2-[(*tert*-butoxycarbonyl)amino]propanoate **2.17**



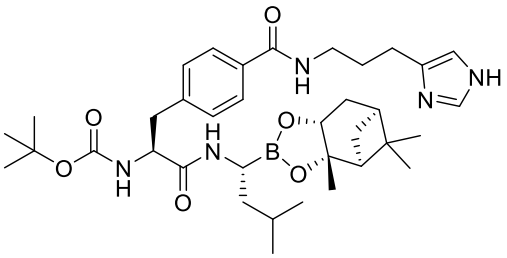
N-Boc-4-carboxy-L-phenylalanine methyl ester **2.10** (100 mg, 0.309 mmol) was coupled with 3-phenylpropylamine (56 mg, 0.281 mmol) according to general procedure A using HBTU, with the residue purified by flash chromatography (silica, 10% MeOH/DCM) to give compound **2.17** as a white solid (105 mg, 87%). R_f (10% MeOH/DCM) = 0.28. ^{1}H NMR (500 MHz, Methanol-d_4) δ 8.08 (s, 1H), 7.75 (d, *J* = 7.9 Hz, 2H), 7.32 (d, *J* = 8.0 Hz, 2H), 7.06 (s, 1H), 5.49 (s, 1H), 4.39 (dd, *J* = 9.2, 5.6 Hz, 1H), 3.70 (s, 3H), 3.43 (t, *J* = 6.9 Hz, 2H), 3.18 (dd, *J* = 13.9, 5.5 Hz, 1H), 2.96 (dd, *J* = 13.8, 9.2 Hz, 1H), 2.73 (t, *J* = 7.4 Hz, 2H), 1.96 (p, *J* = 7.2 Hz, 2H), 1.37 (s, 9H). ^{13}C NMR (126 MHz, Methanol-d_4) δ 173.9, 142.6, 130.5, 128.4, 52.7, 40.2, 38.5, 30.0, 28.6.

(*S*)-3-(4-{[3-(1*H*-imidazol-4-yl)propyl]carbamoyl}phenyl)-2-[(*tert*-butoxycarbonyl)amino]propanoic acid **2.18**



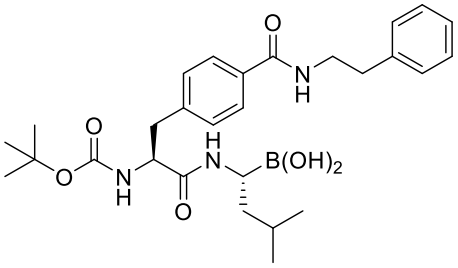
Methyl ester **2.17** (105 mg, 0.244 mmol) was hydrolysed according to general procedure B to give compound **2.18** as a white solid (52 mg, 51%). R_f (10% MeOH/DCM) = 0.00. ^{1}H NMR (500 MHz, Methanol-d_4) δ 8.74 (s, 1H), 7.75 (d, *J* = 7.9 Hz, 2H), 7.37 – 7.32 (m, 3H), 4.37 (dd, *J* = 9.1, 5.1 Hz, 1H), 3.45 (t, *J* = 6.8 Hz, 2H), 3.23 (dd, *J* = 14.0, 5.2 Hz, 1H), 2.97 (dd, *J* = 14.2, 9.4 Hz, 1H), 2.80 (t, *J* = 7.4 Hz, 2H), 2.03 – 1.94 (m, 2H), 1.37 (s, 9H).

tert-butyl [(*S*)-3-(4-{[3-(1*H*-imidazol-4-yl)propyl]carbamoyl}phenyl)-1-({(*R*)-3-methyl-1-[(3a*S*,4*S*,6*S*,7a*R*)-3a,5,5-trimethylhexahydro-4,6-methanobenzo[d][1,3,2]dioxaborol-2-yl]butyl}amino)-1-oxopropan-2-yl]carbamate **2.04**



Carboxylic acid **2.18** (45 mg, 0.105 mmol) was coupled with (*R*)-Boroleucine (1*S*,2*S*,3*R*,5*S*)-(+)-2,3-pinanediol ester trifluoroacetate (36 mg, 0.095 mmol) according to general procedure A using HBTU, with the residue purified by flash chromatography (silica, 10% MeOH/DCM) to give compound **2.04** as a white solid (18 mg, 29%). R_f (10% MeOH/DCM) = 0.22. Analytical HPLC: 40-100% aq. acetonitrile over 20 min, R_t = 8.3 min. ^{1}H NMR (500 MHz, Methanol-d_4) δ 8.20 (s, 1H), 7.79 (d, *J* = 7.9 Hz, 2H), 7.38 (d, *J* = 7.9 Hz, 2H), 7.10 (s, 1H), 4.52 (t, *J* = 7.6 Hz, 1H), 4.21 (dd, *J* = 8.7, 2.3 Hz, 1H), 3.47 – 3.40 (m, 2H), 3.12 (dd, *J* = 13.7, 6.8 Hz, 1H), 3.04 – 2.96 (m, 1H), 2.69 (t, *J* = 7.7 Hz, 1H), 2.40 – 2.31 (m, 1H), 2.17 – 2.10 (m, 1H), 2.01 – 1.92 (m, 2H), 1.91 – 1.84 (m, 1H), 1.84 – 1.76 (m, 1H), 1.54 (s, 1H), 1.45 – 1.22 (m, 21H), 0.94 – 0.81 (m, 9H). ^{13}C NMR (126 MHz, Methanol-d_4) δ 177.4, 169.9, 157.3, 141.7, 134.2, 130.7, 128.5, 84.4, 80.9, 77.5, 54.2, 53.6, 41.7, 41.4, 40.3, 39.2, 38.8, 37.6, 33.1, 30.7, 30.4, 30.3, 30.2, 29.7, 28.6, 28.4, 28.1, 27.8, 27.4, 26.6, 24.6, 23.7, 23.6, 22.2, 14.4. HRMS (ESI+) calcd for $C_{36}H_{54}BN_5O_6$: 686.4065 [$(M+Na)^+$], found 686.4070.

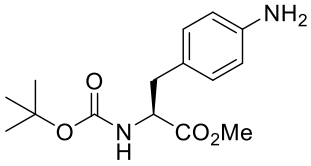
((*R*)-1-{(*S*)-2-[(*tert*-butoxycarbonyl)amino]-3-[4-(phenethylcarbamoyl)phenyl]propanamido}-3-methylbutyl)boronic acid **2.19**



To a solution of boronic ester **2.01** (10 mg, 0.019 mmol) in methanol (2 mL) and hexane (3 mL) was added (2-methylpropyl)boronic acid (3.5 mg, 0.034 mmol) and 1 M aq. HCl. (53 µL, 0.053 mmol) The mixture was stirred overnight at room temperature and the the methanol layer washed with hexane (2 x 10 mL). The aqueous methanol was removed *in vacuo* and the residue taken up in DCM (15 mL), washed with sat. aq. $NaHCO_3$ (2 x 10 mL), dried over $MgSO_4$, filtered and concentrated *in vacuo*. Purification of the crude residue was attempted via preparative TLC, which was developed 3 times in 10% MeOH/DCM with no separation of product from impurities.

5.1.3 Synthesis for Chapter 3

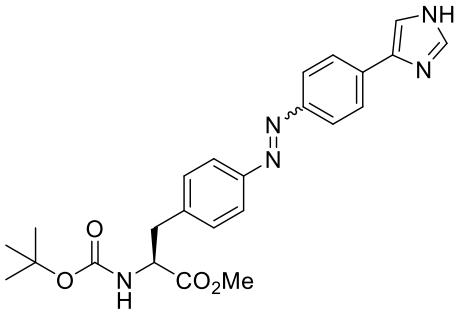
(*S*)-methyl 3-(4-aminophenyl)-2-[(*tert*-butoxycarbonyl)amino]propanoate **3.02**



To a stirring solution of *N*-Boc-4-amino-L-phenylalanine (1.00 g, 3.57 mmol) in methanol (10 mL) at -10 °C was added thionyl chloride (0.26 mL, 3.57 mmol) dropwise. The mixture was stirred for 3 h whilst warming to room temperature. H_2O (20 mL) was added to quench the reaction and the mixture was extracted with ethyl acetate (2 x 20 mL). To the aqueous layer was added 2 M aq. NaOH until it reached pH 5. The aqueous layer was again extracted with ethyl acetate (2 x 20 mL). The combined organic layers were dried over $MgSO_4$, filtered and concentrated in vacuo to give compound **3.02** as a yellow oil (0.95 g, 90%). ^{1}H NMR (500

MHz, Chloroform-*d*) δ 6.93 – 6.87 (m, 2H), 6.64 – 6.58 (m, 2H), 4.93 (d, *J* = 8.2 Hz, 1H), 4.53 – 4.48 (m, 1H), 3.70 (s, 3H), 3.60 (s, 2H), 2.97 (t, *J* = 5.9 Hz, 2H), 1.42 (s, 9H). ^{13}C NMR (126 MHz, Chloroform-*d*) δ 172.7, 145.5, 130.3, 125.9, 115.4, 54.8, 52.2, 37.6, 28.5.

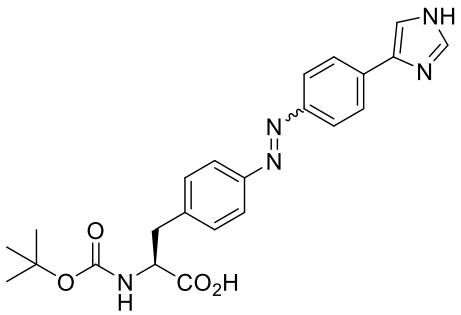
(*S*,*E*/*Z*)-methyl 3-(4-{[4-(1*H*-imidazol-4-yl)phenyl]diazenyl}phenyl)-2-[(*tert*-butoxycarbonyl)amino]propanoate **3.03**



To a solution of *N*-Boc-4-amino-L-phenylalanine methyl ester **3.02** (0.95 g, 3.22 mmol) in DCM (50 mL) was added a solution of Oxone® (1.98 g, 6.44 mmol) in H_2O (50 mL) and the mixture was stirred vigorously for 3 h at room temperature. The mixture was washed with 1 M aq. HCl, dried over $MgSO_4$, filtered and concentrated *in vacuo* to give the crude nitroso as a green solid (500 mg) which was used without further purification. The crude nitroso was dissolved in glacial acetic acid (3 mL) and 4-(1H-imidazol-4-yl)aniline (127 mg, 0.81 mmol) was added. The mixture was stirred overnight at room temperature. The mixture was diluted with water (30 mL) and brine (10 mL) and extracted with DCM (3 x 20 mL). The combined organic layers were washed with 0.1 M HCl (3 x 30 mL) and water (3 x 30 mL), dried over $MgSO_4$, filtered and concentrated *in vacuo* to give the crude product which was purified by flash chromatography (silica, EtOAc) to give compound **3.03** (77:23 *trans*:*cis*) as an orange solid (245 mg, 17% over 2 steps). R_f (EtOAc) = 0.18. ^{1}H NMR (500 MHz, Methanol-d_4) δ 7.96 – 7.88 (m, 4H), 7.85 (d, *J* = 8.1 Hz, 2H), 7.79 (s, 1H), 7.60 (s, 1H), 7.40 (d, *J* = 8.1 Hz, 2H), 4.45 (dd, *J* = 9.0, 5.6 Hz, 1H), 3.72 (s, 3H), 3.21 (dd, *J* = 13.8, 5.6 Hz, 1H), 3.01 (dd, *J* = 13.8, 9.1

Hz, 1H), 1.41 – 1.34 (m, 9H). ^{13}C NMR (126 MHz, Methanol-d_4) δ 174.0, 152.7, 142.0, 137.7, 131.2, 126.4, 124.4, 123.8, 56.4, 52.7, 28.7.

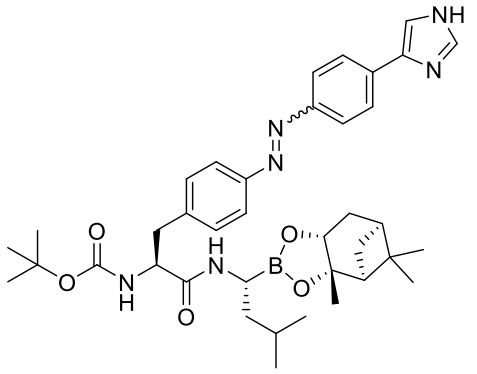
(*S*,*E*/*Z*)-3-(4-{[4-(1*H*-imidazol-4-yl)phenyl]diazenyl}phenyl)-2-[(*tert*-butoxycarbonyl)amino]propanoic acid **3.04**



Successful method: Methyl ester **3.03** (240 mg, 0.534 mmol) was hydrolysed according to general procedure B to give compound **3.04** as an orange solid (60 mg, 26%). ^{1}H NMR (500 MHz, Methanol-d_4) δ 8.34 (s, 1H), 7.98 – 7.94 (m, 2H), 7.90 (dt, *J* = 8.6, 2.2 Hz, 2H), 7.84 (d, *J* = 8.0 Hz, 2H), 7.79 (s, 1H), 7.42 (d, *J* = 8.0 Hz, 2H), 4.38 (dd, *J* = 9.0, 5.1 Hz, 1H), 3.24 (dd, *J* = 13.6, 4.9 Hz, 1H), 2.99 (dd, *J* = 13.9, 9.3 Hz, 1H), 1.38 – 1.29 (m, 9H).

Unsuccessful method: To a solution of *N*-Boc-4-amino-L-phenylalanine (1.00 g, 3.57 mmol) in DCM (50 mL) was added a solution of Oxone® (2.19 g, 7.14 mmol) in H_2O (50 mL) and the mixture was stirred vigorously for 3 h at room temperature. The mixture was washed with 1 M aq. HCl, dried over $MgSO_4$, filtered and concentrated *in vacuo* to give the crude nitroso as a green solid (0.700 g) which was used without further purification. The crude nitroso (700 mg) was dissolved in glacial acetic acid (5 mL) and 4-(1H-imidazol-4-yl)aniline (187 mg, 1.19 mmol) was added. The mixture was stirred overnight at room temperature. The mixture was diluted with water (30 mL) and brine (10 mL) and extracted with DCM (3 x 20 mL). The combined organic layers were washed with 0.1 M HCl (3 x 30 mL) and water (3 x 30 mL), dried over $MgSO_4$, filtered and concentrated *in vacuo* to give the crude product which contained less than 1% desired product.

tert-butyl [(*S*)-3-(4-{(*E*/*Z*)-[4-(1*H*-imidazol-4-yl)phenyl]diazenyl}phenyl)-1-({(*R*)-3-methyl-1-[(3a*S*,4*S*,6*S*,7a*R*)-3a,5,5-trimethylhexahydro-4,6-methanobenzo[d][1,3,2]dioxaborol-2-yl]butyl}amino)-1-oxopropan-2-yl]carbamate **3.01**



Carboxylic acid **3.03** (40 mg, 0.092 mmol) was coupled with (*R*)-Boroleucine (1*S*,2*S*,3*R*,5*S*)-(+)-2,3-pinanediol ester trifluoroacetate (32 mg, 0.084 mmol) according to general procedure A using HBTU, with the residue purified by flash chromatography (silica, 10% MeOH/DCM) to give compound **3.01** as a white solid (42 mg, 74%). R_f (10% MeOH/DCM) = 0.47. Analytical HPLC: 40-100% aq. acetonitrile over 20 min, R_t = 9.9 min (*cis*), 11.1 (*trans*). ^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.95 – 7.86 (m, 4H), 7.83 (d, *J* = 8.3 Hz, 2H), 7.74 (s, 1H), 7.42 (s, 1H), 7.36 (d, *J* = 8.1 Hz, 2H), 6.06 (s, 1H), 5.11 (s, 1H), 4.39 – 4.34 (m, 1H), 4.28 – 4.22 (m, 1H), 3.23 – 3.14 (m, 2H), 3.14 – 3.06 (m, 1H), 2.34 – 2.24 (m, 1H), 2.19 – 2.12 (m, 1H), 2.04 – 1.98 (m, 1H), 1.90 – 1.84 (m, 1H), 1.84 – 1.76 (m, 1H), 1.56 – 1.47 (m, 1H), 1.45 – 1.34 (m, 14H), 1.25 (s, 3H), 1.18 (d, *J* = 10.8 Hz, 1H), 0.90 – 0.83 (m, 6H), 0.80 (s, 3H). ^{13}C NMR (126 MHz, Chloroform-*d*) δ 171.6, 152.0, 151.6, 136.0, 130.3, 125.5, 123.6, 123.2, 86.0, 77.9, 69.4, 54.2, 51.6, 40.7, 40.3, 39.7, 38.5, 38.4, 38.3, 36.6, 35.7, 29.7, 28.8, 28.4, 28.4, 28.2, 28.0, 27.3, 27.2, 26.5, 25.5, 24.3, 24.1, 23.2, 22.1. HRMS (ESI+) calcd for $C_{38}H_{51}BN_6O_6$: 683.4092 [$(M+H)^+$], found 683.4111.

5.3 Enzyme Assays

5.1.1 Rabbit 20S Proteasome

The activity of rabbit 20S proteasome (R&D Systems, Inc., Minneapolis, MN) at a concentration of 2.88 ng/µL was assayed fluorometrically (Synergy H4 Hybrid Multi-Mode Microplate Reader, Bio-Tek Instruments, Inc.) on a 96-well black bottom plate in triplicate using the fluorogenic substrates (Sapphire Bioscience Pty. Ltd., Redfern, NSW) Ac-nLe-Pro-nLe-Asp-amc (β1), Bz-Val-Gly-Arg-amc (β2) and Suc-Leu-Leu-Val-Tyr-amc (β5) at 50 µM. Assay buffer was 20 mM Tris-HCl pH 7.5. A solution of rabbit 20S proteasome in 50 mM pH 7.6 HEPES buffer, 150 mM NaCl and 1 mM DTT was diluted freshly with assay buffer to give a concentration of 16 ng/µL and kept at 0 °C until used. Stock solutions of the substrates in DMSO at 10 mM were diluted with assay buffer to a concentration of 62.5 µM. Stock solutions of the inhibitors at 500 µM were serially diluted into a 96-well PCR plate with DMSO to give 7 concentrations of each compound (25 pM – 25 µM). 10 µL of each inhibitor concentration was added to separate wells of a 96-well PCR plate containing 90 µL of proteasome (16 ng/µL) and incubated at room temperature for 15 min. 10 µL of each inhibitor incubated with the proteasome was added (in triplicate) to a 96-well black bottom plate containing 40 µL of substrate (62.5 µM) and incubated at 37 °C for 2h concealed from light using aluminium foil. Fluorescence from the cleaved AMC was measured at λ_{ex} = 390 nm and λ_{em} = 460 nm. The fluorescence values obtained were thus proportional to enzyme activity. Baseline fluorescence was subtracted from each data set as measured from the average of three wells containing 50 µM of the respective substrate. GraphPad Prism 8 was used for the determination of the mean IC_{50} and standard error by fitting the inhibition curve to the built-in model 'log(inhibitor) vs. response -- Variable slope (four parameters)'.

For the photostationary state in the photoswitch assay, the 7 concentrations of each compound (25 pM – 25 µM) in DMSO and DMSO were irradiated in a 96-well black bottom plate for 1.5 h before being added to the PCR plate containing the enzyme.

5.1.2 Bovine α-Chymotrypsin

The activity of bovine α-chymotrypsin (Sigma Aldrich, Castle Hill, NSW) at a concentration of 11 ng/mL was assayed fluorometrically (Synergy H4 Hybrid Multi-Mode Microplate Reader, Bio-Tek Instruments, Inc.) at 25.0 ± 0.1 °C in a 96-well black bottom plate using the fluorogenic substrate Ala-Ala-Phe-amc (Sigma Aldrich, Castle Hill, NSW) at 500 µM. Assay buffer was 50 mM TES pH 8.0. Bovine α-chymotrypsin was dissolved in 1 mM aq. HCl to give a concentration of 874 µg/mL. The enzyme solution used in the assay was prepared by a 1:40 dilution, followed by a 1:100 dilution of the 874 µg/mL solution of bovine α-chymotrypsin, to give a final enzyme concentration of 0.22 µg/mL. The substrate was dissolved in DMSO to give a stock solution with a concentration of 20 mM. Each inhibitor was dissolved in DMSO to give a stock solution with a concentration of 10 mM and 6 serial dilutions were made on a 96-well PCR plate to give solutions with concentration (0.64 µM – 10 mM). 2.5 µL of substrate solution and 5 µL of each inhibitor concentration was added to 87.5 µL of assay buffer in a 96-well black bottom plate. 5 µL of bovine chymotrypsin solution (0.22 µg/mL) was added to give a final enzyme concentration of 11 ng/mL. Fluorescence from the cleaved AMC was measured at λ_{ex} = 390 nm and λ_{em} = 460 nm continuously over 10 min. The slope of each fluorescence trace is thus proportional to the activity of the enzyme. GraphPad Prism 8 was used for the determination of the mean IC_{50} and standard error by fitting the inhibition curve to the built-in model ‘log(inhibitor) vs. response -- Variable slope (four parameters)’.

References

1. Arora, P.; Ricks, T. K.; Trejo, J. *J. Cell Sci.* **2007,** *120*, 921.

2. Soh, U. J. K.; Dores, M. R.; Chen, B.; Trejo, J. *Br. J. Pharmacol.* **2010,** *160*, 191-203.

3. Turk, B.; Turk, D.; Turk, V. *EMBO J.* **2012,** *31*, 1630-1643.

4. Zetter, B. R.; Chen, L. B.; Buchanan, J. M. *Cell* **1976,** *7*, 407-412.

5. Yadaiah, M.; Sudhamalla, B.; Rao, P. N.; Roy, K. R.; Ramakrishna, D.; Hussain Syed, G.; Ramaiah, K. V. A.; Bhuyan, A. K. *FASEB J.* **2012,** *27*, 803-810.

6. Wells, J. M.; Strickland, S. *Development* **1994,** *120*, 3639.

7. Zeeuwen, P. L. J. M. *Eur. J. Cell Biol.* **2004,** *83*, 761-773.

8. Solary, E.; Eymin, B.; Droin, N.; Haugg, M. *Cell Biol. Toxicol.* **1998,** *14*, 121-132.

9. Stoka, V.; Turk, B.; Schendel, S. L.; Kim, T.-H.; Cirman, T.; Snipas, S. J.; Ellerby, L. M.; Bredesen, D.; Freeze, H.; Abrahamson, M.; Brömme, D.; Krajewski, S.; Reed, J. C.; Yin, X.-M.; Turk, V.; Salvesen, G. S. *J. Biol. Chem.* **2001,** *276*, 3149-3157.

10. Turk, B. *Nat. Rev. Drug Discovery* **2006,** *5*, 785-799.

11. Schechter, I.; Berger, A. *Biochem. Biophys. Res. Commun.* **1967,** *27*, 157-162.

12. Czapinska, H.; Otlewski, J. *Eur. J. Biochem.* **1999,** *260*, 571-595.

13. Stroud, R. M.; Kay, L. M.; Dickerson, R. E. *J. Mol. Biol.* **1974,** *83*, 185-208.

14. Schellenberger, V.; Braune, K.; Hofmann, H.-J.; Jakubke, H.-D. *Eur. J. Biochem.* **1991,** *199*, 623-636.

15. Carrington, J. C.; Dougherty, W. G. *Proc. Natl. Acad. Sci. U. S. A.* **1988,** *85*, 3391.

16. Drag, M.; Salvesen, G. S. *Nat. Rev. Drug Discovery* **2010,** *9*, 690-701.

17. Liu, B.; Schofield, C. J.; Wilmouth, R. C. *J. Biol. Chem.* **2006,** *281*, 24024-24035.

18. Rock, K. L.; Gramm, C.; Rothstein, L.; Clark, K.; Stein, R.; Dick, L.; Hwang, D.; Goldberg, A. L. *Cell* **1994,** *78*, 761-771.

19. Schmidt, M.; Finley, D. *Biochim. Biophys. Acta* **2014,** *1843*, 13-25.

20. Fehling, H. J.; Swat, W.; Laplace, C.; Kuhn, R.; Rajewsky, K.; Muller, U.; von Boehmer, H. *Science* **1994,** *265*, 1234.

21. Hensley, S. E.; Zanker, D.; Dolan, B. P.; David, A.; Hickman, H. D.; Embry, A. C.; Skon, C. N.; Grebe, K. M.; Griffin, T. A.; Chen, W.; Bennink, J. R.; Yewdell, J. W. *J. Immunol.* **2010,** *184*, 4115.

22. Pagano, M.; Tam, S. W.; Theodoras, A. M.; Beer-Romero, P.; Del Sal, G.; Chau, V.; Yew, P. R.; Draetta, G. F.; Rolfe, M. *Science* **1995,** *269*, 682.

23. Bassermann, F.; Eichner, R.; Pagano, M. *Biochim. Biophys. Acta* **2014,** *1843*, 150-162.

24. Auld, K. L.; Silver, P. A. *Cell Cycle* **2006,** *5*, 1503-1505.

25. Orlowski, R. Z. *Cell Death Differ.* **1999,** *6*, 303-313.

26. Sohn, D.; Totzke, G.; Essmann, F.; Schulze-Osthoff, K.; Levkau, B.; Jänicke, R. U. *Mol. Cell. Biol.* **2006,** *26*, 1967.

27. Thibaudeau, T. A.; Smith, D. M. *Pharmacol. Rev.* **2019,** *71*, 170.

28. Tanaka, K. *Proc. Jpn. Acad. Ser. B Phys. Biol. Sci.* **2009,** *85*, 12-36.

29. Arendt, C. S.; Hochstrasser, M. *Proc. Natl. Acad. Sci. U. S. A.* **1997,** *94*, 7156.

30. Schweitzer, A.; Aufderheide, A.; Rudack, T.; Beck, F.; Pfeifer, G.; Plitzko, J. M.; Sakata, E.; Schulten, K.; Förster, F.; Baumeister, W. *Proc. Natl. Acad. Sci. U. S. A.* **2016,** *113*, 7816.

31. Heinemeyer, W.; Fischer, M.; Krimmer, T.; Stachon, U.; Wolf, D. H. *J. Biol. Chem.* **1997,** *272*, 25200-25209.

32. Huber, E. M.; Heinemeyer, W.; Li, X.; Arendt, C. S.; Hochstrasser, M.; Groll, M. *Nat. Commun.* **2016,** *7*, 10900.

33. Beck, P.; Dubiella, C.; Groll, M. *Biol. Chem.* **2012,** *393*, 1101-1120.

34. Hershko, A.; Heller, H.; Elias, S.; Ciechanover, A. *J. Biol. Chem.* **1983,** *258*, 8206-8214.

35. Kisselev, A. F.; Akopian, T. N.; Woo, K. M.; Goldberg, A. L. *J. Biol. Chem.* **1999,** *274*, 3363-3371.

36. Sijts, E. J. A. M.; Kloetzel, P. M. *Cell. Mol. Life Sci.* **2011,** *68*, 1491-1502.

37. Ciechanover, A.; Kwon, Y. T. *Front. Neurosci.* **2017,** *11*, 185.

38. Gilda, J. E.; Gomes, A. V. *J. Physiol.* **2017,** *595*, 4051-4071.

39. Traenckner, E. B.; Wilk, S.; Baeuerle, P. A. *EMBO J.* **1994,** *13*, 5433-5441.

40. Monaco, C.; Andreakos, E.; Kiriakidis, S.; Mauri, C.; Bicknell, C.; Foxwell, B.; Cheshire, N.; Paleolog, E.; Feldmann, M. *Proc. Natl. Acad. Sci. U. S. A.* **2004,** *101*, 5634.

41. Sun, S.-C.; Chang, J.-H.; Jin, J. *Trends Immunol.* **2013,** *34*, 282-289.

42. Chen, D.; Dou, Q. P. *Curr. Protein Pept. Sci.* **2010,** *11*, 459-470.

43. Wilk, S.; Orlowski, M. *J. Neurochem.* **1983,** *40*, 842-849.

44. Dou, Q. P.; Smith, D. M.; Daniel, K. G.; Kazi, A. *Prog. Cell Cycle Res.* **2003,** *5*, 441-446.

45. Vinitsky, A.; Cardozo, C.; Sepp-Lorenzino, L.; Michaud, C.; Orlowski, M. *J. Biol. Chem.* **1994,** *269*, 29860-29866.

46. Masdehors, P.; Omura, S.; Merle-Béral, H.; Mentz, F.; Cosset, J.-M.; Dumont, J.; Magdelénat, H.; Delic, J. *Br. J. Haematol.* **1999,** *105*, 752-757.

47. Adams, J.; Behnke, M.; Chen, S.; Cruickshank, A. A.; Dick, L. R.; Grenier, L.; Klunder, J. M.; Ma, Y.-T.; Plamondon, L.; Stein, R. L. *Bioorg. Med. Chem. Lett.* **1998,** *8*, 333-338.

48. Adams, J.; Kauffman, M. *Cancer Investig.* **2004,** *22*, 304-311.

49. Orlowski, R. Z.; Stinchcombe, T. E.; Mitchell, B. S.; Shea, T. C.; Baldwin, A. S.; Stahl, S.; Adams, J.; Esseltine, D.-L.; Elliott, P. J.; Pien, C. S.; Guerciolini, R.; Anderson, J. K.; Depcik-Smith, N. D.; Bhagat, R.; Lehman, M. J.; Novick, S. C.; O'Connor, O. A.; Soignet, S. L. *J. Clin. Oncol.* **2002,** *20*, 4420-4427.

50. Richardson, P. G.; Barlogie, B.; Berenson, J.; Singhal, S.; Jagannath, S.; Irwin, D.; Rajkumar, S. V.; Srkalovic, G.; Alsina, M.; Alexanian, R.; Siegel, D.; Orlowski, R. Z.; Kuter, D.; Limentani, S. A.; Lee, S.; Hideshima, T.; Esseltine, D.-L.; Kauffman, M.; Adams, J.; Schenkein, D. P.; Anderson, K. C. *N. Engl. J. Med.* **2003,** *348*, 2609-2617.

51. Kane, R. C.; Bross, P. F.; Farrell, A. T.; Pazdur, R. *Oncologist* **2003,** *8*, 508-513.

52. Raedler, L. *Am. Health Drug Benefits* **2015,** *8*, 135-140.

53. Kumar, S.; Rajkumar, S. V. *Blood* **2008,** *112*, 2177.

54. Tan, C. R. C.; Abdul-Majeed, S.; Cael, B.; Barta, S. K. *Clin. Pharmacokinet.* **2019,** *58*, 157-168.

55. Schwartz, R.; Davidson, T. *Oncology* **2005,** *18*, 14-21.

56. Parlati, F.; Lee, S. J.; Aujay, M.; Suzuki, E.; Levitsky, K.; Lorens, J. B.; Micklem, D. R.; Ruurs, P.; Sylvain, C.; Lu, Y.; Shenk, K. D.; Bennett, M. K. *Blood* **2009,** *114*, 3439.

57. Mirabella, A. C.; Pletnev, A. A.; Downey, S. L.; Florea, B. I.; Shabaneh, T. B.; Britton, M.; Verdoes, M.; Filippov, D. V.; Overkleeft, H. S.; Kisselev, A. F. *Chem. Biol.* **2011,** *18*, 608-618.

58. Kisselev, A. F.; Goldberg, A. L. *Chem. Biol.* **2001,** *8*, 739-758.

59. Stein, M. L.; Cui, H.; Beck, P.; Dubiella, C.; Voss, C.; Krüger, A.; Schmidt, B.; Groll, M. *Angew. Chem., Int. Ed.* **2014,** *53*, 1679-1683.

60. Groll, M.; Berkers, C. R.; Ploegh, H. L.; Ovaa, H. *Structure* **2006,** *14*, 451-456.

61. Dou, Q. P.; Zonder, J. A. *Curr. Cancer Drug Targets* **2014,** *14*, 517-536.

62. Corso, A.; Mangiacavalli, S.; Varettoni, M.; Pascutto, C.; Zappasodi, P.; Lazzarino, M. *Leuk. Res.* **2010,** *34*, 471-474.

63. Millenium Pharmaceuticals. VELCADE® Prescribing Information. Apr 2019. Accessed 27 Sep 2019.

64. Oerlemans, R.; Franke, N. E.; Assaraf, Y. G.; Cloos, J.; van Zantwijk, I.; Berkers, C. R.; Scheffer, G. L.; Debipersad, K.; Vojtekova, K.; Lemos, C.; van der Heijden, J. W.; Ylstra, B.; Peters, G. J.; Kaspers, G. L.; Dijkmans, B. A. C.; Scheper, R. J.; Jansen, G. *Blood* **2008,** *112*, 2489.

65. Lü, S.; Wang, J. *Biomark. Res.* **2013,** *1*, 13-13.

66. Zhu, Y. X.; Tiedemann, R.; Shi, C.-X.; Yin, H.; Schmidt, J. E.; Bruins, L. A.; Keats, J. J.; Braggio, E.; Sereduk, C.; Mousses, S.; Stewart, A. K. *Blood* **2011,** *117*, 3847-3857.

67. Muz, B.; Ghazarian, R. N.; Ou, M.; Luderer, M. J.; Kusdono, H. D.; Azab, A. K. *Drug Des. Devel. Ther.* **2016,** *10*, 217-226.

68. Millenium Pharmaceuticals. NINLARO(R) Prescribing Information. Nov 2016. Accessed 27 Sep 2019.

69. Mateos, M.-V.; Masszi, T.; Grzasko, N.; Hansson, M.; Sandhu, I.; Pour, L.; Viterbo, L.; Jackson, S. R.; Stoppa, A.-M.; Gimsing, P.; Hamadani, M.; Borsaru, G.; Berg, D.; Lin, J.; Di Bacco, A.; van de Velde, H.; Richardson, P. G.; Moreau, P. *Haematologica* **2017,** *102*, 1767-1775.

70. Hanada, M.; Sugawara, K.; Kaneta, K.; Toda, S.; Nishiyama, Y.; Tomita, K.; Yamamoto, H.; Konishi, M.; Oki, T. *J. Antibiot.* **1992,** *45*, 1746-52.

71. Meng, L.; Mohan, R.; Kwok, B. H.; Elofsson, M.; Sin, N.; Crews, C. M. *Proc. Natl. Acad. Sci. U. S. A.* **1999,** *96*, 10403-10408.

72. Groll, M.; Kim, K. B.; Kairies, N.; Huber, R.; Crews, C. M. *J. Am. Chem. Soc.* **2000,** *122*, 1237-1238.

73. Arastu-Kapur, S.; Anderl, J. L.; Kraus, M.; Parlati, F.; Shenk, K. D.; Lee, S. J.; Muchamuel, T.; Bennett, M. K.; Driessen, C.; Ball, A. J.; Kirk, C. J. *Clin. Cancer Res.* **2011,** *17*, 2734.

74. Amgen Inc. KYPROLIS® Prescribing Information. Feb 2019. Accessed 28 Sep 2019.

75. Omura, S.; Fujimoto, T.; Otoguro, K.; Matsuzaki, K.; Moriguchi, R.; Tanaka, H.; Sasaki, Y. *J. Antibiot.* **1991,** *44*, 113-6.

76. Craiu, A.; Gaczynska, M.; Akopian, T.; Gramm, C. F.; Fenteany, G.; Goldberg, A. L.; Rock, K. L. *J. Biol. Chem.* **1997,** *272*, 13437-13445.

77. Dick, L. R.; Cruikshank, A. A.; Destree, A. T.; Grenier, L.; McCormack, T. A.; Melandri, F. D.; Nunes, S. L.; Palombella, V. J.; Parent, L. A.; Plamondon, L.; Stein, R. L. *J. Biol. Chem.* **1997,** *272*, 182-188.

78. Fenteany, G.; Standaert, R. F.; Lane, W. S.; Choi, S.; Corey, E. J.; Schreiber, S. L. *Science* **1995,** *268*, 726-731.

79. Koguchi, Y.; Kohno, J.; Nishio, M.; Takahashi, K.; Okuda, T.; Ohnuki, T.; Komatsubara, S. *J. Antibiot.* **2000,** *53*, 105-9.

80. Gallastegui, N.; Beck, P.; Arciniega, M.; Huber, R.; Hillebrand, S.; Groll, M. *Angew. Chem., Int. Ed.* **2012,** *51*, 247-249.

81. Di Giovanni, C.; Ettari, R.; Sarno, S.; Rotondo, A.; Bitto, A.; Squadrito, F.; Altavilla, D.; Schirmeister, T.; Novellino, E.; Grasso, S.; Zappalà, M.; Lavecchia, A. *Eur. J. Med. Chem.* **2016,** *121*, 578-591.

82. Blackburn, C.; Gigstad, K. M.; Hales, P.; Garcia, K.; Jones, M.; Bruzzese, F. J.; Barrett, C.; Liu, J. X.; Soucy, T. A.; Sappal, D. S.; Bump, N.; Olhava, E. J.; Fleming, P.; Dick, L. R.; Tsu, C.; Sintchak, M. D.; Blank, J. L. *Biochem. J.* **2010,** *430*, 461.

83. Bogyo, M.; McMaster, J. S.; Gaczynska, M.; Tortorella, D.; Goldberg, A. L.; Ploegh, H. *Proc. Natl. Acad. Sci. U. S. A.* **1997,** *94*, 6629-6634.

84. Gräwert, M. A.; Gallastegui, N.; Stein, M.; Schmidt, B.; Kloetzel, P.-M.; Huber, R.; Groll, M. *Angew. Chem., Int. Ed.* **2011,** *50*, 542-544.

85. de Bruin, G.; Mock, E. D.; Hoogendoorn, S.; van den Nieuwendijk, A. M. C. H.; Mazurek, J.; van der Marel, G. A.; Florea, B. I.; Overkleeft, H. S. *Chem. Comm.* **2016,** *52*, 4064-4067.

86. Gräwert, M. A.; Groll, M. *Chem. Comm.* **2012,** *48*, 1364-1378.

87. Schrader, J.; Henneberg, F.; Mata, R. A.; Tittmann, K.; Schneider, T. R.; Stark, H.; Bourenkov, G.; Chari, A. *Science* **2016,** *353*, 594.

88. Asai, A.; Hasegawa, A.; Ochiai, K.; Yamashita, Y.; Mizukami, T. *J. Antibiot.* **2000,** *53*, 81-3.

89. Asai, A.; Tsujita, T.; Sharma, S. V.; Yamashita, Y.; Akinaga, S.; Funakoshi, M.; Kobayashi, H.; Mizukami, T. *Biochem. Pharmacol.* **2004,** *67*, 227-234.

90. Korotkov, V. S.; Ludwig, A.; Larionov, O. V.; Lygin, A. V.; Groll, M.; de Meijere, A. *Org. Biomol. Chem.* **2011,** *9*, 7791-7798.

91. Kawamura, S.; Unno, Y.; List, A.; Mizuno, A.; Tanaka, M.; Sasaki, T.; Arisawa, M.; Asai, A.; Groll, M.; Shuto, S. *J. Med. Chem.* **2013,** *56*, 3689-3700.

92. Kawamura, S.; Unno, Y.; Hirokawa, T.; Asai, A.; Arisawa, M.; Shuto, S. *Chem. Comm.* **2014,** *50*, 2445-2447.

93. Kawamura, S.; Unno, Y.; Asai, A.; Arisawa, M.; Shuto, S. *J. Med. Chem.* **2014,** *57*, 2726-2735.

94. Kawamura, S.; Unno, Y.; Asai, A.; Arisawa, M.; Shuto, S. *Org. Biomol. Chem.* **2013,** *11*, 6615-6622.

95. Zhu, M.; Harshbarger, W. D.; Robles, O.; Krysiak, J.; Hull, K. G.; Cho, S. W.; Richardson, R. D.; Yang, Y.; Garcia, A.; Spiegelman, L.; Ramirez, B.; Wilson, C. T.; Yau, J. A.; Moore, J. T.; Walker, C. B.; Sacchettini, J. C.; Liu, W. R.; Sieber, S. A.; Smith, J. W.; Romo, D. *Bioorg. Med. Chem.* **2017,** *25*, 2901-2916.

96. Nakamura, H.; Watanabe, M.; Ban, H. S.; Nabeyama, W.; Asai, A. *Bioorg. Med. Chem. Lett.* **2009,** *19*, 3220-3224.

97. Marastoni, M.; Scotti, A.; Trapella, C.; Ferretti, V.; Sforza, F.; Gavioli, R. *J. Pept. Sci.* **2014,** *20*, 258-265.

98. Marastoni, M.; Trapella, C.; Scotti, A.; Fantinati, A.; Ferretti, V.; Marzola, E.; Eleonora, G.; Gavioli, R.; Preti, D. *J. Enzyme Inhib. Med. Chem.* **2017,** *32*, 865-877.

99. Voss, C.; Scholz, C.; Knorr, S.; Beck, P.; Stein, M. L.; Zall, A.; Kuckelkorn, U.; Kloetzel, P.-M.; Groll, M.; Hamacher, K.; Schmidt, B. *ChemMedChem* **2014,** *9*, 2557-2564.

100. Krishna Deepak, R. N. V.; Sankararamakrishnan, R. *Biochemistry* **2016,** *55*, 3774-3783.

101. Mitsunobu, O.; Yamada, M. *Bull. Chem. Soc. Jpn.* **1967,** *40*, 2380-2382.

102. Manhas, M. S.; Hoffman, W. H.; Lal, B.; Bose, A. K. *J. Chem. Soc., Perkin Trans. 1* **1975**, 461-463.

103. Horatscheck, A.; Wagner, S.; Ortwein, J.; Kim, B. G.; Lisurek, M.; Beligny, S.; Schütz, A.; Rademann, J. *Angew. Chem., Int. Ed.* **2012,** *51*, 9441-9447.

104. Carpino, L. A. *J. Am. Chem. Soc.* **1993,** *115*, 4397-4398.

105. Diaz, D. B.; Yudin, A. K. *Nature Chem.* **2017,** *9*, 731.

106. Leung, I. K. H.; Brown Jr, T.; Schofield, C. J.; Claridge, T. D. W. *MedChemComm* **2011,** *2*, 390-395.

107. Kisselev, A. F.; Goldberg, A. L. Monitoring Activity and Inhibition of 26s Proteasomes with Fluorogenic Peptide Substrates. In *Methods Enzymol.*, Academic Press: 2005; Vol. 398, pp 364-378.

108. Mitscherlich, E. *Ann. Phys.* **1834,** *108*, 225-227.

109. Hartley, G. S. *Nature* **1937,** *140*, 281-281.

110. Masutani, K.; Morikawa, M.-a.; Kimizuka, N. *Chem. Comm.* **2014,** *50*, 15803-15806.

111. Merino, E.; Ribagorda, M. *Beilstein J. Org. Chem.* **2012,** *8*, 1071-1090.

112. Hugel, T.; Holland, N. B.; Cattani, A.; Moroder, L.; Seitz, M.; Gaub, H. E. *Science* **2002,** *296*, 1103.

113. Ichimura, K.; Oh, S.-K.; Nakagawa, M. *Science* **2000,** *288*, 1624.

114. Yamada, S.; Bessho, J.; Nakasato, H.; Tsutsumi, O. *Dyes Pigm.* **2018,** *150*, 89-96.

115. Evangelio, E.; Saiz-Poseu, J.; Maspoch, D.; Wurst, K.; Busque, F.; Ruiz-Molina, D. *Eur. J. Inorg. Chem.* **2008,** *2008*, 2278-2285.

116. Kim, Y.; Phillips, J. A.; Liu, H.; Kang, H.; Tan, W. *Proc. Natl. Acad. Sci. U. S. A.* **2009,** *106*, 6489.

117. Cheng, B.; Shchepakin, D.; Kavanaugh, M. P.; Trauner, D. *ACS Chem. Neurosci.* **2017,** *8*, 1668-1672.

118. Hansen, M. J.; Velema, W. A.; de Bruin, G.; Overkleeft, H. S.; Szymanski, W.; Feringa, B. L. *ChemBioChem* **2014,** *15*, 2053-2057.

119. Blanco, B.; Palasis, K. A.; Adwal, A.; Callen, D. F.; Abell, A. D. *Bioorg. Med. Chem.* **2017,** *25*, 5050-5054.

120. Ueno, K.; Akiyoshi, S. *J. Am. Chem. Soc.* **1954,** *76*, 3670-3672.

121. Marturano, V.; Ambrogi, V.; Bandeira, N. A. G.; Tylkowski, B.; Giamberini, M.; Cerruti, P. Modeling of Azobenzene-Based Compounds. In *Physical Sciences Reviews*, 2017; Vol. 2.

122. Martins, L. M.; Morrison, A.; Klupsch, K.; Fedele, V.; Moisoi, N.; Teismann, P.; Abuin, A.; Grau, E.; Geppert, M.; Livi, G. P.; Creasy, C. L.; Martin, A.; Hargreaves, I.;

Heales, S. J.; Okada, H.; Brandner, S.; Schulz, J. B.; Mak, T.; Downward, J. *Mol. Cell. Biol.* **2004,** *24*, 9848.

123. Gottlieb, H. E.; Kotlyar, V.; Nudelman, A. *J. Org. Chem.* **1997,** *62*, 7512-7515.

Appendices

Appendix A: HPLC Traces for Final Compounds

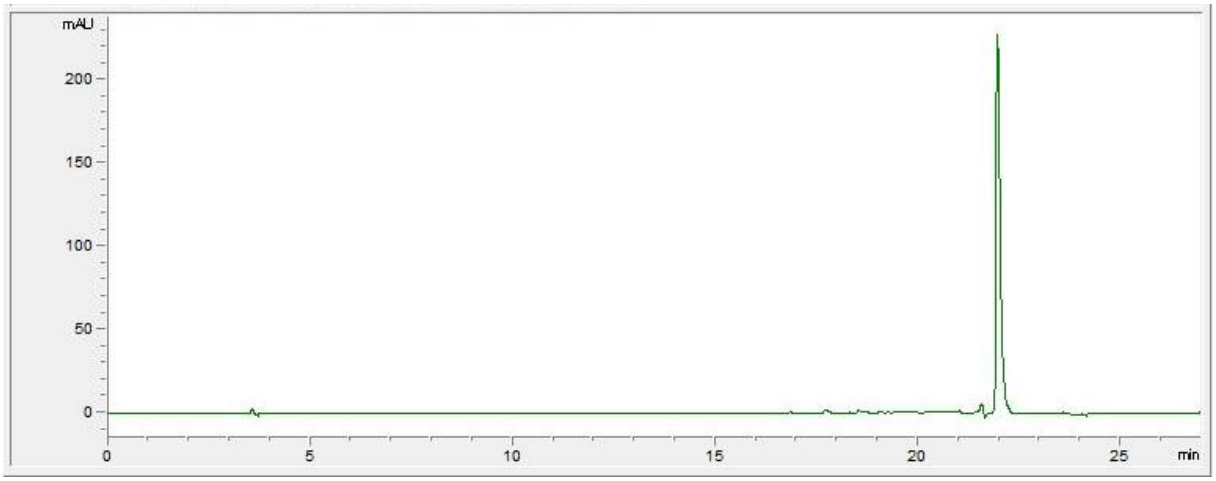


Figure A1 Analytical RP-HPLC C18 spectrum for **2.01**, 20-100% acetonitrile/H_2O over 20 min (5-25 min) visualised at 254 nm.

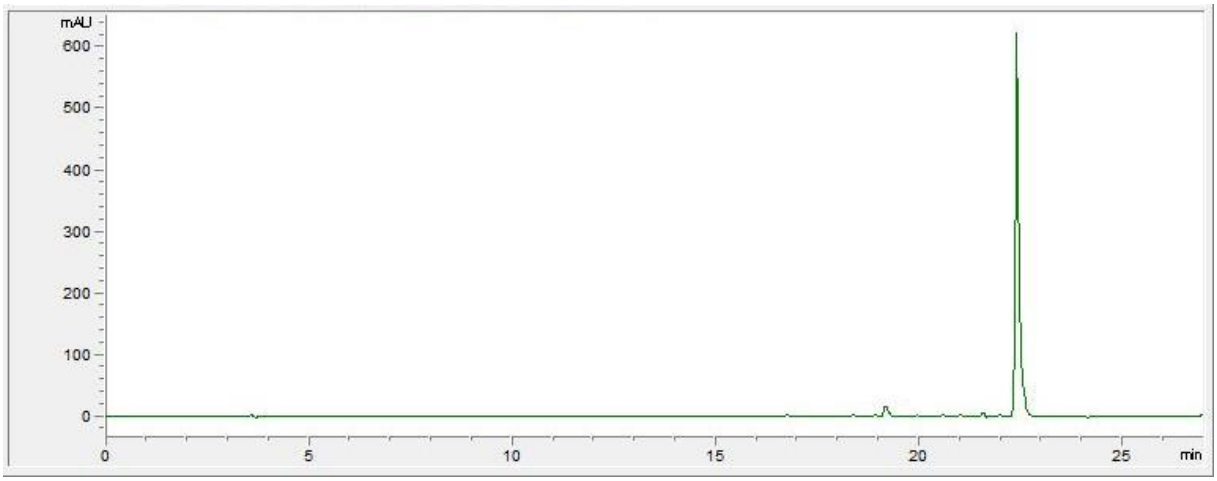


Figure A2 Analytical RP-HPLC C18 spectrum for **2.02**, 20-100% acetonitrile/H_2O over 20 min (5-25 min) visualised at 254 nm.

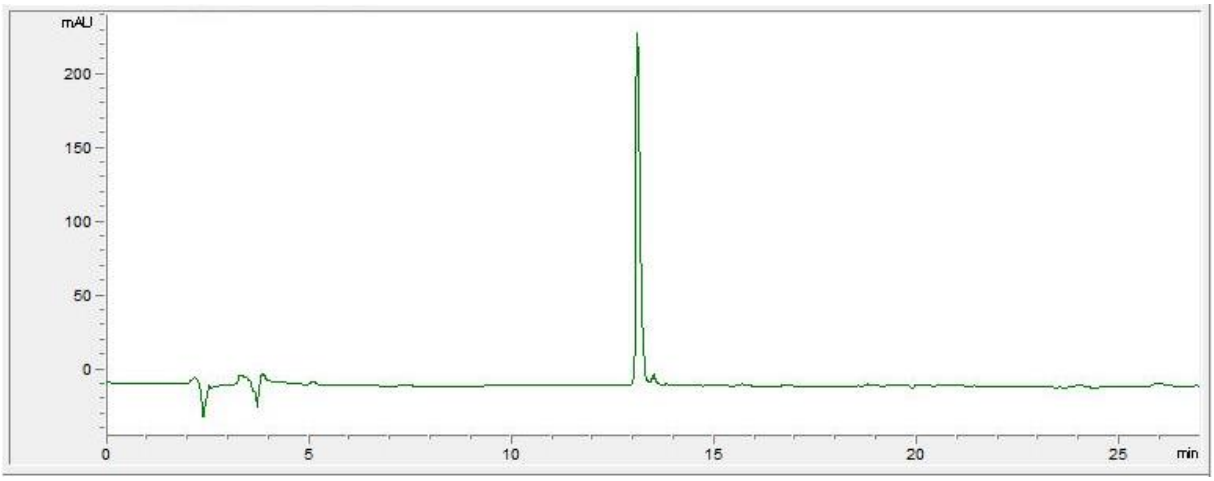


Figure A3 Analytical RP-HPLC C18 spectrum for **2.03**, 40-100% acetonitrile/H_2O over 20 min (5-25 min) visualised at 254 nm.

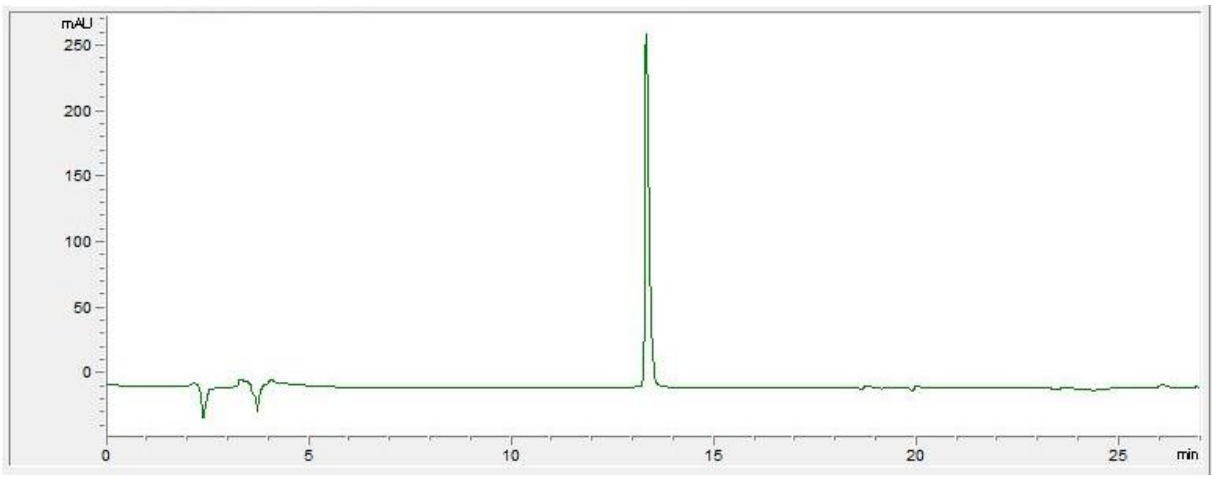


Figure A4 Analytical RP-HPLC C18 spectrum for **2.04**, 40-100% acetonitrile/H_2O over 20 min (5-25 min) visualised at 254 nm.

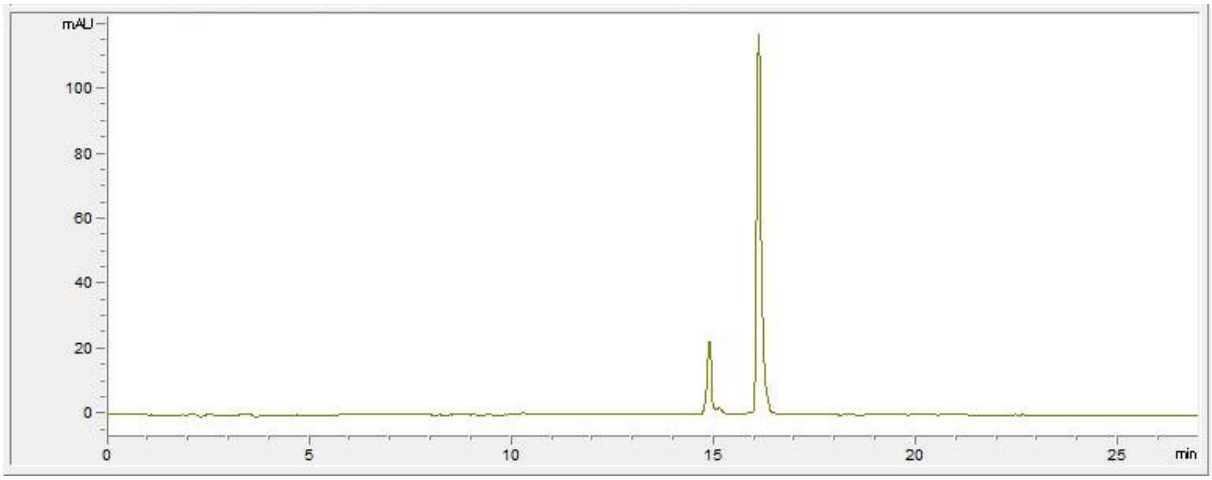


Figure A5 Analytical RP-HPLC C18 spectrum for **3.05**, 40-100% acetonitrile/H_2O over 20 min (5-25 min) visualised at 320 nm.