

**X-RAY CRYSTAL STRUCTURE OF HUMAN 20S PROTEASOME IN
COMPLEX WITH CARFILZOMIB**

A Dissertation

by

WAYNE DANIEL HARSHBARGER

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	James Sacchettini
Committee Members,	David Barondeou
	Daniel Romo
	Tatyana Igumenova
Head of Department,	David Russell

May 2015

Major Subject: Chemistry

Copyright 2015 Wayne Harshbarger

ABSTRACT

20S proteasomes are large, multicatalytic N-terminal threonine proteases which are tasked with maintaining intracellular homeostasis by the breaking down of mis-folded, oxidized, or tagged proteins into small peptides. Proteasomes are essential for cell division and differentiation, and proteasome inhibition represents a successful therapeutic strategy against cancers. There are two main isoforms of the proteasome, the constitutive proteasome and the immunoproteasome. While the constitutive proteasome is emerging as a medical rationale for treatment of solid tumors, the immunoproteasome shows promise in the treatment of autoimmune diseases such as rheumatoid arthritis. Inhibition of both proteasome forms is necessary for the treatment of multiple myeloma, and is accomplished by the FDA approved drugs, bortezomib and carfilzomib. While bortezomib equally targets each isoform, carfilzomib has nearly six-fold selectivity for the constitutive proteasome. Despite the importance of the proteasome in cancer treatment as well as carfilzomib's use in the clinic, there remains no crystal structure of the human 20S proteasome, and no crystal structure of a proteasome:carfilzomib complex. This hampers drug development efforts charged with designing the next generation proteasome inhibitors.

The present dissertation describes the structural and functional characterization of the constitutive proteasome from human, which will facilitate structure-based development of novel drugs. Using X-ray crystallography, the 3D structures of the human proteasome in an unbound and carfilzomib bound state have been determined to high

resolution. The ligand bound structure provides explanations for the chymotrypsin-like selectivity of the drug, as well as to its preferential binding to constitutive proteasomes over immunoproteasome. The caspase-like sites have closed off S3 binding pockets which sterically hamper binding of carfilzomib's P3 phenyl group. Trypsin-like sites have very similar specificity pocket characteristics as chymotrypsin-like sites, however the S1 pockets are spacious in comparison, and do not provide as many interactions to carfilzomib's P1 leucyl group, which may be reflected in the decreased selectivity. Chymotrypsin-like subunits have hydrophobic specificity pockets which form van der Waals interactions with carfilzomib that favor ligand binding. In particular, the P3 and P4 positions are found to be important for overcoming the structural rearrangements of Met45, α -helix H1, and β -sheets S5 and S6 upon carfilzomib binding.

Additionally in this dissertation, high throughput screening techniques were used to identify novel proteasome inhibitors. A natural product polyphenol, PPH-1, was identified from microbial extracts, and a second natural product, α -mangostin, was identified from the National Cancer Institute's library of clinically active compounds. The binding of each to the proteasome's catalytic subunits was confirmed using Western Blotting techniques. Finally, the kinetic characterization of a series of compounds with dual proteasome and fatty acid synthase are described. The inhibition of multiple enzymes by a single drug may decrease the likelihood that diseases such as cancer will develop resistance, due to the need to become resistant along multiple pathways. The structural and functional characterization of the human proteasome described herein, now facilitates the design of novel drugs with therapeutic potential in cancer.

DEDICATION

For my daughter and inspiration, Kayli, who spent many of her precious summers in the heat of Texas helping me with research. For my parents, Cynthia and John Harshbarger, who have been there for me and my siblings through the toughest of times, and make every sacrifice to see their children succeed. For my brothers Elijah and James, and my sister Santana. They define what family means and I am grateful for all their love and support.

ACKNOWLEDGEMENTS

I would like to thank my advisor, Dr. James Sacchetti, for all his guidance and support through my graduate career. I'd like to thank my committee member, Dr. Daniel Romo, Dr. David Barondeau, and Dr. Tatyana Igumenova, for all their advice through the course of my research and encouragement.

Thanks to all my lab members and friends for their constant support both in and out of the lab and for making my time at Texas A&M University a memorable experience. Lastly, I am thankful for to the chemistry department staff, especially Sandy Manning, for all their administration help.

TABLE OF CONTENTS

	Page
ABSTRACT	ii
DEDICATION	iv
ACKNOWLEDGEMENTS	v
TABLE OF CONTENTS	vi
LIST OF FIGURES	viii
LIST OF TABLES	xv
1. INTRODUCTION.....	1
1.1 Protein Degradation.....	1
1.2 Ubiquitin Proteasome System	4
1.3 Proteasomes Assembly and Architecture	6
1.4 Proteasome Proteolysis.....	11
1.5 Immune Functions of the Proteasome	14
1.6 20S Proteasome: Inhibition and Apoptosis	16
1.7 Classes of Proteasome Inhibitors	17
1.8 Clinically Relevant Proteasome Inhibitors.....	20
1.9 Proteasome Inhibitors of Individual Active Subunits	22
1.10 Basic Crystallographic Theory	24
1.10.1 Introduction to Crystallography	24
1.10.2 Phase Problem and Molecular Replacement.....	27
1.10.3 Model Building and Refinement	29
2. STRUCTURAL STUDIES OF THE HUMAN 20S PROTEASOME	31
2.1 Results and Discussion.....	31
2.1.1 Sequence Alignments.....	31
2.1.2 Purification, Crystallization and Structure Determination	36
2.1.3 Architecture of the Human 20S Proteasome Constitutive Core Particle.....	43
2.1.4 β Subunit Post-Translational Processing and Catalytic Residues	50
2.1.5 Cleavage Preferences of Catalytic Subunits.....	52
2.2 Materials and Methods	53
2.2.1 Proteasome Purification	53
2.2.2 Crystallization of Human 20S Proteasome Constitutive Core Particle.....	55
2.2.3 Data Collection, Processing and Structure Determination.....	56

2.2.4 Bioinformatics Tools.....	57
3. CARFILZOMIB BOUND HUMAN 20S PROTEASOME.....	58
3.1 Introduction	58
3.2 Results and Discussion.....	60
3.2.1 Carfilzomib Binding.....	60
3.2.2 Interactions of Carfilzomib with Proteasome Specificity Pockets.....	69
3.2.3 PR-957 Selectivity for Immunoproteasome	73
3.2.4 Predictions of Carfilzomib Interactions with the Immunoproteasome	76
3.3 Conclusion.....	82
3.4 Materials and Methods	85
4. DISCOVERY OF NOVEL PROTEASOME INHIBITORS; HIGH THROUGHPUT SCREENING	88
4.1 Introduction	88
4.2 Results and Discussion.....	90
4.2.1 National Institute of Health Clinical Collection (NCC).....	90
4.2.2 Fungus Microbial Extracts; <i>Zygothiala jamaicensis</i>	98
4.2.3 Lymphoma Hit Plate Screen.....	102
4.3 Conclusions	106
4.4 Materials and Methods	109
4.4.1 Proteasome Fluorogenic Assay for Determination of IC ₅₀ 's.....	109
4.4.2 High Through-put Screening for Novel Proteasome Inhibitors	110
4.4.3 In vitro Binding Assay	112
4.4.4 Cell Based Binding Assay	114
4.4.5 Cell Based Toxicity Assays.....	116
5. DUAL PROTEASOME/FATTY ACID SYNTHASE INHIBITORS.....	117
5.1 Introduction	117
5.2 Results and Discussion.....	124
5.3 Conclusions	129
5.4 Materials and Methods	130
6. CONCLUSIONS	132
REFERENCES.....	135

LIST OF FIGURES

Page

- Figure 1 Ubiquitin proteasome pathway. Adenylation of ubiquitin by an E1 enzyme is followed by its transfer to an E2 ubiquitin-conjugating enzyme. An E3 ligase covalently attaches ubiquitin to a lysine residue in the target protein. The ubiquitin chain is extended by E3, and substrates with at least five ubiquitin molecules are degraded by the 26S proteasome. Adapted from [17].....5
- Figure 2 Comparison of proteasome core particles from archaea (*T. acidophilum*), eukaryotes (*S. cerevisiae*), and vertebrates (*M. musculus*). Structures are depicted in surface representation. The α -subunits are shown in red, catalytic β -subunits in purple, and non-catalytic β -subunits are blue. All β -subunits of *T. acidophilum* are catalytic, whereas six of the fourteen β -subunits from eukaryotes and vertebrates are catalytic, with three distinct cleavage specificities; chymotrypsin-like, trypsin-like, and caspase-like.8
- Figure 3 Proteasome core particle isoforms. Inactive subunits are colored grey. Caspase-like sites ($\beta 1$) are colored green; Trypsin-like sites ($\beta 2$) are colored orange; Chymotrypsin-like sites are colored red for the cCP and iCP, and blue for the tCP. 10
- Figure 4 Peptide bond hydrolysis by the proteasome. The nucleophilic Thr10 γ (red) attacks the electrophilic carbonyl carbon atom of the peptide bond thereby releasing the first cleavage product. In a second step, hydrolysis of the formed acyl-enzyme complex (green bond) frees the N-terminal cleavage product and restores the catalytic Thr1. (Adapted from [16])..... 13
- Figure 5 Schematic representation of the proteasomal substrate binding channel of the proteasomal active sites. The primed (S') pockets and the P' sites of the ligand are colored in blue, and the unprimed specificity (S) pockets and the corresponding substrate residues (P) are colored green. The active site, including Thr1 and the scissile peptide bond are shown in red. (Adapted from [16])..... 14
- Figure 6 Chemical structures of clinically relevant proteasome inhibitors. Bortezomib and carfilzomib are FDA approved for the treatment of multiple myeloma and refractory multiple myeloma, respectively. Oprozomib is in clinical trials for treatment of multiple myeloma as well as solid tumors.....22

Figure 7 Chemical structures of subunit specific proteasome inhibitors. PR-957 targets immunoproteasome subunit $\beta 5i$; PI-1840 targets constitutive proteasome subunit; $\beta 5c$; NC-022 specifically targets trypsin-like subunits of immuno and constitutive proteasomes ($\beta 2c/\beta 2i$).	24
Figure 8 Sequence alignments of proteasomes caspase-like ($\beta 1$) constitutive (c) and immuno (i) subunits from <i>Homo sapien</i> (hu), <i>Mus musculus</i> (mu), and <i>T. acidophilum</i> (y). Secondary structures (S: β -sheet; H: helix) are indicated for the human constitutive subunits. Residues important for the active site are indicated with red boxes. Residues contributing to the substrate-specificity pockets are highlighted by colored boxes: S1 pocket, green; S2 pocket, blue; S3 pocket, yellow.	33
Figure 9 Sequence alignments of proteasomes trypsin-like ($\beta 2$) constitutive (c) and immuno (i) subunits from <i>Homo sapien</i> (hu), <i>Mus musculus</i> (mu), and <i>T. acidophilum</i> (y). Secondary structures (S: β -sheet; H: helix) are indicated for the human constitutive subunits. Residues important for the active site are indicated with red boxes. Residues contributing to the substrate-specificity pockets are highlighted by colored boxes: S1 pocket, green; S2 pocket, blue; S3 pocket, yellow.	34
Figure 10 Sequence alignments of proteasomes chymotrypsin-like ($\beta 5$) constitutive (c) and immuno (i) subunits from <i>Homo sapien</i> (hu), <i>Mus musculus</i> (mu), and <i>T. acidophilum</i> (y). Secondary structures (S: β -sheet; H: helix) are indicated for the human constitutive subunits. Residues important for the active site are indicated with red boxes. Residues contributing to the substrate-specificity pockets are highlighted by colored boxes: S1 pocket, green; S2 pocket, blue; S3 pocket, yellow.	35
Figure 11 One-dimensional SDS-PAGE analysis of purified human proteasome constitutive core particles	38
Figure 12 Proteasome crystal and diffraction from condition containing Peg 8K, sodium chloride, and cesium.	39
Figure 13 Proteasome crystals and diffraction from condition containing 40% MPD and 0.2M Sodium Formate.	39
Figure 14 Ramachandran Plot for human 20S proteasome crystal structure.	41
Figure 15 Quaternary structure of α -subunits. Subunit $\alpha 5$ is shown in ribbon representation, with each α -helix labeled as H, and β -sheets are labeled as S.	45

Figure 16 Quaternary structure of β -subunits. All β -type subunits have a similar arrangement of two α -helices, H1 and H2, followed by two sets of five-stranded anti-parallel β -sheets, then two more α -helices, H3 and H4. An additional two-stranded β -sheet is found in the loop connecting β -strands S2 and S5. Subunit β 5 (shown) has an additional two-turn α -helix at its C-terminal end.	45
Figure 17 Conserved C-terminal loop of β 2 subunits. The C-terminus of subunit β 2, from residue Leu186 to Glu220, wraps around subunit β 3 and forms an additional strand in its anti-parallel β -sheet, S12. The black arrow indicates the C-terminus of β 2.	47
Figure 18 Surface representation of the human 20 proteasome. Subunits are colored according to their symmetrical counterpart.	48
Figure 19 Ribbon representation of top view of the α -subunits. The arrows indicates the entry to the catalytic chamber, which is closed off by the N-termini of the α -subunits. The 20S core can be activated by proteins such as the 19S regulatory particle, which inserts short C-terminal peptide segments into shallow pockets on the α -ring surface, causing structural rearrangements. Each α -subunit is colored uniquely.	48
Figure 20 Structural alignment of proteasome catalytic subunits. Human (wheat), yeast (yellow) and mouse (teal).	49
Figure 21 Active site of catalytic subunits β 1, β 2, and β 5. Residues involved in catalysis are shown in magenta, and are conserved in all three distinct catalytic subunits. Hydrogen bonds are shown as black dashed lines, and the unique network increases the nucleophilicity of Thr1. Thr1 is shown in black, and the active site water, WAT, is shown as a yellow sphere. The water may be involved in intramolecular autolysis and substrate proteolysis by mediating proton transfer between Thr1O γ and Thr1N.	51
Figure 22 S1 specificity pockets of the catalytic β -subunits.	53
Figure 23 Proteasome reaction mechanism against epoxyketone inhibitors. The carbonyl carbon atom is attacked by the proteasomes Thr1O γ nucleophile. The formed hemiketal intermediate can either dissociate to deliberate Thr1O γ or irreversibly cycle under the formation of a secondary amine (morpholine ring system) involving Thr1N.	59
Figure 24 View of the inner cavity of the human 20S proteasome with carfilzomib bound. The α 4 and β 4-subunits have been removed to give a view of the	

core of the 20S proteasome. Carfilzomib, shown as green sticks, is bound to all six active sites.	63
Figure 25 Ramachandran plot for carfilzomib bound human 20S proteasome.....	64
Figure 26 Surface representation and electron density of carfilzomib bound to the chymotrypsin-like active site. Thr1 is colored black; carfilzomib is colored green; electron density is blue and is contoured to 1σ	65
Figure 27 Carfilzomib bound to the trypsin-like subunit, $\beta 2$	66
Figure 28 Carfilzomib bound to caspase-like site, $\beta 1$	66
Figure 29 Hydrogen bonding interactions of carfilzomib in the chymotrypsin like subunit $\beta 5$. Hydrogen bonds are indicated with black dashed lines; the active site water molecules are shown as red spheres. The nucleophilic Thr1 is shown in black.....	67
Figure 30 Hydrogen bonding interactions of carfilzomib in the trypsin-like subunit $\beta 2$. Hydrogen bonds are indicated with black dashed lines; the active site water molecules are shown as red spheres. The nucleophilic Thr1 is shown in black.....	68
Figure 31 Hydrogen bonding interactions of carfilzomib in the caspase-like subunit $\beta 1$. Hydrogen bonds are indicated with black dashed lines; the active site water molecules are shown as red spheres. The nucleophilic Thr1 is shown in black.....	68
Figure 32 Structural alignment of chymotrypsin-like and caspase-like sites with carfilzomib bound. In caspase-like sites, H116 prevents carfilzomib's P4 phenyl group from accessing the S4 pocket (black arrow), resulting in a disordered N-terminus of the drug and decreased selectivity. The N-terminal morpholino ring in chymotrypsin-like sites assumes a fixed conformation due to steric effects of Y107.	69
Figure 33 Structural alignment of the chymotrypsin-like sites of the human:carfilzomib, yeast:bortezomib, and mouse:PR-957 complexes. The black arrow indicates the P2 residue, which is flipped in carfilzomib (green) and PR-957(white) relative to the P2 phenylalanine of bortezomib (pink). Thr1 is shown in black, and the specificity pockets which harbor ligand side chains are labeled as S1-S4.	71
Figure 34 Shifts in the S1 pocket of the human proteasome chymotrypsin-like subunit, $\beta 5$, upon carfilzomib binding. Unbound and carfilzomib bound structures are aligned by their C α atoms. The unbound $\beta 5$ is colored white,	

carfilzomib bound $\beta 5$ is wheat, carfilzomib is colored green, and Thr1 is black. Regions of the protein that shift upon carfilzomib binding are indicated with black arrows.	73
Figure 35 PR-957 bound to mouse constitutive 20S Proteasome chymotrypsin-like site. The same structural rearrangements are observed as in the carfilzomib:proteasome structure, indicated with black arrow. The unbound protein is teal, PR-957 bound is Pink, and PR-957 is yellow (PDB ID unbound, 3UNE; PR-957 bound, 3UNB).....	74
Figure 36 PR-957 bound to mouse 20S immunoproteasome chymotrypsin-like site. Same color scheme as figure 36. No structural rearrangements are observed due to Gln53 stabilizing Met45 resulting in a larger S1 pocket than constitutive proteasomes $\beta 5$ (PDB ID unbound, 3UNH; PR-957 bound, 3UNF).	75
Figure 37 Carfilzomibs P3 and P4 residues contribute to its increased selectivity for constitutive proteasomes chymotrypsin-like sites compared to PR-957. Carfilzomib (left image) has a P3 leucyl group and P4 phenyl group which stabilize ligand binding. PR-957 (right image) has a smaller alanine at P3 and no P4 residue, resulting in decreased selectivity towards constitutive proteasomes chymotrypsin-like sites and more favorable binding to immunoproteasomes.	76
Figure 38 Structural alignment of the chymotrypsin-like subunit $\beta 5c$ of the human proteasome:carfilzomib complex with the corresponding mouse immuno subunit $\beta 5i$. The human proteasome subunit is shown in wheat; carfilzomib is shown in green; the active site Thr1 from human is shown in black; mouse immuno subunit is shown in white. Conserved amino acid substitutions are shown as white sticks and labeled in black. Residue substitutions not conserved with human $\beta 1i$ are shown as black sticks, with the human substitution shown in black in parenthesis. Red labels are of the constitutive particle residue that is substituted.	80
Figure 39 Structural alignment of the trypsin-like subunit $\beta 2c$ of the human proteasome:carfilzomib complex with the corresponding mouse immuno subunit $\beta 2i$. The same color scheme is used as Figure 40.	81
Figure 40 Structural alignment of the human caspase-like subunit $\beta 1c$ with the corresponding mouse immuno subunit $\beta 1i$. The same color scheme is used as Figure 40-41.	82
Figure 41 Chemical structure of α -mangostin	92

Figure 42 In vitro inhibition of chymotrypsin-like and caspase-like activities of the proteasome by α -mangostin. The IC_{50} values against the chymotrypsin-like and caspase-like activities were calculated to be 700 nM, and 1.5 μ M, respectively. Error bars represent standard deviations of triplicate experiments.....	93
Figure 43 In vitro binding assay of α -mangostin against human 20S proteasome. Lane 1 contains only proteasome and probe; lane 2 shows bortezomib as a positive control for probe binding; lane 3 is α -mangostin. The diminished band intensities signify binding of inhibitor to the catalytic site indicated.	94
Figure 44 IC_{50} Graph against Molt4 Leukemia Cells IC_{50} value is approximately 15 μ M.	95
Figure 45 IC_{50} Graph against human dermal fibroblasts.....	96
Figure 46 Western Blot Analysis of α -Mangostin in Molt4 Leukemia Cells.	97
Figure 47 In vitro binding assay of purified human 20S proteasome. Lane 1 is a control with no inhibitor, only probe. Lane 2 had proteasome incubated with bortezomib prior to addition of probe. Lane 3 was proteasome incubated with PPH-1 prior to addition of probe. The diminished band intensities compared to lane 1 indicates binding of the inhibitors.....	99
Figure 48 Chemical structure of PPH-1	99
Figure 49 In vitro enzymatic activity of PPH-1 against chymotrypsin-like activity. The IC_{50} against chymotrypsin-like site was determined to be 15 μ M.	100
Figure 50 In vitro enzymatic activity PPH-1 against caspase-like sites. The IC_{50} against caspase-like sites was determined to be 15 μ M.	100
Figure 51 Activity of PPH-1 against cancer cell lines. Cell lines tested were OVCAR 3(ovarian cancer), MDA-MB-231(breast cancer), and multiple myeloma cell lines 2631, and G3.	102
Figure 52 Binding Assay for hit plate screen hits. Bortezomib (BTZ) was used as a positive control for inhibitor binding. None of the hits from the screen show diminished band intensity compared to the “no inhibitor” control, therefore they were not further pursued as possible proteasome inhibitors. ..	106
Figure 53 Structures of natural product polyphenol proteasome inhibitors.....	108
Figure 54 Proteasome binding probe. (Left) Schematic of the probe. (Right) Probe. ...	114

Figure 55 Proteasome and Fatty Acid Synthase Inhibitors. Orlistat (bottom right) inhibits fatty acid synthase, the remaining compounds inhibit the 20S Proteasome.	121
Figure 56 Design of dual proteasome/fatty acid synthase inhibitors from orlistat and <i>N</i> -CBz- <i>O</i> -Bn homobelactosin C	122

LIST OF TABLES

	Page
Table 1 Chemical Structures of Proteasome Inhibitors.....	19
Table 2 Crystallography Statistics.....	42
Table 3 High Through-Put Screening Hits from Hit Plate.....	103
Table 4 First Generation of Dual Proteasome/FAS Inhibitors	123
Table 5 IC ₅₀ Values of Second Generation Belactosin C Analogs	124

1. INTRODUCTION

1.1 Protein Degradation

Proteases perform many tasks vital to the cell. As cellular structures are continually rebuilt, homeostasis between anabolic and catabolic pathways must be maintained [1]. Mis-folded and malfunctioning proteins, prone to aggregation, must be scavenged and degraded. Otherwise, these abnormal proteins would accumulate in the cell, endangering the cell's functionality. In addition, the activity of many regulatory proteins is also controlled by means of their degradation. Proteases vary from small proteins such as extracellular trypsin and the intracellular caspases, to large ATP-dependent multifunctional proteases called proteasomes.

Lysosomes are specialized membrane-enclosed organelles that contain an array of digestive enzymes. Until the early 1980's lysosomes were believed to be the site of all protein degradation. It is now known that the role of lysosomes is limited to metabolism, including the digestion of extracellular proteins taken up by endocytosis as well as the gradual turnover of cytoplasmic organelles and cytosolic proteins [2]. In lysosomes, proteases termed cathepsins unselectively degrade proteins. The resulting peptide fragments can be loaded on major histocompatibility complex class (MHC) II receptors and presented on the cell surface to immune cells [3].

The task of breaking down damaged, misfolded, or no longer needed proteins, however, is performed by the 26S proteasome. In the cytoplasm and in the nucleus of

eukaryotic cells, the ubiquitin/proteasome system (UPS) provides the main mechanism for a selective and ATP-dependent degradation of short-lived proteins. Proteasome degradation is responsible not only for removal of damaged protein in quality control, but also for precise regulation of the cell cycle [4], transcription [5], and antigen presentation [6]. The 26S proteasome is composed of a multi-subunit 20S core particle (20S proteasome) and two multi-subunit ATP dependent regulatory particles (19S regulatory complex), which selectively cuts polyubiquitinated proteins to peptides of diverse lengths[7].

In the ubiquitin/proteasome system, proteins to be degraded are covalently modified at a lysine residue by a string of ubiquitin molecules, yielding a polyubiquitinated form of the protein which as a substrate for proteasome degradation. This modification leads to the identification of the tagged protein by the 19S regulatory particle of the proteasome, which then unfolds the protein and feeds it into the 20S core particle where proteolysis takes place.

A number of proteins that control fundamental cellular processes, such as gene expression and cell proliferation, are targets for regulated ubiquitination and proteolysis [8]. An example is the controlled degradation of cyclins, which regulate progression through mitosis in eukaryotic cells. The entry of all eukaryotic cells into mitosis is controlled in part by cyclin B, which is a regulatory subunit of a protein kinase called Cdc2. The association of cyclin B with Cdc2 is required for activation of the Cdc2 kinase, which initiates the events of mitosis (including chromosome condensation and nuclear envelope breakdown) by phosphorylating various cellular proteins [9]. Cdc2

also activates a ubiquitin-mediated proteolysis system that degrades cyclin B toward the end of mitosis [10]. This degradation of cyclin B inactivates Cdc2, allowing the cell to exit mitosis and progress to interphase of the next cell cycle. The ubiquitination of cyclin B is a highly selective reaction, targeted by a 9-amino-acid cyclin B sequence called the destruction box [11]. Mutations of this sequence prevent cyclin B proteolysis and lead to the arrest of dividing cells in mitosis, demonstrating the importance of regulated protein degradation in controlling the fundamental process of cell division.

Additionally, many of the peptides produced by proteasome degradation are further broken down to single amino acids; however, in vertebrates a fraction goes on to serve as antigens for the immune system. The peptides transit to the lumen of the endoplasmic reticulum (ER) by the transporter associated with antigen processing (TAP), a member of the ATP-binding cassette transporter family[12]. Once in the ER, peptides can associate with the binding cleft of nascent MHC I receptors based on affinity for the MHC I complex, which is dependent on the peptides C-terminal anchor [13]. Cleavage preferences of the proteasome dictate the C-terminal anchors, and once peptides are bound to the MHC I complex, only stable receptor:ligand complexes adopt a mature structure and are transported in vesicles to the cell membrane for their exposure to the extracellular environment. Cytotoxic T lymphocytes destroy bacterial or viral peptides which are identified by surveying effector cells of the immune system that monitor the peptide cargos of both MHC I and II receptor proteins [14]. Therefore, the proteasome plays a key role in the immune system by shaping the antigenic pool of peptides.

1.2 Ubiquitin Proteasome System

The field of ubiquitin and regulated protein degradation was created in the 1980's, largely through the complementary discoveries by the laboratories of A Hershko (Technion, Haifa, Israel) and A Varshavsky (MIT, Cambridge, MA) [3]. Hershko and his colleagues revealed the initial understandings of ubiquitin conjugation and ubiquitin-mediated proteolysis in cell extracts, while Varshavsky and colleagues uncovered the specific biological necessity of the ubiquitin proteasome system in the cell cycle, DNA repair, protein synthesis, transcriptional regulation, and stress response. These discoveries caused a major expansion of the ubiquitin field in the 1990's, and it became clear that control through regulated protein degradation rivals, and may even surpass in significance, the classical regulation through transcription and translation.

Defective or outlived intracellular proteins undergo posttranslational modification *via* ubiquitination, thereby rendering them targets for degradation by the 26S proteasome. Ubiquitin is a 76-residue polypeptide and is covalently linked to the target protein through the formation of an isopeptide bond between the amino group of a substrate's lysine residue, and ubiquitin's C terminus. This reaction process involves three separate enzymes: E1, an ATP dependent activating enzyme which forms a thiol ester with the carboxyl group of ubiquitin's C-terminal G76 residue, rendering it ready for nucleophilic attack of a substrate's lysine side chain; E2, a conjugating enzyme responsible for carrying the activated ubiquitin molecule to ligase E3; and E3, which is bound specifically to the substrate, and facilitates the attachment of the ubiquitin molecule to the substrate lysine,

as well as the attachment of additional ubiquitins to form a chain [15] (Figure 1). Several classes of E3 enzymes exist, lending to selectivity due to their ability to bind to recognition sites in their target proteins. Recognition of ubiquitin tagged proteins by the 26S Proteasome requires a chain of at least four ubiquitin molecules, whereas proteins tagged with fewer than four are targeted for endocytosis, resulting in proteolysis in the lysosome[16].

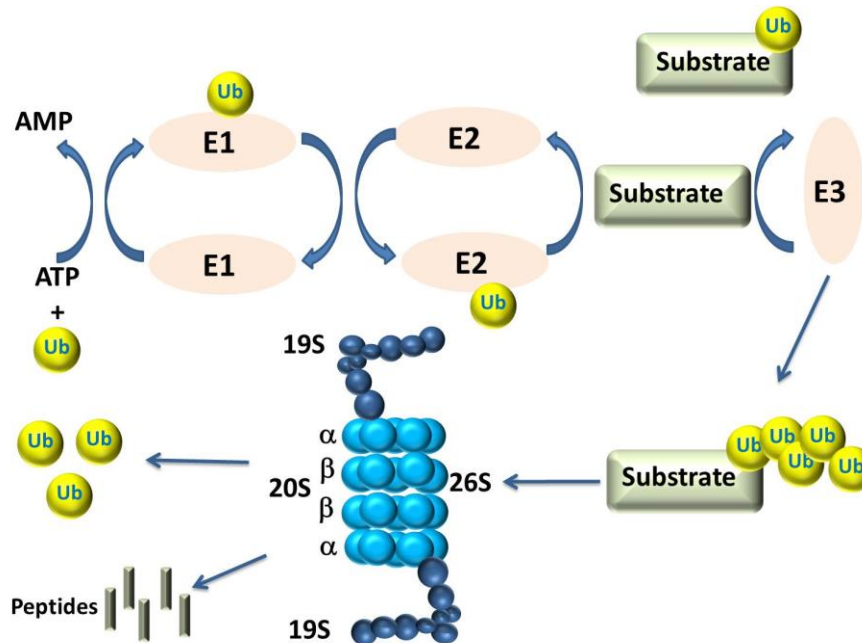


Figure 1 Ubiquitin proteasome pathway. Adenylation of ubiquitin by an E1 enzyme is followed by its transfer to an E2 ubiquitin-conjugating enzyme. An E3 ligase covalently attaches ubiquitin to a lysine residue in the target protein. The ubiquitin chain is extended by E3, and substrates with at least five ubiquitin molecules are degraded by the 26S proteasome. Adapted from [17]

1.3 Proteasomes Assembly and Architecture

Cellular homeostasis is maintained by the proper recognition, unfolding, translocation and cleavage of protein substrates. The success of the proteasome is dependent on proper assembly and regulation of the approximately 2.5 MDa complex, formed by the 28 subunit core particle (CP), or 20S proteasome, and two 19S regulatory particles (RP) which cap each end. The entire complex is termed the 26S proteasome, and each subunit plays an important role in ensuring proper protein function. The importance of the eukaryotic proteasome is exemplified by its early evolutionary origin, with homologous proteins found in archaeon's, and eubacteria. Proteasome CP's have been extensively studied and the following summarizes the current understandings of its evolution and function.

Bacteria, with the exception of actinomycetes, harbor a protease termed heat shock locus V (Hs1V), composed of two hexameric rings, with each subunit having peptide cleavage activity through an N-terminal threonine residue [18]. In contrast, the crystal structure of the 20S proteasome from archaeon *Thermoplasma acidophilum* revealed a hollow cylinder shape, comprised of four stacked rings of seven subunits each, with an overall cylinder length of 148 Å. The approximately 700 kDa protein has two types of subunits present, termed α and β , which are arranged as $\alpha_7\beta_7\beta_7\alpha_7$. With the exception of *Aeropyrum pernix*, all β subunits of archaebacterial proteasomes are catalytically active and share homology with the subunits from Hs1V proteases, and all act through N-terminal threonine residues [19].

In contrast to the low subunit complexity of prokaryotic proteasomes, eukaryotic proteasomes CPs are composed of 14 distinct subunits; seven α type and seven β type. Crystal structures of proteasome CP's from *Saccharomyces cerevisiae*, *Bos taurus* and *Mus musculus*, show that the CP's have a subunit arrangement similar to that of *T. acidophilum*, only they possess C_2 symmetry rather than sevenfold symmetry [20](Figure 2). Furthermore, only subunits β_1 , β_2 , and β_5 have been shown to harbor catalytic activity through N-terminal Thr1 residues.

The α subunits are also attributed with several functions. Their intrinsic ability to form a heptameric ring is pivotal for the assembly of the β subunits and the formation of the CP. Additionally, “shuttling sequences” in the α subunits enable the import of the CP into the nucleus and its re-export. The α subunits also form the entry gate to the interior of the CP. Specifically, the N-terminal extensions of the subunits α_2 , α_3 , and α_4 close the CP on both ends and thereby abolish unmediated degradation of intracellular proteins [19, 21].

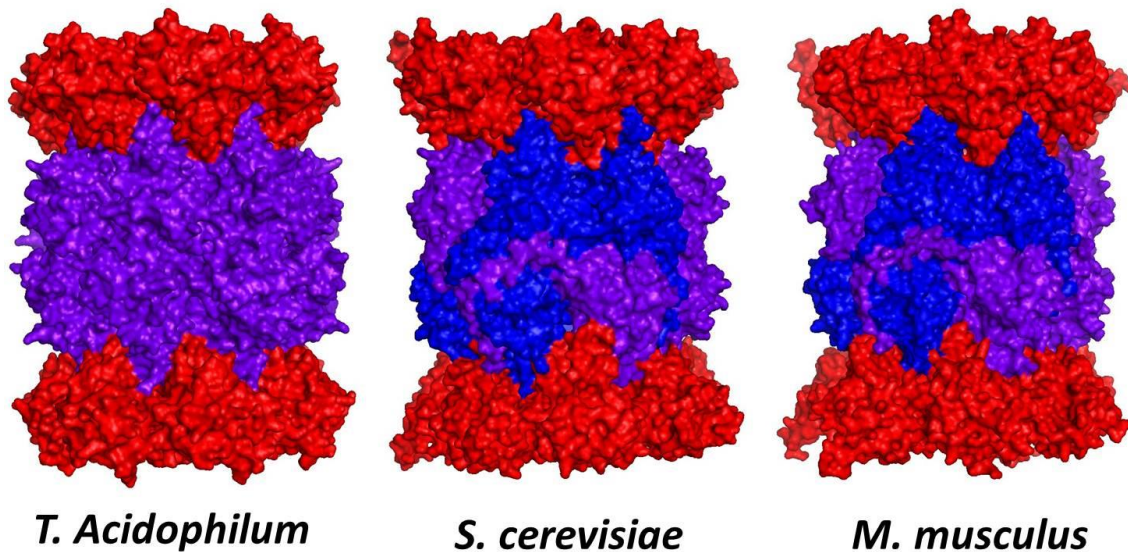


Figure 2 Comparison of proteasome core particles from archaea (*T. acidophilum*), eukaryotes (*S. cerevisiae*), and vertebrates (*M. musculus*). Structures are depicted in surface representation. The α -subunits are shown in red, catalytic β -subunits in purple, and non-catalytic β -subunits are blue. All β -subunits of *T. acidophilum* are catalytic, whereas six of the fourteen β -subunits from eukaryotes and vertebrates are catalytic, with three distinct cleavage specificities; chymotrypsin-like, trypsin-like, and caspase-like.

Construction of the core particle begins with the α rings. Four CP-chaperones, Pba (proteasome biogenesis associated protein) in yeast, and PAC (proteasome assembly chaperone) in human, are responsible for the assembly process. The chaperones are functional as heterodimers, PAC1-PAC2 and PAC3-PAC4. Upon completion of α -ring assembly the β subunits are incorporated and PAC3-PAC4 is released due to steric hindrance with β 4. A fifth chaperone, Ump1/POMP, acts to ensure that the two half CP's do not dimerize until all seven β subunits have been assembled.

Five of the seven β -type subunits, including the three catalytic subunits (β 1, β 2, and β 5) are synthesized as proproteins, which are proteolytically processed to their mature

forms found in the assembled core particle [21]. These propeptides are involved in CP assembly assisted by the proteasome maturation factor Ump1 and protect the catalytic active sites from inactivating N α acetylation prior to proteasome maturation [22]. Upon completion of CP assembly the propeptides are removed by intramolecular autolysis, exposing the proteolytically active Thr1 residues and leaving the proteasome ready to cleave polypeptides that are fed into the 20S core, beginning with the degradation of Ump1 and PAC1-PAC2 [23, 24].

All seven distinct α and β subunits have the same general β -sandwich architecture, which is typical for N-terminal hydrolases [25]. The β -sandwich consists of three α helices H1, H2 and H3, followed by two five-stranded antiparallel β sheets, then three additional α helices, H4, H5 and H6 on the opposite side of the β sheets.

Proteins that activate the 20S core for proteolysis, such as the 19S RP, the 11S regulator, and proteasome-activating nucleotidase (PAN), insert C-terminal peptide segments of ~6 residues into shallow pockets on the α -ring surface [26]. ATPase regulatory particles contain a C-terminal hydrophobic-tyrosine-X (HbYX motif), which has been shown to be crucial for activation of the 20S core. Interaction of the HbYX motif with the proteasome α -chains, results in substantial structural rearrangements which allow access to the interior cavity [27].

Hydrophobic molecules such as rapamycin, have been shown to prevent binding of the 19S RP, thereby acting as allosteric inhibitors of the proteasome [28]. Additionally, prion proteins, which form aggregated β -sheet isoforms, inhibit proteasome activity by stabilizing the entry gate and preventing its opening [29].

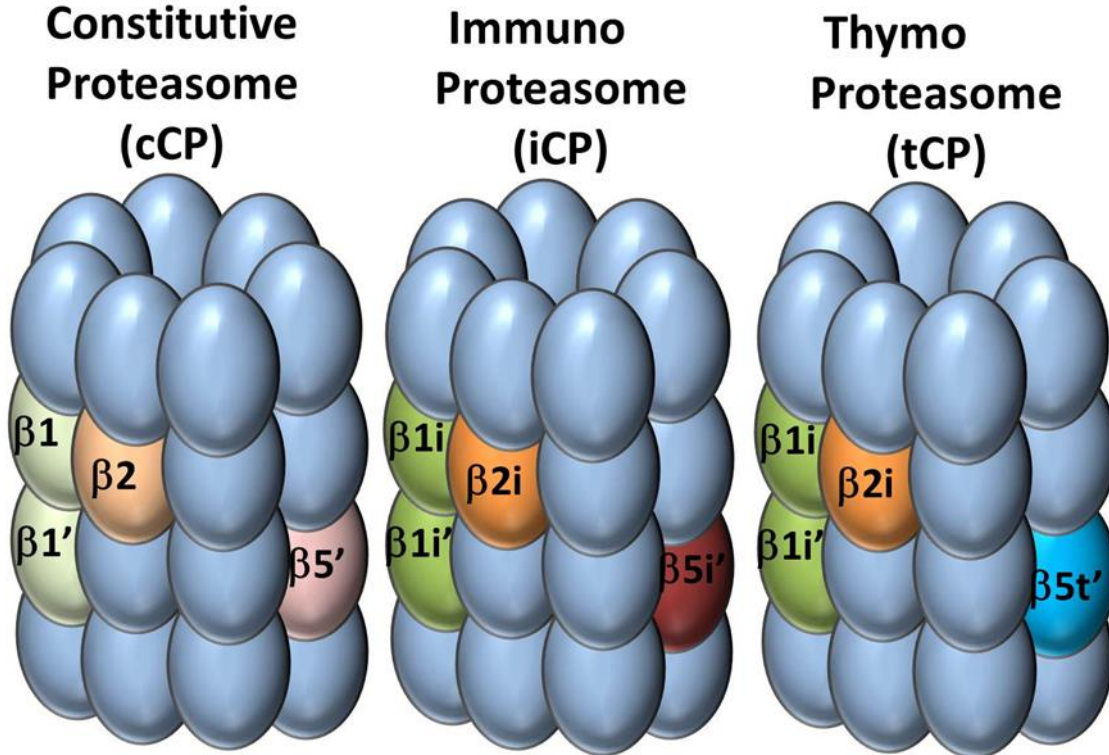


Figure 3 Proteasome core particle isoforms. Inactive subunits are colored grey. Caspase-like sites ($\beta 1$) are colored green; Trypsin-like sites ($\beta 2$) are colored orange; Chymotrypsin-like sites are colored red for the cCP and iCP, and blue for the tCP.

Unlike yeast and other primitive eukaryotes, vertebrates encode for three types of core particles: constitutive proteasome (cCP), immunoproteasome (iCP) and the thymoproteasome (tCP) [30] (Figure 3). The constitutive core particle comprises the catalytic active constitutive subunits $\beta 1c$, $\beta 2c$, and $\beta 5c$, and represents the prevailing proteasome species in cells of non-hematopoietic origin. The immunoproteasome, iCP, is expressed predominantly in immune tissues and cells such as monocytes and lymphocytes [21]. The expression of the iCP catalytic subunits, $\beta 1i$ /LMP2, $\beta 2i$ /MECL-1, and $\beta 5i$ /LMP7, can also be induced by tumor necrosis factor- α and interferon- γ , which are

released by T cells during infection [31]. These immunoproteasome subunits replace their constitutive counterparts, and preferentially hydrolyze proteins after nonpolar amino acids [32]. This results in peptide fragments having hydrophobic C termini, which then anchor to MHC-1 molecules, thus priming the immune system. It has been shown in gene-targeted mice, however, that immunoproteasome subunits are not required for epitope presentation [33], therefore constitutive CPs must also be capable of antigen processing. Interestingly, mixed proteasomes of the composition $\beta 1c$, $\beta 2c$, and $\beta 5i$ or $\beta 1c$, $\beta 2i$, and $\beta 5i$ were reported to account for 30-50% of all cellular CPs, which further broadens the repertoire of antigens presented to CD8 T cells, and suggests that the antigens presented by a given cell depends on their proteasome content [34].

The third type of CP, thymoproteasome (tCP) is specific to vertebrate and is exclusively expressed in cortical thymic epithelial cells. The subunit arrangement in tCP only replaces $\beta 5i$ with $\beta 5t$ (PSMB11), while the remaining catalytic subunits remain that of the immunoproteasome, $\beta 1i$ and $\beta 2i$. This tCP has been implicated to play a pivotal role for the development of CD8⁺ cytotoxic T cells as part of the adaptive immune system [35].

1.4 Proteasome Proteolysis

The proteasome is a threonine protease and belongs to the family of N-terminal nucleophile hydrolases. Six of the proteasomes β -chains contain catalytic N-terminal Thr1 residues, however there are three distinct cleavage specificities: trypsin-like (T-L, chains

$\beta 2/\beta 2'$), chymotrypsin-like (CT-L, chains $\beta 5/\beta 5'$) and caspase-like (C-L, chains $\beta 1/\beta 1'$). Trypsin-like subunits cleave after basic residues such as lysine or arginine, caspase-like sites cleave after acidic residues, and chymotrypsin-like sites cleave after hydrophobic residues.

Despite varying substrate selectivity's, the catalytic mechanism is the same for all three active sites (Figure 4). Thr1O γ is activated by a conserved water molecule and nucleophilically attacks the substrates electrophilic carbonyl carbon atom of the peptide bond [36]. A tetrahedral intermediate is formed, which collapses to release one substrate degradation product. Hydrolysis of the formed acyl-enzyme complex frees the N-terminal cleavage product and restores the catalytic Thr1.

The binding affinity of a substrate for its target is dictated by both entropic and enthalpic contributions, according to Gibbs energy of binding (ΔG), $K_a = e^{\frac{\Delta G}{RT}}$. Several studies using peptide-based inhibitors have shown that it is enthalpic interactions with the proteasomes' primed (S1', S2', S3') and unprimed (S1, S2, S3) substrate-binding pockets that drive the affinity of a substrate toward a particular active site [37, 38]. The binding pockets (S) accommodate the side chains of substrates, or P residues (Figure 5). The active site preferences (chymotrypsin-like, trypsin-like, or caspase-like) correspond to the interactions of the side chain of the last residue before the cleavage site, P1, with the S1 pocket. In particular, amino acid 45 at the base of the S1 pocket is believed to dictate the active sites substrate selectivity [39].

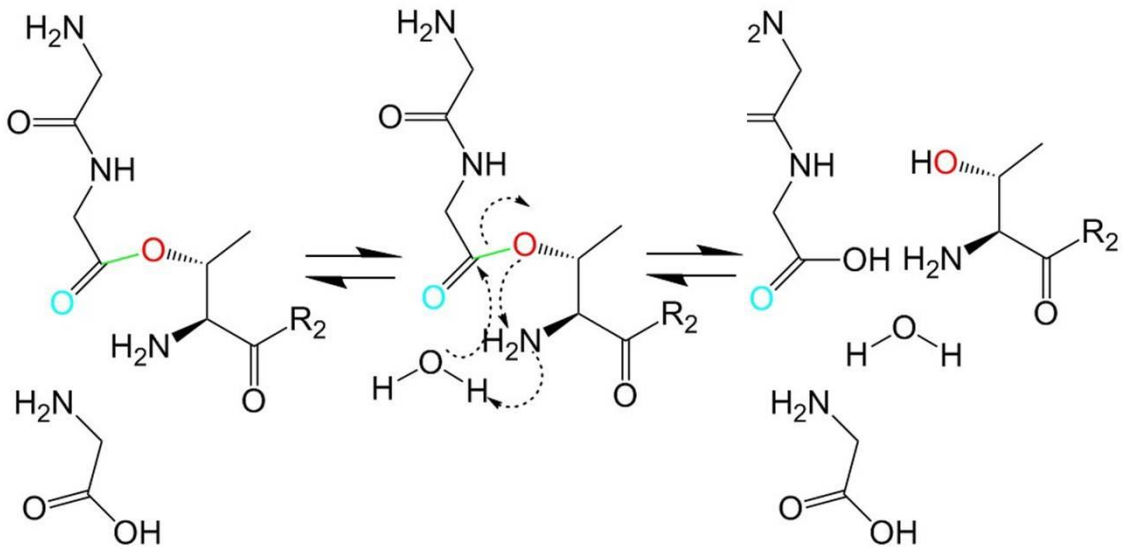
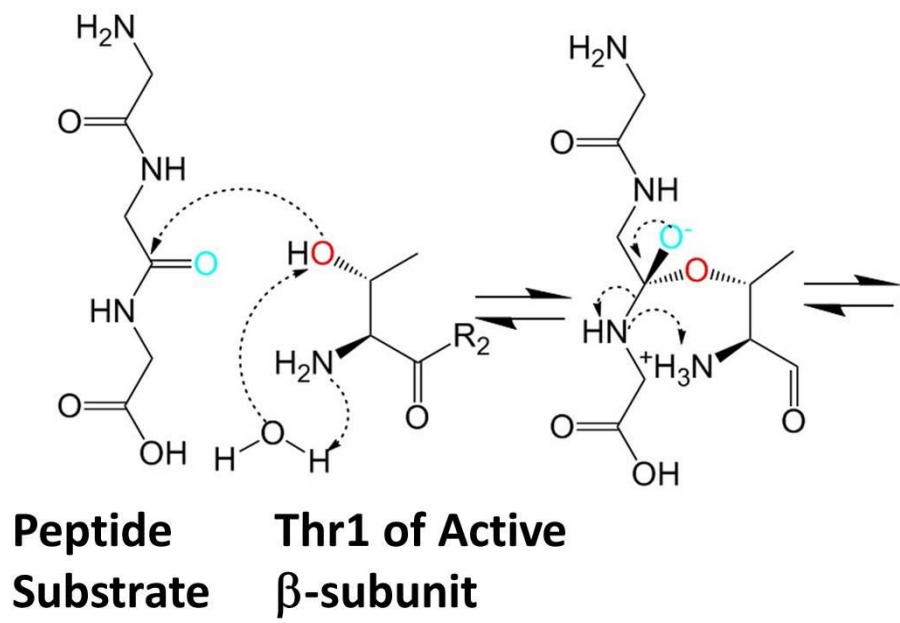


Figure 4 Peptide bond hydrolysis by the proteasome. The nucleophilic Thr1O γ (red) attacks the electrophilic carbonyl carbon atom of the peptide bond thereby releasing the first cleavage product. In a second step, hydrolysis of the formed acyl-enzyme complex (green bond) frees the N-terminal cleavage product and restores the catalytic Thr1. (Adapted from [16])

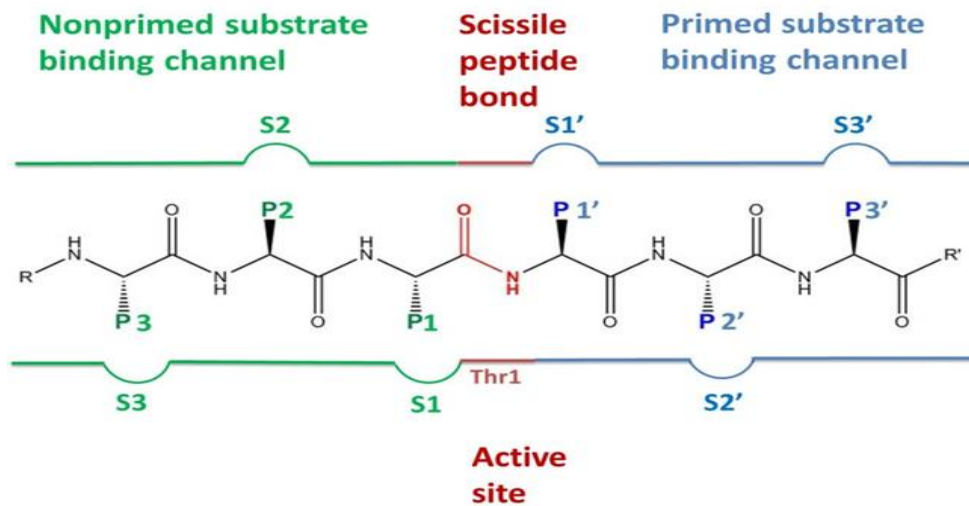


Figure 5 Schematic representation of the proteasomal substrate binding channel of the proteasomal active sites. The primed (S') pockets and the P' sites of the ligand are colored in blue, and the unprimed specificity (S) pockets and the corresponding substrate residues (P) are colored green. The active site, including Thr1 and the scissile peptide bond are shown in red. (Adapted from [16])

1.5 Immune Functions of the Proteasome

The evolution of the immune system in higher vertebrates brought with it a new role for the proteasome and lysosome in facilitating the immune system by screening foreign material in the extracellular and intracellular spaces [40]. Lysosomes are tasked with degrading extracellular proteins that are taken in by endocytosis, which produces antigenic peptides that are presented on MHC class II molecules and elicit antibody production.

Proteasomes, on the other hand, shape the pool of antigenic peptides that are presented by MHC I receptors on the cell surface to immune cells. All three proteasome types, cCP, iCP, and tCP, are capable of degrading intracellular proteins and generating

antigens. To bind to most MHC class I molecules, peptides have to be eight to nine residues long and have to have hydrophobic or basic C-termini. While a majority of the peptides generated are hydrolyzed by cytosolic peptidases to amino acids within a matter of seconds, a small fraction, approximately 30%, escape hydrolysis and are transported by the TAP transporter into the ER. Peptides that are longer than eight or nine residues are further trimmed by the aminopeptidase, ERAP1, and then all bind to MHC class I molecules and are delivered to the cell surface for presentation to circulating lymphocytes.

Although cCP peptides can trigger immune reaction, the iCP is the preferred form due to its ability to a higher fraction of peptides with the appropriate C-terminal anchor. During viral infection, cytokines are released which induce the expression of iCP's thereby enhancing antigen presentation and supporting the clearing of pathogens [33].

The tCP is restricted to the cortical thymic epithelial cells (cTECs) of the thymus, and this specialized organ of the immune system selects T lymphocytes for the optimal interaction strength of their T cell receptor (TCR) with MHC:self-peptide complexes. Interactions between the TCR and MHC can either be tight or weak, resulting in either negative or positive selection, respectively. Immature thymocytes whose TCR tightly binds to MHC receptors loaded with self-antigens are auto-reactive and are eliminated by negative selection, while T cells with weak TCR:MHC interactions are considered as self-tolerant and thus survive, positive selection[21, 41]. Chymotrypsin like activity is decreased in tCPs and this is believed to produce low-affinity self-epitopes for the presentation on MHC receptors, supporting positive selection [42]. To summarize, the tCP helps to establish a tolerant T cell repertoire, the prerequisite for a functional adaptive

immune system and the iCP and cCP promote immune surveillance and the elimination of pathogens. Therefore, all three are key players of the adaptive immune system in vertebrates and their malfunctioning positively correlates with the onset of diverse diseases.

1.6 20S Proteasome: Inhibition and Apoptosis

The 20S proteasome primarily serves to degrade regulatory or aberrant proteins [43], including: oncoproteins [44], transcription factors [45], cell cycle specific cyclins [4], cyclin-dependent kinase inhibitors [46], ornithine decarboxylase [47], and other key regulatory cellular proteins. Altered regulation of these events is linked to the development of cancers as well as autoimmune disorders [48], and thus the proteasome has become an attractive target for the treatment of such diseases.

Due to accelerated cell cycles and metabolism in tumour cells, the CP is often upregulated to increase turnover rates of proteins. This dependence on CP activity renders neoplastic cells highly susceptible to proteasome inhibition. Early studies established that proteasome inhibitors induce apoptosis in leukemic cell lines [49, 50], and additional in vitro studies revealed anti-proliferative and pro-apoptotic activity against hematological and solid tumors. Multiple myeloma cells have been found to be the most sensitive to CP inhibitors [51]. Myeloma cells are derived from plasma cells and their excessive synthesis of immunoglobulins as well as their chromosomal instability leads to many aberrant and misfolded proteins that have to be removed by the proteasome. Proteasome inhibitors

mainly induce apoptosis in proliferating cells and are protective against quiescent or terminally differentiated cells [50]. Inhibition of the proteasome causes the accumulation of protein aggregates and triggers ER stress as well as the unfolded protein response [52].

The exact mechanisms behind the higher sensitivity of malignant cells for proteasome inhibitors is unclear, however, induction of apoptosis can occur through a variety of pathways depending on the cell type. For instance, in MOLT-4 human T cell leukemia, proteasome inhibition leads to an increase in the tumor suppressor p53, which initiates genes that induce cell growth arrest, DNA repair or apoptosis [53]. In multiple myeloma and mantle cell lymphoma cell lines, proteasome inhibition prevents degradation of the inhibitor of the transcription factor NF κ B (I κ B α), blocking the expression of anti-apoptotic target genes [54-56]. This causes the induction of apoptosis in transformed cells by tumor suppressor genes such as cyclin kinase inhibitor p27kip1 [57], however, healthy cells remain unaffected [51], resulting in a therapeutic window for proteasome inhibition in hematological cancers.

1.7 Classes of Proteasome Inhibitors

Proteasome inhibitors can be broadly broken down into either covalent or noncovalent inhibitors. Covalent inhibitors harbor an electrophilic head group that reversibly or irreversibly inhibits the catalytic Thr10 λ . These can be further broken down into several more specific categories based on the type of pharmacophore: aldehydes, vinyl sulfones, vinyl amides (syrbactins), bornic acids, α' , β' epoxyketones, α

ketoaldehydes (glyoxals) and β lactones. With the exception of lactones and syrbactins, proteasome inhibitors consist of a peptide backbone of two to four amino acids attached to the electrophilic head group (Table 1). The crystal structure of the yeast proteasome in complex with calpain inhibitor I (acetyl-Leu-Leu-norleucinal) showed that the peptide backbone completes an antiparallel β -sheet in the substrate binding channel and the inhibitor side chains facilitate with active site selectivity by interacting with the proteasomes specificity pockets [21]. This has been consistent with other peptide based proteasome inhibitors to date, including PR-957 which was reported bound to the mouse immunoproteasome [16].

Peptide aldehydes inhibitors such as MG-132 (Z-Leu-Leu-Leu-H) are not suitable drug candidates because they are non-specific and block a broad range of serine and cysteine proteases. They are widely used, however, for in vitro and in vivo studies due to their cell permeability, and because their inhibition can be reversed [58]. Vinyl sulfones and naturally occurring syrbactins form an ether bond with Thr10 λ in a Michael-type 1,4-addition [59]. They are also non-specific and target cysteine proteases, limiting their use to in vivo and in vitro studies [60]. β lactones can be reversible or irreversible, depending on their P1 substituent [42], and marizomib (salinosporamide A; NPI-0052; Nereus Pharmaceuticals, Inc.) is currently in clinical phase I trials for myeloma, lymphoma and leukemia [41].

Boronic acids, such as bortezomib are among the most potent inhibitors of the proteasome and form a tetrahedral adduct with the active site Thr10 γ [36]. α' , β' epoxyketone inhibitors, such as epoxomicin, are also highly specific for the

proteasome. They irreversibly inhibit the proteasome in a bivalent reaction involving both Thr1O γ and Thr1N [61]. α ketoaldehydes inhibit the proteasome by the formation of a reversible cyclic Schiff base with Thr1O γ and Thr1N[51].

Table 1 Chemical Structures of Proteasome Inhibitors

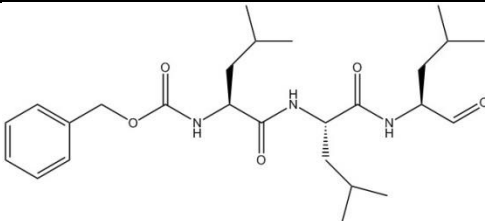
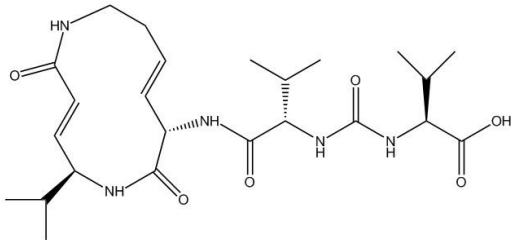
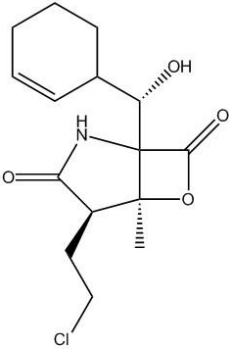
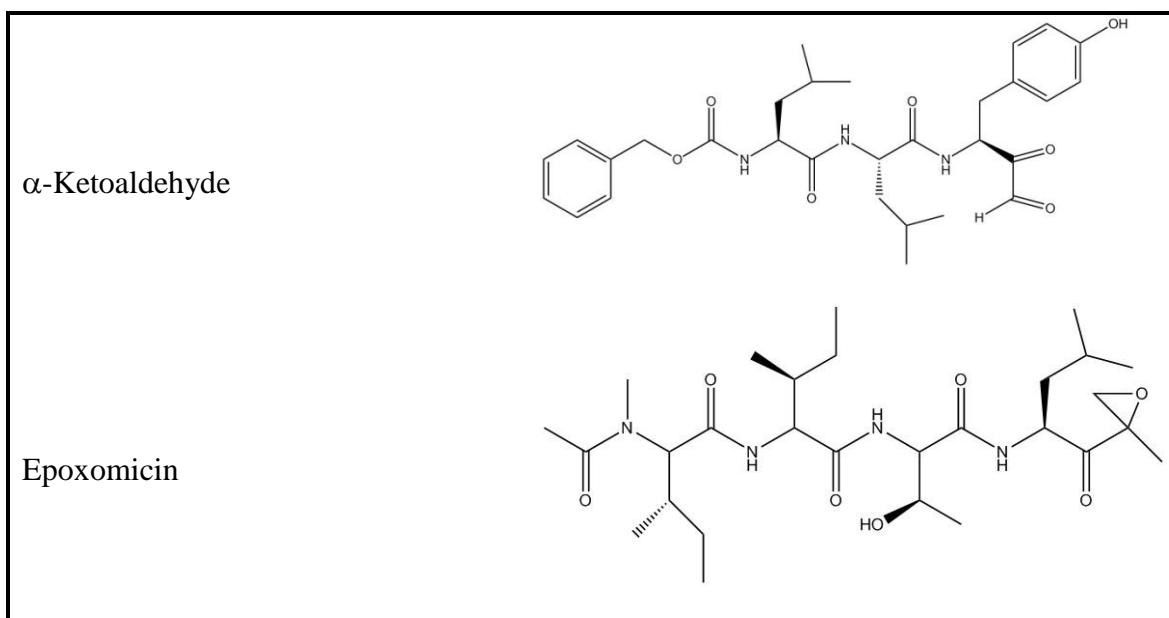
Chemical Name	Structure
MG-132 (Peptide Aldehyde)	
Syrbactin	
Salinosporamide A (Marizomib)	
Chemical Name	Structure

Table 1 continued



1.8 Clinically Relevant Proteasome Inhibitors

The boronic acid proteasome inhibitor bortezomib (Velcade[®], Millenium Pharmaceuticals, Inc.; PS-341) received FDA approval in 2003 for the treatment of multiple myeloma and relapsed or refractory mantle cell lymphoma [52]. The therapeutic benefit of using bortezomib during organ transplantation[62] and solid tumors such as non-small cell lung cancer is currently being investigated [63].

Bortezomib potently inhibits the β 5c, β 5i and β 1i active sites of the cCP and iCP with IC_{50} values of 3-8 nM, and targets the β 1c, β 2c and β 2i subunits to a lesser extent [64]. Bortezomib inhibits *via* a single reversible covalent adduct with its reactive boronate and has moderate affinity for serine proteases, including cathepsin G, cathepsin A, chymase, dipeptidyl peptidase II and HtrA2/omi. HtrA2 is involved in neuronal survival

[65], and inhibition by bortezomib is believed to be the cause of peripheral neuropathy in patients in about 30% of patients treated [65].

In 2012, carfilzomib (Kyprolis[®], PR-171; Onyx Pharmaceuticals) was approved by the FDA for the treatment of refractory multiple myeloma. A twenty-eight day cycle of carfilzomib cost \$10,000, making it the most expensive multiple myeloma drug on the market. Carfilzomib is the result of medicinal chemistry efforts to increase the potency and chymotrypsin like selectivity of the natural product proteasome inhibitor, epoxomicin. The bivalent, irreversible reaction mode of inhibition of α',β' epoxyketones makes them the most specific proteasome inhibitors known [61]. Carfilzomib potently blocks the $\beta5c$ and $\beta5i$ active sites of cCPs and iCPs with IC_{50} values of 6 nM and 33 nM, respectively. Carfilzomib treatment has been linked to neutropenia and thrombocytopenia in patients, however, it does not cause peripheral neurotoxicity, as observed with bortezomib [65, 66].

Oprozomib (Onyx Pharmaceuticals, Inc.; ONX 1912; PR-047) is currently being tested in phase 1b/2 clinical trials for the treatment of multiple myeloma, and a phase 1 study in patients with recurrent or refractory solid tumors (Figure 6). Oprozomib is an orally available epoxyketone analog which targets the chymotrypsin-like activities of constitutive and immunoproteasomes, $\beta5c$ and $\beta5i$ (IC_{50} $\beta5c$ 36nM; IC_{50} $\beta5i$ 82 nM) [67].

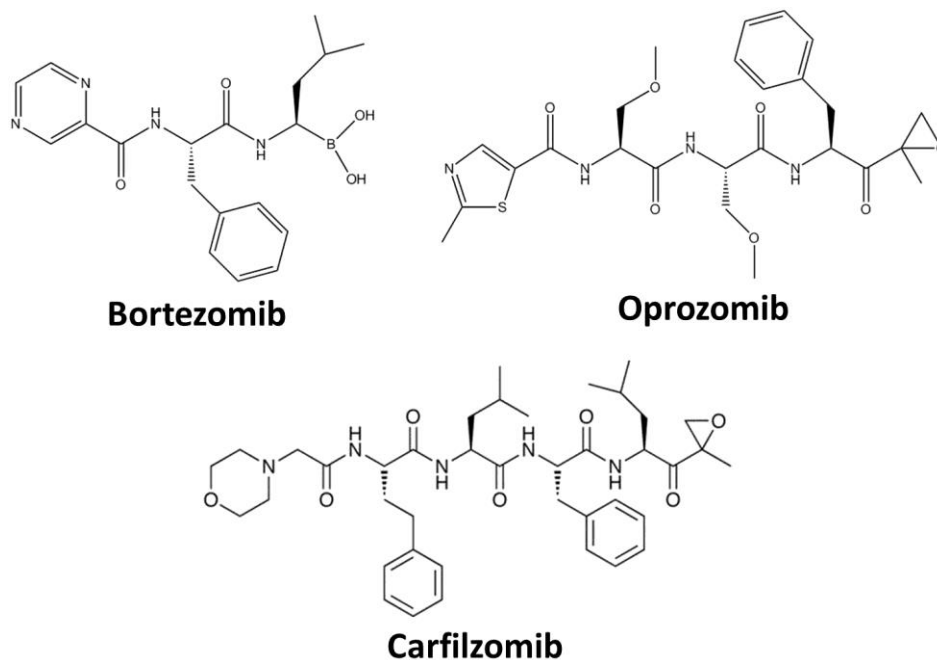


Figure 6 Chemical structures of clinically relevant proteasome inhibitors. Bortezomib and carfilzomib are FDA approved for the treatment of multiple myeloma and refractory multiple myeloma, respectively. Oprozomib is in clinical trials for treatment of multiple myeloma as well as solid tumors.

1.9 Proteasome Inhibitors of Individual Active Subunits

Subunit-specific inhibitory compounds are valuable tools for examining the impact of individual proteasome subunits on cell division and biological signaling pathways. The development of compounds with subunit specificities is often hindered by the strong inhibitory potency of most reactive functional head groups, such as boronic acids, or epoxyketones. Hence, only ligands that undergo optimal enthalpic interactions with the surrounding protein residues exert pronounced subunit selectivity.

Historically, chymotrypsin-like ($\beta 5$) activities have been the focus for design of subunit specific inhibitors. Recently, however, a series of α',β' epoxyketone inhibitors which specifically target either caspase-like ($\beta 1$) or trypsin-like ($\beta 2$) sites of the constitutive and immunoproteasomes have been reported [38, 68] (Figure 7). In particular, the trypsin-like specific inhibitor NC-022 was shown to sensitize malignant cells to the chymotrypsin-like specific inhibitors bortezomib and carfilzomib, rendering them appealing co-targets for antineoplastic therapy.

Inhibition of the immunoproteasomes chymotrypsin like subunit ($\beta 5i$) with the $\beta 5i$ specific inhibitor PR957, has recently shown effectiveness in diseases such as rheumatoid arthritis, [69], experimental colitis [70], lupus erythematosus [71], and Hashimoto's thyroiditis [72]. Whereas selective inhibition of $\beta 5i$ appears effective for the above mentioned diseases, inhibition of both constitutive and immunoproteasome $\beta 5$ subunits was shown to be required to induce an antitumor effect in multiple myeloma, non-hodgkin lymphoma, and leukemia cells [73].

Selective inhibition of the constitutive proteasomes $\beta 5$ subunit, by the non-covalent inhibitor PI-1840, has been shown to sensitize cancer cells to p53 and Bcl2 antagonist [74]. Additionally, in vivo it was shown to suppress the growth in nude mice of human breast tumor xenografts [74]. This suggests that selective inhibition of constitutive $\beta 5$ subunits may have therapeutic efficacy for the treatment of solid tumors.

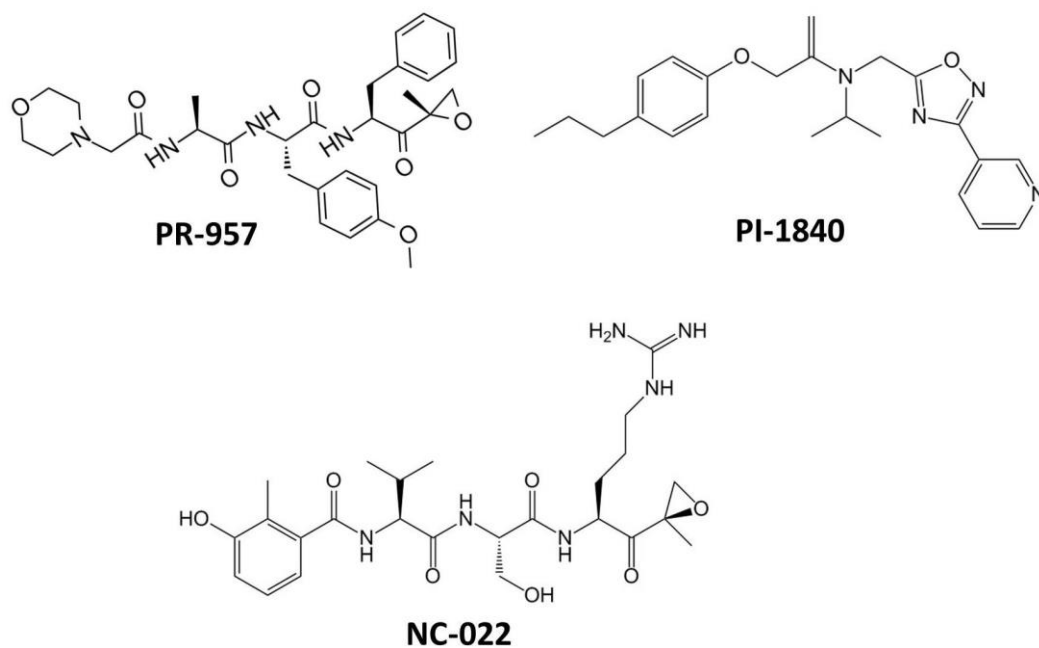


Figure 7 Chemical structures of subunit specific proteasome inhibitors. PR-957 targets immunoproteasome subunit $\beta 5i$; PI-1840 targets constitutive proteasome subunit; $\beta 5c$; NC-022 specifically targets trypsin-like subunits of immuno and constitutive proteasomes ($\beta 2c/\beta 2i$).

1.10 Basic Crystallographic Theory

1.10.1 Introduction to Crystallography

X-ray crystallography is a biochemical technique for determining the three-dimensional structure of macromolecules at atomic resolution. The process of determining the structure is very labor intensive, and several hurdles must be overcome before arriving at your final protein model. The number of structures solved by either NMR, EM, or X-ray crystallography, and deposited to the Protein Data Bank (PDB) has

increased dramatically, with over 100,000 structures deposited as of June 2014; 89,277 of those being X-ray structures. Due to advancements in computing, the volume of structures solved is expected to continue increasing at an exponential rate.

The first step to obtaining a crystal structure is to have a purified, homogeneous, and soluble protein from which you can obtain protein crystals. Protein crystals are necessary because a single protein molecule would not diffract radiation of high enough intensity. A protein crystal is composed of repeating identical unit cells, and the three-dimensional components can be described by their repetition (Rhodes 1993). Six parameters (unit cell parameters), are used to describe the unit cell: three axial lengths a , b , c and three interaxial angles α, β, γ .

A protein crystal is exposed to x-rays over a small angle (typically 0.1-1 degree), and the diffracted x-rays are recorded by a detector, with one frame corresponding to a single rotation of the crystal. Diffraction is a direct consequence of Bragg's law:

$$2d_{hkl}\sin\theta = n\lambda$$

where θ is the angle of the impinging x-ray, n is an integer, λ is the wavelength of the incident x-ray, and d_{hkl} is the distance between two parallel planes of lattice points. When Bragg's Law is satisfied, constructive interference occurs, producing the reflections, or spots, on the detector. An entire data set is collected when all unique reflections have been determined, by rotating the crystal around its center, resulting in several successive frames.

For every such diffracted x-ray that produces a diffraction spot, there is an associated amplitude and phase angle (Rupp, 2005). The amplitude and phase angle for

the diffracted x-rays represent the Fourier transform of the scattering density (electrons, in this case). A Fourier transform is a mathematical expression used to describe the harmonics of any repeating function, for example the repeating density of electrons from identical molecules that make up a crystal.

Electron density is obtained through translation of the diffraction data to obtain a structure factor equation [75]. The structure factor, F_{hkl} , is a function of the electron density distribution in the unit cell. The equation is a Fourier series that describes each reflection and contains one term corresponding to the contribution of each atom in the unit cell:

$$F_{hkl} = \sum f_j e^{2\pi i(hx_j + ky_j + lz_j)}$$

where f_j is called the scattering factor of atom j , and its value depends on what the element is (Rhodes, 1993). The contribution of each atom also depends on the amplitude of the contribution, and its position in the unit cell (x_j, y_j, z_j) which establishes the phase of its contribution. Each reflection should be measured multiple times to insure a high degree of completeness and redundancy, thereby reducing systematic and statistical error. The intensity of the diffracted beam ($h\ k\ l$) is proportional to the square of the amplitude of the structure factor F_{hkl} :

$$I_{hkl} = F_{hkl}^2 LP A$$

where I_{hkl} is the integrated data in the form of intensities for each of the Bragg reflections, F is the structure factor, LP is a combined geometry and polarization factor and A is an absorption correction factor. The Fourier transform is used to convert the structure factors to $\rho(x,y,z)$, the desired electron density equation:

$$\rho(x,y,z) = 1/V \sum \sum \sum F_{hkl} e^{-2\pi i(hx+ky+lz)}$$

where $\rho(x,y,z)$ is the electron density at the point with coordinates (x,y,z) in the unit cell, and V is the volume of the unit cell.

A full description of a diffracted ray must obtain three parameters: amplitude (which is proportional to the square root of the reflection intensity), frequency (which is that of the x-ray source), and phase (Rupp, 2005). The phase of each diffracted ray cannot be measured directly, but must be obtained in order to solve the structure, an issue referred to as the phase problem [76].

1.10.2 Phase Problem and Molecular Replacement

There are several methods that can be employed to circumvent the phase problem, including; molecular replacement (MR), multiple isomorphous replacement (MIR), multi-wavelength anomalous dispersion (MAD), single isomorphous replacement (SIR) and single anomalous dispersion (SAD) [77]. In isomorphous replacement methods, a heavy atom is soaked into the protein crystal, thereby changing the scattered intensity. Both a native (non-soaked) and a heavy metal soaked protein crystal must be collected and the intensities compared to gain a phase estimate. In anomalous dispersion methods, only a single crystal is needed and typically selenomethionine is used in place of methionine residues in the protein. Data can then be collected at several wavelengths near the absorption edge (in this case of selenium), resulting in an estimate of the phase, analogous to IR methods.

Molecular replacement is a popular method for solving the phase angle, because it only relies on having a homologous protein, or search model, whose structure is known. The goal in molecular replacement is to orient and position the search model in the unit cell in a way that it coincides with the position of the unknown protein in the crystal. This requires the determination of six parameters; three angles, and three translational elements. There are two main steps which occur while placing the search model in the unit cell; a rotation function which determines the proteins orientation, and a translation function to determine the absolute position of the protein in the unit cell [78].

For successful molecular replacement, the model should cover at least 50% of the total structure and the C_α r.m.s.d between the core model and the structure to be solved should be less than 2\AA [79]. Additionally, most successful molecular replacements share at least 35% sequence identity between the protein of interest and the model [79]. Several computation programs are available for molecular replacement including Molrep [80] and AmoRe [81] of the ccp4 program suite, and Phaser [78] of the Phenix program suite. With the growing number of submitted protein structures in the protein data bank ($> 86,000$ x-ray structures), and the expectation that there is a limitation to the number of possible folds for globular proteins, it is possible that at some point the phase of any new protein structure will be solvable by molecular replacement.

1.10.3 Model Building and Refinement

Model building and refinement is the iterative process of constructing a protein model consistent with the observed data, and proceeds after phasing [82]. The initial calculated electron density map obtained from either molecular replacement or other phasing techniques is biased due to errors introduced during data collection, processing and phasing.

During refinement, structure factors are calculated (F_c) and compared to the observed structure factors (F_o) from the x-ray experiment. As the phases become more accurate during successive rounds of refinement, the calculated and observed structure factors should converge. This allows a means of assessing the quality of the structural model, indicated by a residual factor (R-factor) based on the following equation:

$$R_{factor} = \frac{\sum hkl ||F_{obs}| - |F_{calc}||}{\sum hkl |F_{obs}|}$$

Values of R range from zero, for a perfect agreement of calculated and observed intensities, to about 0.6, the R factor obtained when a set of measured amplitudes is compared with a set of random amplitudes (Rhodes, 1996).

Another measure of the structure quality is the free R-factor, or R_{free} . This is calculated from a subset of reflections that are set aside prior to refinement, and used to assess the agreement between the calculated and observed models. It measures how well the structure predicts a subset of the measured intensities that were not included in the refinement [83].

Aside from R factors, the chemical, stereo-chemical and conformational aspects of the structure are also monitored. Bond lengths and bond angles should fall near their accepted values, and these are assessed based on the root mean square deviations from ideal or accepted values (Rhodes, 2006). Completed models that are well refined should have root mean square deviations of no more than 0.02 Å for bond lengths, and 4° for bond angles (Rhodes, 2006). Monitoring the Ramachandran plot will insure the model meets conformational standards, assuring that the backbone angles Φ and Ψ fall in allowed regions.

2. STRUCTURAL STUDIES OF THE HUMAN 20S PROTEASOME*

2.1 Results and Discussion

2.1.1 Sequence Alignments

The tertiary and quaternary structure of the 20S proteasome is highly conserved from archaea to mammals and it is assumed that its two types of subunits, termed α and β , evolved from a common ancestor protein [16, 19, 21]. The catalytically active subunits of human 20S constitutive and immunoproteasomes display high sequence identities to each other: $\beta 1c/\beta 1i$: 65%, $\beta 2c/\beta 2i$: 59%, $\beta 5c/\beta 5i$: 55%. The sequence identities between corresponding catalytic immunoproteasome subunits of mouse and human are: 90.5% ($\beta 1i$); 89% ($\beta 2i$); 92% ($\beta 5i$), with majority of the amino acid differences residing outside of the active sites.

Although sequence alignments are indicative of an overall conserved fold for all proteolytically active subunits, defined differences in their substrate binding channels give rise to different cleavage preferences in the cCP, iCP and tCp [21]. In particular, hydrophobic amino acids in the unprimed substrate binding pockets of subunit $\beta 1i$ have been predicted to attenuate the caspase like activity and to increase the chymotrypsin

* Part of the data reported in this chapter is reprinted with permission from “Crystal Structure of the Human 20S Proteasome in Complex with Carfilzomib”, Harshbarger, W., Miller, C., Sacchettini, J. (2015), *Structure*, DOI 10.1016/j.str.2014.11.017, Copyright 2014 by Cell Press.

like activity [21], enhancing the generation of peptides with hydrophobic C-termini. Analysis of the alignments between subunits $\beta 2c/i$ and $\beta 5c/i$ give no indication of chemical changes in the substrates binding regions (Figures 8, 9 and 10).

However, as shown in the immunoproteasome crystal structure from mouse, the S53Q amino acid substitution in iCPs results in a larger S1 pocket by stabilizing the Met45 side chain. The more spacious S1 pocket of immunoproteasomes is more suitable for large hydrophobic residues such as phenylalanine, and is suggested to result in increased catalysis rates compared to cCPs [16].

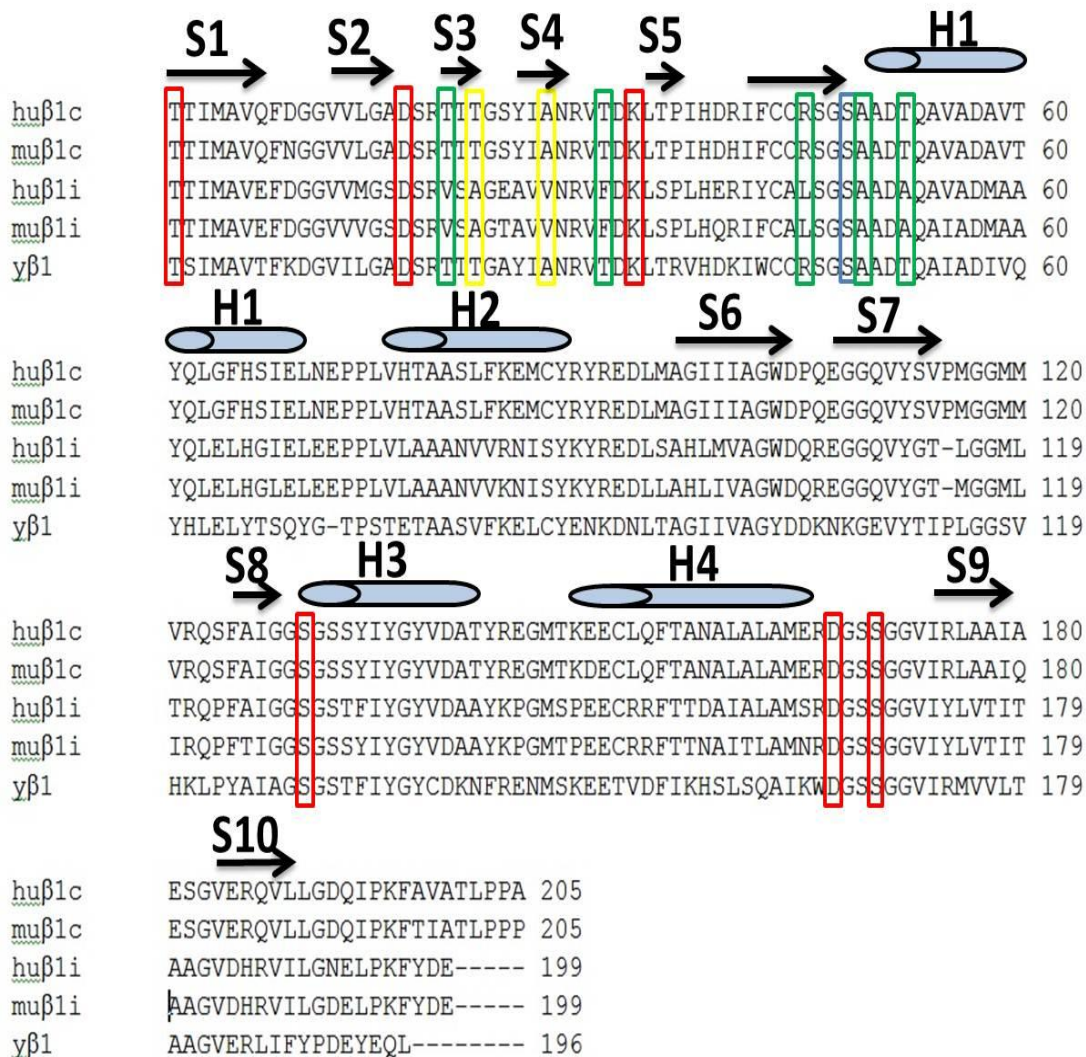


Figure 8 Sequence alignments of proteasomes caspase-like ($\beta 1$) constitutive (c) and immuno (i) subunits from *Homo sapien* (hu), *Mus musculus* (mu), and *T. acidophilum* (y). Secondary structures (S: β -sheet; H: helix) are indicated for the human constitutive subunits. Residues important for the active site are indicated with red boxes. Residues contributing to the substrate-specificity pockets are highlighted by colored boxes: S1 pocket, green; S2 pocket, blue; S3 pocket, yellow.

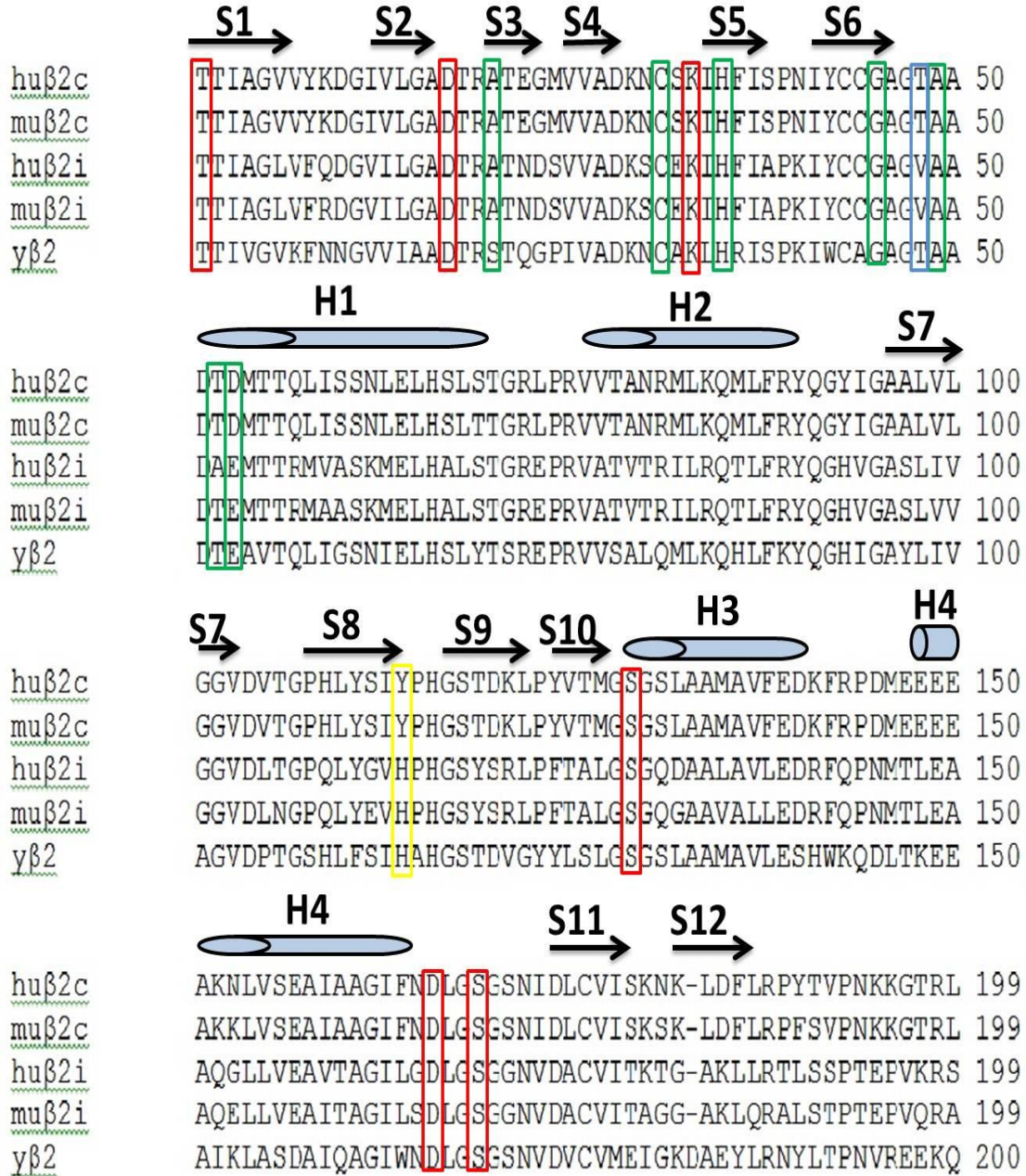


Figure 9 Sequence alignments of proteasomes trypsin-like ($\beta 2$) constitutive (c) and immuno (i) subunits from *Homo sapien* (hu), *Mus musculus* (mu), and *T. acidophilum* (y). Secondary structures (S: β -sheet; H: helix) are indicated for the human constitutive subunits. Residues important for the active site are indicated with red boxes. Residues contributing to the substrate-specificity pockets are highlighted by colored boxes: S1 pocket, green; S2 pocket, blue; S3 pocket, yellow.

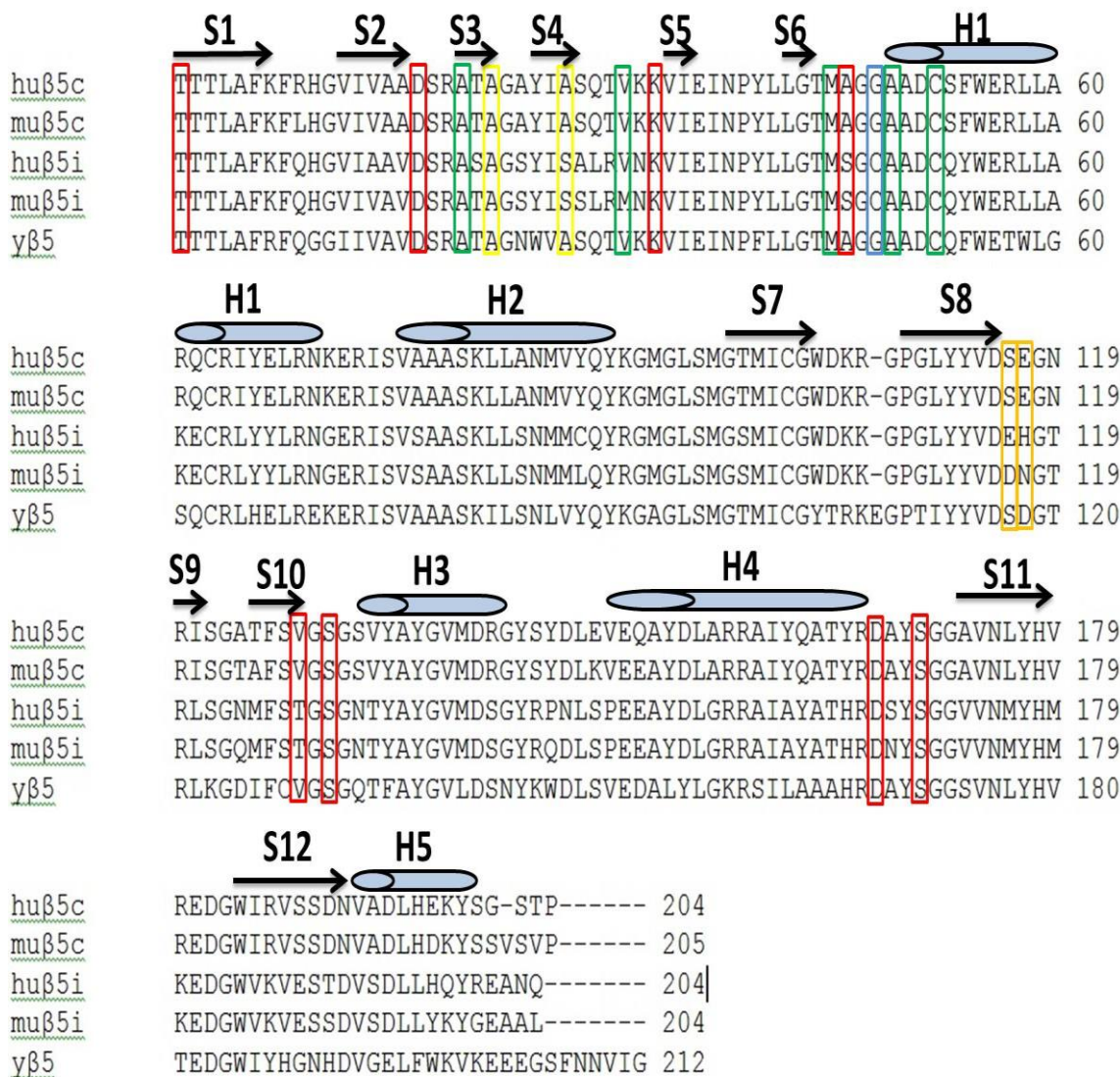


Figure 10 Sequence alignments of proteasomes chymotrypsin-like ($\beta 5$) constitutive (c) and immuno (i) subunits from *Homo sapien* (hu), *Mus musculus* (mu), and *T. acidophilum* (y). Secondary structures (S: β -sheet; H: helix) are indicated for the human constitutive subunits. Residues important for the active site are indicated with red boxes. Residues contributing to the substrate-specificity pockets are highlighted by colored boxes: S1 pocket, green; S2 pocket, blue; S3 pocket, yellow.

2.1.2 Purification, Crystallization and Structure Determination

A strategy was developed in order to purify human 20S proteasome constitutive core particles (cCPs) from red blood cells, and is described in detail in Methods section 2.3.1. The purification strategy was adapted from that used for the purification of the yeast 20S proteasome by Groll et. al. [21] and resulted in greater than 95% purity, proven by SDS-PAGE (Figure 11). To circumvent the large amount of hemoglobin in the initial lysed red blood cells (> 97% of total protein), dialysis with 100 kD dialysis tubing (Spectrum) was carried out overnight against 4L of 10 mM Tris pH 7.5/1mM EDTA prior to affinity chromatography steps. This greatly reduced the concentration of hemoglobin in the protein solution and allowed for easier loading onto an anion exchange column. Early attempts forgoing dialysis resulted in clogged anion exchange columns (HiTrap Q HP, GE Healthcare).

Purified cCPs were concentrated to 30mg/mL, and Mg^{2+} added to a final concentration of 50mM. The Mg^{2+} was added for initial crystal screens because proteasomes from bovine and yeast each crystallized in conditions containing Mg^{2+} [21]. The protein was subjected to sitting drop vapor diffusion crystal trials using a sparse matrix screen including; Emerald Bioscience Wizard I/II; Hampton Research Crystal Screen I/II and Qiagen 2-Methyl-2,4-pentanediol (MPD) screen. The protein:reservoir ratio was 1:1 with a final drop volume of 1 μ L.

The first 20S proteasome crystals identified were in a Wizard condition containing 10% polyethelyene glycol (PEG) 8000 and 0.2M NaCl; however, the crystals did not

diffract beyond 30 Å resolution using a home x-ray source, Rigaku R-Axis IV⁺⁺(Figure 12). In an attempt to optimize the crystal condition to increase resolution, the protein was set up with the addition of the Hampton Research Additive Screen, in a 1:1:0.5 ratio of protein:reservoir:additive, on a hanging drop plate. This was performed manually, with each protein drop containing a different additive. The additive screen is a library of 96 reagents which include multivalent, salt, amino acid, dissociating agent, linker, polyamine, osmolyte, chaotrope, co-factor, reducing agent, chelating agent, polymer, carbohydrate, non-detergent, amphiphile, detergent, non-volatile organic and organic molecules.

Cesium at a final concentration of 1mM changed the crystal morphology from rectangular crystals of approximately 1mm in length, to large squares that were 1mm thick and 3-4mm in length. The addition of cesium improved crystal diffraction to 16 Å. Next, crystals were dehydrated by increasing the percent of Peg 8000 in the reservoir solution by 1% per day to a maximum of 15% Peg 8000. Unfortunately, this method did not improve the diffraction of the crystals, therefore, it was decided to screen for a new condition that may result in better crystal packing and lead to higher resolution diffraction.

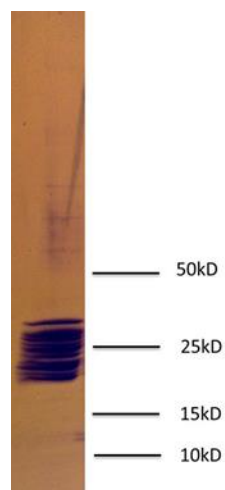


Figure 11 One-dimensional SDS-PAGE analysis of purified human proteasome constitutive core particles

A second round of crystal screening on Intelli 96 well sitting drop plates, without Mg^{2+} in the protein solution, identified several conditions from the Qiagen MPD (2-Methyl-2,4-pentanediol) screen that produced crystals. Each condition contained 40% MPD and a different salt. Diffraction of the various crystals from the home X-ray source identified two conditions with resolution of 3.2 Å; 40% MPD, 0.2M sodium malonate; and 40% MPD, 0.2M sodium formate. Crystals were reproduced by hanging drop vapor diffusion using a 1:1 ratio of protein solution (30 mg/mL, 10mM Tris pH 7.5, 10mM EDTA) and crystallization buffer at 16°C. Crystals appeared in 7-10 days and grew to a maximum size of 0.3 mm (Figure 13). Crystals were then soaked in cryoprotecting buffer (30%MPD, 0.2M Sodium Formate) and flash frozen with liquid nitrogen prior to data collection at Argonne National Labs, beam-lines 19id and 23id. On average, approximately 40 crystals were frozen because many crystals had anisotropic diffraction patterns and were not suitable for data collection.

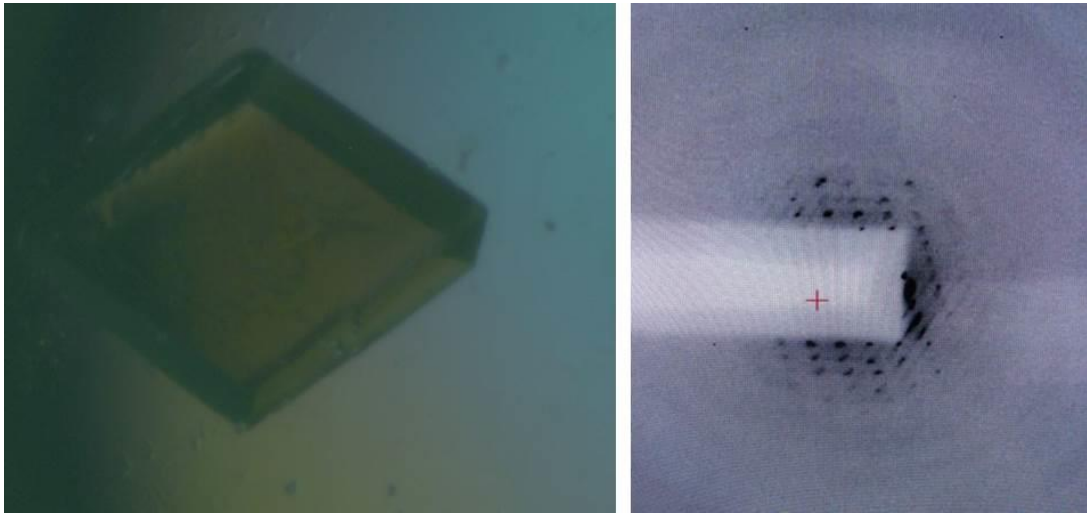


Figure 12 Proteasome crystal and diffraction from condition containing Peg 8K, sodium chloride, and cesium.

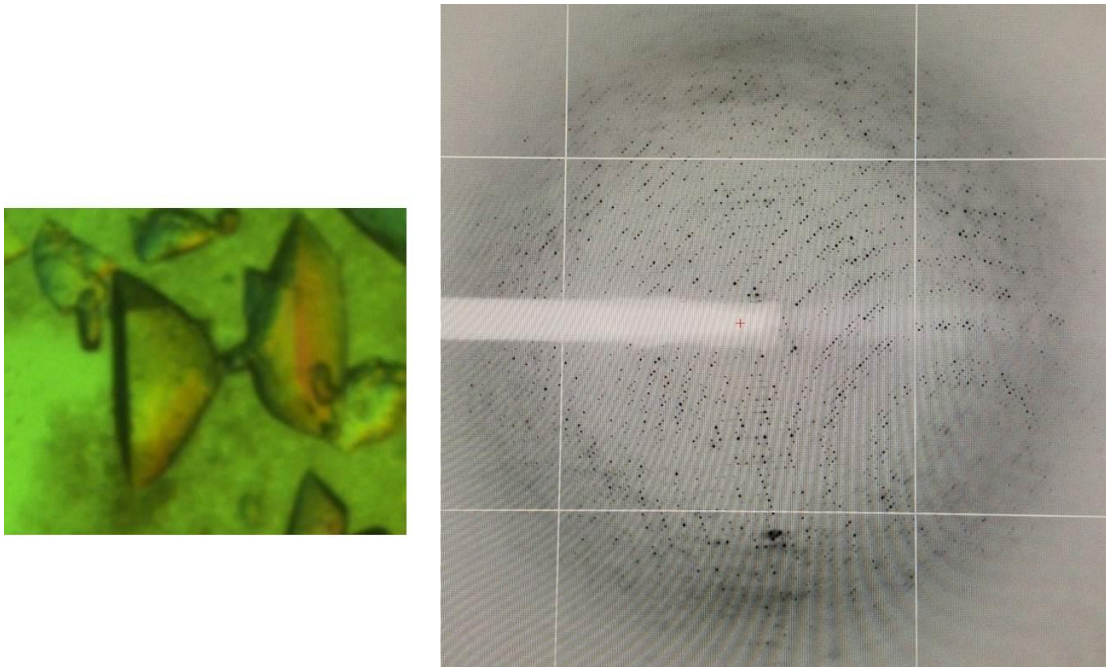


Figure 13 Proteasome crystals and diffraction from condition containing 40% MPD and 0.2M Sodium Formate.

X-ray data for suitable crystals was collected and processed using HKL3000[84], described in Methods at the end of this chapter. The crystals space group was found to be $P2_1$ with unit cell dimensions of $a = 119.2 \text{ \AA}$, $b = 205.1 \text{ \AA}$, $c = 163.1 \text{ \AA}$ and $\beta = 106.4^\circ$. Calculation of the Mathews coefficient in CCP4 [85] indicated a solvent content of 52%, assuming that the asymmetric unit contained only one proteasome core particle. Refinement with Phenix [86] resulted in a final structure with R_{work} and R_{free} values of 0.21 and 0.24, respectively, with root-mean-square (rmsd) bond and angle values of 0.004 \AA and 0.81° , respectively. A Ramachandran plot indicated that greater than 95% of the backbone Ψ and ϕ angles were in acceptable conformation (Figure 14). Outliers occurred in alpha-helices and in regions where the electron density was not well defined. Finalized structure statistics are shown in Table 2.

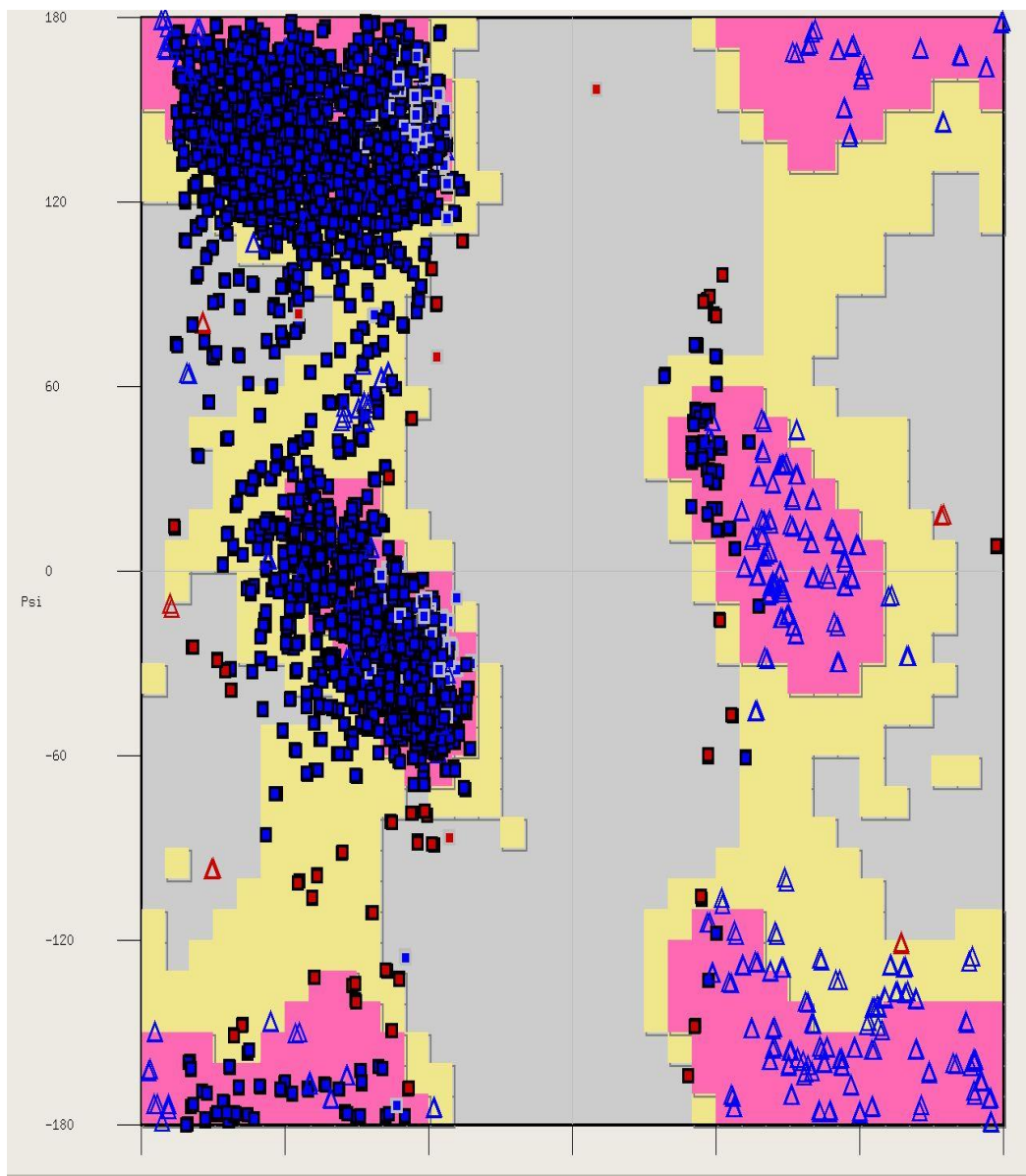


Figure 14 Ramachandran Plot for human 20S proteasome crystal structure.

Table 2 Crystallography Statistics

<u><i>Crystal parameters</i></u>	Human 20S Proteasome	Carfilzomib Bound
Space group	P 2₁	P 2₁
Unit cell		
a (Å)	119.2	172.0
b (Å)	205.1	201.0
c (Å)	163.1	225.6
β	106.4	107.93
<u><i>Data Collection</i></u>		
Wavelength (Å)	0.97918	0.97918
Resolution Range (Å)	50 - 2.6 (2.64 - 2.6)	50 - 2.9 (2.99 - 2.89)
No. observations	621008	1177627
No. unique reflections	228370	321261
Redundancy	2.7 (2.5)	3.7 (3.4)
Completeness (%)	98.2 (91.8)	98.7 (91.7)
I/σ	14.7 (2.1)	9.1 (1.1)
R-merge	6.9 (56.6)	11 (127.3)
<u><i>Refinement</i></u>		
<u>(PHENIX)</u>		
Resolution range (Å)	33.5-2.59	33.5-2.89

Table 2 continued

<u>Crystal parameters</u>	Human 20S Proteasome	Carfilzomib Bound
R_{work}	0.20	0.22
R_{free}	0.24	0.25
Number of atoms	47831	96329
Protein	47646	95258
ligands	-	572
water	185	499
Protein residues	6286	12572
RMSD		
Bond lengths (Å)	0.004	0.004
Bond angles (°)	0.77	0.78
Ramachandran favored (%)	95	95
Average B-factor (Å²)	83.80	75.40

2.1.3 Architecture of the Human 20S Proteasome Constitutive Core Particle

Evaluation of the proteasome in Coot [87] showed that the overall shape is an extended cylinder stretching approximately 150 Å in length with a diameter of about 115

Å (Figure 18). The cylinder consists of twenty-eight total subunits, with seven unique α and seven unique β chains, arranged in rings as $\alpha_{1-7}\beta_{1-7}\beta_{1-7}\alpha_{1-7}$. The organization and overall topology of the human proteasomes constitutive core particle (cCP) subunits is conserved among the eukaryotic proteasomes from yeast (PDB ID 1RYP) and bovine (PDB ID 1IRU).

The entry gate to the central proteolytic chamber of the human cCP is closed by the N termini of the α -subunits, which project into the pore of the α -ring, forming several layers (Figure 19). Access to the core is regulated by ATPase's, such as the 19S regulatory particle, or through an ATP dependent mechanism by association with PA28 [88, 89]. Superposition of the α -ring of the human proteasome with the α -ring from yeast, bovine, and mouse proteasomes illustrates high structural conservation (RMSD C_{α} atoms α -ring < 0.59 Å for all α -rings).

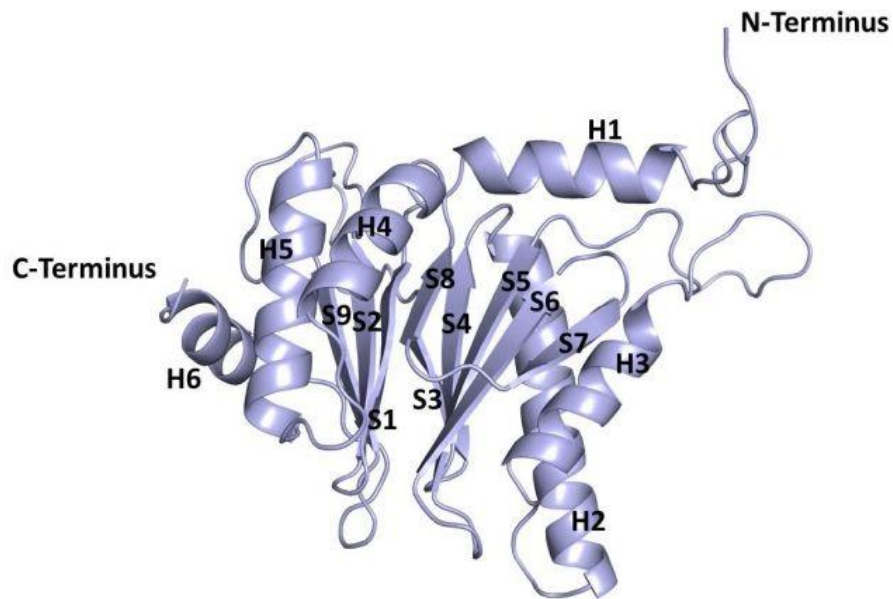


Figure 15 Quaternary structure of α -subunits. Subunit $\alpha 5$ is shown in ribbon representation, with each α -helix labeled as H, and β -sheets are labeled as S.

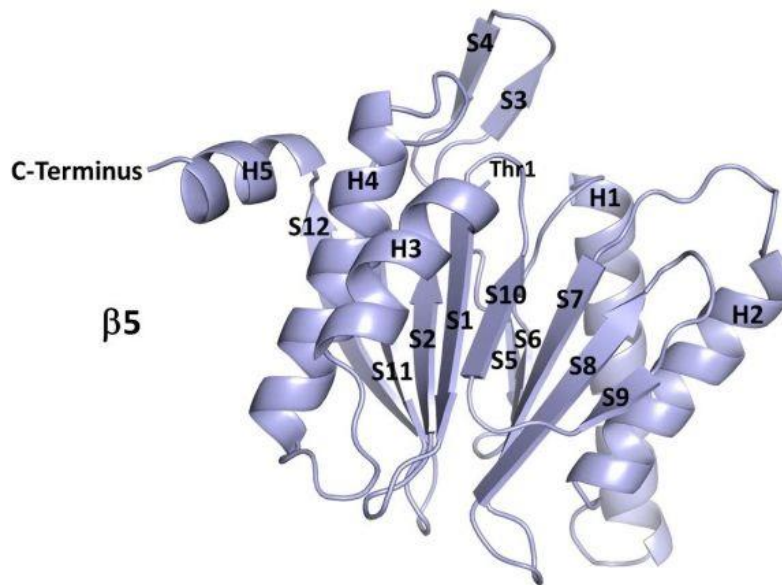


Figure 16 Quaternary structure of β -subunits. All β -type subunits have a similar arrangement of two α -helices, H1 and H2, followed by two sets of five-stranded anti-parallel β -sheets, then two more α -helices, H3 and H4. An additional two-stranded β -sheet is found in the loop connecting β -strands S2 and S5. Subunit $\beta 5$ (shown) has an additional two-turn α -helix at its C-terminal end.

All seven distinct α - and β -subunits have the same general β -sandwich architecture, which is typical for N-terminal hydrolases [25]. In the α -subunits, the β -sandwich consists of three α -helices H1, H2 and H3, followed by two five-stranded anti-parallel β -sheets, then three additional α -helices, H4, H5 and H6 on the opposite side (Figure 15). Differences among the seven α -chains come in the length of the loop regions, and the number of turns in the α -helices. Specifically, subunit $\alpha 5$ has a shorter α -helix, H3, by two turns, followed by a longer insertion loop of nineteen residues from Ala120 to Val139, as opposed to only nine residues in the other six α -type subunits. Subunits $\alpha 3$ and $\alpha 4$ have a six-turn α -helix, H6, whereas they are shorter by one to three turns in the other α -subunits. The C-terminal ends of each α -subunit after helix H6 are disordered and vary in length, containing mainly charged residues.

The β -type subunits have an arrangement of two α helices, H1 and H2, followed by two sets of five-stranded anti-parallel β -sheets, then two more α -helices, H3 and H4. The anti-parallel β -sheets are composed from β -strands S1, S2, S10, S11, and S12 on one side, and by β -strands S5 through S9 on the opposite side (Figure 16). An additional two stranded anti-parallel β -sheet is formed in the loop connecting β -strands S2 and S5. Subunit $\beta 5$ forms an additional two turn α -helix, H5, at the C-terminal end from residues Val192 to Ser200 (Figure 17).

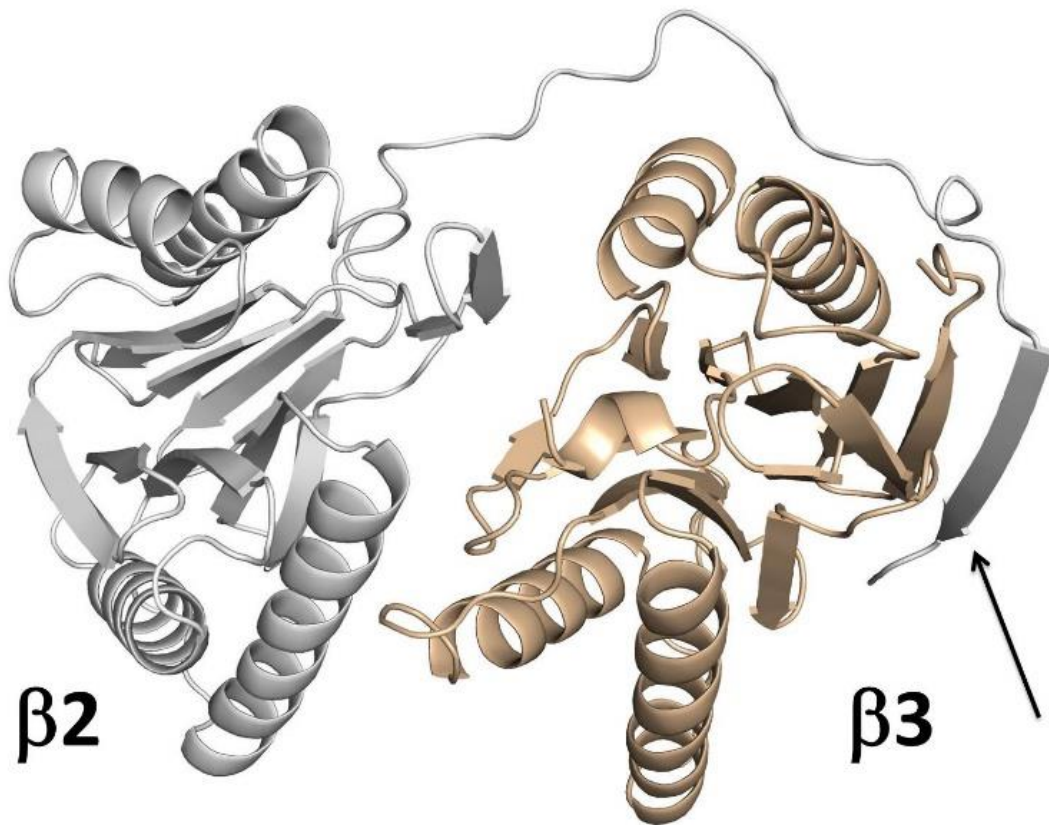


Figure 17 Conserved C-terminal loop of $\beta 2$ subunits. The C-terminus of subunit $\beta 2$, from residue Leu186 to Glu220, wraps around subunit $\beta 3$ and forms an additional strand in its anti-parallel β -sheet, S12. The black arrow indicates the C-terminus of $\beta 2$.

The N-terminal residues of all β -type subunits are buried in the core of the protein, whereas the C-terminal ends extend toward the second β -ring. The exception is the C-terminus of subunit $\beta 2$, from residue Leu186 to Glu220, which wraps around subunit $\beta 3$ and forms an additional strand in its anti-parallel β -sheet, S12. This unique C-terminal loop is conserved among species and is required for CP assembly or stability [90].

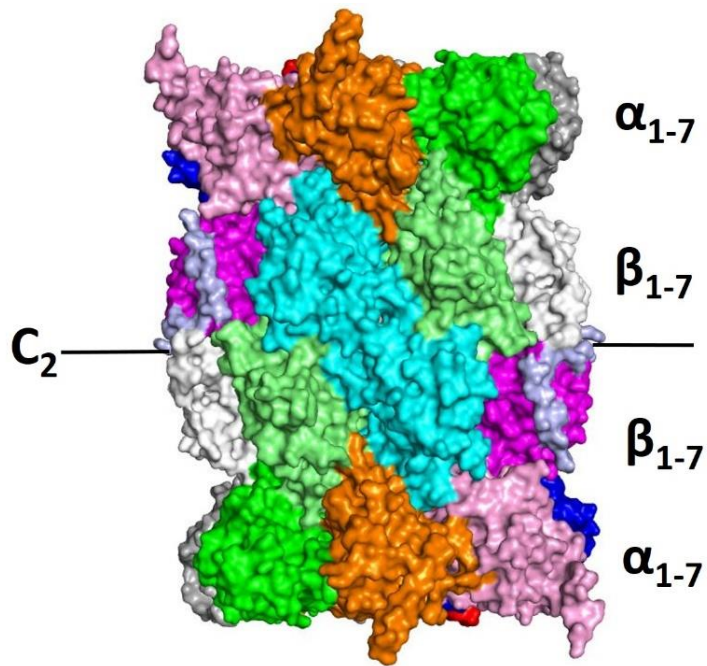


Figure 18 Surface representation of the human 20 proteasome. Subunits are colored according to their symmetrical counterpart.

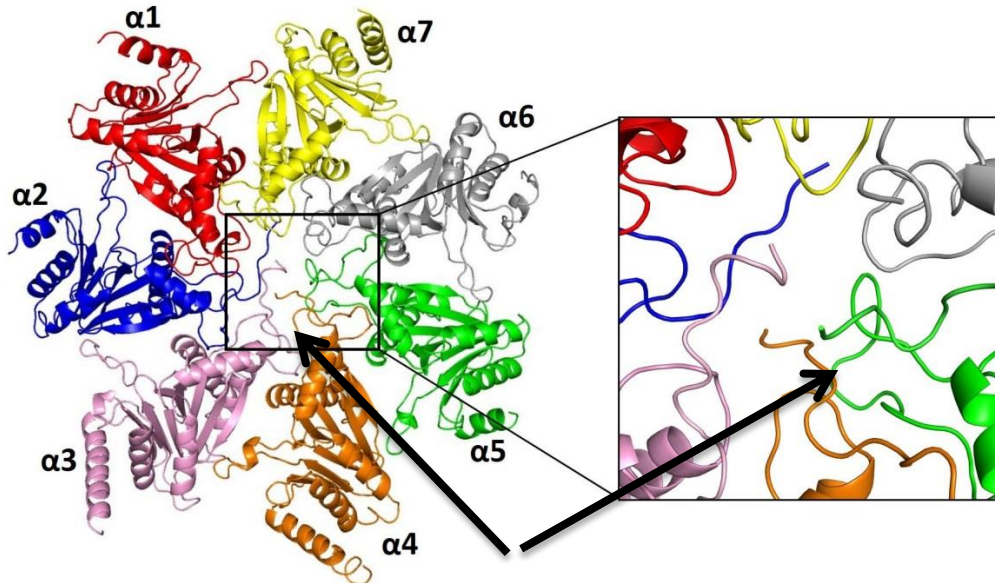


Figure 19 Ribbon representation of top view of the α -subunits. The arrows indicates the entry to the catalytic chamber, which is closed off by the N-termini of the α -subunits. The 20S core can be activated by proteins such as the 19S regulatory particle, which inserts short C-terminal peptide segments into shallow pockets on the α -ring surface, causing structural rearrangements. Each α -subunit is colored uniquely.

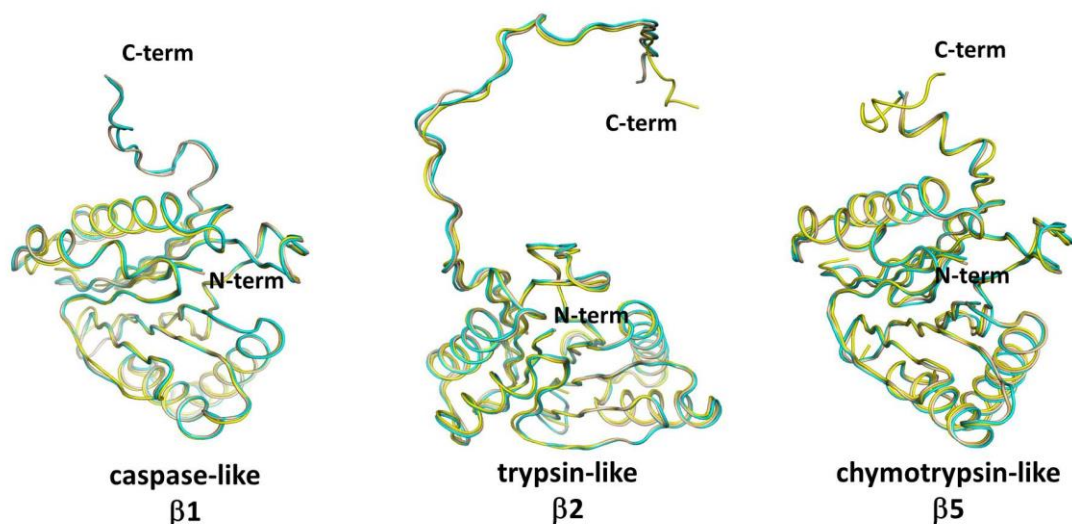


Figure 20 Structural alignment of proteasome catalytic subunits. Human (wheat), yeast (yellow) and mouse (teal).

Superposition of individual catalytic β -subunits of the human cCP with corresponding subunits of the mouse cCP, indicates that they adopt very similar folds (RMSD C_{α} atoms 0.38 Å for $\beta 1$, 0.40 Å for $\beta 2$ and 0.32 Å for $\beta 5$), as does superposition with respective yeast catalytic subunits (RMSD C_{α} atoms 1.5 Å for $\beta 1$, 0.65 Å for $\beta 2$, and 1.6 Å for $\beta 5$) (Figure 20). This was expected, based on the high level of amino acid sequence conservation (> 95% with mouse cCP). The amino acid differences between active constitutive subunits of human and mouse proteasomes are surface-exposed, and PISA analysis indicates that the number and types of contacts between adjacent subunits differ; however, their impact on stability or half-life of the CPs cannot be determined from a structural perspective.

2.1.4 β Subunit Post-Translational Processing and Catalytic Residues

Five of the seven β -type subunits, including the three catalytic subunits (β 1, β 2, and β 5) are synthesized as pro-proteins, which are proteolytically processed to their mature forms found in the assembled core particle [91]. The mature, cleaved form of the proteasome was found for each β subunit in the human CP, indicated by no additional density at the N-terminus of each chain.

Subunit β 6 is cleaved by 28 residues between Leu-1 and Arg1, and the β 7 subunit is cleaved by 45 residues, between Arg-1 and Thr1, exposing an N-terminal threonine. Despite having an N-terminal Thr1, no definitive proof that subunit β 7 is catalytically active has been established, although it was proposed to have hydrolase activity in the bovine crystal structure [92].

The nucleophilicity of Thr1 in each of the catalytic sites is enhanced *via* an elaborate network of hydrogen bonding with and between surrounding residues. The active sites are highly conserved and can be summarized by the chymotrypsin like site as follows: Thr1N is hydrogen bonded to Tyr169 O (2.9 Å), Ser130 O γ (3.1 Å), and Ser170 O γ (3.1 Å); Lys33 N ζ makes hydrogen bonding interactions with Thr1O γ (2.9 Å), Arg19 O (2.7 Å), and Asp17 O δ 1 (2.7Å); Asp17 O δ 2 is hydrogen bonded to Ser18N (3.4 Å), and Gly171N (3.4 Å); and finally Asp17 O δ 1 forms an additional hydrogen bond to Thr2N (2.9 Å) (Figure 21).

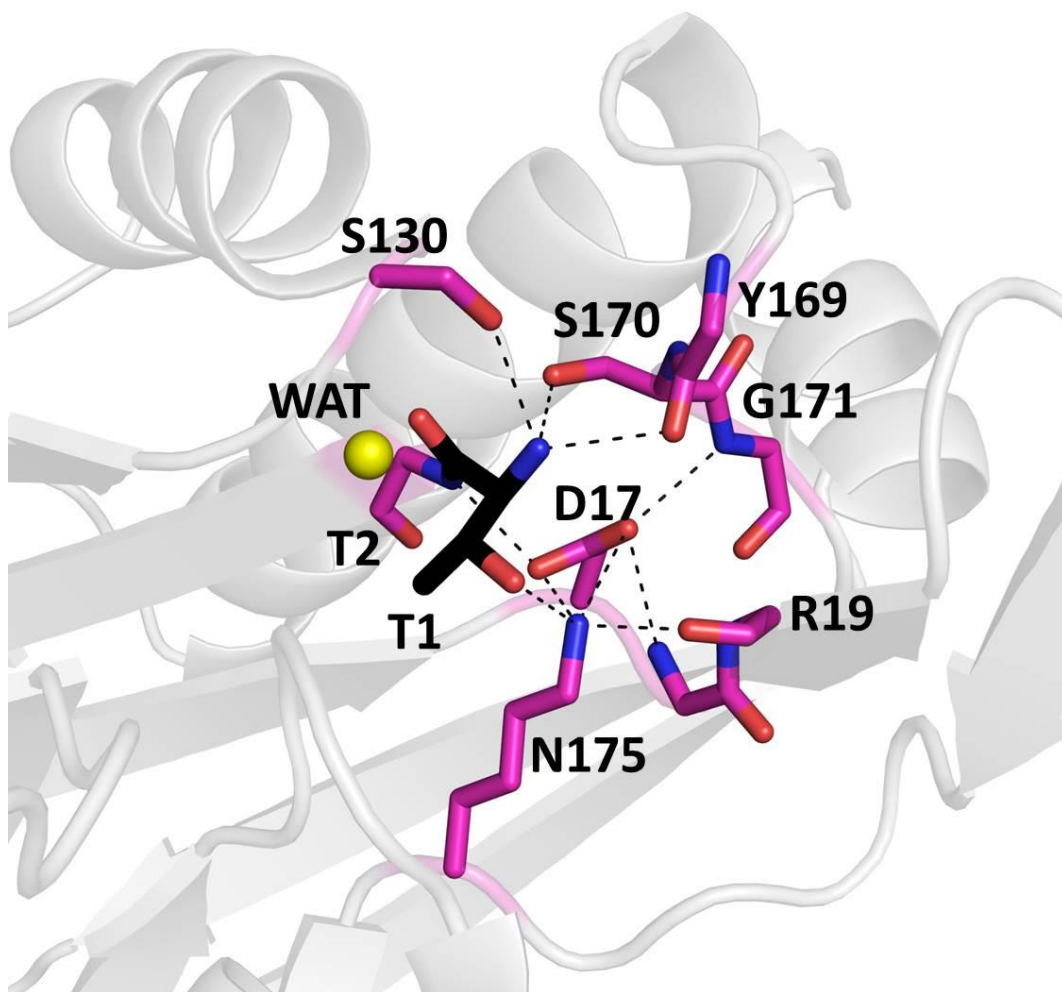


Figure 21 Active site of catalytic subunits $\beta 1$, $\beta 2$, and $\beta 5$. Residues involved in catalysis are shown in magenta, and are conserved in all three distinct catalytic subunits. Hydrogen bonds are shown as black dashed lines, and the unique network increases the nucleophilicity of Thr1. Thr1 is shown in black, and the active site water, WAT, is shown as a yellow sphere. The water may be involved in intramolecular autolysis and substrate proteolysis by mediating proton transfer between Thr1O γ and Thr1N.

2.1.5 Cleavage Preferences of Catalytic Subunits

In subunit $\beta 5$, the residues in the S1 pocket that give rise to chymotrypsin-like activity are Ala20, Met45, Ala49, and Cys52. The S1 pocket in this site was found to be relatively shallow due to the side chain of Met45, located at the back of the pocket and not interacting with any surrounding residues. The S1 pocket of chymotrypsin like sites is able to accommodate hydrophobic residues such as alanine, valine, or tyrosine, consistent with cleavage of the known fluorogenic substrate Suc-LLVY-AMC (*N*-Succinyl-Leu-Leu-Val-Tyr-7-amino-4-methylcoumarin) [93].

The S1 pocket in the caspase like site ($\beta 1$) is also shallow due to the side chains of residues lining the pocket; Thr20, Thr31, Thr35, Arg45, Ala49, Gln53 and $\beta 2$ Asp120. Positively charged Arg45, at the base of the S1 pocket facilitates accommodating acidic residues, and gives rise to the caspase-like activity. Subunit $\beta 1$ can also cleave after some hydrophobic residues, thereby also contributing to the branched-chain amino acid preference activity of the proteasome [94, 95].

Subunit $\beta 2$ is attributed with trypsin like activity. The base of the S1 pocket contains Gly45, resulting in a larger S1 pocket than $\beta 5$ or $\beta 1$ sites. Residues Cys31, His35, Asp53, and $\beta 3$ Cys129 give the region an overall acidic character, leading to a preference for basic residues such as Lysine or Arginine in the P1 position of substrates. This is consistent with trypsin like sites ability to cleave the known fluorogenic substrate Boc-LLR-AMC [93] (Figure 22).

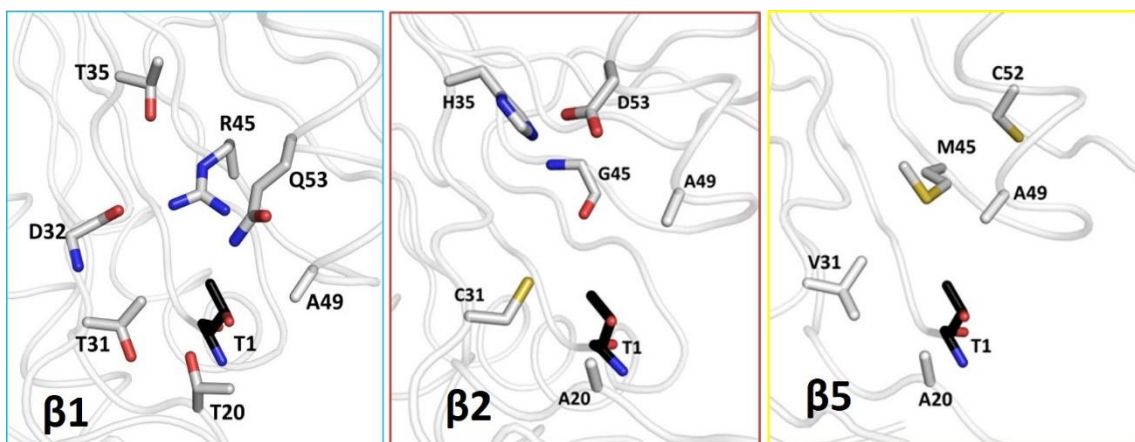


Figure 22 S1 specificity pockets of the catalytic β -subunits.

2.2 Materials and Methods

2.2.1 Proteasome Purification

The constitutive 20S proteasome was purified from human erythrocytes. Red blood cells, isolated from 500 mL of whole blood and suspended in saline solution, were obtained from Gulf Coast Regional Blood Center, College Station, TX. Cells were pelleted by centrifugation for 10 minutes at 35,000 x G in a JA-25.50 fixed-angle Beckman rotor, the supernatant discarded, and the red blood cells resuspended in lysing buffer (20mM Tris pH 7.0, 50mM MgCl₂, 1mM EDTA). The protein was then extracted with French press and the lysate centrifuged at 53,300 x G for 1 hour in the JA-25.50 rotor. The supernatant was dialyzed overnight at 18 °C in 100kD dialysis tubing (Spectrum Labs) against 4L of 10mM Tris-HCl pH 7.5, 1mM EDTA. The dialyzed protein was again centrifuged at 53,300 x G for 1 hr and the supernatant filtered with a 0.2 μ m syringe filter

before being loaded onto a HiTrap QHP anion exchange column (GE Healthcare), pre-equilibrated with buffer A (10 mM Tris pH 7.5, 1 mM EDTA, 350 mM NaCl). The column was washed with at least 5 column volumes of buffer A, then bound protein was eluted with a gradient of Buffer B (0-100%) (10mM Tris pH 7.5, 1mM EDTA, 700 mM NaCl). Fractions were analyzed by SDS-PAGE and those containing proteasome were pooled together and diluted 10 fold in Buffer C (10mM Tris pH 7.5, 1mM EDTA). Solid ammonium sulfate (> 99% pure from Sigma) was added to the pooled protein fractions at a final concentration of 40% saturation (w/v) while stirring with a magnetic rod on a stir plate at 18 °C. The solution was centrifuged for 10 minutes at 1600 x G in a bench top Eppendorf 5702 centrifuge, and the supernatant filtered with a 0.2 µm syringe filter before loading onto a HiTrap Phenyl Sepharose FF column (GE Healthcare), pre-equilibrated with Buffer D (10mM Tris pH 7.5, 1mM EDTA, 1.2M ammonium sulfate). The column was washed with 5 column volumes of Buffer D, then column bound protein was eluted with a gradient of Buffer C (0-100%). Fractions were analyzed by SDS-PAGE and those containing proteasome were pooled and dialyzed overnight at 18 °C in 100 Kd dialysis tubing against Buffer C. Last, the protein was concentrated to 30 mg/mL for crystallization trials, using Spin-X UF 200 20mL concentrating tubes with a 100Kd MWCO purchased from VWR.

2.2.2 Crystallization of Human 20S Proteasome Constitutive Core Particle

Purified proteasome (section 2.2) at a concentration of 30 mg/mL was screened for crystallization conditions in 96-well sitting drop Intelli-Plates (Art Robbins Instruments). The following crystallization buffers were screened: Emerald Bioscience Wizard I/II; Hampton Research Crystal Screen I/II; Qiagen 2-Methyl-2,4-pentanediol (MPD) screen. Screening plates were set up using a Phoenix liquid handling robot (Art Robbins Instruments) to deliver 50 μ L of screening buffer to the appropriate reservoirs, followed by a mosquito (TTPlabtech) which delivered 1 μ L of reservoir solution and 1 μ L of protein onto the sitting drop platform. Plates were manually sealed with ClearSeal Film (Hampton Research), and placed on a vibration free table at 16°C. The plates were checked for the formation of crystals at time intervals of 24 hours, 48 hours, 96 hours, and then once weekly.

Crystals grew in the crystallization buffer of 0.2 M sodium formate, 40% MPD and were optimized by hanging drop vapor diffusion. Hanging drops consisted of either 1:1 or 1:2 ratios of protein to crystallization buffer and placed on a vibration free table at 16°C. Crystals appeared in 7-10 days and grew to a maximum size of 0.3 mm. Crystals were then soaked in cryo-protecting buffer (30%MPD, 0.2M Sodium Formate) and flash frozen with liquid nitrogen prior to data collection. Data was collected using synchrotron radiation at the Advanced Photon Source (Argonne National Labs, IL) with $\lambda = 0.97918$ Å, at beam-lines 23iD and 19iD.

2.2.3 Data Collection, Processing and Structure Determination

X-Ray intensities were indexed and scaled using HKL3000 [84] in the $P2_1$ space group. Data resulted in 621008 reflections, with 228370 unique reflections and overall completion of 98.2%. Calculations of the Matthews coefficient using ccp4 [96] gave an approximate solvent content of 50%, assuming one core particle in the asymmetric unit. The phase of the data was solved by molecular replacement using the default settings for the program Phaser [78] of the Phenix program suite [82], with the coordinates of the core particle from *Bos Taurus* (PDB ID 1IRU) as a starting model, and a sequence identity of 100%, because the *Bos Taurus* structure contains the human proteasomes amino acid sequence in place of its own, due to the amino acid sequence not being known at the time of its publication. This yielded a single solution which was used for structure refinement.

The proteasome structure was refined using Phenix Refine version 1.82 [86]. Three cycles of coordinate, restrained atomic displacement parameters (ADP), translation-libration-screw-rotation (TLS), and occupancy refinement were used with default settings, along with 2-fold non-crystallographic symmetry (NCS) averaging (cartesian). This was followed by manual real space-refinement in COOT [87], then three more rounds of refinement in Phenix with the above settings as well as individual B-factor, and TLS. Final R_{work} and R_{free} values were 0.21 and 0.24, respectively. Bond and angle values were 0.004 Å and 0.81°, and the Ramachandran plot indicated that 96% of backbone dihedral angles were in allowed conformations. The finalized structure was deposited and accepted in the Protein Data Bank under accession numbers 4R3O.

2.2.4 Bioinformatics Tools

Protein sequences were obtained from the “Universal protein resource” (www.uniprot.org). Alignment of protein homologs was performed with ClustalW2 (www.ebi.ac.uk/Tools/msa/clustalw2/).

3. CARFILZOMIB BOUND HUMAN 20S PROTEASOME*

3.1 Introduction

Carfilzomib is the result of medicinal chemistry efforts focused on increasing the potency and chymotrypsin-like selectivity of epoxomicin. It is a synthetic tetrapeptide epoxyketone, and was approved in July 2012 for use in patients with multiple myeloma who have received at least two prior therapies. Structural information on the yeast core particle alone and in complex with epoxomicin strongly supported its development [61].

Drug discovery has also been facilitated by structure activity relationship studies which show that subunit selectivity can be achieved by varying the substituents of a drug at the P1 position, as has been shown for lactacystin [97] and salinosporamides [98], and by varying the P3 substituent, shown by vinylsulfone-inhibitors (Groll et al., 2002; Nazif and Bogyo, 2001). Additionally, the substitution of hydrophobic or aromatic groups in the P4 position has been shown to lead to trypsin like selectivity [99, 100]. Crystal structures of the yeast:epoxomicin (Groll et al., 2000) and yeast:PR-957 (Groll et al., 2012) proteasome complexes have shown that the epoxyketone pharmacophore inhibits *via* covalent adduct, forming a morpholino ring with Thr1N and Thr1O γ . The ring formation occurs in a two-step process. First, a reversible hemiketal is formed by attack

* Part of the data reported in this chapter is reprinted with permission from “Crystal Structure of the Human 20S Proteasome in Complex with Carfilzomib”, Harshbarger, W., Miller, C., Sacchettini, J. (2015), *Structure*, DOI 10.1016/j.str.2014.11.017, Copyright 2014 by Cell Press.

of Thr1O γ with the ketone group of CFZ. Next, the epoxide is nucleophilically attacked by the amine group of Thr1, resulting in an irreversible intermolecular cyclization, which yields the morpholino ring (Figure 23) [16].

The morpholine ring system that results is irreversible and leads to a higher degree of specificity for the proteasome, with minimal off-target effects thereby decreasing side effects such as neurotoxicity. Additionally, the irreversible nature of CFZ inhibition may be a contributing factor to its ability to overcome both primary and secondary resistance to bortezomib [101]. Despite this, some drawbacks of CFZ are its relatively poor bioavailability, and a half-life of less than 30 minutes [66, 102], making it imperative that next generation proteasome inhibitors be designed that have greater stability and enhanced cellular uptake.

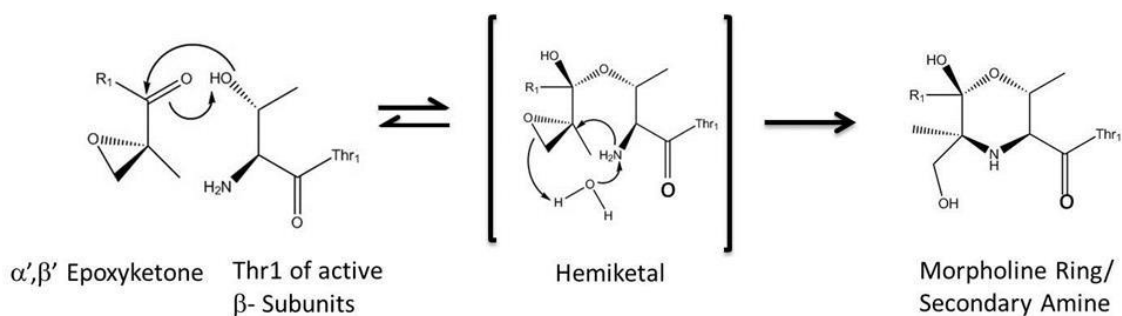


Figure 23 Proteasome reaction mechanism against epoxyketone inhibitors. The carbonyl carbon atom is attacked by the proteasomes Thr1O γ nucleophile. The formed hemiketal intermediate can either dissociate to deliberate Thr1O γ or irreversibly cycle under the formation of a secondary amine (morpholine ring system) involving Thr1N.

CFZ primarily targets the chymotrypsin-like activity of the constitutive proteasome (β 5c), with five times the selectivity over the immuno proteasome subunit β 5i

(IC₅₀ of 6nM against β 5c and 33nM against β 5i) [66]. Caspase-like and trypsin-like sites were shown to be less inhibited, with IC₅₀ values of 2400 nM and 3600 nM, respectively [66]; however, the chymotrypsin-like inhibition is attributed with the drugs anti-cancer effects.

The structural investigation of the proteasome:carfilzomib complex presented in this chapter aims to elucidate the answers to two questions: what are the structural features that give rise to chymotrypsin-like selectivity over caspase-like and trypsin-like sites; and what is the cause for increased selectivity of constitutive proteasome chymotrypsin-like sites over immunoproteasome?

3.2 Results and Discussion

3.2.1 Carfilzomib Binding

In order to determine the molecular basis for the chymotrypsin-like specificity of carfilzomib, human proteasome crystals were soaked with 5mM of the drug for 12 hours before exposure to synchrotron radiation for data collection. Two crystal conditions produced diffraction to greater than 3Å; sodium formate/40% MPD, and sodium malonate/40% MPD. The condition containing sodium formate as the salt produced the more durable crystals, because it was found that the malonate crystals would quickly dissolve once the hanging drops were exposed to air. This made it very difficult to soak

the crystals with inhibitor, therefore, sodium formate crystals were chosen for inhibitor bound structure determination.

Data was collected and processed using HKL300, as described for the apo proteasome structure, at Argonne National Labs, Chicago, IL. Molecular replacement with the program Phenix was used to solve the phase, using the coordinates of the proteasome core particle from *Bos Taurus* (PDB ID 1IRU) as a starting model. Refinement with Phenix led to R_{work} and R_{free} values 0.23 and 0.2, respectively (described in Methods at the end of this chapter). Bonds and angles were also good, with values of 0.002 Å and 0.6°, respectively. Carfilzomib was generated using PRODRG, and fit into the Fo-Fc density manually in COOT. Coordinates for finalized structures were confirmed to have good stereochemistry from the Ramachandran plot, with greater than 98% residing in allowed regions (Figure 25).

The difference electron density map for CFZ soaked crystals indicated that CFZ had targeted all four chymotrypsin-like and trypsin-sites, and three of the four caspase-like sites in the asymmetric unit, through covalent binding to Thr1 (Figure 24). Subunit $\beta 7$, which was proposed to have hydrolase activity, did not have the inhibitor bound, therefore whether it is an active subunit remains elusive.

Carfilzomib modified Thr1 in each inhibited subunit *via* a dual covalent adduct with its C-terminal epoxyketone (Figures 26-28). A morpholino ring is formed with Thr1N and Thr1O γ , which is consistent with proteasome inhibition by other epoxyketone inhibitors such as epoxomicin (Groll et al., 2000) and PR-957 (Groll et al., 2012). The electron density map reveals an ordered water molecule located 3.5 Å from Ser130N in

all three catalytic sites, which forms a 3.0 Å hydrogen bonding interaction with the P1 carbonyl. The water molecule is also suggested to be involved in both intramolecular autolysis and substrate proteolysis by mediating proton transfer between Thr1O γ and Thr1N [21].

The C-terminal end of CFZ completes an anti-parallel β -sheet in each of the proteasome binding channels with the conserved amino acids Thr21, Gly47, and Ala49 (Figures 29, 30, 31). Although the residues that form the hydrogen bonds are conserved in all three catalytic centers, the hydrogen bonding distances differ between subunits. The hydrogen bonds can be summarized for the caspase-like sites as follows; Thr21N (3.3 Å), and Thr21O (2.9 Å) of β -strand S3 on one side of the binding pocket, and Gly47O (2.7 Å) and Ala49N (3.4 Å) from the loop connecting S6 and H1 on the other side of the binding pocket. In trypsin-like sites the hydrogen bond distances are; Thr21N (3.3 Å), Thr21O (2.9 Å), Gly47O (3.0 Å), and Ala49 N (2.7 Å); and in the chymotrypsin-like sites, Thr21N (3.2 Å), Thr21O (2.9 Å), and Gly47O (2.7 Å).

A second water molecule (WAT) found in the chymotrypsin-like site makes a 2.5 Å hydrogen bond with the P3 carbonyl oxygen of carfilzomib, replacing the hydrogen bond from Ala49 N, and results in a shorter hydrogen bond distance by \sim 0.5 Å. The water is also hydrogen bridged with A49N (3.0 Å), A50N (2.9 Å), and β 6D125O δ 2 (3.2 Å). The water molecule was reported in the yeast:bortezomib crystal structure, but was not described in the mouse immunoproteasome or constitutive proteasome complexes with the epoxyketone inhibitor PR-957, therefore whether or not it is important for chymotrypsin-like selectivity of inhibitors is ambiguous.

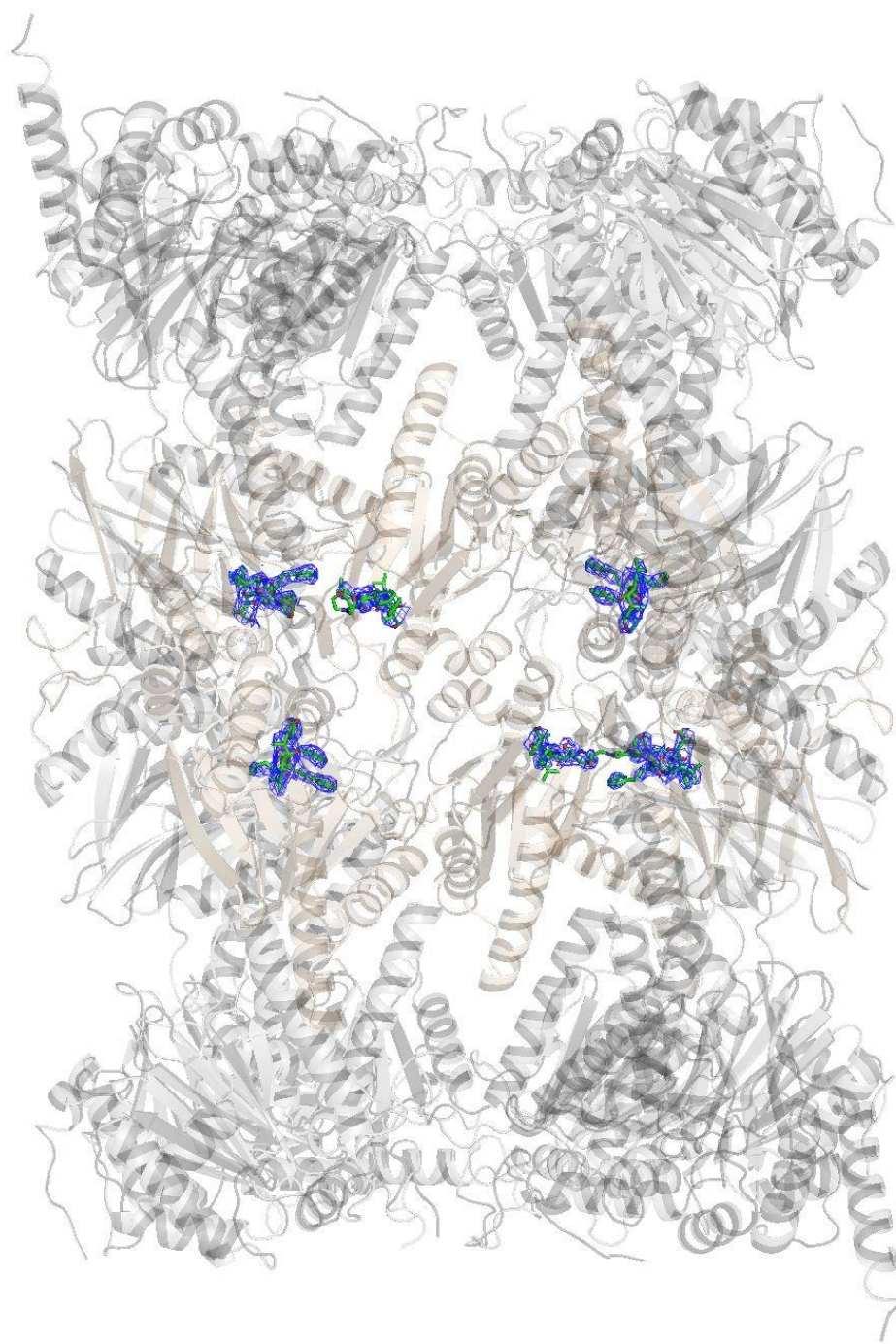


Figure 24 View of the inner cavity of the human 20S proteasome with carfilzomib bound. The $\alpha 4$ and $\beta 4$ -subunits have been removed to give a view of the core of the 20S proteasome. Carfilzomib, shown as green sticks, is bound to all six active sites.

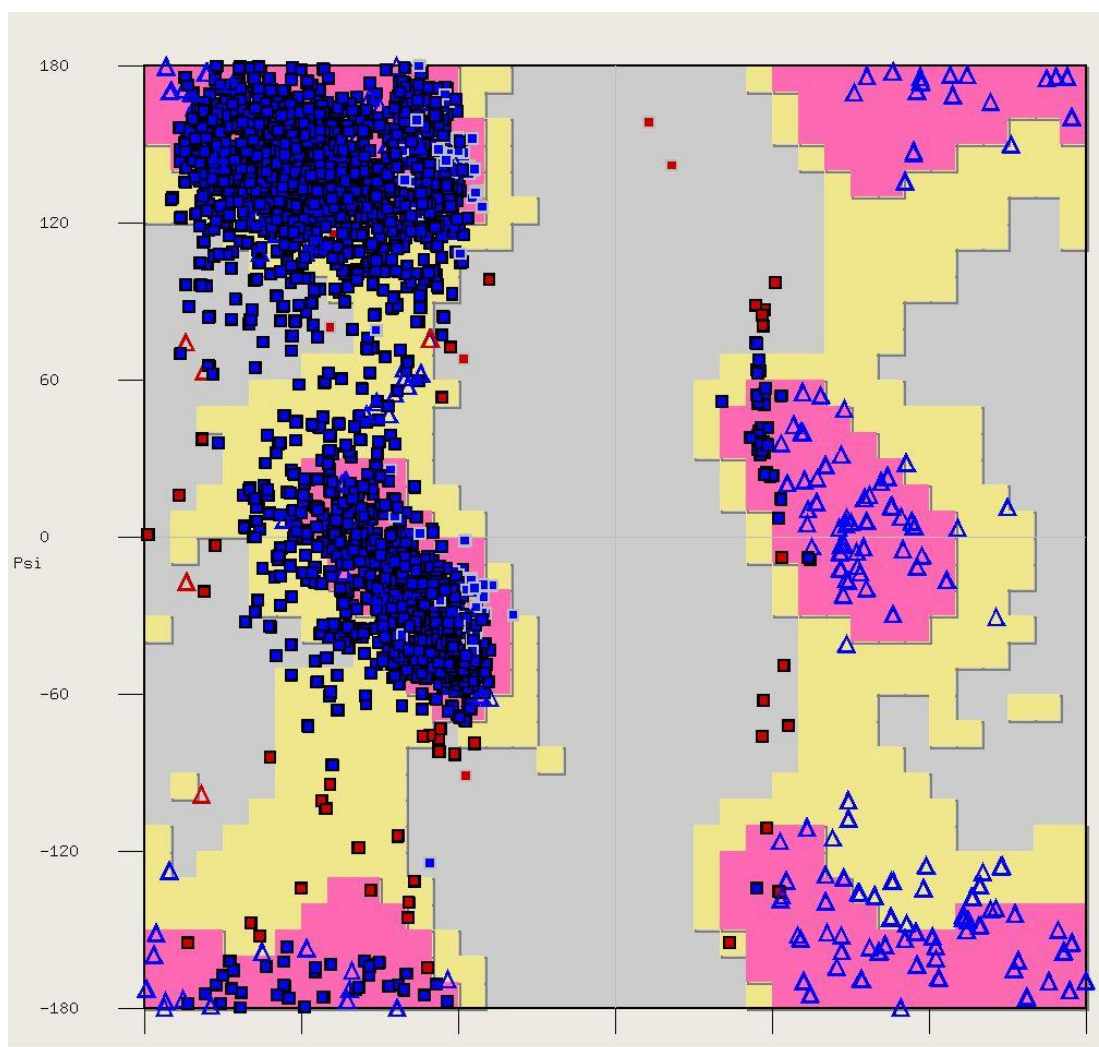


Figure 25 Ramachandran plot for carfilzomib bound human 20S proteasome.

The N-terminal morpholino moiety of CFZ is not engaged in any interactions with the protein in any of the catalytic sites. However, in chymotrypsin-like and trypsin-like sites the backbone of carfilzomib is curved approximately 45 degrees relative to caspase-like sites (Figure 32). This change in orientation is due to steric effect caused by the side chains of β 3Glu106 and β 6Tyr107 in trypsin-like and chymotrypsin-like binding regions, which force the morpholino ring to turn towards β -sheet S3 to avoid clashing.

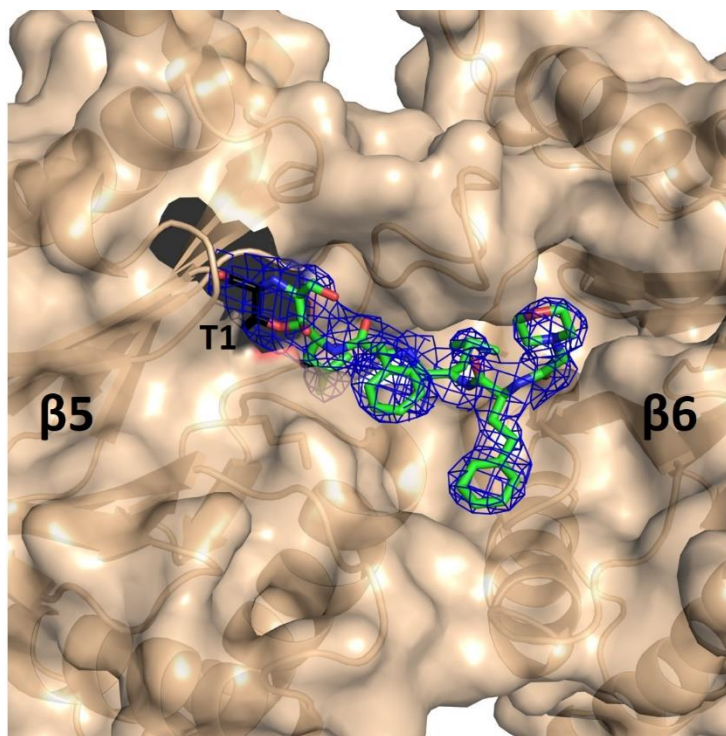


Figure 26 Surface representation and electron density of carfilzomib bound to the chymotrypsin-like active site. Thr1 is colored black; carfilzomib is colored green; electron density is blue and is contoured to 1σ .

An additional hydrogen bond is found in the chymotrypsin-like and trypsin-like sites, between Asp125O $\delta 2$ of the adjoining subunit ($\beta 6$ and $\beta 3$, respectively) and CFZ's P3 peptide bond, with distances of 3.2 Å and 2.7 Å, respectively. This interaction is missing in caspase like sites, where the aspartate residue is replaced by Tyr114, which is flipped away from carfilzomib and points into the S3 pocket.

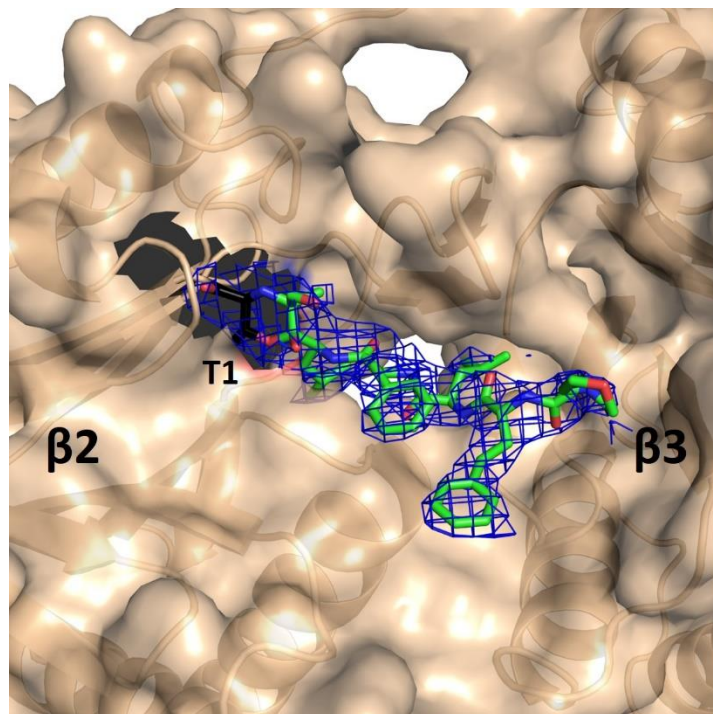


Figure 27 Carfilzomib bound to the trypsin-like subunit, $\beta 2$.

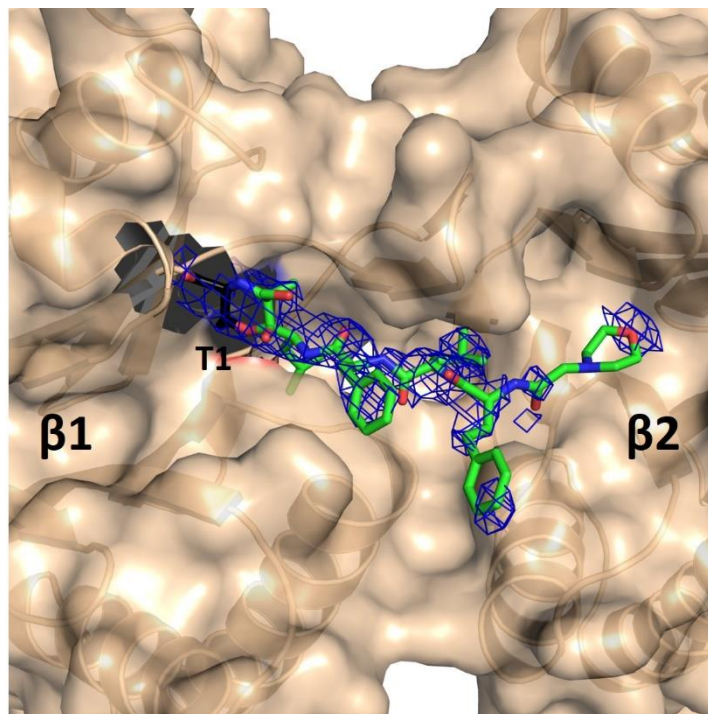


Figure 28 Carfilzomib bound to caspase-like site, $\beta 1$.

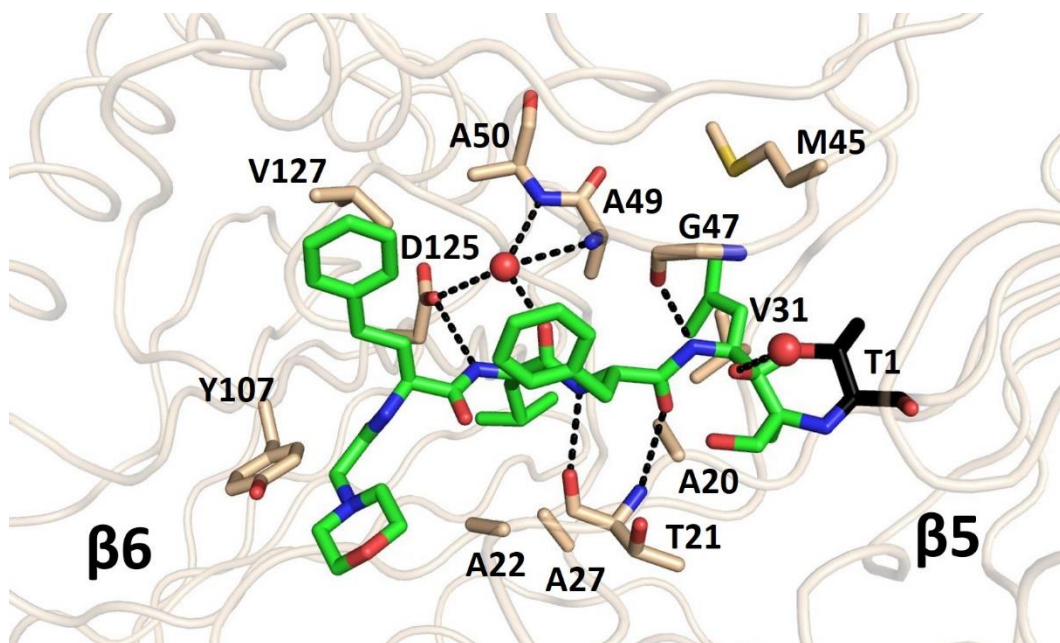


Figure 29 Hydrogen bonding interactions of carfilzomib in the chymotrypsin like subunit $\beta 5$. Hydrogen bonds are indicated with black dashed lines; the active site water molecules are shown as red spheres. The nucleophilic Thr1 is shown in black.

In trypsin like sites the P4 carbonyl oxygen forms a 3.0 Å hydrogen bond with $\beta 3$ Glu106 O δ 2 from the adjoining $\beta 3$ subunit (Figure 30). The analogous residue in chymotrypsin-like sites is $\beta 6$ Tyr107 which does not hydrogen bond with carfilzomib, and in caspase-like sites the nearest residue is $\beta 2$ His116, however at 4.6 Å to the nitrogen, it is also insufficient for hydrogen bonding.

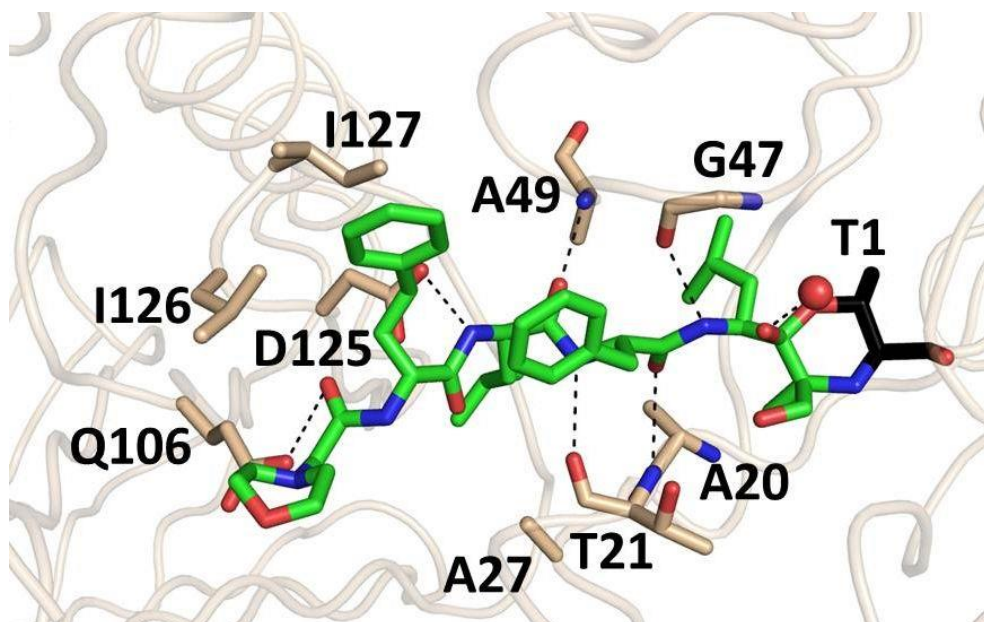


Figure 30 Hydrogen bonding interactions of carfilzomib in the trypsin-like subunit $\beta 2$. Hydrogen bonds are indicated with black dashed lines; the active site water molecules are shown as red spheres. The nucleophilic Thr1 is shown in black.

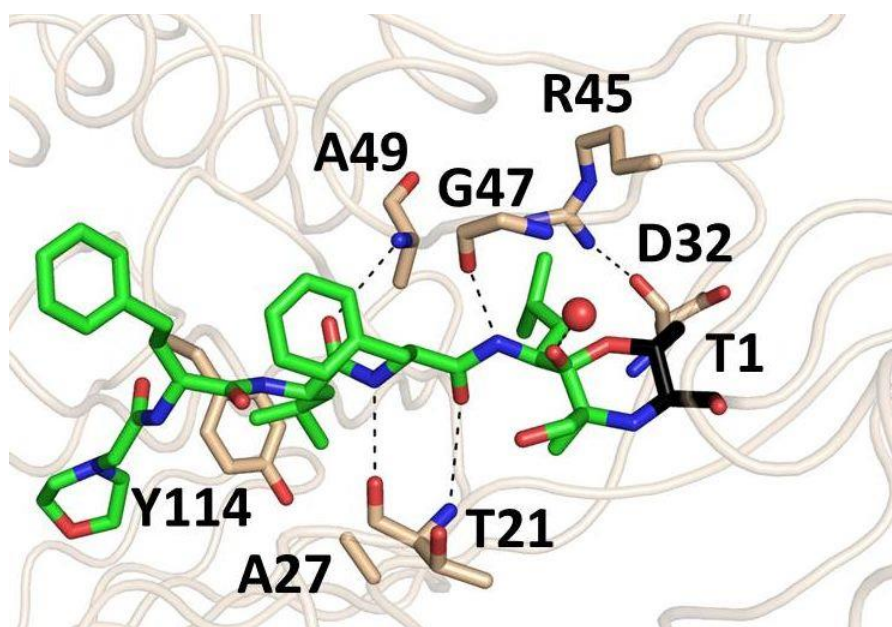


Figure 31 Hydrogen bonding interactions of carfilzomib in the caspase-like subunit $\beta 1$. Hydrogen bonds are indicated with black dashed lines; the active site water molecules are shown as red spheres. The nucleophilic Thr1 is shown in black.

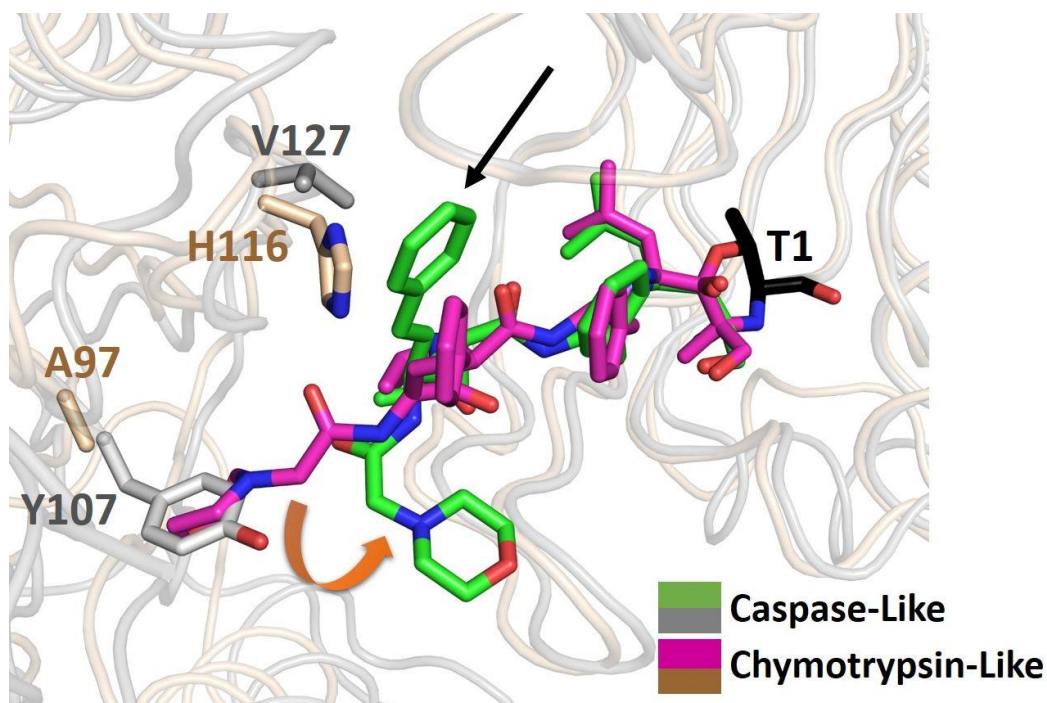


Figure 32 Structural alignment of chymotrypsin-like and caspase-like sites with carfilzomib bound. In caspase-like sites, H116 prevents carfilzomib's P4 phenyl group from accessing the S4 pocket (black arrow), resulting in a disordered N-terminus of the drug and decreased selectivity. The N-terminal morpholino ring in chymotrypsin-like sites assumes a fixed conformation due to steric effects of Y107.

3.2.2 Interactions of Carfilzomib with Proteasome Specificity Pockets

A detailed evaluation of the active subunits in the presence and absence of carfilzomib visualized the structural features that lead to the drug's chymotrypsin-like selectivity. In caspase-like sites, the N-terminus of carfilzomib, beginning at the P4 phenyl group, is not well-defined in the electron density. Comparison with chymotrypsin-like and trypsin-like sites reveals that His116 of chain $\beta 7$ acts to sterically block carfilzomib's P4 phenyl group from entering the S4 pocket of caspase-like sites. Structural alignment with

either chymotrypsin-like or trypsin-like subunits shows that the P4 substituent of CFZ is shifted by up to 3.7 Å away from the S4 pocket. Additionally, the P4 phenyl group of caspase-like sites makes no interactions with the protein. In contrast, chymotrypsin-like and trypsin-like sites each have S4 pockets that favor the hydrophobic P4 phenyl group of carfilzomib. Van der Waals interactions are made with the side chains of β 6Val127 (4.0 Å) in chymotrypsin-like sites and β 3Ile126 (3.8 Å) and β 3Ile127 (4.0 Å) in trypsin-like sites.

The P3 position of CFZ contains leucine which is able to make Van der Waals interactions in each of the three distinct catalytic subunits. In chymotrypsin like sites, the S3 pocket has four hydrophobic amino acid side chains that surround the P3 leucyl and make Van der Waals contacts as follows; Ala20 (3.9 Å), Ala22 (4.4 Å), Ala27 (3.9 Å) and the aliphatic side chain carbon of β 6Asp125 (3.6 Å). Caspase-like sites make Van der Waals contacts with the side chains of Ala27 (3.3 Å), Thr20 (4.0 Å), Thr22 (4.3 Å) and the phenyl ring of β 2Tyr114 (3.7 Å). In trypsin like sites, the Van der Waals interactions come from the side chains of Ala27 (4.3 Å), Ala20 (4.4 Å), and the aliphatic carbon of Glu22. The acidic side chain of β 3Asp125, however, also lines the S3 pocket of trypsin-like sites; therefore binding at the P3 position is less favorable than chymotrypsin-like and caspase-like subunits.

The S2 pockets of all three active sites are able to accommodate the P2 phenylalanine moiety of CFZ, and no contacts are made with the protein. The crystal structure of bortezomib bound to the yeast proteasome shows that its P2 phenylalanine ring is flipped toward Thr1 in both the chymotrypsin-like and caspase-like sites, and away

from Thr1 in the trypsin-like sites [42] (Figure 33). In contrast, the human proteasome:CFZ complex shows the phenyl ring is flipped away from Thr1 in all three distinct catalytic centers. Carfilzomib's P2 substituent assumes the same orientation taken by the P2 methoxyphenyl substituent of PR-957 in the mouse immuno and constitutive proteasome structures [16].

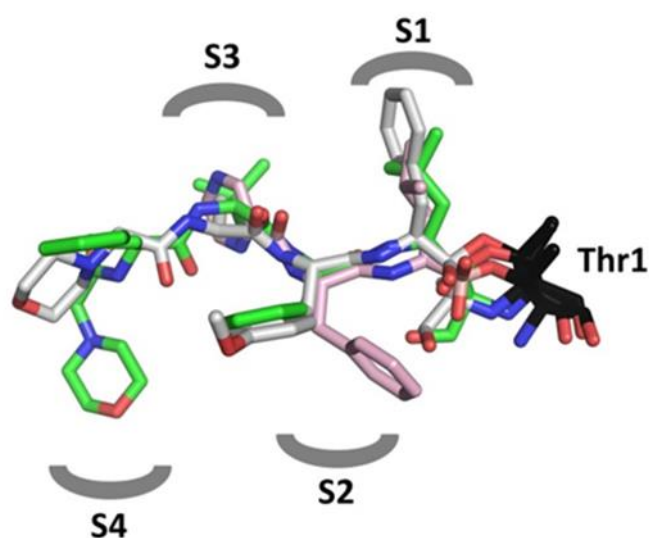


Figure 33 Structural alignment of the chymotrypsin-like sites of the human:carfilzomib, yeast:bortezomib, and mouse:PR-957 complexes. The black arrow indicates the P2 residue, which is flipped in carfilzomib (green) and PR-957(white) relative to the P2 phenylalanine of bortezomib (pink). Thr1 is shown in black, and the specificity pockets which harbor ligand side chains are labeled as S1-S4.

The S1 pocket of chymotrypsin-like subunits makes van der Waals contacts with CFZ's P1 leucyl from the side chains of Ala49 (3.7 Å), Val31 (4.5 Å), Ala20 (3.5 Å), and Met45 (4.3 Å). Structural alignment of unbound and CFZ bound proteasome (rmsd C α : 0.4 Å) reveals that upon binding of CFZ, the P1 leucyl group causes the side chain of Met45 to shift by 2.7 Å towards Ile35, which is then shifted by 1.7 Å (Figure 34). In

contrast to caspase-like subunits where Arg45 is hydrogen bonded and unable to accommodate the leucyl group, the Met45 side chain is not stabilized by surrounding residues and is therefore flexible. Upon shifting of the Met45 and Ile35 side chains, conformational changes involving residues 34 to 72 are also seen. This offsets β -sheets S4 and S5 as well as α -helix H1 compared to the unbound structure.

In summary, caspase-like sites disfavor CFZ binding due to electrostatic and steric effects at the P1 and P4 positions, respectively, while chymotrypsin-like and trypsin-like sites show similar hydrogen bonding and Van der Waals interactions. The most profound difference between chymotrypsin- and trypsin-like sites is the size of the S1 pocket, and the number of van der Waals contacts made with the P1 leucyl of CFZ. This may suggest the importance of a properly stabilized P1 substituent for modification of Thr1 by the epoxyketone pharmacophore. Despite these seemingly minor differences, CFZ is nearly 1000-fold more potent toward chymotrypsin-like subunits. An additional explanation may be accessibility. It is interesting to note that the IC₅₀ for trypsin-like sites of immunoproteasomes is reduced to 620 nM (Screen et al., 2010), which may be a reflection of better accessibility in immunoproteasomes, considering that active site substitutions between the two isoforms do not appear to change the character of the binding region.

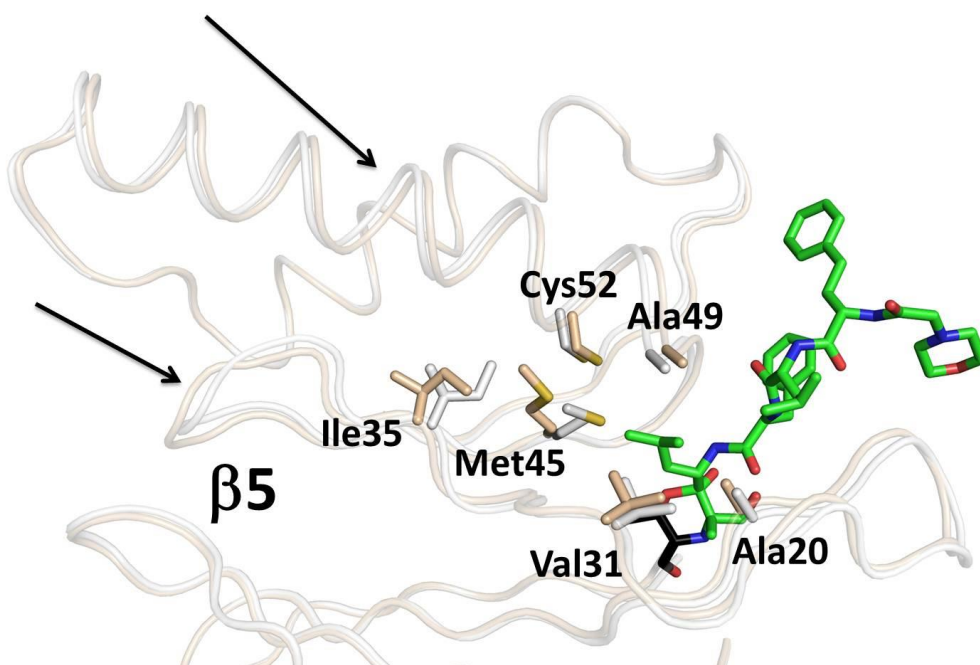


Figure 34 Shifts in the S1 pocket of the human proteasome chymotrypsin-like subunit, $\beta 5$, upon carfilzomib binding. Unbound and carfilzomib bound structures are aligned by their C α atoms. The unbound $\beta 5$ is colored white, carfilzomib bound $\beta 5$ is wheat, carfilzomib is colored green, and Thr1 is black. Regions of the protein that shift upon carfilzomib binding are indicated with black arrows.

3.2.3 PR-957 Selectivity for Immunoproteasome

The residue and loop movements in the chymotrypsin-like S1 pocket of the constitutive proteasome:carfilzomib complex were also reported in the crystal structure of the mouse constitutive proteasome:PR-957 complex, but not in the mouse immunoproteasome:PR-957 structure [16] (Figures 35 and 36, respectively). In the immunoproteasome chymotrypsin-like subunit, the side chain of Met45 in the S1 pocket is stabilized by the aliphatic side chain of the conserved amino acid substitution Q53,

resulting in a larger S1 pocket than constitutive S1, and therefore no need for rearrangement. Although the shifts in the S1 pocket of the mouse constitutive proteasome:PR957 complex were suggested to be the cause for immunoproteasome selectivity [16], carfilzomib is slightly more selective for constitutive proteasomes than immunoproteasome ($IC_{50} = 6nM$ and $34 nM$, respectively) even though this structural rearrangement occurs. To investigate this discrepancy, we compared the mouse constitutive proteasome:PR-957 complex with the human constitutive proteasome:carfilzomib structure to examine the interactions in the entire binding region, and determine the complete nature of the selectivity in each case.

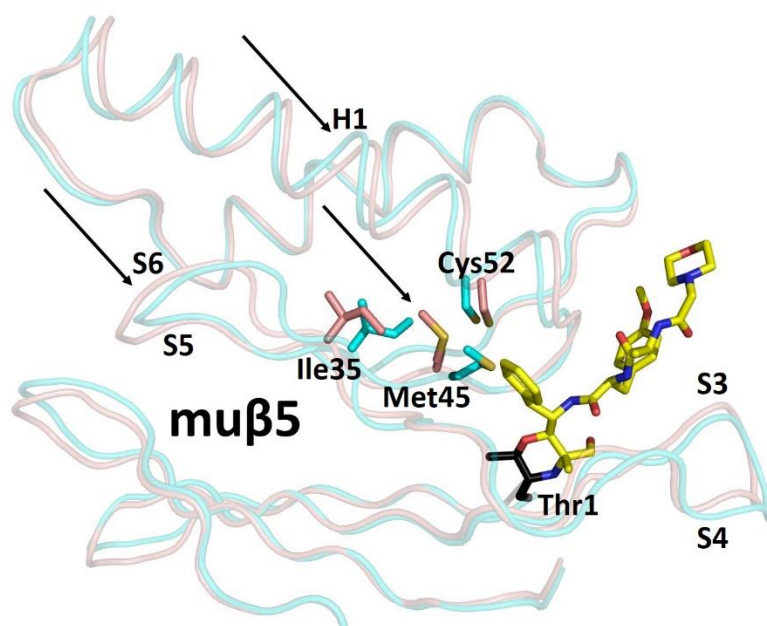


Figure 35 PR-957 bound to mouse constitutive 20S Proteasome chymotrypsin-like site. The same structural rearrangements are observed as in the carfilzomib:proteasome structure, indicated with black arrow. The unbound protein is teal, PR-957 bound is pink, and PR-957 is yellow (PDB ID unbound, 3UNE; PR-957 bound, 3UNB).

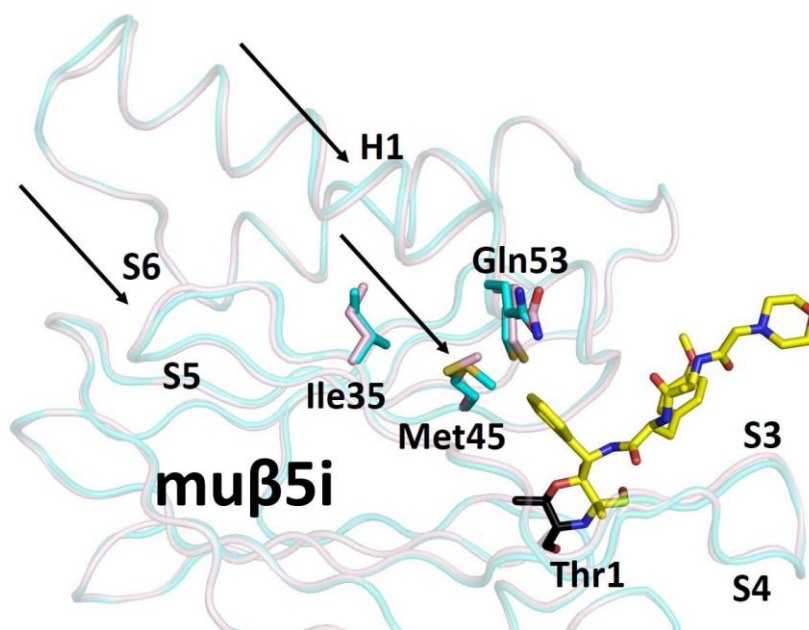


Figure 36 PR-957 bound to mouse 20S immunoproteasome chymotrypsin-like site. Same color scheme as figure 36. No structural rearrangements are observed due to Gln53 stabilizing Met45 resulting in a larger S1 pocket than constitutive proteasomes β 5 (PDB ID unbound, 3UNH; PR-957 bound, 3UNF).

The most profound differences reside at the P4 and P3 positions of the two drugs. PR-957 is a tripeptide epoxyketone and is unable to make interactions with the S4 binding pocket, whereas carfilzomib makes van der Waals interactions with β 6Val127. Additionally, PR-957 contains a P3 alanine, opposed to carfilzomib's P3 leucyl group. While this maintains a hydrophobic nature, the distances of PR-957's alanine to the residue side chains in the S3 pocket are up to 1.9 Å greater (Ala20) than those of carfilzomib's P3 leucyl group (Figure 37). Taken together, this suggests that PR-957's selectivity for immunoproteasomes is only partially explained by immunoproteasomes having a larger S1 pocket and no rearrangements upon binding of PR-957.

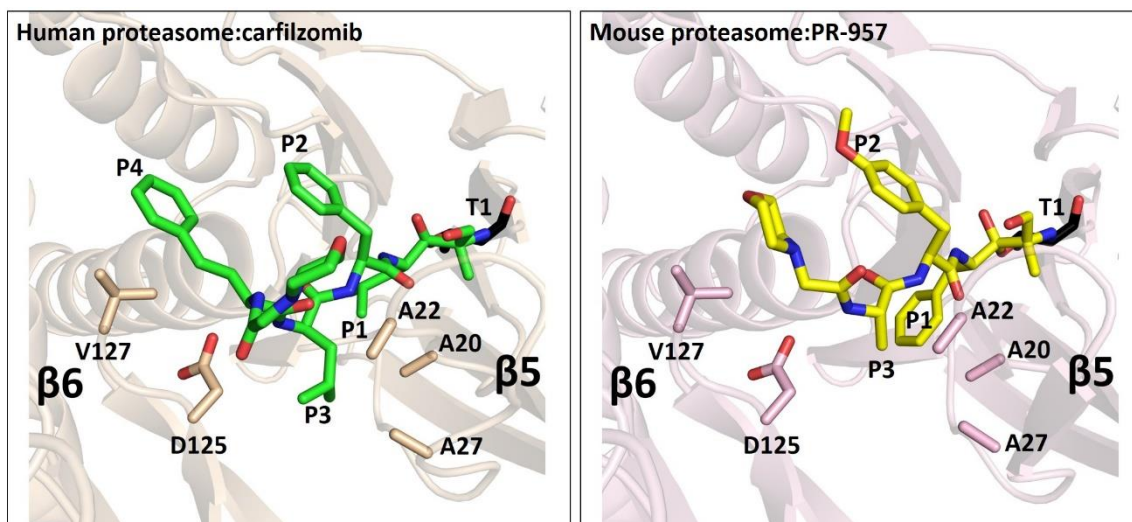


Figure 37 Carfilzomib's P3 and P4 residues contribute to its increased selectivity for constitutive proteasomes chymotrypsin-like sites compared to PR-957. Carfilzomib (left image) has a P3 leucyl group and P4 phenyl group which stabilize ligand binding. PR-957 (right image) has a smaller alanine at P3 and no P4 residue, resulting in decreased selectivity towards constitutive proteasomes chymotrypsin-like sites and more favorable binding to immunoproteasomes.

The human proteasome:carfilzomib structure suggests that it is also an absence of interactions in PR-957's P3 and P4 positions which contributes to its selectivity for immunoproteasomes, whereas CFZ, which does have favorable interactions at the P4 and P3 positions, is able to maintain selectivity for constitutive proteasomes despite structural rearrangements in the S1 pocket.

3.2.4 Predictions of Carfilzomib Interactions with the Immunoproteasome

In most cells, upon exposure to interferons (IFNs), the $\beta 1i$ /LMP2, $\beta 2i$ /MECL-1, and $\beta 5i$ /LMP7 subunits are induced and incorporated into newly synthesized immunoproteasomes, taking the place of the constitutive catalytic subunits $\beta 1$, $\beta 2$, and $\beta 5$

[103]. Cleavage preferences for immuno CPs are after nonpolar amino acids, therefore the peptides produced have hydrophobic C termini which then bind to MHC-1 molecules [104, 105].

The immunoproteasome is a major form of the proteasome expressed in cells of hematopoietic origin, including multiple myeloma tumor cells. Despite being expressed in higher levels than constitutive proteasomes, specific inhibition of either proteasome form alone has been shown to be insufficient at producing an antitumor response in CD138⁺ multiple myeloma cells; rather, inhibition of both forms is necessary [73].

Recent studies also show that high levels of immunoproteasomes are associated not only with various types of cancer, but also in autoimmune disorders and neurodegenerative diseases [106-109]. Therefore, the design of immuno proteasome selective inhibitors is now being explored as novel therapeutics to combat these diseases.

Structural alignments of the $\beta 1$ (caspase-like), $\beta 2$ (trypsin-like), and $\beta 5$ (chymotrypsin-like) subunits of the carfilzomib bound constitutive human proteasome, with the corresponding immuno subunits of *M.mucus* (mouse) shows high structural similarity; rmsd C α : hu $\beta 5c$ /mu $\beta 5i$ 0.59Å; hu $\beta 1c$ /mu $\beta 1i$ 0.56Å; hu $\beta 2c$ /mu $\beta 2i$ 0.63Å. The sequence identities between the human constitutive and mouse immunoproteasome catalytic subunits are 70% ($\beta 5$), 64% ($\beta 1$) and 58% ($\beta 2$). The sequence identities of the catalytic immunoproteasome subunits between the two species are: 90.5% ($\beta 1i$); 89% ($\beta 2i$); 92% ($\beta 5i$), with majority of the amino acid differences residing outside of the active sites.

To investigate why carfilzomib's IC_{50} is lower for binding to the chymotrypsin-like subunits of constitutive proteasomes over immunoproteasomes, the carfilzomib bound chymotrypsin-like subunit ($\beta 5$) was structurally aligned with the mouse immunoproteasome chymotrypsin-like subunit ($\beta 5i$) (rmsd C_{α} = 0.59 Å) (Figure 38). The amino acid sequence similarity between the two subunits is 92%, and they share five conserved amino acid substitutions in the substrate binding region. Conserved amino acid substitutions in the chymotrypsin-like sites of human and mouse constitutive and immunoproteasomes are: V128T, A46S, G48C, S53Q and A27S. The structural alignment revealed that the conserved substitution G48C places the polar cysteine side chain 3.2 Å from carfilzomib's P4 phenyl group. Additionally, the conserved substitution A27S, places the polar side chain of serine within 3.3 Å of the P3 leucyl group. Given the importance of the P3 and P4 positions in the IC_{50} value, as indicated previously for PR-957, it is possible that these minor changes could account for the nearly 6 fold increase in the IC_{50} value between immuno and constitutive proteasomes, respectively.

Amino acid substitutions in the binding region of human immunoproteasome chymotrypsin-like ($\beta 5i$) sites that are not present in mouse immuno proteasomes include S28A, and T21S. These remain in mouse $\beta 5i$ as serine and threonine, respectively. While the T21S substitution maintains similar polarity as constitutive core particles, the S28A substitution is in the S3 pocket and increases the hydrophobicity compared to mouse immunoproteasome $\beta 5i$. Additionally, mouse $\beta 5i$ subunits have a V31M substitution in the back of the S1 pocket, whereas residue 31 remains a valine in human. This maintains hydrophobic character, but results in a larger S1 pocket for human $\beta 5i$.

In caspase like sites of the immunoproteasome ($\beta 1i$), the conserved amino acid substitutions in the S1 pocket are T20V, T31F, R45L, and T52A (Figure 40). This results in a decreased size and increased hydrophobicity of the S1 pocket of $\beta 1c$. This leads to a preference for cleavage after small, hydrophobic and branched residues, consistent with cleavage of the model substrate Ac-PAL-AMC (Acetate-Pro-Ala-Leu-7-amino-methylcoumarin)[110-112]. This results in the generation of more epitopes with nonpolar C termini such as Ile, Leu, or Val, which are presented on MHC-1 molecules. Other conserved substitutions in the binding region of $\beta 1i$ are located in the S3 pocket. T22A increases the S3 pocket's hydrophobicity, and A27V conserves hydrophobic nature but provides a larger surface, providing more contact space. The substitution Y114H of adjoining subunit $\beta 2i$, however, adds polarity to the S3 pocket, but does increase it in size relative to $\beta 1c$, where the tyrosine blocks off the S3 pocket. Whereas mouse proteasomes have the substitution $\beta 2S112E$, which is located in the back of the S3 pocket, humans have the substitution $\beta 2S112G$.

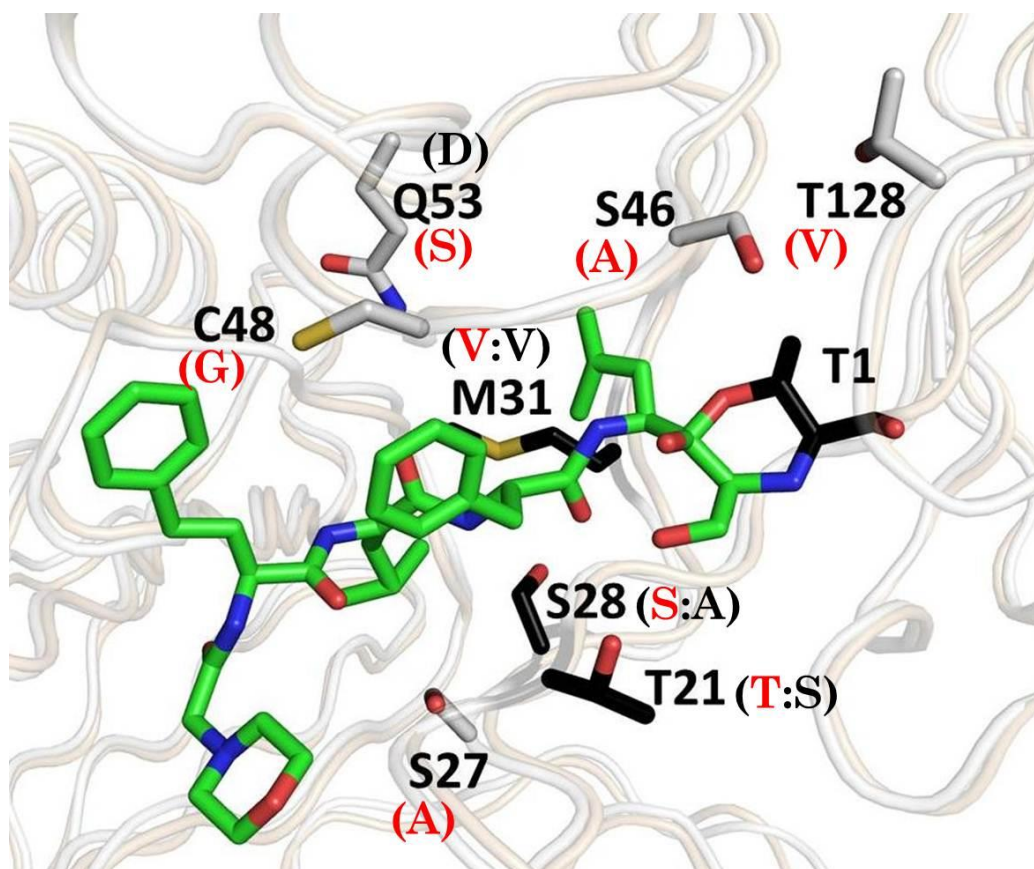


Figure 38 Structural alignment of the chymotrypsin-like subunit $\beta 5c$ of the human proteasome:carfilzomib complex with the corresponding mouse immuno subunit $\beta 5i$. The human proteasome subunit is shown in wheat; carfilzomib is shown in green; the active site Thr1 from human is shown in black; mouse immuno subunit is shown in white. Conserved amino acid substitutions are shown as white sticks and labeled in black. Residue substitutions not conserved with human $\beta 1i$ are shown as black sticks, with the human substitution shown in black in parenthesis. Red labels are of the constitutive particle residue that is substituted.

In trypsin-like subunits ($\beta 1$) the S1 pockets of human and mouse have the conserved substitution D53E, as well as the substitution in the S2 pocket T48V, the later increasing the hydrophobicity of S2 relative to constitutive proteasomes. Structural comparison of the mouse immuno proteasome subunit with the CFZ bound human proteasome, shows that the valine 48 substitution would project the valine side chain

within 3.5 Å of CFZ's P2 phenylalanine ring, and would be able to make Van der Waals interactions (Figure 39). Additionally, the conserved substitutions G23D and E22N are located near the N-terminus of CFZ, and the structural comparison indicates that the aspartate substitution would result in the side chain Oδ2 atom being within 2.9 Å of the CFZ N-terminal morpholino oxygen, which is within hydrogen bonding distance. Human immuno proteasomes have an additional substitution of T52A in the S1 pocket, located 4.5 Å from the P1 leucyl of CFZ, increasing the hydrophobicity of S1 and providing additional van der Waals contacts upon CFZ binding.

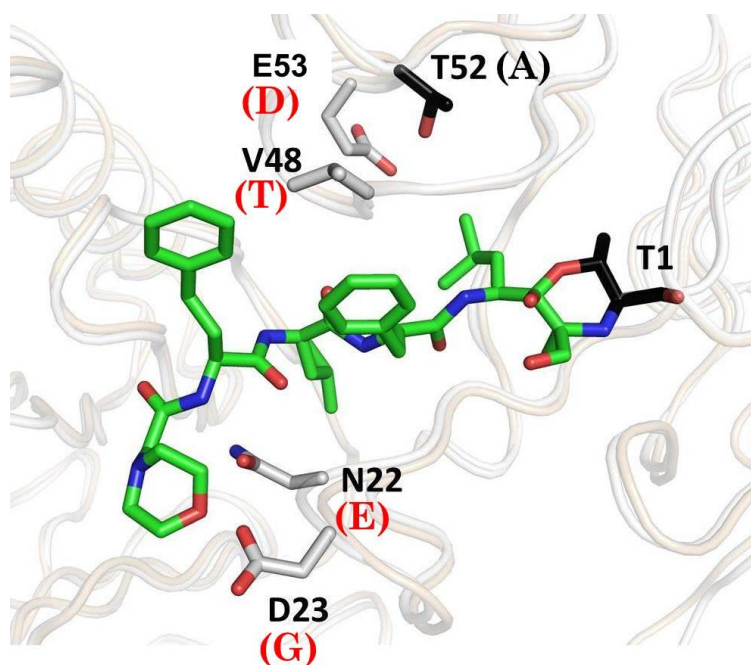


Figure 39 Structural alignment of the trypsin-like subunit $\beta 2c$ of the human proteasome:carfilzomib complex with the corresponding mouse immuno subunit $\beta 2i$. The same color scheme is used as Figure 40.

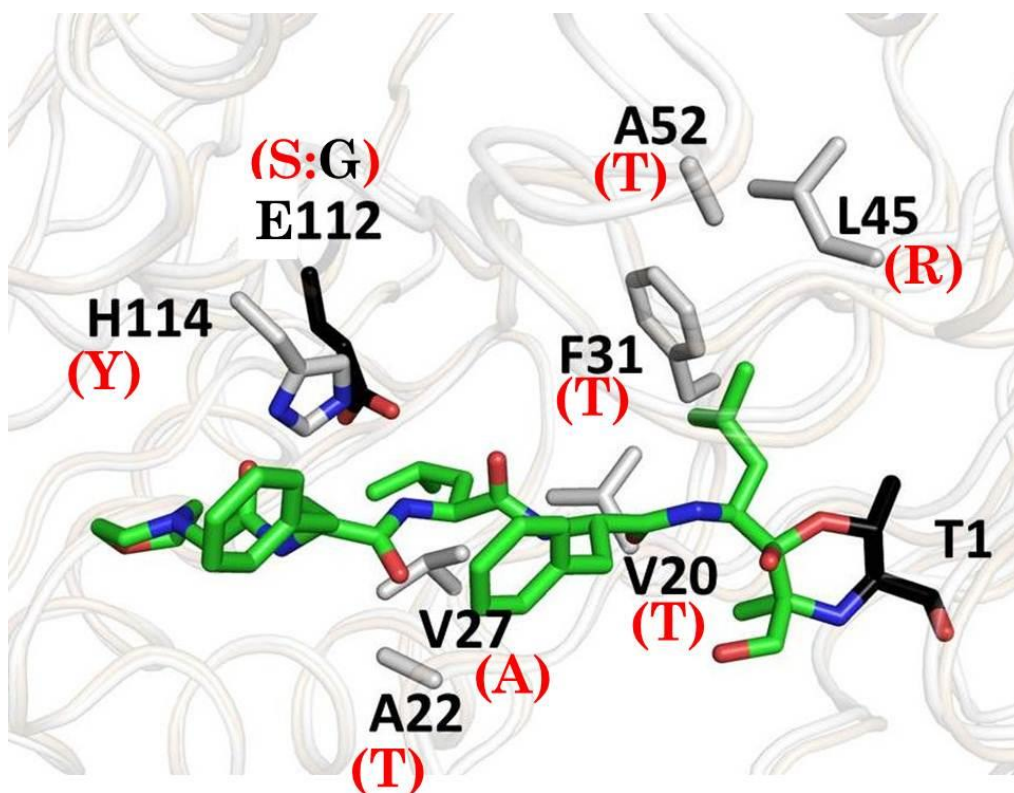


Figure 40 Structural alignment of the human caspase-like subunit $\beta 1c$ with the corresponding mouse immuno subunit $\beta 1i$. The same color scheme is used as Figure 40-41.

3.3 Conclusion

The proteasome is at the center of protein degradation for more than 90% of cytosolic proteins, which are oxidatively damaged, mis-folded, or tagged by ubiquitin for proteolysis. As such, the proteasome has been the focus of drug discovery efforts, which take advantage of the proteasomes essentiality for cellular homeostasis and are proven targets in various forms of cancer, namely multiple myeloma. Previous structural studies

of constitutive proteasome specific inhibitors have only been reported bound to the yeast proteasome.

Reported in this dissertation is the first structure of the human constitutive 20S proteasome, in both an unbound state and bound with the FDA approved proteasome inhibitor carfilzomib, which primarily targets the chymotrypsin-like site of constitutive proteasomes. The structure shows that carfilzomib's selectivity for the chymotrypsin-like subunits can be attributed to distinct van der Waals and hydrogen bonding interactions that are missing in the caspase-like (and to a lesser extent trypsin-like) subunits, as well as steric factors, such as β 2His116 and Arg45 in the caspase-like site. The P2 phenyl alanine of CFZ makes no contacts with the protein in any of the catalytic subunits, and it has been shown that substitution of methyl serine at the P2 position of CFZ demonstrates good oral bioactivity and is more resistant to microsomal degradation while maintaining low nanomolar potency [67]. Taken together, this provides evidence that the P2 substituent likely contributes to the drug's overall pharmacological properties, but not to active site selectivity or to the IC_{50} values.

The P3 and P4 positions of CFZ, however, do appear to play a large role in active site selectivity. Interestingly, CFZ analogs truncated to be dipeptide epoxyketones only show weak potency (>1000 nM) against the chymotrypsin-like subunit, while tripeptide analogs maintain low nanomolar potency [67]. Additionally, comparison of our CFZ bound structure with the mouse constitutive proteasome:PR-957 complex reveals a new explanation for the drug's immunoproteasome selectivity. Similar structural rearrangements in the chymotrypsin-like subunits of constitutive proteasomes are caused

by both CFZ (which has a P1 leucyl group) and PR-957 (which has a P1 phenylalanine). The greater than 6-fold selectivity of CFZ for constitutive proteasomes, therefore, must reside outside the S1 pocket. Indeed, it was found that CFZ is able to make van der Waals interactions at the P3 and P4 positions, thereby regaining selectivity for the constitutive isoform. PR-957, on the other hand, binds to immunoproteasomes without structural rearrangements, however it lacks a P4 substituent and the alanine at P3 is insufficient for making van der Waals interactions, thus favoring immunoproteasomes for energetic reasons. These observations are further supported by the tripeptide epoxyketone inhibitor ONX0912 (PR-047), which also contains a phenylalanine side chain at the P1 position. Based on previous suggestions, a phenylalanine at the P1 position should favor binding to immunoproteasomes, however, the reported IC_{50} values indicate that it is two-fold more selective for the constitutive chymotrypsin-like ($IC_{50} = 36 \text{ nM}$ ($\beta 5c$); 82 nM ($\beta 5i$)) [67]. This can now be explained based on the structures human proteasome:carfilzomib complex structure. Since PR-047 contains a methoxy group at the P3 position, it is suitable for making van der Waals interactions with the hydrophobic S3 pocket of chymotrypsin-like subunits. Only a structure of the proteasome:PR-047 complex could indicate whether the N-terminal 2-methyl-1,3-thiazole cap is associated with the S4 pocket, which would also add favorable hydrophobic interactions, and in each case off-setting the structural rearrangement that is expected to occur in the S1 pocket.

The conservation of quaternary structure between the constitutive and immunoproteasomes, even across species, enabled a detailed look at the binding probabilities of carfilzomib to the human immunoproteasome catalytic subunits.

Carfilzomib is six times more selective for the chymotrypsin-like sites of constitutive proteasomes versus their immunoproteasome counterpart, and structural comparison with the mouse immunoproteasome indicates that the most profound difference is the conserved A27S substitution in the S3 pocket, diminishing its hydrophobic character, and making binding of the P3 leucyl less favorable. PR-825, a structural analogue of PR-957, possesses a leucine at P1 and a methoxy-serine at the P2 and P3 positions. It has been shown to target constitutive chymotrypsin-like sites ($\beta 5$) with 20-fold the selectivity over immunoproteasome chymotrypsin-like sites, further supporting the hypothesis that non-polar residues at P3 preferentially target constitutive chymotrypsin-like sites.

The proteasome remains an attractive target for antineoplastic drugs and the elucidation of the X-ray structure of the human constitutive CP provides the details necessary for it to serve as a useful tool in future drug development. The use of CFZ in the clinic makes it imperative that the binding mode be well understood at the molecular level. The structures reported here will assist in the development of constitutive and immunoproteasome specific inhibitors, with therapeutic potential in cancer, autoimmune, and neurodegenerative diseases.

3.4 Materials and Methods

Purification and crystallization of unbound proteasome is described in sections 2.3.1 and 2.3.2 respectively. For inhibitor bound structures, the proteasome crystals were soaked with carfilzomib at a final concentration of 5 mM in crystallization buffer, for 12

hours. Carfilzomib was solubilized by sonication and vortexing prior to addition to the proteasome crystals. Crystals were then soaked in cryo-protecting buffer (30%MPD, 0.2M Sodium Formate) and flash frozen with liquid nitrogen prior to data collection. Data was collected using synchrotron radiation at the Advanced Photon Source (Argonne National Labs, IL) with $\lambda = 0.97918 \text{ \AA}$, at beam-lines 23iD and 19iD.

X-Ray intensities were indexed and scaled using HKL3000 [84] in the $P2_1$ space group. Carfilzomib bound crystals yielded 1177627 reflections, with 321261 being unique, and an overall completeness of 98.7%. Calculations of the Matthews coefficient using ccp4 [96] gave an approximate solvent content of 50%, assuming two core particles per asymmetric unit.. The phases were solved by molecular replacement using the default settings for the program Phaser [78] of the Phenix program suite [82], with the coordinates of the core particle from *Bos Taurus* (PDB ID 1IRU) as a starting model. A single solution was calculated and used for structure refinement.

Refinement of the carfilzomib bound structure was carried out in a similar manner as the unbound proteasome structure (section 2.3.3), using the program Phenix Refine version 1.82. The initial model obtained from molecular replacement was subjected to three rounds of coordinate, restrained atomic displacement parameters (ADP), translation-libration-screw-rotation (TLS), and occupancy refinement, however, 2-fold and 4-fold non-crystallographic symmetry (NCS) averaging (Cartesian) were used due to two-molecules in the asymmetric unit. Real space refinement of the structure was carried out in COOT. A coordinate file (pdb file extension) for carfilzomib was built using the PRODRG server [113], and a crystallographic information file (CIF file) was generated

from eLBOW in Phenix. The carfilzomib structures were fit to the $|F_o-F_c|$ density manually in COOT to each of the catalytic subunits ($\beta 1$, $\beta 2$, and $\beta 5$), and all pdb files merged to form one pdb in COOT.

Successive rounds of refinement in Phenix included custom geometry restraints of 1.7 Å between the following proteasome:carfilzomib atoms: Thr1N:CBT, Thr1O γ :CBY, for each catalytic site with carfilzomib bound. The structure was refined to final R_{work} and R_{free} values of 0.22 and 0.25, respectively. Bond and angle values were 0.004 Å and 0.78°, and the Ramachandran plot indicated that 96% of backbone dihedral angles were in allowed conformations. The carfilzomib bound structure was deposited in the Protein Data Bank under accession number 4R67.

4. DISCOVERY OF NOVEL PROTEASOME INHIBITORS; HIGH THROUGHPUT SCREENING

4.1 Introduction

Regulated protein turnover centers on the proteasome, and underlies a wide variety of signaling pathways, such as cell cycle control, transcription, and development [114]. Pharmacological inhibition of the proteasome can be efficacious in the treatment of human cancers, as proven by the FDA approved drugs bortezomib (Velcade, Millennium Pharmaceuticals) carfilzomib (Kyprolis, Onyx Pharmaceuticals), both used clinically for the treatment of multiple myeloma. Additionally, several proteasome inhibitors are in clinical trials including Oprozomib (Onyx Pharmaceuticals, Inc.; ONX 1912; PR-047), an epoxyketone, which is being tested in phase 1b/2 clinical trials for the treatment of multiple myeloma, and a phase 1 study in patients with recurrent or refractory solid tumors. Marizomib (Nereus Pharmaceuticals) and ONX-0912 (Onyx Pharmaceuticals) are β -lactone and epoxyketone proteasome inhibitors, respectively, and are each in phase 1 clinical trials for the treatment of multiple myeloma.

Despite the success in developing proteasome inhibitors for use in the clinic, there are several drawbacks to the current therapies. Bortezomib is known to cause peripheral neuropathy in patients due to off-target effects caused by its boronate warhead, which has moderate affinity for serine proteases, in particular HtrA2 [65]. Carfilzomib has poor bioavailability and a half-life of less than 30 minutes, which is inconvenient for patient

care. Drug resistance to bortezomib has been linked to an A49T mutation, which may disrupt hydrogen bonding and β -sheet formation with the inhibitors peptide backbone [115]. Additionally, both bortezomib and carfilzomib are substrates of p-glycoprotein which also leads to drug resistance due to the drug being pumped out of the cell. For these reasons there is a need to identify novel proteasome inhibitors with increased bioavailability, decreased side-effects, and capable of overcoming drug resistance which the current therapies are susceptible to.

High throughput screening (HTS) is a drug discovery processes that is widely used in academia as well as industry, to quickly assay and identify potential drug-like compounds with biochemical activity. Normally, HTS is performed in microliter plates of 96, 384 or 1536 well format. There have been arguments that HTS contributes to the decline in productivity of pharmaceutical companies, and suppresses the creativity of drug discovery [116, 117]. HTS success rates, however, are generally around 50% [118], and while all screening methods have their drawbacks, HTS offers a unique advantage that it is applicable to a broader range of targets. Also, it is now understood that screening hits are simply chemical starting points that need to be optimized, and the expectation of identifying a drug directly from a screen has diminished.

Several HTS techniques have been successfully used in the laboratory to identify novel proteasome inhibitors, including virtual and in vitro screening [119], image-based screening [120], and high-content cell-based screening [121]. In the present thesis, we describe the use of a fluorescence based HTS assay, which relies on the cleavage of an

amino-methyl-coumarin tag ($\lambda_{\text{ex}}= 360\text{nm}$, $\lambda_{\text{em}}= 460\text{nm}$) from a peptide substrate that is specific for the chymotrypsin-like site of the proteasome.

Decrease in fluorescence compared to uninhibited proteasome was evaluated, and those compounds exhibiting greater than 80% inhibition were chosen for a secondary assay to verifying binding to the proteasome, as well as ensure that the decrease in fluorescence was not due to fluorescent quenching or other assay anomalies. The secondary assay utilizes a proteasome activity probe, Ada-Lys(biotinyl)-(Ahx)₃-(Leu)₃-vinyl sulphone [122], which discriminates between compounds that are actually binding to the catalytic sites of the proteasome, and those compounds that give false positives in the HTS assay.

4.2 Results and Discussion

4.2.1 National Institute of Health Clinical Collection (NCC)

The National Institute of Health's Clinical Collection 2 (NCC) contains 450 compounds which have a history of being used in human clinical trials. The plate was tested for inhibition against the chymotrypsin-like activity of the proteasome (Methods 2.3.2). A positive hit was taken as 80% or greater inhibition at 10 μM of test compound. Only one compound met the criteria, α -mangostin, resulting in a hit rate of 0.36%. The low hit rate was expected due to such a high threshold, however only potent proteasome inhibitors were of interest due to the nature of the plate being compounds that are known

to be clinically relevant, versus a fragment library for instance, in which we would set a lower threshold in the hopes of optimizing the hits as lead compounds.

The compound α -mangostin is a natural product that can be isolated from the pericarp of mangosteens, a fruit found natively in Thailand (Figure 41). The fruit has a long history of medicinal use in both Chinese and Ayurvedic medicine. For centuries, people in Southeast Asia have used dried mangosteen pericarp for medicinal purposes; it is used as an antiseptic, an anti-inflammatory, an anti-parasitic, an antipyretic, an analgesic, and as a treatment for skin rashes [123].

Mangosteen has become a popular natural dietary health supplement, and the effects of mangosteen extracts have had recent scientific support. The characteristic anti-proliferative effect of α -mangostin has received a great deal of attention. It is shown to induce cell cycle arrest and apoptosis in colon [124], breast [125], prostate [126], and leukemia HL60 cells [127]. It has also been shown to inhibit cell invasion and migration in mammary and prostate cancer cells. Additionally, it is associated with down-regulation of matrix metalloproteinase-2 and-9, which are involved in disease processes such as arthritis and metastasis [128].

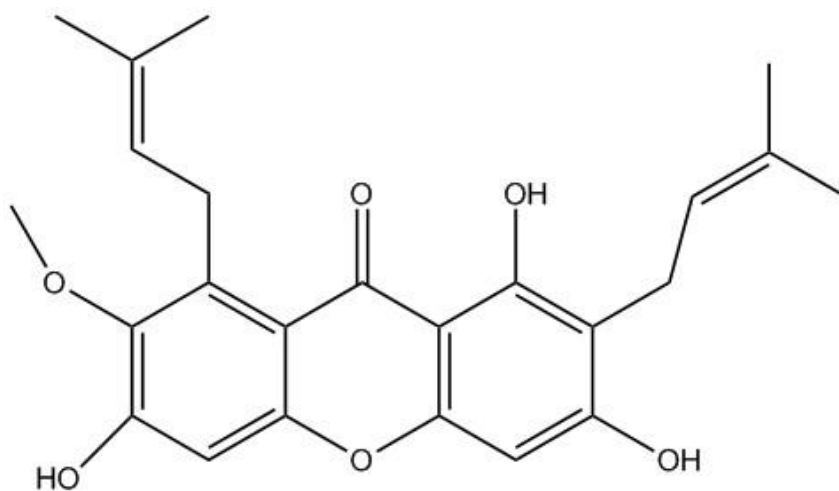


Figure 41 Chemical structure of α -mangostin

To determine the extent of in vitro inhibition of the individual catalytic sites of the proteasome, the IC_{50} value for α -mangostin was determined for each site; caspase-like ($\beta 1$); trypsin-like ($\beta 2$); or chymotrypsin-like ($\beta 5$) (Methods 2.3). Concentrations of α -mangostin were varied from 0-10 μM , and α -mangostin was found to inhibit the chymotrypsin-like activity to the greatest extent ($IC_{50} = 700$ nM), followed by the caspase-like activity ($IC_{50} = 1.5$ μM) (Figure 45). No inhibition of the trypsin-like activity was observed up to 20 μM .

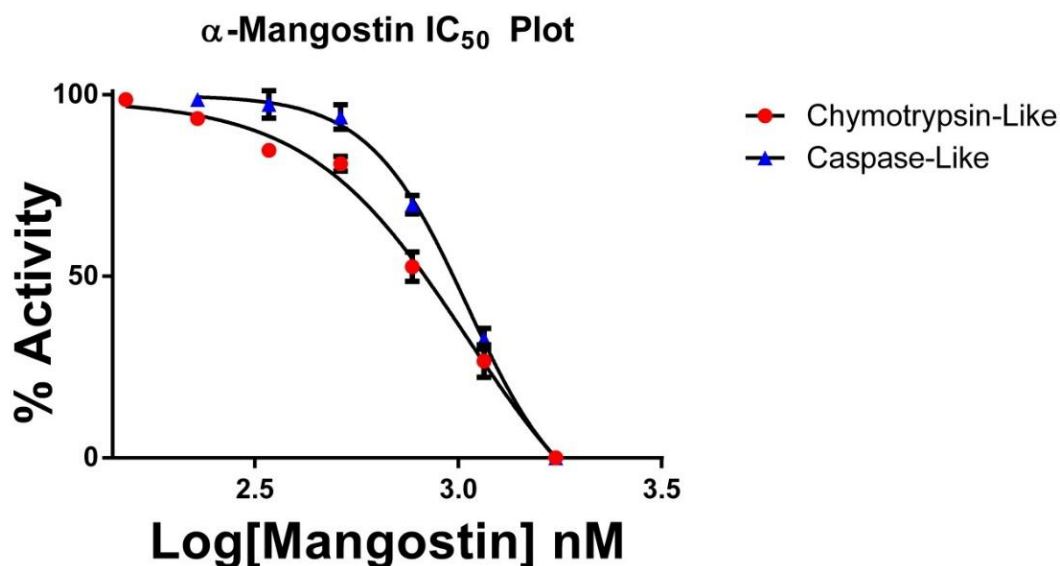


Figure 42 In vitro inhibition of chymotrypsin-like and caspase-like activities of the proteasome by α -mangostin. The IC₅₀ values against the chymotrypsin-like and caspase-like activities were calculated to be 700 nM, and 1.5 μ M, respectively. Error bars represent standard deviations of triplicate experiments.

To confirm the binding capability of α -mangostin to each catalytic site of purified human 20S proteasome, α -mangostin was subjected to an in vitro binding assay utilizing the molecular probe, Ada-Lys(biotinyl)-(Ahx)3-(Leu)3-vinyl sulfone. The probe molecule consists of three parts: a “warhead” which covalently binds to the active site Thr1; a peptide linker which gives subunit specificity; and a visualization tag, biotin. The probe has comparable affinity (IC₅₀ = 1 μ M) for each of the three catalytic subunits of the proteasome (β 1, β 2, and β 5) [122]. A control containing only enzyme and probe but no test compound was used in each trial. Additionally, a sample containing the known inhibitor bortezomib was used in each experiment as a positive control for binding. After thirty minutes the samples were run on an SDS PAGE gel, followed by Western blotting

with streptavidin alkaline phosphatase conjugated. Visualization by the addition of alkaline phosphatase substrate results in three bands on the membrane for no inhibitor control samples; one band for each catalytic subunit that has the probe bound. Bortezomib samples show no visible bands, due to blocking the binding of the probe.

The binding assay confirmed that α -mangostin was binding to both the chymotrypsin-like (β 5) and caspase-like (β 1) sites, while the trypsin-like site (β 2) remained uninhibited, interpreted by a decrease in the band intensity for subunits β 5/1 but not β 2 compared to the experimental control with no inhibitor (Figure 43).

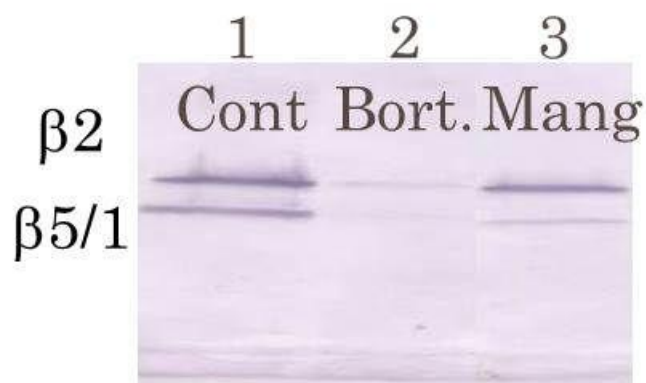


Figure 43 In vitro binding assay of α -mangostin against human 20S proteasome. Lane 1 contains only proteasome and probe; lane 2 shows bortezomib as a positive control for probe binding; lane 3 is α -mangostin. The diminished band intensities signify binding of inhibitor to the catalytic site indicated.

To examine the cellular toxicity of α -mangostin, it was tested in a dose dependent manner against Molt4 Leukemia cells, as well as non-cancerous human dermal fibroblasts,

as described in Methods 2.6. The purpose of the fibroblasts is to establish whether α -mangostin has general cytotoxicity, or if it is specific for cancer cells.

The IC_{50} against the Molt4 leukemia cells was determined to be $15\mu M$ (Figure 44). Interestingly, the leukemia cells reached 100% cell death at $25\mu M$, and showed 25% cell death at $12.5\mu M$ (Figure 45), suggesting that toxicity occurs rapidly between these two concentration.

Human dermal fibroblasts, on the other hand, showed no toxicity against α -mango at concentrations up to $50\mu M$ (Figure 45). Bortezomib was used as a positive control for cell killing, and had an $IC_{50} = 5\mu M$. These results suggested that α -mangostin is selectively toxic against Molt4 Leukemia cells, while non-transformed cells remain unaffected.

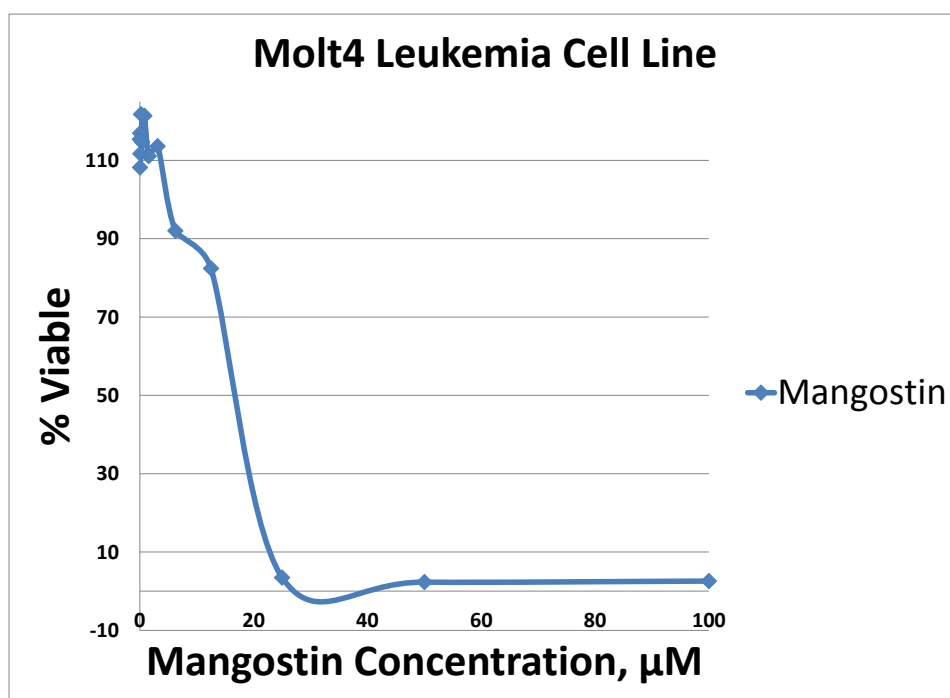


Figure 44 IC_{50} Graph against Molt4 Leukemia Cells IC_{50} value is approximately $15\mu M$.

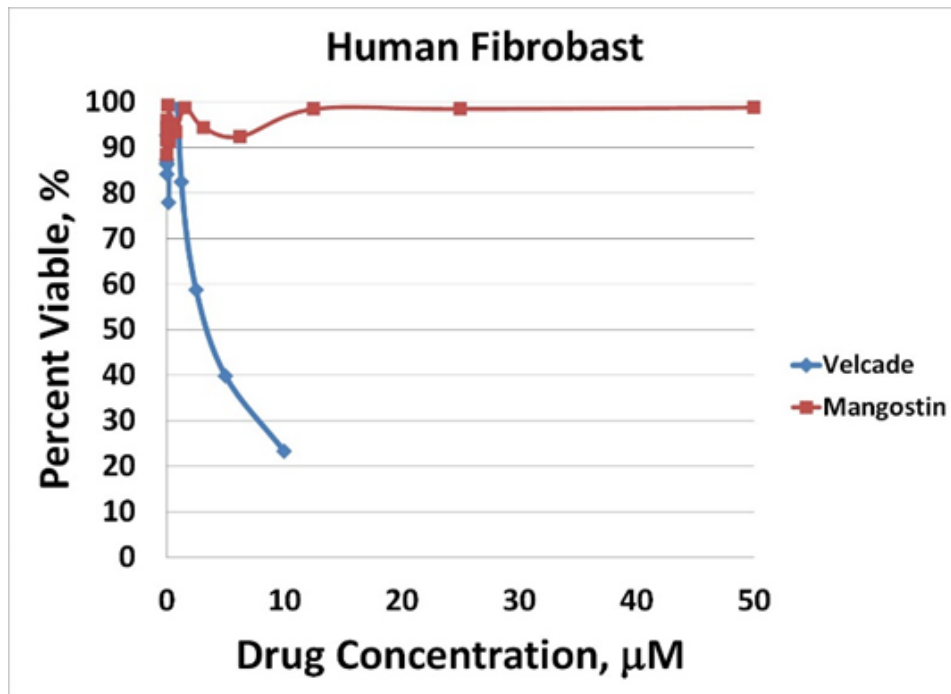


Figure 45 IC₅₀ Graph against human dermal fibroblasts

To determine if α -mangostin targets the proteasome in intact Molt4 Leukemia cells, the cell permeable activity based probe Ada-Lys(biotinyl)-(Ahx)³-(Leu)³-vinyl sulfone was used (Methods 2.5). At concentrations of 50 μ M, α -mangostin completely bound to all three catalytic sites of the proteasome, indicated by a diminished band intensity on the Western Blot compared to a no inhibitor control. At a concentration of 10 μ M of α -mangostin no binding was detectable by Western Blot (Figure 46). Tannic Acid and Enamine were used as a negative control to verify that 50 μ M of a random compound would not alter the expected band intensity leading to a misinterpretation of the blot. Carfilzomib was used as a positive control at a concentration of 10 μ M, and completely eliminated the Western Blot bands for all catalytic sites.

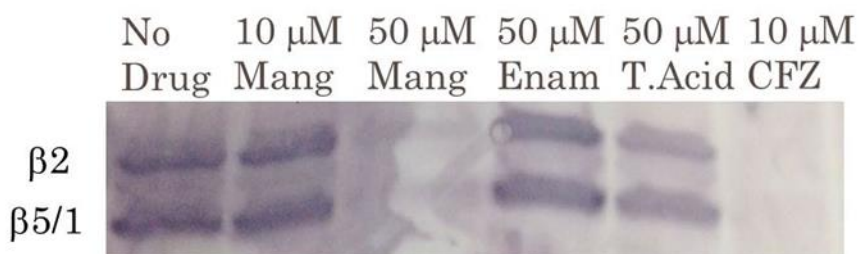


Figure 46 Western Blot Analysis of α -Mangostin in Molt4 Leukemia Cells.

In an attempt to determine the molecular basis for α -Mangostin's proteasome inhibition, proteasome crystals were soaked with 5mM final concentration of α -Mangostin dissolved in crystallization buffer for 24 hours. X-ray data for the crystals were collected at Argonne National Labs and processed as described for the carfilzomib soaked crystals. The resolution extended to 3.0 Å, but unfortunately the difference electron density did not show the presence of α -mangostin in any of the catalytic subunits. This may be due to the low solubility of α -mangostin, as it was not completely dissolved during crystal soaking. Another possibility could be the low affinity, $IC_{50} = 700$ nM against the chymotrypsin-like site, and longer crystal soaking times, or a higher concentration of α -Mangostin may lead to binding in the crystals.

4.2.2 Fungus Microbial Extracts; *Zygothiala jamaicensis*

Microbial extracts from the fungus *Zygothiala jamaicensis* were fractionated by high performance liquid chromatography on a reverse phase C18 column, and a 1mL aliquot of each fraction placed into a separate well in a 96-well plates. These aliquots were subjected to high throughput screening against the chymotrypsin-like subunit of the human 20S proteasome (Section 2.3.2). One fraction demonstrated activity, so it was analyzed by liquid chromatography/mass spectrometry (LC/MS). LC/MS data was obtained for the active fraction on an Agilent 1200 Infinity series HPLC and an in-line Bruker micrOTOF-Q II mass spectrometer. The purified material was subjected to NMR spectroscopy by performing ^1H , ^{13}C , and Dr. Nishant Shetty assigned the structure based on the NMR data. It was determined to be a polyphenol, and therefore named PPH-1, for polyphenol-1 (Figure 48).

Similarly to α -mangostin, the IC_{50} value for PPH-1 was determined for each catalytic site of the proteasome, using concentrations of PPH-1 up to 100 μM . The chymotrypsin-like ($\beta 5$) and caspase-like ($\beta 1$) sites were each inhibited with an IC_{50} of 15 μM (Figures 49 and 50, respectively), while no inhibition was seen against the trypsin-like site ($\beta 2$). Next, the in vitro binding assay was performed to validate PPH-1 as binding to the active sites. The western blot confirmed that PPH-1 does bind to the chymotrypsin-like and caspase-like sites of the proteasome, and also indicated binding to the trypsin-like site, only to a lesser degree, interpreted by diminished band intensity (Figure 47). Due to the nature of the assay, quantitative information could not be obtained, and the binding to

the trypsin-like site is explained by the very high concentration of PPH-1 used for the assay, 100 μ M.

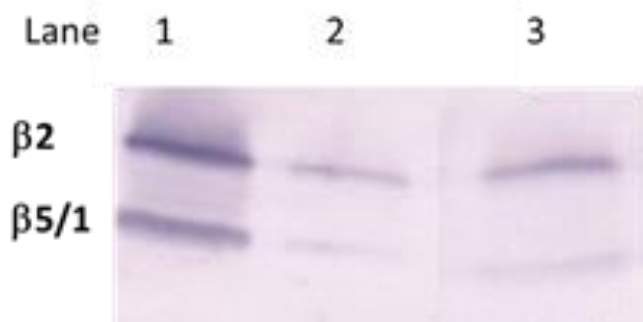


Figure 47 In vitro binding assay of purified human 20S proteasome. Lane 1 is a control with no inhibitor, only probe. Lane 2 had proteasome incubated with bortezomib prior to addition of probe. Lane 3 was proteasome incubated with PPH-1 prior to addition of probe. The diminished band intensities compared to lane 1 indicates binding of the inhibitors.

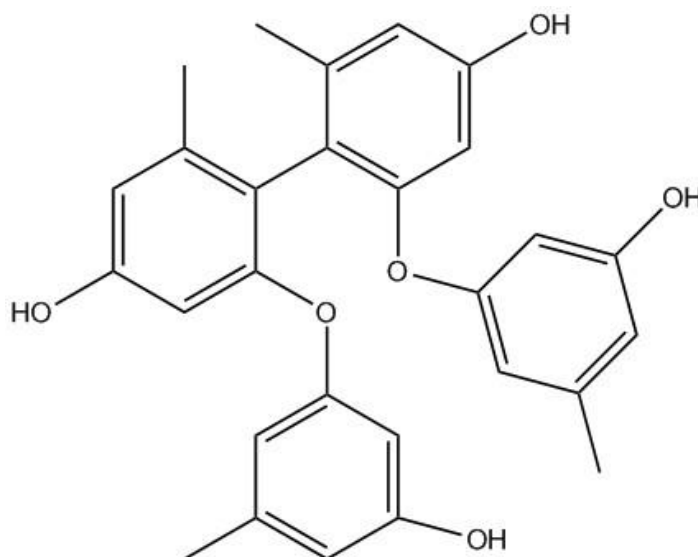


Figure 48 Chemical structure of PPH-1

PPH-1 Chymotrypsin-Like

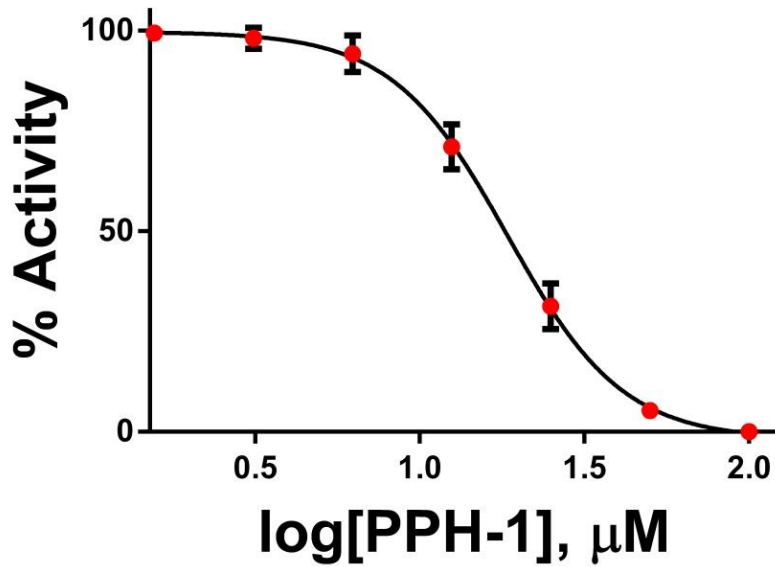


Figure 49 In vitro enzymatic activity of PPH-1 against chymotrypsin-like activity. The IC_{50} against chymotrypsin-like site was determined to be 15 μ M.

PPH-1 Caspase-Like

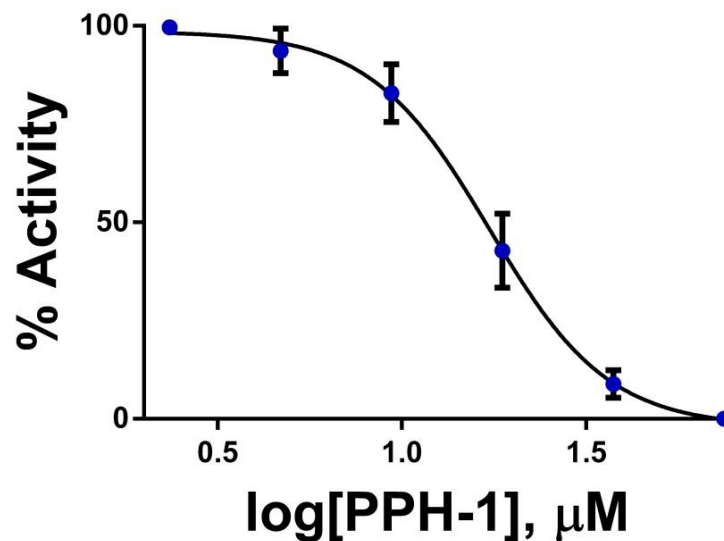


Figure 50 In vitro enzymatic activity PPH-1 against caspase-like sites. The IC_{50} against caspase-like sites was determined to be 15 μ M.

PPH-1 was also tested against a panel of cancer cell lines including; diffuse large B-cell lymphoma line CRL2631, breast cancer cell line MDA-MB-231, and ovarian cancer line OVCAR-3 (Figure 51). Cell lines were cultured in 384-well plates in the presence of PPH-1 at 13 different serial dilutions between 64 μ M to 15.6 nM. After 48 hours resazurin was added. Resazurin is a redox dye that is utilized to measure mitochondrial metabolism in live cells. After overnight incubation with resazurin, plates were read at 570 nM in a POLARstar microplate reader. Graphs were plotted as percent of cell viability as compared to uninhibited cells, versus PPH-1 concentration. Although PPH-1 showed toxicity against the cell lines tested, similar toxicity was also observed against human dermal fibroblasts (data not shown). This suggests that although PPH-1 inhibits the proteasome in vitro, there may be additional targets in intact cells which lead to a general toxicity.

In an attempt to determine the molecular basis for PPH-1's proteasome inhibition, proteasome crystals were soaked with 5mM final concentration dissolved in crystallization buffer for 24 hours. X-ray data for the crystals were collected at Argonne National Labs and processed as described for the carfilzomib soaked crystals. The resolution extended to 3.0 Å, however, the difference electron density did not show the presence of PPH-1 in any of the catalytic subunits. This may be due to the high IC₅₀ value of PPH-1, and higher concentrations during soaking, or longer soaking times, may be necessary to obtain a bound structure.

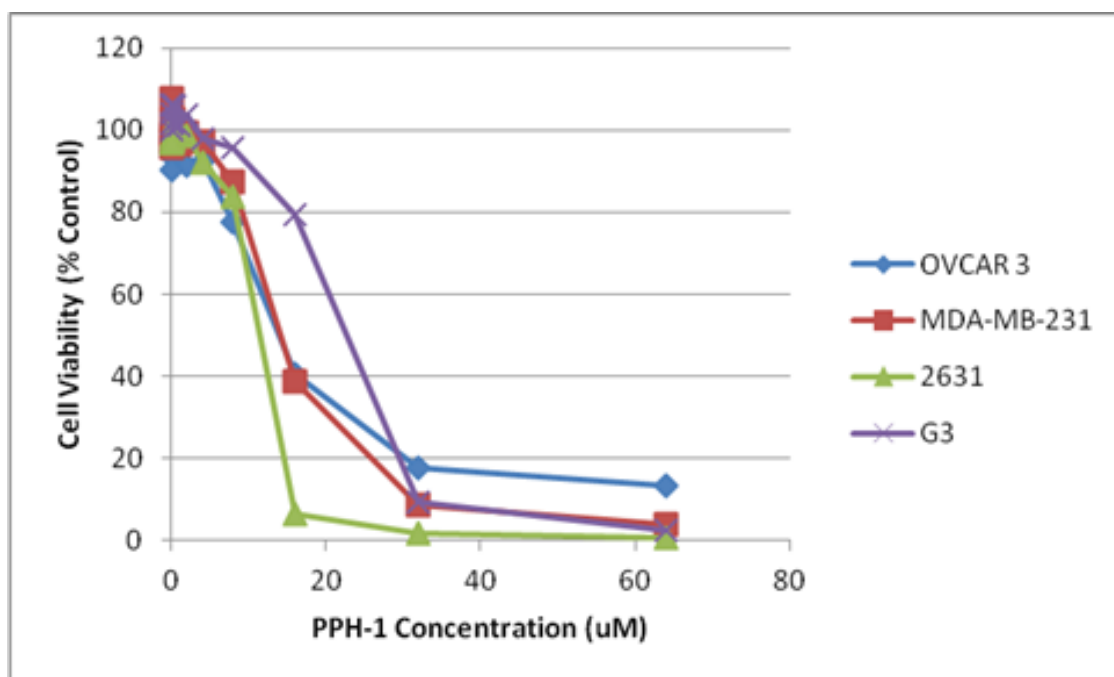


Figure 51 Activity of PPH-1 against cancer cell lines. Cell lines tested were OVCAR 3(ovarian cancer), MDA-MB-231(breast cancer), and multiple myeloma cell lines 2631, and G3.

4.2.3 Lymphoma Hit Plate Screen

The HTS lab at Texas A&M provided a “hit plate” of 313 compounds which show therapeutic potential against the multiple myeloma cell line 2631, and the drug resistant multiple myeloma cell line G3. This was determined by dose response of the compounds against G3 in the presence of a typical therapeutic combination of CHOP (Cyclophosphamide, Hydroxydaunorubicin, Oncovin, and Prednisolone), and a dose response against 2631 in the absence of Chop. The greater the difference between the responses observed on the two cell lines was evaluated, and the top 2% were chosen for the “hit plate”. The compounds screened as candidates for the hit plate were from the Sac1

diversity library of the Sacchettini Research Group at Texas A&M. It contains compounds from ChemBridge and Enamine that generally have $MW \leq 350$ and $cLogP \leq 3$ which make them significantly more water soluble and less lipophilic.

The compounds were screened at a final concentration of 10 μM against the chymotrypsin-like activity of the proteasome (Methods 2.3.2). A threshold of greater than 80% inhibition was used and a total of 7 compounds were identified for a hit rate of 2.2%. These hits are summarized in Table 3.

Table 3 High Through-Put Screening Hits from Hit Plate

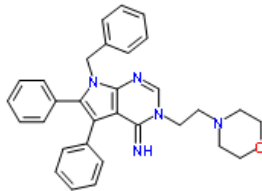
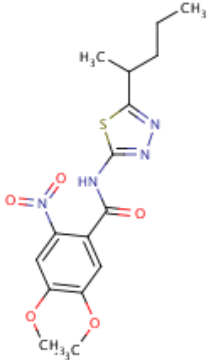
Hit Plate Well	Original Plate ID	Original Plate Well	Compound Structure
C9	N1015-0161	A3	
D9	N1015-0162	H5	

Table 3 continued

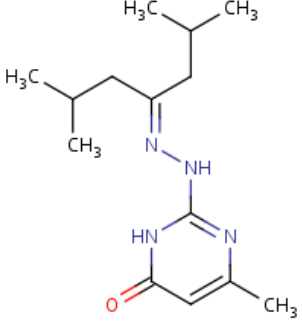
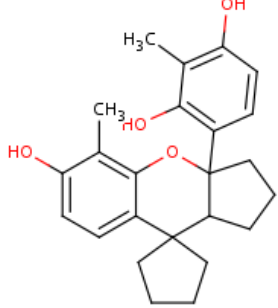
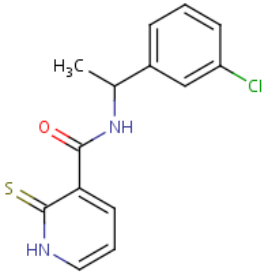
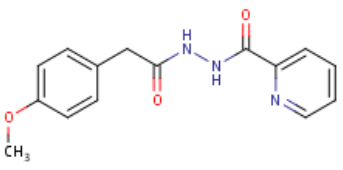
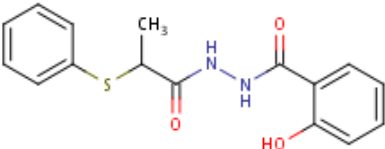
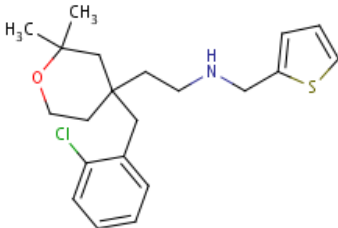
Hit Plate Well	Original Plate ID	Original Plate Well	Compound Structure
D18	EN36478-0016	E3	 <p>The structure shows a 2-methyl-4-oxo-1,2,3,4-tetrahydropyridin-5(1H)-one ring system. Attached to the 3-position of the ring is a secondary amine group (-NH-), which is further substituted with a branched alkyl chain: -CH(CH₃)-CH₂-CH(CH₃)-CH₂-CH₃.</p>
E3	N1015-0005	B11	 <p>The structure is a complex polycyclic molecule featuring a central oxygen atom bridged between two fused ring systems. It includes a cyclopentane ring fused to a six-membered ring, which is further fused to another six-membered ring. Multiple hydroxyl (-OH) groups and methyl (-CH₃) groups are attached to the structure.</p>
E19	EN36478-0147	F5	 <p>The structure consists of a thiopyridine ring system (a six-membered ring with one sulfur and one nitrogen atom). It is substituted with a carbonyl group (-C(=O)-NH-) which is further attached to a 3-chlorophenyl ring.</p>
K19	EN36478-0173	H6	 <p>The structure features a thiopyridine ring system substituted with a methoxy group (-OCH₃) at the 4-position. It is also substituted with a carbonyl group (-C(=O)-NH-) which is further attached to a benzamide group (-NH-C(=O)-C₆H₅).</p>

Table 3 continued

Hit Plate Well	Original Plate ID	Original Plate Well	Compound Structure
L20	EN36478-0225	F3	
P15	CD00000081	C5	

All seven hits were subjected to the secondary binding assay to ascertain if they were truly binding to the proteasome's active site, or if they were false hits. False hits could be due to fluorescent quenching, precipitation, or other experimental anomalies. Bortezomib was used as a positive control for the binding assay which is shown in Figure 53.

None of the hits from the screen showed diminished band intensity compared to the no inhibitor control (Figure 52), indicating that none were binding to the active sites of the proteasome. While we cannot rule out that the hits may be inhibiting the proteasome

allosterically, and therefore may not lead to diminished band intensity, because of the negative result in the binding assay, none of the hits were further pursued for crystallographic studies.

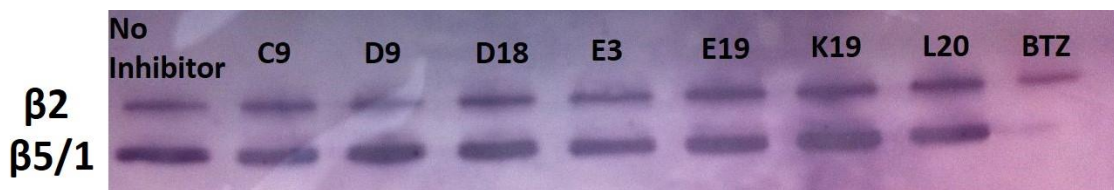


Figure 52 Binding Assay for hit plate screen hits. Bortezomib (BTZ) was used as a positive control for inhibitor binding. None of the hits from the screen show diminished band intensity compared to the “no inhibitor” control, therefore they were not further pursued as possible proteasome inhibitors.

4.3 Conclusions

Since the identification of the proteasome as a target in cancer therapy, several approaches have been utilized for the identification of novel inhibitors, such as image-based screening [120], virtual screening [129], and high-content cell-based screening [121]. Additionally, novel proteasome inhibitors have been synthesized from a peptide core [122], based on structural data obtained from inhibitor complexes with the yeast proteasome.

Here, we show that utilization of a proteasome fluorogenic assay in a high throughput screening manner enabled the identification of two novel proteasome inhibitors, α -mangostin and PPH-1. Verification of binding to the proteasome active sites was confirmed with the molecular probe, Ada-Lys(biotinyl)-(Ahx)₃-(Leu)₃-vinyl

sulphone, and proves that screening with fluorogenic substrates can be a rapid method for identifying compounds with proteasome-inhibiting activity.

α -mangostin is a potent inhibitor of the proteasome in vitro, and was shown to bind to the proteasome in Molt4 leukemia cells, at a concentration which cell death occurs. The results do not, however, conclude that apoptosis is due to proteasome inhibition. Additional experiments must be designed that probe the exact manner and mechanisms of apoptosis for cell death in α -mangostin treated Molt-4 cells.

PPH-1 is a novel polyphenol with potent activity against purified human 20S proteasome. Polyphenols are structurally similar to many plant compounds, that have anti-oxidant properties when taken orally (e.g. flavonoids such as quercetin; common in fruits and vegetables and shown in Figure 53), suggesting favorable bioavailability and low toxicity [130]. Also, compounds in this class are less likely to have activity against other cellular proteases, in contrast to most of the other peptide-based proteasome inhibitors. Polyphenols have been examined before for anti-cancer activity against a number of types of cancer [131]. For example, quercetin has been shown to lead to an immediate inhibition of proteasomal activity in MCF7 and MDA-MB-453 breast cancer cell lines [132]. Additionally, epigallocatechin-3-gallate (EGCG) is in clinical trials as a treatment for chronic lymphocytic leukemia [133].

Polyphenols are known to have a number of liabilities as drugs. For example, some have activity against other (non-protease) targets, such as receptor-tyrosine kinases [134]. Also, most polyphenols are eventually glucuronidated or otherwise metabolized, eventually leading to excretion [135]. Recently however, it has been shown that

peracetylation of the polyphenol proteasome inhibitor epigallocatechin-3-gallate (EGCG) (Figure 53), leads to a prodrug form of the compound [136]. The researchers suggested that esterase's within the cellular milieu convert the prodrug back to its active form, and it was shown to be active against MDA-MB-231 breast cancer cells. These findings indicate therapeutic potential for prodrug forms of polyphenols, giving them potential to be used as either chemosensitizers to reduce the toxicity and boost the effectiveness of current chemotherapeutics such as bortezomib, or as stand-alone treatments.

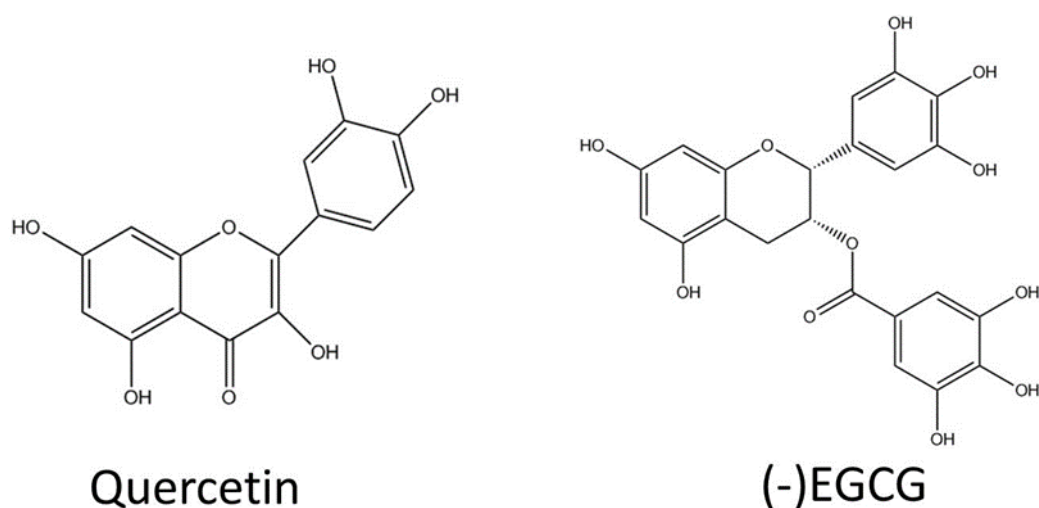


Figure 53 Structures of natural product polyphenol proteasome inhibitors.

Both α -mangostin and PPH-1 are currently being subjected to crystallization trials so that their binding mechanism can be determined. In the case of α -mangostin, a detailed structural description would lead to a new class of proteasome inhibitors, as there are no similar proteasome inhibitor scaffolds described in the literature. While other polyphenol proteasome inhibitors are known, such as the above mentioned EGCG, there are no x-ray

structures to elucidate their binding modes either, therefore both structures could lead to an exciting new avenue of rational drug design against the proteasome.

4.4 Materials and Methods

4.4.1 Proteasome Fluorogenic Assay for Determination of IC₅₀'s

Proteasome activity and the inhibitory potency of ligands are routinely determined by fluorescence spectroscopy using active site specific AMC-peptides (7-Amino-4methylcoumarin) as substrates [137]. For the determination of IC₅₀ values of novel inhibitors, a 96-well Corning black round bottom black polystyrene micro-plate was used to carry out the assay.

In each well, 50 ng of 20S proteasome, purified from human red blood cells (section 2.1), was incubated at 37 °C in assay buffer (20 mM Tris, pH 7.5, 50mM MgCl₂), with inhibitor concentrations serial diluted per column, ranging from 0-100 μM. After 20 minutes of incubation, the reaction was initiated by the addition of a fluorogenic substrate that was specific for the active site being probed (chymotrypsin-like, Suc-Leu-Leu-Val-Tyr-AMC; trypsin-like, Boc-Leu-Arg-Arg-AMC; caspase-like, Z-Leu-Leu-Glu-AMC) (Boston Biochem). The final substrate concentrations of 10 μM, 20 μM, and 20 μM were used for chymotrypsin-like, caspase-like and trypsin-like active sites, respectively, and in each case the substrate was prepared in assay buffer from a 10mM DMSO stock. The rates of hydrolysis of the substrates were monitored using a POLARstar Omega micro plate

reader, with excitation/emission wavelengths of 360 nm/460 nm. The rates of hydrolysis can be directly correlated to the enzymatic activity of the proteasome. The reaction was monitored for 600 seconds, and the linear portion for each curve was used in calculating the IC₅₀ values (inhibitor concentration to cause 50% inhibition of the enzyme active site). This was accomplished by plotting the residual proteasomal activity against the applied inhibitor concentration and fitting the experimental data to the equation: $Y = \text{Bottom} + (\text{Top} - \text{Bottom}) / (1 + 10^{((\text{LogIC}_{50} - X) * \text{HillSlope}))}$ where X is the logarithm of inhibitor concentration and Y is the residual activity. Results are average of triplicate measurements.

4.4.2 High Through-put Screening for Novel Proteasome Inhibitors

Two test compound plates were obtained from the High Throughput Screening Lab (HTS) at Texas A&M University for use in the proteasome enzyme assay to identify novel inhibitors of the chymotrypsin-like activity. The first plate was the National Institute of Health's Clinical Collection 2 (NCC), which is a plate of approximately 450 small molecules that have a history of use in human clinical trials, according to the NIH website (www.nihclinicalcollection.com). The second plate obtained was made by the HTS facility and termed a "hit plate". It is composed of 313 compounds which show therapeutic potential against the multiple myeloma cell line 2631, and the drug resistant multiple myeloma cell line G3. This was determined by a dose response of the compounds against G3 in the presence of a therapeutic combination of CHOP (Cyclophosphamide,

Hydroxydaunorubicin, Oncovin, and Prednisolone), and a dose response against 2631 in the absence of Chop. The greater the difference between the responses observed on the two cell lines was evaluated, and the top 2% were chosen to be included in the “hit plate”.

Additionally, Dr. Nishant Shetty of the Sacchettini Research group at Texas A&M University provided 1mL HPLC fractionated microbial extracts from the fungus *Zygothiala jamaicensis*, which were also tested for their inhibitory potency against the chymotrypsin-like site of the proteasome. Fractions that demonstrated greater than 50% inhibition were re-fractionated until a single agent causing the inhibition could be identified.

In each case, 96-well Corning black round bottom polystyrene micro-plates were used for the assay. 50 ng of purified human 20S proteasome was incubated in assay buffer (20 mM Tris, pH 7.5) at 37 °C in the presence of 10 µM of test compound in each well, or 0.5 µL of the microbial extract fraction. Reactions were initiated by the addition of 20 µL of Suc-Leu-Leu-Val-Tyr-AMC in assay buffer, with a final concentration 50µM. The final reaction volume was 100 µL and the reaction was monitored as described in section 2.3.1. Compounds that showed greater than 80% inhibition compared to control (DMSO) wells (for NCC, and hit plate screen), were selected for an in vitro binding assay (Section 2.4) as a secondary screen to verify the observed inhibition was due to inhibitor binding and not experimental error or anomalies, such as compound absorbance of the fluorescent signal leading to a false positive hit.

4.4.3 In vitro Binding Assay

Ada-Lys(biotinyl)-(Ahx)₃-(Leu)₃-vinyl sulphone is a synthetic proteasome inhibitor, which was designed by Kessler *et al.* to act as a “probe” for the proteasome’s catalytic sites [122], allowing a means to visualize *via* Western blotting whether or not a test compound has displaced the probe, thereby validating the compound as an inhibitor of that active site.

The probe can be divided into three parts: a “warhead” (in this case a vinyl sulphone) which covalently binds to the active site Thr1; a peptide linker which gives subunit specificity; and a visualization tag (biotin) which allows for analysis by Western Blotting with an appropriate antibody (streptavidin)(Figure 54). The probe has comparable affinity ($IC_{50} = 1\mu\text{M}$) for each of the three catalytic subunits of the proteasome ($\beta 1$, $\beta 2$, and $\beta 5$) [122].

A control containing only proteasome and probe but no test compound was used in each trial to verify the presence of three distinct bands after Western Blotting. Subunits $\beta 1$ and $\beta 5$ have approximately the same molecular mass, therefor only two bands were visualized on the gels, with the lower band (indicative of $\beta 1/\beta 5$) appearing thicker than the $\beta 1$ band. Additionally, a sample containing the known inhibitor bortezomib was used in each experiment as a positive control for binding. In the case that an inhibitor is present, the probe is unable to bind to Thr1, and the expected result is either decreased band intensity, or for the bands to be completely removed.

Binding assays were performed on test compounds which showed greater than 80% inhibition against the chymotrypsin-like site of the proteasome during high throughput screening, and against the compound isolated from microbial extracts. The assay was carried out at room temperature (23 °C) in 1.5 mL Eppendorf tubes. 160 µg of purified human 20S proteasome was incubated with 50 µM of the test compound in reaction buffer (20 mM Tris pH 7.5, 50mM MgCl₂). After thirty minutes, Ad-Lys(biotinyl)-(Ahx)₃-(Leu)₃-vinyl sulphone (Enzo Life Sciences) was added at a final concentration of 1 µM to each sample and incubated on ice for 30 minutes. Next, 10µL of SDS loading buffer (10% w/v SDS, 10mM Dithiothreitol, 20% v/v glycerol, 0.2M Tris-HCL, pH 6.8, 0.05% bromophenolblue w/v) was added to each sample, and the samples were run on an SDS-PAGE. The protein was then transferred to a 0.45µm nitrocellulose membrane in transfer buffer (25mM Tris, pH 7.5; 192mM glycine; 20% methanol; 0.1% sodium dodecyl sulfate) for Western Blotting. The membrane was blocked overnight in 6% milk extract in blotting buffer (20mM Tris pH 7.5; 1% NaCl; 0/1% Tween 20). After thoroughly washing the membrane with blotting buffer (three, fifteen minutes soaks of 50mL), streptavidin, alkaline phosphatase conjugated (Sigma Aldrich) was added in 50mL of blotting buffer and 2% milk extract, to a final concentration of 1µg/mL and incubated with the membrane on a shaker for a minimum of 2 hours. The membrane was then washed three times with 50 mL of blotting buffer in 5 minute intervals, and protein bands visualized by the addition of 5mg of alkaline phosphatase substrate (Sigma Aldrich) dissolved in 10mL of water. Bands appeared within 60 seconds, and a diminished band

intensity compared to the control without inhibitor was considered positive for inhibitor binding to the active site.

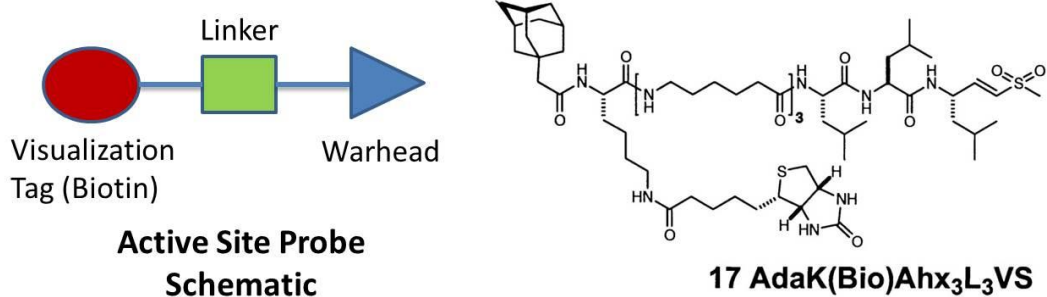


Figure 54 Proteasome binding probe. (Left) Schematic of the probe. (Right) Probe.

4.4.4 Cell Based Binding Assay

Binding of inhibitors to the proteasome in intact Molt4 Leukemia cells was determined using the same proteasome activity probe as described in section 2.4 (Ada-Lys(biotinyl)-(Ahx)₃-(Leu)₃-vinyl sulphone). The probe is cell permeable, however, longer incubation times were necessary to allow for the inhibitors and the probe to each enter the cell and associate with the proteasome. Carfilzomib was used as a positive control for inhibitor binding.

Molt4 Leukemia cells were graciously provided by Dr. Maxwell, Department of Biology, Texas A&M University. The cells were concentrated at 1.5 million cells per mL and 1mL aliquots were used per inhibitor being tested. The cells were incubated with either

10 μ M, or 50 μ M of inhibitor for 2 hours at 37°C. The final concentration of DMSO did not exceed 5%. Controls of 5% DMSO and 10 μ M Carfilzomib were used in each trial. After 2 hours of incubation, the molecular probe, Ada-Lys(biotinyl)-(Ahx)₃-(Leu)₃-vinyl sulphone, was added at a final concentration of 1 μ M, and the cells were incubated for an additional 2 hours at 37°C. The cells were then spun and pelleted at 1600 x G in a bench top Eppendorf 5702 centrifuge for 10 minutes. Cell pellets were re-suspended and lysed with 30 μ L of RIPA buffer (Sigma-Aldrich) on ice for 45 minutes. Lysate was centrifuged at 1600 x G for 10 minutes. Next, 10 μ L of protein loading dye was added to the supernatant, and the protein denatured at 100 °C. The entire supernatant was run on an SDS-PAGE gel. The protein was transferred to a 0.45 μ m nitrocellulose membrane in transfer buffer (25mM Tris, pH 7.5; 192mM glycine; 20% methanol; 0.1% sodium dodecyl sulfate) for Western Blotting. The membrane was blocked overnight in 6% milk extract in blotting buffer (20mM Tris pH 7.5; 1% NaCl; 0.1% Tween 20). After thoroughly washing the membrane with blotting buffer (three, fifteen minutes soaks of 50mL), streptavidin, alkaline phosphatase conjugated (Sigma Aldrich) was added in 50mL of blotting buffer and 2% milk extract, to a final concentration of 1 μ g/mL and incubated with the membrane on a shaker for a minimum of 2 hours. The membrane was then washed three times with 50 mL of blotting buffer in 5 minute intervals, and protein bands visualized by the addition of 5mg of alkaline phosphatase substrate (Sigma Aldrich) dissolved in 10mL of water. Bands appeared within 60 seconds, and a diminished band intensity compared to the control without inhibitor was considered positive for inhibitor binding to the active site.

4.4.5 Cell Based Toxicity Assays

Cell based toxicity assays were performed by Dr. Kim Loesch at Texas A&M University, College Station, Texas. Human dermal fibroblasts (HDF) were purchased from ATCC (Manassas, VA). HDF cells were cultured in DMEM (Lonza) media supplemented with 10% fetal bovine serum (Lonza) and penicillin/streptomycin (Lonza). Test compounds were serially diluted in phosphate buffered saline (PBS) plus 10% DMSO, starting from a maximum concentration of 100mM. HDF cells were trypsinized, counted with a Z1 particle coulter counter (Beckmann) and resuspended in media at a concentration of 64,000 cells/ml. Cells were then plated, over-laid with the inhibitor compound serial dilutions (0-100 μ M), and incubated at 37°C. After 48h, Resazurin dye (7-Hydroxy-3H-phenoxazin-3-one 10-oxide) was added and the assay plates cultured for an additional 24 hours. Resazurin is an oxidation-reduction indicator in cell viability assays. After 24 hours, the absorbance of the Resazurin was measured on a POLARstar Omega plate reader to assess cell death. Cytotoxicity was determined as a percent of dead cells versus living.

5. DUAL PROTEASOME/FATTY ACID SYNTHASE INHIBITORS

5.1 Introduction

The idea of hitting multiple cellular targets with a single drug, or polypharmacology, has emerged as a new concept for drug design by providing a single pharmacokinetic profile, rather than dosing in a “cocktail” fashion with two or more drugs. This strategy allows the design of compounds that purposefully affect multiple expressed targets [138]. The inhibition of multiple sites by a single drug decreases the likelihood that cancer or bacterial cells will develop resistance, due to the need to become resistant along multiple pathways; although, resistance by mechanisms such as drug pumps that remove the drug from the cell prior to it hitting its target, are still possible. The discovery process for polypharmacology drugs is serendipitous, and designing inhibitors aimed at multiple targets with predefined biological profiles is currently a great challenge for medicinal chemists.

There are many classes of proteasome inhibitors, depending on their C-terminal electrophilic group (Figure 55). One such class, are the beta-lactone containing natural and synthetic inhibitor, such as salinosporamide A which is in phase 1 clinical trials for the treatment of multiple myeloma, leukemia, and solid tumors [42]. Another beta-lactone containing proteasome inhibitor, belactosin C and its congener belactosin A, each possess similar inhibitory activity towards the chymotrypsin like activity of the proteasome ($IC_{50} = 0.2 \mu M$). A recent X-ray study of the analog, *N*-CBz-*O*-Bn homobelactosin C, bound to

the yeast 20S proteasome, shows that the *N*-terminal Thr1 in the catalytic pocket of the proteasome is acylated by the beta-lactone moiety, confirming it as a proteasome inhibitor [139].

Fatty acid synthase (FAS), known to catalyze *de novo* lipid synthesis in the fatty acid biosynthetic pathway, is an attractive target for potential anticancer therapy. Found in most human carcinomas, including those of the breast and prostate, a high level of FAS expression and activity is one of the most prevalent changes to the metabolic network in human cancers [140]. FAS inhibition suppresses cell proliferation, adhesion, migration, and invasion. It also leads to suppression of genes involved in production of arachidonic acid and androgen hormones, both of which promote tumor progression [141]. Studies have shown that FAS inhibitors synergize with common anti-cancer therapies (Herceptin® and Taxol®) and in some cases reverse auto-resistance to those therapies, inhibit tumor angiogenesis (new vascularization), and may even act as a chemopreventive for cancer [142]. Healthy cells in the human body are not dependent on FAS activity, rather they take up and use dietary fats, rendering FAS an extremely intriguing therapeutic target.

Orlistat is a beta lactone containing compound in the reduced form of the natural product lipstatin. Being an anti-obesity agent by inhibiting pancreatic lipase, orlistat was approved by the FDA in 1999 as a prescription drug for obesity (Xenical®), and in 2007 as the first over-the-counter weight-loss medication (Alli®). It was found recently that it inhibits the thioesterase (TE) domain of fatty acid synthase (FAS-TE, $IC_{50} = 1.35 \mu M$) and has antitumor activity [143]. X-ray analysis of the orlistat-human FAS-TE complex

reveals the covalent modification of the active site serine via acylation by the beta lactone [144].

The structural similarity between belactosins and orlistat prompted the research group of Dr. Romo at Texas A&M University to design a class of novel dual inhibitors which target both the 20S proteasome and FAS-TE. They synthesized a series of hybrid orlistat/ *N*-CBz-*O*-Bn homobelactosin C analogs, with the hypothesis that they would act as inhibitors of each enzyme. The hypothesis is based on the fact that belactosins and orlistat share a common pharmacophore in the form of a 2*S*,3*S*-bis-substituted beta lactone. The structural hybrids of these two classes of inhibitors could serve as dual inhibitors of both the proteasome and fatty acid synthase and could serve to overcome a major setback associated with anticancer chemotherapies targeting a single target; namely, mutations leading to loss of inhibition to a single target. . The first dual inhibitor **YE-XO2** was designed based on this and premise and is shown in Figure 56.

The crystallographic analysis of bis-benzylhomobelactosin C reveals that its di-peptide side chains play a significant role in dictating which proteasomal active site will be targeted, via interaction with both the S1 specificity pocket and the primed substrate binding site [139]. Belactosins are the only class of proteasome inhibitors which bind to the primed region, whereas all other proteasome inhibitors bind to the nonprimed sites. The amide nitrogen of the C3-side chain of the inhibitor takes a position which blocks the nucleophilic addition of water, and in conjunction with the rings stereochemistry, forces the remainder of the C3-side chain into the primed site of the proteasome. Taking this into consideration, Dr. Romo's group inserted the di-peptide into the hybrid design. Distinct

van der Waals and hydrophobic interactions are made between the ligand and residues in the primed site, and taking advantage of these interactions may lead to better in vitro potency. For example, compound KF33955, a benzyl ester derivative of belactosin A, that is presumably more cell permeable and exhibits greater growth inhibitory activity ($IC_{50} = 0.46 \mu\text{M}$) toward HeLa S3 cells than belactosins A ($IC_{50} = 51 \mu\text{M}$) and C ($IC_{50} = 200 \mu\text{M}$) [145].

In addition, the crystal structure of FAS-TE bound with orlistat, revealed that the C1-C16-carbon skeleton of orlistat, termed the *palmitic core*, binds a predominately hydrophobic channel which they named the *specificity channel* [144]. Since both belactosin and orlistat possess a hydrophobic C3-side chain, Dr. Romo's group used bis-benzyl homobelatocsin C together with orlistat as their lead compounds to design a dual inhibitor.

The C2-hexanoyl chain of orlistat binds in the so-called *short chain pocket*. The mode of binding of belactosins, however, has not been explored for FAS; therefore, it is unknown what role the *sec*-butyl group or the C1' stereocenter might play in enzyme selectivity for this class of inhibitor. Assuming that replacing the *sec*-butyl with a hexanoyl group will not impair the binding to 20S proteasome, they adopted the C2-side chain of orlistat in the initial dual inhibitors. To test this hypothesis along with **YE-OX2** three additional analogs were designed which vary at the C2-side chain, including the one with a *sec*-butyl group. All the compounds were synthesized as mixtures of C2,C3-diastereomers for ease of synthesis. These four compounds were tested for their ability to inhibit the 20S proteasome.

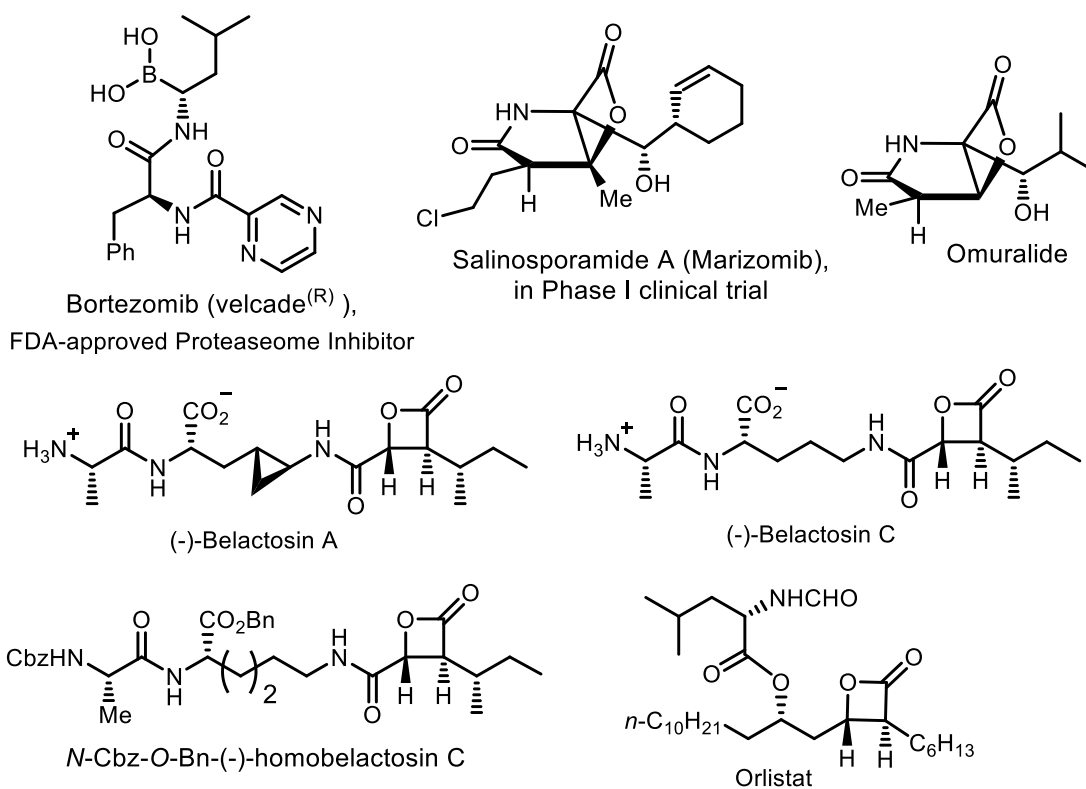


Figure 55 Proteasome and Fatty Acid Synthase Inhibitors. Orlistat (bottom right) inhibits fatty acid synthase, the remaining compounds inhibit the 20S Proteasome.

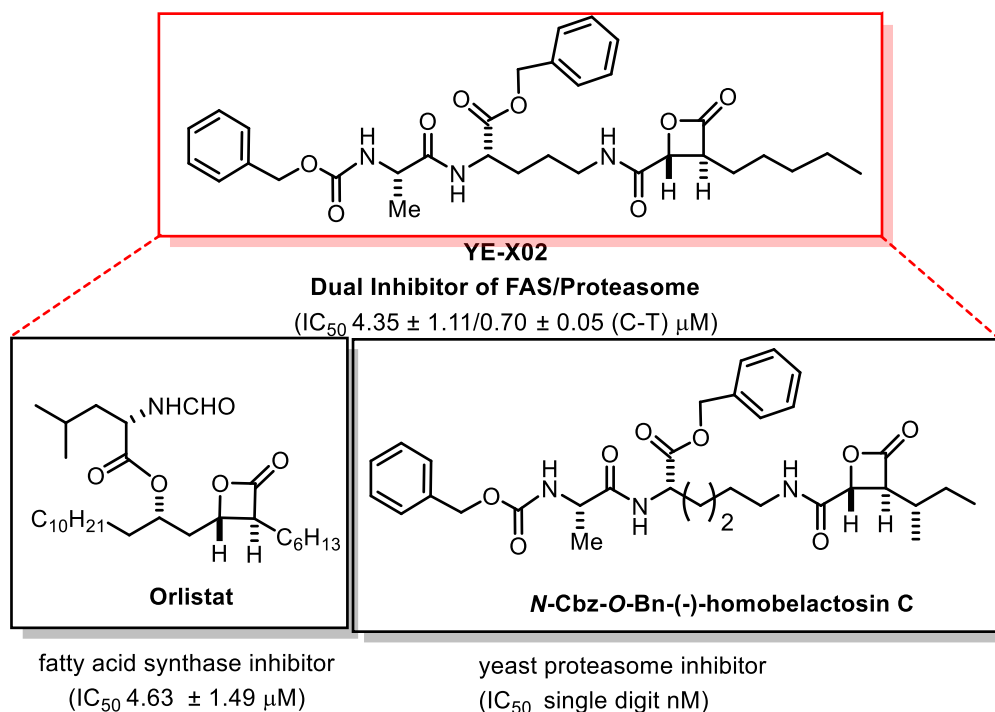


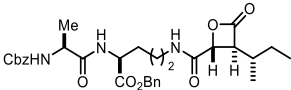
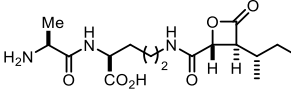
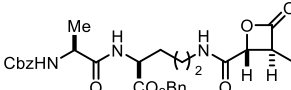
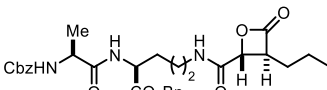
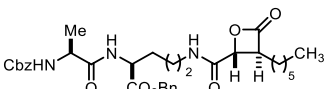
Figure 56 Design of dual proteasome/fatty acid synthase inhibitors from orlistat and N-CBz-O-Bn homobelactosin C

All four compounds were screened for percentage inhibition under the fixed concentration of 0.15 μM of the inhibitor (Table 4). All the compounds inhibited the proteasome at this concentration and the two most potent analogs, DR53 and DR60 demonstrated similar potency to belactosin C (IC₅₀ = 0.21 μM). The results indicate that the C2-side chain tolerates the modifications without significantly deteriorating the binding to the proteasome.

These analogs and belactosin C were then screened for FAS-TE activity and the results are shown in Table 1. To our delight, all four new compounds were more potent than orlistat, with IC₅₀ values ranging from 0.17 to 3.98 μM. Interestingly, belactosin C

was more than 100 times less potent than DR33, indicating that its C3-side chain may enhance the binding to the *specific channel* of FAS-TE. Among the five compounds YE-XO2 and DR33 are the most active. The less bulky C2-side chains of DR59 and DR53 may provide fewer interactions in the *short chain pocket*, which may account for why DR59 and DR53 are more than 10 times less active than DR33 and YE-XO2 against FAS.

Table 4 First Generation of Dual Proteasome/FAS Inhibitors

entry	belactosin derivatives	FAS-TE IC ₅₀ (μM)
1	 (1.3:1 dr) DR33	0.17 ± 0.04
2	 (1.3:1 dr) belactosin C	26.55 ± 0.07
3	 (1:1 dr) DR59	3.98 ± 0.63
4	 (1:1 dr) DR53	2.86 ± 1.75
5	 (1:1 dr) DR60 (YE-XO2)	0.16 ± 0.03

The promising results of the first generation of dual inhibitors prompted Dr. Romo to pursue a second generation of dual inhibitors, which are presented in this thesis. The new inhibitors contain a di-peptide residue and a substituted betalactone with diversified stereocenters and substituents on C2 and C3. The whole molecule can be constructed by amide bond formation between di-peptides and carboxylic acids.

5.2 Results and Discussion

Purified samples of each analog were graciously provided by Dr. Romo of the Department of Chemistry at Texas A&M University, College Station, TX. The analogs were subjected to enzymatic assays against all three catalytic activities of the proteasome (chymotrypsin-like, caspase-like, and trypsin-like) and against the theo-esterase domain of human fatty acid synthase (Methods section 2.3). The results of the assay are expressed as IC₅₀ values and summarized in table 5.

Table 5 IC₅₀ Values of Second Generation Belactosin C Analogs

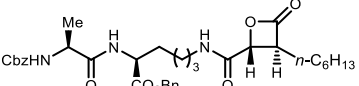
Compound	Chymotrypsin-like (ChTL) IC ₅₀ (μM)	Caspase- like (CL) IC ₅₀ (μM)	Trypsin- like (TL) IC ₅₀ (μM)	FAS-TE IC ₅₀ (μM)
 <p>I-a MZ280</p>	0.497 ± 0.055	3.03 ± 0.053	>100	3.171 ± 0.171

Table 5 continued

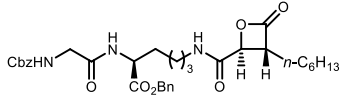
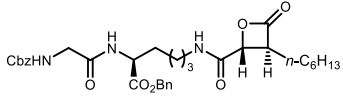
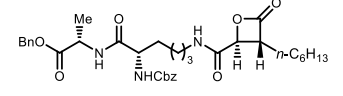
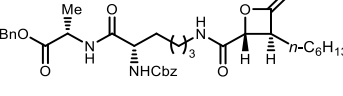
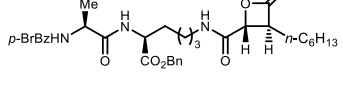
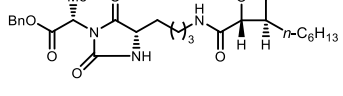
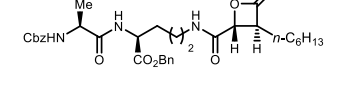
Compound	Chymotrysin-like (ChTL) IC ₅₀ (μM)	Caspase- like (CL) IC ₅₀ (μM)	Trypsin- like (TL) IC ₅₀ (μM)	FAS-TE IC ₅₀ (μM)
 <p>I-b MZ281</p>	1.375 ± 0.025	8.72 ± 0.051	>100	8.370 ± 0.166
 <p>I-c MZ284</p>	0.572 ± 0.032	3.51 ± 0.035	>100	9.563 ± 0.577
 <p>I-d MZ285</p>	2.709 ± 0.169	>100	>100	12.625 ± 0.610
 <p>I-e MZ293</p>	0.374 ± 0.008	> 100	>100	1.504 ± 0.280
 <p>I-f MZ298</p>	0.508 ± 0.013	2.81 ± 0.237	>100	4.812 ± 1.836
 <p>I-g MZ302</p>	0.506 ± 0.015	7.93 ± 0.209	>100	5.546 ± 1.327
 <p>I-h YE-XO2</p>	0.701 ± 0.049	2.16 ± 0.046	>100	4.353 ± 1.109

Table 5 continued

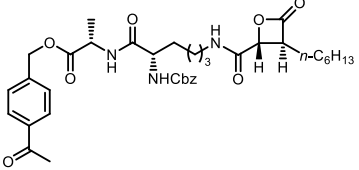
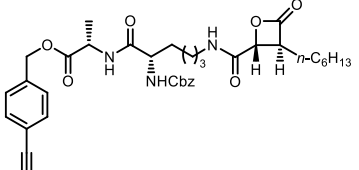
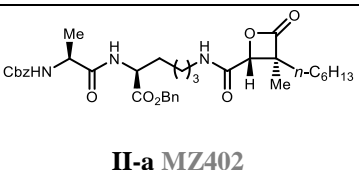
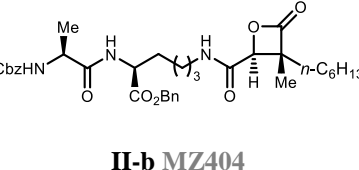
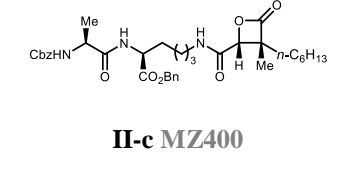
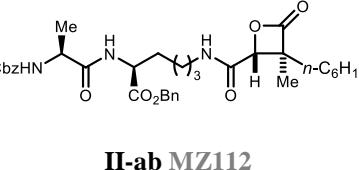
Compound	Chymotrysin-like (ChTL) IC ₅₀ (μM)	Caspase- like (CL) IC ₅₀ (μM)	Trypsin- like (TL) IC ₅₀ (μM)	FAS-TE IC ₅₀ (μM)
 <p>I-i MZ480</p>	0.492 ± 0.002	18.36 ± 0.001	> 100	4.94 ± 0.701
 <p>I-j MZ504</p>	0.265 ± 0.013	> 100	> 100	3.98 ± 0.744
 <p>II-a MZ402</p>	2.43 ± 0.249	3.74 ± 0.045	> 100	2.61 ± 0.944
 <p>II-b MZ404</p>	2.18 ± 0.084	3.98 ± 0.628	> 100	3.11 ± 0.752
 <p>II-c MZ400</p>	3.39 ± 0.076	3.62 ± 0.084	> 100	1.19 ± 0.425
 <p>II-ab MZ112</p>	0.225 ± 0.002	10.11 ± 1.52	>100	0.189 ± 0.151

Table 5 continued

Compound	Chymotrypsin-like (ChTL) IC ₅₀ (μM)	Caspase- like (CL) IC ₅₀ (μM)	Trypsin- like (TL) IC ₅₀ (μM)	FAS-TE IC ₅₀ (μM)
Orlistat	>100	>100	>100	4.63 ± 1.487

All of the analogs were found to have a nanomolar or low micromolar affinity for the chymotrypsin-like subunits of the proteasome, with IC₅₀ values ranging from 0.225 μM for MZ112 to 3.39 μM against MZ400. Caspase-like subunits displayed IC₅₀ values as low as 2.16 μM for YE-XO2, however, no activity was seen against the trypsin-like subunits. Orlistat showed no activity against any of the proteasome active sites.

All analogs also had activity against FAS-TE which was comparable to the parent compound orlistat. The most potent was MZ112 with an IC₅₀ value of 0.189 μM, which is markedly reduced compared to orlistat which inhibited with an IC₅₀ of 4.63 μM.

The only difference in structure between **I-a** and **I-c** is that **I-a** has an alanine residue while **I-c** has a glycine residue. The Lack of a methyl group in **I-c** does not affect its proteasomal inhibitory activity, but makes it about 3 fold less potent against FAS. Replacing the terminus Cbz group of **I-a** with a bromobenzoyl group (**I-f**), conserves the hydrophobic interactions with the binding pocket and the potency was only slightly reduced. With one carbon spacer shorter than **I-a** (replacing L-ornithine with L-lysine), **I-h** has slightly decreased activity against ChTL proteasome and FAS. Compound **I-e** has

an alternative way of assembling the amino acids in comparison with **I-a**, in that the C2-amino group of the lysine residue coupled with alanine. This altered linkage retains the activity and increases the specificity for the chymotrypsin like sites of the proteasome. The potency of **I-e** against FAS was doubled compared to **I-a**. Compound **I-g** has a closer structure to **I-e**, but has a more rigid di-peptide chain. The restricted flexibility did not have a large effect on the chymotrypsin-like activity, but reduced the specificity and the FAS-TE activity. Analogs **I-i** and **I-j** have the similar structure as **I-e**, but have extra substituents on the benzyl ring. The acetylene probe of **I-j** has slightly better chymotrypsin-like activity while maintaining selectivity. The increased polarity imparted by the methyl ketone group on **I-i** is likely the contributing factor for its reduced potency compared to **I-e** and **I-j**, which are presumably each able to make favorable hydrophobic interactions with the primed site and specificity channel of the proteasome and FAS, respectively.

In order to find out how the configuration on the β -lactone affects the activity, we tested analogs **I-b** and **I-c** that have reversed C2- and C3-stereogenic centers relative to naturally occurring belactosins. The change of the configurations in **I-b** decreased its proteasome activity by more than 2 fold compared to its diastereomer **I-a**, but slightly improved FAS activity. On the contrary, both the proteasome and FAS activity of **I-d** was reduced ~8 fold relative to the diastereomer **I-e**. The results validated the importance of the configuration. The mismatched configuration probably makes the molecule fit poorly in the binding pocket. Crystallographic data of bisbenzylbelactosin C [139] reveals that once the molecule binds to the 20S proteasome and the Thr10^γ-CO ester bond is formed,

the di-peptide side chain blocks the attack by water. As to **I-b** and **I-c** the reversed configuration at the beta lactone core probably disables the shielding of the di-peptide chain, so that the covalent binding is short-lived with reduced inhibition.

Replacing the C2-proton of **I-a** with a methyl group (**II-a**) impaired its ability to inhibit the chymotrypsin like sites of the proteasome. The other two isomers **II-b** and **II-c** followed the same trend of reduced potency. Even though the methyl group increased the stability of β -lactone in the medium during the assay process, the steric hindrance caused by the methyl group may impede the covalent attachment to ThrO γ of the proteasome.

5.3 Conclusions

Polypharmacology has emerged as a new concept for drug design, based on the philosophy that one drug can hit multiple targets [146]. A drug with selective polypharmacology provides a single pharmacokinetic profile and potentially improves efficacy [138]. The inhibition of multiple sites makes the likelihood of the cancer cells developing resistance to the drug more difficult, due to the need to become resistant along multiple pathways.

There are several examples of polypharmacological drugs in the literature. Lapatinib, for instance, targets both EGFR and HER2 tyrosine kinase, resulting in down regulation of thymidylate synthase [147], significantly inhibiting the proliferation of cancer cells both in vitro and in vivo [148]. Bromodomains, which have only recently been identified as druggable targets, show compelling rationale for the development of dual kinase-bromodomain inhibitors as therapeutics for both oncology and inflammatory

disease. For example, FLT3 receptor tyrosine kinase and BRD4 are both drivers in acute myelogenous leukemia (AML) [149, 150], JAK kinase and BRD4 inhibitors show complementary tumor and host microenvironment activities in multiple myeloma models [151], and both bromodomain and kinase inhibitors have shown compelling efficacy in inflammatory disease [152, 153].

The structural similarity between the belactosin class of proteasome inhibitors and the fatty acid synthase (FAS) inhibitor orlistat, prompted the design of dual proteasome/FAS inhibitors. The orlistat analogs presented in this chapter show that potency is able to be maintained against FAS, while simultaneously showing nanomolar potency against the proteasome in vitro. Compounds I-e and II-a were also confirmed to have toxicity against the cervical cancer cell line HeLa with an IC_{50} of approximately $40\mu M$, and shown to target both 20S proteasomes and FAS in whole cells using activity-based protein profiling (data not shown). Future experiments will be aimed at determining the complex structures of the orlistat analogs with the 20S proteasome and FAS, to allow for structure guided design of more potent inhibitors which may eventually lead to a clinically relevant drug.

5.4 Materials and Methods

The the-esterase domain of human fatty acid synthase (FAS-TE) was tested alongside the proteasome for its ability to inhibit a series of tetrahydrolipstatin (orlistat) analogs, which were graciously provided by Dr. Daniel Romo of the Department of

Chemistry at Texas A&M University, College Station, TX (See Chapter 6). IC₅₀'s against the three catalytic sites of the proteasome were determined in the same fashion as as described in section 4.4.1.

Purified FAS-TE was provided by Dr. Romo at a stock concentration of 1mg/mL in 20mM Tris pH 7.5, 150mM NaCl, and 10% glycerol. The synthetic fluorogenic substrate for FAS-TE, 4-methylumbelliferyl heptanoate (4-MUH), was purchased from Sigma (St. Louis, MO). The reaction mixture consisted of 500nM FAS- TE in buffer A (100mM Tris-HCl, 50 mM NaCl at pH 7.4) which was pre-incubated with 2.5 μ L test compounds (orlistat or analog) dissolved in DMSO at final concentrations of 0.32-100 μ M and/or 0.08-10 μ M at 37 °C for 30 minutes. The reaction was initiated by addition of 5 μ L of 1.25 mM 4-MUH in 1:1 DMSO:buffer A. The resulting fluorescence from liberated 4-methylumbelliferone was measured every five minutes at 350_{ex}/450_{em}nm for 60 minutes. 4-MUH auto cleaves without enzyme present; therefore, it was incubated without enzyme and served as a background control which was subtracted from all rates. Results are the average of triplicate time points.

6. CONCLUSIONS

For over twenty years the proteasome has been an attractive target for the development of chemotherapeutic agents, due to being a critical regulator of cell growth and apoptosis. The X-ray structure of the *T. acidophilum* 20S proteasome was reported in 1996 [19], and the first X-ray structure of Eukaryotic proteasomes was reported for *S. cerevisiae* in 1998 [21]. Additionally, the X-ray structures of the mammalian proteasomes for bovine [92], and mouse [16], have also been reported. As of June, 2014, there are over 106 ligand complex structures with proteasomes entered in the protein data bank, none of which are the human proteasome.

Carfilzomib was granted FDA approval in 2012, and although the yeast:epoxomicin structure has been reported [61], carfilzomib contains different side chain substituents at the P2, P3 and P4 positions, and its binding to the 20S proteasome has not been described until now. The presented crystallographic analysis of the human constitutive 20S proteasome, alone and in complex with carfilzomib, emphasize the biological impact of distinct interactions in the substrate binding channel that influence the drugs selectivity not only among the different catalytic sites of constitutive proteasomes, but also immunoproteasome isoforms.

The presented x-ray structures do not, however, provide dynamic insights to the 20S proteasome which may have substantial effect on substrate recognition, binding, and inhibition. It is interesting that both chymotrypsin-like and trypsin-like sites bind carfilzomib with very similar interactions. Both sites form six hydrogen bonding

interactions, one of which comes from D125 of the adjoining subunits. Additionally they each have hydrophobic S3 and S4 pockets that make van der Waals interactions with carfilzomib's side chains. While the S1 pocket of chymotrypsin-like sites requires structural rearrangements to accommodate carfilzomib's P1 leucyl group, trypsin-like sites are large and provide several van der Waals interactions without the need for movement. Therefore, based solely on the structural data, it would be tempting to think that trypsin-like sites have better selectivity for carfilzomib; however IC₅₀ values clearly show that chymotrypsin-like sites are 1000-fold better inhibited, highlighting the limitation of structural data on proteasome binding dynamics.

The importance of the interaction of the P1 site and the S1 pocket was demonstrated in the mouse constitutive and immunoproteasome crystal structures [16], however, the carfilzomib complex presented here shows that the S3 and S4 pockets play a critical role as well. While PR-957 has a bulky phenyl P1 side chain and preferentially targets immunoproteasome chymotrypsin-like sites, the epoxyketone oprozomib (ONX 0912) also has a P1 phenyl but displays slight preference for constitutive proteasome chymotrypsin-like sites. This can be explained by the P3 methoxy serine which would form van der Waals interactions with the constitutive proteasome S3 pocket, but not the immunoproteasome, which prefers small P3 residues such as alanine, because the conserved substitution A27S diminishes the size of the S3 pocket and adds polarity [16]. It would be interesting to see if changing the P3 substituent of the immunoproteasome specific inhibitors PR-957 and PR-924 from alanine to leucine would change their selectivity to favor constitutive proteasomes. Additionally, adding a hydrophobic P4

residue to the immunoproteasome specific inhibitors may also shift their selectivity, due to the substitution of G48C which would act to electrostatically hinder binding to the S4 pocket.

Selective targeting of trypsin-like active sites has been accomplished with a peptide based inhibitor NC-022 that contain a P1 arginine and a P3 leucyl [38], which is favored by the large negatively charged S1 pocket and the hydrophobic S3 pocket. Caspase-like specific inhibitors such as NC-001[68] contains a P1 leucyl group and a P3 proline. This carfilzomib bound proteasome structure presented shows that bulky residues at the P3 position sterically clash with H116 of the S3 pocket, possibly lending an explanation for NC-001 selectivity.

The primary work undertaken in this thesis, the high-resolution crystal structure of carfilzomib in complex with the primary target, the human 20S proteasome, represents an important advance for the development of new inhibitor drugs, targeting not only the constitutive core particle, but also immunoproteasome core particles. This thesis provides a platform for a detailed description of binding mechanisms and optimization strategies for future therapeutics with potential applications for the treatment of cancer and autoimmune diseases.

REFERENCES

- [1] Koepp, D.M., *Cell cycle regulation by protein degradation*. Methods Mol Biol, 2014. **1170**: p. 61-73.
- [2] Zhang, L., R. Sheng, and Z. Qin, *The lysosome and neurodegenerative diseases*. Acta Biochim Biophys Sin (Shanghai), 2009. **41**(6): p. 437-45.
- [3] Goldberg, A.L. and K.L. Rock, *Proteolysis, proteasomes and antigen presentation*. Nature, 1992. **357**(6377): p. 375-9.
- [4] Reed, S.I., *The ubiquitin-proteasome pathway in cell cycle control*. Results Probl Cell Differ, 2006. **42**: p. 147-81.
- [5] Muratani, M. and W.P. Tansey, *How the ubiquitin-proteasome system controls transcription*. Nat Rev Mol Cell Biol, 2003. **4**(3): p. 192-201.
- [6] Gaczynska, M., K.L. Rock, and A.L. Goldberg, *Role of proteasomes in antigen presentation*. Enzyme Protein, 1993. **47**(4-6): p. 354-69.
- [7] Etlinger, J.D. and A.L. Goldberg, *A soluble ATP-dependent proteolytic system responsible for the degradation of abnormal proteins in reticulocytes*. Proc Natl Acad Sci U S A, 1977. **74**(1): p. 54-8.
- [8] Voges, D., P. Zwickl, and W. Baumeister, *The 26S proteasome: a molecular machine designed for controlled proteolysis*. Annu Rev Biochem, 1999. **68**: p. 1015-68.
- [9] Choi, H.J. and B.T. Zhu, *Role of cyclin B1/Cdc2 in mediating Bcl-XL phosphorylation and apoptotic cell death following nocodazole-induced mitotic arrest*. Mol Carcinog, 2014. **53**(2): p. 125-37.
- [10] Sudakin, V., et al., *The cyclosome, a large complex containing cyclin-selective ubiquitin ligase activity, targets cyclins for destruction at the end of mitosis*. Mol Biol Cell, 1995. **6**(2): p. 185-97.

- [11] Yamano, H., et al., *Cell cycle-regulated recognition of the destruction box of cyclin B by the APC/C in Xenopus egg extracts*. Mol Cell, 2004. **13**(1): p. 137-47.
- [12] Ritz, U. and B. Seliger, *The transporter associated with antigen processing (TAP): structural integrity, expression, function, and its clinical relevance*. Mol Med, 2001. **7**(3): p. 149-58.
- [13] Engelhard, V.H., *Structure of peptides associated with class I and class II MHC molecules*. Annu Rev Immunol, 1994. **12**: p. 181-207.
- [14] Vyas, J.M., A.G. Van der Veen, and H.L. Ploegh, *The known unknowns of antigen processing and presentation*. Nat Rev Immunol, 2008. **8**(8): p. 607-18.
- [15] Hershko, A. and A. Ciechanover, *The ubiquitin system*. Annu Rev Biochem, 1998. **67**: p. 425-79.
- [16] Huber, E.M., et al., *Immuno- and constitutive proteasome crystal structures reveal differences in substrate and inhibitor specificity*. Cell, 2012. **148**(4): p. 727-38.
- [17] Burger, A.M. and A.K. Seth, *The ubiquitin-mediated protein degradation pathway in cancer: therapeutic implications*. Eur J Cancer, 2004. **40**(15): p. 2217-29.
- [18] Rohrwild, M., et al., *The ATP-dependent HslVU protease from Escherichia coli is a four-ring structure resembling the proteasome*. Nat Struct Biol, 1997. **4**(2): p. 133-9.
- [19] Lowe, J., et al., *Crystal structure of the 20S proteasome from the archaeon T. acidophilum at 3.4 Å resolution*. Science, 1995. **268**(5210): p. 533-9.
- [20] Kish-Trier, E. and C.P. Hill, *Structural biology of the proteasome*. Annu Rev Biophys, 2013. **42**: p. 29-49.
- [21] Groettrup, M., C.J. Kirk, and M. Basler, *Proteasomes in immune cells: more than peptide producers?* Nat Rev Immunol, 2010. **10**(1): p. 73-8.

- [22] Groll, M., et al., *The catalytic sites of 20S proteasomes and their role in subunit maturation: a mutational and crystallographic study*. Proc Natl Acad Sci U S A, 1999. **96**(20): p. 10976-83.
- [23] Matias, A.C., P.C. Ramos, and R.J. Dohmen, *Chaperone-assisted assembly of the proteasome core particle*. Biochem Soc Trans, 2010. **38**(Pt 1): p. 29-33.
- [24] Ramos, P.C. and R.J. Dohmen, *PACemakers of proteasome core particle assembly*. Structure, 2008. **16**(9): p. 1296-304.
- [25] Oinonen, C. and J. Rouvinen, *Structural comparison of Ntn-hydrolases*. Protein Sci, 2000. **9**(12): p. 2329-37.
- [26] Smith, D.M., et al., *Docking of the proteasomal ATPases' carboxyl termini in the 20S proteasome's alpha ring opens the gate for substrate entry*. Mol Cell, 2007. **27**(5): p. 731-44.
- [27] Kusmierczyk, A.R., et al., *A conserved 20S proteasome assembly factor requires a C-terminal HbYX motif for proteasomal precursor binding*. Nature Structural & Molecular Biology, 2011. **18**(5): p. 622-U208.
- [28] Osmulski, P.A. and M. Gaczynska, *Rapamycin allosterically inhibits the proteasome*. Mol Pharmacol, 2013. **84**(1): p. 104-13.
- [29] Deriziotis, P., et al., *Misfolded PrP impairs the UPS by interaction with the 20S proteasome and inhibition of substrate entry*. EMBO J, 2011. **30**(15): p. 3065-77.
- [30] Murata, S., et al., *Regulation of CD8+ T cell development by thymus-specific proteasomes*. Science, 2007. **316**(5829): p. 1349-53.
- [31] Singh, S., et al., *Immunoproteasome expression in a nonimmune tissue, the ocular lens*. Arch Biochem Biophys, 2002. **405**(2): p. 147-53.
- [32] Gaczynska, M., et al., *Peptidase activities of proteasomes are differentially regulated by the major histocompatibility complex-encoded genes for LMP2 and LMP7*. Proc Natl Acad Sci U S A, 1994. **91**(20): p. 9213-7.

- [33] Kincaid, E.Z., et al., *Mice completely lacking immunoproteasomes show major changes in antigen presentation*. Nat Immunol, 2012. **13**(2): p. 129-35.
- [34] Guillaume, B., et al., *Two abundant proteasome subtypes that uniquely process some antigens presented by HLA class I molecules*. Proc Natl Acad Sci U S A, 2010. **107**(43): p. 18599-604.
- [35] Xing, Y., S.C. Jameson, and K.A. Hogquist, *Thymoproteasome subunit-beta5T generates peptide-MHC complexes specialized for positive selection*. Proc Natl Acad Sci U S A, 2013. **110**(17): p. 6979-84.
- [36] Groll, M., et al., *Crystal structure of the boronic acid-based proteasome inhibitor bortezomib in complex with the yeast 20S proteasome*. Structure, 2006. **14**(3): p. 451-6.
- [37] Elofsson, M., et al., *Towards subunit-specific proteasome inhibitors: synthesis and evaluation of peptide alpha',beta'-epoxyketones*. Chem Biol, 1999. **6**(11): p. 811-22.
- [38] Mirabella, A.C., et al., *Specific cell-permeable inhibitor of proteasome trypsin-like sites selectively sensitizes myeloma cells to bortezomib and carfilzomib*. Chem Biol, 2011. **18**(5): p. 608-18.
- [39] Borissenko, L. and M. Groll, *20S proteasome and its inhibitors: crystallographic knowledge for drug development*. Chem Rev, 2007. **107**(3): p. 687-717.
- [40] Rock, K.L., et al., *Protein degradation and the generation of MHC class I-presented peptides*. Adv Immunol, 2002. **80**: p. 1-70.
- [41] Gascoigne, N.R. and E. Palmer, *Signaling in thymic selection*. Curr Opin Immunol, 2011. **23**(2): p. 207-12.
- [42] Murata, S., Y. Takahama, and K. Tanaka, *Thymoproteasome: probable role in generating positively selecting peptides*. Curr Opin Immunol, 2008. **20**(2): p. 192-6.

- [43] Salomons, F.A., et al., *Selective accumulation of aggregation-prone proteasome substrates in response to proteotoxic stress*. Mol Cell Biol, 2009. **29**(7): p. 1774-85.
- [44] Bossis, G., et al., *c-Fos proto-oncoprotein is degraded by the proteasome independently of its own ubiquitinylation in vivo*. Mol Cell Biol, 2003. **23**(20): p. 7425-36.
- [45] Auld, K.L., et al., *Genomic association of the proteasome demonstrates overlapping gene regulatory activity with transcription factor substrates*. Mol Cell, 2006. **21**(6): p. 861-71.
- [46] Glickman, M.H. and A. Ciechanover, *The ubiquitin-proteasome proteolytic pathway: destruction for the sake of construction*. Physiol Rev, 2002. **82**(2): p. 373-428.
- [47] Hoyt, M.A., M. Zhang, and P. Coffino, *Probing the ubiquitin/proteasome system with ornithine decarboxylase, a ubiquitin-independent substrate*. Methods Enzymol, 2005. **398**: p. 399-413.
- [48] Frankland-Searby, S. and S.R. Bhaumik, *The 26S proteasome complex: an attractive target for cancer therapy*. Biochim Biophys Acta, 2012. **1825**(1): p. 64-76.
- [49] Imajoh-Ohmi, S., et al., *Lactacystin, a specific inhibitor of the proteasome, induces apoptosis in human monoblast U937 cells*. Biochem Biophys Res Commun, 1995. **217**(3): p. 1070-7.
- [50] Kim, M., et al., *Activation of the programmed cell death pathway by inhibition of proteasome function in plants*. J Biol Chem, 2003. **278**(21): p. 19406-15.
- [51] Voorhees, P.M., et al., *The proteasome as a target for cancer therapy*. Clin Cancer Res, 2003. **9**(17): p. 6316-25.
- [52] McConkey, D.J. and K. Zhu, *Mechanisms of proteasome inhibitor action and resistance in cancer*. Drug Resist Updat, 2008. **11**(4-5): p. 164-79.

- [53] Bhatia, U., et al., *Induction of apoptosis and cell cycle-specific change in expression of p53 in normal lymphocytes and MOLT-4 leukemic cells by nitrogen mustard*. Clin Cancer Res, 1995. **1**(8): p. 873-80.
- [54] Kiliccioglu, I., et al., *Apoptotic effects of proteasome and histone deacetylase inhibitors in prostate cancer cell lines*. Genet Mol Res, 2014. **13**(2): p. 3721-31.
- [55] Traenckner, E.B., S. Wilk, and P.A. Baeuerle, *A proteasome inhibitor prevents activation of NF-kappa B and stabilizes a newly phosphorylated form of I kappa B-alpha that is still bound to NF-kappa B*. EMBO J, 1994. **13**(22): p. 5433-41.
- [56] Amschler, K., et al., *NF-kappaB inhibition through proteasome inhibition or IKKbeta blockade increases the susceptibility of melanoma cells to cytostatic treatment through distinct pathways*. J Invest Dermatol, 2010. **130**(4): p. 1073-86.
- [57] Nickeleit, I., et al., *Argyria reveals a critical role for the tumor suppressor protein p27(kip1) in mediating antitumor activities in response to proteasome inhibition*. Cancer Cell, 2008. **14**(1): p. 23-35.
- [58] Tsubuki, S., et al., *Differential inhibition of calpain and proteasome activities by peptidyl aldehydes of di-leucine and tri-leucine*. J Biochem, 1996. **119**(3): p. 572-6.
- [59] Groll, M., et al., *A plant pathogen virulence factor inhibits the eukaryotic proteasome by a novel mechanism*. Nature, 2008. **452**(7188): p. 755-8.
- [60] Ettari, R., et al., *Development of novel peptidomimetics containing a vinyl sulfone moiety as proteasome inhibitors*. ChemMedChem, 2011. **6**(7): p. 1228-37.
- [61] Groll, M., et al., *Crystal structure of epoxomicin : 20S proteasome reveals a molecular basis for selectivity of alpha ',beta '-epoxyketone proteasome inhibitors*. Journal of the American Chemical Society, 2000. **122**(6): p. 1237-1238.

- [62] Tzvetanov, I., et al., *The use of bortezomib as a rescue treatment for acute antibody-mediated rejection: report of three cases and review of literature*. *Transplant Proc*, 2012. **44**(10): p. 2971-5.
- [63] Escobar, M., et al., *The role of proteasome inhibition in nonsmall cell lung cancer*. *J Biomed Biotechnol*, 2011. **2011**: p. 806506.
- [64] Mujtaba, T. and Q.P. Dou, *Advances in the understanding of mechanisms and therapeutic use of bortezomib*. *Discov Med*, 2011. **12**(67): p. 471-80.
- [65] Arastu-Kapur, S., et al., *Nonproteasomal targets of the proteasome inhibitors Jbortezomib and carfilzomib: a link to clinical adverse events*. *Clin Cancer Res*, 2011. **17**(9): p. 2734-43.
- [66] Demo, S.D., et al., *Antitumor activity of PR-171, a novel irreversible inhibitor of the proteasome*. *Cancer Res*, 2007. **67**(13): p. 6383-91.
- [67] Zhou, H.J., et al., *Design and synthesis of an orally bioavailable and selective peptide epoxyketone proteasome inhibitor (PR-047)*. *J Med Chem*, 2009. **52**(9): p. 3028-38.
- [68] Britton, M., et al., *Selective inhibitor of proteasome's caspase-like sites sensitizes cells to specific inhibition of chymotrypsin-like sites*. *Chem Biol*, 2009. **16**(12): p. 1278-89.
- [69] Muchamuel, T., et al., *A selective inhibitor of the immunoproteasome subunit LMP7 blocks cytokine production and attenuates progression of experimental arthritis*. *Nat Med*, 2009. **15**(7): p. 781-7.
- [70] Basler, M., et al., *Prevention of experimental colitis by a selective inhibitor of the immunoproteasome*. *J Immunol*, 2010. **185**(1): p. 634-41.
- [71] Ichikawa, H.T., et al., *Beneficial effect of novel proteasome inhibitors in murine lupus via dual inhibition of type I interferon and autoantibody-secreting cells*. *Arthritis Rheum*, 2012. **64**(2): p. 493-503.

- [72] Nagayama, Y., et al., *Prophylactic and therapeutic efficacies of a selective inhibitor of the immunoproteasome for Hashimoto's thyroiditis, but not for Graves' hyperthyroidism, in mice*. Clin Exp Immunol, 2012. **168**(3): p. 268-73.
- [73] Parlati, F., et al., *Carfilzomib can induce tumor cell death through selective inhibition of the chymotrypsin-like activity of the proteasome*. Blood, 2009. **114**(16): p. 3439-47.
- [74] Kazi, A., et al., *Discovery of PI-1840, a novel noncovalent and rapidly reversible proteasome inhibitor with anti-tumor activity*. J Biol Chem, 2014. **289**(17): p. 11906-15.
- [75] Bethanis, K., et al., *Convergence study of a Schrodinger-equation algorithm and structure-factor determination from the wavefunction*. Acta Crystallogr A, 2008. **64**(Pt 4): p. 450-8.
- [76] Spek, A.L., *Structure validation in chemical crystallography*. Acta Crystallogr D Biol Crystallogr, 2009. **65**(Pt 2): p. 148-55.
- [77] Hauptman, H., *Phasing methods for protein crystallography*. Curr Opin Struct Biol, 1997. **7**(5): p. 672-80.
- [78] McCoy, A.J., et al., *Phaser crystallographic software*. J Appl Crystallogr, 2007. **40**(Pt 4): p. 658-674.
- [79] Abergel, C., *Molecular replacement: tricks and treats*. Acta Crystallogr D Biol Crystallogr, 2013. **69**(Pt 11): p. 2167-73.
- [80] Vagin, A. and A. Teplyakov, *An approach to multi-copy search in molecular replacement*. Acta Crystallographica Section D-Biological Crystallography, 2000. **56**: p. 1622-1624.
- [81] Navaza, J., *Amore - an Automated Package for Molecular Replacement*. Acta Crystallographica Section A, 1994. **50**: p. 157-163.

- [82] Adams, P.D., et al., *PHENIX: a comprehensive Python-based system for macromolecular structure solution*. Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 2): p. 213-21.
- [83] Holton, J.M., et al., *The R-factor gap in macromolecular crystallography: an untapped potential for insights on accurate structures*. FEBS J, 2014. **281**(18): p. 4046-60.
- [84] Otwinowski, Z. and W. Minor, *Processing of X-ray diffraction data collected in oscillation mode*. Macromolecular Crystallography, Pt A, 1997. **276**: p. 307-326.
- [85] Winn, M.D., et al., *Overview of the CCP4 suite and current developments*. Acta Crystallogr D Biol Crystallogr, 2011. **67**(Pt 4): p. 235-42.
- [86] Afonine, P.V., et al., *Towards automated crystallographic structure refinement with phenix.refine*. Acta Crystallogr D Biol Crystallogr, 2012. **68**(Pt 4): p. 352-67.
- [87] Emsley, P., et al., *Features and development of Coot*. Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 4): p. 486-501.
- [88] Rabl, J., et al., *Mechanism of gate opening in the 20S proteasome by the proteasomal ATPases*. Mol Cell, 2008. **30**(3): p. 360-8.
- [89] Wilk, S., W.E. Chen, and R.P. Magnusson, *Modulators of the activation of the proteasome by PA28 (11S reg)*. Mol Biol Rep, 1999. **26**(1-2): p. 39-44.
- [90] Ramos, P.C., et al., *Role of C-terminal extensions of subunits beta2 and beta7 in assembly and activity of eukaryotic proteasomes*. J Biol Chem, 2004. **279**(14): p. 14323-30.
- [91] Chen, P. and M. Hochstrasser, *Autocatalytic subunit processing couples active site formation in the 20S proteasome to completion of assembly*. Cell, 1996. **86**(6): p. 961-72.

- [92] Unno, M., et al., *The structure of the mammalian 20S proteasome at 2.75 Å resolution*. Structure, 2002. **10**(5): p. 609-18.
- [93] Kisselev, A.F. and A.L. Goldberg, *Monitoring activity and inhibition of 26S proteasomes with fluorogenic peptide substrates*. Methods Enzymol, 2005. **398**: p. 364-78.
- [94] Dick, T.P., et al., *Contribution of proteasomal beta-subunits to the cleavage of peptide substrates analyzed with yeast mutants*. Journal of Biological Chemistry, 1998. **273**(40): p. 25637-25646.
- [95] McCormack, T.A., et al., *Kinetic studies of the branched chain amino acid preferring peptidase activity of the 20S proteasome: Development of a continuous assay and inhibition by tripeptide aldehydes and clasto-lactacystin beta-lactone*. Biochemistry, 1998. **37**(21): p. 7792-7800.
- [96] Kantardjieff, K.A. and B. Rupp, *Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals*. Protein Sci, 2003. **12**(9): p. 1865-71.
- [97] Fenteany, G., et al., *Inhibition of proteasome activities and subunit-specific amino-terminal threonine modification by lactacystin*. Science, 1995. **268**(5211): p. 726-31.
- [98] Macherla, V.R., et al., *Structure-activity relationship studies of salinosporamide A (NPI-0052), a novel marine derived proteasome inhibitor*. J Med Chem, 2005. **48**(11): p. 3684-7.
- [99] Bogyo, M., et al., *Substrate binding and sequence preference of the proteasome revealed by active-site-directed affinity probes*. Chem Biol, 1998. **5**(6): p. 307-20.
- [100] Marastoni, M., et al., *P3 and P4 position analysis of vinyl ester pseudopeptide proteasome inhibitors*. Bioorg Med Chem Lett, 2006. **16**(12): p. 3125-30.

- [101] Kuhn, D.J., et al., *Potent activity of carfilzomib, a novel, irreversible inhibitor of the ubiquitin-proteasome pathway, against preclinical models of multiple myeloma*. *Blood*, 2007. **110**(9): p. 3281-90.
- [102] Yang, J., et al., *Pharmacokinetics, pharmacodynamics, metabolism, distribution, and excretion of carfilzomib in rats*. *Drug Metab Dispos*, 2011. **39**(10): p. 1873-82.
- [103] Kingsbury, D.J., T.A. Griffin, and R.A. Colbert, *Novel propeptide function in 20 S proteasome assembly influences beta subunit composition*. *Journal of Biological Chemistry*, 2000. **275**(31): p. 24156-24162.
- [104] Gaczynska, M., et al., *Peptidase Activities of Proteasomes Are Differentially Regulated by the Major Histocompatibility Complex-Encoded Genes for Lmp2 and Lmp7*. *Proceedings of the National Academy of Sciences of the United States of America*, 1994. **91**(20): p. 9213-9217.
- [105] Romero, P., et al., *H-2Kd-restricted antigenic peptides share a simple binding motif*. *J Exp Med*, 1991. **174**(3): p. 603-12.
- [106] Ho, Y.K., et al., *LMP2-specific inhibitors: chemical genetic tools for proteasome biology*. *Chem Biol*, 2007. **14**(4): p. 419-30.
- [107] Diaz-Hernandez, M., et al., *Neuronal induction of the immunoproteasome in Huntington's disease*. *J Neurosci*, 2003. **23**(37): p. 11653-61.
- [108] Puttapparthi, K. and J.L. Elliott, *Non-neuronal induction of immunoproteasome subunits in an ALS model: possible mediation by cytokines*. *Exp Neurol*, 2005. **196**(2): p. 441-51.
- [109] Visekruna, A., et al., *Expression of catalytic proteasome subunits in the gut of patients with Crohn's disease*. *Int J Colorectal Dis*, 2009. **24**(10): p. 1133-9.
- [110] Cardozo, C. and R.A. Kohanski, *Altered properties of the branched chain amino acid-preferring activity contribute to increased cleavages after branched chain residues by the "immunoproteasome"*. *J Biol Chem*, 1998. **273**(27): p. 16764-70.

- [111] Orlowski, M., C. Cardozo, and C. Michaud, *Evidence for the presence of five distinct proteolytic components in the pituitary multicatalytic proteinase complex. Properties of two components cleaving bonds on the carboxyl side of branched chain and small neutral amino acids.* Biochemistry, 1993. **32**(6): p. 1563-72.
- [112] Blackburn, C., et al., *Characterization of a new series of non-covalent proteasome inhibitors with exquisite potency and selectivity for the 20S beta5-subunit.* Biochem J, 2010. **430**(3): p. 461-76.
- [113] Schuttelkopf, A.W. and D.M. van Aalten, *PRODRG: a tool for high-throughput crystallography of protein-ligand complexes.* Acta Crystallogr D Biol Crystallogr, 2004. **60**(Pt 8): p. 1355-63.
- [114] Pickart, C.M., *Back to the future with ubiquitin.* Cell, 2004. **116**(2): p. 181-90.
- [115] Oerlemans, R., et al., *Molecular basis of bortezomib resistance: proteasome subunit beta5 (PSMB5) gene mutation and overexpression of PSMB5 protein.* Blood, 2008. **112**(6): p. 2489-99.
- [116] Lahana, R., *Who wants to be irrational?* Drug Discov Today, 2003. **8**(15): p. 655-6.
- [117] Garnier, J.P., *Rebuilding the R&D engine in big pharma.* Harv Bus Rev, 2008. **86**(5): p. 68-70, 72-6, 128.
- [118] Fox, S., et al., *High-throughput screening: Update on practices and success.* Journal of Biomolecular Screening, 2006. **11**(7): p. 864-869.
- [119] Basse, N., et al., *Novel organic proteasome inhibitors identified by virtual and in vitro screening.* J Med Chem, 2010. **53**(1): p. 509-13.
- [120] Rickardson, L., et al., *Image-based screening for the identification of novel proteasome inhibitors.* J Biomol Screen, 2007. **12**(2): p. 203-10.

- [121] Lavelin, I., et al., *Discovery of novel proteasome inhibitors using a high-content cell-based screening system*. PLoS One, 2009. **4**(12): p. e8503.
- [122] Kessler, B.M., et al., *Extended peptide-based inhibitors efficiently target the proteasome and reveal overlapping specificities of the catalytic beta-subunits*. Chem Biol, 2001. **8**(9): p. 913-29.
- [123] Shibata, M.A., et al., *alpha-Mangostin extracted from the pericarp of the mangosteen (Garcinia mangostana Linn) reduces tumor growth and lymph node metastasis in an immunocompetent xenograft model of metastatic mammary cancer carrying a p53 mutation*. BMC Med, 2011. **9**: p. 69.
- [124] Matsumoto, K., et al., *Xanthenes induce cell-cycle arrest and apoptosis in human colon cancer DLD-1 cells*. Bioorg Med Chem, 2005. **13**(21): p. 6064-9.
- [125] Moongkarndi, P., et al., *Antiproliferation, antioxidation and induction of apoptosis by Garcinia mangostana (mangosteen) on SKBR3 human breast cancer cell line*. J Ethnopharmacol, 2004. **90**(1): p. 161-6.
- [126] Nakagawa, Y., et al., *Characterized mechanism of alpha-mangostin-induced cell death: caspase-independent apoptosis with release of endonuclease-G from mitochondria and increased miR-143 expression in human colorectal cancer DLD-1 cells*. Bioorg Med Chem, 2007. **15**(16): p. 5620-8.
- [127] Matsumoto, K., et al., *Preferential target is mitochondria in alpha-mangostin-induced apoptosis in human leukemia HL60 cells*. Bioorg Med Chem, 2004. **12**(22): p. 5799-806.
- [128] Lee, Y.B., et al., *alpha-Mangostin, a novel dietary xanthone, suppresses TPA-mediated MMP-2 and MMP-9 expressions through the ERK signaling pathway in MCF-7 human breast adenocarcinoma cells*. J Food Sci, 2010. **75**(1): p. H13-23.
- [129] Basse, N., et al., *Novel Organic Proteasome Inhibitors Identified by Virtual and in Vitro Screening*. Journal of Medicinal Chemistry, 2010. **53**(1): p. 509-513.

- [130] Scalbert, A. and G. Williamson, *Dietary intake and bioavailability of polyphenols*. J Nutr, 2000. **130**(8S Suppl): p. 2073S-85S.
- [131] Asensi, M., et al., *Natural polyphenols in cancer therapy*. Crit Rev Clin Lab Sci, 2011. **48**(5-6): p. 197-216.
- [132] Klappan, A.K., et al., *Proteasome inhibition by quercetin triggers macroautophagy and blocks mTOR activity*. Histochem Cell Biol, 2012. **137**(1): p. 25-36.
- [133] Shanafelt, T.D., et al., *Phase I trial of daily oral Polyphenon E in patients with asymptomatic Rai stage 0 to II chronic lymphocytic leukemia*. J Clin Oncol, 2009. **27**(23): p. 3808-14.
- [134] Larsen, C.A., R.H. Dashwood, and W.H. Bisson, *Tea catechins as inhibitors of receptor tyrosine kinases: mechanistic insights and human relevance*. Pharmacol Res, 2010. **62**(6): p. 457-64.
- [135] Scalbert, A., et al., *Absorption and metabolism of polyphenols in the gut and impact on health*. Biomed Pharmacother, 2002. **56**(6): p. 276-82.
- [136] Landis-Piwowar, K.R., et al., *A novel prodrug of the green tea polyphenol (-)-epigallocatechin-3-gallate as a potential anticancer agent*. Cancer Res, 2007. **67**(9): p. 4303-10.
- [137] Stein, R.L., F. Melandri, and L. Dick, *Kinetic characterization of the chymotryptic activity of the 20S proteasome*. Biochemistry, 1996. **35**(13): p. 3899-908.
- [138] Boran, A.D. and R. Iyengar, *Systems approaches to polypharmacology and drug discovery*. Curr Opin Drug Discov Devel, 2010. **13**(3): p. 297-309.
- [139] Groll, M., et al., *Inhibitor-binding mode of homobelactosin C to proteasomes: new insights into class I MHC ligand generation*. Proc Natl Acad Sci U S A, 2006. **103**(12): p. 4576-9.

- [140] Lupu, R. and J.A. Menendez, *Pharmacological inhibitors of Fatty Acid Synthase (FASN)--catalyzed endogenous fatty acid biogenesis: a new family of anti-cancer agents?* *Curr Pharm Biotechnol*, 2006. **7**(6): p. 483-93.
- [141] Yoshii, Y., et al., *Fatty acid synthase is a key target in multiple essential tumor functions of prostate cancer: uptake of radiolabeled acetate as a predictor of the targeted therapy outcome.* *PLoS One*, 2013. **8**(5): p. e64570.
- [142] Abdel-Magid, A.F., *Fatty Acid synthase inhibitors as possible treatment for cancer: patent highlight.* *ACS Med Chem Lett*, 2012. **3**(8): p. 612-3.
- [143] Kridel, S.J., et al., *Orlistat is a novel inhibitor of fatty acid synthase with antitumor activity.* *Cancer Res*, 2004. **64**(6): p. 2070-5.
- [144] Pemble, C.W.t., et al., *Crystal structure of the thioesterase domain of human fatty acid synthase inhibited by Orlistat.* *Nat Struct Mol Biol*, 2007. **14**(8): p. 704-9.
- [145] Hasegawa, M., et al., *Affinity labeling of the proteasome by a belactosin A derived inhibitor.* *Bioorg Med Chem Lett*, 2008. **18**(20): p. 5668-71.
- [146] Reddy, A.S. and S. Zhang, *Polypharmacology: drug discovery for the future.* *Expert Rev Clin Pharmacol*, 2013. **6**(1): p. 41-7.
- [147] Kim, H.P., et al., *Lapatinib, a dual EGFR and HER2 tyrosine kinase inhibitor, downregulates thymidylate synthase by inhibiting the nuclear translocation of EGFR and HER2.* *PLoS One*, 2009. **4**(6): p. e5933.
- [148] Rusnak, D.W., et al., *The effects of the novel, reversible epidermal growth factor receptor/ErbB-2 tyrosine kinase inhibitor, GW2016, on the growth of human normal and tumor-derived cell lines in vitro and in vivo.* *Mol Cancer Ther*, 2001. **1**(2): p. 85-94.
- [149] Mertz, J.A., et al., *Targeting MYC dependence in cancer by inhibiting BET bromodomains.* *Proc Natl Acad Sci U S A*, 2011. **108**(40): p. 16669-74.

- [150] Smith, C.C., et al., *Validation of ITD mutations in FLT3 as a therapeutic target in human acute myeloid leukaemia*. Nature, 2012. **485**(7397): p. 260-3.
- [151] Delmore, J.E., et al., *BET bromodomain inhibition as a therapeutic strategy to target c-Myc*. Cell, 2011. **146**(6): p. 904-17.
- [152] Nicodeme, E., et al., *Suppression of inflammation by a synthetic histone mimic*. Nature, 2010. **468**(7327): p. 1119-23.
- [153] Kyttaris, V.C., *Kinase inhibitors: a new class of antirheumatic drugs*. Drug Des Devel Ther, 2012. **6**: p. 245-50.