Activity-based proteasome profiling

Li, N.

Activity-based proteasome profiling

PROEFSCHRIFT

ter verkrijging van
de graad van Doctor aan de Universiteit van Leiden,
op gezag van Rector Magnificus prof. Mr. C.J.J.M.Stolker,
volgens besluit van het College voor Promoties
te verdedigen op maandag 16 december 2013
klokke 13:45 uur

door

Nan Li

geboren te Fushun, China in 1984

Promotiecommissie

Promotor:	Prof. Dr. H. S. Overkleeft
Co-promotor:	Dr. B. I. Florea
Overige leden:	Prof. Dr. J. Brouwer
	Prof. Dr. C. Driessen
	Prof. Dr. H. Ovaa
	Dr. M. van der Stelt

The front cover shows the structure of yeast 20S proteasome in complex with bortezomib (not shown)-top view (PDB 2f16).

山重水复疑无路，柳暗花明又一村。

Every cloud has a silver lining.

Table of contents

General Introduction

1.1 The ubiquitin proteasome system and the activity of the proteasome

The ubiquitin proteasome system (UPS) is the central protein turnover machinery in eukaryotic cells(*1*). Proteins that are at the end of their lifecycle are labeled by a small protein post translational modification (PTM), about 8 KDa, called ubiquitin. Protein substrates are modified on their lysine residues by ubiquitin through its C terminus, under the catalysis of ubiquitin activating enzyme (E1), ubiquitin conjugating enzyme (E2), and ubiquitin ligase (E3). Following addition of a single ubiquitin moiety to a protein substrate, further ubiquitin molecules can be added to the first, yielding a polyubiquitin chain(*2*). The polyubiquitin chain is recognized as the signal for proteasomal degradation of the substrate, also known as the "kiss of death". The ubiquitin molecules are removed from the protein substrates by deubiquitinating enzymes and recycled for ubiquitinylation of new substrates (Fig 1).

Polyubiquitinylated proteins can be transferred by ubiquitin adapter/receptor proteins to the proteasome for controlled degradation(*2, 3*). Ubiquitin receptor proteins recognize polyubiquitinylated proteins by their UBA (ubiquitin associate) domains, and transport them to the proteasome through the binding by UBL (ubiquitin like) domains to subunits of the 19S proteasome. Ubiquitin receptor proteins Rad23 and Dsk2 do not get degraded by the proteasome due to their special sequences and have long lifetime(*4*).

The proteasome is a multi-subunit protease cluster, which is evolutionarily conserved. In mammalian cells the proteolytically active forms of proteasome are the 26S or 30S proteasomes(*1, 5*). Structurally, the 26S proteasome contains one barrel like shaped 20S core particle and one 19S regulatory particle on one side of the 20S barrel, while 30S contains two 19S particles on both sides of the 20S. The 19S particles recognize polyubiquitinylated proteins, remove the polyubiquitin chains from the substrates, unfold the substrates and open the gate of the 20S particle(*6*).

Linearized polypeptides are degraded by the 20S core particles into peptides of about 4-25 amino acids(*7*). These peptides can be further degraded by downstream amino-

peptidases into single amino acids. The 20S particle has a barrel like shape, involving four hetero heptameric subunit rings, two identical β rings in the middle and two α rings on the outside. Among seven different α subunits (α1-α7) and seven different β subunits (β1-β7), 3 subunits β1, β2, β5 have proteolytic activity. The three active subunits have distinguished cleavage preferences. Generally, β1 cleaves at the C terminus of acidic amino acids; β2 cleaves after basic amino acids; and β5 cleaves after hydrophobic amino acids(*5*, *8*). This kind of difference makes the proteasome capable of cutting most of the poly-peptide sequences. The difference mostly results from their different sequences and shapes of the binding pockets. The catalytic amino acid residue in the active proteasome subunits is the N-terminal threonine.

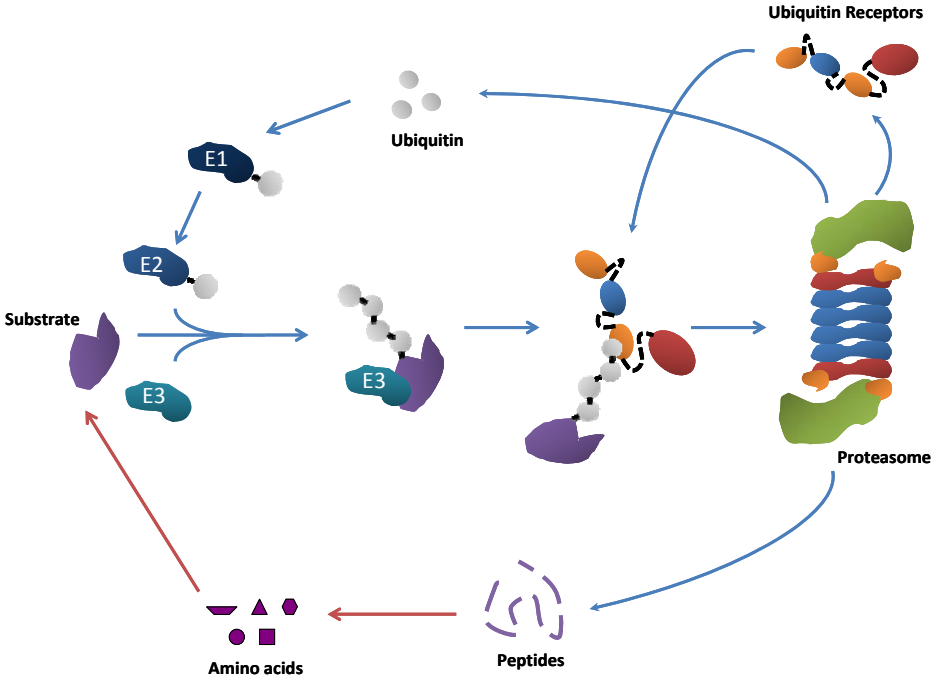


Figure 1: General scheme of the Ubiquitin-Proteasome system. Proteins (substrates) are modified by ubiquitin under catalysis of E1, E2 and E3 ubiquitin ligases and then transported by ubiquitin receptors to the proteasome for controlled degradation. The ubiquitin molecules are recycled for labeling of new substrates. Peptides resulted from proteasome degradation are hydrolyzed into amino acids. The amino acids are used for synthesis of new proteins.

Vertebrates have developed immune system, in which proteasomes play an important role. Peptides resulting from proteasomal degradation can be processed and presented as antigenic peptides on the cell surface by MHC (major histocompatibility complex) I and verified by CD8+ T cells(*9*, *10*). Besides constitutive proteasome, immune cells express a second type of 20S proteasome, known as immuno proteasome. In the

immuno proteasome, the active subunits β1, β2, β5 are replaced for another three active subunits β1i, β2i, β5i, which show slightly different substrate preferences with their constitutive homologues(*11*).

In vertebrates, a third type of 20S particle was identified recently as thymo proteasome for its excluding existence in the cortical thymus epithelial cells (cTEC)(*12*). In the thymo proteasome, a different active subunit β5t replaces the β5i in the immuno proteasome to form the new particle. β5t-deficient mice showed less efficient positive selection of CD8+ T cells(*13*).

1.2 Activity-based proteasome profiling

The activity-based protein profiling (ABPP) is applied to determine proteasome activity specifically in complex biological systems, for instance a cell lysate or living cells(*14, 15*). Synthetic, covalent and irreversible enzyme inhibitors that are modified with reporter groups like fluorophores or biotin, named activity-based probes (ABP), are used to label the active enzymes in the ABPP experiments. The probe-enzyme reaction ensures that only the catalytically active enzyme is targeted by the ABP.

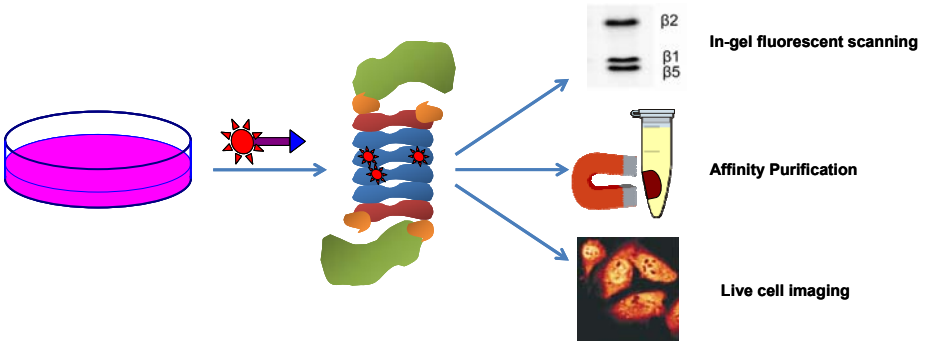


Figure 2: Schematic presentation of the application of several activity-based proteasome profiling approaches. Proteasomes are labeled by Activity-based probes in living cells and the fluorescent probe labeled proteasome active subunits are resolved by SDS-PAGE and visualized by in-gel fluorescent scanning. Affinity purification is used to enrich the active proteasome subunits that are labeled by biotinylated ABP. Fluorescent proteasome ABPs can also be used for live cell imaging to monitor the proteasome localization in the living cells by fluorescent microscopy.

A typical proteasome ABP involves three parts, the warhead, the linker and the tag(*14, 16*). The warhead is normally an electrophilic trap which can react with the active N-terminal threonine of the active proteasome subunits, for instance the epoxyketone or vinyl sulphone groups(*17*). The linker is generally a short peptide which can be recognized by the proteasome subunits as their endogenous substrates. The tag can be either a fluorophore or a biotin group, for visualization by in gel fluorescent scanning or biotin-streptavidin

immuno blot(*18*). The fluorescent group can also be used for live cell imaging by fluorescent microscopy, or staining of the active proteasome in FACS (fluorescence-activated cell sorting) experiments(*19, 20*). Biotin-streptavidin affinity purification is used to enrich the active proteasome subunits for protein identification or quantification (Fig 2). In some cases, the specificity and cell permeability of the probes can be altered by chemical groups like fluorophores or biotin. To circumvent these effects, an azide or alkyne or tetrazin group can be installed on the probes instead of the tag. After the first step labeling of the proteasome with the probes, the tags can be installed onto the probes by bioorthogonal chemistry(*21, 22*).

Active proteasome subunits were identified as the binding target of Streptomyces metabolite lactacystin using its radioactive ^{3}H-labeled derivatives, in murine neuroblast cells(*23*). The anti-tumor natural product epoxomicin was shown to be a proteasome inhibitor by labeling active proteasome subunits with biotin-epoxomicin(*24*).Since that time, several generations of activity-based proteasome probes have been developed, broad spectrum and subunit specific(*17*). For instance, ^{125}I-labeled peptide vinyl sulfone L_3VS (^{125}I-NIP-L_3VS) and the N-terminal extended biotinylated AdaK(bio)Ahx$_3$L$_3$VS derivative were applied to reveal the proteasome activity in either bacteria or EL4 murine T cells(*25, 26*). DansylAhx$_3$L$_3$VS was synthesized for in vivo profiling of the specificity of proteasome inhibitor bortezomib via dansyl fluorescence or western blotting detection with antibodies against dansyl(*27*). The pan-reactive MV151 equipped with lighter bodipy fluorophore was used for revealing proteasome activity in bortezomib adapted myeloma cells(*28*).

1.3 Aim and outlines of this thesis

The work described in this thesis is mainly focusing on setting up and application of a quantitative activity-based proteasome profiling method.

Chapter 1 provides a general introduction on the ubiquitin proteasome system (UPS) and activity-based proteasome profiling.

Chapter 2 is a literature review of some new achievements in the activity-based protein profiling field in the recent years, focusing on application in biochemistry, molecular and cellular biology, medicinal chemistry, pathology, physiology and pharmacology research.

Chapter 3 is a protocol for performing quantitative activity-based proteasome profiling experiments. In the protocol, both high throughput fluorescent ABPP and biotinylated probe plus LC/MS approaches are described.

Chapter 4 is a brief technical report about bioorthogonal chemistry in ABPP. The commonly used secondary azide group is compared with a primary azide group in proteasome ABPs performing Cu(I) catalyzed azide-alkyne cycloaddition and Staudinger-Bertozzi reaction under native/denatured protein conditions.

Chapter 5 is focusing on the application of quantitative activity-based proteasome profiling in the prognosis of cancer therapeutics. A combination of ABPP and global proteomics is performed to elucidate the bortezomib sensitivity and resistance mechanisms in leukemia and solid tumor cells.

Chapter 6 describes the characterization of the newly discovered proteasome subunit β5t by ABPP and LC/MS proteomics. The subunit is proven to be catalytically active. A hydrophilic Thr residue on the P2 position of the proteasome inhibitor improves the inhibitory efficiency of β5t, which indicates it might prefer to cleave hydrophilic peptides.

Chapter 7 describes the identification of O-GlcNAcylation modifications on the ubiquitin receptor protein hHR23B and characterization of how the sugar moiety influences the conformation and functions of the protein.

References and notes

1. Jung, T., Catalgol, B., and Grune, T. (2009) The proteasomal system, *Mol Aspects Med 30*, 191-296.

2. Finley, D. (2009) Recognition and processing of ubiquitin-protein conjugates by the proteasome, *Annu Rev Biochem 78*, 477-513.

3. Dantuma, N. P., Heinen, C., and Hoogstraten, D. (2009) The ubiquitin receptor Rad23: at the crossroads of nucleotide excision repair and proteasomal degradation, *DNA Repair (Amst) 8*, 449-460.

4. Heinen, C., Acs, K., Hoogstraten, D., and Dantuma, N. P. (2011) C-terminal UBA domains protect ubiquitin receptors by preventing initiation of protein degradation, *Nat Commun 2*, 191.

5. Marques, A. J., Palanimurugan, R., Matias, A. C., Ramos, P. C., and Dohmen, R. J. (2009) Catalytic mechanism and assembly of the proteasome, *Chem Rev 109*, 1509-1536.

6. Besche, H. C., Peth, A., and Goldberg, A. L. (2009) Getting to first base in proteasome assembly, *Cell 138*, 25-28.

7. Voges, D., Zwickl, P., and Baumeister, W. (1999) The 26S proteasome: A molecular machine designed for controlled proteolysis, *Annu Rev Biochem 68*, 1015-1068.

8. Borissenko, L., and Groll, M. (2007) 20S proteasome and its inhibitors: crystallographic knowledge for drug development, *Chem Rev 107*, 687-717.

9. Kloetzel, P. M., and Ossendorp, F. (2004) Proteasome and peptidase function in MHC-class-I-mediated antigen presentation, *Curr Opin Immunol 16*, 76-81.

10. Rock, K. L., and Goldberg, A. L. (1999) Degradation of cell proteins and the generation of MHC class I-presented peptides, *Annu Rev Immunol 17*, 739-779.

11. Huber, E. M., Basler, M., Schwab, R., Heinemeyer, W., Kirk, C. J., Groettrup, M., and Groll, M. (2012) Immuno- and Constitutive Proteasome Crystal Structures Reveal Differences in Substrate and Inhibitor Specificity, *Cell 148*, 727-738.

12. Murata, S., Sasaki, K., Kishimoto, T., Niwa, S., Hayashi, H., Takahama, Y., and Tanaka, K. (2007) Regulation of CD8+ T cell development by thymus-specific proteasomes, *Science 316*, 1349-1353.

13. Murata, S., Takahama, Y., and Tanaka, K. (2008) Thymoproteasome: probable role in generating positively selecting peptides, *Curr Opin Immunol 20*, 192-196.

14. Cravatt, B. F., Wright, A. T., and Kozarich, J. W. (2008) Activity-based protein profiling: from enzyme chemistry to proteomic chemistry, *Annu Rev Biochem 77*, 383-414.

15. Li, N., Overkleeft, H. S., and Florea, B. I. (2012) Activity-based protein profiling: an enabling technology in chemical biology research, *Curr Opin Chem Biol 16*, 227-233.

16. Evans, M. J., and Cravatt, B. F. (2006) Mechanism-based profiling of enzyme families, *Chem Rev 106*, 3279-3301.

17. Verdoes, M., Florea, B. I., van der Marel, G. A., and Overkleeft, H. S. (2009) Chemical Tools To Study the Proteasome, *Eur J Org Chem*, 3301-3313.

18. Florea, B. I., Verdoes, M., Li, N., van der Linden, W. A., Geurink, P. P., van den Elst, H., Hofmann, T., de Ru, A., van Veelen, P. A., Tanaka, K., Sasaki, K., Murata, S., den Dulk, H., Brouwer, J., Ossendorp, F. A., Kisselev, A. F., and Overkleeft, H. S. (2010) Activity-based profiling reveals reactivity of the murine thymoproteasome-specific subunit beta5t, *Chem Biol 17*, 795-801.

19. Verdoes, M., Florea, B. I., Menendez-Benito, V., Maynard, C. J., Witte, M. D., van der Linden, W. A., van den Nieuwendijk, A. M., Hofmann, T., Berkers, C. R., van Leeuwen, F. W., Groothuis, T. A., Leeuwenburgh, M. A., Ovaa, H., Neefjes, J. J., Filippov, D. V., van der Marel, G. A., Dantuma, N. P., and Overkleeft, H. S. (2006) A fluorescent broad-spectrum proteasome inhibitor for labeling proteasomes in vitro and in vivo, *Chem Biol 13*, 1217-1226.

20. Chang, J. T., Ciocca, M. L., Kinjyo, I., Palanivel, V. R., McClurkin, C. E., De Jong, C. S., Mooney, E. C., Kim, J. S., Steinel, N. C., Oliaro, J., Yin, C. C., Florea, B. I., Overkleeft, H. S., Berg, L. J., Russell, S. M., Koretzky, G. A., Jordan, M. S., and Reiner, S. L. (2011) Asymmetric Proteasome Segregation as a Mechanism for Unequal Partitioning of the Transcription Factor T-bet during T Lymphocyte Division, *Immunity 34*, 492-504.

21. Willems, L. I., Li, N., Florea, B. I., Ruben, M., van der Marel, G. A., and Overkleeft, H. S. (2012) Triple bioorthogonal ligation strategy for simultaneous labeling of multiple enzymatic activities, *Angew Chem Int Ed Engl 51*, 4431-4434.

22. Willems, L. I., van der Linden, W. A., Li, N., Li, K. Y., Liu, N., Hoogendoorn, S., van der Marel, G. A., Florea, B. I., and Overkleeft, H. S. (2011) Bioorthogonal chemistry: applications in activity-based protein profiling, *Acc Chem Res 44*, 718-729.

23. Fenteany, G., Standaert, R. F., Lane, W. S., Choi, S., Corey, E. J., and Schreiber, S. L. (1995) Inhibition of proteasome activities and subunit-specific amino-terminal threonine modification by lactacystin, *Science 268*, 726-731.

24. Meng, L., Mohan, R., Kwok, B., Elofsson, M., Sin, N., and Crews, C. (1999) Epoxomicin, a potent and selective proteasome inhibitor, exhibits in vivo antiinflammatory activity, *Proc Natl Acad Sci U S A 96*, 10403-10411.

25. Bogyo, M., McMaster, J. S., Gaczynska, M., Tortorella, D., Goldberg, A. L., and Ploegh, H. (1997) Covalent modification of the active site threonine of proteasomal beta subunits and the Escherichia coli homolog HslV by a new class of inhibitors, *Proc Natl Acad Sci U S A 94*, 6629-6634.

26. Kessler, B. M., Tortorella, D., Altun, M., Kisselev, A. F., Fiebiger, E., Hekking, B. G., Ploegh, H. L., and Overkleeft, H. S. (2001) Extended peptide-based inhibitors efficiently target the proteasome and reveal overlapping specificities of the catalytic beta-subunits, *Chem Biol 8*, 913-929.

27. Berkers, C. R., Verdoes, M., Lichtman, E., Fiebiger, E., Kessler, B. M., Anderson, K. C., Ploegh, H. L., Ovaa, H., and Galardy, P. J. (2005) Activity probe for in vivo profiling of the specificity of proteasome inhibitor bortezomib, *Nat Methods 2*, 357-362.

28. Ruckrich, T., Kraus, M., Gogel, J., Beck, A., Ovaa, H., Verdoes, M., Overkleeft, H. S., Kalbacher, H., and Driessen, C. (2009) Characterization of the ubiquitin-proteasome system in bortezomib-adapted cells, *Leukemia 23*, 1098-1105.

2

Activity-based protein profiling:

an enabling technology in chemical biology research

Li N, Overkleeft HS, Florea BI, Curr Opin Chem Biol. 2012, 16, 227-33.

2.1 Introduction

Amongst the toolkit of functional proteomic techniques, ABPP has proved powerful and attractive for its remarkable ability to label and enrich variable enzymatic activities. Activity-based probes (ABPs) can be viewed as chemical antibodies to report on the expression of a protein, but at the same time as probes to detect the target active enzymes in a living system (see Figure 1 for a general representation of the workflow). ABPP has been used on protein extracts, on living cells, and sometimes even on animal models. In this review some highlights of ABPP in chemical biological research during the past two years will be discussed, with a specific focus on applications in the fields of biochemistry, molecular and cellular biology, medicinal chemistry, pathology, physiology and pharmacology. The structures of the activity-based probes discussed here are compiled in Figure 2 and cross-referenced in bold numbers in the text.

2.2 ABPP in biochemistry

Since the invention of the methodology, ABPP has proven highly useful for the identification and annotation of enzymatic activities and their role in important biochemical pathways. One research area that has benefited in particular from ABPP is that involving the ubiquitin-proteasome system. For instance, recently a thymus specific proteasome was identified in which the β5t subunit replaces β5i in the immunoproteasome to create a new particle called the thymoproteasome(*1*). Affinity purification of the active site fragment coupled to biotin-epoxomicin followed by LC/MS protein identification demonstrated that β5t is catalytically active in murine thymus, and an ensuing ABPP-based competition assay employing proteasome ABPs **1** and **2** pointed towards a preference of β5t for neutral, hydrophilic substrates, in contrast to the hydrophobic substrates preferred by β5/β5i(*2*). In a second example of ABPP application to the UPS system, C-terminally modified, HA-tagged ubiquitin derivatives were applied as activity-based probe (**3**) for the profiling of a whole range of deubiquitinating enzymes (DUBs) and ubiquitin ligases,

demonstrating their catalytic activity and delivering the research tools for probing their involvement in protein ubiquitilation events(3).

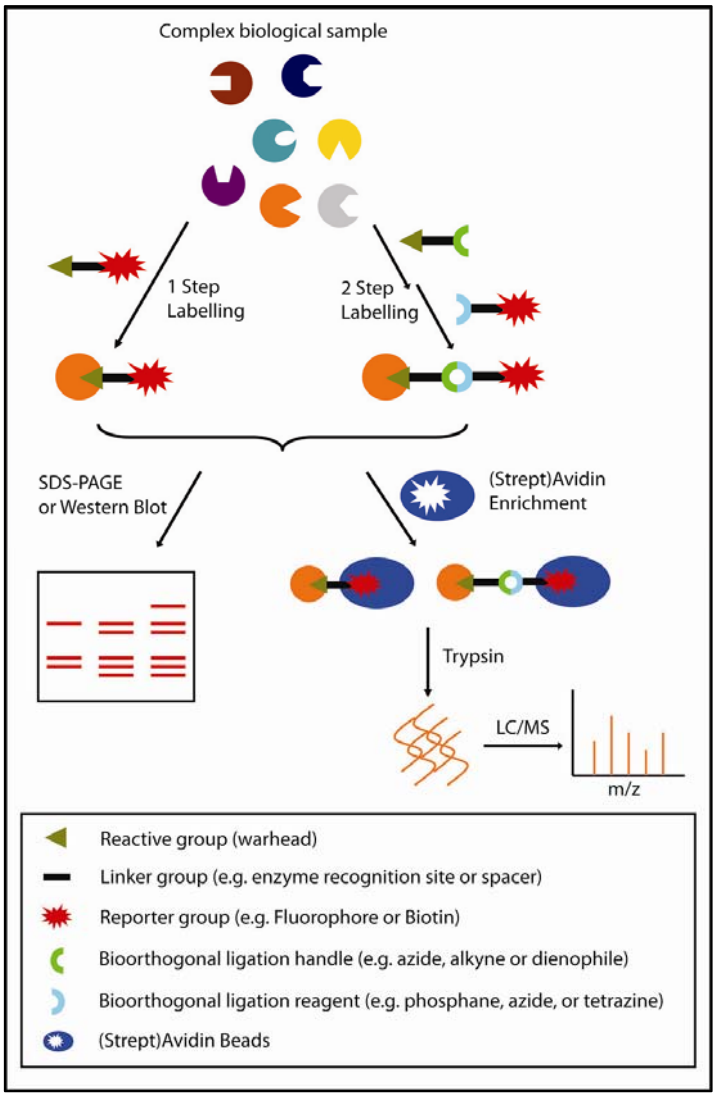


Figure 1: General scheme of activity-based protein profiling experiments. Complex proteomes are exposed to a type of activity-based probe either in vitro or in situ and the labeled protein targets are affinity-purified, separated, visualized and identified with a proteomics analysis system of choice: SDS-PAGE, western blotting or liquid chromatography hyphenated to mass spectrometry (LC-MS). A two-step labeling strategy is optional in case that the reporter tag obstructs the cell permeability of the ABP or the interaction between the ABP and target protein.

In multiple human diseases such as rheumatoid arthritis, cancer and colitis the enzymatic activity of protein arginine deiminase 4 (PAD_4), which catalyzes the hydrolysis of peptidyl-arginine to peptidyl-citrulline was found to be dysfunctional, possibly due to autodeimination. In order to test this hypothesis, a PAD_4 selective probe (**4**) was used to affinity purify PAD_4 from living cells together with several binding partners including histone H_3, the histone deacetylase HDAC1 and p53(4). It was shown that PAD_4 autodeimination does not alter its activity, substrate specificity, or calcium dependence. Autodeimination however modulates the ability of PAD_4 to interact with its previously identified binding partners(5). The same researchers investigated the regulation of PRMT1 protein arginine methyltransferase 1 (PRMT1). In the presence of estrogen, PRMT1 methylates Arg260 of the estrogen receptor, triggering the activation of protein kinase B (PKB/Akt) and promoting cell survival. A substrate-based ABP (**5**) showed that the PRMT1

activity is regulated, both temporally and spatially, in response to estrogen(*6*).

ABPP can be combined with LC/MS-based protein quantification platforms, allowing for accurate quantification of enzymatic activities. In a recent and highly innovative application, iodoacetamide (IA) alkyne (**6**) was used to determine the global reactivity profile of cysteine thiols across the entire human proteome(*7*). Substoichiometric amounts of the probe relative to the total number of cysteines were applied to modify the most reactive, and hence functionally involved (as for instance in enzyme active sites) cysteine thiols. In a comparative experiment extracts from the same biological source were saturated with **6** to modify all cysteines. The pool of hyper-reactive cysteines was click-ligated to a heavy N_3-stable isotopic-TEV tag and the comparative pool to the light form of the tag. Both pools were combined and ensuing enzymatic digestion, affinity purification and LC/MS analysis of peptides resulted in a global map of cysteine reactivity that is instrumental for the prediction and annotation of functional cysteines. In another example, Cravatt and co-workers combined stable isotopic labeling in cell culture (SILAC) with their Fluorophosphonate (FP) ABPP methodology to quantify the inhibition of serine hydrolyses(*8*). The ABP-enriched enzymes from samples with or without inhibitor (containing differentially labeled lysines and arginines) were mixed and analyzed by LC/MS. By comparing the intensity of the assigned peaks of the target enzyme, enzyme inhibition was quantified.

2.3 ABPP in molecular and cellular biology

Visualizing active proteins in living systems requires a robust signal to noise ratio involving fluorescent labels that emit at higher wavelengths compared to the biologic auto-fluorescence background of around 480-500 nm. Cysteine dependent cathepsins function mainly in the endo-lysosomal compartments catalyzing the hydrolysis of intra- and extra-cellular proteins but are also associated with tumor formation, growth, invasiveness and metastasis. Cathepsins B and L in particular are highly expressed in various tumors and are thus promising targets for tumor diagnosis and monitoring of therapy. The Bogyo lab developed a fluorescently quenched cathepsin probe (**7**) for the noninvasive optical imaging of subcutaneously grafted tumors in mice(*9*). Upon intravenous (i.v.) administration, the quencher is cleaved by active cathepsins and a sharp increase in fluorescence signal of the near infrared fluorophore Cy5 is detected in and around the tumors. A second example is the development of a potent and selective ABP (**8**) bearing a near infrared fluorophore for in vivo imaging of legumain(*10*). Legumain is a lysosomal protease involved in antigen processing and matrix degradation, but is also upregulated during tumorigenesis. The ABP enabled monitoring legumain activity in normal tissues, in solid tumors by high contrast shortly after i.v administration and the tracking of whole body distribution of the probe as well as the level of active legumain in organs by ex vivo imaging and SDS-PAGE.

In macrophages and dendritic cells, elevated cathepsin activities are required for antigen processing and presentation. In order to target and monitor cysteine cathepsins in professional antigen presenting cells, a mannose cluster was clicked to the Bodipy-TMR (tetramethylrhodamine) equipped cathepsin probe DCG04 (**9**). Probe **9** was taken up by specific receptor mediated transport via the mannose receptor and selectively labeled active cathepsins in cell culture(*11*).

Caspase activities are early mediators of apoptosis. Imaging and quantification of caspase activities holds promise for early diagnosis or disease monitoring, for instance after exposure to anti-tumor drugs in clinic. A cell permeable, caspase 3 and 7 sensitive ABP (**10**) equipped with a near-infrared fluorophore revealed dexamethasone-induced apoptosis in murine thymi and in tumor grafted mice treated with the apoptosis-inducing monoclonal antibody Apomab(*12*). Maximum fluorescence signal in live mice coincided with peak caspase activity assessed by SDS-PAGE analysis, emphasizing the potential of the probes for in vivo non-invasive optical imaging in preclinical and perhaps clinical settings.

Recently, quinone methide chemistry was utilized to develop a series of scaffolds in a highly modular fashion and with facile interchangeability of the moieties for making a series of caspase and phosphatase activity-based probes (**11**). Several active caspases were labeled both in vitro and in digitonin permeated cells and the quenched, two photon activatable fluorescent labels allowed sensitive detection of the target enzymes(*13*).

ABPs specific for the proteasome were instrumental in providing evidence for a remarkable discovery in the field of immune biology(*14*). Lymphocytes involved in an immune response undergo vigorous cell division to amplify their numbers and small fluctuations in the critical T-box transcription factor (T-bet) severely impacts on T-cell progeny where naïve CD8+ T-cells differentiate towards effector but not memory fate, while CD_4+ T-cells developed towards T helper 1 (Th_1) and less Th_2 or Th_{17} cells. Flow cytometry and fluorescence microscopy using proteasome ABPs **1**, **2** and **12** showed that T-bet levels are under proteasomal control and that during mitosis, proteasome activity is asymmetrically distributed between daughter cells as shown in Figure 3. Unequal proteasome distribution determines T-cell lineage fate and thus the direction of immune response development.

2.4 ABPP in medicinal chemistry

Both the FDA approved antiobesity drug orlistat, potentially an antitumor agent, and the nucleoside antibiotic showdomycin are covalently binding to their cellular targets(*15, 16*). For target identification, an alkyne group was introduced to minimally change the chemical structure and that enabled post-lysis bio-orthogonal ligation of reporter or affinity tags followed by either fluorescent imaging or LC/MS based protein identification (**13**, **14**). Orlistat showed 8 off-targets next to the known thioesterase domain

of fatty acid synthase (FAS). The antibiotic effect of showdomycin against Staphylococcus aureus in turn might involve inhibition of the essential enzymes MurA1 and MurA2 that are required for cell wall biosynthesis.

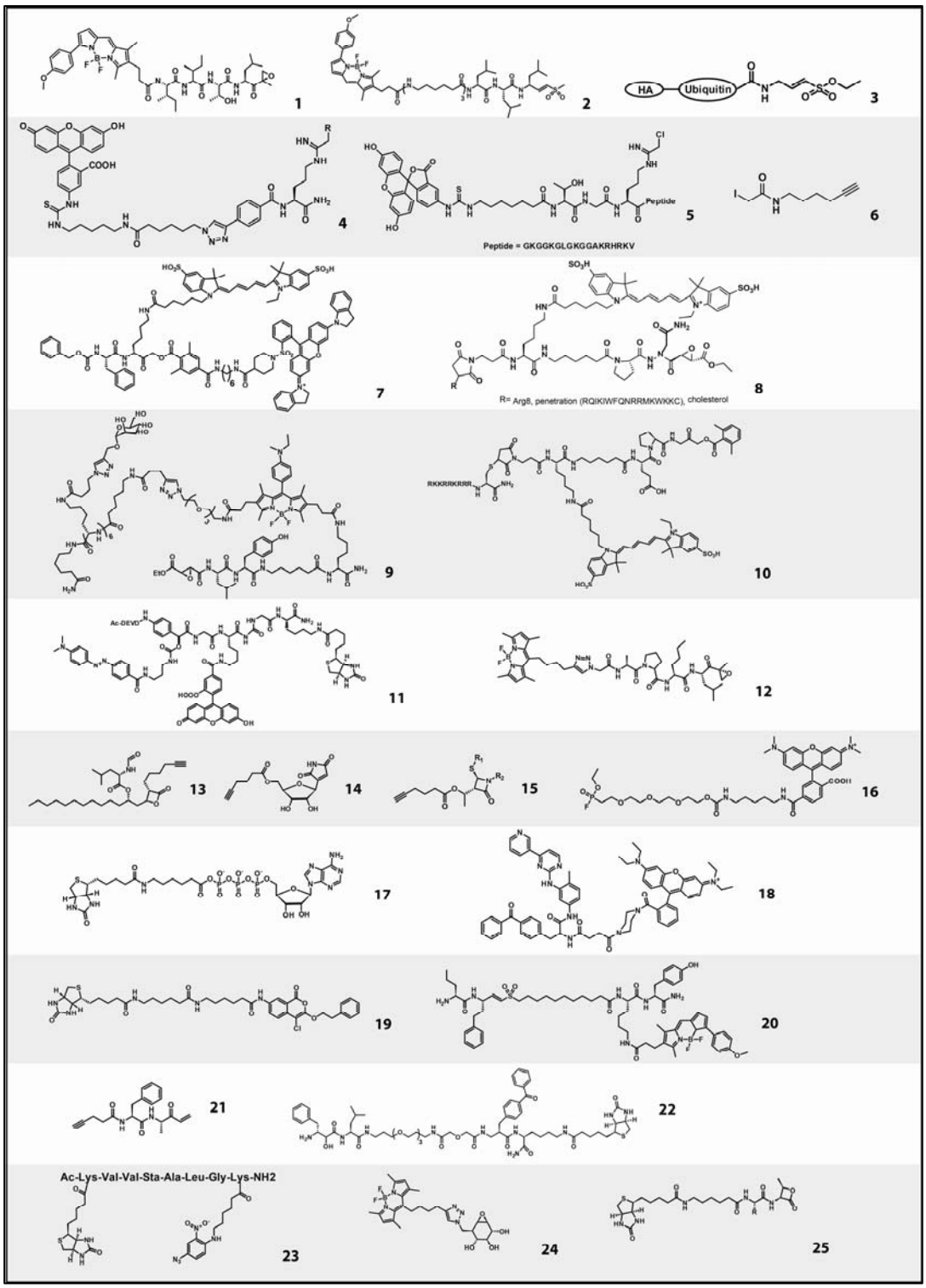


Figure 2: Chemical structure of Activity-Based Probes: 1 reference (2, 14); 2 reference (2, 14, 24); 3 reference (3); 4 reference (4,5); 5 reference (6); 6 reference (7); 7 reference (9); 8

reference (10); 9 reference (11); 10 reference (12); 11 reference (13); 12 reference (14); 13 reference (15); 14 reference (18); 15 reference (18); 16 reference (19, 23, 29, 31); 17 reference (20, 21); 18 reference (22); 19 reference (25); 20 reference (25); 21 reference (26); 22 reference (27); 23 reference (28); 24 reference (30); 25 reference (32).

The non-covalent antibiotic vancomycin was equipped by the Sieber lab with an alkyne group and a photo-activatable crosslinker(*17*). As expected, the vancomycin probe binds to the D-Ala-D-Ala motif of nascent peptidoglycan disrupting cell wall biosynthesis followed by autolysin (Atl) triggered cell rupture and death. Surprisingly, the probe also binds and inhibits the Atl amidase domain that causes massive defects in cell morphology and enhances the tolerance of S. aureus to low concentrations of vancomycin.

Not only the drug target proteins but also the resistance related proteins can be identified by ABPP. Small synthetic β-lactam probes (**15**) were applied to comparatively profile in situ enzyme activities of wild type and methicilin resistant S. aureus (MRSA) strains revealed unique MRSA features as known resistant associated targets, involved in cell wall biosynthesis and antibiotic sensing, but also uncharacterized enzymes capable of hydrolyzing β-lactam moieties(*18*). Such tools might prove their value for the identification of resistance genes and help to discover new drug targets for customized therapeutic interventions.

The design, synthesis and screening of compound libraries are important activities in medicinal chemistry and ABPP comes in useful when screening for enzyme inhibitors is the subject of study. A high-throughput screening protocol based on the FP-rhodamine probe (**16**) was developed to identify selective and potent inhibitors for two unrelated and poorly characterized serine hydrolases, namely retinoblastoma-binding protein-9 (RBBP9) and thioltransferase glutathione S-transferase omega 1 (GSTO1) both suggested to be cancer related genes(*19*). The enzymes were incubated with a compound library in a 384 well format, followed by FP-rhodamine labeling of residual enzyme activity. The bioactive alkaloid emetine was identified as a selective inhibitor of RBBP9. GSTO1 was found to be a target of several electrophilic compounds (including omeprazole and rifampicin) present in public libraries.

The KiNativ high-throughput screening platform from ActivX employs an ATP-analogue probe (**17**) for capturing ATP processing enzymes and profiling several well studied kinase inhibitors against >200 kinases in native cell proteomes to reveal biological targets for some of the inhibitors(*20, 21*). The authors found several striking differences between native and recombinant kinase inhibitory profiles, in particular, for the Raf kinases. This highlights the complexities of protein kinase behavior in the cellular context and demonstrates that profiling results based on recombinant/purified enzymes can be misleading. An Abelson (Abl) tyrosine kinase, the molecular target linked to the

development of chronic myelogenous leukemia (CML), specific ABP (**18**) was synthesized inspired by the clinically used Imatinib drug and equipped with a photo-activatable crosslinker(22). These examples indicate the power of ABPP for screening and development of new therapeutic strategies.

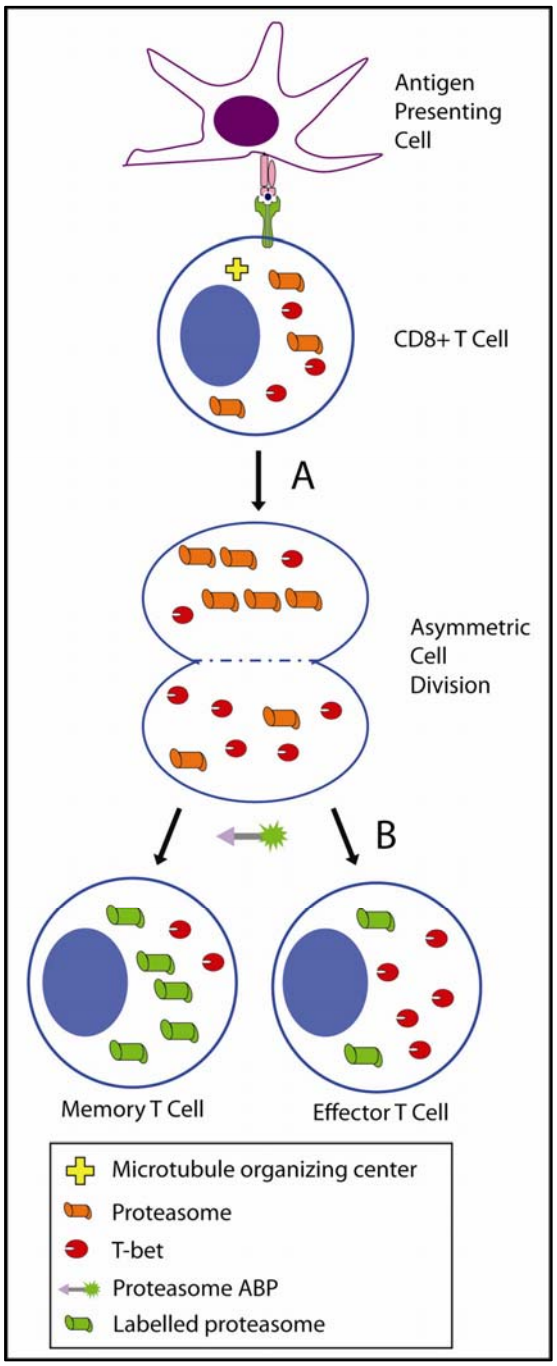


Figure 3: Asymmetric T-cell division revealed by activity-based protein profiling of the proteasome activity (14). Upon stimulation by professional antigen presenting cells, CD8+ and CD_4+ T-cells can undergo asymmetric cell division producing daughters of different types. ABPP assays showed that asymmetric proteasome segregation during mitosis dictates the T-bet transcription factor concentration that changes the fate of the daughter cells.

2.5 ABPP in pathology, physiology and pharmacology

Since ABPP is a powerful tool to observe the activity of a wide range of functional enzymes, it can be used as a sensitive method for diagnosis and prognosis of a series of diseases, which are related to some abnormal enzymatic activities, such as cancer,

pathogen infection, and metabolic disorders. Tumor cells display progressive changes in metabolism that correlate with malignancy, including development of a lipogenic phenotype. The Cravatt lab measured the activity of a series of monoacylglycerol lipases (MAGL) with ABPP (**16**) in both aggressive and nonaggressive human cancer cell lines(*23*). Elevated MAGL activity was found in the aggressive cancer lines and primary tumors where MAGL regulates a fatty acid network enriched in oncogenic signaling lipids that promotes migration, invasion, survival, and in vivo tumor growth. An example of ABPP in studying tumor pharmacology is the use of MV151 (**2**) for profiling the proteasome activity of bortezomib sensitive and resistant multiple myeloma (MM) cells(*24*). Multiple myeloma is an aggressive malignance of plasma B-cells, which can be treated with bortezomib, a proteasome inhibitor that blocks the β5 and β1 subunits. Driessen and co-workers found that elevated transcription rates, activities and polypeptide levels of β5, β1 and β2 subunits combined with increased expression and proteasome association of the 11S proteasome activator were the main pathways for resistant MM cells to cope with Bortezomib stress.

The ABPP approach was used in the Bogyo lab to identify essential proteases required for the proliferation of parasites. ABPs **19** and **20** were used to identify the malaria proteases pfSUB1 and DPAP3 as the key regulators of erythrocyte rupture(*25*). Recently, the same group identified the parasite protease TgDJ-1 in Toxoplasma gondii (**21**), which plays a key role in the pathogen attachment and invasion of host cells(*26*). Harbut and colleagues used bestatin analogues (**22**) to identify the malaria parasite aminopeptidases pfA-M1 and pf-LAP, which are necessary peptidases for hemoglobin digestion and parasite early life cycle(*27*). Various peptidase activities were profiled by ABPP (**23**) in tick GI-tract, which are often carriers of parasites. The study revealed the way hemoglobin from human blood was digested in tick GI-tract by a multipeptidase pathway(*28*). Hepatitis c virus (HCV) infection is a global harmful disease with unclear pathogenesis. Pezacki and co-workers used FP-Rhodamine (**16**) to visualize the differential host enzyme activation during the HCV replication and identified carboxylesterase 1 to play an important role in HCV propagation(*29*).

Gaucher disease, a common lysosomal storage disorder, is often underlined by the deficiency of glucocerebrosidase (GBA). Coupling the GBA inhibitor cyclophellitol to different Bodipy fluorescent groups provided ABPs (**24**) of ultra-high sensitivity and specificity for GBA(*30*). The probes were tested and validated both in vitro and in vivo and will find application in screens for new GBA inhibitors or chemical chaperones in living cells and for the diagnosis or therapy progression in Gaucher disease by quantifying the GBA activity in patient materials.

ABPP has also been used to study plant pathology and physiologic problems. FP probes (**16**) were used to visualize the differential activation of serine hydrolases in the unchallenged and botrytis-infected Arabidopsis thaliana(*31*). In another work, beta-lactone

probes (**25**) were used to identify a papain-like peptide ligase in the same organism(*32*).

2.6 Conclusion and outlook

Thanks to the discovery of new warheads, the design of improved enzyme targeting moieties and reporter/affinity tags, in conjunction with the remarkable increase in sensitivity, resolution and dynamic range of detection instruments, ABPP has become a powerful functional proteomics tool in the field of the biochemistry, (molecular) cell biology, medicinal chemistry, physiology, pathology and pharmacology. However, to date many enzyme families defy modification by ABPs and future research will reveal the extent by which ABPP can be applied to different enzyme families and indeed also non-enzymatic protein families.

References and notes

1. Murata, S., Sasaki, K., Kishimoto, T., Niwa, S., Hayashi, H., Takahama, Y., and Tanaka, K. (2007) Regulation of CD8+ T cell development by thymus-specific proteasomes, *Science 316*, 1349-1353.

2. Florea, B. I., Verdoes, M., Li, N., van der Linden, W. A., Geurink, P. P., van den Elst, H., Hofmann, T., de Ru, A., van Veelen, P. A., Tanaka, K., Sasaki, K., Murata, S., den Dulk, H., Brouwer, J., Ossendorp, F. A., Kisselev, A. F., and Overkleeft, H. S. (2010) Activity-Based Profiling Reveals Reactivity of the Murine Thymoproteasome-Specific Subunit beta 5t, *Chem Biol 17*, 795-801.

3. Love, K. R., Pandya, R. K., Spooner, E., and Ploegh, H. L. (2009) Ubiquitin C-Terminal Electrophiles Are Activity-Based Probes for Identification and Mechanistic Study of Ubiquitin Conjugating Machinery, *Acs Chem Biol 4*, 275-287.

4. Slack, J. L., Causey, C. P., Luo, Y., and Thompson, P. R. (2011) Development and Use of Clickable Activity Based Protein Profiling Agents for Protein Arginine Deiminase 4, *Acs Chem Biol 6*, 466-476.

5. Slack, J. L., Jones, L. E., Bhatia, M. M., and Thompson, P. R. (2011) Autodeimination of Protein Arginine Deiminase 4 Alters Protein-Protein Interactions but Not Activity, *Biochemistry-Us 50*, 3997-4010.

6. Obianyo, O., Causey, C. P., Jones, J. E., and Thompson, P. R. (2011) Activity-Based Protein Profiling of Protein Arginine Methyltransferase 1, *Acs Chem Biol 6*, 1127-1135.

7. Weerapana, E., Wang, C., Simon, G. M., Richter, F., Khare, S., Dillon, M. B. D., Bachovchin, D. A., Mowen, K., Baker, D., and Cravatt, B. F. (2010) Quantitative reactivity profiling predicts functional cysteines in proteomes, *Nature 468*, 790-U779.

8. Adibekian, A., Martin, B. R., Wang, C., Hsu, K. L., Bachovchin, D. A., Niessen, S., Hoover, H., and Cravatt, B. F. (2011) Click-generated triazole ureas as ultrapotent in vivo-active serine hydrolase inhibitors, *Nat Chem Biol 7*, 469-478.

9. Blum, G., von Degenfeld, G., Merchant, M. J., Blau, H. M., and Bogyo, M. (2007) Noninvasive optical imaging of cysteine protease activity using fluorescently quenched activity-based probes, *Nat Chem Biol 3*, 668-677.

10. Lee, J., and Bogyo, M. (2010) Development of Near-Infrared Fluorophore (NIRF)-Labeled Activity-Based Probes for in Vivo Imaging of Legumain, *Acs Chem Biol 5*, 233-243.

11. Hillaert, W., Verdoes, M., Florea, B. I., Saragliadis, N., Habets, K. L. L., Kuiper, J., Van Calenbergh, S., Ossendorp, F., van der Marel, G. A., Driessen, C., and Overkleeft, H. S. (2009) Receptor-Mediated Targeting of Cathepsins in Professional Antigen Presenting Cells, *Angew Chem Int Edit 48*, 1629-1632.

12. Edgington, L. E., Berger, A. B., Blum, G., Albrow, V. E., Paulick, M. G., Lineberry, N., and Bogyo, M. (2009) Noninvasive optical imaging of apoptosis by caspase-targeted activity-based probes, *Nat Med 15*, 967-U177.

13. Hu, M. Y., Li, L., Wu, H., Su, Y., Yang, P. Y., Uttamchandani, M., Xu, Q. H., and Yao, S. Q. (2011) Multicolor, One- and Two-Photon Imaging of Enzymatic Activities in Live Cells with Fluorescently Quenched Activity-Based Probes (qABPs), *J Am Chem Soc 133*, 12009-12020.

14. Chang, J. T., Ciocca, M. L., Kinjyo, I., Palanivel, V. R., McClurkin, C. E., De Jong, C. S., Mooney, E. C., Kim, J. S., Steinel, N. C., Oliaro, J., Yin, C. C., Florea, B. I., Overkleeft, H. S., Berg, L. J., Russell, S. M., Koretzky, G. A., Jordan, M. S., and Reiner, S. L. (2011) Asymmetric Proteasome Segregation as a Mechanism for Unequal Partitioning of the Transcription Factor T-bet during T Lymphocyte Division, *Immunity 34*, 492-504.

15. Yang, P. Y., Liu, K., Ngai, M. H., Lear, M. J., Wenk, M. R., and Yao, S. Q. (2010) Activity-Based Proteome Profiling of Potential Cellular Targets of Orlistat - An FDA-Approved Drug with Anti-Tumor Activities, *J Am Chem Soc 132*, 656-666.

16. Bottcher, T., and Sieber, S. A. (2010) Showdomycin as a Versatile Chemical Tool for the Detection of Pathogenesis-Associated Enzymes in Bacteria, *J Am Chem Soc 132*, 6964-6972.

17. Eirich, J., Orth, R., and Sieber, S. A. (2011) Unraveling the Protein Targets of Vancomycin in Living S. aureus and E. faecalis Cells, *J Am Chem Soc 133*, 12144-12153.

18. Staub, I., and Sieber, S. A. (2009) beta-Lactam Probes As Selective Chemical-Proteomic Tools for the Identification and Functional Characterization of Resistance Associated Enzymes in MRSA, *J Am Chem Soc 131*, 6271-6276.

19. Bachovchin, D. A., Brown, S. J., Rosen, H., and Cravatt, B. F. (2009) Identification of selective inhibitors of uncharacterized enzymes by high-throughput screening with fluorescent activity-based probes, *Nat Biotechnol 27*, 387-394.

20. Patricelli, M. P., Nomanbhoy, T. K., Wu, J. Y., Brown, H., Zhou, D., Zhang, J. M., Jagannathan, S., Aban, A., Okerberg, E., Herring, C., Nordin, B., Weissig, H., Yang, Q. K., Lee, J. D., Gray, N. S., and Kozarich, J. W. (2011) In Situ Kinase Profiling Reveals Functionally Relevant Properties of Native Kinases, *Chem Biol 18*, 699-710.

21. Patricelli, M. P., Szardenings, A. K., Liyanage, M., Nomanbhoy, T. K., Wu, M., Weissig, H., Aban, A., Chun, D., Tanner, S., and Kozarich, J. W. (2007) Functional interrogation of the kinome using nucleotide acyl phosphates, *Biochemistry-Us 46*, 350-358.

22. Kalesh, K. A., Sim, D. S. B., Wang, J. G., Liu, K., Lin, Q. S., and Yao, S. Q. (2010) Small molecule probes that target Abl kinase, *Chem Commun 46*, 1118-1120.

23. Nomura, D. K., Long, J. Z., Niessen, S., Hoover, H. S., Ng, S. W., and Cravatt, B. F. (2010) Monoacylglycerol Lipase Regulates a Fatty Acid Network that Promotes Cancer Pathogenesis, *Cell 140*, 49-61.

24. Ruckrich, T., Kraus, M., Gogel, J., Beck, A., Ovaa, H., Verdoes, M., Overkleeft, H. S., Kalbacher, H., and Driessen, C. (2009) Characterization of the ubiquitin-proteasome system in bortezomib-adapted cells, *Leukemia 23*, 1098-1105.

25. Arastu-Kapur, S., Ponder, E. L., Fonovic, U. P., Yeoh, S., Yuan, F., Fonovic, M., Grainger, M., Phillips, C. I., Powers, J. C., and Bogyo, M. (2008) Identification of proteases that regulate erythrocyte rupture by the malaria parasite Plasmodium falciparum, *Nat Chem Biol 4*, 203-213.

26. Hall, C. I., Reese, M. L., Weerapana, E., Child, M. A., Bowyer, P. W., Albrow, V. E., Haraldsen, J. D., Phillips, M. R., Sandoval, E. D., Ward, G. E., Cravatt, B. F., Boothroyd, J. C., and Bogyo, M. (2011) Chemical genetic screen identifies Toxoplasma DJ-1 as a regulator of parasite secretion, attachment, and invasion, *P Natl Acad Sci USA 108*, 10568-10573.

27. Harbut, M. B., Velmourougane, G., Dalal, S., Reiss, G., Whisstock, J. C., Onder, O., Brisson, D., McGowan, S., Klemba, M., and Greenbaum, D. C. (2011) Bestatin-based chemical biology strategy reveals distinct roles for malaria M1-and M17-family aminopeptidases, *P Natl Acad Sci USA 108*, E526-E534.

28. Horn, M., Nussbaumerova, M., Sanda, M., Kovarova, Z., Srba, J., Franta, Z., Sojka, D., Bogyo, M., Caffrey, C. R., Kopacek, P., and Mares, M. (2009) Hemoglobin Digestion in Blood-Feeding Ticks: Mapping a Multipeptidase Pathway by Functional Proteomics, *Chem Biol 16*, 1053-1063.

29. Blais, D. R., Lyn, R. K., Joyce, M. A., Rouleau, Y., Steenbergen, R., Barsby, N., Zhu, L. F., Pegoraro, A. F., Stolow, A., Tyrrell, D. L., and Pezacki, J. P. (2010) Activity-based Protein Profiling Identifies a Host Enzyme, Carboxylesterase 1, Which Is Differentially Active during Hepatitis C Virus Replication, *J Biol Chem 285*, 25602-25612.

30. Witte, M. D., Kallemeijn, W. W., Aten, J., Li, K. Y., Strijland, A., Donker-Koopman, W. E., van den Nieuwendijk, A. M. C. H., Bleijlevens, B., Kramer, G., Florea, B. I., Hooibrink, B., Hollak, C. E. M., Ottenhoff, R., Boot, R. G., van der Marel, G. A., Overkleeft, H. S., and Aerts, J. M. F. G. (2010) Ultrasensitive in situ visualization of active glucocerebrosidase molecules, *Nat Chem Biol 6*, 907-913.

31. Kaschani, F., Gu, C., Niessen, S., Hoover, H., Cravatt, B. F., and van der Hoorn, R. A. L. (2009) Diversity of Serine Hydrolase Activities of Unchallenged and Botrytis-infected Arabidopsis thaliana, *Mol Cell Proteomics 8*, 1082-1093.

32. Wang, Z., Gu, C., Colby, T., Shindo, T., Balamurugan, R., Waldmann, H., Kaiser, M., and van der Hoorn, R. A. L. (2008) beta-Lactone probes identify a papain-like peptide ligase in Arabidopsis thaliana, *Nat Chem Biol 4*, 557-563.

3

Relative quantification of proteasome activity by activity-based protein profiling and LC-MS/MS

Li N, Kuo CL, Paniagua G, van den Elst H, Verdoes M, Willems LI, van der Linden WA, Ruben M, van Genderen E, Gubbens J, van Wezel GP, Overkleeft HS and Florea BI, Nat Protoc . 2013, 8, 1155-1168.

3.1 Introduction

3.1.1 Focusing on proteasome activities

The proteasome is an evolutionarily conserved proteolytic complex that is responsible for the degradation of most proteins in eukaryotic cells ranging from yeast to human. It is essential for protein homeostasis and production of MHC class I restricted epitopes. Protein degradation is necessary for the turnover of damaged or misfolded proteins and regulation of biochemical pathways by lowering enzyme activity or messenger concentration. Obviously, the proteasome is a central protease in various cellular processes, including transcription, translation, DNA repair, cell division and antigen presentation(*1, 2*). In the last decade, the proteasome has become an attractive clinical target since the approval of the proteasome inhibitor bortezomib (Velcade®) by the US Food and Drug Administration (FDA) for the therapy of multiple myeloma(*3*). Encouraged by the clinical success of bortezomib, series of new generation proteasome inhibitors are being investigated as therapeutics of various diseases thus both (pre)clinical and fundamental knowledge on the activity of proteasomes is required.

The challenge of determining the proteasome activity by a robust and high-throughput method is significant. The proteasome is not a single protease, but a multi-subunit protease cluster which in eukaryotes contains active subunits with different cleavage preferences. Mammalian 30S proteasomes contain the catalytic 20S barrel-shaped core particle (CP) capped on both sides by 19S regulatory particles (RP) (Fig 1a). The 20S core particle consists of four heptameric rings assembled from α or β subunits (α1-7,β1-7,β1-7,α1-7) and harbors three different peptidase activities on each β ring(*4, 5*). Crystallographic and substrate specificity studies show that the active-site pockets of the β5 subunits can accommodate and cut at the C-terminus of bulky, hydrophobic amino acid residues in a manner that resembles chymotrypsin activity(*6*). β2 subunits prefer cleaving after basic residues and are referred to as bearing the "trypsin-like" activities, whereas β1

cuts after acidic residues and is known as the "caspase-like" activity.

This constitutive 20S proteasome core particle is present in all eukaryotic cells. In immune-competent tissues, three additional catalytically active β subunits are expressed: β1i (Low Molecular weight Protein 2, LMP2), β2i (multicatalytic endopeptidase complex-like-1, MECL1) and β5i (LMP7)2. The immuno β subunits show comparable substrate cleavage preference, share around 50% protein sequence identity and have different functional roles compared to the constitutive β1, β2 and β5 when replacing them in newly assembled 20S CP yielding the so-called immuno-proteasomes(*1*). Recently, the β5t subunit was identified that is exclusively expressed in cortical thymus epithelial cells where it is incorporated in immuno-proteasomes instead of β5i, yielding the thymo-proteasome(*7*). Despite this diversity, the activity of the β subunits is conveyed by the same mechanism of nucleophilic attack of the N-terminal threonine (Thr1) γ-hydroxyl on the peptide backbone(*8*).

A commonly used technique to determine the proteasome activity is by means of fluorogenic substrates; short peptides that optically quench the amino coumarin at their C-terminus which upon cleavage by the proteasome is released and starts to fluoresce in solution(*9*). Excellent subunit specific fluorogenic substrates for each of the three constitutive β subunit activities are commercially available, however, these can not discriminate between constitutive and immuno-proteasome activities when present in the same cellular system. This problem is encountered during immunological studies of antigen presentation or (pre)clinical research in immune cells, such as leukemia and myeloma cells.

3.1.2 Activity-based Protein profiling

ABPP is a chemical proteomics technology, which allows capturing, visualization, identification and quantification of the target enzymatic activities either in a test tube or in living systems(*10*). It requires no protein purification step or specific labeled substrates and enables simultaneous labeling of multiple activities providing a robust and high-throughput platform to observe the activity of target enzyme (family). In ABPP workflows, organic compounds are used, the so called activity-based probes (ABP), which are mechanism-based covalent and irreversible inhibitors of the target enzymes, equipped with a reporter group. A typical ABP consists of three parts, a reactive group (electrophilic trap or warhead) that targets the catalytic site of the enzyme, a spacer (linker) that serves as the recognition element for the enzyme binding pockets, and a reporter group (tag) for detection (Fig 1b).

Application of ABPP to proteasome activity assays embraces a rich history, from the early reports of natural products inspired tritium-labeled lactacystin and biotin labeled epoxomicin, to more rational design synthetic probes like the ^{125}I-labeled nitrophenol derivative of peptide vinyl sulfone L3VS (^{125}I-NIP-L3VS) and the commercially available N-terminal extended AdaAhx3L3VS inhibitor and its biotinylated AdaK(bio)Ahx3L3VS

derivative(*11-14*). Advances in fluorescence detection methods allowed the development of fluorescent probes to replace the radioactivity-based probes which are more technically challenging to handle. Berkers et al synthesized DansylAhx3L3VS that could be used for dual readout both via (weak) dansyl fluorescence or western blot analysis with antibodies against dansyl(*15*). The bright, pan-reactive MV151 and Bodipy-TMR-epoxomicin (MVB003, probe **3**) probes were followed by subunit specific probes for the β5 and β1 activities, and several generations of ABPs equipped with different warheads, linkers and tags have been developed in our lab(*16-18*). For low tech but high-throughput detection and quantification DansylAhx3L3VS was used to determine proteasome activity in hematologic malignancies, MV151 was instrumental to determine the proteasome activity in bortezomib adapted cells and MVB003 highlighted the β5t activity(*17, 19, 20*).

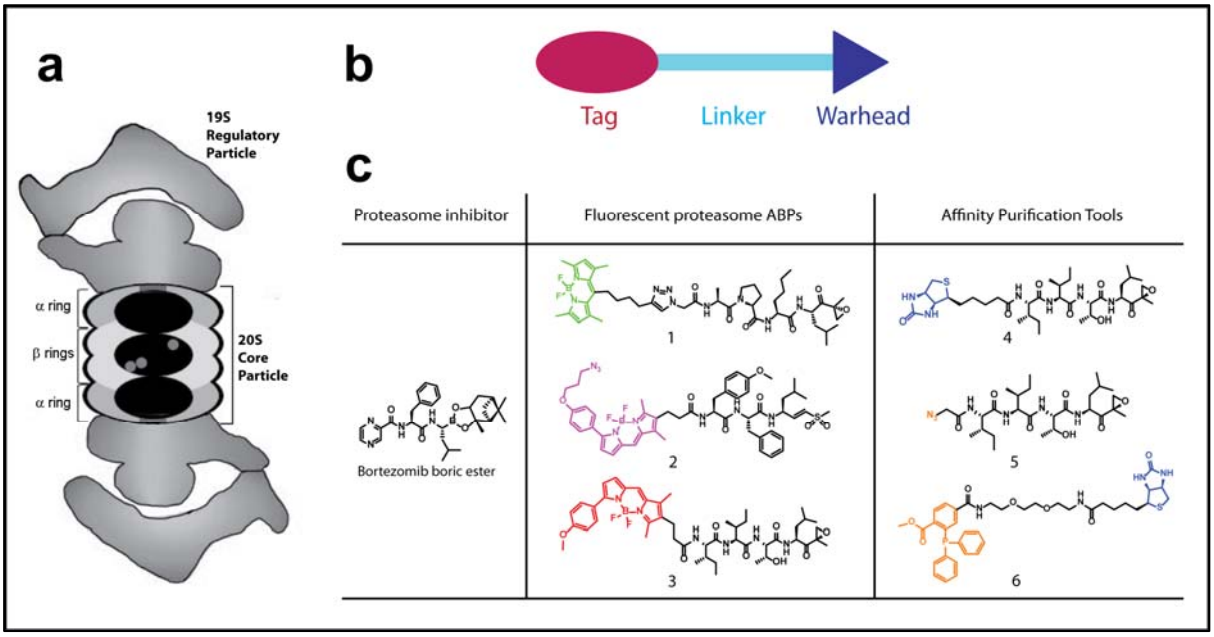


Figure 1:Proteasome ABPP. Model of the 30S proteasome where active β subunits are represented by grey spheres (a); schematics of a typical activity-based probe (b); Activity-based proteasome profiling tool box, LW124 (1) is a β1/β1i specific probe, MVB127 (2) is a β5/β5i specific probe, MVB003 (3) is a pan reactive probe, Biotin-Epoxomicin (4) is a pan reactive probe, N3-epoxomicin (5) is a two-step pan reactive probe, Biotin-Phosphine (6) is a bio-orthogonal ligation compound for installing a biotin tag (c). Synthesis of the compounds is described in refs 17-18, 21-23.

MVB003 contains the epoxomicin sequence (Ile-Ile-Thr) and electrophilic trap (Leu-epoxyketone), which makes the compound pan-reactive to all active proteasome subunits, it binds covalently, essentially labeling them with a fluorescent tag, and after SDS-PAGE separation and fluorescence scanning of the wet gel slab, a characteristic banding pattern appears (Figure 3, MCF7 cells) in which the β2 subunit runs higher than the doublet

consisting of β1 and β5. However, a general shortcoming of the pan-reactive probes is that in the case of immuno-proteasomes the molecular weights and thus the electrophoretic gel shift of the β1/β1i/β5/β5i subunits on a standard 8 cm tall 1D SDS-PAGE gel is limited thus their signals will overlap, show poor resolution and cannot be individually quantified (Figure 3, AMO1 cells). To overcome this problem, a rational design and screening program have been initiated to synthesize the β1/β1i selective sequence (Ala-Pro-dNorLeu) equipped with a Leu-epoxyketone warhead for probe **1** and the β5/β5i selective sequence (metTyr-Phe) with the Leu-vinylsuplhone electrophilic trap for probe **2**(*18*). The subunit specific probes **1** and **2** (Figure 1c) proved to be a remarkable technical improvement that could solve this problem, allowing for separation of β1i from β1, and β5i from β5 on SDS-PAGE with superior resolution, although the β subunits have a difference of some 500 Da in molecular weight. Furthermore, relative quantification of an individual activity between different samples can be done by fluorescent signal analysis in gel. The schematics of this experiment are shown in Figure 2 (lower route).

Alternatively, proteasome ABPP can be done with direct one-step biotinylated ABP (**4**) or two-step bioorthogonal tools (**5** and **6**)(*17, 21*). For one-step ABPP, proteasome active subunits in cell extracts are labeled with biotinylated probe **4** followed by affinity purification with streptavidin coated paramagnetic beads. Because biotin hampers probe permeability in living cells, the two-step approach is used where the azide equipped probe **5** labels the active proteasome in living cells and after cell lysis the reporter or affinity tag of choice is installed via a bio-orthogonal reaction(*22, 23*). After the pull down step, enriched proteins are identified by on-bead tryptic digest and LC/MS-based proteomics analysis (Figure 2, upper route). For relative quantification of proteasome activity, the stable isotope dimethyl labeling method has been combined with the affinity purification ABPP platform because it is a general labeling method that can be applied to peptide samples of various origin ranging from tissue cultures to biopsies taken from patients(*24, 25*). After on-bead digest, tryptic peptides were loaded on a C18 solid support on which both the di-methyl labeling and the desalting were performed, and after elution were mixed and analyzed by LC/MS. The open source MaxQuant software was then used for peptide identification and quantification(*26, 27*). In addition, active site peptides can also be analyzed by LC/MS after elution from the streptavidin beads.

Bortezomib shows much higher affinity for the β5/β5i and β1/β1i subunits because it is stabilized by hydrogen bonding in these active sites and the potency and the subunit specificity of bortezomib can be quantified accurately(*6*). This platform can find other applications for instance the screening for novel proteasome inhibitors. With this combination of methods and using the MVB003 probe it has been demonstrated that the newly discovered β5t subunit, expressed only in the thymus, is indeed catalytically active(*17*).

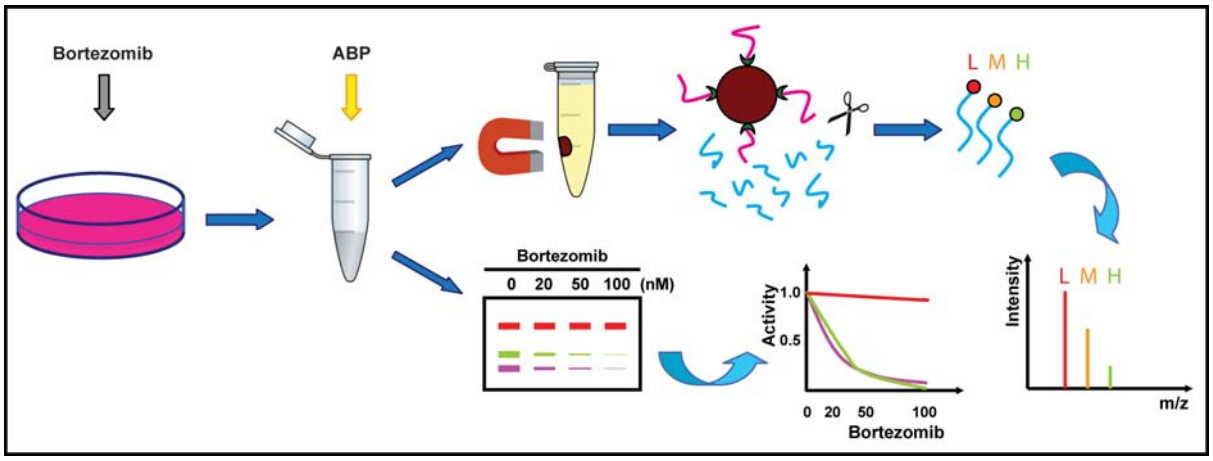


Figure 2: Schematic overview of the quantitative activity-based proteasome profiling experiments. Upper pathway describes the one or two step affinity purification protocol using the biotin/Streptavidin system, followed by stable isotope labeling, LC/MS-based proteomics analysis and quantification. Active site peptides (in pink) are analyzed after elution from the beads. Lower pathway shows the SDS-PAGE analysis and quantification of a competitive ABPP experiment.

3.1.3 Advantages and limitations with respect to other methods

Study of the individual function of active β-subunits by genetic knockdown in yeast proved to be difficult because they are essential for survival of the cell and for that reason chemical modulation and readout of activity is preferred(*28*). Compared to the popular proteasome activity determination by fluorogenic substrate assay, the ABPP method is high-throughput (no need for protein purification), robust and simple (no specific substrates needed and little off-target labeling), of superior resolution (capability to distinguish constitutive and immuno activities) and living cell compatible (good cell-permeability shown by ABPs).

There are some drawbacks due to the chemical and mechanism based nature of the method. It is possible that the chemicals used as probes can also hit other targets than the proteasome that might limit their use in living cells. Because the probes bind covalently to the catalytic sites, the enzymatic activity is consumed during the measurement necessitating the analysis of time dependency when enzyme kinetics needs to be assayed. For accurate kinetic measurements it is necessary to determine enzyme activity by taking increasing concentration of the probe and incubate for increasing amount of time. A rate constant (kobs) for each probe concentration is extracted from the semi-logarithmic plot of enzyme activity versus time and plotted against the probe concentration to determine the affinity of the probe for the enzyme (Ki value). For an excellent review on this methodology see Singh et al(*29*). However, these extensive studies are not necessary for global screens using a standard incubation time and probe concentration which should make the inter-

experiment results comparable. The ABPP assays presented here are not easily scalable to high throughput 96 well format and require relatively high (5-10x10^6) cell numbers. An ELISA based, high throughput method for the analysis and quantification of active site peptides of the proteasome β-subunits is described by Muchamuel and coworkers(*30*).

Despite the few drawbacks, the ABPP platform shows to be a powerful method that facilitates quantitative comparison between different treatments, e.g. increasing concentrations of bortezomib, allows analysis of all proteasome activities in one experiment in living cells, can be used for identification of the active site peptides and is applicable to clinical samples. Although it has been described here for the analysis of proteasome activity, this protocol can be used as guideline for setting up comparable ABPP workflows(*10*). Recently, this method has been adapted to determine the activity of the membrane bound cerebrosidase GBA1 and GBA2 enzymes that are expressed at very low abundance (1:300,000)(*31*). The fluorescent probes can also find application in fluorescence microscopy and flow cytometry analysis as exemplified by asymmetric T lymphocyte division caused by unequal proteasome segregation between daughter cells(*32*). For background information on the use of Click chemistry (TOP-ABPP) for performing two-step labeling and the use of the stable isotope in cell culture labeling (ABPP-SILAC) approach the work from the Cravatt lab can be read(*33, 34*). As future outlook, the proteasome can be an interesting target for rational drug design of novel antibiotics35 and the ABPP method might find application in this field. In this manuscript a step by step description of a competitive ABPP protocol is given, in which the β subunits activity of the proteasome is determined after treatment with bortezomib and in non-treated cells.

3.1.4 Experimental Design

High resolution fluorescent ABPP This method can be easily implemented in any laboratory that has basic cell culture and protein analysis equipment. The method has been used successfully in many different mammalian cell culture systems ranging from immortal cell lines to primary cells and from adhering to suspension cultures. Typical applications are for drug discovery programs of novel proteasome inhibitors, determination of the effect of a drug on the proteasome activity, antibiotic discovery or basic understanding of proteasome biology in plants(*35, 36*). In this protocol, human and mouse cell lines are incubated with increasing concentrations of bortezomib and the remaining activity of individual β-subunits of both the constitutive as the immuno proteasome in the respective cell lysates is determined. The advantage of post-lysis ABPP is that the same lysate can be interrogated with different probes and compared. Alternatively, because probes **1, 2** and **3** are cell permeable they can be added in parallel to cells after bortezomib treatment to study competition in the context of living cells. This assay can be performed in 6 or 12 well format but not higher because protein isolation from small cell numbers would prove too

difficult. The experiments should be performed at least in duplicate and incubation with 0 nM bortezomib serves as positive control for determination of the maximum probe binding signal. Incubation with a potent and pan-reactive inhibitor to all proteasome β-subunits (e.g. N3-epoxomicin, **5**) or boiling the cell lysate with 1% SDS prior to labeling with the activity-based probes can serve as negative controls. Commercially available protease inhibitor cocktails are tolerated during lysis. The protease inhibitor cocktail tablets (Roche) and PMSF (Sigma) have been tested and found out that they do not compete for the covalent binding between ABP and proteasome subunits.

For quantification purposes it is important that the cell lysis buffer is compatible with the protein determination method used after cell lysis as it can be sensitive to some detergents (SDS, Triton X100) or reducing agents (e.g. DTT) that can skew the measured protein concentration. SDS-PAGE analysis is best performed in 12.5% acrylamide gels followed by imaging the ABP labeled β-subunits simply by scanning on a fluorescence bed scanner. Active β-subunits of the constitutive proteasome show a typical banding pattern of β2 above the 25kDa marker and β1 and β5 subunits as a doublet below the 25 kDa. Commercial dual or multicolor protein markers contain a red band at 25 kDa that appears in the "red" channel used for probes **2** and **3** however up to 40x dilution of the marker might be necessary to prevent signal saturation. Mixed proteasomes (constitutive and immuno proteasome) show an additional β2i band between the β2 and β1 band and because the molecular weights of β1, β1i, β5, β5i subunits are in close proximity they are poorly separated on gel. High resolution separation can though be achieved by using the subunit specific probes **1** for β1β1i and **2** for β5β5i allowing for quantification of the individual band signals. Coomassie blue staining is a facile and reliable method to correct and normalize for pipetting errors and should be performed after the fluorescence scanning.

The outcome of a typical competitive ABPP experiment is that the signal of one or several β-subunit bands decreases with increasing competitor concentrations, which highlights the superior sensitivity and resolution of the ABPP method compared with fluorogenic assays because the detected signal provides both the identity as well as the intensity of the active β-subunit. By plotting the signal intensity against the concentration of the competitor compound, one can estimate the potency and selectivity of the competitor for the proteasome subunits.

When exploring organisms with poorly annotated genomes, ABPP can reveal the identity of the active proteasome genes, which is particularly interesting for profiling pathogenic organisms that cause disease in humans and animals. The ABPP method can be used with virtually all proteasome inhibitors given that they display an interaction strong enough to outcompete the covalent ABP from binding. The method is also adaptable to other organisms than mammals, it works in all types of cell lysates, however, probe optimization and identification of gel bands by affinity purification and LC/MS-based

proteomics is necessary for organisms that display cell walls or show extensive efflux pump activities like pathogens, yeasts or plants(*36*).

Quantitative LC/MS based ABPP Protein identification by LC/MS is very sensitive to background pollutants like keratins from dust, plasticizers from plastics or tubing and polymeric detergents (e.g. Triton, Tween). For this work it is advisable to create a dedicated lab space with a laminar flow cabinet, dedicated pipette sets, tubes, tips and very importantly dedicated solutions. Protect your work from dust and wear gloves that should be rinsed regularly to avoid electrostatic dust accumulation. Keep organic solvents in glass containers and use chemically pure water (18 MΩ) either freshly tapped from the de-ionizing apparatus or commercially available (Biosolve). Distrust autoclaved water because it is full of plasticizers or PEG's that pollute the LC/MS. The use of low-binding tips and tubes (e.g. Sarstedt, Eppendorf) is advisable to prevent loss of proteins/peptides due to adsorption to plastics.

In order to determine the proteasome activity in living cells or cell lysates by LC/MS based ABPP, it is necessary to first capture the active subunits by covalent binding to the activity-based probe, followed by affinity purification of the active subunits. Activity-based probes equipped with a biotin tag are the most popular because they facilitate the use of robust biotin-streptavidin affinity purification work flows. However, presence of biotin reduces the cell permeability of ABP's making it unsuitable for use in living cells. A solution to this problem is presented in this protocol where epoxomicin equipped with an N-terminal azide (N_3) ligation handle (compound **5**) that retains perfect cell permeability, was used to capture residual proteasome activity after bortezomib treatment in living cells, followed by post-lysis introduction of biotin via a bio-orthogonal reaction with biotin-phosphine (**6**) in the aqueous environment. Several other bio-orthogonal ligation procedures have been explored like the Cu+ catalyzed Huisgen 1,3-dipolar cycloaddition (Click chemistry) or the copper free approaches and empirically discovered that the azide-phosphine ligation pair under native (non-denaturing) conditions proved to be the most robust for the proteasome activity determination, because the phosphine reagent is simply added to the lysate, there is not need for a catalyst and we have experienced less protein precipitation(*34*, *37*).

High concentrations of biotin-phosphine (250 µM) and removal of this excess after the reaction by precipitating the proteins is necessary for the protocol. Protein precipitation with the chloroform/methanol method has several advantages compared with acetone precipitation or size exclusion cleanup: speed, it takes only 10-15 min, allows choice in partition of chemical impurities either to the aqueous or organic phase, adaptable for a wide range of protein concentrations from 10 µg to several mg and low losses of material(*38*). Drawback of the method is that protein pellets should be dried not longer than 5 minutes and the use of 1-4% SDS containing buffers is necessary for dissolving the

proteins. SDS concentration needs to be decreased before the pull down because it can interfere with and lower the biotin-streptavidin interaction. Several strategies can be adopted: dispersion of protein pellets in a low volume (25µl) of 2% SDS buffer followed by stepwise addition of buffer increasing the volume of solvation but slowly diluting out the SDS to keep its solvation power, as described in this protocol or dissolving the pellet in a higher volume (500 µl) of 2-4% SDS followed by dilution to 0.05-0.1% SDS in several tens of milliliter buffer, use of larger amount of beads (1mg or higher) and increased pull down time like overnight at 4°C as described in Kallemeijn et al(*31*).

Several sources and types of paramagnetic streptavidin coated beads have been tested and for this protocol mono-disperse, high binding capacity, 1 µm diameter beads showed superior pull-down efficiency. The advantage of paramagnetic beads compared to agarose beads is the ease of liquid handling because the beads stick to the tube wall when a magnetic field is applied. The high binding affinity of biotin for streptavidin facilitates stringent washing that decreases protein background but cannot discriminate for endogenously biotinylated proteins. Elution of biotinylated proteins is not trivial and two possibilities are presented: elution by boiling SDS-PAGE sample buffer containing SDS under denaturing conditions and in the presence of an excess of free biotin to displace the bound proteins from the streptavidin matrix, or on bead digest with trypsin for LC/MS proteomics analysis. Elution in sample buffer affords the possibility of SDS-PAGE analysis that yields valuable information about protein molecular weight and after in-gel tryptic digest, protein identity can be determined by LC/MS analysis(*39*). For on-bead tryptic digest an estimation of the trypsin quantity should be made as most protocols use a 1:100 ratio of trypsin:protein. Empirical evaluation showed that when starting with 1 mg of protein and 500 µg of beads, using some 500 ng trypsin yields reproducible results.

For comparison of treated versus non treated samples by relative quantification using LC/MS-based proteomics the "light" "medium" and "heavy" labels with stable isotope by reductive amination of primary amines at the N-termini and ε-position of lysine residues using formaldehyde have been chosen(*24*, *25*). This robust technique, applicable to virtually all types of protein specimens ranging from clinical samples to cell culture lysates, has been adapted and optimized in this protocol for small protein amounts that yield after affinity purification, by using the stage tip method(*40*). Finally, after on-bead tryptic digest, the active site peptides bound to the biotinylated probe can be eluted with acetonitrile/formic acid mixtures (containing free biotin) and used for unraveling the binding mechanism of the probe to the protein by MS/MS or MS/MS/MS analysis of the ion fragments.

3.2 Materials

3.2.1 Reagents

<CRITICAL> The chemical compounds from Fig 1 are available upon request, the primary

papers that describe their synthesis are given in the references(*17, 18, 21, 22*).

MCF7 (HTB-22) and EL4 (TIB-39) cells are commercially available from ATCC (USA). See Box 1 for information regarding cell culture and protein extraction.

AMO-1 (ACC-538) was a kind gift from the Driessen lab and it is commercially available from the Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures (DSMZ, Braunschweig, Germany). See Box 1 for information regarding cell culture and protein extraction(*19, 20*).

ACN (Acetonitrile, Sigma-Aldrich cat.no. 14261)

ATP (Adenosine triphosphate, Sigma-Aldrich cat.no. A2383)

NH_4HCO_3 (Ammonium Bicarbonate, Sigma-Aldrich cat.no. 09830)

APS (Ammonium Persulfate, Sigma-Aldrich cat.no. A3678)

β-mercaptoethanol (Sigma-Aldrich cat.no. M3148)

BSA (Bovine serum albumin, Sigma-Aldrich cat.no. A7906)

$CaCl_2$ (Calcium Chloride, Sigma-Aldrich, cat.no. 223506)

$CHCl_3$ (Chloroform, Sigma-Aldrich, cat.no. 650498)

Colloidal Blue Staining Kit (Invitrogen, cat.no. LC6025)

Digitonin (Sigma-Aldrich, cat.no. D141)

Na_2HPO_4 (Di-sodium hydrogen phosphate, Sigma-Aldrich, cat.no. S3264)

KH_2PO_4 (Potassium phosphate monobasic, Sigma-Aldrich, cat no. P5655-500G)

K_2HPO_4 (Potassium phosphate dibasic, Sigma-Aldrich, cat no. P2222-500G)

DTT (Dithiothreitol, Sigma-Aldrich, cat.no. 43815)

DMSO (Dimethyl Sulfoxide, Sigma-Aldrich, cat.no. D8418)

ECL Plus Western Blotting System (GE Healthcare, cat.no. RPN2132)

CH_2O (Light Formaldehyde, Sigma-Aldrich, cat.no. F8775)

CD_2O (Intermediate Formaldehyde, Sigma-Aldrich, cat.no. 492620)

$^{13}CD_2O$ (Heavy Formaldehyde, Sigma-Aldrich, cat.no. 596388)

! CAUTION work with formaldehyde in a fume hood, because the solutions and vapours are toxic.

FA (Formic acid, Sigma-Aldrich, cat.no. F0507)

Glycerol (Sigma-Aldrich, cat.no. G5516)

Glycine (Sigma-Aldrich, cat.no. G8898)

HCl (Hydrochloric acid, Sigma-Aldrich, cat.no. 84415)

IAA (Iodoacetamide, Sigma-Aldrich, cat.no. I1149)

$MgCl_2$ (Magnesium Chloride, Sigma-Aldrich, cat.no. M8266)

Tergitol NP-40 (Sigma-Aldrich, cat.no. NP40)

KCl (Potassium Chloride, Sigma-Aldrich, cat.no. P9541)

Trypsin (Promaga, cat.no. V5111)

SDS (Sodium dodecyl sulfate, Sigma-Aldrich, cat.no. L4390)

SilverQuest Staining Kit (Invitrogen, cat.no. LC6070)
Bradford assay reagent (BioRad, cat.no. 500-0205)
NaCl (Sodium Chloride, Sigma-Aldrich, cat.no. S3014)
$NaBH_3CN$ (Sodium cyanoborohydride, Sigma-Aldrich, cat.no. 156159)
$NaBD_3CN$ (Sodium cyanoborodeuteride, Sigma-Aldrich, cat.no. 190020)
NaH_2PO_4 (Sodium dihydrogen phosphate, Sigma-Aldrich, cat.no. S3139)
Sucrose (Sigma-Aldrich, cat.no. 84097)
TEMED (N,N,N′,N′-Tetramethylethylenediamine, Sigma-Aldrich, cat.no. T9281)
Tris (Sigma-Aldrich, cat.no. T6066)
Tween 20 (Sigma-Aldrich, cat.no. P9416)
Urea (Sigma-Aldrich, cat.no. U5378)
Biotin (Sigma-Aldrich, cat.no. B4639)
Dynabeads MyOne Streptavidin C1 (Invitrogen, cat.no. 650-01)
EDTA (Ethylenediaminetetraacetic acid, Sigma-Aldrich, cat.no. E6758)
Ethanol (VWR, cat.no. 20816.298)
Methanol (VWR, cat no. 83638.290)
Bortezomib ester (Fig. 1, S1013, Selleckchem.com, Houston, USA)
LW124 (**1**, Fig. 1)[18]
MVB127 (**2**, Fig. 1)[18]
MVB003 (**3**, Fig. 1)[17]
Biotin-Epoxomicin (**4**, Fig. 1)[17]
N_3-epoxomicin (**5**, Fig. 1)[17]
Biotin-Phosphine (**6**, Fig. 1)[21,22]
DMEM (Dulbecco's modified eagle's medium, PAA, cat.no.E15-883)
RPMI-1640 (PAA, E15-041)
FCS (Fetal calf serum, PAA, cat.no. A15-104)
Stable L-Glutamine (PAA, cat no. M11-006)
Penicillin (Duchefa Biochemie, cat no. P0142.0100)
Streptomicin (Duchefa Biochemie, cat no. S0148.0100)
BPB (Bromophenol Blue, Sigma-Aldrich, cat no. B0126-25G)
DC Marker (Precision Plus Protein standards, Bio-Rad cat.no. 161-0374)

3.2.2 Equipment

Cell culture

Galaxy 170R CO_2 Incubater (New Brunswick)
Clean air cabinet (Thermo)
37°C Water bath (Grant)
Fridge Freezer (Bosch)

Ependorf centrifuge 5702 (Ependorf, cat.no. 5702 000.019)
Ependorf centrifuge 5424 (Ependorf, cat.no. 5424 000.410)
Cell culture Flasks (PAA, cat.no. PAA70325X and PAA70375X)
Cell culture Dishes (PAA, cat.no. PAA20035X, PAA20060X, PAA20101X and PAA20151X)
Ependorf pipettes (Ependorf, cat.no. 3120 000.011, 3120 000.038, 3120 000.054 and 3120 000.062)
Accu-jet pro Pipette Controller (BrandTech, cat.no. 2026330)
Cell lifter (Corning, cat.no. 3008)
Pipette tips (Sarstedt, cat.no. 70.1130, 70.760.002, and 70.762)
Serological pipettes (Sarstedt, cat.no. 86.1252.025, 86.1253.025, 86.1254.025 and 86.1685.020)

Protein extraction

Ependorf cool centrifuge 5402 (Ependorf)
Vortex Genie 2 mixer (Scientific industries, cat no. SI-0256)
Sonicator (Sonics, cat no. VC 505)
GENios plate reader (Tecan)
Ice maker (Hoshizaki)
Forma -86°C ULT Freezer (Thermo)
10-100µl Multichannel pipette (Eppendorf, cat no. 3120 000.046)
96 well plate (Greiner Bio One, cat no. 655101-ORT)

Affinity Purification

Thermomixer compact (Ependorf)
Test tube rotator (Labinco, cat no. 29000)
QBD2 metal heating block (Fisher, cat no. BLD-560-020X)
BioRad mini protean 3 gel running system (with power supply and blot transfer unit)
SSM4 Rocker (Stuart)
DynaMag-2 (invitrogen, cat no. 123.21D)

Imaging

BioRad ChemDoc system (with GS-800 scanner)
Typhoon 9400 scanner (GE Healthcare)

LC/MS system

LTQ-Orbitrap mass-spectrometer (thermo-fisher scientific)

3.2.3 Reagent setup

TBS, 10x Dissolve 12.1g Tris and 88g NaCl in 800ml H_2O, and then adjust the pH to 7.5, add H_2O up to 1 liter and sterile filter it. This can be stored for several months.

TBST Dilute 100ml 10x TBS to 1 liter in H_2O, add 1ml Tween 20, mix well. Make fresh every week.

PBS, 10x Dissolve 6.8g KH_2PO_4 in 50ml H_2O and 26.1g K_2HPO4 in 150ml H2O, mix the solution, dilute it to 800ml, and use it to dissolve 87.7g NaCl, adjust the pH to 7.6, add H_2O up to 1 liter and sterile filter it. This can be stored for several months.

Lysis buffer Mix 50 mM TrisHCl pH 7.5, 250 mM sucrose, 5mM $MgCl_2$, 1mM DTT, 2mM ATP, 0.025% digitonin, 0.1% NP40. Make fresh. Additionally, protease inhibitor cocktails (Roche) can be added because they do not interfere with the proteasome activity.

Assay buffer Mix 50 mM TrisHCl pH 7.5, 250 mM sucrose, 5mM $MgCl_2$, 1mM DTT, 2mM ATP, 50mM KCl. Make fresh.

PD buffer Mix 50 mM TrisHCl pH 7.5 and 150 mM NaCl. This can be stored at 4°C for several weeks.

PD wash buffer I Dissolve 4 M urea in 50 mM NH_4HCO_3. This should be made freshly.

PD wash buffer II Mix 50 mM TrisHCl pH 7.5 and 10 mM NaCl. This can be stored at 4°C for several weeks.

On bead digest buffer Mix 100 mM TrisHCl pH 7.8, 100 mM NaCl, 1mM $CaCl_2$, 2% ACN and 10ng/µl Trypsin. This should be made freshly.

Active site elution solution Add 100 µM biotin in a mixture of 5% formic acid, 25% acetonitrile and 70% H_2O. This should be made freshly.

Sample buffer, 4x Mix 1ml 20% SDS, 1ml 0.625M TrisHCL pH6.8, 2.1ml 87% glycerol, 0.4ml BME, 0.1ml 10%BPB and 0.4ml H_2O. This can be stored at -20°C for several months.

Running buffer, 5x Mix 42.6g Tris, 214g Glycine and 300ml 10%SDS in 4 liters of H_2O, when dissolved well add H_2O up to 5 liters. This can be stored for several weeks.

In gel digest buffer Mix 10 mM NH_4HCO_3, pH 8, 5% ACN, 1 mM $CaCl_2$ and 10ng/µl Trypsin. This should be made freshly.

Gel block extraction solution 2% FA in 67% ACN/H_2O. Make fresh.

Stage tip solution A 0.5% FA in H_2O. Make fresh.

Stage tip solution B 0.5% FA in 80 % ACN/H_2O. Make fresh.

LC/MS sample solution Mix 95ml H_2O, 3ml ACN, and 0.1ml FA. This can be stored for several weeks.

PB 7.5 (Phosphate buffer) Mix 2ml 50µM Na_2HPO_4 and 7ml 50µM NaH_2PO_4. This can be stored for several weeks.

[L] Light Dimethyl labeling buffer (100µl)

Composition	Stock	Volume
PB 7.5		90µl
CH_2O	4% (v/v)	5µl
$NaBH_3CN$	0.6M	5µl

<CRITICAL> Prepare fresh.

[M] Medium Dimethyl labeling buffer (100µl)

Composition	Stock	Volume
PB 7.5		90µl
CD_2O	4% (v/v)	5µl
$NaBH_3CN$	0.6M	5µl

<CRITICAL> Prepare fresh.

[H] Heavy Dimethyl labeling buffer (100µl)

Composition	Stock	Volume
PB 7.5		90µl
$^{13}CD_2O$	4% (v/v)	5µl
$NaBD_3CN$	0.6M	5µl

<CRITICAL> Prepare fresh.

Proteasome inhibitors: While the use of bortezomib ester is described here, this ABPP method is also compatible with other proteasome inhibitors. It is used in the Biosyn lab as a screening technique to determine the activity, potency and subunit specificity of novel and potential proteasome inhibitors. Most compounds are soluble in DMSO and 500x stock dilutions is used prior to addition to the biological sample (like cell culture) to keep the DMSO concentration at <1% of the sample.

3.2.4 Equipment setup

Typhoon settings

The Typhoon 9400 scanner was used to scan the gel with fluorescent ABP labeled proteins. The scanning setting for MVB003 (Probe **3**) and MVB127 (Probe **2**) is Fluorescence Cy3/TAMRA, 600 PMT at 50 microns pixel resolution. The setting for LW124 (Probe **1**) is Fluorescence Cy2 (with Blue Laser), 600 PMT and 50 microns pixel resolution. The ImageJ software is used for quantifying the intensity of the bands after subtracting the background.

ChemDoc settings

The ChemDoc system for imaging the western blots and the GS800 scanner for imaging the commassie stained SDS-PAGE gels. The program Quanty One was used to operate the imaging systems, and subsequently used for quantifying the intensity of the bands and/or lanes. Setting for imaging the western blots is chemiluminescence.

Orbitrap settings

General settings of the mass spectrometer were: an electrospray voltage of 1.5 kV was

applied to the emitter, no sheath and auxiliary gas flow, ion transfer tube temperature 150°C, capillary voltage 15V, tube lens voltage 150V. Internal mass calibration was performed with air-borne protonated polydimethylcyclosiloxane (m/z=445.12002) and the plasticizer protonated dioctyl phthalate ions (m/z= 391.28429) as lock mass(*41*). For shotgun proteomics analysis, 10 µl sample was pressure loaded on the trap column at 10 µl/min flow for 5 min followed by peptide separation with a gradient of 35 min 5-30% B, 15 min 30-60% B and 5 min 100% A at a flow of 300 µl/min split to 250 nl/min by the LTQ divert valve. For each data dependent cycle, one full MS scan (300-2000 m/z) acquired at high mass resolution (60,000 at 400 m/z, AGC target $1x10^6$, maximum injection time 1,000 ms) in the Orbitrap was followed by 3 MS/MS fragmentations in the LTQ linear ion trap (AGC target $5x10^3$, maximum injection time 120 ms) from the three most abundant ions. MS^2 settings were: collision gas pressure 1.3 mT, normalized collision energy 35%, ion selection threshold of 500 counts, activation q = 0.25 and activation time of 30 ms. Fragmented precursor ions that were measured twice within 10 s were dynamically excluded for 60s and ions with z<2 or unassigned were not analyzed(*17*).

BOX1 | Cell culture and protein extraction ● **Timing 1 d**

In this study we have used adherent and suspension growing cells from human and mouse to show that the activity-based protein profiling method can be used in virtually all in vitro cultured mammalian cell types including primary cells. In (micro)organisms bearing a cell wall (e.g. yeast, plants) the method works well in cell lysates but should be optimized (mostly by using higher probe concentrations) for living cells.

1 | Cell culture MCF7(Human breast carcinoma), and EL4 (Murine T-lymphocyte) are grown in DMEM medium with 10% fetal calf serum (FCS), 0.1mg/ml penicillin and 0.1mg/ml streptomycin, in a 37°C, 7% CO_2 incubator. AMO1 (plasmacytoma) cells are grown in RPMI 1640 medium with 10% fetal calf serum (FCS), 0.1mg/ml penicillin, 0.1mg/ml streptomycin, and 2mM stable L-Glutamine and in 37°C, 5% CO_2 incubator.

2 | Harvesting Harvest MCF7 cells by scrapping them in cold PBS. Harvest EL4 and AMO1 cells by centrifugation directly at 1200rpm for 5 min. Wash the cells with ice cold PBS for three times to remove serum and free inhibitor molecules.

■ **PAUSE POINT**: The cells can be stored in -80°C freezer for a couple of months.

3 | Protein extraction Thaw cell pellets from -80°C on ice, add 4 volumes of lysis buffer to the cell pellet and leave on ice for 15-30 min. Sonicate the cells on ice for 3x10 seconds at 12 watts, with 5 second pulse/pause. Centrifuge the lysate for 10 minutes at top speed (13Krpm) at 4°C

! CAUTION; Check by eye whether the supernatant is clearly transparent.

4 | Lysate storage Transfer the supernatant to new Eppendorf tubes. Determine the protein concentration by Bradford Protein Assay, and freeze the lysate at -80°C for future use.

■ **PAUSE POINT**: The lysate can be kept in -80°C for a couple of weeks.

! CAUTION: Aliquot the lysate in small portions to avoid freezing and thawing for multiple times.

3.3 Procedure

There are two options depending on whether you want to do high-resolution fluorescent ABPP on cell lysates (option A) or quantitative LC/MS based ABPP (option B). For option A, three different fluorescent probes for the ABPP are recommended, LW124 (**1**) to detect β1 (β1i), MV127 (**2**) to detect β5 (β5i), MV003 (**3**) to detect all the subunits (Fig. 1)(*17, 18*). This is done in order to get perfect separation between the constitutive and immunoproteasome subunits on SDS-PAGE for quantification purposes. For option B, to quantify proteasome in living cells, N_3-epoxomicin (**5**) is recommended because it is cell permeable, and biotin-phosphine (**6**) for the bioorthorgonal ligation to install the biotin tag; for in vitro experiments biotin-epoxomicin (**4**) can be used directly (Fig. 1)(*17, 21-23*).

Option A High resolution fluorescent ABPP ● Timing 3 d

(i) Make 10 mM stocks of all probes and inhibitors in DMSO and use DMSO as negative control for your experiments.

(ii) Seed $2.5x10^6$ MCF7 cells in 10 cm petri dishes, culture to 70-80% confluency and treat with 0, 20, 50 and 100 nM bortezomib for 16 hours. From logarithmic growing AMO1 cultures use $10x10^6$ cells and treat with 0, 10, 20, 50 nM bortezomib for 16 hrs(*20*).

! CAUTION: The number of cells, concentration of bortezomib and time of treatment should be determined according to the sensitivity of the cells to the proteasome inhibitor.

(iii) Lyse the cells and determine the protein concentration as described in BOX1. Typical cell pellets of 10-20 µl are lysed in 40-80 µl lysis buffer and the protein concentration is in the range of 10-20 µg/µl.

(iv) Prepare 20 µg of total protein for each sample; bring the total volume up to 9 µl by adding assay buffer. Then add 1µl 10x working stock of the ABP, yielding final ABP concentrations of 0.5 µM LW124 (**1**), 0.5 µM MVB127 (**2**) or 2 µM MVB003 (**3**). Incubate the mixture at 37°C for 1 hour. Add 3 µl 4x sample buffer, and boil the sample for 5min at 100°C.

(v) The lanes with 0 nM bortezomib were used as positive controls, as shown in Figure 3, there is clear proteasome inhibition by the treatment. If any other proteasome inhibitor is used in a similar experiment, as negative controls deactivated lysate by boiling with 1% SDS or blocking the proteasome with 10 µM N3-epoxomicin (**5**) prior to incubation with the activity-based probe were suggested. In these lanes no clear proteasome bands were expected to appear.

(vi) Separate the proteins on 12.5% SDS-PAGE.

(vii) Scan the gels on a typhoon scanner with correct settings (as described in the Equipment Setup).

(viii) As loading control, coomassie stain the gels, scan on the BioRad GS-800 scanner and determine the signal intensity for the total amount of protein in each lane using the QuantityOne program. Normalize all lanes to the highest intensity value to

correct for pipetting errors.

(viii) Quantify the fluorescence image with the ImageJ software by drawing a region of interest (ROI) around individual bands followed by background subtraction prior to calculating the proteasome activity and correction for pipetting errors. The fluorescent gel image and the quantification graphs from MCF7 and AMO1 are shown in Figure 3.

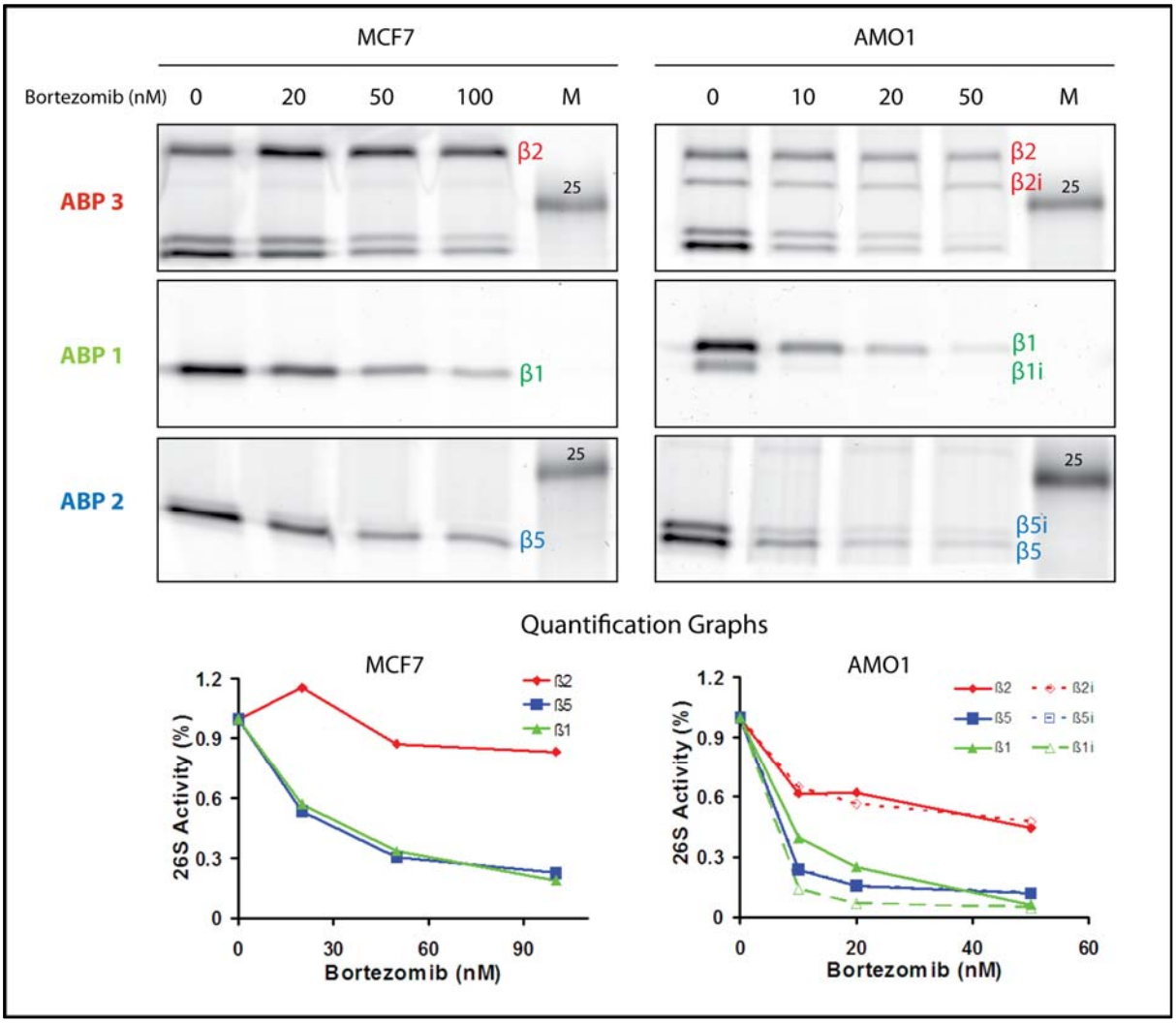


Figure 3: Results of fluorescence based quantitative proteasome ABPP in adherent growing breast carcinoma cell line MCF7 expressing constitutive proteasome and suspension growing cells AMO1 expressing both constitutive and immuno proteasomes incubated with increasing bortezomib concentrations. Top panel global proteasome assay with MVB003 (**3**), used for quantification of β2/ β2i activities, middle panel β1/β1i quantification with LW124 (**1**), lower panel β5/β5i quantification with MVB127 (**2**). The graphs show the proteasome activity, quantified from the gel bands, after treatment with increasing concentrations bortezomib.

Option B Quantitative LC/MS based ABPP ● Timing 8 d

(i) Cell treatment Treat 1 x 10^8 EL4 cells with the proteasome inhibitor bortezomib at following concentrations: 0, 10nM, 100nM, 1μM, and 10μM. DMSO can be used as negative control and the concentration of DMSO is kept below 1%. After 2 hours of incubation, add 10μM N_3-epoxomicin (**5**) to each sample, and incubate again for 2 hours. Check the cells by light microscopy to observe signs of necrosis, shrinking, apoptosis or other toxicities. Lyse the cells as described in BOX1.

(ii) **ABP reaction** For direct labeling, cell lysate (1-2 mg protein in total) plus 10 μM biotin-epoxomicin (**4**) is incubated at 37°C for 1 hour. For two step labeling, add 250 μM biotin-phosphine (**6**) to cell lysates from N_3-epoxomicin (**5**) treated living cells and incubate at 37°C for 1 hour. Empirically, the 2-step labeling of the proteasome in native (non-denatured) lysates with a high excess of biotin-phosphine (**6**) shows the best reaction and pull down efficiency compared to the Cu(I) catalyzed Huisgen 1,3-dipolar cycloaddition (click chemistry) or the copper free click(*34*, *37*).

! CAUTION: Use lysis buffer to adjust total volume to 180 μl.

(iii) **Denaturation** Add 20μl 10% SDS to the reaction mixture and vortex slightly, and then boil it at 100°C for 5 minutes for denaturation. Vortex the solution homogeneously, and centrifuge 5 minutes at 12 krpm to pellet any aggregates. Transfer the clear supernatant into a new 2 ml low-binding eppendorf tube.

(iv) **Chloroform/Methanol precipitation** (steps B iv-ix) (*38*) Add 800 μl Methanol to the 200 μl protein solution, vortex vigorously.

(v) Add 200 μl Chloroform, vortex vigorously.

(vi) Add 600 μl H_2O (to get phase separation), vortex vigorously and centrifuge (2 min, 9000 g). A 3-layer separation is derived, upper layer of water, a film of protein in the middle, and the lower layer of chloroform.

(vii) Carefully remove and discard upper layer.

(viii) Add 600 μl methanol, vortex at low speed, thoroughly and centrifuge (2 min, 9000 g).

! CAUTION: Avoid disrupting the protein membrane in too many small pieces because they will not sediment well during the next centrifugation step.

(ix) Carefully discard supernatant by pippetting and dry the pellet for exactly 5 minutes to the air in the flow air cabinet, tube up side down.

▲ **CRITICAL STEP**. Both too wet or too dry pellets do not hydrate well. The protein pellet should be dried for exactly 5 min. if it is dryer or wetter, it will be hard to dissolve in the buffer for further steps.

(x)Rehydrate the protein pellet in 180 μl urea buffer (8 M Urea/100 mM NH_4HCO_3) for 15 min at room temperature.

(xi) **Reduction and alkylation** (to disable cysteine bridge formation)(Steps B xi-xiii)

Add 10 µl fresh 90 mM DTT (in 8 M Urea/100 mM NH_4HCO_3) to the 180 µl sample and incubate this for 30 min at 37°C. Note that 1%SDS or the acid labile surfactant RapiGest® (Waters) can be used instead of 8 M Urea.

! CAUTION: Temperatures > 37°C can give carbamylation of proteins by urea.

(xii) Add 15 µl 200 mM Iodoacetamide (in 8 M Urea/100 mM NH_4HCO_3) solution to the 190 µl sample and incubate for 30 min at room temperature in the dark.

! CAUTION: The reaction is light sensitive.

(xiii) Centrifuge 5 min at top speed and transfer the supernatant to a new 2 ml eppendorf tube.

(xiv) **Chloroform/Methanol precipitation** Repeat steps B (iv-ix). This extra precipitation step is suggested in order to remove the excess free ABP or free tag in the case of the 2-step labelling protocol.

(xv)**Stepwise dilution of SDS** (steps B xv-xviii) Add 25 µl of 2% SDS in PD buffer to the pellet and vortex to help dissolve. Dilute the SDS with normal PD buffer step by step. Heating to 70°C can also enhance protein solubility

(xvi) Dilute 3 times with 25 µl PD buffer, vortex after each addition of PD buffer, and make sure the solution becomes clear.

(xvii) Add the rest in 9 x100 µl steps of PD buffer, to make the SDS end concentration 0.05% and obtain a clear solution.

! CAUTION: Be patient here, it may take a long time to homogeneously disperse the entire pellet. But it should be dissolved as much as possible. If necessary, use a heated water bath and/or sonic bath to help dissolve. It is important to avoid lumps of protein here because they will a-specifically stick to the beads, yielding high background

▲ **CRITICAL STEP**. This is the hardest step in the whole protocol. This tiny volume might not be enough to solve your protein completely but more is not possible otherwise it will be ended up with too high SDS concentration in the pull down step and lose efficiency there. Disperse the protein in this volume by vortexing vigorously. Do not pipette up and down, because lumps of protein can stick in the tips, plus there is a potential for foam formation. The proteins will dissolve once the PD buffer is stepwise added. This increases the volume while keeping the concentration of SDS relatively high.

(xviii)Centrifuge at top speed for 5 min, and transfer the supernatant to a new 2 ml Eppendorf tube.

▲ **CRITICAL STEP**. After taking up the volume of the PD mixture to 1 ml by adding PD buffer, centrifuge the mixture at top speed for 5 min to remove all the insoluble particles. Or, there might be non-specific protein binding to the beads, and increase of background during analysis.

(xix) **Pull down with paramagnetic beads** Add 50µl streptavidin magnetic beads (pre-washed with water, PD buffer and 0.05% SDS PD buffer). Pull down for 1 hour at room

temperature with vigorous shaking.

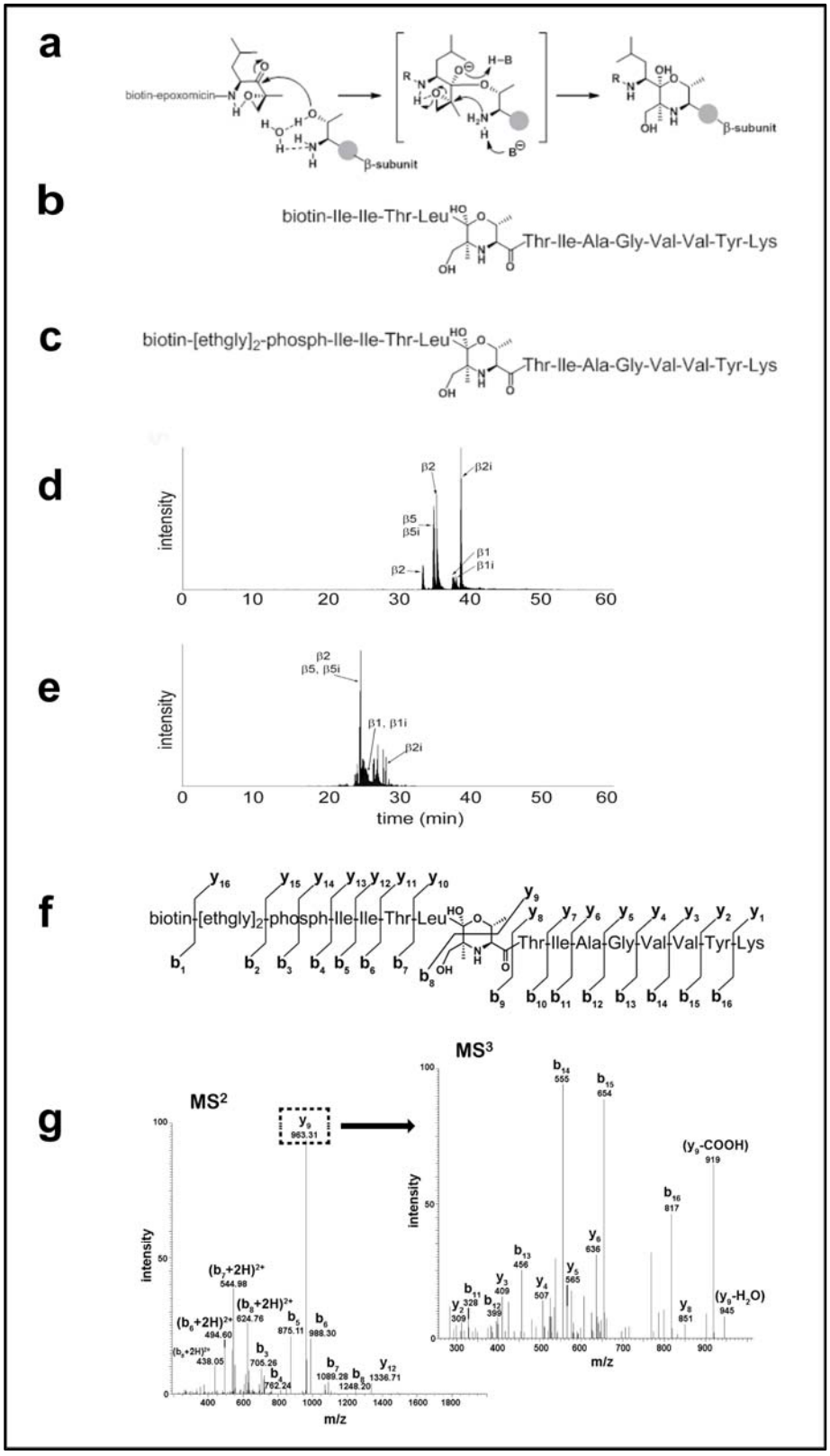


Figure4: Active site identification of the murine β2 subunit by one step and two-step ABPP with LC-MS analysis. Reaction mechanism of biotin-epoxomicin (**4**) with the catalytic T1

residue (**a**); schematic representation of the direct [biotin-epoxomicin]- [β2 active-site peptide] (**b**) and the two-step [biotin-phosphine-epoxomicin]- [β2 active-site peptide] (**c**) construct obtained after on-bead tryptic digest; LC/MS analysis and identification of the [biotin-epoxomicin]-[active site peptides] (**d**) and two-step [biotin-phosphine-epoxomicin]-[active-site peptides] (**e**) from constitutive and immuno proteasomes; theoretical MS/MS fragmentation pattern for the two-step [biotin-phosphine-epoxomicin]- [β2 active-site peptide] construct (**f**), experimental data for f showing the diagnostic ions b6, b7, b8, b3, b4 after MS/MS fragmentation and subsequently the MS^3 fragmentation of the y9 ion revealing the amino acid sequence of the [β2 active-site peptide] part of the construct (**g**).

(xx) Leave tubes with pull down sample on the DynaMag, wait for 3 minutes, and pipette the supernatant to another tube.

(xxi) Run a sample of the supernatant on a SDS-PAGE gel to assess the pull down efficiency (Fig 6).

(xxii) **Washing** Wash the beads as shown in the in-text table, then divide them: 2/3 for in gel detection, and 1/3 for on bead digest.

Once	300µl PD buffer (containing 0.1% SDS)
Twice	300µl PD buffer
Once	300µl wash buffer I
Once	300µl wash buffer II
Twice	300µl water

An interesting observation is that after the pull-down, the beads might behave slightly different in solution because their surface properties have changed by the binding of the biotinylated proteins.

(xxiii) **In-gel detection** (steps B xxiii-xxxiii) Elute the proteins from the beads with 100µl 1x sample buffer containing 10µM biotin, by boiling for 5 min at 100°C.

(xxiv) Put samples on DynaMag for 3 min, transfer supernatant while still warm to another tube.

▲ **CRITICAL STEP**. Keep the boiled sample tubes on the DynaMag, when loading the samples onto the gel, to prevent loading the beads. If the beads are also loaded, background can be increased.

(xxv) Resolve samples on 12.5% SDS-PAGE and silver stain. After staining, scan the stained gel on the GS-800 scanner, excise protein bands and perform in-gel digestion steps(*39*). A typical silver stained gel image of biotin-epoxomicin (**4**) pull down in EL4 cells is shown in Figure 6a. Cut desired gel bands into 1 mm^3 blocks with a surgery knife, and keep them in a 1.5ml Eppendorf tube.

▲ **CRITICAL STEP**. In case coomassie stain is used, destaining is necessary, because the blue dye is not removed by stage tip desalting and it is unwanted in the LC/MS system.

■ **PAUSE POINT**: The gel blocks can be kept in -20°C for a couple of weeks.

(xxvi) Wash the gel blocks with 100µl Milli-Q water.

(xxvii) Discard water and add 500µl ACN to shrink the gel blocks, vortex slightly (the gel blocks will shrink, become white and sticky).

(xxviii) Discard ACN, dry gel blocks for 5 minutes in the air flow.

(xxix) Swell gel blocks with 25µl in gel digest buffer on ice for at least 2 hr.

(xxx) Digest in 37°C incubator (with shaker) overnight.

(xxxi) Add 100µl gel block extraction solution and incubate at 37°C for 15 min.

(xxxii) Spin down and transfer the supernatant to new Eppendorf tubes and evaporate the ACN from the samples at 50°C in a SpeedVac until volume is <30µl.

(xxxiii) Add 50µl stage tip solution A to adjust pH <5.

! CAUTION: The peptides should be desalted by stage tips before LC/MS analysis.

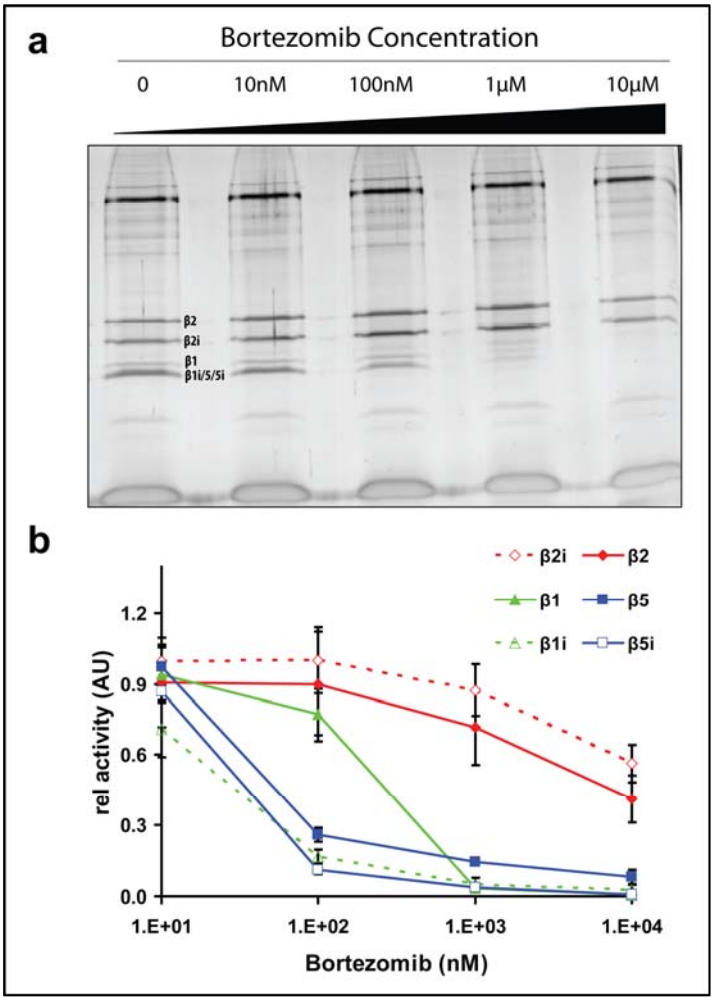


Figure 5: Competitive ABPP assay in EL4 cells (constitutive and immuno proteasome) incubated with increasing bortezomib concentrations for 2 hrs and residual β-subunit activity captured with 10 µM N_3-epoxomicin (**5**) and biotin phosphine (**6**).

Qualitative analysis by SDS-PAGE and silver stained gel after pulldown with Streptavidin beads (**a**); quantification of the β-subunits activity by on-bead tryptic digest, dimethyl stable isotope labeling and LC/MS based proteomics (**b**).

(xxxiv) **On-bead digestion** (steps B xxxiv-xxxv) Add 100µl on-bead digest buffer to the beads and digest proteins at 37°C with shaking overnight.

(xxxv) Place tubes on DynaMag for 3 min, transfer supernatants to new Eppendorf tubes and add 5µl FA for adjusting the pH <4 to inhibit trypsin.

! CAUTION: The peptides should be desalted by stage tips before LC/MS analysis.

(xxxvi) Eluting the active site peptides After on bead digest, wash the beads with 100 µl water, incubate with 100 µl active-site elution solution for 30 min at 37°C to release the active site peptide from the beads, evaporate the ACN in speedvac and add 50ul stage tip solution A. The active site peptides from murine proteasome obtained after tryptic digest of EL4 cell lysates are shown in the in-text table. The reaction mechanism and typical results for the active site identification are shown in Figure 4.

Murine proteasome subunit	Active site peptide sequence after tryptic digest
β1	TTIMAVQFNGGVVLGADSR
β2	TTIAGVVYK
β5	TTTLAFK
β1i	TTIMAVEFDGGVVVGSDSR
β2i	TTIAGLVFR
β5i	TTTLAFK

! CAUTION: The peptides should be desalted by stage tips before LC/MS analysis.

▲ **CRITICAL STEP**.All the tryptic digested peptides should be desalted by Stage Tips before loaded on the LC/MS for analysis. For the protein identification samples (like in gel digest and active site peptide identification samples), a single Stage Tip desalting is enough for the LC/MS sample preparation, which is described in step B (xxxvii). For the samples that need to be quantified, for instance the on bead digest samples treated with different bortezomib concentrations, should be both dimethyl labeled by stable isotopes and desalted on Stage Tip before LC/MS analysis, go directly to step B (xl).

(xxxvii) **Desalting using stage tips** (steps B xxxvii-xl) The Stage Tips are made of C18 material inserted in 200µl pipette tips(*40*). Desalt the peptide samples as shown in the in-text table.

Conditioning	50µl methanol
Conditioning	50µl Stage Tip solution B
Conditioning	50µl Stage Tip solution A
Loading	Load samples on Stage Tips
Washing	100µl Stage Tip solution A
Elution	100µl Stage Tip solution B

! CAUTION: Run the solutions through the stage tips by low speed centrifugation (400-800 g).

(xxxviii) Evaporate the ACN from samples until the volume is < 20µl.

(xxxix) Add 50µl LC/MS sample solution, and then the samples are ready for LC/MS analysis.

(xl) **On Stage Tip dimethyl labeling** (steps B xl-xlvi) Follow the protocol step B (xxxvii) till washing. After washing, load 20µl of L, M, or H reagents into each of the assigned stage tips.

(xli) Centrifuge at 100 g for 5 min in an eppendorf centrifuge (adjust the speed such that 20µl of reagent flows through the stage tip in around 5 min). Prior to this step you can determine the optimal centrifuge speed by using a conditioned stage tip and loaded with 20µl PB7.5.

(xlii) Repeat step B (xli) four times.

▲ **CRITICAL STEP**. Keep the dimethyl labeling buffer on stage tips for longer than 10 min is important, or there might be insufficient labeling.

(xliii) Wash with 100µl Stage Tip solution A.

(xliv) Elute peptides with 100µl Stage Tip solution B.

(xlv) Mix the three samples 1:1:1 in a new Eppendorf tube and evaporate ACN in Speedvac at 50°C until the volume is <60µl.

(xlvi) Add 100 ul LC/MS sample solution and perform LC/MS analysis.

! CAUTION: It is important to keep solutions and reagents on ice to prevent degradation. The isotopic reagents are light sensitive; use brown tubes to store the solutions.

■ **PAUSE POINT**: The prepared LC/MS samples can be kept at 4°C for several days.

(xlvii) **LC/MS analysis, Mascot search and MaxQuant processing** Extract peak file lists from the .raw data files of the ABPP quantification results using the DTA supercharge module from MSQuant and submitted to automated peptide identification by the Mascot search engine using a false discovery rate (FDR) of 1% and ion cutoff scores of >25. Alternatively, the MaxQuant software can be used for the relative quantification of the proteomics data. The competitive inhibition profiles of bortezomib were normalized compared to the negative control and presented in Figure 5b. For an elaborate protocol of the search engines details please see references of Boersema et al. and Cox et al.(*24*, *26*, *27*).

● **Timing**

Option A High resolution fluorescent ABPP (including Box I): 3 d

Option B Quantitative LC/MS based ABPP:

Step B (i) (including Box I), protein extraction: 2 d

Step B (ii-xxii), affinity purification: 1 d

Step B (xxiii-xxxiii), in gel digestion: 2 d

Step B (xxxiv-xlvi), on bead digestion and isotopic labeling: 2 d

Step B (xlvii), LC/MS analysis and data processing: 1 d

3.4 Anticipated results

Figure 3 shows the results for a fluorescence, gel-based ABPP quantification experiment where the β-subunit activities were assayed after treatment of constitutive and immuno proteasome expressing cell lines with bortezomib. Proteins were resolved on a 12.5% SDS-PAGE gel and the separation between the β2 and β2i subunits is sufficient to use the pan-reactive probe MVB003 (**3**) to determine their activities. When only constitutive proteasome is present MVB003 can also be used to quantify the closer running β1 and β5 bands. However, in cells expressing both constitutive and immuno proteasomes the resolution between the β1, β1i, β5, β5i is poor and MVB003 cannot be used for reliable quantification. In this case we use the subunit specific probes LW124 (**1**) for β1/β1i and MVB127 (**2**) for β5/β5i separation and quantification. The results show that adherent MCF7 cells are less sensitive to bortezomib compared to the myeloma cell line AMO1. Interestingly, bortezomib lowers also the β2/β2i activity in AMO1 cells, which is not observed in MCF7 cells. A possible explanation might be that the immuno proteasomes expressed by the AMO1 but not by MCF7 cells, are more active and are thus more efficiently blocked by bortezomib. This method is considered to be a fast, robust and medium throughput (2-8 assays/day) screen for the proteasome activity that can be easily implemented in any biochemical laboratory.

Figure 4 shows a study for the active site identification of the murine β2 subunit using ABPP and LC/MS analysis. The epoxyketone electrophilic trap reacts with the catalytic Thr1 residue yielding a stable morpholine ring. For a two step ABPP protocol, biotin-phosphine (**6**) is then ligated to N_3-epoxomicin (**5**) followed by affinity purification with Streptavidin beads, on-bead tryptic digest and elution of ABP-active-site constructs. Figure 4b and 4c show the one step and two-step constructs for the β2 subunit, respectively. Figure 4d and 4e show chromatograms of the active site constructs of all β subunits after one step and two step ABPP, respectively. Figure 4f shows the theoretical fragmentation of the two-step ABP-active site construct of the β2 subunit after collision induced dissociation (CID) of the peptide bonds in b and y ions. Figure 4g shows the experimental data of 4f. The N-terminus of the molecule is labile and yields a typical fragmentation pattern of b ions and the prominent appearance of the intact y9 ion containing the peptide sequence originating from the β2 protein. Subsequently, the y9 ion was isolated in the LTQ iontrap and subdued to a new round of fragmentation (MS^3) revealing the β2 peptide sequence. Analysis of active-site peptides is interesting from a toxicological perspective because covalent binding inhibitors might cause idiosyncratic drug related toxicity. Expression of an epitope of the active site peptide attached to a xenobiotic compound on MHC class I molecules might induce an adverse immunological response and lead to allergy against the compound which will exclude it from therapy. The method presented here can be used to determine epitope expression on MHC class I

molecules.

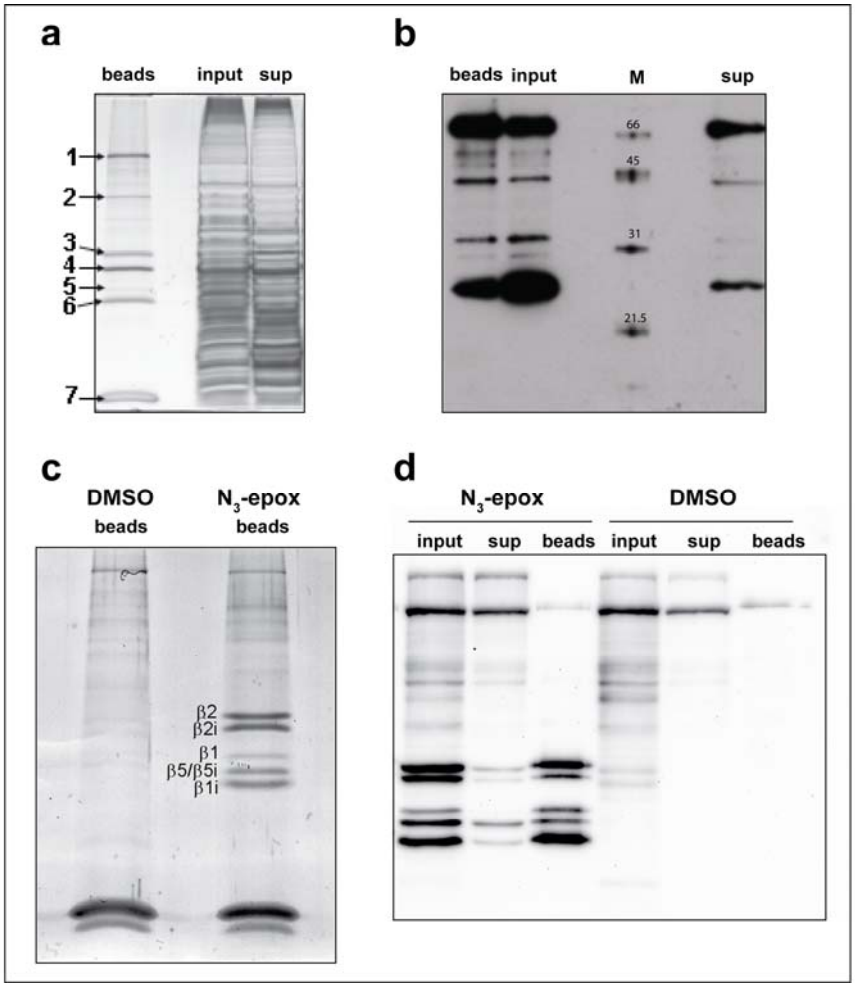


Figure 6: Direct and two-step activity-based affinity purification of the proteasome β-subunits with Streptavidin beads from EL4 lysate. SDS-PAGE/silver stain analysis after treatment with biotin-epoxomicin (**4**) showing proteins eluted from the beads, input and supernatant, bands indicated by arrows were identified by LC/MS (1: Propionyl-CoA carboxylase, 2: 60 S ribosomal protein, 3:β2, 4:β2i, 5:β1, 6:β5β5iβ1i, 7: Streptavidin) (**a**); western blot probed with Streptavidin HRP of the same samples from **a** (M, molecular marker) (**b**); SDS-PAGE/silver stain analysis after treatment with DMSO or N_3-epoxomicin (**5**) and ligation with biotin-phosphine (**6**) showing proteins eluted from the beads (**c**); western blot probed with Streptavidin HRP of the samples from **c** next to the input and supernatant samples (**d**).

The experiment of Figure 5 entails a competitive ABPP experiment between bortezomib and N_3-epoxomicin (**5**) in living EL4 cells that express both constitutive and immuno proteasomes. Cells were treated with 0, 10, 10^2, 10^3 and 10^4 nM bortezomib concentrations in parallel, the residual proteasome activity was captured by two-step ABPP using the N_3-epoxomicin/biotin-phosphine couple, proteins were affinity purified with

Streptavidin beads and analysed qualitatively by SDS-PAGE or subdued to on-bead digest with trypsin and the resulting peptides were captured on stage tips. The first group contains peptides from the 0, 10 and 10^2 nM bortezomib treatment that were labelled light, medium and heavy respectively, eluted from stage tips, mixed 1:1:1 and analysed by LC/MS. The second group contains peptides from the 0, 10^3 and 10^4 nM bortezomib treatment that were labelled light, medium and heavy respectively, eluted, mixed 1:1:1 and analysed by LC/MS. The 0 nM bortezomib condition was used as benchmark to combine the two groups resulting in a quantification report on the proteasome activity. The relative quantitative analysis shows that bortezomib first hits the β1i β5 and β5i subunits and then the β1 activity. At higher concentrations (>500 nM) also the β2 and β2i subunits are inhibited. This method is suitable for quantitative determination of proteasome activity in living cells and it is more accurate than the fluorescence gel-based method described previously because all active subunits are determined in the same LC/MS run.

Figure 6 is an illustration of the efficacy and robustness of the affinity purification protocol used for qualitative or quantitative determination of the proteasome activity. The results show typical pull-down efficiencies for one-step and two-step ABPP work flows followed by Streptavidin purification, SDS-PAGE separation of proteins and western blot analysis of biotinylated β-subunits visualised with Streptavidin-HRP. This method is routinely used to determine the pull-down efficiency.

References and notes

1. Rock, K. L., and Goldberg, A. L. (1999) Degradation of cell proteins and the generation of MHC class I-presented peptides, *Annu Rev Immunol 17*, 739-779.

2. Kloetzel, P. M., and Ossendorp, F. (2004) Proteasome and peptidase function in MHC-class-I-mediated antigen presentation, *Curr Opin Immunol 16*, 76-81.

3. Kane, R. C., Bross, P. F., Farrell, A. T., and Pazdur, R. (2003) Velcade: U.S. FDA approval for the treatment of multiple myeloma progressing on prior therapy, *Oncologist 8*, 508-513.

4. Jung, T., Catalgol, B., and Grune, T. (2009) The proteasomal system, *Mol Aspects Med 30*, 191-296.

5. Marques, A. J., Palanimurugan, R., Matias, A. C., Ramos, P. C., and Dohmen, R. J. (2009) Catalytic mechanism and assembly of the proteasome, *Chem Rev 109*, 1509-1536.

6. Borissenko, L., and Groll, M. (2007) 20S proteasome and its inhibitors: crystallographic knowledge for drug development, *Chem Rev 107*, 687-717.

7. Murata, S., Sasaki, K., Kishimoto, T., Niwa, S., Hayashi, H., Takahama, Y., and Tanaka, K. (2007) Regulation of CD8+ T cell development by thymus-specific proteasomes, *Science 316*, 1349-1353.

8. Verdoes, M., Florea, B. I., van der Marel, G. A., and Overkleeft, H. S. (2009) Chemical Tools To Study the Proteasome, *Eur J Org Chem*, 3301-3313.

9. Kisselev, A. F., and Goldberg, A. L. (2005) Monitoring activity and inhibition of 26S proteasomes with fluorogenic peptide substrates, *Methods Enzymol 398*, 364-378.

10. Li, N., Overkleeft, H. S., and Florea, B. I. (2012) Activity-based protein profiling: an enabling technology in chemical biology research, *Curr Opin Chem Biol 16*, 227-233.

11. Fenteany, G., Standaert, R. F., Lane, W. S., Choi, S., Corey, E. J., and Schreiber, S. L. (1995) Inhibition of proteasome activities and subunit-specific amino-terminal threonine modification by lactacystin, *Science 268*, 726-731.

12. Meng, L., Mohan, R., Kwok, B., Elofsson, M., Sin, N., and Crews, C. (1999) Epoxomicin, a potent and selective proteasome inhibitor, exhibits in vivo antiinflammatory activity, *Proceedings of the National Academy of Sciences of the United States of America 96*, 10403-10411.

13. Bogyo, M., McMaster, J. S., Gaczynska, M., Tortorella, D., Goldberg, A. L., and Ploegh, H. (1997) Covalent modification of the active site threonine of proteasomal beta subunits and the Escherichia coli homolog HslV by a new class of inhibitors, *Proceedings of the National Academy of Sciences of the United States of America 94*, 6629-6634.

14. Kessler, B. M., Tortorella, D., Altun, M., Kisselev, A. F., Fiebiger, E., Hekking, B. G., Ploegh, H. L., and Overkleeft, H. S. (2001) Extended peptide-based inhibitors efficiently target the proteasome and reveal overlapping specificities of the catalytic beta-subunits, *Chem Biol 8*, 913-929.

15. Berkers, C. R., Verdoes, M., Lichtman, E., Fiebiger, E., Kessler, B. M., Anderson, K. C., Ploegh, H. L., Ovaa, H., and Galardy, P. J. (2005) Activity probe for in vivo profiling of the specificity of proteasome inhibitor bortezomib, *Nat Methods 2*, 357-362.

16. Verdoes, M., Florea, B. I., Menendez-Benito, V., Maynard, C. J., Witte, M. D., van der Linden, W. A., van den Nieuwendijk, A. M., Hofmann, T., Berkers, C. R., van Leeuwen, F. W., Groothuis, T. A., Leeuwenburgh, M. A., Ovaa, H., Neefjes, J. J., Filippov, D. V., van der Marel, G. A., Dantuma, N. P., and Overkleeft, H. S. (2006) A fluorescent broad-spectrum proteasome inhibitor for labeling proteasomes in vitro and in vivo, *Chem Biol 13*, 1217-1226.

17. Florea, B. I., Verdoes, M., Li, N., van der Linden, W. A., Geurink, P. P., van den Elst, H., Hofmann, T., de Ru, A., van Veelen, P. A., Tanaka, K., Sasaki, K., Murata, S., den Dulk, H., Brouwer, J., Ossendorp, F. A., Kisselev, A. F., and Overkleeft, H. S. (2010) Activity-based profiling reveals reactivity of the murine thymoproteasome-specific subunit beta5t, *Chem Biol 17*, 795-801.

18. Verdoes, M., Willems, L. I., van der Linden, W. A., Duivenvoorden, B. A., van der Marel, G. A., Florea, B. I., Kisselev, A. F., and Overkleeft, H. S. (2010) A panel of subunit-selective activity-based proteasome probes, *Org Biomol Chem 8*, 2719-2727.

19. Kraus, M., Ruckrich, T., Reich, M., Gogel, J., Beck, A., Kammer, W., Berkers, C. R., Burg, D., Overkleeft, H., Ovaa, H., and Driessen, C. (2007) Activity patterns of proteasome subunits reflect bortezomib sensitivity of hematologic malignancies and are variable in primary human leukemia cells, *Leukemia 21*, 84-92.

20. Ruckrich, T., Kraus, M., Gogel, J., Beck, A., Ovaa, H., Verdoes, M., Overkleeft, H. S., Kalbacher, H., and Driessen, C. (2009) Characterization of the ubiquitin-proteasome system in bortezomib-adapted cells, *Leukemia 23*, 1098-1105.

21. Saxon, E., and Bertozzi, C. (2000) Cell surface engineering by a modified Staudinger reaction, *Science (New York, N.Y.) 287*, 2007-2017.

22. Ovaa, H., van Swieten, P. F., Kessler, B. M., Leeuwenburgh, M. A., Fiebiger, E., van den Nieuwendijk, A. M., Galardy, P. J., van der Marel, G. A., Ploegh, H. L., and Overkleeft, H. S. (2003) Chemistry in living cells: detection of active proteasomes by a two-step labeling strategy, *Angew Chem Int Ed Engl 42*, 3626-3629.

23. Willems, L. I., Li, N., Florea, B. I., Ruben, M., van der Marel, G. A., and Overkleeft, H. S. (2012) Triple bioorthogonal ligation strategy for simultaneous labeling of multiple enzymatic activities, *Angew Chem Int Ed Engl 51*, 4431-4434.

24. Boersema, P. J., Raijmakers, R., Lemeer, S., Mohammed, S., and Heck, A. J. (2009) Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics, *Nat Protoc 4*, 484-494.

25. Raijmakers, R., Dadvar, P., Pelletier, S., Gouw, J., Rumpel, K., and Heck, A. J. R. (2010) Target Profiling of a Small Library of Phosphodiesterase 5 (PDE5) Inhibitors using Chemical Proteomics, *Chemmedchem 5*, 1927-1936.

26. Cox, J., and Mann, M. (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification, *Nat Biotechnol 26*, 1367-1372.

27. Cox, J., Matic, I., Hilger, M., Nagaraj, N., Selbach, M., Olsen, J. V., and Mann, M. (2009) A practical guide to the MaxQuant computational platform for SILAC-based quantitative proteomics, *Nat Protoc 4*, 698-705.

28. Kisselev, A. F., Songyang, Z., and Goldberg, A. L. (2000) Why does threonine, and not serine, function as the active site nucleophile in proteasomes?, *J Biol Chem 275*, 14831-14837.

29. Singh, J., Petter, R. C., Baillie, T. A., and Whitty, A. (2011) The resurgence of covalent drugs, *Nat Rev Drug Discov 10*, 307-317.

30. Muchamuel, T., Basler, M., Aujay, M. A., Suzuki, E., Kalim, K. W., Lauer, C., Sylvain, C., Ring, E. R., Shields, J., Jiang, J., Shwonek, P., Parlati, F., Demo, S. D., Bennett, M. K., Kirk, C. J., and Groettrup, M. (2009) A selective inhibitor of the immunoproteasome subunit LMP7 blocks cytokine production and attenuates progression of experimental arthritis, *Nat Med 15*, 781-787.

31. Kallemeijn, W. W., Li, K. Y., Witte, M. D., Marques, A. R., Aten, J., Scheij, S., Jiang, J., Willems, L. I., Voorn-Brouwer, T. M., van Roomen, C. P., Ottenhoff, R., Boot, R. G., van den Elst, H., Walvoort, M. T., Florea, B. I., Codee, J. D., van der Marel, G. A., Aerts, J. M., and Overkleeft, H. S. (2012) Novel activity-based probes for broad-spectrum profiling of retaining beta-exoglucosidases in situ and in vivo, *Angew Chem Int Ed Engl 51*, 12529-12533.

32. Chang, J. T., Ciocca, M. L., Kinjyo, I., Palanivel, V. R., McClurkin, C. E., De Jong, C. S., Mooney, E. C., Kim, J. S., Steinel, N. C., Oliaro, J., Yin, C. C., Florea, B. I., Overkleeft, H. S.,

Berg, L. J., Russell, S. M., Koretzky, G. A., Jordan, M. S., and Reiner, S. L. (2011) Asymmetric Proteasome Segregation as a Mechanism for Unequal Partitioning of the Transcription Factor T-bet during T Lymphocyte Division, *Immunity 34*, 492-504.

33. Adibekian, A., Martin, B. R., Wang, C., Hsu, K. L., Bachovchin, D. A., Niessen, S., Hoover, H., and Cravatt, B. F. (2011) Click-generated triazole ureas as ultrapotent in vivo-active serine hydrolase inhibitors, *Nat Chem Biol 7*, 469-478.

34. Weerapana, E., Speers, A. E., and Cravatt, B. F. (2007) Tandem orthogonal proteolysis-activity-based protein profiling (TOP-ABPP) - a general method for mapping sites of probe modification in proteomes, *Nat Protoc 2*, 1414-1425.

35. Lin, G., Li, D., de Carvalho, L. P., Deng, H., Tao, H., Vogt, G., Wu, K., Schneider, J., Chidawanyika, T., Warren, J. D., Li, H., and Nathan, C. (2009) Inhibitors selective for mycobacterial versus human proteasomes, *Nature 461*, 621-626.

36. Gu, C., Kolodziejek, I., Misas-Villamil, J., Shindo, T., Colby, T., Verdoes, M., Richau, K. H., Schmidt, J., Overkleeft, H. S., and van der Hoorn, R. A. (2010) Proteasome activity profiling: a simple, robust and versatile method revealing subunit-selective inhibitors and cytoplasmic, defense-induced proteasome activities, *Plant J 62*, 160-170.

37. van der Linden, W. A., Li, N., Hoogendoorn, S., Ruben, M., Verdoes, M., Guo, J., Boons, G. J., van der Marel, G. A., Florea, B. I., and Overkleeft, H. S. (2012) Two-step bioorthogonal activity-based proteasome profiling using copper-free click reagents: A comparative study, *Bioorgan Med Chem 20*, 662-666.

38. Wessel, D., and Flugge, U. I. (1984) A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids, *Anal Biochem 138*, 141-143.

39. Shevchenko, A., Tomas, H., Havlis, J., Olsen, J. V., and Mann, M. (2006) In-gel digestion for mass spectrometric characterization of proteins and proteomes, *Nat Protoc 1*, 2856-2860.

40. Rappsilber, J., Mann, M., and Ishihama, Y. (2007) Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips, *Nat Protoc 2*, 1896-1906.

41. Olsen, J. V., de Godoy, L. M., Li, G., Macek, B., Mortensen, P., Pesch, R., Makarov, A., Lange, O., Horning, S., and Mann, M. (2005) Parts per million mass accuracy on an Orbitrap mass spectrometer via lock mass injection into a C-trap, *Mol Cell Proteomics 4*, 2010-2021.

4

Copper-catalyzed Huisgen's azide-alkyne cycloaddition and Staudinger-Bertozzi reaction in a two-step activity-based proteasome profiling experiment,

a comparative study

4.1 Introduction

Activity-based protein profiling (ABPP) is a chemically promoted method that allows researchers to label active enzymes selectively in a complex biological system, such as full cell protein extracts, living cells, or even animals(*1*). To perform an ABPP experiment, chemical tools, named activity-based probes (ABP), which react specifically with the active sites of the targeted enzymes are necessary. A typical ABP consists of a reactive group (warhead), a recognition site (linker), and a reporter group (tag). Normally, the warhead is an electrophilic trap, which reacts with the active site amino acid of the enzyme resulting in a covalent and irreversible bond. The linker is usually a chemical motif which is recognized by the targeted enzymes as their endogenous substrate. The tag can either be a fluorophore or a biotin group, to visualize the enzymes on gel or on immunoblots(*2*).

In some cases, the reporter group can influence the properties of the probe, such as the potency and selectivity of the probe or its cell permeability. To avoid these effects, a two-step ABPP strategy can be applied, which entails incorporation of a bioorthogonal ligation handle into the ABP (*3, 4*). By the use of this strategy, the tags are installed after labeling of the enzymes. There were several successful two-step ABPP examples published, since the pioneering reports using Staudinger-Bertozzi reaction and copper-catalyzed Huisgen's azide-alkyne cycloaddition (also known as click chemistry) for ABPP of proteasome and glutathione S-transferases, respectively(*5, 6*).

The azide group is commonly used for two step ABPP experiments. However, when a secondary azide is installed in a proteasome ABP, inefficient bioorthogonal ligation by biotin-phosphine **5** was observed in native cell lysate, possibly due to steric effects. To address this problem, comparison of click chemistry and Staudinger-Bertozzi reaction to install the biotin tag on the protein-probe complexes was performed, using the proteasome as the model enzyme on both primary and secondary azide groups, under both native and denatured lysate conditions.

4.2 Results and discussion

The chemical tools used in this study are shown in Figure 1. To compare the efficiency of bioorthogonal ligations on primary and secondary azides, comparative ABPP experiments were performed using proteasome probes equipped with both types of azide, respectively. ABP **1** has a primary azide, while **2** has a more sterically hindered secondary azide on the N terminus. A competitive ABPP experiment was performed in HEK293T lysate. Active proteasome subunits were first labeled with either ABP**1**, **2** or **3**, and the residual proteasome activity was labeled with fluorescent proteasome probe **4**(*7, 8*). The results are shown in Figure 2. It shows that, both **1** and **2** prefer β5 and β2 over β1. Both inhibitors completely block β5 and β2 at the concentration of about 100μM, however, β1 is still partially active. As expected, **3** shows great potency of inhibition. At 5 μM concentration of **3**, no residual activity could be labeled by **4**.

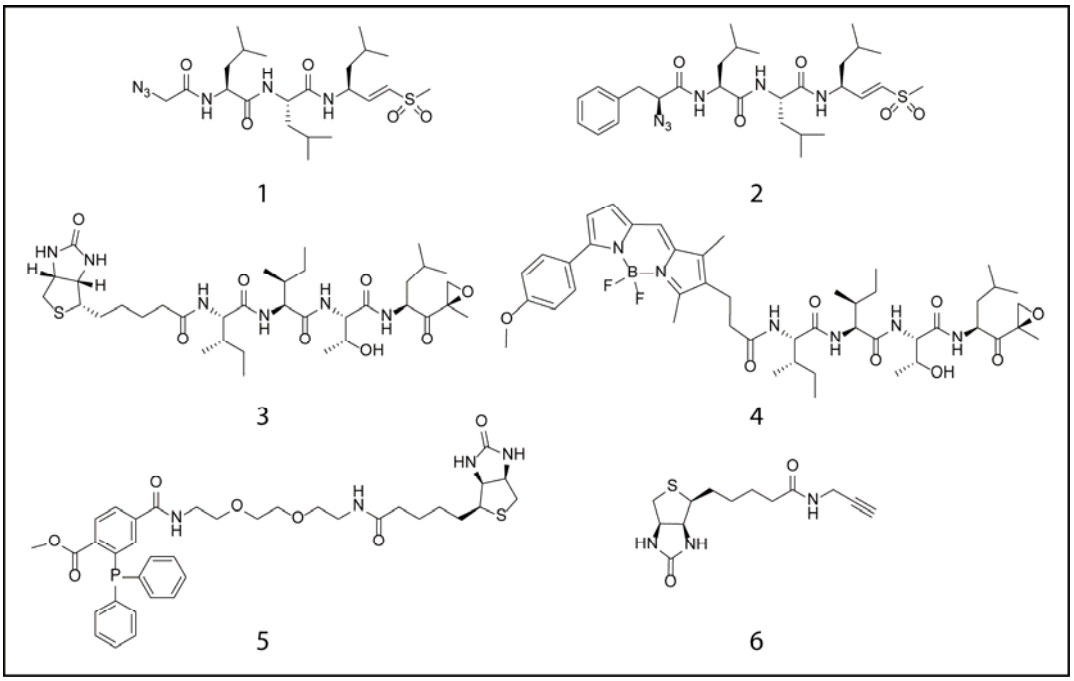


Figure 1: Structures of the chemical tools in this study.

According to these results, 100 μM of **1** and **2** were used to compare the bioorthogonal ligation efficiency. Alongside, 10 μM of **3** was chosen as a positive control to show the fully biotin labeling of the active proteasome subunits. However, the free **1** and **2** probes might label more active proteasome subunits during the time necessary for the bioorthogonal ligations. This extra labeling with **1** and **2** could change the results of the two-step ABPP experiments. To avoid that, all the residual active proteasome subunits were blocked by the potent fluorescent probe **4** after proteasome labeling by **1**, **2** and **3**, before the bioorthogonal ligations (Fig 3). This can also be seen as a control step showing the amount of active proteasome subunits that were labeled by **1** and **2**. While **4** shows the

same amount of competitive proteasome labeling against **1** and **2**, the difference shown in the anti-biotin western blot is due to the different efficiency of the bioorthogonal ligations. Following the control step, the biotin tag was installed via either click chemistry or Staudinger-Bertozzi reaction on the azide labeled proteins, with or without denaturing the proteins. Proteins were separated on SDS-PAGE, in-gel fluorescent imaging and anti-biotin chemiluminescent imaging were performed subsequently.

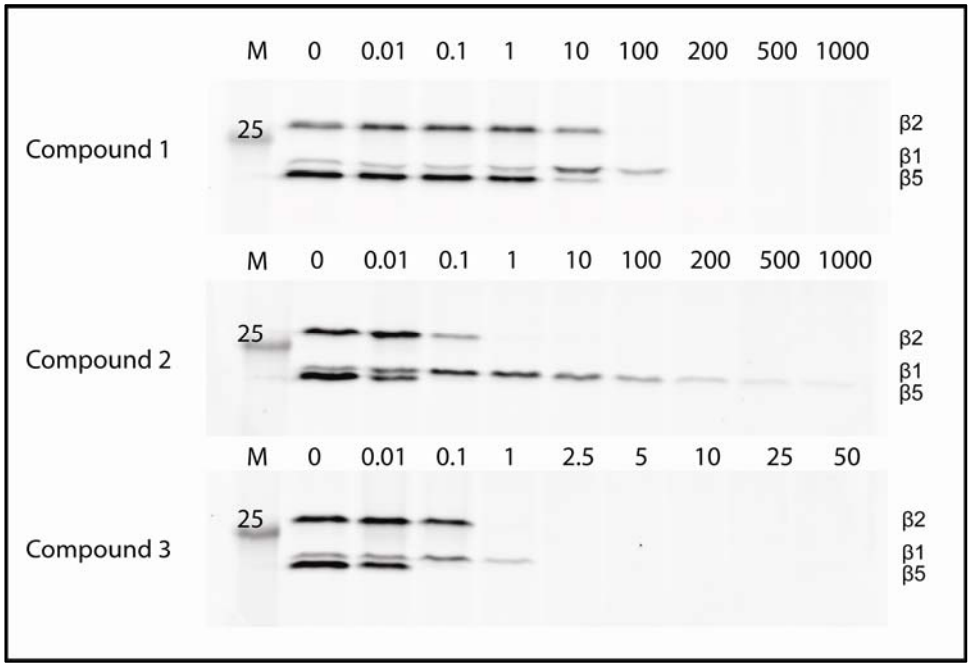


Figure 2: Determination of the inhibitor concentrations for the later experiments, by Competitive ABPP of proteasome inhibitors **1**, **2**, **3** and ABP **4**. HEK293T lysate was first treated with different concentrations of **1**, **2**, or **3**, and then labeled by **4**. Proteins were separated by SDS-PAGE and visualized by fluorescent scanning.

Fluorescent imaging in Figure 4 shows that both N_3-Leu_3-VS probes **2** (lanes B and E) and **1** (lanes C and F) blocked the proteasome activity almost completely displayed by disappearance of the bands of **4**. This indicates the difference shown on the anti-biotin blot is due to the different efficiency of bioorthogonal ligations. In the native condition labeling experiment (upper panel, Fig 4), both biotin-alkyne **6** (lane C) and biotin-phosphine **5** (lane F) could label the primary azide of **1**. With respect to the secondary azide **2**, neither **6** (lane B) nor **5** (lane E) showed very promising labeling. Click chemistry mediated ligation seemed to label the proteasome subunits more efficiently than Staudinger-Bertozzi reaction, comparing lanes B to E and C to F. However, **6** (lane D) showed more background labeling than **5** (lane G), where only bioorthogonal reagents but no azide equipped probes were added. Considering the heavier background labeling by **6**, it is more difficult to access the proteasome labeling, compared to labeling by **5**.

In the denatured condition labeling experiment (lower panel, Fig 4), both biotin-alkyne **6** (lanes B) and biotin-phosphine **5** (lanes E) labeled secondary azide efficiently, which is the most significant difference from the native condition experiment (upper panel, Fig 4). This is perhaps because under native conditions, the secondary azide group attached to the N terminus of the proteasome subunit is confined in the relatively small binding pockets and less accessible for the bioorthogonal ligation reagents. Denaturation of the proteins removed the folding of the proteasome subunits and made the azide groups more accessible. Another unexpected finding is that biotin-alkyne **6** labeled less background proteins under denatured condition (lanes B and C in lower panel) than native condition (lanes B and C in higher panel). A possible explanation is that active cysteins lost their reactivity to alkyne group after protein denaturation(9).

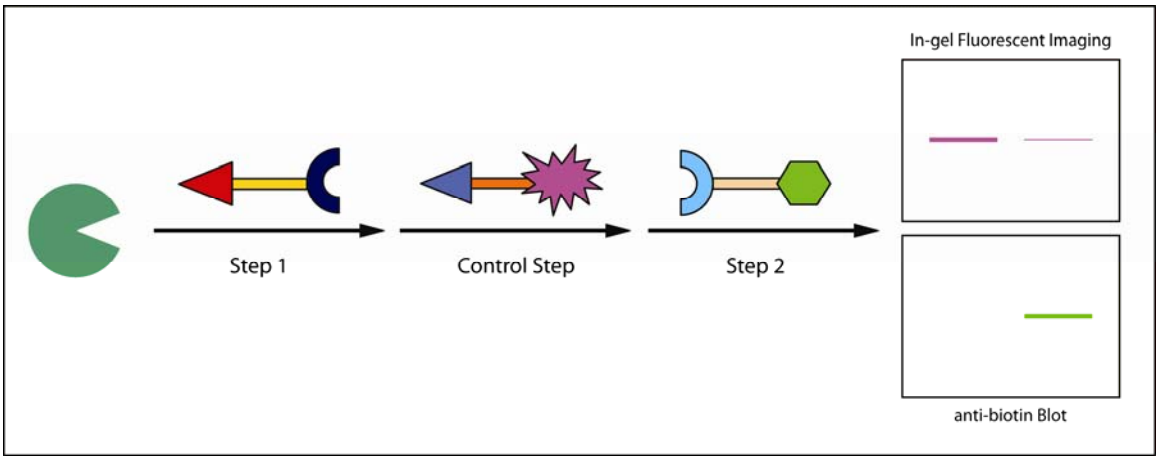


Figure 3: Scheme of the three-step ABPP experiment. Step 1, labeling of the proteasome by azide equipped inhibitors. Control step, labeling the remained proteasome activity by fluorescent probe. Step 2, bioorthorgonal ligations to install the biotin tag. Visualize the differently labeled proteasome by both in-gel fluorescent imaging and anti-biotin blot.

4.3 Conclusion

In conclusion, it might be an additional effect of the steric-effect of the secondary azide and the folded proteasome subunits surrounding the azide with their binding pockets, which will together prevent it from reacting with the bioorthogonal ligation reagents under native condition. This also explains why better labeling efficiency is achieved under denaturing conditions, in which all the proteins have been unfolded by boiling with SDS followed by a chloroform/methanol precipitation. Secondly, labeling of azide proteins seems to be more efficient using biotin-alkyne **6** than biotin-phosphine **5**, with the same concentration of ligation reagents and reaction time. However, it shows more background labeling by performing click chemistry, which might be due to the possible reactions between alkyne group and active cysteins(9). This has to be considered, when one wants to

combine the two-step ABPP with a pull down experiment for affinity purification of the target enzymes, because the background proteins covalently modified by biotin-alkyne **6** will not be washed away after binding to the streptavidin beads. In that case, the Staudinger-Bertozzi reaction, through which it labels less background protein, might be an advisable choice. Finally, bioorthogonal ligations on proteasome probes equipped with secondary azide can be efficiently processed under denatured conditions. For the ligation on primary azide, both conditions are suitable. It suggests that while making azide equipped proteasome probes, one should consider which type of azide is needed for the downstream visualization work.

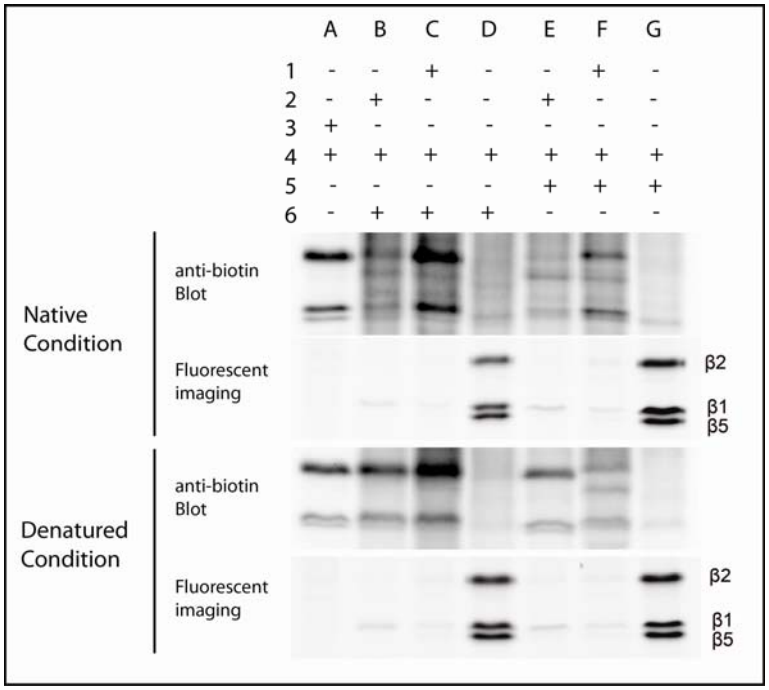


Figure 4: anti-biotin blots and in-gel fluorescent imaging to visualize the proteasome labeling patterns. The bioorthogonal ligations were performed on either native proteins (shown in upper panels) or denatured proteins (lower panels).

4.4 Experimental procedures

4.4.1 Synthesis of the chemical tools

Dichloromethane (DCM), N,N-dimethylformamide (DMF), methanol (MeOH), diisopropylethylamine (DiPEA) and trifluoroacetic acid (TFA) were of peptide synthesis grade, purchased at Biosolve, and used as received. All general chemicals (Fluka, Acros, Merck, Aldrich, Sigma) were used as received. Column chromatography was performed on Screening Devices b.v. Silica Gel, with a particle size of 40-63 μm and pore diameter of 60 Å. TLC analysis was conducted on Merck aluminium sheets (Silica gel 60 F254). Compounds were visualized by UV absorption (254 nm), by spraying with a solution of $(NH_4)_6Mo_7O_{24}\cdot 4H_2O$ (25 g/L) and $(NH_4)_4Ce(SO_4)_4\cdot 2H_2O$ (10 g/L) in 10% sulfuric acid, a

solution of $KMnO_4$ (20 g/L) and K_2CO_3 (10 g/L) in water, or ninhydrin (0.75 g/L) and acetic acid (12.5 mL/L) in ethanol, where appropriate, followed by charring at ca. 150 oC. ^{1}H- and ^{13}C-NMR spectra were recorded on a Bruker AV-400 (400 MHz) spectrometer. Chemical shifts are given in ppm (δ) relative to CD_3OD as internal standard. High resolution mass spectra were recorded by direct injection (2 µL of a 2 µM solution in water/acetonitrile 50/50 (v/v) and 0.1% formic acid) on a mass spectrometer (Thermo Finnigan LTQ Orbitrap) equipped with an electrospray ion source in positive mode (source voltage 3.5 kV, sheath gas flow 10, capillary temperature 250 °C) with resolution R = 60,000 at m/z 400 (mass range m/z = 150-2,000) and dioctylpthalate (m/z = 391.28428) as a "lock mass". The high resolution mass spectrometer was calibrated prior to measurements with a calibration mixture (Thermo Finnigan). . $BocLeu_3VS$ and Tris-[1-(3-hydroxypropyl)-1H-[1,2,3]triazol-4-yl)methyl]amine (MDW999, ligand for CuAAC) were synthesized according to literature procedures(*10, 11*).

GB123 **1** and WLB426A **2** were synthesized from $BocLeu_3VS$. To synthesize **1**, $BocLeu_3VS$ (110 mg, 213 µmol) was dissolved in 1:1 TFA/DCM (2 mL). After stirring for 30 min., the reaction mixture was concentrated and co-evaporated with toluene (3x), providing the free amine as TFA salt, which was directly used in the next step. To a solution of the TFA salt in DMF was added DiPEA (44.57 µL, 255 µmol) followed by the addition of N_3Gly-OSu (51 mg, 255 µmol). After 3 hours, TLC showed completion of the reaction and the reaction mixture was diluted with EtOAc, washed with 1N HCl (1x), sat. $NaHCO_3$ (2x) and brine (1x). The organic layer was dried over $NaSO_4$ and concentrated. The product was precipitated from DCM, filtered off and isolated as a white solid (49 mg, 46%). ^{1}H NMR (400 MHz, Methanol-d_4): δ ^{1}H NMR (400 MHz, Methanol-d4) δ 6.77 (dd, J = 15.1, 5.1 Hz, 1H), 6.47 (dd, J = 15.1, 1.6 Hz, 1H), 4.78 – 4.48 (m, 1H), 4.48 – 4.19 (m, 2H), 3.89 (s, 2H), 2.93 (s, 3H), 1.69 – 1.10 (m, 9H), 1.10 – 0.62 (m, 18H). ^{13}C NMR (101 MHz, MeOD) δ 173.07, 172.70, 168.63, 148.05, 129.36, 52.40, 52.37, 48.15, 42.77, 42.65, 41.09, 40.51, 25.10, 25.03, 25.00, 22.96, 21.87, 21.81. LC-MS: Rt (min): 8.08 (linear gradien 10-90% in 15 min, m/z 501.13 [M+H+]. HRMS: calcd. for $C_{22}H_{40}N_6O_5S$ 501.38537 $[M+H]^+$; found 501.28531.

For synthesis of **2**, $BocLeu_3VS$ (46 mg, 89 µmol) was dissolved in 4N HCl in dioxane (1 mL). After stirring for 40 min., the reaction mixture was concentrated and co-evaporated with toluene (3x), providing the free amine as HCl salt, which was directly used in the next step. To a solution of the HCl salt in DMF was added N_3PheOH (288 µL of a 0.34 M sol. in DMF, 98 µL), HBTU (41 mg, 107 µmol) and DiPEA (51 µL, 312 µmol). After 2 hours, TLC showed completion of the reaction. The reaction mixture was diluted with DCM and washed with 1N HCl (2x), sat. $NaHCO_3$ (4x). The organic layer was dried over $NaSO_4$ and concentrated. Purification by column chromatography (0.5—1% MeOH in DCM) provided the product as a white solid.(47 mg, 89%). ^{1}H NMR (400 MHz, Methanol-d_4) δ 7.35 – 7.21 (m, 5H), 6.80 (dd, J = 15.2, 5.2 Hz, 1H), 6.60 (dd, J = 15.2, 1.3 Hz, 1H), 4.67 (dt, J = 9.2, 5.0 Hz, 1H),

4.44 – 4.34 (m, 2H), 4.16 (dd, J = 8.5, 4.9 Hz, 1H), 3.20 (dd, J = 14.0, 4.9 Hz, 1H), 3.04 – 2.93 (m, 4H), 1.73 – 1.40 (m, 9H), 1.03 – 0.87 (m, 18H). ^{13}C NMR (101 MHz, MeOD) δ 174.26, 174.18, 171.69, 148.47, 137.78, 130.77, 130.41, 130.29, 129.59, 128.04, 65.42, 53.39, 53.34, 49.14, 43.28, 42.80, 41.80, 41.57, 38.62, 25.93, 25.78, 23.43, 23.39, 23.34, 22.12, 22.09, 21.91. HRMS: calcd. for $C_{22}H_{40}N_6O_5S$ 591.33232 $[M+H]^+$; found 591.33246.

Compounds **3, 4, 5** and **6** were synthesized through literature procedures(*5, 7, 12*).

4.4.2 Competitive activity-based proteasome profiling

HEK293T cells (derived from ATCC) were grown in DMEM supplemented with 10% fetal calf serum and 10mg/ml penicillin and streptomycin in a humid CO_2 (5%) incubator at 37°C. At 90% confluence, the cells were harvested by scrapping in pre-chilled PBS. The cell pellet was lysed with a mild lysis buffer containing Tris pH 7.5 (50 mM), sucrose (250 mM), $MgCl_2$ (5 mM), dithiothreitol (DTT; 1 mM), ATP (2mM), digitonin (0.025%)(*13*).

The protein concentration was determined by Qubit protein assay (Invitrogen). 20µg of lysate was first incubated with different concentration of **1, 2** or **3** at 37°C for 1 hour, and then labeled with **4** for 1 hour at 37°C. Subsequently, the samples were boiled at 95°C for 5min after addition of 3x sample buffer (6% SDS, 3% β-mercaptomethanol, 30% v/v Glycerol, 0.1% Bromophenol Blue, 150mM Tris pH 6.8).

The boiled samples were separated on a 12.5% SDS-PAGE, and then imaged with a ChemiDoc MP system (BioRad). The gels were further stained by coomassie blue, and imaged and quantified as loading controls.

4.4.3 Two step activity-based proteasome profiling

100µg of HEK293T lysate was incubated with **1, 2** or **3**, at 37°C for 1 hour. Afterwards, fluorescent probe **4** was added in label the remained proteasome activity. The mixture was again incubated at 37°C for 1 hour. For click chemistry under native condition, the mixture was first diluted by lysis buffer to 50µl, then the same volume of click cocktail (10mM CuSO4, 10mM MDW999, 10mM Sodium Ascorbate and 0.8mM **6** in 50mM Tris pH 8.0) was added. For Staudinger-Bortozzi reaction under native condition, the mixture volume after labeling with **4** was taken up by lysis buffer to 100µl while adding 0.4mM **5**.

For performing click chemistry and Staudinger-Bortozzi reactions under denatured condition, the lysate after **4** labeling was denatured by boiling with 1% SDS at 95°C for 5 min followed by a chloroform/methanol precipitation (C/M (*13*)). The pellet was dissolved by 8M Urea in 50mM Tris pH8.0, and then the protein solutions can be processed as described above in the native reaction protocols.

All of the reaction mixtures were incubated at 37°C for 1 hour. C/M precipitation was performed to quench the reaction and remove the excess biotin. All the samples were dissolved in 3x sample buffer and boiled at 95°C for 5 min. The boiled samples were

separated on SDS-PAGE and transferred to western blot. The labeled proteasome subunits were detected by either fluorescent scanning or chemiluminascent scanning after probing the blot with Streptavidin-HRP.

Acknowlegement

We thank Gerjan de Bruin and Wouter van der Linden for the synthesis of proteasome inhibitors, and Martin de Witte for the synthesis of click ligand.

References and notes

1. Li, N., Overkleeft, H. S., and Florea, B. I. (2012) Activity-based protein profiling: an enabling technology in chemical biology research, *Curr Opin Chem Biol 16*, 227-233.

2. Cravatt, B. F., Wright, A. T., and Kozarich, J. W. (2008) Activity-based protein profiling: from enzyme chemistry to proteomic chemistry, *Annu Rev Biochem 77*, 383-414.

3. Willems, L. I., Li, N., Florea, B. I., Ruben, M., van der Marel, G. A., and Overkleeft, H. S. (2012) Triple bioorthogonal ligation strategy for simultaneous labeling of multiple enzymatic activities, *Angew Chem Int Ed Engl 51*, 4431-4434.

4. Willems, L. I., van der Linden, W. A., Li, N., Li, K. Y., Liu, N., Hoogendoorn, S., van der Marel, G. A., Florea, B. I., and Overkleeft, H. S. (2011) Bioorthogonal chemistry: applications in activity-based protein profiling, *Acc Chem Res 44*, 718-729.

5. Ovaa, H., van Swieten, P. F., Kessler, B. M., Leeuwenburgh, M. A., Fiebiger, E., van den Nieuwendijk, A. M., Galardy, P. J., van der Marel, G. A., Ploegh, H. L., and Overkleeft, H. S. (2003) Chemistry in living cells: detection of active proteasomes by a two-step labeling strategy, *Angew Chem Int Ed Engl 42*, 3626-3629.

6. Speers, A. E., Adam, G. C., and Cravatt, B. F. (2003) Activity-based protein profiling in vivo using a copper(I)-catalyzed azide-alkyne [3+2] cycloaddition, *J Am Chem Soc 125*, 4686-4687.

7. Florea, B. I., Verdoes, M., Li, N., van der Linden, W. A., Geurink, P. P., van den Elst, H., Hofmann, T., de Ru, A., van Veelen, P. A., Tanaka, K., Sasaki, K., Murata, S., den Dulk, H., Brouwer, J., Ossendorp, F. A., Kisselev, A. F., and Overkleeft, H. S. (2010) Activity-based profiling reveals reactivity of the murine thymoproteasome-specific subunit beta5t, *Chem Biol 17*, 795-801.

8. Meng, L., Mohan, R., Kwok, B. H., Elofsson, M., Sin, N., and Crews, C. M. (1999) Epoxomicin, a potent and selective proteasome inhibitor, exhibits in vivo antiinflammatory activity, *Proc Natl Acad Sci U S A 96*, 10403-10408.

9. Ekkebus, R., van Kasteren, S. I., Kulathu, Y., Scholten, A., Berlin, I., Geurink, P. P., de Jong, A., Goerdayal, S., Neefjes, J., Heck, A. J., Komander, D., and Ovaa, H. (2013) On terminal alkynes that can react with active-site cysteine nucleophiles in proteases, *J Am Chem Soc 135*, 2867-2870.

10. Bogyo, M., McMaster, J. S., Gaczynska, M., Tortorella, D., Goldberg, A. L., and Ploegh, H. (1997) Covalent modification of the active site threonine of proteasomal β subunits and the

Escherichia coli homolog HslV by a new class of inhibitors, *Proceedings of the National Academy of Sciences 94*, 6629-6634.

11. Peacock, H., Maydanovych, O., and Beal, P. A. (2010) N2-Modified 2-aminopurine ribonucleosides as minor-groove-modulating adenosine replacements in duplex RNA, *Organic Letters 12*, 1044-1047.

12. Meier, J. L., Mercer, A. C., Rivera, H., Jr., and Burkart, M. D. (2006) Synthesis and evaluation of bioorthogonal pantetheine analogues for in vivo protein modification, *J Am Chem Soc 128*, 12174-12184.

13. Li, N., Kuo, C. L., Paniagua, G., van den Elst, H., Verdoes, M., Willems, L. I., van der Linden, W. A., Ruben, M., van Genderen, E., Gubbens, J., van Wezel, G. P., Overkleeft, H. S., and Florea, B. I. (2013) Relative quantification of proteasome activity by activity-based protein profiling and LC-MS/MS, *Nat Protoc 8*, 1155-1168.

5

Combination of activity-based proteasome profiling and global proteomics to elucidate mechanism of bortezomib resistance

5.1 Introduction

In eukaryotic cells the ubiquitin proteasome system (UPS) plays a central role in the cellular protein turnover(*1*). Proteins that are at the end of their life cycle or wrongly folded are modified by a small protein marker (around 8KDa) called ubiquitin. The ubiquitin ligases can install more ubiquitins on the former ones to form a polyubiquitin chain(*2*). The polyubiquitinylated proteins are delivered to the proteasome for proteolytic degradation, while the ubiquitin molecules are recycled for modification of new protein substrates.

The 26S proteasome that is evolutionarily conserved from yeast to mammals is a large protease cluster (2.5 MDa) (*3*). Structurally, the catalytically active proteasome involves two types of main components, the 20S core particle and the 19S or 11S regulatory particles. The 20S particle (Fig 1A) contains 4 heptamerical rings containing either subunits α1-α7 (outer layers) or β1-β7 (inner layers)(*4*). Among these subunits, three subunits are catalytically active for hydrolysis of peptide bonds, β1 cleaving C-terminal of acidic amino acid residues, β2 cleaving after basic amino acid residues, and β5 cleaving after hydrophobic residues. The regulatory particles are responsible for the recognition of substrates, removal of polyubiquitin chains, linearization of the polypeptides, and opening of the gates of the outer α rings of the 20S particles(*4*).

The proteasome is responsible for degradation of over 90% of the cellular proteins and involved in several important biological processes, such as protein turnover, cell cycle progression, transcription, and DNA repair(*1*). In higher organisms that have immune systems, the proteasome is crucial for the production of MHC I (major histocompatibility complex class I) epitopes for antigen presentation(*5, 6*). In immune cells, three extra catalytically active subunits β1i, β2i, β5i are expressed that replace the β1, β2, β5 subunits in the constitutive proteasome to form the immuno-proteasome. The immuno β subunits

show comparable substrate cleavage preference and share around 50% protein sequence identity compared to the constitutive β1, β2 and β5.

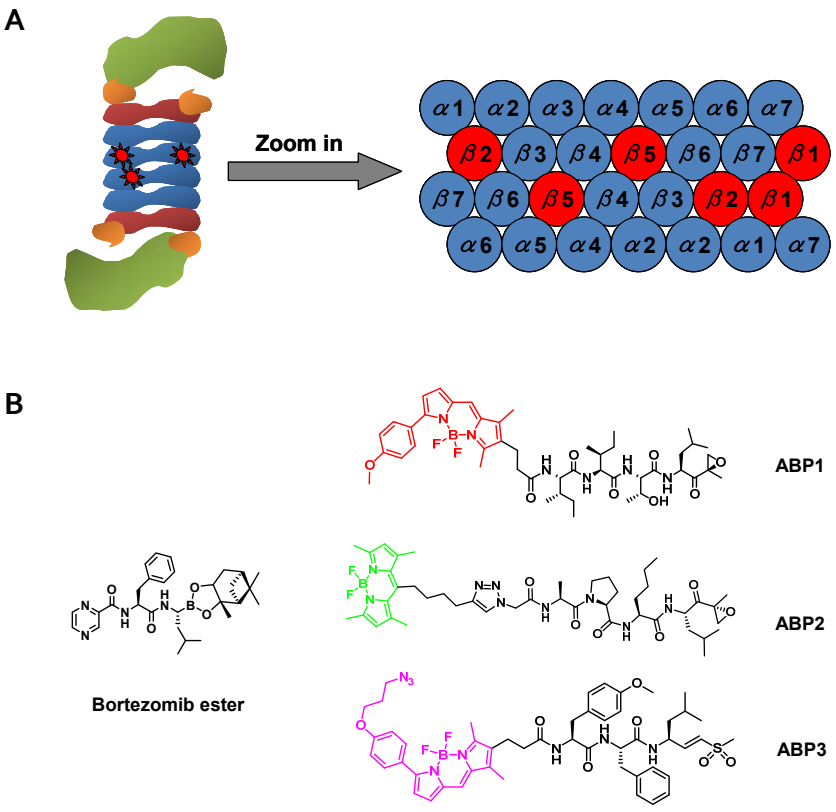


Figure 1: structures of Proteasome, Bortezomib and activity-based proteasome probes.
(A) Scheme of 30S proteasome is shown on the left. A cut view of the 20S proteasome core particle is present on the right. Subunits indicated in red are the active subunits.
(B) Structures of proteasome inhibitor Bortezomib and Activity-based proteasome probes. ABP**1** is broad spectrum probe MVB003. ABP**2** is β1/1i specific probe LW124. ABP**3** is β5/5i specific probe MVB127.

Proteasome inhibitor bortezomib (Velcade®, Fig 1B) targeting proteasome active subunits β1, β1i, β5 and β5i has been approved as an anti-tumor drug for treatment of multiple myeloma about ten years ago(*7*). Multiple myeloma is a malignancy of plasma cells that are B cells producing multiple antibodies for the immunity, happening in the bone marrow. When plasma cells transform into multiple myeloma, they will continue secreting one single type of monoclonal antibody, which is named M-protein. Clinically, the level of M-protein is used as one of the biomarkers for the early diagnosis of multiple myeloma. During the pathogenesis of the disease, multiple myeloma cells degrade bone by interfering with the crucial process of bone remodelling in bone marrow leading to bone lesion in the later stage of the disease(*8*).

Bortezomib can extend the lives of myeloma patients significantly. Pharmacologists and clinicians have been attempting to elucidate the molecular mechanism of the anti-cancer effect of proteasome inhibition through different experiments. A number of theories were put forward. For instance, multiple myeloma cells are B cells that work as antibody factories and thus have a high rate of protein synthesis and degradation, so proteasome inhibition induces apoptosis in myeloma cells. Proteasome inhibition blocks the degradation of IKB, which is the inhibitor of the transcription factor NFKB essential for cancer progression in B cells. Thus, the inactivation of NFKB will lead to apoptosis of the cancer cells, for they need NFKB to initiate the transcription of key proteins for cell proliferation(*9, 10*). Additionally, accumulation of proteasome substrates induces unfolded protein response (UPR) and ER stress in the cells, which if remained unresolved also results in apoptosis(*11, 12*). On the other hand, it is suggested that proteasome inhibition also causes shortage of several essential amino acid from the blockage of proteasomal degradation of proteins, which can direct the cells to death(*13*).

Many patients develop bortezomib resistance after a period of treatment. A series of point mutations were found by sequencing of the β5 cDNA of bortezomib adapted myeloma or other cancer cell lines(*14, 15*). However, no meaning point mutations were found in patients who developed bortezomib resistance(*16-18*). A few genes were found to be up or down-regulated by gene expression profiling (GEP) in bortezomib resistant cell lines or patient samples(*19-21*). Because, the messenger RNA levels do not correlate perfectly to the protein levels in many cases, one might argue that a quantitative proteomics analysis of protein levels would be more accurate(*22*). Quantitative phosphoproteomics revealed that in U266 myeloma cells, bortezomib treatment increased the Ser38 phosphorylation level of the protein stathmin, which is a key controller of cell cycle and proliferation. S38A point mutation of stathmin in the tested multiple myeloma cells increased their resistance to bortezomib treatment(*23*). The proteasome inhibition by MG132 or bortezomib resulted in minor change of protein expression and significant change in protein ubiquitinylation in Jurkat (T cell lymphoma) or HCT116 (colon carcinoma) cells(*24, 25*). A quantitative proteomic comparison of the clinical relevant cancer cells and their bortezomib adapted derivatives is missing(*26*).

Another aspect to observe the effect of bortezomib treatment is to determine the residual activity of the proteasome after treatment(*27*). That might also indicate why some cancer cells are sensitive to the drug whereas others are not. Since most clinical relevant cells are lymphocytes, they contain a mixture of proteasome subtypes, constitutive-proteasome, immuno-proteasome and intermediate-proteasome(*5, 6, 28*). This complexity makes the analysis of proteasome activity more difficult, because the classical tool of fluorogenic substrates can not distinguish the constitutive subunits and immuno subunits due to their similar cleavage preference(*29*). To determine the residual proteasome

activity, activity-based proteasome profiling can be applied, which allows relative quantification of the activity of each subunit in a high throughput mode(*30, 31*). Activity-based protein profiling (ABPP) is a chemical biology method to label and visualize simultaneous active enzymes selectively in a complicated biological system, using activity-based probes (ABP)(*32, 33*).

Chapter 3 in this thesis is a relative quantitative ABPP method, with which one can quantify the activity of each active proteasome subunit using fluorescent proteasome ABPs(*31*). Cells were treated with different concentrations of bortezomib and after lysis, the residual activity of the proteasome subunits were labeled by fluorescent ABPs that are either broad-spectrum or subunit specific, resolved by SDS-PAGE and visualized by in gel fluorescence scanning. The fluorescent signal was quantified and corrected by the loading control of the coomassie stained gel. In order to quantify the activity of each proteasome subunit with high resolution, three probes were used for the proteasome activity profiling. As shown in Figure 1B, the broad spectrum ABP MVB003 (**1**) was used for β2/2i activity profiling. Subunit specific probes LW124 (**2**) and MVB127 (**3**) labeled β1/1i and β5/5i respectively.

This chapter describes the combination of activity-based proteasome profiling and quantitative global proteomics to observe the anti-cancer effect of bortezomib in various cancer cell lines. The proteasome activity alteration was linked to the global protein expression levels and cell proliferation rates. Bortezomib sensitive leukemia cells were compared to bortezomib insensitive solid tumor cells on all these aspects before and after treatment. Sensitive myeloma and leukemia cells were also compared to their bortezomib adapted derivative cell lines.

5.2 Results

5.2.1 Sensitivity and resistance of cancer cells to bortezomib

Eight different human cancer cell lines were compared in parallel. The cancer cells were treated with different concentrations of bortezomib for 16 hours. The metabolic activity of the cells was determined by commercially available CellTiter-Glo luminescent cell viability assays (Promega). Briefly, in the assay mono-oxygenation of luciferin, which generates luminescence, is catalyzed by luciferase in the presence of Mg^{2+}, ATP and molecular oxygen. So by measuring the luminescence intensity, the total ATP amount in cells can be determined that reflects the ratio of living cells in the test well through comparison to the non-treated control.

The results are shown in Figure 2A. Cervical cancer HeLa S3 and breast cancer MCF7 cells were extremely resistant to bortezomib treatment. They survived under exposure of up to 500nM bortezomib, which is more than 20 times of the clinical concentration for multiple myeloma treatment, with no clear alteration of the metabolic activity. Lung cancer

cell line A549 was slightly more sensitive to bortezomib than HeLa S3 and MCF7, but still showed 70% survival at 500nM concentration. However, the ovarian carcinoma OvCar3 showed significant bortezomib sensitivity. These cells appear almost as sensitive as the HL60 cells, which is an acute myeloid leukemia cell line. At the bortezomib concentration of 100nM, they both showed 50% survival ratio. It was to be expected that, the clinical relevant myeloma cell lines AMO1 and RPMI8226 and mantle cell lymphoma Granta519 were all extremely sensitive to bortezomib. Especially for AMO1 and Granta519, when treated with 20nM, which is approximately the clinical concentration of bortezomib, only 20% metabolic activity remained and for RPMI8226 some 30%.

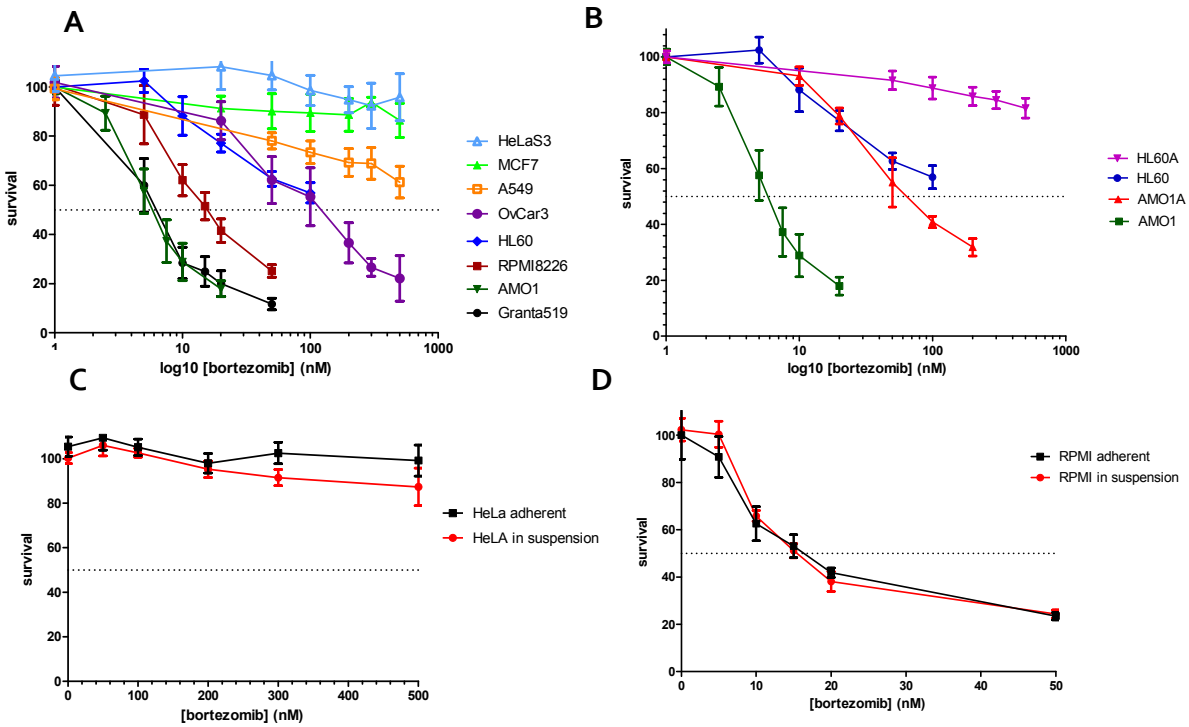


Figure 2: Cell proliferation assays after bortezomib treatment (for all the assays, n≥5).
(A) Eight different cell lines were compared in the assay, four solid tumor cell lines: HeLa S3 (cervical cancer), MCF7 (breast cancer), A549 (lung cancer), OvCar3 (ovarian cancer) and four lymphoblast cell lines: HL60 (acute myeloid leukemia), RPMI8226 (multiple myeloma), AMO1 (plasmacytoma), Granta519 (mantle cell lymphoma).
(B) Wild type HL60 and AMO1 were compared with their bortezomib adapted derivative cell lines HL60A and AMO1A on cell proliferation ratio after bortezomib treatment. The adapted cells were ten times more resistant to bortezomib than their parental cells.
(C) and (D) HeLa S3 and RPMI8226 cells in suspension or adhesion were compared on proliferation ratio. Cells under different growth conditions showed similar proliferation after bortezomib treatment.

To mimic the situation of the tumor cells from bortezomib-resistant patients, the bortezomib adapted cell lines of HL60 and AMO1 were generated as described in a previous report(*14*). The resistant cells were compared with their wild type parental cell lines under bortezomib exposure in the metabolic activity assay (Fig 2B). HL60A was significantly more resistant to the treatment than the parental cell line, which showed 60% metabolic activity at 50nM bortezomib concentration. The HL60A cells were almost as resistant as the HeLa S3 and MCF7 cells, without clear survival decrease at 500nM bortezomib concentration. AMO1A also showed over ten times higher bortezomib resistance than its parental cell line. It showed about 30% survival ratio at 200nM concentration, while the AMO1 shows less than 20% survival at 20nM bortezomib concentration. This shows that the bortezomib adapted cell lines were well adapted and could be used for the study of the molecular mechanism of bortezomib resistant effects.

As seen from Figure 2A, the solid tumor cells are generally more resistant than the lymphoblast (leukemia, lymphoma and myeloma) cells to bortezomib. One possible explanation for this difference is the growing condition of the cells. Most solid tumor cells in culture are growing adherent, while lymphocytes are growing in suspension. This difference might influence the efficiency of the bortezomib permeation into the cells, and further change the sensitivity or resistance of the cells. To test this possibility, the cervical carcinoma HeLa S3 and multiple myeloma RPMI8226 were grown both in suspension and adherent conditions. The metabolic activity assay was performed on the cells growing under both conditions, respectively (Fig 2C and D). The results show that the same cell line growing under different conditions did not change its resistance to bortezomib, which also indicates the bortezomib permeation efficiency does not depend on the shape of the tumor cells. On the other hand, the bortezomib sensitivity might be depending on the proteasome subtypes expressed in the cells, because most solid tumor cells express only constitutive proteasome, whereas the lymphocytes express both constitutive and immuno proteasome.

5.2.2 Proteasome inhibition by bortezomib in sensitive and resistant cells

Among the solid tumor models, HeLa S3 and MCF7 cells showed similar proteasome inhibition patterns after bortezomib treatment (Fig 3A and B). The activity of the target subunit β1 decreased sharply to almost zero at clinical relevant concentration up to 100nM. The β2 activity decreased some 20% in MCF7 cells and 50% in HeLa S3 cells at 500nM bortezomib. However, for β5 the other main target of bortezomib, only 50% activity was inhibited at 500nM. The β5 subunit has more broad cleavage preference compared to β1 and β2, and might degrade the most peptides among the three active subunits. This might explain why these cells are not sensitive to bortezomib treatment. The β1 and β5 subunits in A549 and OvCar3 were both inhibited nearly completely at 100nM bortezomib (Fig 3C

and D). Interestingly, in A549 the β2 activity dropped to 50%, while in OvCar3 it was increased by 50%. Considering the results of the cell survival assays where A549 was more resistant to bortezomib than OvCar3, the conclusion might be drawn that the β2 activity itself does not reflect the fate of the cells, but plays an assistant role in the proteasomal degradation.

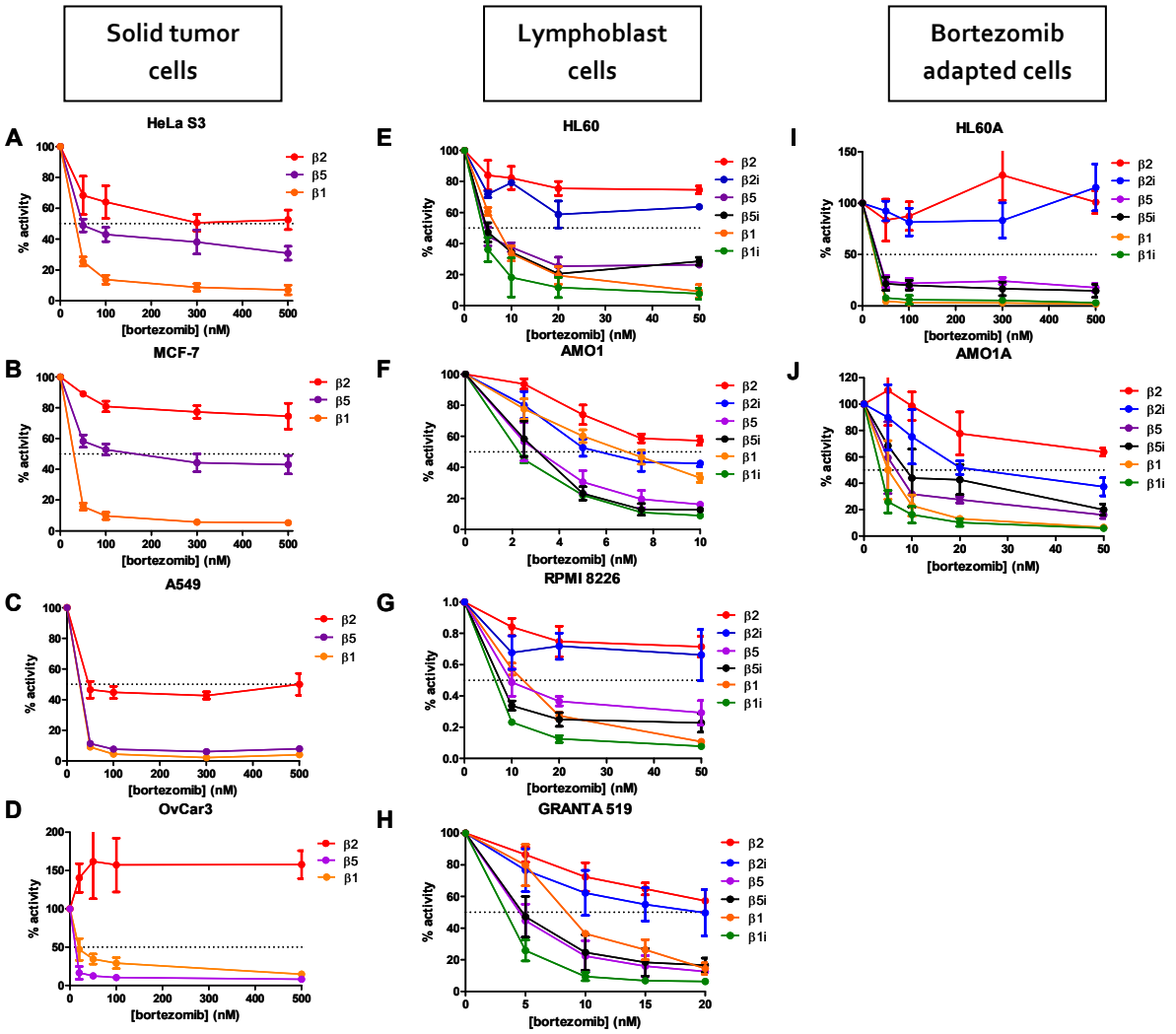


Figure 3: Proteasome activity profiling by fluorescence ABPP after bortezomib treatment of different cell lines.
(A) HeLa S3, (B) MCF7, (C) A549, (D) OvCar3, (E) HL60, (F) AMO1, (G) RPMI8226, (H) Granta519, (I) HL60A and (J) AMO1A.

In all tested lymphoblast cells (HL60, AMO1, Granta519 and RPMI8226), the target subunits involving β1, β1i, β5 and β5i, lost over half of the activity, while treated by clinical concentrations of bortezomib, up to 20nM (Fig 3E-H). Additionally, the activity of β2(i) decreased to 60-80%. This shows significant difference from the solid tumor cells except the relatively sensitive OvCar3 (Fig 3D), which might link the bortezomib inhibition pattern

to the sensitivity of the cells. Arguably, if the activity of the target subunits especially the β5(i) decreases to lower than 50% in range of clinical concentration of bortezomib (<20nM), the cell line might be bortezomib sensitive, otherwise the cells could be resistant.

To test this hypothesis, bortezomib adapted cell lines were analyzed by ABPP for proteasome activity for comparison with their parental cell lines. At 50nM bortezomib treatment, the activity of target subunits in HL60A (Fig 3I) decreased to similar levels as in HL60 (Fig 3E), for β1/1i around 10% and for β5/5i 25%. From 50 to 500nM bortezomib concentration, these subunits kept similar activity. However, these subunits might show some different activity levels with the ones in HL60, if the HL60A cells were treated with lower concentration of bortezomib as their parental cells. In AMO1A cells, when treated with 5nM bortezomib, only the β1i activity decreased to 25%, the activity of β1/5/5i remained no less than 50% (Fig 3J). This showed clear difference with the wild type AMO1 cells that, when treated with same concentration bortezomib, the activity of β1i/5/5i all decreased by 70-80%, only β1 activity was higher than 50%. Therefore, when treated with clinical concentrations of bortezomib, the resistant cells might have higher proteasome activity that helps them escaping from the apoptosis induced by proteasome inhibition.

To interrogate bortezomib resistant mechanism of the two adapted cell lines, they were compared directly with their parental cells in the same ABPP experiment (Fig 4A). Strikingly, the total signal of proteasome activity labeling by probe **1** in the untreated HL60A cells was almost double that of the amount in the HL60 cells. This indicates that when the same percentage of proteasome activity was blocked by bortezomib treatment, a much higher activity remained in the adapted cells than the wild type cells. This could be a good explanation for bortezomib resistance of HL60A. In the AMO1 and AMO1A cells, the observation is similar. AMO1A also showed higher proteasome activity than AMO1. The difference between cells was that, the proteasome activity up-regulation in HL60A was more significant than in the AMO1A, possibly because the two cell lines were generated from two different types of disease. The myeloma and acute myeloid leukemia, from which AMO1 and HL60 were generated respectively, might have different molecular mechanisms to overcome the proteasome inhibition treatment.

The question is then, whether the difference in proteasome activity between the bortezomib adapted cells and the wild type cells is related to proteasome expression levels. The total proteasome amount in different cells was compared in a native PAGE experiment(*34*), in which only the intact proteasome particles (20S, 26S and 30S) were assayed, while the free subunits would be excluded (Fig 4B). The expression of proteasome was detected by anti alpha subunits western blot and quantified and corrected by loading control (Fig 4C). The western blotting results correlated with the ABPP results (Fig 4A). The HL60A and AMO1A both expressed more proteasome than their parental cell lines. This might explain how the adapted cells become resistant to the drug. There are also

possibilities that the adapted cells have generated point mutations in the main target subunit β5 as described in several previous reports(*14*, *15*). The PSMB5 cDNA from the adapted cells and their parental cells, reversed transcribed from their total mRNA was sequenced (data not shown). No point mutations were found, which meant the cells have other ways to counterpart the proteasome inhibition induced apoptosis by bortezomib.

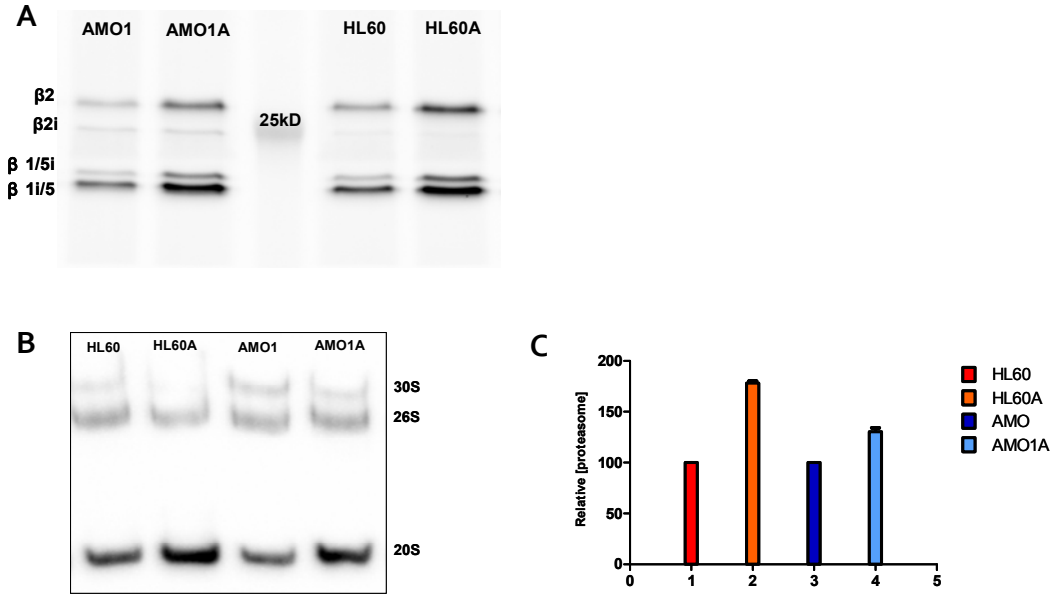


Figure 4: The bortezomib adapted HL60A and AMO1A cell lines express more proteasome than wild type HL60 and AMO1, respectively.

(A) Proteasome active subunits in non-treated HL60, HL60A, AMO1 and AMO1A were labeled by the broad spectrum probe **1**. More proteasome activity was probed in the Bortezomib adapted cells HL60A and AMO1A.

(B) Lysates from the four cell lines were resolved by a native PAGE and proteasome species were detected by a general antibody against α1-7 subunits on a western blot. More 20S proteasome was observed in HL60A and AMO1A cells.

(C) Quantification of results from (B). The proteasome species signal was added up and corrected for loading control by coomassie stained SDS-PAGE. The results of HL60A and AMO1A were normalized to the parental cells HL60 and AMO1, respectively. The quantification data is from three independent experiments.

5.2.3 Global proteomic profiling to reveal the bortezomib resistance mechanism

The bortezomib adapted cell lines show higher expression of proteasome, but no point mutation were found in the PSMB5 gene that encodes its main target subunit β5. However, whether the higher level of proteasome expression is the only reason of the bortezomib resistance was still to be determined. Previous reports have shown that there are genes up or down-regulated on the mRNA levels in the bortezomib resistant patient

samples, compared to the control samples(*19-21*). But mRNA levels are not always equal to the protein levels(*22*). So it is envisioned to have a quantitative proteomic comparison that is global and systematic, between the resistant cells and their sensitive parental cells.

Clinical relevant, bortezomib adapted AMO1A cells were compared to their parental AMO1 cells in quantitative proteomics experiments. Untreated cells from each cell line were lysed in a buffer containing strong denaturants (4%SDS) to extract as many proteins as possible. The cellular proteins were digested by trypsin. The tryptic peptides were labeled by stable isotopes using the dimethyl labeling method, resistant as heavy and wild type as light(*35*). The labeled peptides were mixed and fractionated by strong cation exchange (SCX) HPLC. Fractions were analyzed by 2 hour gradients reversed phase C18 nano-LC/MS. The raw data was processed and quantified with the MaxQuant program(*36*). Two biological replicates were performed for the experiment. In each replicate, around 4000 proteins were identified. The proteins were only counted for quantification when at least three different peptides were assigned and quantified. Protein quantification was done with the median value of the peptide quantifications. In each replicate, about 2500 proteins were quantified. The reproducible quantifications were extracted from the two replicates. The cut off was set to be $\log_2$(H/L) values higher than 0.5 or lower than -0.5, which means the up or down-regulated proteins show at least 50% higher or 50% lower concentration in the adapted cell line than in the parental cell line (Fig 5G).

In the comparison of AMO1A and AMO1 cells, around 500 proteins were either up or down-regulated in the resistant cells(Fig 5A). Approximately 300 proteins were up-regulated and 350 proteins were down-regulated. The proteins were classified and grouped by protein-protein interactions into clusters to identify groups of proteins with similar functions(Fig 5B-F) (*37*).

Among the up-regulated proteins, proteasome subunits formed a significant cluster (Fig 5B). Fourteen subunits of the 20S and the 19S proteasome particles were found to be up-regulated, twelve of which were 20S proteasome subunits, including the active subunits β1 (PSMB6), β2 (PSMB7) and β5 (PSMB5). Interestingly, the proteasome active subunit β2i (PSMB10) was found to be down-regulated. However, β5i (PSMB8) level was not changed and β1i (PSMB9) was not identified in proteomics experiments, perhaps due to the low abundance. This suggests that, the resistant cells expressed more constitutive and intermediate proteasome to circumvent the proteasome inhibition, but expressed less immuno proteasome(*28*). Considering the results of the cell proliferation assay and proteasome activity assay, there might be an indication that cells containing immuno proteasome are more sensitive to the bortezomib treatment, or that the immuno proteasome is more preferred by the bortezomib molecule. The AMO1A cells expressed more constitutive proteasome but less immuno proteasome, which ended up with a lower ratio of immuno proteasome in the total cellular proteasome population and were less

affected by Bortezomib. The level of PAK2 (p21 activated kinase 2) protein that is a proteasomal substrate was down-regulated in the AMO1A cells. PAK2 is known as an important pro-apoptotic protein that could be activated by cleavage during caspase-mediated apoptosis(*38*). Therefore, the higher proteasomal activity might also contribute to degrade the pro-apoptotic proteins such as PAK2.

In the proteasome related cluster (Fig 5B), the DEAD box protein 49 (DDX49) is observed to be down-regulated, which links the cluster to another one shown in Figure 5C including the DEAD box protein 23 (DDX23). Both DEAD box proteins are ATP dependent RNA helicases essential for transcription(*39*) and this cluster (Fig 5C) of proteins was found to be down-regulated and found to be important for transcription. For instance, RPB1 (POLR2A) is the largest and catalytic component of RNA Polymerase II (RNA Pol II), which is the major RNA polymerase to synthesize mRNA precursors and many functional non-coding RNAs in the eukaryotic cells(*40*). RPB1 contains a carboxy terminal domain (CTD) composed of up to 52 heptapeptide repeats (YSPTSPS) that are essential for polymerase activity(*41*). CCNT1 that was strongly down-regulated is also called positive transcription elongation factor B (P-TEFb), which is reported to facilitate the transition from abortive to productive elongation by phosphorylating the CTD of the large subunit of RNA Pol II(*42*, *43*). The down-regulation of the transcription regulatory proteins might result in global lowering of the protein expression rate. Then the lower rate of protein synthesis in the Bortezomib adapted cells might also contribute to the drug resistance by providing less protein for proteasomal degradation.

Another up-regulated protein cluster in AMO1A cells was made up of the HLA (human leukocyte antigen) proteins, which are the human MHC proteins involved in antigen presentation, and related proteins (Fig 5D). It is known that, bortezomib induces down-regulation of the HLA protein levels in various myeloma cells and enhances the natural killer cell induced lysis of the myeloma cells(*44*). So it is to be expected that the resistant cells over-expressed the HLA complexes to overcome the down regulation induced by bortezomib treatment. As seen, there are several proteins with the same name HLA-DRB1, but they all have different accession numbers. This suggests they might be different isoforms of the HLA-DRB1 molecule. Importantly, the small GTP binding protein Rap-1A (RAP1A, Ras-related protein 1A) is shown to be up-regulated in the cluster. As a small membrane bound GTPase, Rap-1A plays critical roles in multiple cellular pathways and is known as an important oncogene in various cancers. Rap-1A shares almost 50% sequence identity with the classical Ras proteins and several structural features in common. However, Rap-1A counteracts the mitogenic functions of the Ras proteins by competitively binding to their interaction partners such as the c-Raf. Recently, the Ras proteins and Rap1GDS1, which is the Rap1 GDP dissociation stimulator (agonist), were found to be key activators in the protein secretion pathway(*45*). That indicates Rap-1A might also be

involved in the pathway and work as an inhibitor of the secretory activity of extracellular signaling proteins and antibodies.

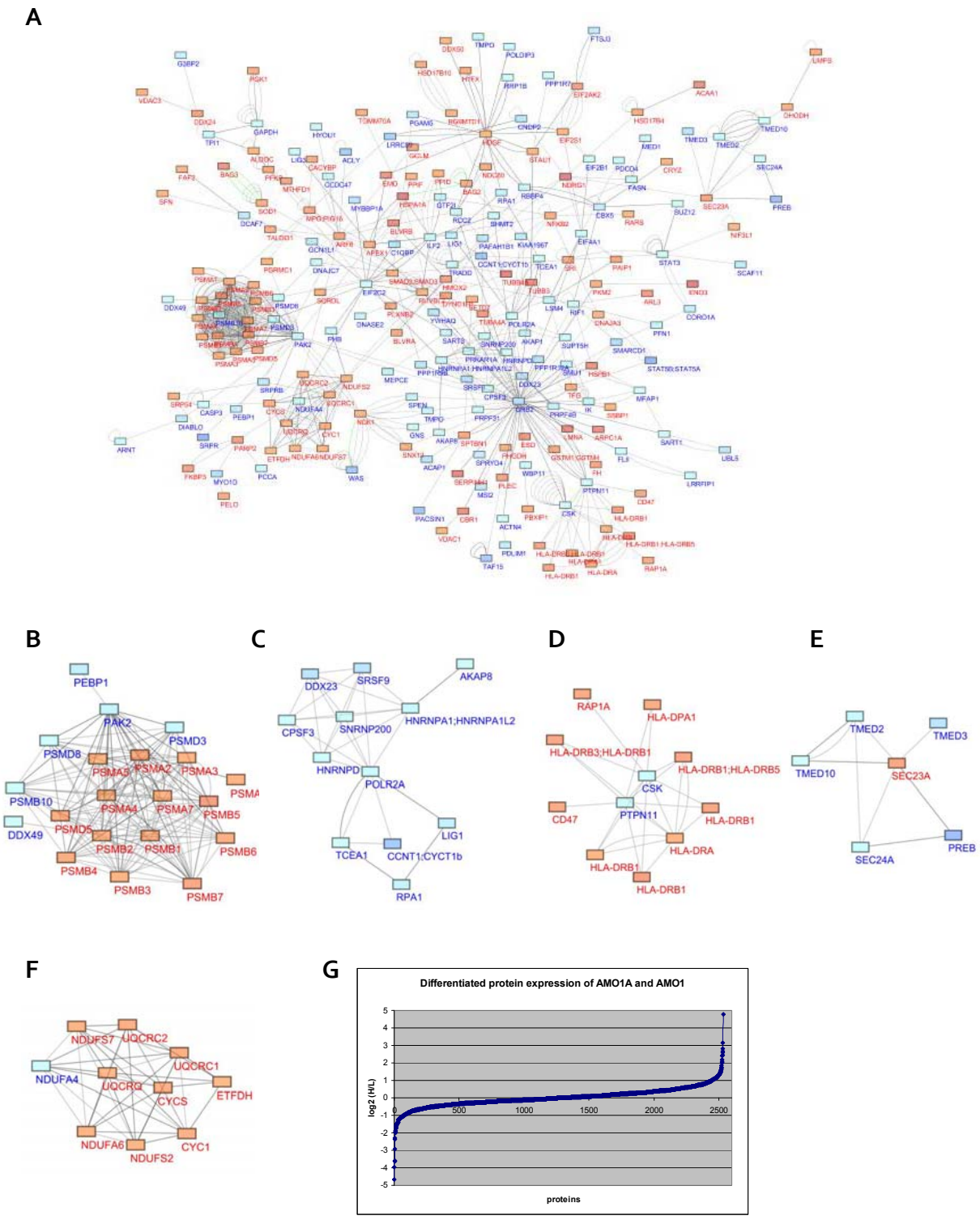


Figure 5: Networks of the up and down-regulated proteins in the AMO1A cells compared to AMO1 cells.

(A) The total network of up and down-regulated proteins identified in the global proteomics experiment (two biological replicates). The proteins in orange or blue color are the up or down-regulated proteins respectively. Proteins that have related functions and protein-protein interactions were grouped into small clusters and highlighted in (B-F).
(B) Proteasome subunits and proteasome related proteins.
(C) Transcription factors and transcription regulatory proteins.
(D) HLA proteins and related proteins.
(E) Protein secretion regulators.
(F) Respiratory chain proteins.
(G) Differentiated protein expression of AMO1A and AMO1.

Together with the possible secretion inhibitor Rap-1A that was up-regulated, several proteins that are involved in the protein secretory pathway were found to be down-regulated (Fig 5E). For instance, SEC12 is a guanine nucleotide exchange factor crucial for the formation of COPII transport vesicles from the ER to the Golgi. Myeloma is an aggressive malignance of the plasma cells that keep generating a single type of monoclonal antibody, the so called M-protein. By down-regulating the activity of the protein secretory pathway, myeloma cells might also lower the protein turnover stress, owing to lower synthesis and secretion of the M-protein.

Several proteins involved in the respiratory chain were up-regulated in AMO1A cells compared to in AMO1 cells(Fig 5F). Most of these proteins are part of either the cytochrome c complex or the NADH:quinone oxidoreductase complex. These complexes are responsible for the production of ATP and the maintenance of the redox status in the cell by reducing oxidized proteins. It is reported that bortezomib induces ER-stress, which produces more reactive oxygen species (ROS) as well as to induce UPR (unfolded protein response) activation. If ROS species are not quenched effectively they may trigger mitochondrial potential malfunction and activate apoptosis cascades. Having a more active respiratory chain (up-regulated) might help counteracting redox stress thus conferring resistance to redox stress promotors.

5.3 Discussion

As a proteasome inhibitor, bortezomib has been used for therapeutic treatment of multiple myeloma for ten years. However, the reasons why some cancer cells are less sensitive to bortezomib treatment than the myeloma cells and how the myeloma cells develop resistance to bortezomib remain unclear. In this chapter, several methods (cell proliferation assay, activity-based proteasome profiling and global proteomics) were combined to investigate the answers.

From the cell proliferation assay, it is seen that solid tumor cells are generally ten times less sensitive to bortezomib treatment than the lymphoblast cells that are B cell cancers in this experiment (Fig 2A). This might be because solid tumor cells express mainly constitutive proteasome, whereas B cells express both constitutive and immuno proteasome. Proteasome inhibition by bortezomib treatment was quantified for each active subunit in different cells by fluorescence and SDS-PAGE based activity-based proteasome profiling (Fig 3). The results of the activity assays correlate to the cell proliferation results. Generally, the active subunits in the solid tumor cells, expressing only constitutive proteasome, were less inhibited by bortezomib than the ones in lymphoblast cells, which might explain why the solid tumor cells are less sensitive. The target active immuno proteasome subunits (β1i and β5i) were more sensitive to bortezomib treatment compared to the constitutive subunits (β1 and β5). Interestingly, in all tested immuno proteasome presenting cell lines, the β1i subunit was always the most sensitive subunit to bortezomib treatment, whose activity decreased most significantly compared to other target subunits. It correlates with the results of Berkers et al., in EL4 murine T lymphocytes bortezomib inhibits β1i prior to other active proteasome subunits(*30*). Normally, the main target of bortezomib is claimed to be β5. However, the observation from these experiments indicates that bortezomib prefers to inhibit the active immuno proteasome subunits β1i and β5i first.

There is a possibility that bortezomib kills myeloma cells but not solid tumors due to lower cell permeation into the solid tumor cells, which are normally adherently growing. It was also reported that the adhesion of myeloma cells induced drug resistance to several anti-cancer drugs(*46*). In order to test whether the growing conditions (adherent or in suspension) of cancer cells affect the cell permeation and the anti-cancer effect of bortezomib, HeLa S_3 and RPMI8226 from both growing conditions were compared on proteasome activity (data not shown) and cell proliferation rate (Fig 2C and D). No significant difference was observed in either experiment, which demonstrates that the adhesion of cancer cells is not likely to change the penetration efficiency of bortezomib or its anti-cancer effect. However, these experiments were performed in vitro; the situation in vivo is much more complicated, especially for the permeation into solid tumors with diameter of centimeters. For this kind of comparison, animal models for multiple myeloma and other in-sensitive solid tumors are necessary(*47*).

One important question in the anti-myeloma research is how the myeloma cells generate resistance to bortezomib. Several previous reports have been published to discuss the molecular mechanism of how myeloma cells became bortezomib-resistant. Point mutations were found in the β5 subunit, which was also up-regulated in some bortezomib adapted cell lines(*14*, *15*). However, no such mutations were found in the bone marrow samples from myeloma patients who develop bortezomib resistance as yet(*16-18*). So there

might be other mechanisms that help myeloma cells circumvent the proteasome inhibition. In order to find these, the bortezomib adapted cells (HL60A and AMO1A) were compared to their sensitive parental cells (HL60 and AMO1) on cell proliferation (Fig 2B) and proteasome activity (Fig 3E and H, F and J). As seen, the adapted cells were at least ten times more resistant than their parental cells. At 50nM bortezomib concentration, the inhibition patterns of HL60 and HL60A were comparable. In AMO1 cells, at 5nM bortezomib concentration, the activity of three target subunits (β1i/5/5i) decreased by over 70%, however in AMO1A cells only β1i decreased so much, the rest target subunits remain over 40% active. Nonetheless, at these bortezomib concentrations the wild type cells have proliferation ratios of around 60%, while the adapted cells have over 90% metabolic activity (Fig 2B). This indicates that, the residual proteasome activity ratio is not the only explanation for the bortezomib resistance in the adapted cells, especially for HL60A cells.

The proteasome activity of non-treated wild type and resistant cells were then compared by ABPP on the same gel (Fig 4A). Apparently, the bortezomib adapted cells have elevated proteasome activity than the wild type cells. This indicates that with the bortezomib treatment at clinical concentration (<20nM), the adapted cells might still have equal or higher proteasome activity compared to the non-treated wild type cells. This might also be explained by just higher expression of the proteasome. Then, the resistant and wild type cells were also compared on proteasome expression instead of activity using native PAGE and anti α subunits immunoblot (Fig 4B). The proteasome expression in both HL60A and AMO1A was up-regulated significantly by 75% and 25%, respectively (Fig 4C). There were reports that identified proteasome genes up-regulation in the bortezomib-resistant patient materials using the GEP (gene expression profiling) analysis, which are also suggested by the results shown here(*20*). These evidences together could partially explain why these cells are resistant to bortezomib treatment, HL60A and AMO1A cells express more proteasome compared to the parental cell lines.

However, is the high proteasome expression level the only reason for the bortezomib-resistance of the adapted cells? The former GEP experiments presented several other genes which do not have direct interactions with proteasome, which were either up or down-regulated on the mRNA levels in the bortezomib resistant cells, though protein levels nowadays have been found to be differentiated from transcription levels of genes(*22*).

The myeloma AMO1 cells were compared to their bortezomib-adapted derivative cells AMO1A in a global proteomics experiment. The quantitative proteomics results show that there are several hundred proteins which were either up or down-regulated in the adapted cells, rather than only the up-regulation of proteasome compared to the wild type cells (Fig 5A). The up and down-regulated proteins were then grouped and classified by

their functions and protein-protein interactions. Several significant up or down-regulated protein groups are presented in Figure 5B-F.

The picture unfolds that the mechanism to generate bortezomib resistance is not just a minor change of one single biological process but a comprehensive modulation of multiple cellular processes, such as faster protein turn over, slower gene transcription, higher antigen presentation, lower protein secretion and enhanced reduction of reactive oxygen species. From the five clusters shown in Figure 5B-F a global picture emerges. For example, up-regulation of the degradation machinery (proteasome) to overcome protein accumulation combined with repressed protein synthesis will decrease the amount of proteins targeted for degradation (Fig 5B and C). Higher expression of HLA molecules will counterpart its down-regulation induced by bortezomib treatment and make the cells less susceptible for recognition and cell lysis by natural killer cells (Fig 5D). Another group of proteins being down-regulated is the secretion machinery (Fig 5E). Myeloma cells have a high production among others of the M-protein for secretion. This secretory pathway needs a continuous and highly active protein turnover, which make cells more sensitive for proteasome inhibitors. The lower levels of secretory proteins may decrease the need for degradation and production of proteins leading the cells to be less sensitive against proteasome inhibition. On the other hand, the altered protein secretion might also lead to altered tumor microenvironment in vivo, through the secreted proteins for extracellular signaling. Bortezomib treatment is also linked to ER-stress and as a last state mitochondrial potential malfunction which finally promotes apoptosis and cell death. AMO1A cells escape this death route by increasing the control on the redox state in mitochondria (Fig 5F). Because the proteasome is also involved in degradation of oxidized proteins, its up-regulation and the increase of the reductive capacity may confer cells a higher admittance of redox-induced stress and promote survival. According to this, other proteins involved in redox homeostasis were found in the global network being over expressed, like oxidoreductases, heme oxigenases, enolases and superoxide dismutases, which suggests this route might be critical for the well being of the cells (Fig 5A).

Apart from these five main clusters other proteins or groups are found in the general network (Fig 5A). There are some chaperones like heat-shock proteins or BAG proteins being over expressed. These chaperone proteins are important for the shuttling of the polyubiquitinylated proteins to either proteasome or autophagy. It might suggest a more tight regulation of protein degradation and maybe also promotion of autophagy as an alternative pathway for protein degradation. Several important nucleic acid binding proteins were found to be down-regulated. General transcription factor II-I (GTF2I) is down-regulated as well as its binding partner RNA Pol II subunit A shown in Figure 5C possibly leading to lowering the general transcription(*48*). Eukaryotic initiation factor 4A1 (EIF4A1) is a key regulator of translation for its function of assisting mRNA binding to ribosomes(*49*).

EIF4A1 down-regulation suggests that the resistant cells lower the protein synthesis rate by repressing not only transcription but also translation rates.

However, there are several short-comings of the global proteomics method. For instance, proteins of low abundance will not be detected. The method does not discriminate between free proteasome subunits and the subunits assembled in active 20S particles. It suggests that down-regulation of β2i observed in the global proteomics experiments might result from lower expression of the protein, instead of lower assembly of the immuno proteasome. In order to verify this observation, more experiments are still needed. Proteasomes should be isolated from cell lysate and analyzed by quantitative proteomics using AQUA (absolute quantification) peptides, in which the proteasome concentration in different cells can be quantified absolutely(*50*).

Overall, the proteomics results provide quantitative comparison of the global protein levels in the bortezomib adapted and wild type cells. In combination with the proteasome activity assay using activity-based proteasome profiling, a bortezomib resistant myeloma cell line (AMO1A) was characterized by proteomics for the first time. In conclusion, the cells adapt to bortezomib treatment due to a comprehensive regulation on expression of genes involved in multiple cellular processes. Notably, several pieces of evidence suggest that the immuno proteasome is preferred by bortezomib compared to constitutive proteasome and that the main target of the drug might be the immuno proteasome. Additionally, the methods used in this chapter can be adapted to clinical setting, which might be used for characterization of clinical relevant samples or other drugs.

5.4 Experimental procedures

5.4.1 Cell culture

HeLa S3 (cervical cancer), MCF7 (breast cancer), A549 (lung cancer) and OvCar3 (ovarian cancer) were grown in DMEM. HL60 (acute myeloid leukemia), Granta519 (mantle cell lymphoma), AMO1 (plasmacytoma), RPMI8226 (multiple myeloma), HL60A (bortezomib adapted HL60) and AMO1A (bortezomib adapted AMO1) were grown in RPMI-1640 medium. All the medium was supplemented with 10% Fetal Calf Serum and 0.1mg/ml penicillin and 0.1mg/ml Streptomycin. All the cells were grown in a 37°C, 5%CO_2 humid incubator. HeLa S3, MCF7, A549, OvCar3, HL60 and RPMI8226 were purchased from ATCC. AMO1 and Granta519 were purchased from the Leibniz Institue DSMZ- German collection of microorganisms and cell cultures. The procedure to create the HL60A and AMO1A was described in previous report(*26*).

5.4.2 Metabolic activity assay and apoptotic assay

The procedure is slightly modified from the manufacture's protocol. 20000 to 30000 cells were seeded in each well of the 96-well white cell culture plate. The cells were treated for fixed time with different concentrations of bortezomib. The luciferase and lysis buffer was added to the cell culture directly. The mixture was incubated in dark for 10 minutes, and mixed on a shaker for 5 minutes. The luminescence could then be measured by a plate reader. The Annexin V apoptotic assay by FACS was performed according to the manufacture's protocol.

5.4.3 Quantitative Activity-based proteasome Profiling

Detailed protocol was described in **Chapter 3** of this thesis. Minor modification was done to ensure that the cells were growing in the Logarithmic phase, while receiving the treatment. HL60, Granta519, AMO1, RPMI8226, HL60A and AMO1A cells were treated at cell concentration of $0.5x10^6$ cells/ml. HeLa S3, MCF7, A549, OvCar3 and F44 were treated at cell confluence of approximately 70%.

5.4.4 Native PAGE for Proteasome

10^7 cells were used to make the lysate. Gentle lysis method (douncing) was used for this aim. Same amount of lysate from each cell line was loaded on the Native PAGE to compare the proteasome expression level. The protocol of making and running native PAGE for proteasome was described in detail by Finely and coworkers(*34*). After running the gel, the proteins were denatured by soaking the gel in 2% SDS, and then transferred to a PVDF membrane for western blotting. Proteasomes were detected by anti alpha subunits antibody on the blot. The chemilluminascence signal was quantified by the ImageLab program.

5.4.5 Global Proteomics

$20x10^6$ cells in Logarithmic growing phase from each cell line were harvested for the global proteomics experiments. The cells were washed with pre-chilled PBS and lysed in SDT buffer (4%SDS, 100mM DTT, 100mM Tris, pH 8.0). The protein concentration was determined, and 500 µg lysate was used from each cell line. Small molecules were removed by a Chloroform/Methanol precipitation(*51*). The proteins were dissolved by 8M Urea. The disulfide bonds reduced by 90mM DTT were then alkylated by 200mM Iodoacetamide. Subsequently, the proteins were digested by Trypsin at 37°C overnight, with an enzyme to protein ratio of 1:100. The digested peptides were desalted on Sep Pak C18 columns, and labeled by the Dimethyl isotopic labeling(*35*). The light and heavy labeled peptides were mixed after the elution from C18 columns. The mixed peptides were fractioned by HPLC equipped with a strong cation exchange (SCX) column. From each SCX run, around 20

fractions were derived. All the fractions were desalted by Stage Tips, and analyzed by an LTQ-Orbitrap nano-LC/MS. The raw data was calculated by MaxQuant program against the uniprot human proteome database to present the quantitative protein identification list(*36*).

5.4.6 cDNA sequencing of PSMB5 gene

Cellular mRNA was purified by Oligo dT Dynabeads, following the manufacture's protocol with minor modification. $2x10^6$ of HL60, HL60A, AMO1 and AMO1A cells were harvested and lysed by lysis buffer containing lithium dodecyl sulfate . The whole lysate was used for the affinity purification of the mRNA. cDNA library was made by a reversed transcription from the total mRNA on beads. The PSMB5 gene was cloned through a PCR with primers 2930 and 2931 from the cDNA library and sequenced.

Primer sequences: 2930: 5' atggcgcttgccagcgtgttggagagacc 3' (forward),

2931: 5' tcaggggtagagccactatacttctcatg 3' (reverse).

Acknowledgement

We thank Marianne Kraus and Christoph Driessen for the providing the bortezomib resistant cell lines and for extensive discussion. We thank Vanessa Frodermann and Johan Kuiper for pursuing the FACS analysis for apoptotic assay. We thank Hans den Dulk and Jaap Brouwer for the cDNA sequencing of PSMB5.

References and Notes

1. Jung, T., Catalgol, B., and Grune, T. (2009) The proteasomal system, *Mol Aspects Med 30*, 191-296.

2. Finley, D. (2009) Recognition and processing of ubiquitin-protein conjugates by the proteasome, *Annu Rev Biochem 78*, 477-513.

3. Borissenko, L., and Groll, M. (2007) 20S proteasome and its inhibitors: crystallographic knowledge for drug development, *Chem Rev 107*, 687-717.

4. Marques, A. J., Palanimurugan, R., Matias, A. C., Ramos, P. C., and Dohmen, R. J. (2009) Catalytic mechanism and assembly of the proteasome, *Chem Rev 109*, 1509-1536.

5. Kloetzel, P. M., and Ossendorp, F. (2004) Proteasome and peptidase function in MHC-class-I-mediated antigen presentation, *Curr Opin Immunol 16*, 76-81.

6. Rock, K. L., and Goldberg, A. L. (1999) Degradation of cell proteins and the generation of MHC class I-presented peptides, *Annu Rev Immunol 17*, 739-779.

7. Kane, R. C., Bross, P. F., Farrell, A. T., and Pazdur, R. (2003) Velcade: U.S. FDA approval for the treatment of multiple myeloma progressing on prior therapy, *Oncologist 8*, 508-513.

8. Bataille, R., Chappard, D., Marcelli, C., Dessauw, P., Baldet, P., Sany, J., and Alexandre, C. (1991) Recruitment of New Osteoblasts and Osteoclasts Is the Earliest Critical Event in the Pathogenesis of Human Multiple-Myeloma, *J Clin Invest 88*, 62-66.

9. Palombella, V. J., Rando, O. J., Goldberg, A. L., and Maniatis, T. (1994) The ubiquitin-proteasome pathway is required for processing the NF-kappa B1 precursor protein and the activation of NF-kappa B, *Cell 78*, 773-785.

10. Schwartz, A. L., and Ciechanover, A. (2009) Targeting proteins for destruction by the ubiquitin system: implications for human pathobiology, *Annu Rev Pharmacol Toxicol 49*, 73-96.

11. Kim, I., Xu, W., and Reed, J. C. (2008) Cell death and endoplasmic reticulum stress: disease relevance and therapeutic opportunities, *Nat Rev Drug Discov 7*, 1013-1030.

12. Ron, D., and Walter, P. (2007) Signal integration in the endoplasmic reticulum unfolded protein response, *Nat Rev Mol Cell Biol 8*, 519-529.

13. Suraweera, A., Munch, C., Hanssum, A., and Bertolotti, A. (2012) Failure of amino acid homeostasis causes cell death following proteasome inhibition, *Mol Cell 48*, 242-253.

14. Ri, M., Iida, S., Nakashima, T., Miyazaki, H., Mori, F., Ito, A., Inagaki, A., Kusumoto, S., Ishida, T., Komatsu, H., Shiotsu, Y., and Ueda, R. (2010) Bortezomib-resistant myeloma cell lines: a role for mutated PSMB5 in preventing the accumulation of unfolded proteins and fatal ER stress, *Leukemia 24*, 1506-1512.

15. Oerlemans, R., Franke, N. E., Assaraf, Y. G., Cloos, J., van Zantwijk, I., Berkers, C. R., Scheffer, G. L., Debipersad, K., Vojtekova, K., Lemos, C., van der Heijden, J. W., Ylstra, B., Peters, G. J., Kaspers, G. L., Dijkmans, B. A., Scheper, R. J., and Jansen, G. (2008) Molecular basis of bortezomib resistance: proteasome subunit beta5 (PSMB5) gene mutation and overexpression of PSMB5 protein, *Blood 112*, 2489-2499.

16. Chapman, M. A., Lawrence, M. S., Keats, J. J., Cibulskis, K., Sougnez, C., Schinzel, A. C., Harview, C. L., Brunet, J. P., Ahmann, G. J., Adli, M., Anderson, K. C., Ardlie, K. G., Auclair, D., Baker, A., Bergsagel, P. L., Bernstein, B. E., Drier, Y., Fonseca, R., Gabriel, S. B., Hofmeister, C. C., Jagannath, S., Jakubowiak, A. J., Krishnan, A., Levy, J., Liefeld, T., Lonial, S., Mahan, S., Mfuko, B., Monti, S., Perkins, L. M., Onofrio, R., Pugh, T. J., Rajkumar, S. V., Ramos, A. H., Siegel, D. S., Sivachenko, A., Stewart, A. K., Trudel, S., Vij, R., Voet, D., Winckler, W., Zimmerman, T., Carpten, J., Trent, J., Hahn, W. C., Garraway, L. A., Meyerson, M., Lander, E. S., Getz, G., and Golub, T. R. (2011) Initial genome sequencing and analysis of multiple myeloma, *Nature 471*, 467-472.

17. Lichter, D. I., Danaee, H., Pickard, M. D., Tayber, O., Sintchak, M., Shi, H., Richardson, P. G., Cavenagh, J., Blade, J., Facon, T., Niesvizky, R., Alsina, M., Dalton, W., Sonneveld, P., Lonial, S., van de Velde, H., Ricci, D., Esseltine, D. L., Trepicchio, W. L., Mulligan, G., and Anderson, K. C. (2012) Sequence analysis of beta-subunit genes of the 20S proteasome in patients with relapsed multiple myeloma treated with bortezomib or dexamethasone, *Blood 120*, 4513-4516.

18. Politou, M., Karadimitris, A., Terpos, E., Kotsianidis, I., Apperley, J. F., and Rahemtulla, A. (2006) No evidence of mutations of the PSMB5 (beta-5 subunit of proteasome) in a case of myeloma with clinical resistance to bortezomib, *Leukemia Res 30*, 240-241.

19. Kuhn, D. J., Berkova, Z., Jones, R. J., Woessner, R., Bjorklund, C. C., Ma, W. C., Davis, R. E., Lin, P., Wang, H., Madden, T. L., Wei, C. M., Baladandayuthapani, V., Wang, M., Thomas, S. K., Shah, J. J., Weber, D. M., and Orlowski, R. Z. (2012) Targeting the insulin-like growth factor-1 receptor to overcome bortezomib resistance in preclinical models of multiple myeloma, *Blood 120*, 3260-3270.

20. Shaughnessy, J. D., Qu, P. P., Usmani, S., Heuck, C. J., Zhang, Q., Zhou, Y. M., Tian, E. M., Hanamura, I., van Rhee, F., Anaissie, E., Epstein, J., Nair, B., Stephens, O., Williams, R., Waheed, S., Alsayed, Y., Crowley, J., and Barlogie, B. (2011) Pharmacogenomics of bortezomib test-dosing identifies hyperexpression of proteasome genes, especially PSMD4, as novel high-risk feature in myeloma treated with Total Therapy 3, *Blood 118*, 3512-3524.

21. Stessman, H. A. F., Baughn, L. B., Sarver, A., Xia, T., Deshpande, R., Mansoor, A., Walsh, S. A., Sunderland, J. J., Dolloff, N. G., Linden, M. A., Zhan, F. H., Janz, S., Myers, C. L., and Van Ness, B. G. (2013) Profiling Bortezomib Resistance Identifies Secondary Therapies in a Mouse Myeloma Model, *Mol Cancer Ther 12*, 1140-1150.

22. Rogers, S., Girolami, M., Kolch, W., Waters, K. M., Liu, T., Thrall, B., and Wiley, H. S. (2008) Investigating the correspondence between transcriptomic and proteomic expression profiles using coupled cluster models, *Bioinformatics 24*, 2894-2900.

23. Ge, F., Xiao, C. L., Bi, L. J., Tao, S. C., Xiong, S., Yin, X. F., Li, L. P., Lu, C. H., Jia, H. T., and He, Q. Y. (2010) Quantitative phosphoproteomics of proteasome inhibition in multiple myeloma cells, *PLoS One 5*.

24. Kim, W., Bennett, E. J., Huttlin, E. L., Guo, A., Li, J., Possemato, A., Sowa, M. E., Rad, R., Rush, J., Comb, M. J., Harper, J. W., and Gygi, S. P. (2011) Systematic and quantitative assessment of the ubiquitin-modified proteome, *Mol Cell 44*, 325-340.

25. Udeshi, N. D., Mani, D. R., Eisenhaure, T., Mertins, P., Jaffe, J. D., Clauser, K. R., Hacohen, N., and Carr, S. A. (2012) Methods for quantification of in vivo changes in protein ubiquitination following proteasome and deubiquitinase inhibition, *Mol Cell Proteomics 11*, 148-159.

26. Ruckrich, T., Kraus, M., Gogel, J., Beck, A., Ovaa, H., Verdoes, M., Overkleeft, H. S., Kalbacher, H., and Driessen, C. (2009) Characterization of the ubiquitin-proteasome system in bortezomib-adapted cells, *Leukemia 23*, 1098-1105.

27. Kraus, M., Ruckrich, T., Reich, M., Gogel, J., Beck, A., Kammer, W., Berkers, C. R., Burg, D., Overkleeft, H., Ovaa, H., and Driessen, C. (2007) Activity patterns of proteasome subunits reflect bortezomib sensitivity of hematologic malignancies and are variable in primary human leukemia cells, *Leukemia 21*, 84-92.

28. Guillaume, B., Chapiro, J., Stroobant, V., Colau, D., Van Holle, B., Parvizi, G., Bousquet-Dubouch, M. P., Theate, I., Parmentier, N., and Van den Eynde, B. J. (2010) Two abundant proteasome subtypes that uniquely process some antigens presented by HLA class I molecules, *Proc Natl Acad Sci U S A 107*, 18599-18604.

29. Kisselev, A. F., and Goldberg, A. L. (2005) Monitoring activity and inhibition of 26S proteasomes with fluorogenic peptide substrates, *Methods Enzymol 398*, 364-378.

30. Berkers, C. R., Verdoes, M., Lichtman, E., Fiebiger, E., Kessler, B. M., Anderson, K. C., Ploegh, H. L., Ovaa, H., and Galardy, P. J. (2005) Activity probe for in vivo profiling of the specificity of proteasome inhibitor bortezomib, *Nat Methods 2*, 357-362.

31. Li, N., Kuo, C. L., Paniagua, G., van den Elst, H., Verdoes, M., Willems, L. I., van der Linden, W. A., Ruben, M., van Genderen, E., Gubbens, J., van Wezel, G. P., Overkleeft, H. S., and Florea, B. I. (2013) Relative quantification of proteasome activity by activity-based protein profiling and LC-MS/MS, *Nat Protoc 8*, 1155-1168.

32. Li, N., Overkleeft, H. S., and Florea, B. I. (2012) Activity-based protein profiling: an enabling technology in chemical biology research, *Curr Opin Chem Biol 16*, 227-233.

33. Cravatt, B. F., Wright, A. T., and Kozarich, J. W. (2008) Activity-based protein profiling: from enzyme chemistry to proteomic chemistry, *Annu Rev Biochem 77*, 383-414.

34. Elsasser, S., Schmidt, M., and Finley, D. (2005) Characterization of the proteasome using native gel electrophoresis, *Methods Enzymol 398*, 353-363.

35. Boersema, P. J., Raijmakers, R., Lemeer, S., Mohammed, S., and Heck, A. J. (2009) Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics, *Nat Protoc 4*, 484-494.

36. Cox, J., and Mann, M. (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification, *Nat Biotechnol 26*, 1367-1372.

37. Koh, G. C. K. W., Porras, P., Aranda, B., Hermjakob, H., and Orchard, S. E. (2012) Analyzing Protein-Protein Interaction Networks, *J Proteome Res 11*, 2014-2031.

38. Jakobi, R., McCarthy, C. C., Koeppel, M. A., and Stringer, D. K. (2003) Caspase-activated PAK-2 is regulated by subcellular targeting and proteasomal degradation, *J Biol Chem 278*, 38675-38685.

39. Mathew, R., Hartmuth, K., Mohlmann, S., Urlaub, H., Ficner, R., and Luhrmann, R. (2008) Phosphorylation of human PRP28 by SRPK2 is required for integration of the U4/U6-U5 tri-snRNP into the spliceosome, *Nat Struct Mol Biol 15*, 435-443.

40. Wintzerith, M., Acker, J., Vicaire, S., Vigneron, M., and Kedinger, C. (1992) Complete Sequence of the Human Rna Polymerase-Ii Largest Subunit, *Nucleic Acids Res 20*, 910-910.

41. Akhtar, M. S., Heidemann, M., Tietjen, J. R., Zhang, D. W., Chapman, R. D., Eick, D., and Ansari, A. Z. (2009) TFIIH Kinase Places Bivalent Marks on the Carboxy-Terminal Domain of RNA Polymerase II, *Mol Cell 34*, 387-393.

42. Jang, M. K., Mochizuki, K., Zhou, M. S., Jeong, H. S., Brady, J. N., and Ozato, K. (2005) The bromodomain protein Brd4 is a positive regulatory component of P-TEFb and stimulates RNA polymerase II-dependent transcription, *Mol Cell 19*, 523-534.

43. Yang, Z. Y., Yik, J. H. N., Chen, R. C., He, N. H., Jang, M. K., Ozato, K., and Zhou, Q. (2005) Recruitment of P-TEFb for stimulation of transcriptional elongation by the bromodomain protein brd4, *Mol Cell 19*, 535-545.

44. Shi, J., Tricot, G. J., Garg, T. K., Malaviarachchi, P. A., Szmania, S. M., Kellum, R. E., Storrie, B., Mulder, A., Shaughnessy, J. D., Jr., Barlogie, B., and van Rhee, F. (2008) Bortezomib down-regulates the cell-surface expression of HLA class I and enhances natural killer cell-mediated lysis of myeloma, *Blood 111*, 1309-1317.

45. Simpson, J. C., Joggerst, B., Laketa, V., Verissimo, F., Cetin, C., Erfle, H., Bexiga, M. G., Singan, V. R., Heriche, J. K., Neumann, B., Mateos, A., Blake, J., Bechtel, S., Benes, V., Wiemann, S., Ellenberg, J., and Pepperkok, R. (2012) Genome-wide RNAi screening identifies human proteins with a regulatory function in the early secretory pathway, *Nat Cell Biol 14*, 764-774.

46. Damiano, J. S., Cress, A. E., Hazlehurst, L. A., Shtil, A. A., and Dalton, W. S. (1999) Cell adhesion mediated drug resistance (CAM-DR): Role of integrins and resistance to apoptosis in human myeloma cell lines, *Blood 93*, 1658-1667.

47. Yaccoby, S., Barlogie, B., and Epstein, J. (1998) Primary myeloma cells growing in SCID-hu mice: A model for studying the biology and treatment of myeloma and its manifestations, *Blood 92*, 2908-2913.

48. Egloff, A. M., and Desiderio, S. (2001) Identification of phosphorylation sites for Bruton's tyrosine kinase within the transcriptional regulator BAP/TFII-I, *J Biol Chem 276*, 27806-27815.

49. Quinn, C. M., Wiles, A. P., El-Shanawany, T., Catchpole, I., Alnadaf, T., Ford, M. J., Gordon, S., and Greaves, D. R. (1999) The human eukaryotic initiation factor 4Al gene (EIF4A1) contains multiple regulatory elements that direct high-level reporter gene expression in mammalian cell lines, *Genomics 62*, 468-476.

50. Kettenbach, A. N., Rush, J., and Gerber, S. A. (2011) Absolute quantification of protein and post-translational modification abundance with stable isotope-labeled synthetic peptides, *Nat Protoc 6*, 175-186.

51. Wessel, D., and Flugge, U. I. (1984) A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids, *Anal Biochem 138*, 141-143.

6

Activity-based protein profiling reveals reactivity of the murine thymoproteasome-specific subunit β5t

Florea BI, Verdoes M, Li N, van der Linden WA, Geurink PP, van den Elst H, Hofmann T, de Ru A, van Veelen PA, Tanaka K, Sasaki K, Murata S, den Dulk H, Brouwer J, Ossendorp FA, Kisselev AF, Overkleeft HS, Chem Biol. 2010, 17, 795-801.

6.1 Introduction

The ability to recognise non-self oligopeptides is a key feature of mammalian immunity. T cells that recognise antigenic oligopeptides elicit a directed adaptive immune response aimed at the identification and eventual eradication of the invading pathogen that is the source of the non-self protein from which the antigenic oligopeptide is derived(*1*, *2*). T cell recognition is effected by binding of specific T cell receptors to the antigenic peptides that are complexed to either major hisocompatibility complex (MHC) class I or MHC class II molecules(*3*, *4*). MHC I molecules present oligopeptides derived from cytosolic and nuclear proteins to CD8+ cytotoxic T lymphocytes (CTL) and by this virtue report on the presence of virally encoded proteins(*5*). T cells specific for non-self peptides are produced by thymic selection. The generation in the thymus of non-self peptide selective CTL proceeds in two discreet events(*6*). Positive selection is mediated by cortical thymic epithelial cells. In this process, thymocytes expressing T cell receptors are confronted with tissues expressing MHC I molecules loaded with oligopeptides. Current understanding is that the peptide antigens produced by cortical thymic epithelial cells are low affinity MHC I binders. Thymocytes, passing through the thymic cortex, that bind to MHC I molecules carrying a peptide load are selected from thymocytes expressing non-binding receptors. In the ensuing negative selection step, mediated by medullary thymic epithelial cells, thymocytes from the positively selected pool that are responsive to MHC I molecules exposing self peptides are eliminated.

Recently, Tanaka and co-workers made a major breakthrough towards understanding how positive selection proceeds(*7*). They found that epithelial cells at the thymic cortex express, next to the constitutive proteasome and the immunoproteasome, a third 20S proteasome particle which was dubbed the thymoproteasome. The 20S core particle of the proteasome is assembled from α and β subunits in a pattern of four, stacked,

heptameric rings (α1-7, β1-7, β1-7, α1-7) generating a barrel-shaped structure that contains 2 copies of the catalytically active β-subunits: β1 (post acidic), β2 (tryptic-like), β5 (chymotriptic-like) peptidase activities(*8*). The thymoproteasome contains the β1i and β2i subunits just like the immunoproteasome, with the important exception that the unique subunit β5t replaces the immunoproteasome specific subunit, β5i.

The thymoproteasome is the most abundant proteasome species in cortical thymic epithelial cells (cTEC). Thymoproteasome expression may have implications for the repertoire of oligopeptides presented by MHC I molecules on the surface of cTEC's that might significantly differ from to the repertoire produced by medullary thymic epithelial cells. Closer inspection of the thymoproteasome 20S particle revealed that, in contrast to the constitutive and the immunoproteasome, it possessed little chymotryptic activity, a finding that seems to correlate with the hydrophilic nature of the putative substrate-binding site of β5t compared to β5/β5i(*7*). In theory β5t can contribute in two ways to the generation of specific MHC I peptides used in positive T cell selection(*9*). It could act as an impassive, catalytically inactive bystander, in which case β1i/β2i produce the majority of MHC I peptides with a bias towards their substrate preferences. Alternatively, it could actively participate in protein degradation and assist in producing 'non-self' peptides thanks to its intrinsic substrate preference, which then must be distinct from that of β5/ β5i.

Activity-based probes are synthetic compounds bearing a reporter or affinity tag and an enzyme reactive group that can covalently bind to the active site of an enzyme(*10*). The tagged enzymatic activities can than be visualized by fluorescence or affinity purified, digested with trypsin and identified by LC/MS analysis. This Chapter demonstrates, by making use of activity-based proteasome probes(*11*), that β5t is in fact a catalytically active subunit, and show that its preference towards established proteasome inhibitors differs substantially from those of β5/ β5i.

6.2 Results and discussion

6.2.1 Activity-based profiling reveals β5t activity

As the first experiment, whole tissue thymus homogenate from 3 weeks old mice was incubated with the fluorescent broad-spectrum ABP's **1**, **2**, **4** and MV151 shown in Figure 1 (for the synthesis of probes **2** and **4** see supplemental methods) (*12*, *13*). Proteins were resolved by SDS-PAGE under reducing conditions and fluorescently labeled proteasome subunits were visualized by in-gel fluorescence scanning. In Figure 2A, MV151 shows the typical band pattern of staining that is similar to that of the EL4 cell line expressing the constitutive and the immunoproteasome (see supplemental Figure S1) indicating that both particles are expressed in the thymus(*14*). Peptide vinyl sulphone **1**, the biotinylated derivative of MV151, shows a similar pattern as MV151. Interestingly, the peptide epoxyketones **2** and **4** show two new bands that run below and above the

constitutive and immunoproteasome subunits. Of these, the lower band corresponds to β1i.

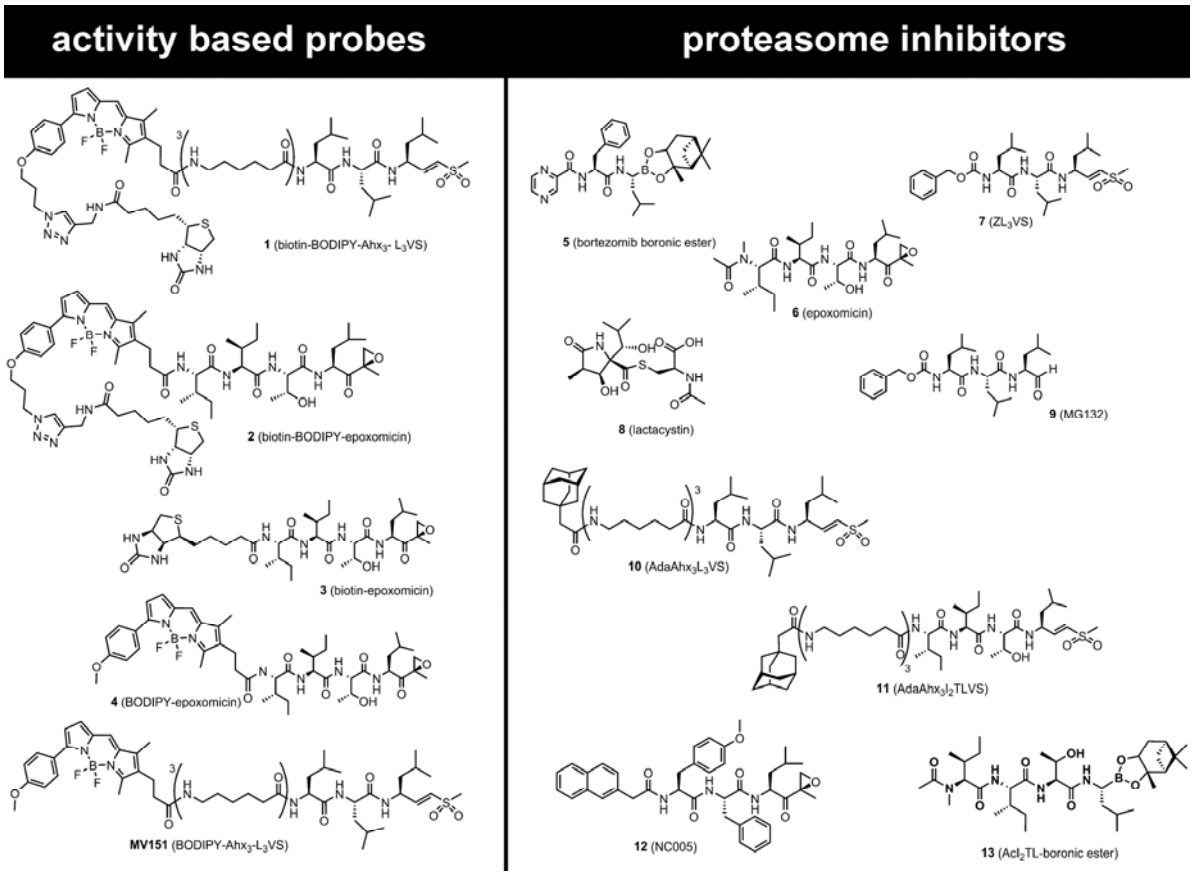


Figure 1: Activity-based probes and proteasome inhibitors used in this study.
In addition to the enzyme reactive group (warhead) and targeting sequence of the inhibitors, activity-based probes are equipped with a fluorophore for in-gel detection, a biotin tag for affinity purification or with both.

To ascertain whether the new, epoxyketone sensistive protein, with a gel mobility corresponding to the predicted molecular weight of β5t(7), is indeed β5t and not a thymus-specific gene product unrelated to the thymoproteasome, a pull-down experiment was performed by making use of the biotin moiety present in ABPs **1** and **2**. Biotinylated proteins from thymus homogenate were captured by streptavidin-coated magnetic beads, resolved by SDS-PAGE and detected both by fluorescence and silver staining. Figure 2B shows the specific purification of several proteins that run in a pattern similar with that of Figure 2A. Protein ID, indicated by arrows in Figure 2B, was determined by on-bead (Table S1) and in-gel tryptic digestion followed by LC-MS/MS analysis. Oligopeptides corresponding to the expected constitutive proteasome (β1/β2/β5) and immunoproteasome (β1i/β2i/β5i) were captured by ABP **1** but no evidence for β5t was

found. Peptides derived from β5t were found by affinity purification with ABP **2**, indeed in the band running higher than the other active β subunits. However, the protein yield achieved by pull-down with ABP **1** and **2** was low. Then, the short biotinylated epoxomicin ABP **3** might increase the pull-down efficiency was synthesized to enhance the pull down efficiency (for the synthesis of probe **3** see supplemental methods). ABP **3** performed as expected, showing bands of similar pattern as ABP **2**, stronger signal in silver stained gels and reliable LC-MS identification of proteins that is presented in Table 1.

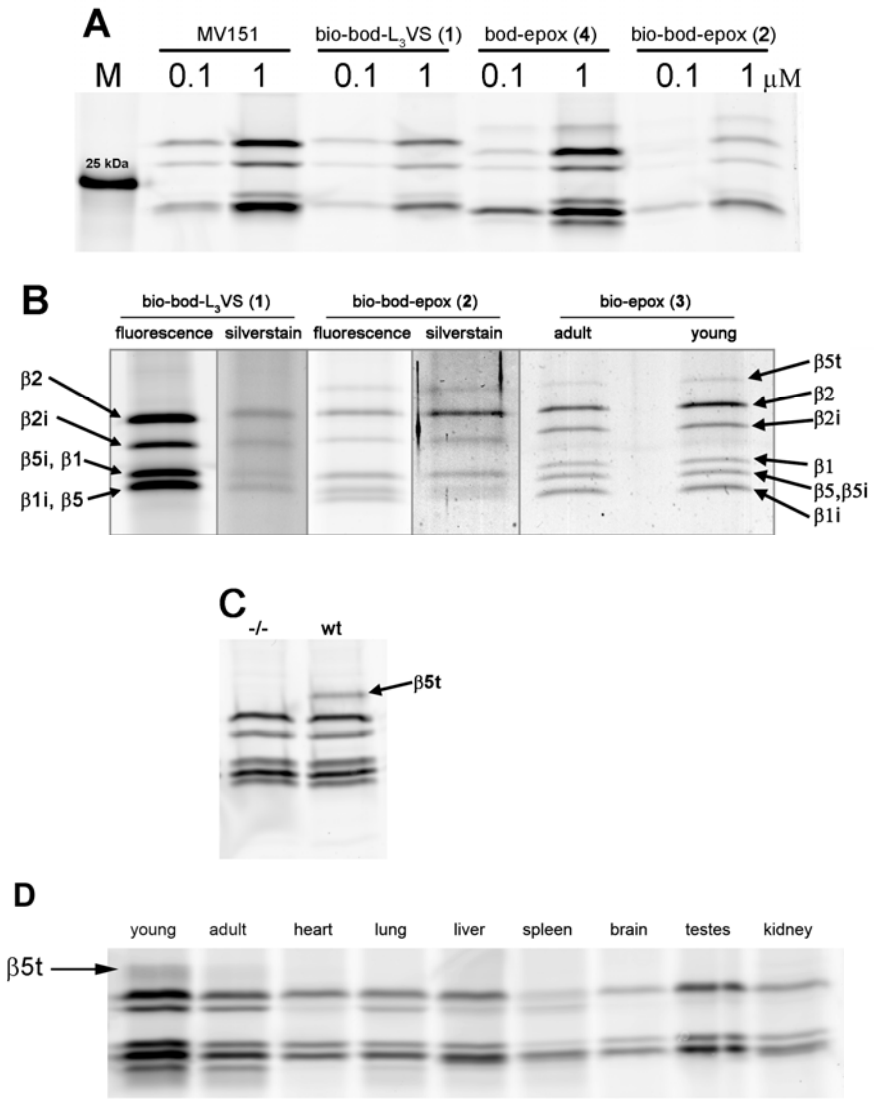


Figure 2: Activity-based protein profiling, affinity purification and LC-MS identification of proteasome β subunits in murine tissues lysates.

(A) In-gel fluorescence detection of active proteasome β subunits in 3 weeks old wild type murine thymus homogenate after labeling with MV151, ABP **1**, **2** and **4**.

(B) In-gel fluorescence and silver stain detection of active proteasome β subunits in young and adult thymus after labeling with ABP **1**, **2**, **3** and affinity purification. Protein

identification by LC-MS analysis of in-gel digested silver stained bands (indicated by arrows).
(C) In-gel fluorescence detection with ABP **4** of β5t activity in wild type and absence of activity in the (-/-) β5t knock down thymus from 3 weeks old mice.
(D) Activity-based proteasome profiling using ABP **4** shows β5t activity in murine thymus (young and adult) but not in heart, lung, liver, spleen, brain, testes and kidney.

Table 1: Protein identification of silver stained bands captured by probe **3** in Figure 2B, by in-gel digestion and LC-MS analysis.

		mass	cover			pept	
prot acc		(Da)	% AA	z	ppm	score	peptide sequence
Psmb11	(β5t)	27834	20	2	-0.29	50	SLEQELEAK
IPI00221461				2	-0.54	39	ESGWEYVSR
				2	0.05	34	LLGTTSGTSADCATWYR
				3	0.00	25	GYHYDMTIQEAYTLAR
Psmb7	(β2)	25235	57	2	-1.20	41	GTTAVLTEK
IPI00136483				2	0.39	64	DGIVLGADTR
				3	0.14	42	FRPDMEEEEAK *
				2	-0.63	45	LDFLRPFSVPNK *
				3	0.58	34	LDFLRPFSVPNKK **
				2	2.14	130	LPYVTMGSGSLAAMAVFEDK
				3	-0.19	64	VTPLEIEVLEETVQTMDTS #
				4	5.34	100	IHFISPNIYCCGAGTAADTDMTTQLISSNLELHSLTTGR
Psmb10	(β2i)	24789	18	2	0.20	25	DGVILGADTR
IPI00316736				2	-0.42	44	ALSTPTEPVQR
				2	1.54	85	EVRPLTLELLEETVQAMEVE #
Psmb6	(β1)	21982	48	2	0.09	56	QVLLGDQIPK
IPI00119239				2	-2.05	76	LAAIQESGVER
				2	-0.46	132	DECLQFTANALALAMER
				2	1.65	100	QSFAIGGSGSSYIYGYVDATYR
				4	4.96	36	SGSAADTQAVADAVTYQLGFHSIELNEPPLVHTAASLFK

Psmb8	(β5i)	22635	42	2	0.71	72	ATAGSYISSLR
IPI00116712				2	0.00	42	FQHGVIVAVDSR
				2	-1.03	73	VESSDVSDLLYK
				2	-0.35	63	GPGLYYVDDNGTR
				2	-0.34	60	QDLSPEEAYDLGR
				2	0.95	72	VIEINPYLLGTMSGCAADCQYWER
Psmb5	(β5)	22514	13	2	0.23	24	ATAGAYIASQTVK
IPI00317902				2	-0.28	48	GPGLYYVDSEGNR
Psmb9	(β1i)	21313	17	2	0.67	79	FTTNAITLAMNR
IPI00309379				2	-0.43	101	DGSSGGVIYLVTITAAGVDHR

Table 1. Protein name, mass of the active β subunit, % coverage of the protein by amino acids identified by LC-MS, charge of the peptide (z), measurement error (ppm), Mascot peptide scores, one (*) or two (**) miss cleavages, and C-terminal peptides (#). Mascot identifications were manually validated.

Thymus from adult animals treated in the same fashion shows β5t activity as well, which suggests that the murine thymoproteasome remains active for at least 6 months. Next to the active proteasome β subunits, only four endogenously biotinylated background proteins were recovered with this method, a result that reflects the selectivity of ABPs **1**, **2**, **3**, and **4** towards proteasomes. Thymus lysates of 2 weeks old mice in which the β5t protein expression was genetically knocked down show normal activity of immuno- and constitutive proteasome compared with the wild type, but complete absence of β5t activity (Figure 2C). To characterize the expression of β5t in murine tissues a tissue scan was performed with ABP probe **4**. Figure 2D shows that β5t activity is exclusively present in the young thymus and at lower activity in thymus of 6 months old mice. Integration of the fluorescent signal from young thymus indicated that β5t contributes to some 4% of the total active subunits signal in this full thymus lysate. Heart, lung, liver, spleen, brain, testes and kidney do not show β5t activity. The presence of immunoproteasome bands in the heart, lung, liver and spleen tissues is explained by the presence of lymphocytes in these organs.

6.2.2 LC-MS^3 analysis of the β5t active-site peptide

Isolation and analysis of the active-site peptide covalently bound to ABP probe **3** would be the ultimate proof for the β5t acitivity. Biotin-epoxomicin binds to the catalytic N-terminal threonine via an irreversible morpholino ring formation shown in Figure 3A. The β5t active-site peptide (Figure 3B) is generated after denaturation and tryptic digest of the

thymoproteasome. Given that biotin-epoxomicin binds to all active β subunits, 6 different active-site peptides were expected because the tryptic peptides derived from β5 and β5i are identical (see Table S2). After LC-MS analysis, the active-site peptides were identified from the high resolution full MS scans by their exact mass and charge (Figure 3C). Further evidence was provided by the MS/MS (MS^2) fragmentation that revealed the presence of the biotin-epoxomicin signature ions b1, b2, b3, and b4 from Figure 3D. In fact, the favored fragmentation of the morpholino ring, due to push-pull radical stabilization of the ions, yields mainly two major ions b4 and y7 where y7 contains the peptide sequence of the β subunit active-site(*15*). By electrostatic trapping and further MS^3 fragmentation of the y7 ion, the LAFR sequence of the β5t active-site peptide was identified (Figure 3E). Taken together, this data set demonstrates that β5t is, indeed, a catalytically active proteasome subunit.

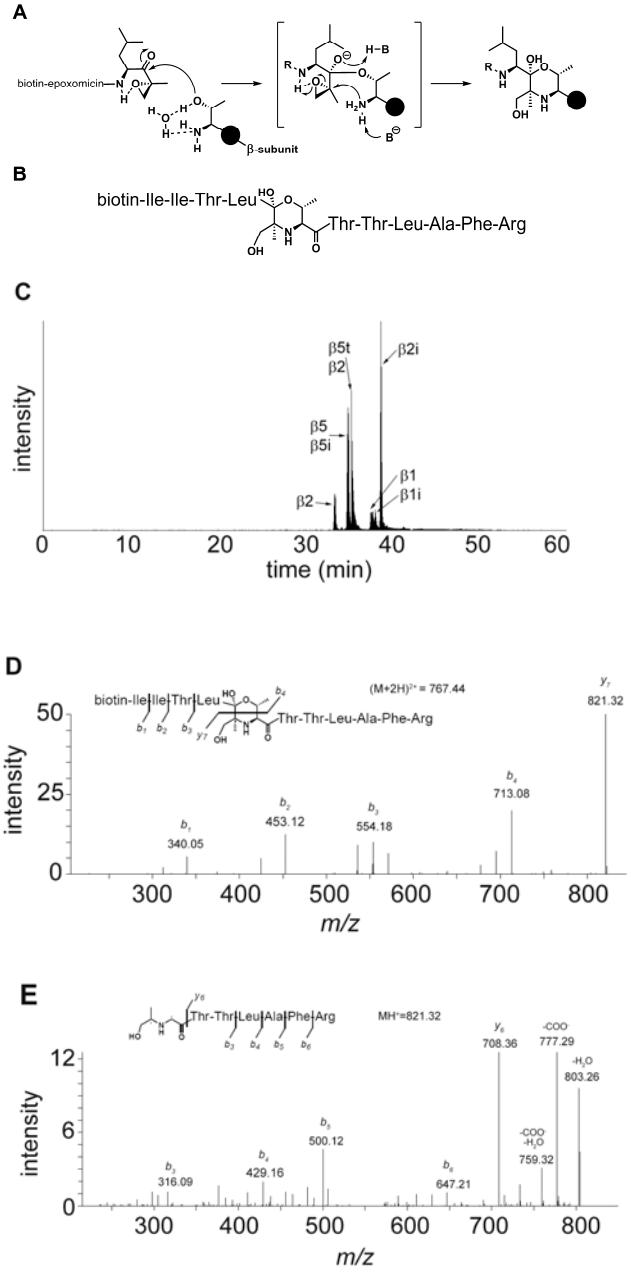


Figure 3: Active-Site Peptide Identification and Determination of Proteasome β subunits by Affinity Purification, Tryptic digest and LC-MS analysis.

(A) Reaction mechanism of biotin-epoxomicin **3** with the catalytically active N-terminal Thr residue of active proteasome β subunits. The morpholino ring formation results in a covalent and irreversible binding.

(B) Schematic representation of the biotin-epoxomicin modified, N-terminal active site tryptic peptide of β5t. Amino acid residues are represented in a 3-letter code.

(C) LC-MS elution profile of the six unique biotinylated tryptic peptides derived from the active sites. Notice that β5 and β5i active site peptides are identical (see Table S2)

(D) LC-MS^2 determination of the β5t active site fragmentation pattern. The parent ion (m/z $(M+2H)^{2+}$ = 767.44) was fragmented. The b1, b2, b3 and b4 ions are signature ions of the biotin-

epoxomicin N-terminal part. The abundant y7 ion containing the β5t active site peptide sequence was selected for further (MS^3) fragmentation (see panel (E)).
(E) LC-MS_3 determination of the y7 ion (MH^+ = 821.32) revealing the β5t active site peptide amino acid sequence.

6.2.3 Competitive activity-based protein profiling reveals β5t substrate specificity

The finding that β5t reacts with epoxyketones **2, 3,** and **4** but not with peptide vinyl sulphones **1** and MV151**,** gives a first indication of an altered substrate specificity compared to β5/β5i. With probe **4** in hand as read-out**,** investigation was set out to reveal the β5t substrate preference by competitive activity-based studies with established proteasome inhibitors of diverse chemical characteristics.

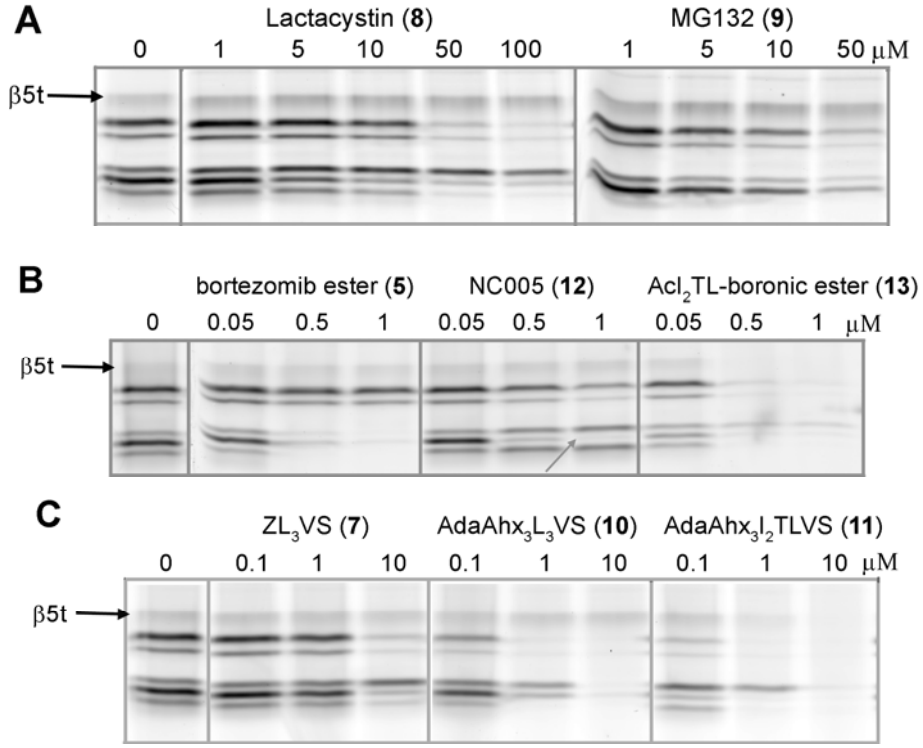


Figure 4: Analysis of β5t substrate specificity in Juvenile murine thymus lysates by competitive activity-based protein profiling with ABP **4**.
(A) Lysates were exposed to increasing concentrations Lactacystin or MG132**,** residual β5t activity was stained with ABP **4** and visualized by in-gel fluorescence detection. The inhibitors are not reactive towards the β5t activity.
(B) Bortezomib efficiently inhibits β5, β5i**,** β1 and β1i activity but not β5t. NC005 specifically targets β5 (indicated by the grey arrow) but not β5t. The mixed inhibitor containing the boronic ester warhead equipped with the epoxomicin tail efficiently blocks β5t.
(C) In analogy to (B), installation of the epoxomicin IleIleThrLeu peptide targeting motif to the vinyl sulphone warhead affords a potent inhibitor of the β5t activity.

Figure 4A shows the results of the most commonly used proteasome inhibitors lactacystin and MG132(*16, 17*). Both require concentrations higher that 10 μM for broad-spectrum proteasome inhibition with marked affinity for β5 and β2, but do not inhibit the β5t activity. Figure 4B shows that the Bortezomib boronic ester **5** effectively blocks β1, β1i, β5 and β5i as previously described(*11, 18*), while the subunit specific inhibitor NC005 selectively inhibits the β5/β5i subunits(*19*), but neither interacts with the β5t. However, mixing of the potent boronic ester warhead with the AcI2TL peptide sequence inherent to epoxomicin as in compound **13**, abolished the subunit preference of Bortezomib and efficiently inhibited β5t. A similar effect is revealed in Figure 4C where the vinyl sulphone warhead was equipped with the AdaAhx3 extended I2TL motif. Apparently, the presence of the hydrophilic threonine side chain at P2 in an inhibitor or ABP probe is favorable for affinity to the β5t subunit(*20*). From the results, some interesting trends pointing towards a substrate preference of β5t that is rather distinct to that of β5/β5i appear. Whereas β5t is sensitive towards the broad-spectrum proteasome epoxomicin **6**, it is quite unreactive towards the β5/β5i biased compounds. Bortezomib boronic ester **5** (which at the concentrations used disables β1/β1i/β5/β5i) and lactacystin **8** are inreactive towards β5t, as is the case with vinyl sulfone **7**, peptide aldehyde **9** and epoxyketone **12**. Altogether, our data, revealing that β5t is catalytically active towards inhibitors more hydrophilic than those recognized by β5/β5i, point towards the involvement of β5t in the generation of a unique set of oligopeptides complexed to MHC I molecules for optimal positive T cell selection.

6.3 Conclusion

In summary, it is proved for the first time that β5t is catalytically active. Interestingly, active β5t is also found in adult thymi. The first insight into the nature of the substrate preference of β5t is also provided. This body of evidence was made possible by the direct action of activity-based probes with emphasis on the bi-functional ABPs that facilitate both read-out and affinity purification. A more thorough investigation is needed to establish the nature of substrates accepted by β5t, and thus the nature of the MHC I peptides produced by the thymoproteasome in the positive T cell selection process.

6.4 Experimental procedure

6.4.1 Animals and tissues

Thymus and other organs were isolated from young (3 weeks) or adult mice and kindly provided by Ine Tijdens, Chantal Pont and Prof Dr. Bob van de Water. Thymus from β5t knock-out mice was kindly provided by Dr Tanaka. Organ isolation was approved by the animal experimentation ethical committee of the Leiden University and Tokyo Metropolitan Institute of Medical Science.

6.4.2 Compounds

Design, synthesis and mechanism of action of the activity-based probes **1**-**4**, MV151 and the proteasome inhibitors **5**-**11**, **13** is reviewed in ((*12*) and the references therein). NC005 is described in Britton et al.(*19*). Lactacystin, MG132 and all other compounds of analytical grade were purchased from Sigma-Aldrich.

Fmoc-Ile-Thr(tBu)-OMe

L-threonine(tBu) methyl ester HCl salt (2.5 g, 11 mmol) was dissolved in DCM (60 mL). To this solution were added Fmoc-L-isoleucine (4.7 g, 13.3 mmol, 1.2 equiv.), HCTU (5.5 g, 13.3 mmol, 1.2 equiv.) and DiPEA (6.0 mL, 36 mmol, 3.3 equiv.). The mixture was stirred for 2 hours after which TLC analysis indicated a completed reaction. The mixture was concentrated in vacuo, dissolved in EtOAc and extracted with 1 M HCl (2x), saturated $NaHCO_3$ (2x) and brine. The organic layer was dried ($MgSO_4$) and concentrated under reduced pressure. Purification of the product by column chromatography (10% → 15% EtOAc/petroleum ether) gave the title compound as a colorless solid (yield: 5.16 g, 9.83 mmol, 89%). 1H NMR (400 MHz, $CDCl_3$) δ = 7.76 (d, J = 7.48 Hz, 2H), 7.60 (d, J = 7.41 Hz, 2H), 7.39 (t, J = 7.46, 7.46 Hz, 2H), 7.31 (dt, J = 7.43, 7.43, 0.98 Hz, 2H), 6.48 (d, J = 8.84 Hz, 1H), 5.58 (d, J = 8.70 Hz, 1H), 4.49 (dd, J = 9.00, 1.68 Hz, 1H), 4.44-4.33 (m, 2H), 4.28-4.15 (m, 3H), 3.71 (s, 3H), 1.94-1.83 (m, 1H), 1.65-1.53 (m, 1H), 1.33-1.21 (m, 1H), 1.17 (d, J = 6.27 Hz, 3H), 1.11 (s, 9H), 1.03-0.93 (m, 6H) ppm. 13C NMR (100 MHz, $CDCl_3$) δ = 171.426, 170.868, 156.074, 143.910, 143.784, 141.249, 127.635, 127.017, 125.077, 119.904, 74.215, 67.193, 66.969, 59.307, 57.832, 52.135, 47.173, 38.179, 28.272, 24.820, 21.046, 15.085, 11.521 ppm.

Boc-Ile-Ile-Thr(tBu)-$NHNH_2$

Fmoc-Ile-Thr(tBu)-OMe (5.16 g, 9.83 mmol) was dissolved in DMF (50 mL) and DBU (1.57 mL, 10.3 mmol, 1.05 equiv.) was added. The reaction was stirred for 5 minutes after which TLC analysis showed complete removal of the Fmoc group. Next, HOBt (1.98 g, 14.7 mmol, 1.5 equiv.) was added and the reaction mixture was stirred for another 30 minutes. To this mixture were added Boc-L-isoleucine (2.73 g, 11.8 mmol, 1.2 equiv.), HCTU (4.88 g, 11.8 mmol, 1.2 equiv.) and DiPEA (4.87 mL, 29.5 mmol, 3 equiv.). The mixture was stirred for 16 hours after which TLC analysis indicated a completed reaction. The mixture was concentrated in vacuo, dissolved in DCM and extracted with 1 M HCl (2x), saturated $NaHCO_3$ (2x) and brine. The organic layer was dried ($MgSO_4$) and concentrated under reduced pressure. Purification of the product by column chromatography (10% → 50% EtOAc/petroleum ether) gave Boc-Ile-Ile-Thr(tBu)-OMe as a colorless solid (yield: 3.69 g, 7.15 mmol, 73%). LC-MS: gradient 10% → 90% ACN/(0.1% TFA/H_2O): Rt (min): 9.88 (ESI-MS (m/z): 516.13 (M + H+)). The obtained product was dissolved in MeOH (50 mL) and hydrazine hydrate (10.4 mL, 214.5 mmol, 30 equiv.) was added. The reaction mixture was refluxed for 16 hours after which TLC analysis indicated complete conversion. Toluene was added and the mixture was concentrated under reduced pressure. Traces of hydrazine were

removed by co-evaporating the mixture with toluene (3x) and the title compound was obtained as a colorless solid (yield: 6.67 g, 7.15 mmol, quant.). ^{1}H NMR (400 MHz, MeOD) δ = 4.36 (d, J = 3.53 Hz, 1H), 4.32 (d, J = 8.12 Hz, 1H), 4.07-4.00 (m, 1H), 3.94 (d, J = 7.90 Hz, 1H), 1.93-1.84 (m, 1H), 1.83-1.73 (m, 1H), 1.61-1.50 (m, 2H), 1.44 (s, 9H), 1.19 (s, 9H), 1.19-1.16 (m, 2H), 1.10 (d, J = 6.32 Hz, 3H), 0.94-0.87 (m, 12H) ppm. ^{13}C NMR (100 MHz, MeOD) δ = 174.839, 173.393, 171.301, 157.910, 80.568, 75.849, 68.522, 60.624, 59.227, 58.566, 37.949, 37.852, 28.772, 28.668, 25.941, 19.781, 16.231, 15.951, 11.392, 11.325 ppm. LC-MS: gradient 10% → 90% ACN/(0.1% TFA/H2O): R_t (min): 6.08 (ESI-MS (m/z): 516.4 ($M + H^+$)).

Boc-Ile-Ile-Thr(tBu)-leucinyl-(R)-2-methyloxirane

Boc-Ile-Ile-Thr(tBu)-$NHNH_2$ (2.0 g, 3.87 mmol) was dissolved in DCM (40 mL) and cooled to -30°C under an argon atmosphere. tBuONO (566 µL, 4.25 mmol, 1.1 equiv.) and HCl (2.8 equiv., 10.8 mmol, 2.7 mL of a 4 M solution in 1,4-dioxane) were added and the mixture was stirred at -30 °C for 3 hours. (Boc-leucinyl)-(R)-2-methyloxirane (1.16 g, 4.25 mmol, 1.1 equiv.) was deprotected with DCM/TFA (1:1 v/v, 20 mL) for 30 minutes followed by co-evaporation with toluene (3x). The resulting TFA salt was dissolved in DMF (5 mL) and added to the former reaction mixture together with DiPEA (3.31 mL, 20 mmol, 5 equiv.). The reaction mixture was slowly warmed to ambient temperature and stirred for 16 hours. Next, the mixture was extracted with 1 M HCl (2x), H_2O and brine, dried ($MgSO_4$) and concentrated in vacuo. The title compound was obtained after column chromatography (20% → 50% EtOAc/petroleum ether) as a colorless solid (yield: 2.25 g, 3.43 mmol, 89%). ^{1}H NMR (400 MHz, $CDCl_3$) δ = 7.64 (d, J = 7.47 Hz, 1H), 6.99 (d, J = 5.64 Hz, 1H), 6.45 (d, J = 8.20 Hz, 1H), 5.22 (d, J = 7.85 Hz, 1H), 4.46 (ddd, J = 10.45, 7.55, 2.94 Hz, 1H), 4.40-4.32 (m, 2H), 4.14-4.07 (m, 1H), 3.94 (t, J = 7.34, 7.34 Hz, 1H), 3.38 (d, J = 5.07 Hz, 1H), 2.89 (d, J = 5.06 Hz, 1H), 1.93-1.77 (m, 2H), 1.74-1.64 (m, 1H), 1.60-1.55 (m, 1H), 1.52 (s, 3H), 1.51-1.46 (m, 2H), 1.44 (s, 9H), 1.28 (s, 9H), 1.27-1.24 (m, 1H), 1.17-1.08 (m, 2H), 1.06 (d, J = 6.44 Hz, 3H), 0.96 (d, J = 6.54 Hz, 6H), 0.92-0.86 (m, 12H) ppm. ^{13}C NMR (100 MHz, $CDCl_3$) δ = 208.062, 171.593, 170.738, 169.515, 155.807, 79.761, 75.492, 66.143, 59.249, 57.686, 56.956, 52.395, 50.746, 39.809, 37.300, 36.971, 28.280, 28.082, 25.423, 24.879, 24.695, 23.358, 21.359, 16.754, 15.532, 15.405, 11.285 ppm. LC-MS: gradient 10% → 90% ACN/(0.1% TFA/H_2O): R_t (min): 11.31 (ESI-MS (m/z): 655.27 ($M + H^+$)).

Biotin-epoxomicin (3)

Boc-Ile-Ile-Thr(tBu)-leucinyl-(R)-2-methyloxirane (13.2 mg, 20.2 µmol) was dissolved in 2 mL DCM. TFA (2 mL) was added and the mixture was stirred for 20 min. The reaction mixture was co-evaporated with toluene (3x). The residue was dissolved in 1 mL DMF. Biotin-OSu (7 mg, 21 µmol, 1.01 equiv.) and DiPEA (8.3 µL, 50 µmol, 2.5 equiv.) were added and the mixture was stirred for 2 hr. The volatiles were removed in vacuo and the title compound was obtained after HPLC purification (yield: 5 mg, 6.9 µmol, 34%). ^{1}H NMR (400 MHz, MeOD) δ = 4.55 (dd, J = 10.63, 3.03 Hz, 1H), 4.48 (dd, J = 7.72, 4.85 Hz, 1H), 4.32-

4.20 (m, 4H), 4.06-3.99 (m, 2H), 3.25 (d, J = 5.07 Hz, 1H), 3.23-3.16 (m, 1H), 2.95-2.89 (m, 2H), 2.69 (d, J = 12.71 Hz, 1H), 2.33-2.20 (m, 2H), 1.90-1.77 (m, 2H), 1.78-1.30 (m, 13H), 1.24-1.11 (m, 5H), 0.95-0.86 (m, 18H) ppm. LC-MS: gradient 10% → 90% ACN/(0.1% TFA/H_2O): R_t (min): 6.30 (ESI-MS (m/z): 725.7 $(M + H^+)$).

Azido-BODIPY-epoxomicin

Boc-Ile-Ile-Thr(tBu)-leucinyl-(R)-2-methyloxirane (7.9 mg, 12 µmol) was dissolved in TFA (1 mL) and stirred for 30 min., before being coevaporated with toluene (3). The residue was dissolved in DMF (2 mL) and azido-BODIPY-OSu (6.6 mg, 12µmol, 1 equiv.) and DiPEA (8 µL, 48 µmol, 4 equiv.) were added and the reaction mixture was stirred for 12 hr. Concentration in vacuo, followed by purification by column chromatography (DCM → 2% MeOH/DCM) yielded the title compound as a brown/red solid (yield: 5.4 mg, 5.7 µmol, 47%). ^{1}H NMR (600 MHz, MeOD) δ = 7.88 (d, J = 8.7 Hz, 2H), 7.41 (s, 1H), 7.06 (d, J = 3.9 Hz, 1H), 6.99 (d, J = 8.7 Hz, 2H), 6.60 (d, J = 3.9 Hz, 1H), 4.55 (dd, J1 = 10.7, J2 = 2.8 Hz, 1H), 4.30 (d, J = 5.0 Hz, 1H), 4.22 (d, J = 7.8 Hz, 1H), 4.15-4.12 (m, 3H), 4.02 (p, J = 6.1 Hz, 1H), 3.54 (t, J = 6.7 Hz, 2H), 3.25 (d, J = 5.1 Hz, 1H), 2.92 (d, J = 5.1 Hz, 1H), 2.81 (m, 1H), 2.71 (m, 1H), 2.51 (s, 3H), 2.45-2.40 (m, 2H), 2.25 (s, 3H), 2.07 (p, J = 6.3 Hz, 2H), 1.89-1.79 (m, 1H), 1.75-1.66 (m, 2H), 1.65-1.52 (m, 2H), 1.53-1.41 (m, 5H), 1.41-1.21 (m, 15H), 1.20-1.06 (m, 5H), 1.05-0.97 (m, 1H), 0.97-0.85 (m, 16H), 0.82 (d, J = 6.7 Hz, 3H), 0.76 (t, J = 7.4 Hz, 3H) ppm. ^{13}C NMR (150 MHz, MeOD) δ = 209.51, 174.86, 174.06, 173.59, 172.23, 161.03, 160.67, 156.57, 141.79, 136.67, 135.83, 132.45, 131.92, 131.89, 131.86, 131.67, 131.65, 129.91, 129.28, 127.16, 124.70, 119.10, 115.27, 115.19, 69.14, 68.55, 65.98, 60.13, 59.82, 59.42, 59.41, 53.10, 51.84, 40.38, 38.02, 37.71, 36.45, 30.82, 29.90, 26.26, 26.03, 23.81, 21.52, 21.21, 20.02, 17.05, 15.92, 15.86, 11.47, 11.22, 9.67 ppm.

Biotin-BODIPY(Tmr)-epoxomicin (2)

Azido-BODIPY(Tmr)-epoxomicin (4.1 mg, 4.3 µmol) and Biotin-propargylamide (2.4 mg, 8.6 µmol, 2 equiv.) were dissolved in tBuOH (0.25 mL) and toluene (0.25 mL) before CuSO4 (125 µL 3.4 mM, 10 mol%) and sodium ascorbate (125 µL 6.9 mM, 20 mol%) were added. The reaction mixture was stirred at 80 °C for 12 hr., before being cooled to room temperature and concentrated in vacuo. Purification by column chromatography (petroleum ether → 50% acetone/petroleum ether) yielded the title compound as a brown/red solid (4.5 mg, 3.7 µmol, 85%). ^{1}H NMR (600 MHz, MeOD) δ = 7.95-7.78 (m, 3H), 7.42 (s, 1H), 7.07 (d, J = 4.1 Hz, 1H), 6.95 (d, J = 8.9 Hz, 2H), 6.61 (d, J = 4.1 Hz, 1H), 4.70-4.52 (m, 5H), 4.46-4.39 (m, 2H), 4.34-4.26 (m, 1H), 4.25-4.19 (m, 1H), 4.17-4.11 (m, 1H), 4.08-3.99 (m, 3H), 3.95 (t, J = 2.2 Hz, 1H), 3.25 (d, J = 5.0 Hz, 1H), 3.16-3.10 (m, 1H), 2.92 (d, J = 5.1 Hz, 1H), 2.71-2.64 (m, 2H), 2.60-2.56 (m, 1H), 2.51 (s, 3H), 2.46-2.37 (m, 4H), 2.26 (s, 3H), 2.24-2.17 (m, 2H), 1.95-1.21 (m, 32H), 1.21-1.10 (m, 5H), 1.06-0.85 (m, 17H), 0.82 (d, J = 6.8 Hz, 3H), 0.76 (t, J = 7.3 Hz, 3H) ppm.

6.4.3 Activity-based protein profiling

Tissues were homogenized in 3 volumes of ice cold lysis buffer (50 mM TrisHCl pH 7.5, 250 mM sucrose, 5mM $MgCl_2$, 1mM DTT, 2mM ATP, 0.025% digitonin, 0.2% NP40, (*21*)) with a tissue homogenizer and further disrupted by 2 x 30 sec sonication. Lysates were cleared by cold centrifugation at 13,000 g, protein concentrations determined by Bradford assay and kept at -80°C until use. For comparative activity based profiling, equal amounts of protein were incubated with ABPs for 1 hr at 37°C, resolved by 12.5% SDS-PAGE and the wet gel slab was scanned on a Thyphoon scanner (GE Healthcare) with the TAMRA settings (λex=530 nm, λem=560 nm). Competitive activity based profiling was done by first incubating thymus lysates with increasing concentrations of various proteasome inhibitors for 1 hr at 37°C, followed by 1 hr incubation with 0.5 µM ABP **4** for the in-gel detection of the residual proteasome activity. Images were acquired, processed and quantified with Image Quant (GE Healthcare).

6.4.4 Affinity purification

Some 1 or 2 mg of protein was incubated with 10 µM biotinylated ABPs **1**, **2** or **3** for 1 hr at 37°C, denatured by boiling for 5 min with 1% SDS and precipitated with chloroform/ methanol (C/M, (*22*)). The protein pellet was rehydrated in 180 µl 8M urea/100 mM NH_4HCO_3, reduced with 10 µl 90 mM DTT for 30 min at 37°C, alkylated with 15 µl 200 mM iodoacetamide at RT in the dark, cleared by centrifugation at 13,000 g and desalted by C/M. The pellet was dispersed in 25 µl PD buffer (50 mM TrisHCl pH7.5, 150 mM NaCl) with 2% SDS in a heated (37°C) sonic bath. Stepwise (3 x 25 µl, 4 x 100 µl, 1 x 500 µl) addition of PD buffer afforded a clear solution that was incubated with 50 µl MyOne T1 Streptavidin grafted beads (Invitrogen) at RT with vigorous shaking for 2 hr. The beads were stringently washed with 2 x 300µl PD buffer with 0.1% SDS, 2 x 300 µl PD buffer, 2 x 300µl wash buffer I (4M urea/50 mM NH_4HCO_3), 2 x 300µl wash buffer II (50 mM TrisHCl pH7.5, 10 mM NaCl) and 2 x 300 µl water. For in-gel analysis, 2/3 of the beads was eluted with 100 µl 1x sample buffer containing 10 µM biotin by boiling for 5 min at 90°C and resolved by 12.5% SDS-PAGE. Proteins were visualized by fluorescence and silverstain, in-gel digested and desalted (*23, 24*). For on/bead digest, 1/3 of the beads was digested with 300 ng trypsin in 100 µl digest buffer (100 mM TrisHCl pH 7.8, 100 mM NaCl, 1mM $CaCl_2$, 2% ACN) o.n. at 37°C. Peptides were collected and desalted on stage tips. The active-site peptides were eluted with 2 x 80 µl 10 µM biotin in 5% formic acid/25% ACN/70% H_2O for 30 min at 37°C and desalted after ACN evaporation.

6.4.5 LC-MS analysis

Trytic peptides were analyzed on a Surveyor nanoLC system (Thermo) hyphenated to a LTQ-Orbitrap mass spectrometer (Thermo). Gold and carbon coated emitters

(OD/ID=360/25µm tip ID=5 µm), trap column (OD/ID=360/100 µm packed with 25 mm robust Poros®10R2/ 15 mm BioSphere C18 5 µm 120Å) and analytical columns (OD/ID=360/75µm packed with 20 cm BioSphere C18 5 µm 120Å) were from Nanoseparations (Nieuwkoop, The Netherlands). The mobile phases (A: 0.1% FA/H_2O, B: 0.1%FA/ACN) were made with ULC/MS grade solvents (Biosolve). The emitter tip was coupled end-to-end with the analytical column via a 15 mm long TFE teflon tubing sleeve (OD/ID 0.3x1.58 mm, Supelco, USA) and installed in a stainless steel holder mounted in a nano-source base (Upchurch scientific, Idex, USA).

General mass spectrometric conditions were: an electrospray voltage of 1.8 kV was applied to the emitter, no sheath and auxiliary gas flow, ion transfer tube temperature 150°C, capillary voltage 41V, tube lens voltage 150V. Internal mass calibration was performed with air-borne protonated polydimethylcyclosiloxane (m/z = 445.12002) and the plasticizer protonated dioctyl phthalate ions (m/z = 391.28429) as lock mass(*25*).

For shotgun proteomics analysis, 10 µl of the samples was pressure loaded on the trap column with a 10 µl/min flow for 5 min followed by peptide separation with a gradient of 35 min 5-30% B, 15 min 30-60% B, 5 min A at a flow of 300 µl/min split to 250 nl/min by the LTQ divert valve. For each data dependent cycle, one full MS scan (300-2000 m/z) acquired at high mass resolution (60,000 at 400 m/z, AGC target $1x10^6$, maximum injection time 1,000 ms) in the Orbitrap was followed by 3 MS/MS fragmentations in the LTQ linear ion trap (AGC target $5x10^3$, max inj time 120 ms) from the three most abundant ions(*26*). MS^2 settings were: collision gas pressure 1.3 mT, normalized collision energy 35%, ion selection threshold of 500 counts, activation q = 0.25 and activation time of 30 ms. Fragmented precursor ions that were measured twice within 10 s were dynamically excluded for 60s and ions with z<2 or unassigned were not analyzed.

A parent ion list of the m/z ratios of the active-site peptides was compiled and used for LC-MS^3 analysis in a data dependent protocol. The parent ion was electrostatically isolated in the ion trap of the LTQ, fragmented by MS^2 and the most intense peak was isolated and further fragemented in MS^3 to reveal the amino acid sequence of the active-site peptide. Data from MS^2 and MS^3 was validated manually.

References and notes

1. Janeway, C. A., Jr., and Bottomly, K. (1994) Signals and signs for lymphocyte responses, *Cell 76*, 275-285.

2. Medzhitov, R. (2007) Recognition of microorganisms and activation of the immune response, *Nature 449*, 819-826.

3. Huseby, E. S., White, J., Crawford, F., Vass, T., Becker, D., Pinilla, C., Marrack, P., and Kappler, J. W. (2005) How the T cell repertoire becomes peptide and MHC specific, *Cell 122*, 247-260.

4. Takahama, Y., Tanaka, K., and Murata, S. (2008) Modest cortex and promiscuous medulla for thymic repertoire formation, *Trends Immunol 29*, 251-255.

5. Kloetzel, P. M., and Ossendorp, F. (2004) Proteasome and peptidase function in MHC-class-I-mediated antigen presentation, *Curr Opin Immunol 16*, 76-81.

6. Nitta, T., Murata, S., Ueno, T., Tanaka, K., and Takahama, Y. (2008) Thymic microenvironments for T-cell repertoire formation, *Adv Immunol 99*, 59-94.

7. Murata, S., Sasaki, K., Kishimoto, T., Niwa, S., Hayashi, H., Takahama, Y., and Tanaka, K. (2007) Regulation of CD8+ T cell development by thymus-specific proteasomes, *Science 316*, 1349-1353.

8. Baumeister, W., Walz, J., Zuhl, F., and Seemuller, E. (1998) The proteasome: paradigm of a self-compartmentalizing protease, *Cell 92*, 367-380.

9. Murata, S., Takahama, Y., and Tanaka, K. (2008) Thymoproteasome: probable role in generating positively selecting peptides, *Curr Opin Immunol 20*, 192-196.

10. Cravatt, B. F., Wright, A. T., and Kozarich, J. W. (2008) Activity-based protein profiling: from enzyme chemistry to proteomic chemistry, *Annu Rev Biochem 77*, 383-414.

11. Verdoes, M., Florea, B. I., van der Linden, W. A., Renou, D., van den Nieuwendijk, A. M., van der Marel, G. A., and Overkleeft, H. S. (2007) Mixing of peptides and electrophilic traps gives rise to potent, broad-spectrum proteasome inhibitors, *Org Biomol Chem 5*, 1416-1426.

12. Verdoes, M., Florea, B. I., van der Marel, G. A., and Overkleeft, H. S. (2009) Chemical Tools To Study the Proteasome, *Eur J Org Chem*, 3301-3313.

13. Verdoes, M., Florea, B. I., Menendez-Benito, V., Maynard, C. J., Witte, M. D., van der Linden, W. A., van den Nieuwendijk, A. M., Hofmann, T., Berkers, C. R., van Leeuwen, F. W., Groothuis, T. A., Leeuwenburgh, M. A., Ovaa, H., Neefjes, J. J., Filippov, D. V., van der Marel, G. A., Dantuma, N. P., and Overkleeft, H. S. (2006) A fluorescent broad-spectrum proteasome inhibitor for labeling proteasomes in vitro and in vivo, *Chem Biol 13*, 1217-1226.

14. Kessler, B. M., Tortorella, D., Altun, M., Kisselev, A. F., Fiebiger, E., Hekking, B. G., Ploegh, H. L., and Overkleeft, H. S. (2001) Extended peptide-based inhibitors efficiently target the proteasome and reveal overlapping specificities of the catalytic beta-subunits, *Chem Biol 8*, 913-929.

15. Carey, F. A., and Sundberg, R. J. (2007) *Advanced organic chemistry*, 5th ed., Springer, New York.

16. Fenteany, G., Standaert, R. F., Lane, W. S., Choi, S., Corey, E. J., and Schreiber, S. L. (1995) Inhibition of proteasome activities and subunit-specific amino-terminal threonine modification by lactacystin, *Science 268*, 726-731.

17. Rock, K. L., and Goldberg, A. L. (1999) Degradation of cell proteins and the generation of MHC class I-presented peptides, *Annu Rev Immunol 17*, 739-779.

18. Adams, J., Behnke, M., Chen, S., Cruickshank, A. A., Dick, L. R., Grenier, L., Klunder, J. M., Ma, Y. T., Plamondon, L., and Stein, R. L. (1998) Potent and selective inhibitors of the proteasome: dipeptidyl boronic acids, *Bioorg Med Chem Lett 8*, 333-338.

19. Britton, M., Lucas, M. M., Downey, S. L., Screen, M., Pletnev, A. A., Verdoes, M., Tokhunts, R. A., Amir, O., Goddard, A. L., Pelphrey, P. M., Wright, D. L., Overkleeft, H. S., and Kisselev, A. F. (2009) Selective inhibitor of proteasome's caspase-like sites sensitizes cells to specific inhibition of chymotrypsin-like sites, *Chem Biol 16*, 1278-1289.

20. Bogyo, M., McMaster, J. S., Gaczynska, M., Tortorella, D., Goldberg, A. L., and Ploegh, H. (1997) Covalent modification of the active site threonine of proteasomal beta subunits and the Escherichia coli homolog HslV by a new class of inhibitors, *Proc Natl Acad Sci U S A 94*, 6629-6634.

21. Kisselev, A. F., and Goldberg, A. L. (2005) Monitoring activity and inhibition of 26S proteasomes with fluorogenic peptide substrates, *Methods Enzymol 398*, 364-378.

22. Wessel, D., and Flugge, U. I. (1984) A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids, *Anal Biochem 138*, 141-143.

23. Shevchenko, A., Tomas, H., Havlis, J., Olsen, J. V., and Mann, M. (2006) In-gel digestion for mass spectrometric characterization of proteins and proteomes, *Nat Protoc 1*, 2856-2860.

24. Rappsilber, J., Mann, M., and Ishihama, Y. (2007) Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips, *Nat Protoc 2*, 1896-1906.

25. Olsen, J. V., de Godoy, L. M., Li, G., Macek, B., Mortensen, P., Pesch, R., Makarov, A., Lange, O., Horning, S., and Mann, M. (2005) Parts per million mass accuracy on an Orbitrap mass spectrometer via lock mass injection into a C-trap, *Mol Cell Proteomics 4*, 2010-2021.

26. Baek, D., Villen, J., Shin, C., Camargo, F. D., Gygi, S. P., and Bartel, D. P. (2008) The impact of microRNAs on protein output, *Nature 455*, 64-71.

Supplementary data

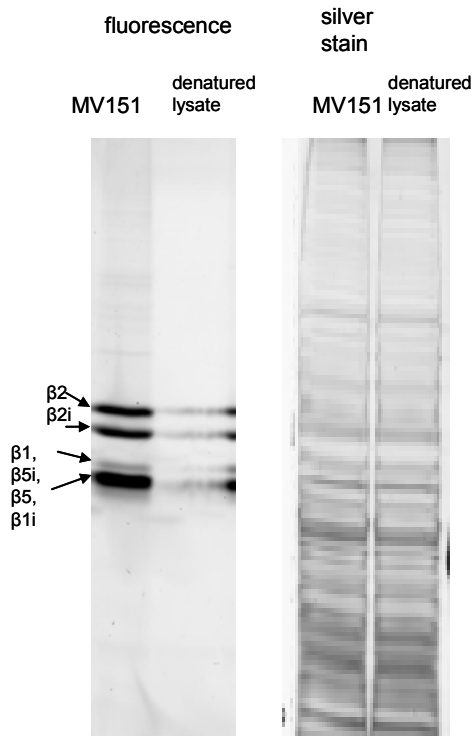


Figure *S1*. Fluorescence and silver stain detection of EL4 (murine B cell lymphoma cell line) cell lysate incubated with the fluorescent, broad-spectrum proteasome activity-based probe MV151. Some 20 µg protein was incubated with 0.5 µM MV151 for 60 min at 37°C, resolved by 12.5% SDS-PAGE and imaged by fluorescence scanning followed by silver staining of the same gel. In the denatured lane, the lysate was deactivated by boiling with 1% SDS prior to the MV151 incubation.

		mass	cover			pept	
prot acc		(Da)	% AA	z	ppm	score	peptide sequence
Psmb11	(β5t)	27834	34	2	-0.89	47	HGVIAAADTR
IPI00221461				2	1.30	21	EGQLPSVAGTAK
				2	0.41	76	LLAAMMSCYR
				2	-2.14	93	SSCGSYVACPASR
				2	0.76	71	ACGIYPEPATPQGAR
				2	1.67	130	LLGTTSGTSADCATWYR
				2	1.89	99	ELFVEQEEVTPEDCAIIMK
Psmb7	(β2)	25235	50	2	0.43	27	QMLFR
IPI00136483				2	1.09	70	FRPDMEEEEAK
				2	3.42	55	LDFLRPFSVPNK

				2	-0.80	57	FRPDMEEEEAKK *
				3	0.13	45	LDFLRPFSVPNKK *
				3	2.19	38	SKLDFLRPFSVPNK *
				2	-4.89	62	LPYVTMGSGSLAAMAVFEDK
				2	-1.17	111	VTPLEIEVLEETVQTMDTS #
				2	-0.11	129	LVSEAIAAGIFNDLGSGSNIDLCVISK
				3	3.18	57	KLVSEAIAAGIFNDLGSGSNIDLCVISK *
				3	0.28	173	IHFISPNIYCCGAGTAADTDMTTQLISSNLELHSLTTGR
Psmb10	(β2i)	24789	65	2	-0.29	61	ATNDSVVADK
IPI00316736				2	0.00	40	MELHALSTGR
				2	-0.60	39	FAPGTTPVLTR
				2	-1.12	121	IYCCGAGVAADTEMTTR
				2	0.35	167	LPFTALGSGQGAAVALLEDR
				2	0.73	93	EVRPLTLELLEETVQAMEVE #
				4	0.77	53	YQGHVGASLVVGGVDLNGPQLYEVHPHGSYSR
Psmb6	(β1)	21982	74	2	0.35	58	DGSSGGVIR
IPI00119239				2	-0.21	50	FTIATLPPP #
				2	0.55	78	TTTGSYIANR
				2	1.28	91	LAAIQESGVER
				2	-1.59	55	LTPIHDHIFCCR
				2	1.59	132	DECLQFTANALALAMER
				2	1.27	129	QSFAIGGSGSSYIYGYVDATYR
				3	1.68	32	EGMTKDECLQFTANALALAMER *
				3	1.68	106	YREDLMAGIIIAGWDPQEGGQVYSVPMGGMMVR *
				3	4.81	149	SGSAADTQAVADAVTYQLGFHSIELNEPPLVHTAASLFK
Psmb8	(β5i)	22635	55	2	-0.27	92	ATAGSYISSLR
IPI00116712				2	2.60	66	LLSNMMLQYR
				2	-0.68	72	FQHGVIVAVDSR
				2	-0.74	84	VESSDVSDLLYK
				2	0.56	80	GPGLYYVDDNGTR
				2	1.18	75	DNYSGGVVNMYHMK
				2	1.35	99	GMGLSMGSMICGWDK
				2	-0.75	121	LSGQMFSTGSGNTYAYGVMDSGYR

				3	0.81	60	VIEINPYLLGTMSGCAADCQYWER
Psmb5	(β5)	22514	18	2	0.38	47	VEEAYDLAR
IPI00317902				2	0.91	56	GPGLYYVDSEGNR
Psmb9	(β1i)	21313	58	2	0.21	56	VSAGTAVVNR
IPI00309379				2	-0.71	61	VILGDELPK
				2	-0.07	91	FTTNAITLAMNR
				2	0.05	96	DGSSGGVIYLVTITAAGVDHR
				3	-0.15	99	QPFTIGGSGSSYIYGYVDAAYKPGMTPEECR
				3	-1.11	122	IFCALSGSAADAQAIADMAAYQLELHGLELEEPPLVLAAANVVK

Table *S1*. Protein identification after affinity purification with probe **3**, on-bead digestion with trypsin and LC-MS analysis.

Protein name, mass of the active β subunit, % coverage of the protein by amino acids identified by LC-MS, charge of the peptide (z), measurement error (ppm), Mascot peptide scores, miss cleavage (*), and C-terminal peptides (#). Mascot identifications were manually validated.

		Exact mass		*z*=2		*z*=3	
	y_7 ion sequence	mono-iso	High-peak	mono-iso	High-peak	mono-iso	High-peak
β1	TTIMAVQFNGGVVLGADSR	2659.40773	2660.41061	1330.71114	1331.21258	887.47652	887.81081
β1i	TTIMAVEFDGGVVVGSDSR	2663.35502	2664.35793	1332.68479	1333.18624	888.79228	889.12659
β2	TTIAGVVYK	1674.96302	1674.96302	838.48878	838.48878	559.32828	559.32828
β2i	TTIAGLVFR	1700.98990	1700.98990	851.50223	851.50223	568.00391	568.00391
β5	TTTLAFK	1504.85749	1504.85749	753.43602	753.43602	502.64644	502.64644
β5i	TTTLAFK	1504.85749	1504.85749	753.43602	753.43602	502.64644	502.64644
β5t	TTTLAFR	1532.86364	1532.86364	767.43909	767.43909	511.96182	511.96182

Table *S2*: Calculated exact (*m/z*) masses of the active-site peptides bound to biotin-epoxomicin (probe **3**).

The mono-isotopic mass (mono-iso) and the mass of the most abundant isotope peak (High-peak) are shown at charge (*z*) of 0, 2, and 3. The active site peptide sequence of β5 and β5i is identical.

O-GlcNAcylation and hHR23B functions

7.1 Introduction

Protein O-linked GlcNAcylation (O-GlcNAc) is a post-translational modification (PTM) characterized by the covalent and reversible bonding of a N-acetylglucosamine moiety to the γ-hydroxyl of serine or threonine residues of a protein via a β-C2 linkage. It bears resemblance to protein phosphorylation and O-GlcNAcylation levels can respond rapidly to intracellular or environmental cues(*1*). The level of O-GlcNAc is regulated by O-GlcNAc transferase (OGT) and O-GlcNAcase (OGA), and is highly dependent on the level of its donor substrate, UDP-N-acetylglucosamine (UDP-GlcNAc)(*1, 2*). UDP-GlcNAc levels are regulated by an interplay of the hexosamine biosynthetic pathway (HBP) and the metabolic pathways of amino acid, nucleotide, fatty acid, and glucose(*3*). Therefore, protein O-GlcNAcylation levels can vary significantly according to the cell or tissue type, and the nutrition state of the cell. Hundreds of proteins are found to be O-GlcNAcylated, and the list is still increasing(*4, 5*). They belong to diverse families, including metabolic enzymes, transcription factors, heat-shock proteins, and architectural proteins.

The consequence of protein O-GlcNAcylation is complicated and highly dependent on the specific biological network that the protein is being involved in. The physical presence of O-GlcNAcylation alone could affect protein-protein interaction (PPI). For example, O-GlcNAcylation on transcription factor SP1 affects its binding to NF-Y, which is another important transcription factor in complex with SP1 to synergistically activate transcription of many genes, and further interrupts the transcription regulated by SP1 and NF-Y(*6*). Oligomerization of tau is a key process contributing to the progressive death of neurons in Alzheimer's disease. Treatment of tau transgenic mice with an O-GlcNAcase inhibitor increased O-GlcNAcylation level of tau in vivo and decreased neuronal cell loss. It has been proven that O-GlcNAcylation on Tau prevents it from aggregation(*7*). Alternatively, O-GlcNAcylation can affect the PTM state of a given protein. O-GlcNAcylation might show overlapping modification sites with protein phosphorylation on the same or adjacent amino acid residues. In the case of transcription regulation,

phosphorylation activates δ-lactoferrin and P53 but also makes them susceptible to ubiquitination. O-GlcNAcylation inhibits δ-lactoferrin and P53 phosphorylation under normal conditions, thus suppressing them from activation and protecting them from degradation by the proteasome(*8*, *9*). O-GlcNAcylation has also been shown to reduce ubiquitination through the recruitment of deubiquitin enzymes (DUBs) on the target proteins, as shown in the case of Bmal1/Clock in the regulation of the circadian clock(*10*, *11*). The lectin property of heat-shock protein Hsp70 was proposed to allow Hsp70 bind to O-GlcNAcs on the damaged protein, thus prevent them from ubiquitination and degradation under hyperthermia and UV stress(*12*, *13*). Due to the diverse outcomes of protein O-GlcNAcylation, the effect of O-GlcNAcylation on a given protein should be examined by the change in interaction between that protein and its binding partners, directly or indirectly through the involvement of the regulation of other PTMs.

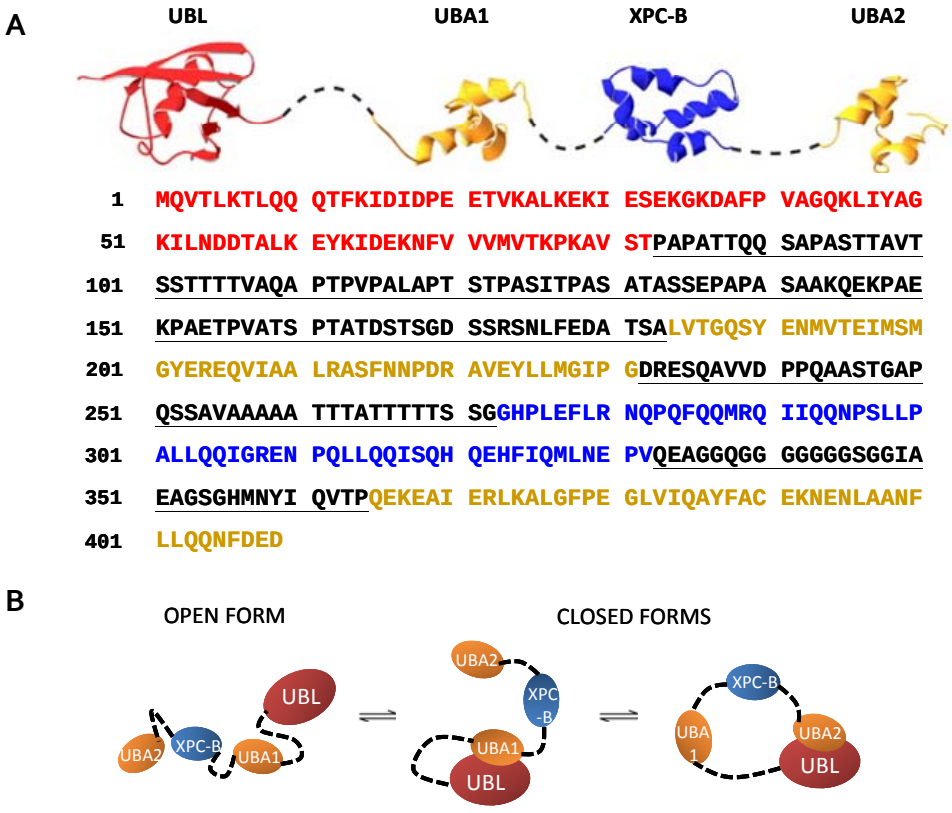


Figure 1: Schematic presentation of hHR23B protein.

(A) The upper panel is a scheme of the hHR23B functional domains, UBL (ubiquitin like domain), UBA1(ubiquitin associate domain 1), XPC-B (XPC binding domain) and UBA2 (ubiquitin associate domain 2). The lower panel is protein sequence of hHR23B. The linker sequences are shown underlined.

(B) Models of open or closed conformations of hHR23B. The UBL domain can interact with UBA1 or UBA2 domains intra-molecularly. When they do not interact with each other, the protein shows an open form. Or, it shows the closed form.

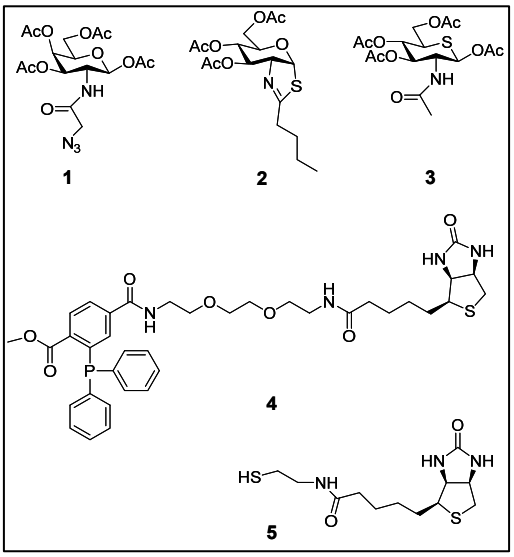
In previous O-GlcNAcylation profiling experiments using a metabolic labeling method(*4*), one of the modified proteins identified was the ubiquitin receptor protein hHR23B. hHR23B is a multi-functional scaffold protein involved in two important biological pathways, the global genomic nucleotide excision repair (GG-NER) and the ubiquitin proteasome system (UPS)(*15*). Functionally, it is often referred to as "NER accessory protein" for its role in stabilizing XPC during the recognition of nucleotide lesions, or "ubiquitin shuttle protein" for its role in trafficking Lys-48 linked polyubiquitinated proteins to the proteasome for their controlled degradation. hHR23B is evolutionarily conserved, as it shares an overall structural and functional similarity to its human paralog hHR23A, and the yeast ortholog Rad23(*15*). Starting from the N-terminus, hHR23B consists of four intra-molecular domains, namely the UBL (ubiquitin like), UBA1 (ubiquitin associate 1), XPC binding, and UBA2 (ubiquitin associate 2) domains (Fig 1A). The functional domains are interconnected by flexible linkers, with the one linking UBL and UBA1 being exceptionally long (78 amino acids). This long linker has a protease resistant amino acid sequence (with very few trypsin (K and R) or chymotrypsin (F, W, L and M) cleavage sites). Meanwhile, it has high Pro, Ala, Ser, Thr content, which is the OGT preference sequence for installing an O-GlcNAc moiety. Presence of a Pro/Ser/Thr rich N terminal and Thr/Ser/Ala rich C terminal adjacent to an O-GlcNAcylation site has been commonly found in O-GlcNAcylated peptides identified before, and it is comparable to the linker sequence in the hHR23B(*2*). Thanks to the long flexible linker regions in the hHR23B, it might adopt an open/closed conformation change just as its paralog hHR23A(*16*). As shown in Fig 1B, the UBL and UBA domains might have some intra-molecular interactions, by which the open/close model has been set up.

The ubiquitin proteasome system is responsible for tightly regulated degradation of specific protein substrates, which is critical in protein turnover, cell cycle progression, antigen presentation, and regulation of oncogenes. Proteins are first tagged with polyubiquitin chains by an enzyme cascade consisting of ubiquitin activating enzyme (E1), ubiquitin conjugating enzyme (E2), and ubiquitin ligase (E3)(*17*). Polyubiquitinated proteins are either recognized and degraded by the proteasome, or shuttled to the proteasome for degradation by the assist of ubiquitin receptor proteins like hHR23B. All four domains of hHR23B might contribute to its role in the UPS. The UBL domain resembles ubiquitin structurally, and can bind to Rpn10 (s5a in yeast) and Rpn1 subunits of the 19S regulatory particle of the 26S proteasome, the main protein degradation machinery in the UPS(*15*). The two UBA domains serve to bind to poly-ubiquitin chains on substrate proteins of the proteasome(*17*). The binding of UBL and UBAs to their binding partners ensures a one-directional shuttling of the poly-ubiquitinated protein substrates to the proteasome for degradation. The XPC binding domain is also involved in the UPS. Although its major function is to bind to XPC and thus stabilize it during the lesion recognition phase of GG-

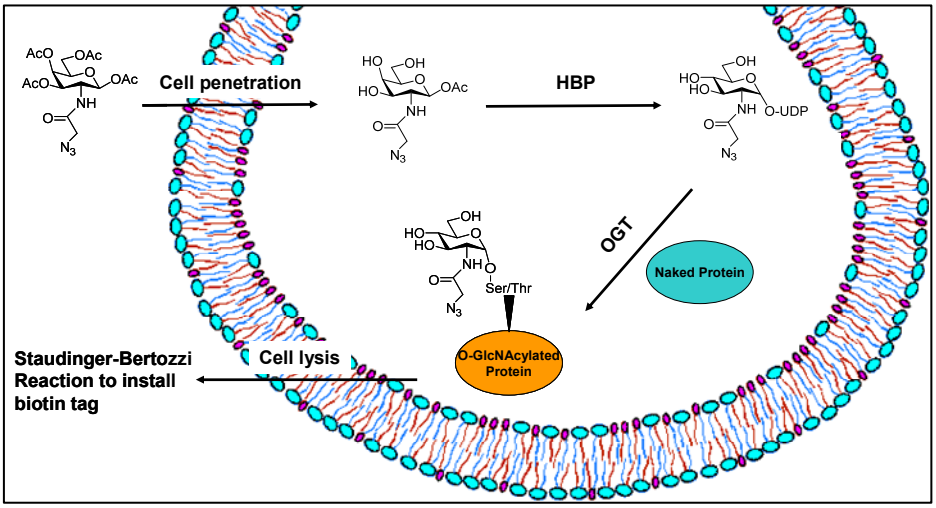
NER, it was found that in yeast the Rad4 (ortholog of XPC in yeast) binding domain of Rad23 also binds to protein glycanase Png1, a protein responsible for the removal of sugar-tagged substrates in the ER-associated degradation (ERAD), and that Rad23 is an essential protein in the ERAD pathway(*18*). Therefore, the O-GlcNAc modification on the hHR23B protein might influence the efficiency of the whole UPS.

In this chapter, a series of chemical biology tools are applied for the visualization, site identification, and unraveling the physiology of O-GlcNAcylation on hHR23B. It proves that, hHR23B is indeed modified by O-GlcNAc. Six O-GlcNAcylation sites were identified on the protein, among which four were only observed after chemical OGA inhibition. While, the cellular O-GlcNAcylation level is increased by an OGA inhibitor treatment, there is almost 50% more free hHR23B molecules existing in the cells. It demonstrates that, the O-GlcNAcylation affects the protein-protein interaction of the hHR23B protein.

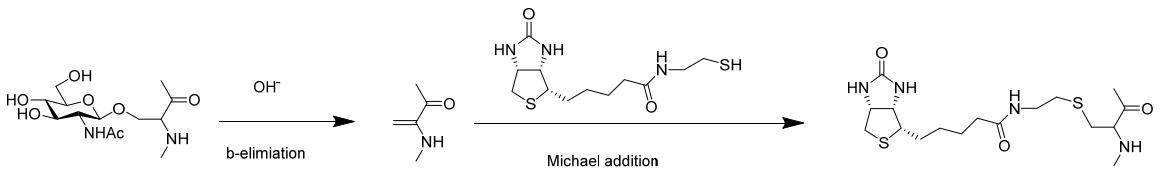
A



B



C



D

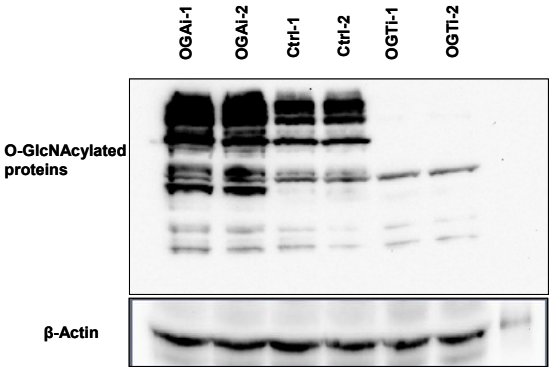


Figure 2: Chemical tools used in this chapter for O-GlcNAcylation study.
(A) Structures of the chemical tools used for the O-GlcNAcylation study.
(B) Schematic presentation of metabolic labeling of O-GlcNAcylated proteins. Cells take up the Ac_4GalNAz in the medium and deacetylate the molecule. GalNAz is converted to UDP-GlcNAz HBP biochemical synthesis enzymes. OGT uses the UDP-GlcNAz as the donor substrate to modify protein substrates. Biotin tag can be installed on the O-GlcNAz modified proteins by bio-orthogonal ligation following lysing the cells.
(C) Schematic presentation of the BEMAB replacement of O-GlcNAc. The labile O-GlcNAc modification on the amino acid residues is removed by a β-elimination under basic conditions, and then biotin-cystamine is added to the amino acid residue by a Michael addition to result in the stable S-linked biotin modification for replacing the O-GlcNAc in LC/MS analysis.
(D) Western blots to test the OGA and OGT inhibitors (**2** and **3**). Cells were treated with or without 50μM **2** or **3** for 24hrs. Full cell lysate were resolved by SDS-PAGE and transferred to PVDF membrane. The O-GlcNAcylated proteins were detected by O-GlcNAc antibody CTD110.6 (upper panel). Anti β-actin blot was performed as loading control (lower panel). Two independent experiments were done and compared on the same blot.

7.2 Results

7.2.1 hHR23B is O-GlcNAcylated

Immunoblot detection is the most straightforward method for proving the existence of O-GlcNAc on the target proteins. To circumvent the disadvantage of sensitivity and specificity of direct visualization by O-GlcNAc antibody, for instance the commonly used O-

GlcNAc antibody CTD110.6 cross reacts with N-GlcNAc$_2$-modified proteins in some cases(*19*), a chemical biological approach was taken. HEK293T cells were transfected with a plasmid (pFLAG-6His-hHR23B) for the expression of Flag tagged hHR23B, followed by treatment with high concentration (250μM) of Ac4-GalNAz (**1** in Fig 2A). The Ac4-GalNAz was deacetylated and converted by the HBP enzymes into UDP-GlcNAz intracellularly, and then OGT used the UDP-GlcNAz as the donor substrate to modify protein substrates. Consequently, proteins got modified by GlcNAz instead of the endogenous GlcNAc(*3*). Then, cells were harvested and lysed. A Staudinger-Bertozzi reaction was performed with biotin-phosphine **4** in the lysate to install a biotin tag on GlcNAz modified proteins and as control for background labeling on cell lysate without GlcNAz exposure (Fig 2B). Afterwards, one immunoprecipitation (IP) against the FLAG tag to purify FLAG-6His-hHR23B and one biotin-streptavidin affinity purification (AP) to enrich for O-GlcNAz modified proteins were done with the lysate. The eluted samples from the IP and AP experiments were detected by both anti-FLAG antibody and streptavidin-HRP, respectively.

As shown in Figure 3A, successful transfection and expression of FLAG-tagged hHR23B was achieved. In all four samples, the bands have sizes of around 57 kD, which is consistent with the molecular weight of FLAG-6His-hHR23B. Lane 1 and lane 2 are the input and elute samples of the anti-FLAG IP without the GalNAz treatment. Lane 3 and lane 4 are the input and elute samples of the IP with GalNAz treatment. The possible presence of O-GlcNAz and biotin on hHR23B does not interfere with antibody binding against FLAG tag (lane 3 and 4). Additionally, the band intensity is slightly stronger in the IP elutes (lane 4), which indicates efficient enrichment. Detection of the same blot with Streptavidin-HRP (Fig 3B) shows clear bands only in the GalNAz treated samples (lane 3 and 4), which indicates that O-GlcNAz was metabolically incorporated as PTM on proteins and that the Staudinger-Bertozzi reaction showed minor background labeling as seen in lanes 1 and 2. The IP sample showed only one intensive band correlating to the same electrophorestic shift of FLAG-6His-hHR23B (lane 4), whereas the input sample showed multiple bands (lane 3). That is because in these two samples, GlcNAz was successfully incorporated as PTM on many cellular proteins, and that, the anti-FLAG IP specifically and efficiently enriched the O-GlcNAcylated hHR23B. Therefore, the anti-biotin immuno blot (IB) detected numerous proteins in lane 3 and only hHR23B in lane 4.

For the affinity purification with streptavidin beads, Figure 3C shows the anti-FLAG detected proteins. Lane 1 and lane 2 are the input and AP elute samples without the GalNAz treatment. Lane 3 and lane 4 are the input and AP samples with GalNAz treatment. In both input samples with or without the GalNAz treatment, the FLAG-6His-hHR23B bands are comparable (lane 1 and 3). The AP elute with GalNAz treatment shows a clear band (lane 4), nonetheless the elution sample without the treatment does not (lane 2),

because no biotin was installed on the proteins. The hHR23B protein shown by the AP is less intensive than the band in the input samples, which might be because the O-GlcNAz modified protein is only a small portion of the whole population or the efficiency of the AP is not 100%. The anti-biotin detection shown in Figure 3D illustrates the enrichment efficiency of the AP experiment, and the efficiency of the metabolic labeling method. The slight bands shown in lane 1 and 2 are just endogenous biotinylated proteins instead of aspecific binding of the biotin-phosphine, considering the results in lane 1 and 2 of Figure 3B. The bands in lane 4 are much more intensive than the ones in lane 3, which is because of the enrichment of biotinylated proteins by the AP.

In summary, the cis-trans pull down experiments show that hHR23B is O-GlcNAcylated. The described strategy of metabolic labeling coupled to bio-orthogonal ligation provided a foundation for target enrichment and allowed straightforward visualization of O-GlcNAcylation on the target protein.

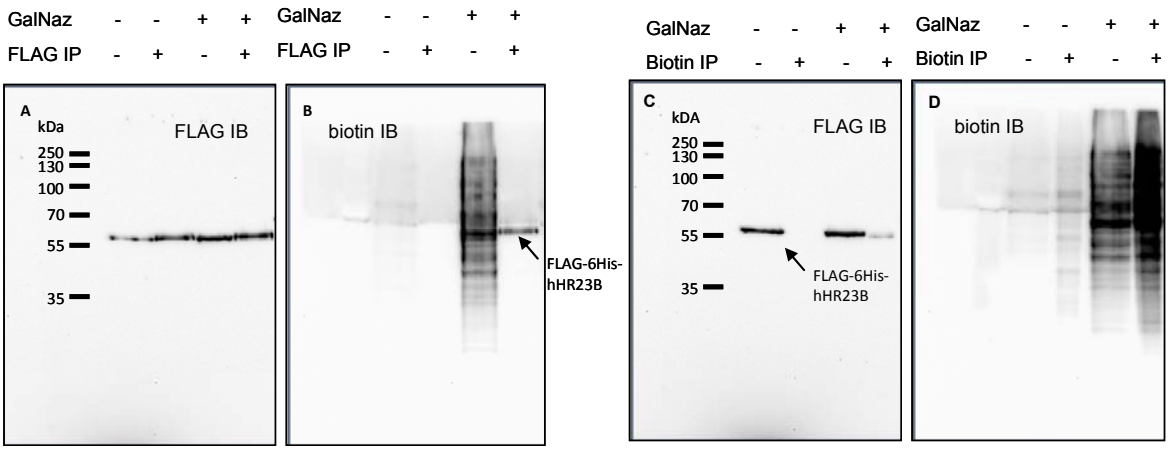


Figure 3: IP and AP to confirm O-GlcNAcylation on hHR23B. Cells tranfected with pFLAG-6His-hHR23B (with or without GalNAz) treatment were lysed. The biotin tag was installed by Staudinger-Bertozzi reaction. The FLAG tagged hHR23B was immunoprecipitated by anti-FLAG beads, and detected by FLAG antibody (A) or Strep-HRP (B). The biotin tagged proteins were affinity purified by Streptavidin beads, and detected by FLAG antibody (C) or Strep-HRP (D). IB demonstrates immuno blot.

7.2.2 Mapping of O-GlcNAcylation sites on hHR23B

To identify the exact O-GlcNAc modification sites on the hHR23B protein, the unstable O-GlcNAc moiety was replaced for a stable biotin-cystamine (**5**, Fig 2A) for the LC/MS analysis. Cells were transfected with plasmid pFLAG-6His-hHR23B, and then grown with or without the OGA inhibitor **2** for 24 hours(*20*) (Fig 2A). As seen in Figure 2D, OGA inhibitor **2** increased and OGT inhibitor **3** decreased the cellular O-GlcNAcylation levels significantly in the living cells. Following harvest and lysis of the cells, an anti-FLAG IP was

performed to enrich the expressed protein. The eluted protein was digested by proteinase K (ProtK) instead of the generally used trypsin, due to the special sequence of the hHR23B protein, which contains long linker parts with very few trypsin cleavage sites(C terminal of Lys and Arg). The digested peptides were dephosphorylated by alkaline phosphatase to prevent β-elimination of the phosphate. The O-GlcNAc groups on the peptides were replaced by biotin-cystamine (**5**) through a β-elimination followed by a Michael addition reaction under basic condition at 52°C(*21*) (Fig 2C). The peptides were then analyzed by nano-LC/MS for identification of the modification sites.

Through this experiment, two O-GlcNAcylation sites were identified in both samples, and four extra modification sites were identified only in the sample with OGA inhibition treatment (Table 1). This suggests the presence of constitutive O-GlcNAc modifications on the hHR23B protein, even under OGA activity in the cell. It is hypothesized that the constitutive O-GlcNAc are for protein functions and the differentially added O-GlcNAc are for regulating the functions by switching between open and closed states. Interestingly, the modification sites identified are mostly located around the linker part between UBL and UBA1, which is indeed the longest linker in the protein. It might play a role in the open-closed mode changes, according to the structural studies on the paralogue protein hHR23A(*16*).

Table 1: O-GlcNAcylation sites identification on hHR23B

samples	peptide identification	mascot score	O-GlcNAc sites
hHR23B-nt	^{29}KIE$_{BiCy}$**S**EKGKDAFPVAGQ^{44}	30	**Ser^{32}**
	^{73}MVTKPKAV$_{BiCy}$**S**TPAPATTQ^{89}	44	**Ser^{81}**
hHR23B-OGAi	^{29}KIE$_{BiCy}$**S**EKGKDAFPVAGQ^{44}	21	**Ser^{32}**
	^{73}MVTKPKAV$_{BiCy}$**S**TPAPATTQ^{89}	45	**Ser^{81}**
	^{74}VTKPKAVSTPAPAT$_{BiCy}$**T**QQ^{90}	23	**Thr^{88}**
	^{79}AVS$_{BiCy}$**T**PAPATTQ^{89}	23	**Thr^{82}**
	^{248}GAPQS$_{BiCy}$**S**AV^{255}	21	**Ser^{253}**
	^{339}GGGGGGG$_{BiCy}$**S**GGIA^{350}	33	**Ser^{346}**

What should also be considered is that a substantial part of the linker sequence Gln^{90}-Lys^{144} was not found by LC/MS analysis. Although the aggressive protease ProtK with little sequence preference was used for the proteolytic degradation of the protein, the linker part of the protein might still be incompletely digested due to the special sequence, and was not analyzed by the mass spectrometer. This could also be the reason that very few reports were published about the O-GlcNAcylation on hHR23B, when trypsin was used for digestion. To solve this problem, studies should be performed with pure hHR23B to find appropriate ProtK digestion condition for increasing the hHR23B coverage and possibly identify additional O-GlcNAcylation sites.

7.2.3 O-GlcNAcylation alters protein-protein interaction of hHR23B

hHR23B functions through binding to other proteins via its four intra-molecular domains. These functions could be influenced by conformational alterations, such as the open/closed model suggested by the hHR23A structural research(*16*). As determined, the O-GlcNAc modification sites are mainly positioned in the linker part between UBL and UBA1, which is the largest linker in the protein and might contribute to the open-closed model. It is to be expected that the O-GlcNAc modifications might influence the changing of the mode from one to the other.

To prove this hypothesis, the first step is to find out whether the interactome of hHR23B is changing depending on the O-GlcNAcylation levels. When hHR23B binds to its interaction partners, the fraction of free hHR23B is expected to be lower. By detecting the free fraction, an estimate of the bound hHR23B can be made, which indirectly reflects on the fraction of hHR23B that is bound to its interaction partners. HEK293T cells were transfected with pFLAG-6His-hHR23B and cultured with or without the OGA and OGT inhibitors (**2** and **3**) for 24 hours(*20, 22*). Subsequently, cellular proteins were cross linked by formaldehyde or not. After quenching the residual free formaldehyde with glycine, the cells were harvested and lysed. The same amount of lysate was loaded on SDS-PAGE from each condition. Samples were run on two gels, one was stained with coomassie blue as loading control correction and visualization of the total protein content, the other one was transferred to a western blot. Free hHR23B protein was detected by anti-hHR23B antibody (Fig 4).

The anti hHR23B western blot shows two bands around 55kD, because there are two populations of the hHR23B protein. The upper band is the transfected FLAG-6His-hHR23B, and the lower one is the endogenous hHR23B. The free hHR23B proteins stays on approximately the same level in the non-cross linked samples with or without inhibitor treatment (lane 1-3, Fig 4A). Because protein analysis by SDS-PAGE is performed under both denaturing and reductive conditions, it is obvious that most non-covalent PPI between hHR23B and its binding partners were lost. The results also demonstrate that the OGA/OGT inhibition did not influence the hHR23B expression level. However, the cross linked samples show some difference between the control and OGA/OGT inhibitions (lane 4-6). The results were then quantified and normalized by the total protein amount quantified from the coomassie stained gel (Fig 4B). The quantification results were shown in Figure 4C. After OGA inhibition, there was around 50% more free hHR23B than in the control cells. Whereas, about 10% less hHR23B monomer was detected in the OGT inhibited samples. These results in total illustrate that, the high O-GlcNAcylation level in cells altered the protein-protein interaction of hHR23B. The evidence indeed correlates to the hypothesis that O-GlcNAc modifications on the hHR23B could change its conformation and further alter its interactome and regulate its functions.

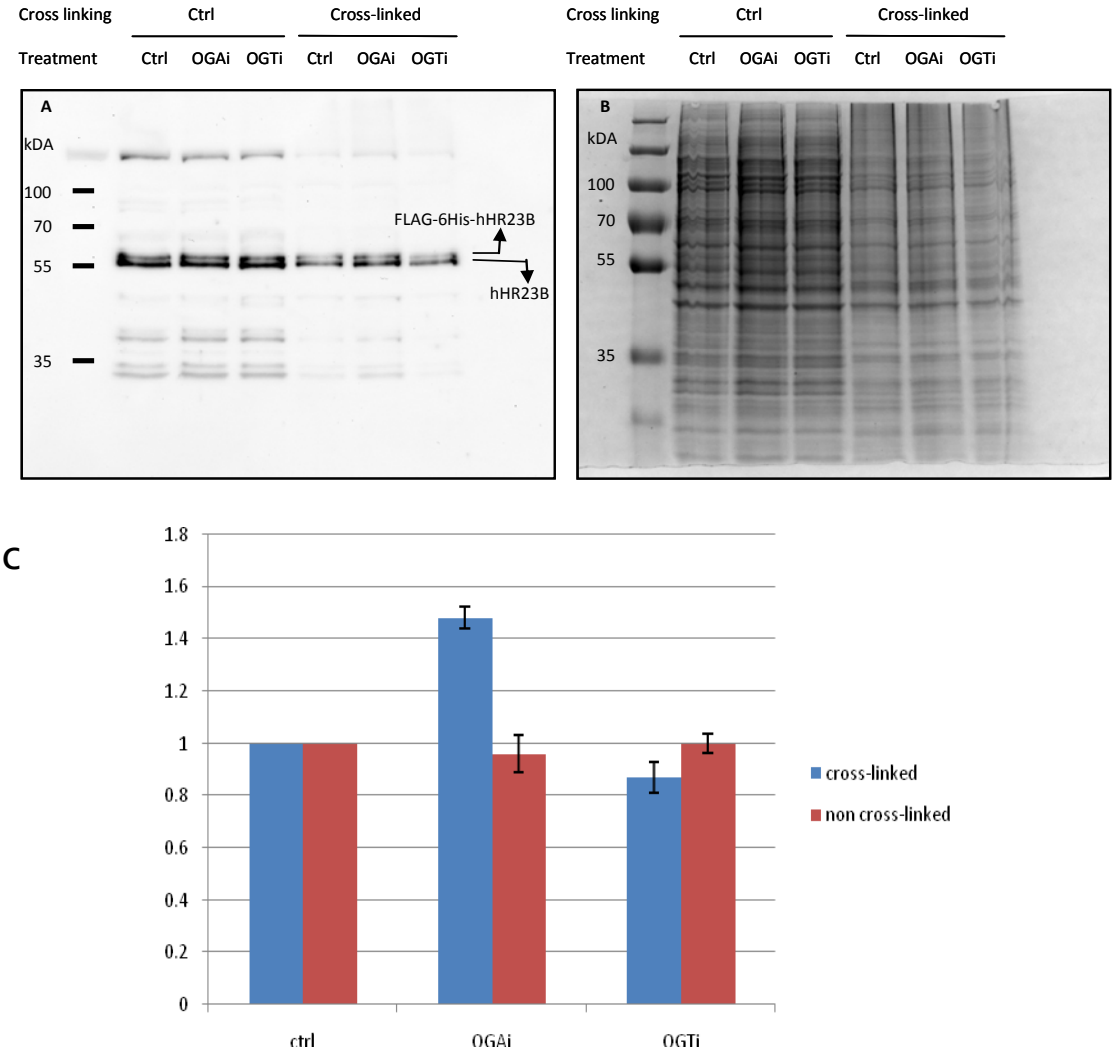


Figure 4: Detection of hHR23B monomer after O-GlcNAcylation regulation. Cellular proteins with or without OGA/OGT inhibitions were cross-linked by a formaldehyde treatment. Control samples and cross-linked samples were separated on SDS-PAGE, and detected by anti-hHR23B immunoblot after transferred to a PVDF membrane (A) or by Coomassie Blue staining (B). The hHR23B level is quantified from the blot (A) and normalized by the loading control (B) to get the quantification graph (C).

7.3 Discussion

O-GlcNAcylation is an essential regulator of many cellular processes, especially by modulation of protein-protein interactions and interplay with other post-translational modifications. The results in this chapter demonstrate that, hHR23B is indeed modified by O-GlcNAc under normal growth condition, which is confirmed by metabolic labeling of GlcNAz followed by the cis-trans pull down experiments. This result was subsequently reconfirmed by LC/MS mapping of the modification sites. Two out of six modification sites were identified on hHR23B in control cells without OGA or OGT inhibition treatment. It

indicates that the modification might be important for the protein functions or that the protein is O-GlcNAcylated under normal condition. Four extra O-GlcNAcylation sites were identified after OGA inhibition, which suggests that the modification is actually highly dynamic on hHR23B and the OGA activity was altered by the chemically synthesized inhibitor. Considering the localization of the O-GlcNAcylation sites identified, the linker between UBL and UBA1 might be the most possible heavily modified region. Moreover, LC/MS coverage of that part of the linker is low due to the difficult digestion, which implies that more O-GlcNAcylation sites might be identified. This also explains why hHR23B has been hardly picked up in previous O-GlcNAcylated protein identification studies, where trypsin was used as the protease for digestion(4).

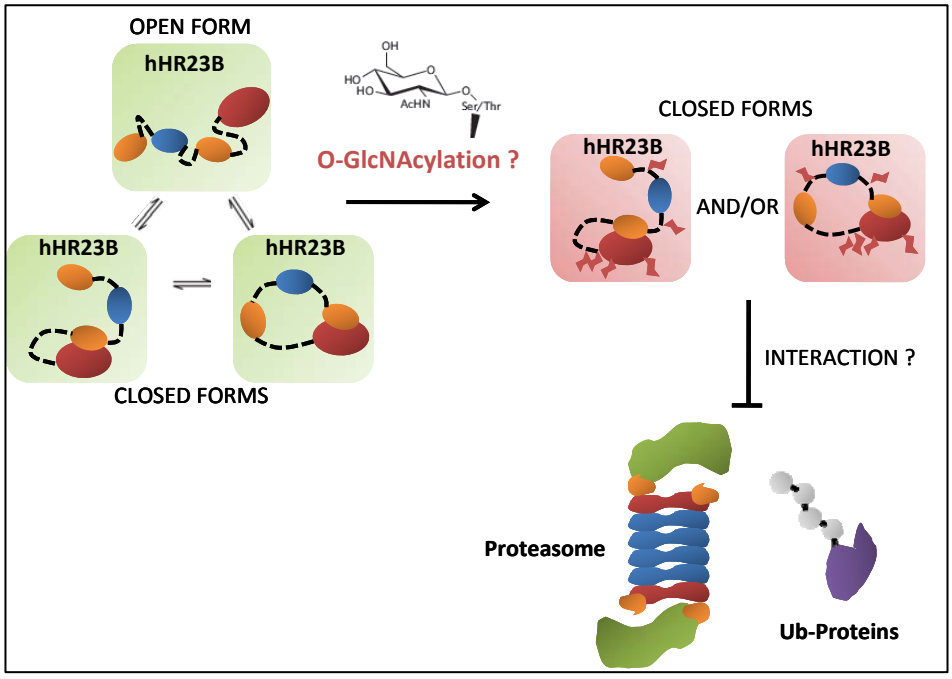


Figure 5: Hypothesis of the O-GlcNAc regulation of hHR23B. The O-GlcNAc moieties on the hHR23B protein might force it to present the closed mode of conformation. Further, the hHR23B interactome is altered by this closed conformation.

hHR23B is known as a key regulator of the global genome nucleotide excision repair pathway (GG-NER) and of the ubiquitin proteasome system (UPS)(15). It functions through binding to the interaction partner XPC that is in charge of the DNA damage recognition and recruitment of other NER factors, and stabilizes it against ubiquitination and proteasomal degradation. In the UPS, hHR23B is one of the major ubiquitin receptor proteins that shuttles polyubquitinated proteins to the proteasome for controlled degradation, by binding to the proteasomal substrates with the UBA domains and interacting with the proteasome using the UBL domain. So, research on the protein-protein interactions of hHR23B is the major focus of characterizing its functions. As an initial experiment, free

hHR23B level in cells with or without OGA/OGT inhibitions was compared. 50% more free hHR23B was observed in the OGA inhibited cells, while 10% less of that was detected in cells with OGT inhibition than in the control cells. This indicates that high O-GlcNAcylation level alters the PPI of hHR23B significantly, which then results in a low free protein ratio in total. Less O-GlcNAcylation does not affect this property of the protein so significantly. That might display the extra sugar moieties on the hHR23B after the OGA inhibition are important regulators for the PPI. These results correlate with previous reports on O-GlcNAcylation research of Sp1 and Tau(*6*, *7*).

The complete molecular mechanism why O-GlcNAcylation level can alter the PPI of hHR23B is still not clear. The hypothesis is that multiple sugars attached to the flexible linker change its property and possibly put the protein conformation into a closed mode, which does not allow the binding partners like proteasome or polyubiquintinated proteins to interact with it anymore (Fig 5). Possibly O-GlcNAcylation is used as a switch to rapidly activate or deactivate a certain pathway, in response to metabolic stimuli. The hypothesis is still to be proven. But recent observation opens new insights into the O-GlcNAcylation functions, that the modification can change the functions of scaffold protein which does not have enzymatic activity by altering its interactome. Furthermore, O-GlcNAcylation can influence the efficiency of the UPS by manipulating the functions of ubiquitin receptor proteins, so that the environmental glucose level and nutrition condition can be linked with the global protein turn over rate.

7.4 Experimental Procedures

7.4.1 Cell culture, transfection and treatment by chemicals

HEK293T cells were purchased from ATCC. The cells were grown in DMEM medium with 10% Fetal Calf Serum and 0.1 mg/ml penicillin and 0.1 mg/ml streptomycin, at 37°C, in a 5% CO_2 humid incubator. The cells were transfected with the PEI transfection reagent to express the target constructs. Ac4-GalNAz (**1**)(*3*), OGA inhibitor (**2**)(*20*) and OGT inhibitor (**3**)(*22*) were dissolved in DMSO before use. 1000x stock solution of the compounds were added to the cell culture for treatment to have the DMSO concentration lower than 1% in the culture medium. All treatments (250μM Ac4-GalNAz, 50 μM OGA inhibitor or 50 μM OGT inhibitor) described were done for 24 hours.

7.4.2 Immuno-precipitation and affinity purification

The cells were lysed in a mild lysis buffer containing 0.1% TX100, 50 mM Tris (pH 7.5), 5 mM EDTA, 250 mM NaCl and 10% Glycerol, supplemented with Protease inhibitor cocktail, phosphostop, and 100 μM OGA inhibitor. 1mg of the lysate was used for Staudinger-Bertozzi reaction with 250 μM biotin-phosphine. After the bioorthogonal ligation, the lysate could directly be used in an anti-FLAG immuno-precipitation

experiment with M2 magnetic agrose beads (Sigma-Aldrich). To do the biotin-streptavidin affinity purification, the protocol described in **Chapter 3** can be used with minor modifications. Reduction and alkylation steps can be skipped, because no LC/MS analysis is done afterwards. The lysis buffer described here can be used for the washing steps.

7.4.3 O-GlcNAcylation sites mapping with BEMAB

The FLAG-6His-hHR23B protein was purified by an anti-FLAG precipitation as described above. After elution, the protein was digested by Proteinase K for 30 min at 37°C. The digested peptides were acidified with Formic Acid to pH<3, cooled on ice. The peptides were desalted by a StageTip. The eluted peptides were evaporated to dryness in a SpeedVac. The dry peptides were dissolved in dephosphorylation buffer (20mM Tris pH7.5 and 10mM $MgCl_2$) and dephosphorylated by TSAP (thermosensitive alkaline phosphatise, Promega) for 3 hours at 37°C. The biotin-cystamine solution (dissolved in 0.4% NaOH and 4% triethylamine) was added at final concentration of 5mM. The reaction mixture was kept at 52°C for 2 hours(*21*). Then, the mixture was acidified and desalted by a StageTip again. The eluted sample was evaporated in a SpeedVac, and reconstituted with LC/MS sample solution (95ml H_2O, 3ml ACN and 0.1ml Fomic Acid). The samples could then be analyzed by nano-LC/MS.

The MS data was searched by Mascot against a self defined database containing only the hHR23B protein sequence to simplify the difficulty of calculations.

7.4.4 Formaldehyde cross linking

Formaldehyde solution was added to the medium of cells with or without OGA/OGT inhibitors treatment to reach a final concentration of 0.75%. The plates were shaken gently for several times to mix well and then incubated at room temperature for 15 minutes. 125mM Glycine was added to quench the cross linking reaction, for 5 minutes at room temperature. The medium was removed and cells were harvested by scraping.

7.4.5 Synthesis of biotin-cystamine

2-(tritylthio)ethan-1-amine A mixture of cysteamine hydrochloride (464mg, 4.1 mmol) and triphenylmethanol (1.06g, 4.1mmol) in 8 mL TFA was stirred for 24h at room temperature. After evaporation of all solvent in vacuum, water (40 mL) was added and pH was adjusted to 9 upon addition of $NaHCO_3$. The product was extracted with chloroform (3×20mL), the combined organic layer was washed by brine, dried over $MgSO_4$, filtered and concentrated in vacuum to product crude product and purification by flash chromatography (2%MeOH in DCM with 0.1% TEA) gave compound (200mg, yield 15.3%).

5-(2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)-N-(2-(tritylthio)ethyl) pentanamide Dissolve compound 2-(tritylthio)ethan-1-amine (75.5mg, 0.24mmol) in 5 mL DMF and then

add Biotin 70.8 mg (1.2 equi., 0.29mmol), HCTU 120mg(1.2 equi., 0.29mmol) and DiPEA 0.13mL(3 equi., 0.72mmol) and then stir at room temperature overnight. Remove DMF in vacum and dissolve the residue in EtOAc. The organic layer was washed by 1M HCl(2×20mL), Sat. $NaHCO_3$(3×20mL), and Brine(1×20mL) and dried over $MgSO_4$, giving crude product (126.7mg, yield %). The crude product was used without further purification.

N-(2-mercaptoethyl)-5-(2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)pentanamide

Dissolve crude compound 5-(2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)-N-(2-(tritylthio)ethyl) pentanamide 126.7mg in 6 mL DCM and then add 4 mL TFA and 7 drops of triisopropylsilane. TLC shows complete conversion of the starting material after stirring at room temperature overnight. After evaporation of all solvent in vacuum, product (40mg, yield 56.8%) was obtained by HPLC purification.

^{1}HNMR (400 MHz, CD_3OD): 4.40(dd, 1H, J = 3.2, 4.8Hz), 4.21 (dd, 1H, J = 4.4, 4.4Hz), 3.12 (ddd, 1H, J = 4.4, 3.2, 3.2 Hz), 2.84 (dd, 1H, J = 4.8, 5.2Hz), 2.61 (d, 1H, J= 12.8 Hz), 2.52 (t, 2H, J = 4.8 Hz), 2.12 (t, 2H, J = 7.6 Hz), 1.67-1.28(m, 6H). 13C NMR (CD_3OD): 176.19 (C), 166.13 (C), 63.36 (CH), 61.62 (CH), 57.00 (CH), 43.84 (CH_2), 41.04 (CH_2), 36.72 (CH_2), 29.76 (CH_2), 29.47 (CH_2), 26.33 (CH_2), 24.50 (CH_2). MS $(M+H)^+$: 304.07.

Acknowledgement

We thank Francesco Marchese and Nico Dantuma for making the pFLAG-6His-hHR23B construct. We thank Hendrick Gold, Wouter van der Linden and Bo-Tao Xin for synthesis of the chemical compounds.

References and notes

1. Hart, G. W., Slawson, C., Ramirez-Correa, G., and Lagerlof, O. (2011) Cross talk between O-GlcNAcylation and phosphorylation: roles in signaling, transcription, and chronic disease, *Annu Rev Biochem 80*, 825-858.

2. Lazarus, M. B., Nam, Y., Jiang, J., Sliz, P., and Walker, S. (2011) Structure of human O-GlcNAc transferase and its complex with a peptide substrate, *Nature 469*, 564-567.

3. Boyce, M., Carrico, I. S., Ganguli, A. S., Yu, S. H., Hangauer, M. J., Hubbard, S. C., Kohler, J. J., and Bertozzi, C. R. (2011) Metabolic cross-talk allows labeling of O-linked beta-N-acetylglucosamine-modified proteins via the N-acetylgalactosamine salvage pathway, *Proc Natl Acad Sci U S A 108*, 3141-3146.

4. Hahne, H., Sobotzki, N., Nyberg, T., Helm, D., Borodkin, V. S., van Aalten, D. M., Agnew, B., and Kuster, B. (2013) Proteome wide purification and identification of O-GlcNAc-modified proteins using click chemistry and mass spectrometry, *J Proteome Res 12*, 927-936.

5. Hahne, H., Moghaddas Gholami, A., and Kuster, B. (2012) Discovery of O-GlcNAc-modified proteins in published large-scale proteome data, *Mol Cell Proteomics 11*, 843-850.

6. Lim, K., and Chang, H. I. (2009) O-GlcNAcylation of Sp1 interrupts Sp1 interaction with NF-Y, *Biochem Biophys Res Commun 382*, 593-597.

7. Yuzwa, S. A., Shan, X., Macauley, M. S., Clark, T., Skorobogatko, Y., Vosseller, K., and Vocadlo, D. J. (2012) Increasing O-GlcNAc slows neurodegeneration and stabilizes tau against aggregation, *Nat Chem Biol 8*, 393-399.

8. Hardiville, S., Hoedt, E., Mariller, C., Benaissa, M., and Pierce, A. (2010) O-GlcNAcylation/phosphorylation cycling at Ser10 controls both transcriptional activity and stability of delta-lactoferrin, *J Biol Chem 285*, 19205-19218.

9. Yang, W. H., Kim, J. E., Nam, H. W., Ju, J. W., Kim, H. S., Kim, Y. S., and Cho, J. W. (2006) Modification of p53 with O-linked N-acetylglucosamine regulates p53 activity and stability, *Nat Cell Biol 8*, 1074-1083.

10. Guinez, C., Mir, A. M., Dehennaut, V., Cacan, R., Harduin-Lepers, A., Michalski, J. C., and Lefebvre, T. (2008) Protein ubiquitination is modulated by O-GlcNAc glycosylation, *Faseb J 22*, 2901-2911.

11. Li, M. D., Ruan, H. B., Hughes, M. E., Lee, J. S., Singh, J. P., Jones, S. P., Nitabach, M. N., and Yang, X. (2013) O-GlcNAc signaling entrains the circadian clock by inhibiting BMAL1/CLOCK ubiquitination, *Cell Metab 17*, 303-310.

12. Guinez, C., Mir, A. M., Leroy, Y., Cacan, R., Michalski, J. C., and Lefebvre, T. (2007) Hsp70-GlcNAc-binding activity is released by stress, proteasome inhibition, and protein misfolding, *Biochem Biophys Res Commun 361*, 414-420.

13. Ruan, H. B., Nie, Y., and Yang, X. (2013) Regulation of protein degradation by O-GlcNAcylation: crosstalk with ubiquitination, *Mol Cell Proteomics*.

14. Rexach, J. E., Clark, P. M., and Hsieh-Wilson, L. C. (2008) Chemical approaches to understanding O-GlcNAc glycosylation in the brain, *Nat Chem Biol 4*, 97-106.

15. Dantuma, N. P., Heinen, C., and Hoogstraten, D. (2009) The ubiquitin receptor Rad23: at the crossroads of nucleotide excision repair and proteasomal degradation, *DNA Repair (Amst) 8*, 449-460.

16. Wang, Q., Goh, A. M., Howley, P. M., and Walters, K. J. (2003) Ubiquitin recognition by the DNA repair protein hHR23a, *Biochemistry 42*, 13529-13535.

17. Finley, D. (2009) Recognition and processing of ubiquitin-protein conjugates by the proteasome, *Annu Rev Biochem 78*, 477-513.

18. Kim, I., Ahn, J., Liu, C., Tanabe, K., Apodaca, J., Suzuki, T., and Rao, H. (2006) The Png1-Rad23 complex regulates glycoprotein turnover, *J Cell Biol 172*, 211-219.

19. Isono, T. (2011) O-GlcNAc-specific antibody CTD110.6 cross-reacts with N-GlcNAc2-modified proteins induced under glucose deprivation, *PLoS One 6*, e18959.

20. Macauley, M. S., Whitworth, G. E., Debowski, A. W., Chin, D., and Vocadlo, D. J. (2005) O-GlcNAcase uses substrate-assisted catalysis: kinetic analysis and development of highly selective mechanism-inspired inhibitors, *J Biol Chem 280*, 25313-25322.

21. Overath, T., Kuckelkorn, U., Henklein, P., Strehl, B., Bonar, D., Kloss, A., Siele, D., Kloetzel, P. M., and Janek, K. (2012) Mapping of O-GlcNAc sites of 20 S proteasome subunits and Hsp90 by a novel biotin-cystamine tag, *Mol Cell Proteomics 11*, 467-477.

22. Gloster, T. M., Zandberg, W. F., Heinonen, J. E., Shen, D. L., Deng, L., and Vocadlo, D. J. (2011) Hijacking a biosynthetic pathway yields a glycosyltransferase inhibitor within cells, *Nat Chem Biol* *7*, 174-181.

Summary and future perspectives

The work described in this thesis focused on developing chemical biology tools for understanding the ubiquitin proteasome system (UPS). **Chapter 1** is a general overview of the UPS model system and the main methods used in this thesis about activity-based proteasome profiling. The UPS is the central system of the cellular protein turn over processes, in which ubiquitin works as the signal targeting the protein to destruction and proteasome is the protein recycling center. The proteasome is the major protein degradation machinery in eukaryotic cells and has been validated as an important drug target for treatment of cancer and autoimmune diseases. Methods to reveal the proteasome activity are therefore necessary for both fundamental and (pre)clinical researches.

Chapter 2 is a brief literature review of the application of activity-based protein profiling (ABPP) in biological research in recent years. ABPP was combined with LC/MS based protein quantification platforms, to determine the global reactive cysteine thiols throughout the whole human proteome(*1*). In cellular biology, ABPP is often used as a imaging tool. A fluorescently quenched cathepsin probe was applied for noninvasive optical imaging in living mice for the early diagnosis of grafted tumors(*2*). In another case, fluorescent proteasome probes were used to observe the T cell asymmetric division(*3*). As a powerful tool in medicinal chemistry, ABPP has been broadly used in target identification. The KiNativ high-throughput screening platform from the ActivX company profiled several well defined kinase inhibitors against over 200 kinases in the human proteome to reveal the cellular targets of the inhibitors(*4*). Activity-based probes (ABP) for serine hydrolases and glucocerebrosidase were applied to study tumor migration, growth and therapy progression in Gaucher disease(*5, 6*). In conclusion, the ABPP method has been widely used in various biological research disciplines, and it could be developed for more enzyme families and applied in more fields.

In **Chapter 3**, a relative quantification activity-based proteasome profiling method is described to quantify the activity of different proteasome subunits in a complex biological

sample. The method employs fluorescent proteasome ABPs, including subunit specific probes. Active β1, β2, β5 proteasome subunits were labeled with fluorescent probes in cell lysate, and separated on a 1D SDS-PAGE. The labeled subunits were visualized by in-gel fluorescent scanning, and the fluorescent signal was quantified. This protocol using the fluorescent probes allows one to perform the quantification of the proteasome activity of every active subunit, in a medium throughput mode, which solves the problem that the fluorogenic substrates can not distinguish constitutive subunits and their immuno homologues. The second part of the procedure describes the use of a biotin-epoxomicin probe in a LC/MS-based proteasome activity quantification experiment. All the proteasome active subunits were labeled by biotin-epoximicin, then affinity purified with streptavidin paramagnetic beads through biotin-streptavidin interaction. Subsequently, the proteasome subunits were digested by trypsin on the beads and differentially labeled with dimethyl stable isotopes. The labeled peptides were mixed and analyzed by reversed phase nano-LC/MS. Through this MS analysis, the activity of all proteasome subunits can be quantified in one single LC/MS run. This protocol has a lower throughput than the gel based protocol; however, it has a lower chance of systematic error as different subunits are quantified in one single experiment. The two protocols can be chosen for different applications according to the aims, and can be adapted to a clinical setting.

Bioorthogonal chemistry has been a valuable chemical tool in the biological research for its property that, it labels marco-molecules selectively, through the introduction of unnatural, small ligation handles in a complex biological sample. Because reporter groups sometimes influence the property of the probes, the bioorthogonal ligation handles are introduced instead of the reporter, and then the reporter group can be installed after the first enzymatic labeling step. **Chapter 4** is a technical report, comparing primary and secondary azide groups in two step ABPP experiments. The two kinds of azide were compared for their ligation efficiency with Staudinger-Bertozzi biotin-phosphine and Copper-catalyzed Huisgen's azide-alkyne cycloaddition biotin-alkyne, under either denatured or native protein condition. Under denatured condition both reactions can label both kinds of azide quite efficiently, however under native condition the secondary azide is less accessible for labeling by both reactions. It demonstrates that the environment composed by the folded probed protein and the chemical groups around the secondary azide decreased the efficiency of the bioorthogonal ligation under native conditions.

Bortezomib is the first proteasome inhibitor approved by FDA as a drug and is clinically used as an anti-cancer therapy. During the ten years of clinical application, bortezomib has been proven to be a great success. However, the molecular mechanisms how bortezomib is able to kill myeloma cells instead of other cancer cells, and how myeloma cells develop resistance to the drug are still unclear. In **Chapter 5**, the quantitative activity-based proteasome profiling method (described in **Chapter 3**) is employed in

combination with a quantitative global proteomics method to elucidate the molecular mechanisms of the questions. In solid tumor cells that are less sensitive to bortezomib treatment, the proteasome β5 subunit is less susceptible for bortezomib inhibition than in lymphoblast cells. In the bortezomib adapted cells tested in this project, the proteasome expression has been up-regulated by different levels, especially the constitutive proteasome. Moreover, comprehensive modulations of multiple cellular processes were observed in bortezomib adapted cell lines HL60A and AMO1A by global proteomics analysis. For instance, in the acute myeloid leukemia cell line HL60A, together with proteasome up-regulation it shows up-regulated PLK1 (polo like kinase 1) expression, which is an activator of the proteasome through phosphorylating the enzyme complex and also an important regulator of mitosis(*7, 8*). HL60A cells down-regulate the ribosomal proteins probably to synthesize less protein, lowering the ER stress by slower protein turnover (data not shown). However, in AMO1A myeloma cells, except the higher proteasome levels, lower RPB1 level was observed, which is the key subunit of the essential eukaryotic gene transcription machinery RNA polymerase II. So these cells evade bortezomib toxicity by up-regulating the proteasome and repressing protein synthesis. The proteomic changes observed might be used as prognosis biomarkers to predict whether bortezomib will be suitable for the treatment of the cancer or as an inspiration source for a different therapy.

More experiments are necessary to validate the conclusions. Native PAGE of proteasome and in-gel fluorescent ABPP were performed to compare the proteasome activity and expression levels in the parental and bortezomib adapted cells. The LC/MS-based quantitative activity-based proteasome profiling experiments will be performed to provide a complete overview of the proteasome activity in the adapted and the parental cells. The active proteasome from untreated parental (HL60/AMO1) and bortezomib adapted cells (HL60A/AMO1A) can be labeled by biotin-epoxomicin. The biotinylated active subunits can be then affinity purified and quantified by LC/MS.

Up-regulation of the constitutive proteasome was observed in the bortezomib adapted cells. One hypothesis is that the adapted cells up-regulate the constitutive proteasome and perhaps down-regulate the immuno proteasome by the competitive assembly of the 20S core particles. To prove the hypothesis, western blots against β2 and β2i are going to be performed in adapted and wild type cells. In combination with the Native PAGE and LC/MS based ABPP data, the expression/assembly and the activity of the proteasome in resistant and wild type cells will be elucidated.

Proteasomes were found to be less sensitive to bortezomib treatment in solid tumor cells than in lymphoblast cells (especially the clinical relevant plasma cells). One possibility to explain this phenomenon is that, in the solid tumor or resistant myeloma cells more proteasome is expressed. The ABPP technique described here is a suitable method for relative quantification of the proteasome activity between samples. However, for the large

scale screening of various cell lines to profile their proteasome expression, ABPP might not have enough throughput. To solve this problem, a combination of MRM (multiple reaction monitoring) LC/MS technique and AQUA (absolute quantification) peptides can be an efficient solution (*9-11*). The tryptic peptide sequences of the proteasome subunits that show robust LC/MS identification properties can be synthesized as AQUA peptides by incorporating Valine ($^{13}C_5$) and Leucine ($^{13}C_6$). And then, fixed amount of the AQUA peptides can be spiked in the trypsin digested full cell lysate prior to LC/MS analysis. The Mass Spectrometer can be programmed to fragment and quantify the peptide pairs consisted of the heavy AQUA peptides and their light endogenous parental peptides. In this way, the proteasome subunits expression levels can be quantified absolutely in a high throughput mode. This method can also be adapted to (pre)clinical setting for screening of cell lines or patient materials.

Various cell lines were compared with or without bortezomib treatment in global proteomics experiments (data not shown). However, the results were not conclusive. Several up or down-regulated proteins were observed, however no clear network was achieved after bioinformatics data processing. There might be two possibilities. One is that the cells did not react to the bortezomib treatment by regulating their protein expression levels, but by changing posttranslational modification (PTM) levels(*12*). It was found that Ser38 phosphorylation on stathmin, a key regulator in the control of proliferation and cell cycle increased upon bortezomib treatment in multiple myeloma cells and mutation of Ser38 of stathmin decreased sensitivity of the cells to bortezomib treatment. The other possibility is that the cells did change the protein expression levels, but the change was blurred by the accumulated proteasomal substrates after the bortezomib treatment. The former possible problem could be solved by quantification of the protein PTM levels, using specific PTM enrichment tools (antibodies or resins) and proteomics(*13, 14*). The latter problem could be addressed by quantifying the newly synthesized proteins in the cells after bortezomib treatment(*15, 16*). The strategy would be to grow cells in light SILAC medium (stable isotopic labeling of amino acid in cell culture) with or without bortezomib for 2 hour, and then culture in either intermediate (without treatement) or heavy (with treatment) SILAC medium supplemented with AHA (azidohomoalanine) instead of methionine for 4 hours. The AHA and the heavier stable isotopic Lys and Arg residues will be incorporated into newly synthesized proteins instead of methionine and light Lys and Arg respectively. The cells are harvested, mixed (1:1) and lysed. All the newly synthesized proteins can be enriched by performing the CuAAC in the presence of Alkyne coated resins. Proteins containing AHA will react and covalently bind to the beads and as such enriched from solution. After tryptic on bead digest, quantitative proteomics can be performed on the enriched newly synthesized proteins. In this way, protein expression levels can be compared quantitatively in cells with or without bortezomib treatment.

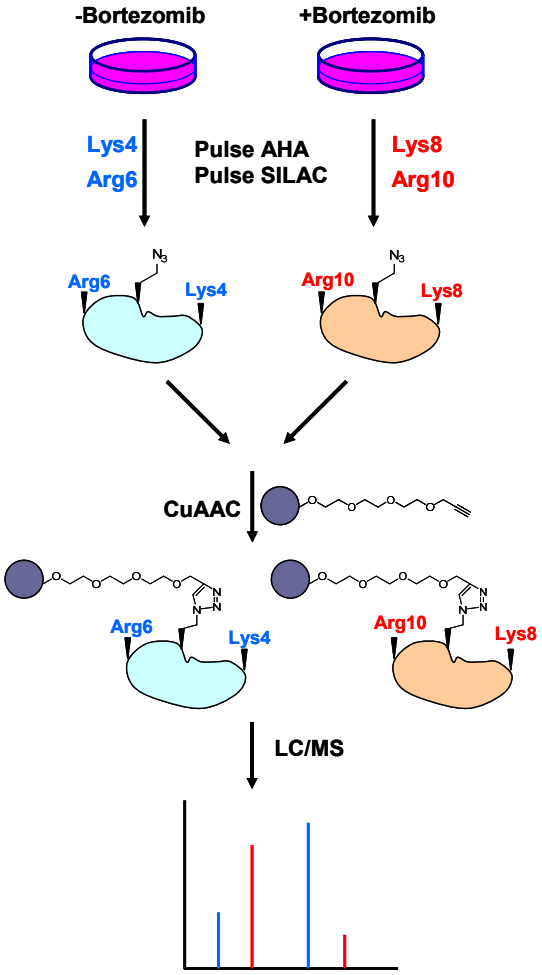


Figure 1: Quantification of newly synthesized proteins with/without bortezomib treatment. Cells are treated with or without bortezomib for 2hours in the light SILAC medium. And then, the light medium is removed and non-treated cells are cultured in intermediate (M) SILAC medium while bortezomib treated cells receive heavy (H) SILAC medium for 4 hours. Both M and H media contain AHA to label newly synthesized proteins with an azide. After mixing the cells (1:1) and lysis, the AHA incorporated proteins can be enriched through coupling to Alkyne coated resins under Cu(I) catalysis. Enriched newly synthesized proteins will be digested and analyzed by LC/MS.

Several cell skeleton proteins were found to be up-regulated in the bortezomib adapted cells, which might indicate that more of the adapted cells than the parental cells are in mitosis. A flow-cytometric experiment is going to be applied to compare this. The parental and adapted cells can be fixed by ethanol, and then stained with anti-tubulin (green) and Propidium Iodide (red) to stain the DNA (nucleus), including RNAse treatment to prevent PI staining of RNA. The stained cells are quantified by flow cytometer to show the 2N/4N ratio, which illustrates the ratio of non-dividing to dividing cells. If the adapted cells show a lower 2N/4N ratio than the parental cells, it means that they are more engaged in mitosis. These results will answer the question why more cell skeleton proteins were observed in bortezomib adapted cells.

Proteins that showed significant changes in their expression levels in the global proteomics experiments, for instance PLK1, might be used as therapeutic targets for the treatment of bortezomib resistant patients. To model this, HL60A cells can be treated with a combination of bortezomib and PLK1 inhibitor volasertib and compare with the treatment with either single inhibitor(17). If synergistic effects are observed from the

comparison, the global proteomics approach might be considered as a useful strategy to search for new therapeutic targets against bortezomib resistant cancer cells.

In **Chapter 6,** activity-based proteasome profiling is used for the characterization of the newly identified active proteasome subunit β5t that plays a crucial role in positive T cell selection. The β5t was found to replace β5i in the immuno proteasome in the cortical thymus epithelial cells to form the newly found 20S particle thymo proteasome. However, the direct proof whether β5t is catalytically active or just an inactive bystander was missing. Through affinity purification followed by active site peptide identification, β5t was proved to be active indeed. The competitive ABPP shows that the β5t has different cleavage preference compared to β5/5i that cleaves hydrophobic peptides. Proteasome inhibitors with a hydrophilic P2 amino acid have more chances to inhibit β5t. This might illustrate how to make a likely β5t specific inhibitor or fluorogenic substrate, which could benefit the immunological research.

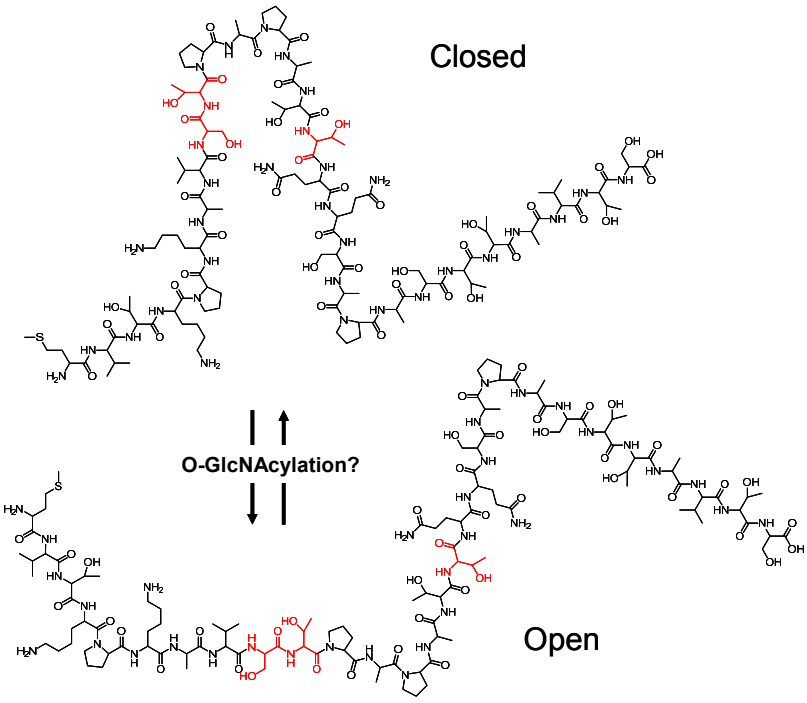


Figure 2: Chemical structure of the synthetic peptide and O-GlcNAcylated Thr and Ser.

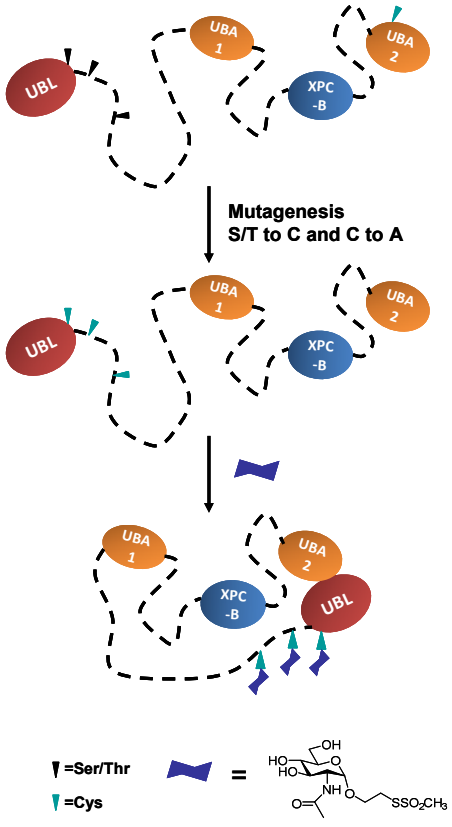
hHR23B peptide (Met^{73}-Ser^{101}) can be made by solid phase synthesis. The synthetic O-GlcNAcylated Ser or Thr can be incorporated into the peptide during the synthesis to replace the Ser^{51}, Thr^{52} and Thr^{58} that are in the red color.

Besides the protein expression level, post-translational modifications (PTM) are also important for cellular functions of proteins. **Chapter 7** focuses on the O-GlcNAcylation of the ubiquitin receptor protein hHR23B and how the PTM regulates protein function. It is proved that, there is O-GlcNAcylation on the hHR23B protein by the cis-trans immunoprecipitation experiments where protein was pulled down and the PTM was detected, or the other way round. Multiple GlcNAc groups have been found on the hHR23B protein by LC/MS protein identification, by replacing the unstable GlcNAc with a stable biotin-cystamine. After increasing or decreasing the cellular O-GlcNAcylation levels with

chemical inhibitors for OGA and OGT, followed by formaldehyde cross linking of proteins, the free hHR23B amounts differ in samples with higher or lower O-GlcNAcylation. These results indicate that, the PTM might influence the protein functions by changing its conformation and protein-protein interactions.

It still remains to be proved that, the O-GlcNAcylation of hHR23B changes its conformation and protein-protein interactions. The conformation study will start with the synthesis of a naked peptide of hHR23B that is the O-GlcNAcylation rich region, and compare its structure to that of the O-GlcNAc modified peptide using circular dichroism spectrometry and NMR(*18, 19*). To do this, the peptide shown in Figure 2 will be made by solid phase peptide synthesis. O-GlcNAcylated Ser and Thr need to be synthesized and used for synthesis of the modified peptide for comparison.

If the experiment above presents clear difference between the synthetic peptides, the conformation change should also be measured on the protein level. The hHR23B gene can be cloned into a bacterial expression vector, and then the Thr/Ser residues as the O-GlcNAc modification sites in the gene can be mutated into cysteines. The mutant recombinant protein can be purified. Meanwhile, the GlcNAc molecule with a thiol group to replace the 1-hydroxyl group needs to be synthesized. Subsequently, the artificial GlcNAcylation can be installed on the protein according to van Kasteren et al(*20, 21*) (Fig 2). The fully modified hHR23B protein can be separated from the free sugar molecules by a size exclusion column, and the purity can be checked by MALDI-TOF. NMR spectroscopy or AFM imaging will be used to reveal the conformation change.



to

Figure 3: Schemes of chemical installation of GlcNAc on hHR23B.

Ser^{51}, Thr^{52} and Thr^{58} in hHR23B can be mutated Cys in the bacteria expression construct, while Cys^{390} can be mutated to Ala. The mutated hHR23B is purified from E.coli. Subsequently, GlcNAc-MTS compound was used to install GlcNAc on the mutated hHR23B in vitro.

To get a complete O-GlcNAcylation sites atlas on hHR23B, it is advisible to start with a substantial amount of the protein, to test which in vitro modification mapping experiment could be a feasible strategy. The idea is to purify recombinant hOGT and hHR23B from E.coli(*22*). The synthetic UDP-GlcNAc can be used as the donor substrate and pure hHR23B as the protein substrate in the hOGT promoted in vitro glycosylation. Following protein digestion, the labile O-GlcNAc can be switched into a stable S-biotin by the BEMAB reaction as described in **Chapter 7**. The peptides can be identified by nano-LC/MS.

In order to reveal the different protein-protein interactions of hHR23B with high or low O-GlcNAcylation, a quantitative proteomics experiment is designed. A stable cell line will be made by transfecting HEK293T with pFLAG-6His-hHR23B and growing the transfectants under G418 selection. Single clones can be picked up and amplified for the immunoblot test. Cells from the best clone can be grown in light (L), intermediate (M) and heavy (H) SILAC medium for at least five passages. The L, M and H labeled cells can then be treated with OGT/OGA/no inhibitors. The same number of cells from different treatment can be mixed and lysed in a gentle buffer. The hHR23B interactome can be enriched by co-immunoprecipitation using the anti-FLAG beads or Ni-NTA beads and quantified by LC/MS. The quantitative protein lists can be compared to extract the promising data.

In conclusion, this thesis brings experimental evidence that ABPP is a powerful method to study chemical biological questions. The ABPP method has been applied to proteasome as a model system and showed results that shed light on mechanism of bortezomib resistance in myeloma cells, compared accessibility primary and secondary azides, proved the activity of thymus specific β5t proteasome subunit. Additionally, a new research line has been initiated on O-GlcNAcylation in ubiquitin receptor protein hHR23B.

Experimental procedures

Synthesis of O-GlcNAcylated Ser/Thr

Acetyl-3,4,6-tri-O-acetyl-2-N-acetamide-2-deoxy-D-glucopyranoside (1.56 g, 4 mmol) was dissolved in dry dichloroethane (20 ml) under argon atmosphere. TMSOTf (0.82 ml, 4.4 mmol) was added portion wise and the resulting solution was stirred at 65°C for 20 h. After completion, the mixture was quenched with Et_3N, concentrated, and purified by column chromatography (pentane : EtOAc = 1 : 2, 0.5% Et_3N). 3,4,6-tri-O-acetyl-2-amino-2-deoxy-1-O,2-N-methylidene-α-D-glucopyranoside (1.1 g, 84%)

3,4,6-tri-O-acetyl-2-amino-2-deoxy-1-O,2-N-methylidene-α-D-glucopyranoside (296 mg, 0.9 mmol) and FmocSerOH (196 mg, 0.6 mmol) were dissolved in a mixture of CH_2Cl_2-acetonitrile (2 : 1, 10 ml). After cooling to 0°C, freshly distilled $BF_3.Et_2O$ (0.094 ml, 0.9 mmol) were added dropwise to the solution. The reaction mixture was left to stand at room temperature and periodically monitored by TLC ($CHCl_3$: MeOH : AcOH = 80 : 10 : 1).

When the reaction was complete (from 48-150 h), the crude was neutralized at 0°C with Et_3N, diluted with CH_2Cl_2 and filtered through Celite. The filtrate was concentrated and the crude purified by column chromatography (silica gel, $CHCl_3$: MeOH from 40 : 1 to 10 : 1). Following purification, FmocSer(β-D-pGlcNAc)OH was obtained in a yield of 38%. A similar procedure was used for the synthesis of FmocThr(β-D-pGlcNAc)OH, which was obtained in a yield of 27%.

Construction of the hHR23B expression plasmid

pFLAG-6His-hHR23B was linearized by digestion with StuI, as the template of the hHR23B sequence. PCR was performed with primers 2934 (forward) and 2935 (reverse). The PCR product and E.coli expression vectors pET16B and pET20B were digested with NcoI and XhoI. The ligation was performed with the digested PCR product and both digested vectors. Sequencing was done to check the new constructs.

Primer sequences: 5′ CATGGATCCGAATTCCATGGACTACAAGGACGATGACG 3′ (2934), 5′ CATCTCGAGAAGCTTCTAGACTAATCTTCATCAAAGTTCTGCTG 3′ (2935).

Acknowledgement

We thank Qingju Zhang for synthesizing the O-GlcNAcylated Thr and Ser.

References and notes

1. Weerapana, E., Wang, C., Simon, G. M., Richter, F., Khare, S., Dillon, M. B., Bachovchin, D. A., Mowen, K., Baker, D., and Cravatt, B. F. (2010) Quantitative reactivity profiling predicts functional cysteines in proteomes, *Nature 468*, 790-795.

2. Blum, G., von Degenfeld, G., Merchant, M. J., Blau, H. M., and Bogyo, M. (2007) Noninvasive optical imaging of cysteine protease activity using fluorescently quenched activity-based probes, *Nat Chem Biol 3*, 668-677.

3. Chang, J. T., Ciocca, M. L., Kinjyo, I., Palanivel, V. R., McClurkin, C. E., Dejong, C. S., Mooney, E. C., Kim, J. S., Steinel, N. C., Oliaro, J., Yin, C. C., Florea, B. I., Overkleeft, H. S., Berg, L. J., Russell, S. M., Koretzky, G. A., Jordan, M. S., and Reiner, S. L. (2011) Asymmetric proteasome segregation as a mechanism for unequal partitioning of the transcription factor T-bet during T lymphocyte division, *Immunity 34*, 492-504.

4. Patricelli, M. P., Nomanbhoy, T. K., Wu, J., Brown, H., Zhou, D., Zhang, J., Jagannathan, S., Aban, A., Okerberg, E., Herring, C., Nordin, B., Weissig, H., Yang, Q., Lee, J. D., Gray, N. S., and Kozarich, J. W. (2011) In situ kinase profiling reveals functionally relevant properties of native kinases, *Chem Biol 18*, 699-710.

5. Nomura, D. K., Long, J. Z., Niessen, S., Hoover, H. S., Ng, S. W., and Cravatt, B. F. (2010) Monoacylglycerol lipase regulates a fatty acid network that promotes cancer pathogenesis, *Cell 140*, 49-61.

6. Witte, M. D., Kallemeijn, W. W., Aten, J., Li, K. Y., Strijland, A., Donker-Koopman, W. E., van den Nieuwendijk, A. M., Bleijlevens, B., Kramer, G., Florea, B. I., Hooibrink, B., Hollak, C. E., Ottenhoff, R., Boot, R. G., van der Marel, G. A., Overkleeft, H. S., and Aerts, J. M. (2010) Ultrasensitive in situ visualization of active glucocerebrosidase molecules, *Nat Chem Biol 6*, 907-913.

7. Feng, Y., Longo, D. L., and Ferris, D. K. (2001) Polo-like kinase interacts with proteasomes and regulates their activity, *Cell Growth Differ 12*, 29-37.

8. Strebhardt, K., and Ullrich, A. (2006) Targeting polo-like kinase 1 for cancer therapy, *Nat Rev Cancer 6*, 321-330.

9. Schubert, O. T., Mouritsen, J., Ludwig, C., Rost, H. L., Rosenberger, G., Arthur, P. K., Claassen, M., Campbell, D. S., Sun, Z., Farrah, T., Gengenbacher, M., Maiolica, A., Kaufmann, S. H., Moritz, R. L., and Aebersold, R. (2013) The Mtb proteome library: a resource of assays to quantify the complete proteome of Mycobacterium tuberculosis, *Cell Host Microbe 13*, 602-612.

10. Mallick, P., Schirle, M., Chen, S. S., Flory, M. R., Lee, H., Martin, D., Ranish, J., Raught, B., Schmitt, R., Werner, T., Kuster, B., and Aebersold, R. (2007) Computational prediction of proteotypic peptides for quantitative proteomics, *Nat Biotechnol 25*, 125-131.

11. Kettenbach, A. N., Rush, J., and Gerber, S. A. (2011) Absolute quantification of protein and post-translational modification abundance with stable isotope-labeled synthetic peptides, *Nat Protoc 6*, 175-186.

12. Ge, F., Xiao, C. L., Bi, L. J., Tao, S. C., Xiong, S., Yin, X. F., Li, L. P., Lu, C. H., Jia, H. T., and He, Q. Y. (2010) Quantitative phosphoproteomics of proteasome inhibition in multiple myeloma cells, *PLoS One 5*.

13. Mann, M., and Jensen, O. N. (2003) Proteomic analysis of post-translational modifications, *Nat Biotechnol 21*, 255-261.

14. Mertins, P., Qiao, J. W., Patel, J., Udeshi, N. D., Clauser, K. R., Mani, D. R., Burgess, M. W., Gillette, M. A., Jaffe, J. D., and Carr, S. A. (2013) Integrated proteomic analysis of post-translational modifications by serial enrichment, *Nat Methods 10*, 634-+.

15. Eichelbaum, K., Winter, M., Berriel Diaz, M., Herzig, S., and Krijgsveld, J. (2012) Selective enrichment of newly synthesized proteins for quantitative secretome analysis, *Nat Biotechnol 30*, 984-990.

16. Howden, A. J., Geoghegan, V., Katsch, K., Efstathiou, G., Bhushan, B., Boutureira, O., Thomas, B., Trudgian, D. C., Kessler, B. M., Dieterich, D. C., Davis, B. G., and Acuto, O. (2013) QuaNCAT: quantitating proteome dynamics in primary cells, *Nat Methods 10*, 343-346.

17. Schoffski, P. (2009) Polo-Like Kinase (PLK) Inhibitors in Preclinical and Early Clinical Development in Oncology, *Oncologist 14*, 559-570.

18. Wu, W. G., Pasternack, L., Huang, D. H., Koeller, K. M., Lin, C. C., Seitz, O., and Wong, C. H. (1999) Structural study on O-glycopeptides: Glycosylation-induced conformational changes of O-GlcNAc, O-LacNAc, O-sialyl-LacNAc, and O-sialyl-lewis-X peptides of the mucin domain of MAdCAM-1, *J Am Chem Soc 121*, 2409-2417.

19. Liang, F. C., Chen, R. P., Lin, C. C., Huang, K. T., and Chan, S. I. (2006) Tuning the conformation properties of a peptide by glycosylation and phosphorylation, *Biochem Biophys Res Commun 342*, 482-488.

20. van Kasteren, S. I., Kramer, H. B., Jensen, H. H., Campbell, S. J., Kirkpatrick, J., Oldham, N. J., Anthony, D. C., and Davis, B. G. (2007) Expanding the diversity of chemical protein modification allows post-translational mimicry, *Nature 446*, 1105-1109.

21. van Kasteren, S. I., Kramer, H. B., Gamblin, D. P., and Davis, B. G. (2007) Site-selective glycosylation of proteins: creating synthetic glycoproteins, *Nat Protoc 2*, 3185-3194.

22. Lazarus, M. B., Nam, Y., Jiang, J., Sliz, P., and Walker, S. (2011) Structure of human O-GlcNAc transferase and its complex with a peptide substrate, *Nature 469*, 564-567.

Samenvatting

Activiteits-gebaseerde proteasoom profilering

Het werk beschreven in dit proefschrift beslaat het ontwikkelen van chemisch biologische gereedschappen en methodes voor het bestuderen van het Ubiquitine Proteasoom Systeem (UPS). In **hoofdstuk 1** wordt een algemeen overzicht gegeven van het UPS en de activiteits-gebaseerde proteasoom profileringsmethodes beschreven in dit proefschrift. Het UPS is het centrale eiwit afbraak systeem in de cel dat zowel in het cytoplasma als in de cel kern te vinden is. Ubiquitine is een klein eiwit (8 kDa) dat als ketens aan de te af te breken eiwitten gekoppeld wordt waardoor deze naar het proteasoom vervoerd worden voor afbraak. Het proteasoom is evolutionair geconserveerd in eukaryoten van gist tot mens en is in het afgelopen decennium gevalideerd als een belangrijke drug target voor de behandeling van verschillende bloed kanker types. Methodes om de proteasoom activiteit te bepalen zijn nodig voor zowel fundamenteel als (pre) klinisch onderzoek.

In **hoofdstuk 2** wordt een literatuur overzicht gegeven van de toepassing van activiteits-gebaseerde eiwitprofilering (activity-based protein profiling, ABPP) in biologisch onderzoek van de afgelopen 2 jaar. ABPP is een chemisch biologische methode waarbij enzym activiteiten bepaald kunnen worden door middel van activiteits-gebaseerde sondes (activity-based probes, ABP). Een ABP is een stof die een enzym herkenningsmotief en een reporter groep (bijvoorbeeld een fluorescerende groep) draagt en die in staat is met een enzym te reageren waarbij een covalente en irreversibele binding met het katalytische aminozuur aangegaan wordt. Enkele voorbeelden van ABPP platformen worden gegeven waarbij ABPP gekoppeld werd aan LC/MS analyses om de globale reactiviteit van cysteines in het humane proteoom te bepalen; ABPP platformen voor visualisatie van enzym activiteit in gel of non-invasief in levende dieren voor de vroege diagnosis van tumoren; voor observatie van asymmetrische T cel deling of als methode voor screening in medicinale chemie voor de profilering van verschillende enzym klassen als kinases, serine hydrolases en glucocerebrosidases.

In **hoofdstuk 3** wordt een relatieve kwantificatie ABPP methode voor het proteasoom beschreven dat uitgevoerd kan worden in een complexe matrix. Het proteasoom bevat drie actieve subunits die eiwitten kunnen afbreken met voorkeur voor bepaalde aminozuren. De zogenaamde β1 subunit knipt C-terminaal van zure aminozuren, β2 heeft een voorkeur voor basische aminozuren en de β5 subunit knipt na hydrofobe, kleine of grote aminozuren. Immuun competente cellen brengen naast deze constitutieve subunits ook de immunoproteasoom versies daarvan tot expressie die β1i, β2i en respectievelijk β5i genoemd worden. Om de activiteit van ieder individuele subunit te bepalen zijn subunit specifieke ABPs gebruikt voor de β1/1i, β2/2i en β5/5i paren die

afgelezen werden met fluorescentie waarbij de intensiteit van het signaal een maat is voor de hoeveelheid activiteit van de subunit. Daarnaast is een affiniteitzuivering protocol ontwikkeld waarbij een pan-reactieve ABP voorzien van een biotine reporter gebruikt werd om de actieve proteasoom subunits te binden aan de oppervalk van paramagnetische streptavidine beads. De beads werden gevangen in een magnetisch veld, gewassen van aspecifiek bindende eiwitten en de proteasoom subunits werden gedigesteerd tot peptiden met trypsine. De peptiden werden gelabeld met formaldehyde waardoor twee methyl groepen geïnstalleerd werden op zowel de vrije N-terminus als op de ε-amine van lysine. Door "licht" en "zwaar" (^{13}C en deuterium verijkt) formaldehyde te gebruiken kunnen twee proteasoom samples met elkaar vergeleken worden tijdens de LC/MS analyse. Het voordeel van deze methode is dat alle actieve proteasoom subunits in één analyse run bepaald kan worden.

Hoofdstuk 4 is een technisch rapport omtrent het gebruik van azide ligatie hengsels op primaire en secundaire posities in een ABP. In geval dat de reporter in de ABP de binding aan het enzym of de cel permeabiliteit verlaagd kan er via bio-orthogonale ligatie de reporter op een later tijdstip geïntroduceerd worden. De reporter wordt dus vervangen door een azide dat biologisch inert is. Na cel lysis kan de azide reageren met een alkyn in een zogeheten klik reactie, of met een phosphine groep in geval voor Staudinger-Bertozzi ligatie gekozen wordt. Deze bio-orthogonale reacties vinden plaats in waterige milieus en geven niet veel achtergrond labeling van andere functionaliteiten als SH of COOH groepen. Het is gebleken dat onder eiwit denaturerende condities beide bio-orthogonale reacties goed verlopen zowel met primaire als secundaire azides, hoewel onder natieve omstandigheden de secundaire azides een stuk minder goed reageerden dan azides op een primaire positie.

In **hoofdstuk 5** is de kwantitatieve ABPP methode uit hoofdstuk 3 gebruikt in een poging om de moleculaire mechanismes voor de werking en het ontwikkelen van resistentie tegen de proteasoom inhibitor Bortezomib te verklaren. Daarnaast is ook gekeken naar het proteoom wijd effect van Bortezomib adaptatie van gevoelige B-cell kanker types in zogeheten HL60 en AMO1 cel kweek modellen. Het is gebleken dat Bortezomib aangepaste cellen (HL60A en AMO1A) meer constitutief proteasoom tot expressie brengen en daarnaast de eiwit synthese onderdrukken zowel op transcriptie (van DNA naar mRNA) als op translatie (van mRNA naar eiwit) niveau. Naast deze effecten zijn tal van andere cellulaire aanpassingen gevonden die interessante vraagstukken oproepen voor verder onderzoek.

In **hoofdstuk 6** wordt de proteasoom ABPP methode waarbij biotine-epoxomicin als ABP gebruikt wordt om aan te tonen dat de nieuw ontdekte β5t proteasoom subunit, die exclusief tot expressie wordt gebracht in de thymus, daadwerkelijk katalytisch actief is. T cellen worden door de thymus zowel positief als negatief geselecteerd en de β5t subunit is waarschijnlijk betrokken bij positieve selectie omdat het "lichaamsvreemde" peptides genegeerd. Met behulp van massaspectrometrie werd het peptide fragment dat de reactieve beta5t threonine covalent

gebonden aan de ABP gedetecteerd, waarmee de intrinsieke reactiviteit van beta5t onomstotelijk bewezen werd.

In **hoofdstuk 7** wordt ingegaan op de invloed van de post-translationele modificatie (PTM) O-GlcNAcelyring op de functie van de ubiquitine receptor en shuttle eiwit hHR23B. hHR23B is betrokken zowel bij DNA schade herstel door binding en stabilisering van het XPC herstel eiwit dat DNA laesies herkent als bij het vervoeren van geubiquitineerde eiwitten naar het proteasoom. Tijdens een globale screen naar O-GlcNAcelyring werd het hHR23B eiwit prominent aanwezig gevonden en vervolgens werd gepoogd de functie van de PTM te begrijpen. Om de positie van de PTM te bepalen werden zowel chemisch biologische als LC-MS trucs toegepast waaruit bleek dat O-GlcNAcelyring op hHR23B vooral in de verbindingstukken die de functionele domeinen verbindt voorkwamen. hHR23B bevat een ubiquitine-achtig (ubiquitin-like, UBL) en twee ubiquitine-geassocieerde (ubiuitin-associated, UBA) domeinen naast een XPC bindingsdomein. Er werd gespeculeerd dat hHR23B inter-moleculaire bindingen aan kan gaan tussen de UBL en UBA domeinen en de resultaten suggereren dat de aanwezigheid van O-GlcNAcelyring het evenwicht tussen deze twee vormen kan verstoren ten gunste van de ene over de andere. Dit preliminaire resultaat suggereert dat de cel in staat is om met een metabole switch de mate van proteasoom afbraak te reguleren.

总结

蛋白酶体活性表达谱

本书主要讨论了研究泛素-蛋白酶体系统的化学生物学方法。

第一章简要概括了泛素-蛋白酶体系统的生物学功能和蛋白酶体活性表达谱方法。泛素-蛋白酶体系统是真核细胞中最重要的蛋白质代谢机制。其中泛素是用于标记细胞内需要被降解的蛋白质；而蛋白酶体则通过识别蛋白上的泛素，对蛋白质进行有规则的水解。真核细胞内大部分的蛋白质都可以被该系统降解成多肽，进而被各种氨肽酶分解成氨基酸而用于新蛋白质的合成。基于该系统的重要作用，蛋白酶体已经被开发成用于治疗诸如癌症和自身免疫等疾病的药物靶点。因此，研究对蛋白酶体活性进行检测的方法已经成为临床和基础研究中的重要议题。

第二章是一篇关于最近几年内蛋白质活性表达谱（ABPP）方法在生物医学及相关学科中应用的文献综述。蛋白质活性表达谱是一种用于直接检测有活性的酶在细胞或其他生物体系中存在量的功能蛋白质组学方法。它是化学生物学发展最重要的驱动力之一，也是有机化学对于生物学研究贡献最明显的领域。这项技术是基于抑制物的反应活性基团与酶活性位点进行不可逆的共价结合，利用事先修饰在抑制物上的荧光基团或亲和标记物对被目标酶进行标记或者亲和纯化。近几年，蛋白质活性表达谱研究从比较易于标记的目标酶（比如丝氨酸水解酶，半胱氨酸和苏氨酸蛋白水解酶）向更加难以标记的目标酶转移。这些较难标记的目标酶包括在生物体中存在丰度比较低的酶、活性中心不是亲核氨基酸残基的酶以及底物比较特殊的酶。同时，蛋白质活性表达谱已经对临床医学研究产生了明显的推动作用。

第三章介绍了一个已经建立完善的用于相对定量检测蛋白酶体活性的方法。一方面，可以利用特异识别蛋白酶体亚基的小分子荧光探针（ABP）对活细胞中有活性的蛋白酶体进行定量检测，进而测定被蛋白酶体抑制物硼替佐米处理过的细胞中残留的蛋白酶体活性的比例。另一方面，可以利用生物素修饰的小分子探针标记细胞蛋白提取液中的蛋白酶体活性亚基并进行亲和纯化。而对于活细胞内的蛋白酶体，可以利用一个易于穿过细胞膜的叠氮探针进行标记，在提取细胞蛋白以后，再通过生物正交化学的方法将生物素偶联到探针上，并进行亲和纯化。然后利用高效液相色谱质谱联用技术（LC-MS）准确地对富集的蛋白酶体亚基的活性位点进行鉴定。同时，通过稳定同位素标记的方法，可以在同一个实验中对蛋白酶体的所有活性亚基进行相对定量。

第四章是关于生物正交化学在蛋白酶体活性表达谱实验中应用的方法介绍。由于在某些情况下，标签基团（如荧光基团或生物素）会影响小分子探针的活性、选择性或细胞渗透性等性质，限制了探针的应用。我们用生物正交反应基团在探针分子中代替标签基团，在对酶进行第一步活性标记以后，通过生物正交反应加装标记物，这种方法也被称作两步法蛋白质活性表达谱。叠氮作为一种重要的生物正交官能团被广泛的应用在两步法之中。通过这个章节，我们比较了铜催化 Huisgen 叠氮-炔环加成和 Staudinger-Bertozzi 这两种生物正交反应，在蛋白变性和未变性两种条件下对不同位置叠氮的标记效率。

第五章主要介绍了用定量蛋白酶体活性表达谱（第三章）和定量蛋白质组学相结合的方法对骨髓瘤细胞形成硼替佐米抗药性分子机制进行的探索。从 2003 年开始硼替佐米被应用于多发性骨髓瘤临床治疗并在疗效上取得了巨大的成功。但是另一方面在大量病人接受治疗的过程中，体内的癌细胞也对硼替佐米产生了抗药性。所以发现并解释骨髓瘤细胞产生硼替佐米抗药性的分子机理变得十分重要。我们发现在我们所研究的骨髓瘤细胞的抗药性形成机制是复杂的细胞通路间此消彼长的相互作用，而不是简单的靶基因突变。

第六章介绍了利用蛋白酶体活性表达谱方法研究一种最近发现的蛋白酶体亚基的工作。2007 年日本科学家在小鼠的胸腺中发现了一种新的 20S 蛋白酶体，在这种蛋白酶体中，免疫蛋白酶体亚基 β5i 被一个新亚基替代。这个新的亚基被命名为 β5t（t 代表胸腺），而这种新发现的蛋白酶体被命名为胸腺蛋白酶体。随后他们又发现胸腺蛋白酶体对于 CD8+ T 淋巴细胞的阳性选择具有重要作用。这个章节揭示了 β5t 是具备蛋白水解活性的，而且它的蛋白水解活性与其他两个同源的蛋白水解酶 β5 和 β5i 有所不同。β5t 更倾向于水解那些包含较多亲水氨基酸的多肽序列。

第七章描述了对于 O-GlcNAc 糖基化修饰和泛素化蛋白转运子 hHR23B 功能之间关系的研究。O-GlcNAc 糖基化是一种广泛存在于高等真核生物细胞核与细胞质中的蛋白质翻译后修饰。与多数糖基化修饰不同，O-GlcNAc 为单糖修饰，而且比较活跃不稳定。在这一点上它的性质与磷酸化更相似。另一方面由于 O-GlcNAc 也修饰在蛋白质裸露的丝氨酸和苏氨酸残基上，它有时也与磷酸化表现出一种竞争性的关系。我们利用一系列化学生物学方法证明了 hHR23B 上确实存在 O-GlcNAc 修饰。同时实验结果也表明通过这些糖基化修饰，hHR23B 的构型可能会发生变化，从而影响它与其他蛋白之间的相互作用和它的功能。

第八章对整本论文的实验工作进行了归纳和总结，并对未来的研究工作进行了展望。

Curriculum Vitae

Nan Li was born on 8th October 1984 in Fushun, China. In 2003, he graduated from the second high school of Fushun. During this period, he got interested in chemistry and biology subjects. Thereafter, he followed a BSc program in the Department of Chemical Biology, Xiamen University, China and obtained his bachelor degree in 2007. He moved to Leiden University, the Netherlands to pursue his MSc study in Biological Chemistry with financial support of a traveling grant from the Chinese government and scholarship awarded by the Leiden University Fund (LUF). During his master study, he worked on the role of transcription factor TFB5 in yeast nucleotide excision repair in the Molecular Genetics group under the supervision of Dr. Riekje Brandsma.

In a project in the Bio-organic Synthesis group under the supervision of Dr. Bobby Florea, he revealed the activity of thymoproteasome specific subunit β5t with activity-based probes in. He obtained the master degree in 2009.

From September 2009 to December 2013, he worked as a PhD student in the group of Bio-organic Synthesis, Leiden Institute of Chemistry under supervision of Prof. Dr. Herman Overkleeft and Dr. Bobby Florea. In 2010 and 2012, he presented posters at the 3rd and 4th EMBO Chemical Biology meetings in Heidelberg, Germany.

List of Publications

Relative quantification of proteasome activity by activity-based protein profiling and LC-MS/MS.
Li N, Kuo CL, Paniagua G, van den Elst H, Verdoes M, Willems LI, van der Linden WA, Ruben M, van Genderen E, Gubbens J, van Wezel GP, Overkleeft HS and Florea BI.
Nat Protoc. 2013 Jun;8(6):1155-68.

The antimalarial natural product symplostatin 4 is a nanomolar inhibitor of the food vacuole falcipains.
Stolze SC, Deu E, Kaschani F, **Li N**, Florea BI, Richau KH, Colby T, van der Hoorn RA, Overkleeft HS, Bogyo M, Kaiser M.
Chem Biol. 2012 Dec 21;19(12):1546-55.

Peptido sulfonyl fluorides as new powerful proteasome inhibitors.
Brouwer AJ, Jonker A, Werkhoven P, Kuo E, **Li N**, Gallastegui N, Kemmink J, Florea BI, Groll M, Overkleeft HS, Liskamp RM.
J Med Chem. 2012 Dec 27;55(24):10995-1003.

Identification and isolation of lantibiotics from culture: a bioorthogonal chemistry approach.
Li J, Girard G, Florea BI, Geurink PP, **Li N**, van der Marel GA, Overhand M, Overkleeft HS, van Wezel GP.
Org Biomol Chem. 2012 Nov 21;10(43):8677-83.

Triple bioorthogonal ligation strategy for simultaneous labeling of multiple enzymatic activities.
Willems LI, **Li N**, Florea BI, Ruben M, van der Marel GA, Overkleeft HS.
Angew Chem Int Ed Engl. 2012 Apr 27;51(18):4431-4.

Activity-based protein profiling: an enabling technology in chemical biology research.
Li N, Overkleeft HS, Florea BI.
Curr Opin Chem Biol. 2012 Apr;16(1-2):227-33.

Discovery of a potent and highly β1 specific proteasome inhibitor from a focused library of urea-containing peptide vinyl sulfones and peptide epoxyketones.
van der Linden WA, Willems LI, Shabaneh TB, **Li N**, Ruben M, Florea BI, van der Marel GA, Kaiser M, Kisselev AF, Overkleeft HS.
Org Biomol Chem. 2012 Jan 7;10(1):181-94.

Two-step bioorthogonal activity-based proteasome profiling using copper-free click reagents: a comparative study.
van der Linden WA, **Li N**, Hoogendoorn S, Ruben M, Verdoes M, Guo J, Boons GJ, van der Marel GA, Florea BI, Overkleeft HS.
Bioorg Med Chem. 2012 Jan 15;20(2):662-6.

Bioorthogonal chemistry: applications in activity-based protein profiling.
Willems LI, van der Linden WA, **Li N**, Li KY, Liu N, Hoogendoorn S, van der Marel GA, Florea BI, Overkleeft HS.
Acc Chem Res. 2011 Sep 20;44(9):718-29.

The natural product hybrid of Syringolin A and Glidobactin A synergizes proteasome inhibition potency with subsite selectivity.
Clerc J, **Li N**, Krahn D, Groll M, Bachmann AS, Florea BI, Overkleeft HS, Kaiser M.
Chem Commun (Camb). 2011 Jan 7;47(1):385-7.

Activity-based profiling reveals reactivity of the murine thymoproteasome-specific subunit beta5t.
Florea BI, Verdoes M, **Li N**, van der Linden WA, Geurink PP, van den Elst H, Hofmann T, de Ru A, van Veelen PA, Tanaka K, Sasaki K, Murata S, den Dulk H, Brouwer J, Ossendorp FA, Kisselev AF, Overkleeft HS.
Chem Biol. 2010 Aug 27;17(8):795-801.

A cleavable linker based on the levulinoyl ester for activity-based protein profiling.
Geurink PP, Florea BI, **Li N**, Witte MD, Verasdonck J, Kuo CL, van der Marel GA, Overkleeft HS.
Angew Chem Int Ed Engl. 2010 Sep 10;49(38):6802-5.