Probing proteasome activity and function : cancer diagnostics and mechanism of antigen processing

Berkers, C.R.

Probing proteasome activity and function

Cancer diagnostics and mechanism of antigen processing

Celia Rosita Berkers

Probing proteasome activity and function

Cancer diagnostics and mechanism of antigen processing

PROEFSCHRIFT

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof.mr. P.F. van der Heijden,
volgens besluit van het College voor Promoties
te verdedigen op dinsdag 5 oktober 2010
klokke 13.45 uur

door

Celia Rosita Berkers

geboren te Amsterdam
in 1980

Promotiecommissie

Promotores:	Prof.dr. H.S. Overkleeft Prof.dr. J.J. Neefjes
Copromoter:	Dr. H. Ovaa
Overige leden:	Prof.dr. T. Schumacher Prof.dr. J. Brouwer Dr. E. Reits Dr. R. Wolthuis

Cover by Buro buurvrouw (www.burobuurvrouw.nl)

The work described in this thesis was performed at the Divisions of Cellular Biochemistry and Cell Biology II of the Netherlands Cancer Institute, Amsterdam, The Netherlands. Financial support was provided by the Dutch Cancer Society (KWF), grant nr. NKI 2005-3368, the Koninklijke Hollandse Maatschappij der Wetenschappen and Stichting Fonds Doctor Catharine van Tussenbroek.

Publication of this thesis was financially supported by the Dutch Cancer Society (KWF) and the Netherlands Cancer Institute. Printed by Wöhrmann Print Service, Zutphen.

Table of contents

Preface

In cells, proteins are continuously synthesized and degraded to control protein levels and thereby regulate a wide variety of biochemical processes. The proteasome is the main cellular degradation machinery, responsible for the degradation of key proteins involved in the regulation of a wide range of cellular processes, including quality control, cell cycle progression, cell differentiation, signal transduction and apoptosis. In addition, the proteasome is involved in immunity. A broad repertoire of antigenic peptides is presented at the cell surface to the immune system by major histocompatibility complexes (MHC) class I, enabling the immune system to monitor changes in intracellular protein content and sense malignant transformation or the presence of viral proteins. Recognition of foreign or abnormal peptides by CD8+ cytotoxic T cells ultimately results in the elimination of the antigen presenting cell. Both foreign and self MHC class I antigens are generated by the proteasome.

The proteasome has been identified as an attractive pharmacological target. Inhibition of the proteasome causes disruption of many regulatory processes, eventually leading to cell death. The observation that many cancer cells are more sensitive to proteasome inhibition than normal cells has led to the development of several proteasome inhibitors for the treatment of cancer. The drug velcade® with the proteasome inhibitor bortezomib as its active ingredient is used for the treatment of multiple myeloma and mantle cell lymphoma, while many clinical trials in other malignancies are ongoing. The clinical success of bortezomib therapy has initiated the development of several second-generation proteasome inhibitors, which differ in their mode of inhibition and downstream mechanism of action. With several proteasome inhibitors becoming available for clinical use, the outstanding challenge is to define optimal treatment regimes and to choose the most suitable proteasome inhibitor for treatment.

Proteasome activity in its broadest sense is the central theme of this thesis. In **chapter 1**, we describe the development and characterization of different chemical proteasome activity probes that can be used to study proteasome activity in a wide range of tissue types using a variety of assays. With the proteasome emerging as a therapeutic target for cancer treatment, accurate tools for monitoring proteasome activity are in demand. Most methods however require prior cell lysis before activity measurements can be performed, or can only be used in genetically modified cells or organisms. **chapter 1.2** describes the development of a dansylated proteasome activity probe and SDS-PAGE based assays to monitor proteasome activity. This probe is the first in its class that could be used to study the activity of velcade® in intact cells conveniently without the need for laborious radiolabeling procedures. In **chapter 1.3**, fluorescent versions of this probe are described that allow a direct readout in SDS-PAGE based assays, circumventing the need for an antibody, while they can also be used for confocal microscopy and flow cytometry-based assays

to detect proteasome activity in cells directly.

Chapter 2 focuses on different uses of these fluorescent proteasome activity probes. We assess the effects of different proteasome inhibitors in preclinical and clinical studies and use these probes to purify proteasome and characterize its activity and composition. In **chapter 2.1**, we compare the effects of bortezomib and the novel proteasome inhibitor CEP-18770 in mouse tissues and in cell lines with acquired resistance to bortezomib due to a specific mutation in one of the proteasome active sites. In **chapter 2.2** we study the effects of marizomib therapy on patient blood samples and in the following chapter (**2.3**), we describe an activity-guided proteasome purification and determine the relative subunit composition in purified bovine liver and spleen proteasomes by mass spectrometry.

Proteasome activity is studied on a molecular level in **chapter 3** and the role of the proteasome in immunity is addressed. Recent studies have identified novel MHC class I epitopes in which two distant parts of a protein are excised and ligated together to form a novel peptide. These non-contiguous antigens are thought to be produced by the proteasome via a transpeptidation mechanism. In **chapter 3.2**, the molecular mechanism of proteasomal antigen splicing is studied, and rules that predict proteasomal antigen splicing are deduced. In contrast to current thoughts we show that this process can be remarkably efficient. Finally, in **chapter 3.3**, we show that proteasome splicing can produce a novel type of antigen, containing an isopeptide linkage.

Chapter 1

Development of chemical proteasome activity probes

Chapter 1.1

Drug discovery and assay development in the ubiquitin–proteasome system

Biochemical Society Transactions
Volume 38, 2010, Pages 14-20

Drug discovery and assay development in the ubiquitin–proteasome system

Celia R. Berkers[*] & Huib Ovaa[*]

Division of Cell Biology II, The Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands.

The observation that tumour cells are more sensitive to pharmacological inhibition of the proteasome than normal cells has led to the development of the proteasome inhibitor bortezomib. To date, this is the only proteasome inhibitor that has been approved for clinical use. The clinical success of bortezomib, combined with the occurrence of adverse effects and the development of clinical resistance against this compound, has initiated the development of a broad range of second-generation proteasome inhibitors as well as of assays that can be used to establish a relationship between the extent and type of proteasome inhibition and the effectiveness of a particular drug. In the present paper, we discuss new strategies that may be used in the future to overcome drug resistance and to broaden the use of proteasome inhibitors for the treatment of both cancer and infectious and autoimmune disease.

THE UPS (UBIQUITIN–PROTEASOME SYSTEM)

The UPS is responsible for the degradation of redundant and misfolded proteins, as well as for the turnover of many key regulatory proteins that control a wide range of cellular processes such as cell-cycle progression, cell differentiation, apoptosis, stress response, DNA repair and signal transduction.[1,2] Proteins that are destined for destruction are first post-translationally modified with a polyubiquitin chain via the consecutive action of three different types of ubiquitinating enzymes[3] (Figure 1). First, an E1 ubiquitin-activating enzyme activates ubiquitin in an ATP-dependent manner and then transfers it to an E2 ubiquitin-conjugating enzyme. In a final step, ubiquitin is conjugated to the target protein or a previously attached ubiquitin via the combined action of the E2 and an E3 ubiquitin ligase. The latter ligase is capable of interacting with both the E2 and the target protein and thereby confers substrate specificity on the ubiquitination cascade.

Sequential ubiquitination cycles lead to the formation of polyubiquitinated proteins, which are subsequently degraded by the 26S proteasome, a large threonine protease complex. The proteolytic activity of the 26S proteasome resides within its barrel-shaped 20S core, which is formed by four stacked rings of seven subunits each and has an overall architecture of $\alpha(1-7)\beta(1-7)\beta(1-7)\alpha(1-7)$. Within the two inner β-rings, three different catalytically active subunits, termed $\beta1$, $\beta2$ and $\beta5$, are responsible for the proteasome's caspase-like, trypsin-like and chymotrypsin-like activity, respectively. In lymphoid tissues and upon induction with interferon-γ, these constitu-

[*]Correspondence may be addressed to either author (c.berkers@nki.nl or h.ovaa@nki.nl).

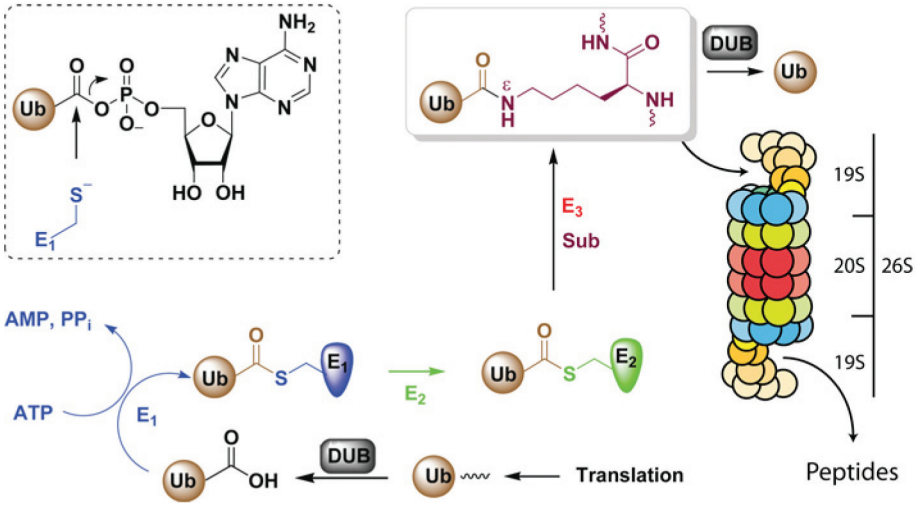


Figure 1 | Ubiquitin conjugation and deconjugation, and proteasomal protein breakdown. DUB, deubiquitinating enzyme; Ub, ubiquitin.

tive subunits can be replaced by their immunoproteasome counterparts, termed β1i, β2i and β5i, to form the immunoproteasome, whereas mixed-type proteasomes have also been reported.[4,5] The two outer α-rings provide stability to the 20S complex and serve as docking stations for the attachment of 19S regulatory complexes, which consist of a base that binds directly to the α-rings and which contain several ATPases. These 19S caps recognize and unfold the substrates and subsequently thread them into the 20S catalytic chamber for degradation.

PROTEASOME INHIBITORS AS ANTICANCER THERAPEUTICS

Proteasome inhibition is more cytotoxic to proliferating malignant cells than to quiescent normal cells, making the proteasome an attractive pharmacological target for the treatment of cancer.[6] Bortezomib[7] (velcade, PS341) (Figure 2) was the first proteasome inhibitor to enter the clinic, and has now been approved as a single agent for the treatment of relapsed/refractory MM (multiple myeloma) and mantle cell lymphoma, whereas many clinical trials in solid and haematological tumours are ongoing. Although the clinical success of bortezomib is evident, bortezomib treatment has been associated with serious adverse effects.[8] In addition, the occurrence of both primary and acquired resistance to bortezomib has been reported and although many haematological malignancies seem to respond well to proteasome inhibitor treatment, results in treatment of solid tumours have been disappointing.[9] This has initiated an ongoing search for novel proteasome inhibitors with distinct subunit specificity, lower toxicity or improved bioavailability.

Several second-generation proteasome inhibitors that vary in their mode of binding are now in clinical trials, including carfilzomib (PR-171),[10,11] NPI-0052 (salinosporamide A)[12] and CEP-18770[13,14] (Figure 2). For all inhibitors, the first step in proteasome inhibition is the deprotonation of the proteasomal Thr1 O^{γ}, which is thought to occur by the N-terminal amino group acting as a base, either di-

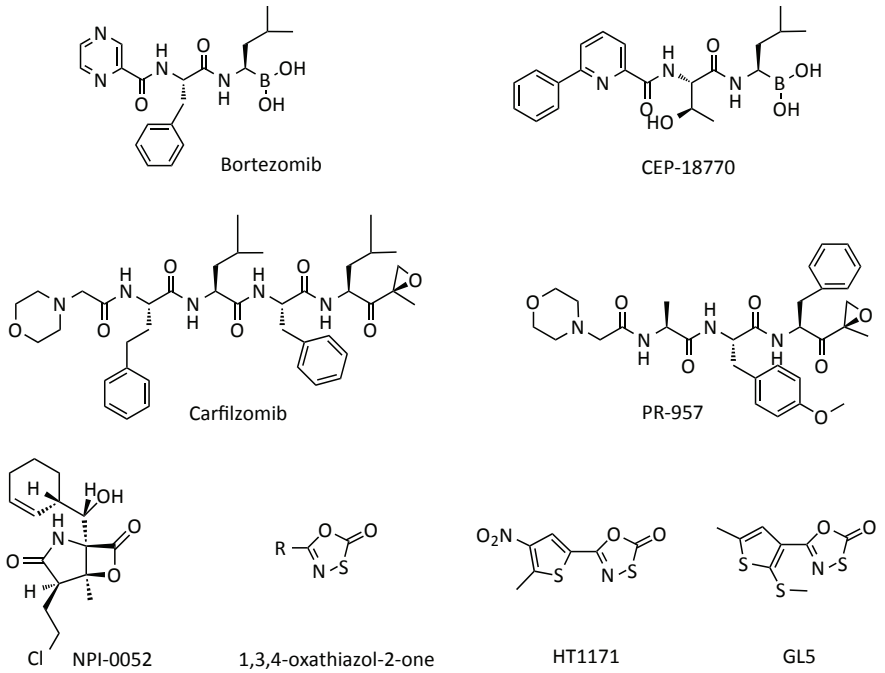


Figure 2 |Structures of (clinically relevant) 20S proteasome inhibitors.

rectly or via a neighbouring water molecule.[15] The structure of bortezomib complexed to the yeast 20S proteasome[16] has shown that for boronic acid-based inhibitors such as bortezomib and CEP-18770, subsequent nucleophilic attack of the Thr1 O^{γ} results in a covalent interaction between the boron atom and the Thr1 O^{γ} (Figure 3A). The proteasome–inhibitor complex is stabilized further by the Thr1 N, which forms a tight hydrogen bond with one of the acidic boronate hydroxyl groups.[16] In the case of epoxyketone-based inhibitors, such as carfilzomib, nucleophilic attack of Thr1 O^{γ} is thought to result in hemiacetal formation, followed by subsequent cyclization of Thr1 N onto the epoxide, resulting in the formation of a morpholino adduct (Figure 3B), as observed in the crystal structure of the 20S proteasome–epoxomicin complex.[17] NPI-0052 is a proteasome inhibitor isolated from the marine actinomycete *Salinispora*. The crystal structure of the 20S proteasome–NPI-0052 complex[18] suggests that nucleophilic attack by the catalytic Thr1 O^{γ} on the β-lactone ring results in the formation of an ester between the Thr1 O^{γ} and the carbonyl derived from the β-lactone ring (Figure 3C). Subsequent chlorine substitution then gives rise to a cyclic ether,[18] resulting in an adduct resistant to further hydrolysis.

The proteasome inhibitors that are currently in use in the clinic also vary in their subunit specificity, downstream mechanisms of action and bioavailability. Bortezomib and CEP-18770 both reversibly inhibit β1 and β5 subunits,[7,14,19] carfilzomib predominantly targets the β5 subunit,[11] whereas NPI-0052 irreversibly inhibits all constitutive and immunoproteasome subunits.[12] As a consequence, carfilzomib and CEP-18770 induce apoptosis via similar downstream mechanisms to that of bortezomib,[11,14,20] whereas NPI-0052 induces

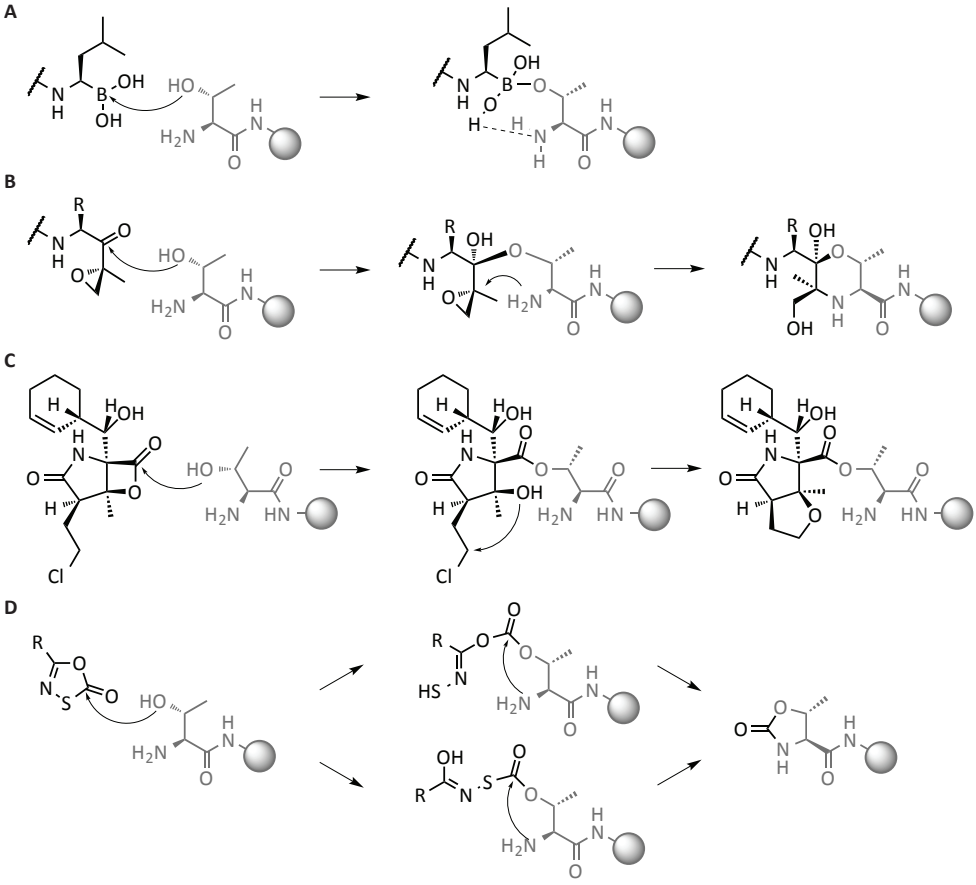


Figure 3 | Proposed mechanisms of proteasome inactivation by (clinically relevant) 20S proteasome inhibitors. Inhibitors are indicated in black, N-terminal threonine residues of active proteasomal β-subunits are in grey. Inactivation mechanism of **A)** boronic acid-based proteasome inhibitors (bortezomib and CEP-18770),[16] **B)** epoxyketone-based proteasome inhibitors (carfilzomib and PR-957),[17] **C)** β-lactone-based NPI-0052[18] and **D)** 1,3,4-oxathiazol-2-one compounds (mycobacterial-specific proteasome inhibitors).[47]

apoptosis also via different pathways.[12] Both carfilzomib and NPI-0052 were shown to be able to overcome bortezomib resistance, both in cell lines and in samples from patients with clinical bortezomib resistance.[11,12] NPI-0052 and CEP-18770 are orally bioavailable.[12,14]

PROTEASOME ACTIVITY ASSAYS

For effective treatment with proteasome inhibitors, it is essential to establish a relationship between the effectiveness of the drug in a particular tumour type and the extent and type of proteasome inhibition.[6] There is therefore an emerging need for proteasome activity assays that can profile the efficacy and specificity of proteasome inhibitors, both *in vitro* and in a cellular context.

Traditionally, different fluorogenic substrates, which typically consist of a short peptide and

a C-terminally attached fluorogenic reporter, are used to measure the activity of the different proteasome active sites. The proteasome cleaves these substrates between the last amino acid and the reporter, resulting in the release of a fluorescent molecule.[21] Fluorogenic substrates are widely used to measure effects of proteasome inhibitors,[22] but often only the chymotrypsin-like activity is measured.[21] It should be noted, however, that measuring only the chymotrypsin-like activity may not accurately reflect the degree of proteasome inhibition, and that the activities of all active sites need to be assayed in order to reliably evaluate the efficacy of inhibitors.[21,23] Most fluorogenic substrates cannot be used in cells, and prior cell lysis is required before activity measurements can be performed. This treatment isolates the proteasome from its regulatory environment,[6] and may lead to dissociation of the 19S caps. Efects of the 19S regulatory particle or other associated proteins may thus not be accounted for.[24] Furthermore, fluorogenic substrates cannot distinguish between constitutive and immunoproteasome activity, as immunoproteasome-specific substrates have not been described to date.[6]

To measure proteasome activity *in vivo*, both cell-based assays[25] and transgenic mouse models[24,26] have been developed that are based on recombinant proteasome substrates, typically ubiquitin–GFP (green fluorescent protein) or ubiquitin–luciferase fusion protein. Under normal cellular conditions, these fusion proteins are rapidly degraded by the proteasome, and no fluorescence or luciferase activity can be detected, but when the proteasome is inhibited, degradation is impaired and intracellular fluorescence increases. These methods and models are useful in screening for new inhibitors and for testing bioavailability, but their activity readout depends on the balance between synthesis and degradation of fusion proteins, which involves many cellular factors other than the proteasome, including the rate of fusion protein synthesis. Moreover, because they are based on reporter proteins, their use remains confined to genetically altered cells and organisms and does not allow profiling of patient material.

A third type of assay uses chemical proteasome probes that allow profiling of proteasome activity in cell lysates, cells and *ex vivo*.[11,19,27-30] In general, these probes consist of a reactive group that covalently modifies all active-site N-terminal threonine residues, which is attached to a peptide-recognition motif to target the probe to the proteasome. A tag allows for visualization or isolation of the proteasome–probe complex. An example of such a probe is dansyl-Ahx_3Leu_3vinylsulfone (where Ahx is aminohexanoic acid),[19] which was used successfully to profile bortezomib,[19] NPI-0052[12] and carfilzomib[11] in cells using an antibody against the dansyl moiety. More recently, chemical probes in which the dansyl moiety has been replaced with a fluorescent tag have been developed.[27,30] These probes are more sensitive and allow for easier visualization of proteasome activity either using SDS/PAGE or directly using confocal microscopy. Importantly, the incorporation of a fluorescent tag also enables the use of these probes in flow-cytometry-based assays,[27] which can be applied in a high-throughput fashion to screen for direct and indirect regulators of proteasome activity. Chemical proteasome probes have also been used for patient tumour profiling,[31] thereby confirming that these probes can be used to quantitatively and qualitatively monitor proteasome blockage in patients receiving proteasome inhibitor therapy.[6] Finally, proteasome active-site probes were used for the development of an ELISA-based activity assay.[11,28,29] This assay combines the use of a biotinylated epoxyketone-based proteasome

probe with subunit-specific antibodies to measure proteasome activities and has been used to profile the activity of carfilzomib and several recently developed immunoproteasome-specific inhibitors.

NOVEL ANTICANCER STRATEGIES

With the availability of a panel of different proteasome inhibitors, the outstanding clinical challenge will be to define treatment schedules that offer a therapeutic advantage over bortezomib treatment.[32] Different proteasome inhibitors may find use in different therapeutic settings. Specific inhibitors of the β5 subunit, such as bortezomib, carfilzomib or CEP-18770, may offer a large therapeutic window in some tumours, whereas pan-proteasome inhibitors that block all three activities may lead to improved proteasome inhibitor efficacy in others,[29,33] depending on subunit activity profile, tissue type and disease.

An interesting strategy to overcome resistance to the current generation of proteasome inhibitors may be the development of inhibitors that do not interfere with the 20S core, but with other elements of the UPS. An obvious strategy would be the development of inhibitors of the 19S regulatory particle, especially inhibitors of the 19S ATPases, which also play non-proteolytic roles.[34] The first compound of this type, RIP-1 (regulatory particle inhibitor peptoid-1) was recently identified and the 19S ATPase Sug2/Rpt4 was identified as its molecular target.[34,35] A more advanced UPS-modulating compound that holds promise for the treatment of cancer is MLN4924,[36] an inhibitor of NEDD8 (neural-precursor-cell-expressed developmentally down-regulated 8)-activating enzyme, an essential component of the NEDD8 conjugation pathway that, in turn, controls the activity of a subset of E3 ubiquitin ligases. MLN4924 treatment induced apoptosis in human tumour cells, and suppressed tumour growth in pre-clinical models.[36] Currently, MLN4924 is being evaluated in clinical trials.

Proteasome inhibitors can be combined with other proteasome inhibitors or chemotherapeutics to increase their therapeutic potential. The combination of bortezomib and NPI-0052 remarkably induced synergistic cytotoxicity,[12,33,37] indicating that combination therapy of proteasome inhibitors might improve patient outcome.[6] It is tempting to speculate that combining proteasome inhibitors that induce apoptosis via distinct mechanisms, such as bortezomib and NPI-0052, may be particularly beneficial.[6] Interesting in this respect is the recent identification of argyrin A,[38] a cyclic peptide derived from the mycobacterium *Archangium gephyra* that inhibits all proteasome activities. Induction of apoptosis by argyrin A was shown to be dependent on $p27^{kip1}$, but not on IκBα (inhibitor of nuclear factor κB α) stabilization, whereas the cytotoxic effects of bortezomib are mediated via IκBα stabilization, but are not dependent on p27.[38] Argyrin A may therefore represent another class of proteasome inhibitors that can be successfully combined with other proteasome inhibitors in anticancer therapy.

Many of the pathways used by tumour cells to survive chemotherapy are blocked by proteasome inhibition, which provided the rationale for combining bortezomib with a variety of chemotherapeutics.[9,39] The combination of bortezomib and PEGylated doxorubicin has been approved for the treatment of relapsed MM,[40,41] and many treatment regimens combining bortezomib and other proteasome inhibitors with conventional or novel anti-MM agents have yielded promising results.[6,9,42,43] Another strategy currently being evaluated is the simultaneous targeting of the proteasomal and lysosomal degradation pathways using a combination of bortezomib and HDAC (histone deacetylase) inhibitors.[6]

NEW FRONTIERS IN UPS DRUG DISCOVERY

Proteasome inhibitor therapy may not remain restricted to the treatment of cancer, but may also be used to treat a variety of other disorders. Recently, PR-957 (Figure 2), an epoxyketone-based inhibitor specific for the immunoproteasomal β5i subunit, was developed.[28] Immunoproteasome is found predominantly in cells of haemopoietic origin. Specific immunoproteasome inhibitors may therefore be used for the treatment of haemopoietic tumours, while causing fewer side effects in tissues that express only low levels of immunoproteasome subunits, such as neurotoxicity or gastrointestinal side effects.[9] Immunoproteasome inhibitors may also be useful for the treatment of other disorders. Increased immunoproteasome expression has been observed in autoimmune diseases such as rheumatoid arthritis[44] and has also been correlated to neurodegenerative disease states, including Huntington's and Alzheimer's disease.[45,46] Consistent with this, Muchamuel et al.[28] showed that PR-957 reversed signs of disease in mouse models of rheumatoid arthritis, providing a rationale for development of immunoproteasome inhibitors for the treatment of autoimmune diseases. Furthermore, various infectious diseases may be treatable with proteasome inhibitors. Lin et al.[47] identified two related 1,3,4-oxathiazol-2-one compounds, HT1171 and GL5 (Figure 2), as selective inhibitors of the proteasome of *Mycobacterium tuberculosis*. Nucleophilic attack of the proteasomal Thr1 O^{γ} on the oxathiazol-2-one is thought to result in the formation of a carbonated or carbonothioated enzyme intermediate on Thr1.[47] Subsequent nucleophilic attack of the Thr1 N results in cyclocarbonylation of the active-site threonine residue to form an oxazolidin-2-one and prevents regeneration of the active protease (Figure 3D), thereby killing *M. tuberculosis*.[47] It is important that these oxathiazol-2-one compounds did not significantly inhibit the human proteasome nor a panel of cysteine, serine or metallo-proteases. The selectivity of these compounds for *M. tuberculosis* over mammalian proteasomes seemed to be conferred by residues distant from the active site, suggesting that a functionally exploitable diversity exists between *M. tuberculosis* and human proteasomes and that compounds that are specific for different types of mycobacteria or a variety of other bacteria could also be developed, although *in vivo* data are required. This opens the way for the treatment of a variety of (myco)bacterial infections with proteasome inhibitors, either alone or in combination with other antibiotic strategies.

CONCLUSION

Proteasome inhibition has emerged as a novel strategy for the treatment of cancer, as is evident from the clinical success of bortezomib, the first proteasome inhibitor to pass clinical trials. The occurrence of side effects, the development of resistance and low success rates in the treatment of solid tumours, however, limits the use of this agent. The recent development of a broad range of second-generation proteasome inhibitors is likely to contribute to a more widespread use of this class of agents for the treatment of both haematological and solid tumours. The development of compounds that interfere with the UPS via different mechanisms and the development of proteasome inhibitors that are specific for proteasome subtypes or other species may broaden the clinical application of UPS-modulating compounds in cancer and infectious and autoimmune diseases.

ACKNOWLEDGEMENTS

We thank Farid El Oualid for help with Figures. This work was funded by the Dutch Cancer Society [grant number NKI 2005-3368] and Netherlands Organisation for Scientific Research: Earth and Life Sciences [grant number 58248].

REFERENCES

1. Adams, J. The development of proteasome inhibitors as anticancer drugs. *Cancer Cell* **5**, 417-421 (2004).
2. Nalepa, G., Rolfe, M. & Harper, J.W. Drug discovery in the ubiquitin-proteasome system. *Nat. Rev. Drug Discov.* **5**, 596-613 (2006).
3. Pickart, C.M. Mechanisms underlying ubiquitination. *Annu. Rev. Biochem.* **70**, 503-533 (2001).
4. Dahlmann, B., Ruppert, T., Kuehn, L., Merforth, S. & Kloetzel, P.M. Different proteasome subtypes in a single tissue exhibit different enzymatic properties. *J. Mol. Biol.* **303**, 643-653 (2000).
5. Dahlmann, B., Ruppert, T., Kloetzel, P.M. & Kuehn, L. Subtypes of 20S proteasomes from skeletal muscle. *Biochimie* **83**, 295-299 (2001).
6. Chauhan, D., Bianchi, G. & Anderson, K.C. Targeting the UPS as therapy in multiple myeloma. *BMC Biochem.* **9 (Suppl 1)**, -S1 (2008).
7. Hideshima, T. et al. The proteasome inhibitor PS-341 inhibits growth, induces apoptosis, and overcomes drug resistance in human multiple myeloma cells. *Cancer Res.* **61**, 3071-3076 (2001).
8. Richardson, P.G. et al. A phase 2 study of bortezomib in relapsed, refractory myeloma. *N Engl. J. Med.* **348**, 2609-2617 (2003).
9. Orlowski, R.Z. & Kuhn, D.J. Proteasome inhibitors in cancer therapy: lessons from the first decade. *Clin. Cancer Res.* **14**, 1649-1657 (2008).
10. Demo, S.D. et al. Antitumor activity of PR-171, a novel irreversible inhibitor of the proteasome. *Cancer Res.* **67**, 6383-6391 (2007).
11. Kuhn, D.J. et al. Potent activity of carfilzomib, a novel, irreversible inhibitor of the ubiquitin-proteasome pathway, against preclinical models of multiple myeloma. *Blood* **110**, 3281-3290 (2007).
12. Chauhan, D. et al. A novel orally active proteasome inhibitor induces apoptosis in multiple myeloma cells with mechanisms distinct from Bortezomib. *Cancer Cell* **8**, 407-419 (2005).
13. Dorsey, B.D. et al. Discovery of a potent, selective, and orally active proteasome inhibitor for the treatment of cancer. *J. Med. Chem.* **51**, 1068-1072 (2008).
14. Piva, R. et al. CEP-18770: A novel, orally active proteasome inhibitor with a tumor-selective pharmacologic profile competitive with bortezomib. *Blood* **111**, 2765-2775 (2008).
15. Groll, M. et al. The catalytic sites of 20S proteasomes and their role in subunit maturation: a mutational and crystallographic study. *Proc. Natl. Acad. Sci. U S A* **96**, 10976-10983 (1999).
16. Groll, M., Berkers, C.R., Ploegh, H.L. & Ovaa, H. Crystal structure of the boronic acid-based proteasome inhibitor bortezomib in complex with the yeast 20S proteasome. *Structure* **14**, 451-456 (2006).
17. Groll, M., Kim, K.B., Kairies, N., Huber, R. & Crews, C.M. Crystal structure of epoxomicin : 20S proteasome reveals a molecular basis for selectivity of alpha ',beta '-epoxyketone proteasome inhibitors. *J. Am. Chem. Soc.* **122**, 1237-1238 (2000).
18. Groll, M., Huber, R. & Potts, B.C. Crystal structures of Salinosporamide A (NPI-0052) and B (NPI-0047) in complex with the 20S proteasome reveal important consequences of beta-lactone ring opening and a mechanism for irreversible binding. *J. Am. Chem. Soc.* **128**, 5136-5141 (2006).
19. Berkers, C.R. et al. Activity probe for in vivo profiling of the specificity of proteasome inhibitor bortezomib. *Nat. Methods* **2**, 357-362 (2005).
20. Stapnes, C. et al. The proteasome inhibitors bortezomib and PR-171 have antiproliferative and proapoptotic effects on primary human acute myeloid leukaemia cells. *Br. J. Haematol.* **136**, 814-828 (2007).
21. Kisselev, A.F. & Goldberg, A.L. Monitoring activity and inhibition of 26S proteasomes with fluorogenic peptide substrates. *Methods Enzymol.* **398**, 364-378 (2005).
22. Lightcap, E.S. et al. Proteasome inhibition measurements: clinical application. *Clin. Chem.* **46**, 673-683 (2000).
23. Kisselev, A.F., Callard, A. & Goldberg, A.L. Importance of the different proteolytic sites of the proteasome and the efficacy of inhibitors varies with the protein substrate. *J. Biol. Chem.* **281**, 8582-8590 (2006).
24. Luker, G.D., Pica, C.M., Song, J., Luker, K.E. & Piwnica-Worms, D. Imaging 26S proteasome activity and inhibition in living mice. *Nat. Med.* **9**, 969-973 (2003).
25. Dantuma, N.P., Lindsten, K., Glas, R., Jellne, M. & Masucci, M.G. Short-lived green fluorescent proteins for quantifying ubiquitin/proteasome-dependent proteolysis in living cells. *Nat. Biotechnol.* **18**, 538-543 (2000).
26. Lindsten, K., Menendez-Benito, V., Masucci, M.G. & Dantuma, N.P. A transgenic mouse model of the ubiquitin/proteasome system. *Nat. Biotechnol.* **21**, 897-902 (2003).
27. Berkers, C.R. et al. Profiling proteasome activity in tissue with fluorescent probes. *Mol. Pharm.* **4**, 739-748 (2007).

28. Muchamuel, T. et al. A selective inhibitor of the immunoproteasome subunit LMP7 blocks cytokine production and attenuates progression of experimental arthritis. *Nat. Med.* **15**, 781-787 (2009).
29. Parlati, F. et al. Carfilzomib can induce tumor cell death through selective inhibition of the chymotrypsin-like activity of the proteasome. *Blood* **114**, 3439-47 (2009).
30. Verdoes, M. et al. A fluorescent broad-spectrum proteasome inhibitor for labeling proteasomes in vitro and in vivo. *Chem. Biol.* **13**, 1217-1226 (2006).
31. Kraus, M. et al. Activity patterns of proteasome subunits reflect bortezomib sensitivity of hematologic malignancies and are variable in primary human leukemia cells. *Leukemia* **21**, 84-92 (2007).
32. Hoeller, D. & Dikic, I. Targeting the ubiquitin system in cancer therapy. *Nature* **458**, 438-444 (2009).
33. Chauhan, D. et al. Combination of proteasome inhibitors bortezomib and NPI-0052 trigger in vivo synergistic cytotoxicity in multiple myeloma. *Blood* **111**, 1654-1664 (2008).
34. Lim, H.S., Archer, C.T. & Kodadek, T. Identification of a peptoid inhibitor of the proteasome 19S regulatory particle. *J. Am. Chem. Soc.* **129**, 7750-7751 (2007).
35. Lim, H.S., Cai, D., Archer, C.T. & Kodadek, T. Periodate-triggered cross-linking reveals Sug2/Rpt4 as the molecular target of a peptoid inhibitor of the 19S proteasome regulatory particle. *J. Am. Chem. Soc.* **129**, 12936-12937 (2007).
36. Soucy, T.A. et al. An inhibitor of NEDD8-activating enzyme as a new approach to treat cancer. *Nature* **458**, 732-736 (2009).
37. Roccaro, A.M. et al. Dual targeting of the proteasome regulates survival and homing in Waldenstrom macroglobulinemia. *Blood* **111**, 4752-4763 (2008).
38. Nickeleit, I. et al. Argyrin a reveals a critical role for the tumor suppressor protein p27(kip1) in mediating antitumor activities in response to proteasome inhibition. *Cancer Cell* **14**, 23-35 (2008).
39. Nencioni, A., Grunebach, F., Patrone, F., Ballestrero, A. & Brossart, P. Proteasome inhibitors: antitumor effects and beyond. *Leukemia* **21**, 30-36 (2007).
40. Orlowski, R.Z. et al. Phase 1 trial of the proteasome inhibitor bortezomib and pegylated liposomal doxorubicin in patients with advanced hematologic malignancies. *Blood* **105**, 3058-3065 (2005).
41. Orlowski, R.Z. et al. Randomized phase III study of pegylated liposomal doxorubicin plus bortezomib compared with bortezomib alone in relapsed or refractory multiple myeloma: combination therapy improves time to progression. *J. Clin. Oncol.* **25**, 3892-3901 (2007).
42. Fenical, W. et al. Discovery and development of the anticancer agent salinosporamide A (NPI-0052). *Bioorg. Med. Chem.* **17**, 2175-2180 (2009).
43. Sanchez, E. et al. The Novel Proteasome Inhibitor CEP-18770 Inhibits Myeloma Tumor Growth In Vitro and In Vivo and Enhances the Anti-MM Effects of Melphalan. *ASH Annual Meeting Abstracts* **112**, 843 (2008).
44. Egerer, T. et al. Tissue-specific up-regulation of the proteasome subunit beta5i (LMP7) in Sjogren's syndrome. *Arthritis Rheum.* **54**, 1501-1508 (2006).
45. Diaz-Hernandez, M. et al. Neuronal induction of the immunoproteasome in Huntington's disease. *J. Neurosci.* **23**, 11653-11661 (2003).
46. Mishto, M. et al. Immunoproteasome and LMP2 polymorphism in aged and Alzheimer's disease brains. *Neurobiol. Aging* **27**, 54-66 (2006).
47. Lin, G. et al. Inhibitors selective for mycobacterial versus human proteasomes. *Nature* **461**, 621-626 (2009).

Chapter 1.2

Activity probe for *in vivo* profiling of the specificity of proteasome inhibitor bortezomib

Nature Methods
Volume 2, 2005, Pages 357-362

Activity probe for *in vivo* profiling of the specificity of proteasome inhibitor bortezomib

Celia R. Berkers,[a,b,§] Martijn Verdoes,[a,§] Eben Lichtman,[a] Edda Fiebiger,[a] Benedikt M. Kessler,[a] Kenneth C. Anderson,[c] Hidde L. Ploegh,[a] Huib Ovaa[a,b,*] & Paul J. Galardy[a,d]

[a]Department of Pathology, Harvard Medical School, 77 Avenue Louis Pasteur, Boston, Massachusetts 02115, USA. [b]Division of Cellular Biochemistry, Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX Amsterdam, the Netherlands. [c]Department of Medical Oncology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, Massachusetts 02115, USA. [d]Program in Hematology/Oncology, Children's Hospital Boston and Dana-Farber Cancer Institute, 44 Binney Street, Boston, Massachusetts 02115, USA.

Proteasome inhibitors, such as the dipeptide boronic acid bortezomib, are emerging as important tools in the treatment of the fatal hematologic malignancy multiple myeloma. Despite the recent US Food and Drug Administration approval of bortezomib (PS341, Velcade) for the treatment of refractory multiple myeloma, many of the basic pharmacologic parameters of bortezomib and its mode of action on myeloma cells remain to be determined. We describe the synthesis and use of a cell-permeant active site–directed probe, which allows profiling of proteasomal activities in living cells. When we compared proteasome activity patterns in cultured cells and crude cell extracts with this probe, we observed substantial differences, stressing the importance for bioassays compatible with live cells to ensure accuracy of such measurements. Using this probe, we investigated the *in vivo* subunit specificities of bortezomib and another inhibitor, MG132.

INTRODUCTION

The proteasome inhibitor bortezomib (**2**) (PS341, Velcade; Figure 1) is among the most promising pharmaceuticals developed for the treatment of multiple myeloma.[1] The ability to target the proteasome pharmacologically without prohibitive toxicity is in itself remarkable: the proteasome is a large proteolytic complex, which is abundant in the cytosol and the nucleus, and it is considered central to many different aspects of cellular physiology, as polypeptides are continuously synthesized and destroyed to maintain steady-state protein levels.[2] Protein turnover by the proteasome has many critical regulatory roles, including quality control by removing abnormal proteins that arise by mutation, metabolic damage or misfolding.

Eukaryotic proteasomes consist of a 20S core, which is most often the central part of higher-order complexes, of which the 26S proteasome is the major active species (Figure 2). In 26S proteasomes, the 20S core particle is complexed at one or both ends with a 19S regulatory complex that modulates the entrance of substrates to the core and is required for the recognition of ubiquitin-tagged substrates and the unfolding of substrates prior to their threading into the proteolytic chamber.[3] Hy-

[§]These authors contributed equally to this work. [*]Correspondence should be addressed to H.O. (h.ovaa@nki.nl).

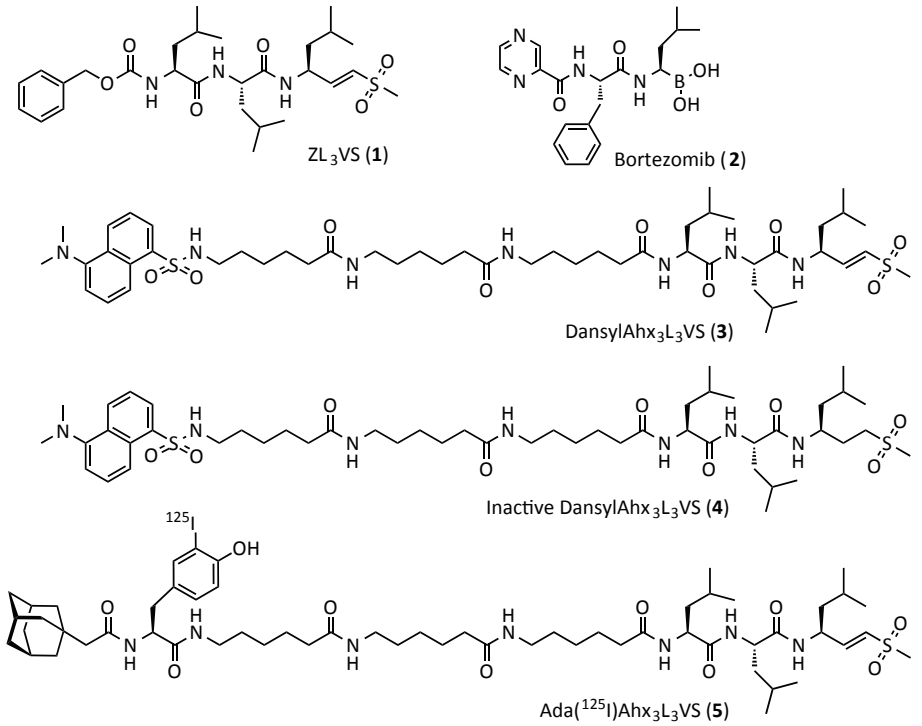


Figure 1 | Representative structures of bortezomib and vinyl sulfone–based proteasome probes.

brid proteasomes, in which different regulatory caps are attached to the 20S core, also exist. The 20S core is a hollow, barrel-shaped protein complex[4] composed of 14 α and 14 β subunits arranged in four stacked rings with an overall architecture of α(1–7)β(1–7)β(1–7)α(1–7). The two outer α rings form gated channels through which substrates enter and products exit the catalytic core. In addition, the outer α rings provide binding sites for regulatory complexes such as the 19S caps. The proteolytic activity of the 20S particle resides in the two inner β rings. The proteasome has three distinct catalytic activities, caspase-like, tryptic-like and chymotryptic-like, each of which can be roughly correlated to the action of an individual catalytic β subunit. These are termed β1, β2 and β5, respectively, and can be replaced by three interferon-inducible subunits, termed β1i, β2i and β5i to form immunoproteasomes—proteasomes with altered catalytic specificity favoring production of antigenic peptides. The secondary alcohol of the N-terminal threonine of the active β subunits acts as the nucleophilic species.

The dipeptide boronic acid bortezomib is a potent inhibitor of the proteasome's chymotryptic-like activity.[5] But variations in proteasomal composition alter the proteolytic specificities of the proteasome[6–8] and may therefore affect the sensitivity of the proteasome to the action of bortezomib. Little is known about the composition of active proteasomes in living multiple myeloma cells. Furthermore, most investigators measure proteasomal activity in cell extracts using fluorogenic substrates, and as a result, it has been difficult to assess the extent of inhibition of individual

catalytic activities *in vivo* after treatment with bortezomib. This underscores the need for improved tools that allow the accurate assessment of proteasomal activity *in vivo*.

To facilitate proteomic research, activity-based chemical probes have been developed for the study of serine, cysteine and threonine proteases.[9–11] These probes form covalent adducts with active proteases of the class that is targeted, leading to a convenient readout of enzyme activity. Most of these probes, however, can not be used in intact living cells.[12,13] Tags that are incorporated to provide an enzymatic activity readout, such as [^{125}I]iodotyrosine, often render probes impermeant to cell membranes, thereby restricting their use to cell extracts. Furthermore, they are nonspecific and target many proteases at the same time. Thus, such probes do not provide information on *in vivo* activities of specific proteases.

Peptide vinyl sulfone–based proteasome inhibitors[9,14] such as ZL_3VS (**1**) (Figure 1) modify catalytically active threonine residues through a Michael addition reaction, resulting in the formation of a covalent β-sulfonyl ether linkage. When a detectable tag is attached to such inhibitors, they can be used to label and detect proteasome active sites. Recently we set out to profile proteasome activity in multiple biological samples. A previously developed peptide vinyl sulfone–based inhibitor proved to be cell-permeant,[15] but the need for secondary chemical labeling procedures made this method unsuitable as a rapid screen of proteasome inhibition in multiple samples. In need for fast, reproducible activity readouts we reasoned that the use of small haptens, which can be recognized by antibodies, may allow retention of cell permeability of probes and allow detection of labeled enzymes directly by SDS-PAGE and Western blot analysis. Therefore we synthesized an inhibitor with a dansyl-sulfonamidohexanoyl hapten as a fast and easy method to monitor proteasome activity in living cells. Using this probe, we could profile active proteasome subunits in cell extracts and cultured cells. Furthermore, we investigated the *in vivo* subunit specificities of the proteasome inhibitors bortezomib and MG132. We found that the β1 and β1i subunits are additional targets for bortezomib and that MG132 inhibits all subunits.

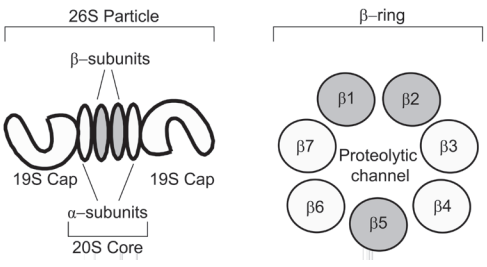


Figure 2 | Schematic of the 26S proteasome.

RESULTS

Synthesis and characterization of proteasome probes

We synthesized compound **3** (Figure 1) containing a dansyl-sulfonamidohexanoyl hapten, as well as control **4** in which the reactive vinyl sulfone moiety is reduced by hydrogenation, rendering the overall probe unreactive. Using a cell-based assay,[16] we tested proteasome inhibition by compounds **3** and **4**, and compared the inhibition to that of other known inhibitors. We exposed cells expressing a substrate fusion protein comprising ubiquitin, the amino acid arginine and green fluorescent protein (ubiquitin-Arg-GFP) to proteasome inhibitors. Efficient proteasome inhibition results in the accumulation of GFP, which can be quantitated by fluorescence activated cell sorting (FACS; Figure 3A,B). Whereas treatment with increasing amounts of **3** or Ada-Ahx_3Leu_3VS, the parent compound, resulted in a substantial increase in fluorescence, treatment with

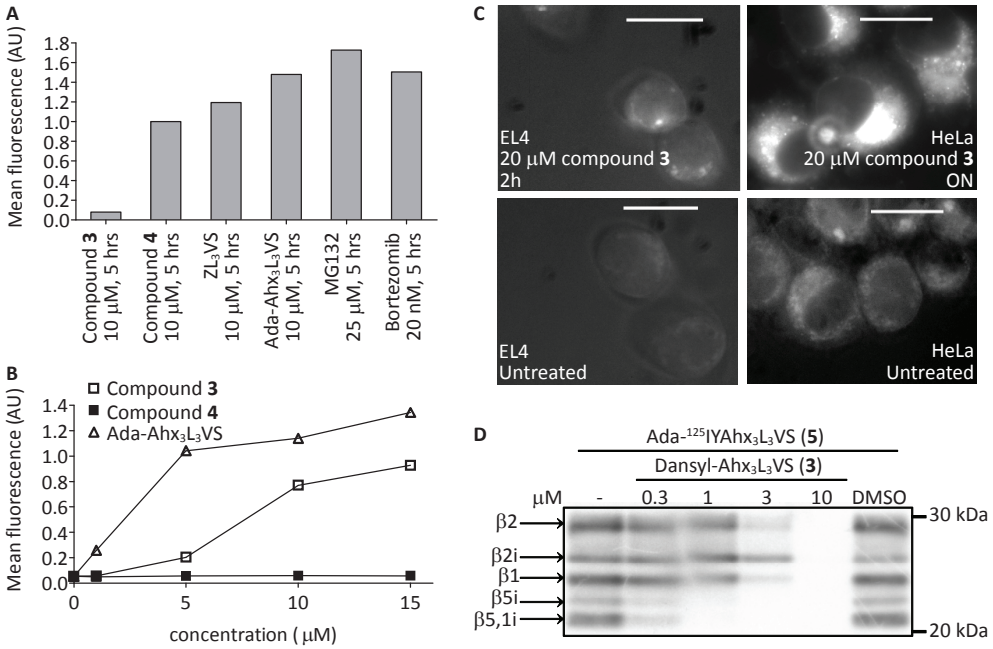


Figure 3 | Characterization of compounds 3 and 4. A,B) FACS analysis results of ubiquitin-Arg-GFP–expressing U373[12] cells upon incubation with proteasome probes and proteasome inhibitors. Incubation with compound **3** resulted in an increase in fluorescence, comparable to that of known proteasome inhibitors, whereas incubation with inactive compound **4** did not **(A)**. **B)** Graph shows the increase in fluorescence after a 5-h incubation with the indicated concentrations of compound **3** and the parent compound Ada-Ahx_3L_3VS, from which an approximate IC_{50} value could be estimated. Graphs are of representative results obtained in two independent experiments. **C)** Visualization by fluorescence microscopy of the distribution of compound **3** in EL4 and HeLa cells. Bar, 10 μm. **D)** Competition experiment between inhibitor **3** and inhibitor **5**. Autoradiograph showing the proteasome subunits that have reacted with **5**.

the control inhibitor **4** was ineffective. The approximate half-maximal inhibitory concentration (IC_{50}) values for Ada-Ahx_3Leu_3VS and **3** were 3 μM and 8 μM, respectively (Figure 3B). Thus, although the potency and/or bioavailability of compound **3** is somewhat lower than that of the parent compound, it is by far sufficient as a probe for activity measurements. By observing the relatively weak fluorescence emitted by the dansyl moiety of inhibitor **3**, we confirmed cell permeability of **3** by fluorescence microscopy (Figure 3C). After incubation with 20 μM of **3**, we visualized fluorescence throughout the cytosol and nucleus of cultured HeLa and EL4 cells. The microscopy results conclusively demonstrate the cell permeability of inhibitor **3**. We next determined the subunit specificity of inhibitor **3** by its ability to compete for proteasome subunit labeling with the fully characterized radiolabeled inhibitor Ada-$^{125}IYAhx_3Leu_3VS$ (**5**) (Figures 1 and 3D), that we have previously shown to react with all active subunits of the proteasome.[14,15] In a dose-dependent manner inhibitor **3** could compete with Ada-^{125}IY-Ahx_3Leu_3VS for labeling of all active protea-

some subunits in extracts from the multiple myeloma cell line MM.1S. We therefore conclude that compound **3** is a cell-permeant proteasome inhibitor that can be used to label all active proteasome subunits.

Probing proteasome activity in cell extracts

To directly visualize the proteins labeled by **3** in EL4 cell extracts, we took advantage of the probe's dansyl-sulfonamidohexanoyl tag. We incubated cell extracts with increasing concentrations of **3** followed by SDS-PAGE and immunoblotting for the dansyl-sulfonamidohexanoyl moiety (Figure 4A). We observed a clear labeling profile, devoid of spurious cross-reactivity attributable to the antibody used. At a concentration of 300 nM, inhibitor **3** labeled all active proteasome subunits to a similar extent, despite weaker reactivity of the probe towards the β2i subunit. In the presence of increasing concentrations of inhibitor, the proteasomal active sites were rapidly saturated, and we observed a nonspecific highmolecular-weight labeling profile at the higher concentrations. This underscores the need for a careful titration of directed activity-based probes. Notably, subunits migrate slightly differently on the gels (Figure 4A compared to Figure 3D); the migration of a particular subunit-probe complex on gel is probe-dependent, which explains the subtle differences. To validate the accuracy of the method, we performed proteasome activity measurements with fluorogenic substrates in parallel. We incubated extracts with increasing concentrations of inhibitor **3** and then monitored the cleavage of SucLeuLeuVal-TyrAMC (chymotryptic-like activity), ZAlaArgArgAMC (tryptic-like activity) and ZLeuLeuGluβNA (caspase-like activity) to determine the remaining activities (Figure 4B). The activity measurements with fluorogenic substrates corresponded well to the immunoblot profile. At a concentration of 10 μM of **3**, labeling of subunits with this inhibitor reached saturation, and activity against fluorogenic substrates was abolished almost completely, thus confirming the accuracy of the use of inhibitor **3** as a probe to measure proteasome activity *in vitro*.

Probing proteasome activity in living cells

We next probed proteasomal activity in intact cells. We incubated cultured EL4 cells with increasing concentrations of **3**, collected the cells, and immunoblotted for the dansyl group as described above (Figure 4C). We observed a dose-dependent labeling of active proteasome subunits from EL4 cells, which confirms that probe **3** permeates cell membranes and modifies all active proteasome subunits. When we compared the labeling pattern obtained in experiments with EL4 extracts to that obtained in experiments with cultured cells, we found that elevated concentrations of probe are needed to efficiently label active proteasome subunits, probably owing to kinetics of membrane permeation. Furthermore, the *in vivo* labeling profile differed in comparison to the labeling pattern in cell extracts. In EL4 cell lysates, the β1 and β1i subunits appear to be quite active. But *in vivo* the β1 subunit was labeled only in the presence of high concentrations of inhibitor **3**, whereas the β1i subunit was not labeled at all. In contrast, the β5 subunits seem to be more active and therefore the target for the majority of the observed *in vivo* active-site labeling. The exact reason for the observed activity difference is not clear. Possible explanations include partial dissociation of associated (19S) regulatory proteasomal cofactors upon preparation of extracts, as well as disruption of transient associations of active proteasomes with cell membranes in detergent extracts. In addition, when we compared *in vivo* labeling experiments to experiments in lysates, virtually no highmolecular-weight adducts were observed despite

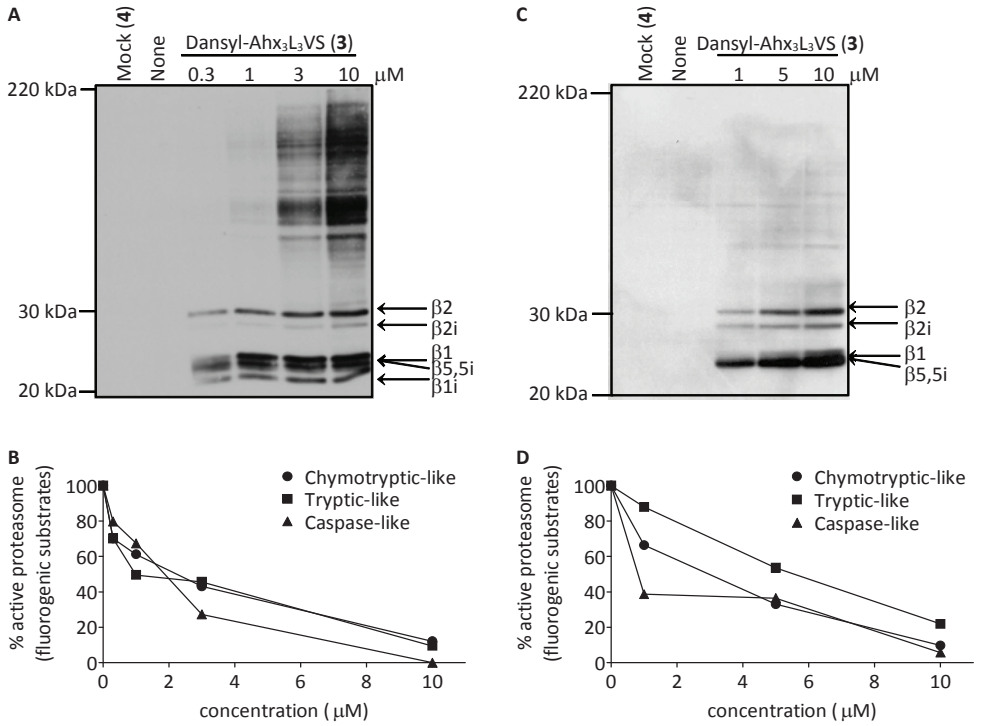


Figure 4 | Profiling of proteasome activity in EL4 cell extracts and cells. A,C) Immunoblots originating from the same gel showing results of a 1-h labeling experiment in EL4 cell extracts **(A)** or a 2-h labeling experiment in living EL4 cells **(C)** with the indicated concentrations of inhibitor **3** or 10 μM inactive control **4**. **B,D)** Measurement with fluorogenic substrates of residual proteasome activity in EL4 cell extracts **(B)** or living EL4 cells **(D)** after a 1-h **(B)** or 2-h **(D)** treatment with the indicated concentrations of inhibitor **3**. Results are representative of two independent experiments.

the presence of equal concentrations of inhibitor. Thus, inhibitor **3** labels proteasomes with greater selectivity *in vivo*. Next, we validated the *in vivo* activity pattern obtained with probe **3** by performing activity measurements with fluorogenic substrates in extracts prepared from identical samples (Figure 4D). *In vivo*, the saturation of subunit labeling corresponded to the inhibition of proteasome activity. At 10 μM, compound **3** almost completely inhibited proteasome activity, as was the case *in vitro*, confirming that it is unlikely that an altered bioavailability of compound **3** *in vivo* is responsible for the altered labeling pattern. Next, we used inhibitor **3** to profile the proteasome activity in HeLa cells *in vivo*. Again, the β5 subunits seemed to be most active and therefore the targets for the majority of the *in vivo* active-site labeling (Supplementary Figure 1). We conclude that proteasome activity profiles obtained by performing measurements in cell extracts are not necessarily representative of the *in vivo* activity patterns, stressing the need for live cell–based assays.

In vivo profiling of bortezomib and MG132

Next, we determined whether the probe could be used to accurately measure the remaining proteasome activity after inhibition by other proteasome inhibitors. We cultured EL4 cells

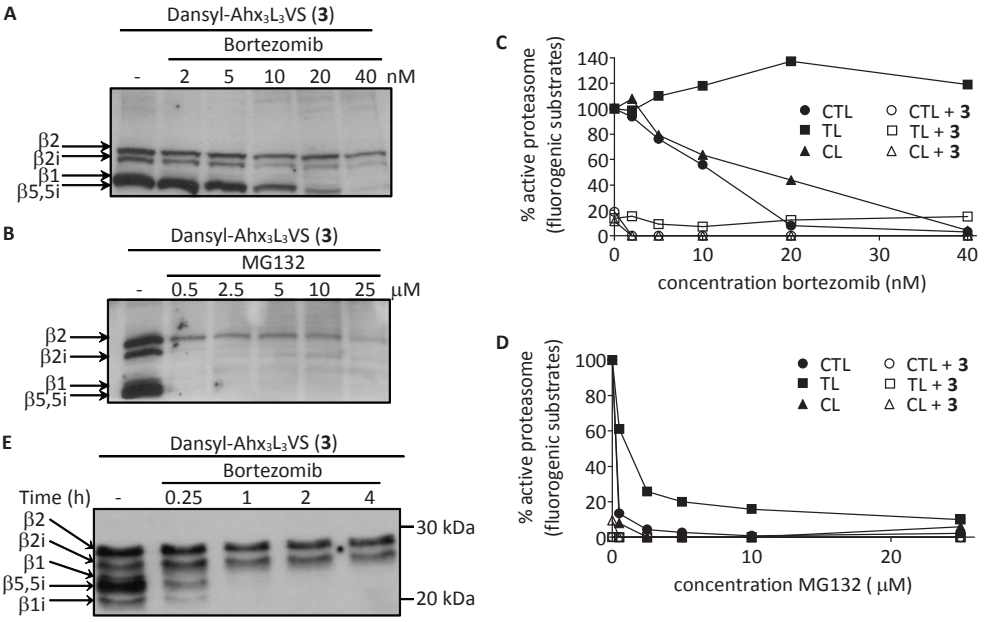


Figure 5 | *In vivo* profiling of proteasome inhibitors. A,B) Competition experiment between inhibitor **3** and bortezomib **(A)** or MG132 **(B)** in EL4 cells. **C,D)** Measurement of residual proteasome activity with fluorogenic substrates after treatment of EL4 cells with bortezomib or a combination of bortezomib and inhibitor **3 (C)** or MG132 or a combination of MG132 and inhibitor **3 (D)**. Results are representative of two independent experiments. CTL: chymotryptic-like, TL, tryptic-like; CL, caspase-like. **E)** Competition experiment between bortezomib and inhibitor **3** in MM.1S cells.

in the presence of bortezomib or MG132, then incubated with inhibitor **3** and immunoblotted as described above. We performed activity measurements with fluorogenic substrates in parallel before and after addition of inhibitor **3**. The addition of bortezomib resulted in a very specific reduction in the ability of compound **3** to modify the proteasomal β5, β5i and β1 subunits, with maximum inhibition of these subunits by bortezomib reached at 40 nM (Figure 5A). The β2 and β2i subunits remained fully available for modification with compound **3** at all time points, indicating the lack of a substantial interaction of bortezomib with these subunits *in vivo*. Bortezomib was previously reported to be a specific inhibitor for the chymotryptic-like activity largely provided by the β5 subunits. Here, we identified the β1 subunits as additional targets for bortezomib.

After incubation of EL4 cells with MG132, compound **3** only labeled the β2 subunit suggesting that all other subunits are inhibited by MG132 at high-nM concentrations (Figure 5B). Notably, MG132 completely inhibited the β2i subunit, which suggests a difference in the active-site chemistry between the β2 and β2i subunits. The results we obtained with fluorogenic substrates correlated in both cases to the immunoblot experiments using inhibitor **3** (Figure 5C,D). Inhibitor **3**, however, can be used to distinguish between constitutive and immunoproteasome subunit activity, whereas fluorogenic substrates cannot. Therefore, we did not observe the difference in labeling between the β2i subunit and the

β2 subunit using fluorogenic substrates. Combined, these results validate inhibitor **3** as a probe for measuring proteasome activity as well as for characterization of subunit specificity of proteasome inhibitors *in vivo*. Using inhibitor **3**, we next determined the composition of the active proteasome in the multiple myeloma cell line MM.1S, as well as the proteasomal targets of bortezomib in these cells. We cultured MM.1S cells in the presence or absence of 20 nM bortezomib, a concentration previously demonstrated to be toxic to this cell line.[17] In this cell line, we observed that all constitutive and immunoproteasome catalytic subunits were active in the absence of interferon γ (Figure 5E). The labeling pattern in MM.1S cells resembled the profile that was observed in EL4 cells. A notable difference was the activity of the β1i subunit in living MM1.S cells: we detected this subunit in extracts of EL4 cells, but not when we performed the assay in living cells. As quickly as in 15 min, 20 nM bortezomib substantially inhibited the activity of all β5 and β1 subunits with maximum inhibition reached within 2 h (Figure 5E).

DISCUSSION

Of the proteasome inhibitors described to date, each has a varying ability to inhibit each of the three known catalytic activities.[14,18,19] With proteasome inhibition becoming widely used in the treatment of multiple myeloma, the composition of proteasomes may be an important determinant of clinical response to treatment with such agents. Most methodologies in common practice for measuring proteasome activity use fluorogenic peptidyl substrates and often involve procedures to reduce potentially cross-reacting protease activities.[5,20,21] More importantly, procedures that rely on the hydrolysis of fluorogenic substrates do not distinguish between constitutive and immunoproteasome activity and therefore do not reflect the relative activity of different proteasome subunits accurately.

Recently, cell-based assays[16,22] as well as mouse models[23,24] have been developed for *in vivo* monitoring of proteasome activity. These methods may be useful in screening for new inhibitors and testing bioavailability, respectively, but their activity readout depends on the balance between synthesis and degradation of fusion proteins, which involves many cellular factors other than the proteasome, including the rate of fusion-protein synthesis. Therefore these methods may not accurately reflect true proteasomal activity. Moreover, the use of expressed fusion proteins makes these methods unfit for monitoring human samples. In addition, they do not provide insight in the constitutive/immunoproteasome activity ratio and cannot be used to characterize the specific activity of proteasome inhibitors *in vivo*. The small molecule–based assay described here overcomes these problems.

We have devised proteasome activity–profiling probe **3** that freely reaches cellular targets in intact cells and that can modify covalently and irreversibly all of the proteasome's exposed catalytic subunits. Use of antibodies directed against the dansylsulfonamidohexanoyl hapten present in **3** allows direct measurement of activity of all proteasome subunits in living cells without the need for prior lysis or purification steps. This ensures not only a fast but also accurate activity readout, as under these conditions proteasomes remain in their native state. Using this probe, we characterized the anti-neoplastic agent bortezomib and revealed that the β1 and β1i subunits are targets for this drug, in addition to the previously identified β5 subunits.[20] We also characterized the widely used proteasome inhibitor MG132 and found that it inhibited all subunits, although the β2 subunit was inhibited to a lesser extent. Notably,

MG132 was found to fully inhibit the immunoproteasome subunit β2i.

We have described an easy method for characterizing the specificity of (new) proteasome inhibitors that may be helpful in a clinical setting and could be used to study resistance to this new class of drugs. Differences in labeling profiles observed in experiments with extracts and living cells, which can not be detected using methodology now used to measure proteasome activity, underscores the need for reliable *in vivo* labeling techniques. The technique described here could be adapted for use with other proteases and optimized. Switching to a more chemoselective reactive group could improve both probe reactivity and selectivity, and improving the fluorescence yield may allow for *in vivo* localization studies of active proteasomes and a truly high-throughput, in-gel readout.

MATERIALS AND METHODS

Materials

Unless indicated otherwise, chemicals were obtained commercially of the highest available grade. Cell culture. The cell lines EL4 (mouse, thymoma) and MM.1S (human, multiple myeloma) were cultured in RPMI 1640 medium. HeLa cells (human, cervix epitheloid carcinoma) and U373 cells (human, astrocytoma) were cultured in DMEM. Both media were supplemented with 10% fetal calf serum (FCS), 100 units/ml penicillin and 100 μg/ml streptomycin.

Synthesis and characterization of proteasome probes

Detailed information on synthetic procedures and characterization can be found in the Supplementary Methods and Supplementary Figures 2–4 online. Compounds **3** and **4** were tested and compared to other inhibitors as described previously.[16] U373 cells expressing an ubiquitin-Arg-GFP fusion protein were exposed to various concentrations of compounds **3**, **4**, Ada-Ahx$_3$Leu$_3$VS and various other proteasome inhibitors. After 5 h, we quantitated the accumulation of GFP by FACS. The specificity for proteasomal targets was determined by assaying competition for proteasome inhibition with compound **5**. We incubated equal amounts of protein, obtained by lysis of EL4 cells by glass beads (see below), with the indicated amounts (Figure 3D) of inhibitor **3** or DMSO for 1 h at 37 °C, and then incubated with compound **5** for 2 h at 37 °C. Proteins were denatured by boiling in reducing sample buffer, separated by 12.5% SDS-PAGE and autoradiographed.

Probing proteasome activity in cell extracts

To obtain lysates for proteasome activity assays, we lysed cells with glass beads as previously described.[18] In brief, cells were washed twice with cold phosphate-buffered saline, pelleted and lysed with one volume of glass beads (<106 microns, acid-washed; Sigma) and an equal volume of homogenization buffer (50 mM Tris-HCl pH 7.4, 1 mM dithiothreitol, 5 mM $MgCl_2$, 2 mM ATP and 250 mM sucrose) by vortexing at high speed for 15–30 min at 4 °C. Beads, membrane fractions, nuclei and cell debris were then removed from the supernatant by centrifugation at 16,000g for 5 min. We quantitated the protein content of extracts using the Bradford assay (BioRad) and then incubated equal amounts of protein with the indicated amounts (Figure 4A) of probe **3**, 10 μM control probe **4** or buffer for 1 h at 37 °C. Proteasome activity was assayed as described below.

Probing proteasome activity in living cells

We incubated 10 x 10^6 EL4 cells in 10 mL RPMI 1640 medium supplemented with 10% FCS and penicillin and streptomycin with the indicated amounts (Figure 4C) of probe **3** or 10 μM probe **4**, at 37 °C for 2 h. HeLa cells were incubated with 10 μM probe **3** for the indicated time periods (Supplementary Figure 1). Cells were collected by centrifugation. We lysed cells using glass beads and assayed protein content using the Bradford assay. Proteasome activity was assayed as described below.

In vivo profiling of bortezomib and MG132

We incubated 10 x 10^6 EL4 cells in 10 mL RPMI 1640 medium supplemented with 10% FCS and penicillin and streptomycin with the indicated amounts of bortezomib (Figure 5A) or MG132 (Figure 5B) for 1 h at 37 °C. When indicated, this was followed by a 2-h incubation at 37 °C with 10

μM probe **3**. We incubated 5 x 10^6 MM.1S cells in 5 mL RPMI 1640 medium supplemented with 10% FCS and penicillin and streptomycin with 20 nM bortezomib for the indicated time periods at 37 °C, followed by a 1-h incubation with 10 μM probe **3** at 37 °C. We lysed cells using glass beads or by incubating them for 30 min in NP-40 lysis buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 1% NP-40])and centrifuged the lysates for 5 min to remove membrane fractions, nuclei and cell debris. Proteasome activity was assayed as described below.

Assaying proteasome activity by immunoblotting

Equal amounts (typically 60 μg) of protein were denatured by boiling in reducing sample buffer, separated by 12.5% SDS-PAGE and electrotransferred onto polyvinylidene difluoride (PVDF) membranes. Immunoblotting was performed using a dansyl-sulfonamidohexanoyl polyclonal antibody (1:7,500, rabbit; Molecular Probes) and horseradish peroxidase–coupled goat or swine anti-rabbit secondary antibody (Southern Biotech) followed by enhanced chemiluminescence (Western Lightning, Perkin-Elmer).

Assaying proteasome activity with fluorogenic substrates

Assays were performed as previously described[25] with minor changes. 10 μg of lysate was added to 100 μL of substrate buffer, containing 20 mM HEPES (pH 8.2), 0.5 mM EDTA, 1% DMSO, 1 mM ATP and 30 μM ZAlaArgArg-AMC (tryptic-like), 60 μM ZGlyGlyLeuAMC (chymotryptic-like), 60 μM SucLeuLeuValTyr-AMC (chymotryptic-like) or 60 μM ZLeuLeuGlu-βNA (caspase-like). SDS was omitted from the substrate buffer. Fluorescence was measured every minute for 25 min at 37 °C using a Fluostar Optima 96-well plate reader (BMG lab technologies) (λ ex/em = 355/450 nm for AMC and 320/405 nm for βNA) and the increase in fluorescence per minute was used to calculate specific activities of each sample. Cells were incubated with 1 μM epoxomicin (Affiniti) at 37 °C for 1 h to determine the nonspecific activity, which was subtracted from each measurement.

ACKNOWLEDGMENTS

The authors thank M.A. Leeuwenburgh for providing Ada-Ahx_3Leu_3VS. This work was financially supported by the National Institutes of Health (H.L.P. and P.J.G.), a grant from AstraZeneca (H.L.P. and P.J.G.), the Netherlands Organization for Scientific Research (H.O.), a National Cancer Institute SPORE grant Career Developmental Award (H.O.), the Dutch Cancer Society (H.O.), a Multiple Myeloma Research Foundation Senior Research Award (B.M.K.), the Dr. Saal van Zwanenberg Stichting (M.V.), the Austrian Academy of Sciences (E.F.), the Koninklijke Hollandsche Maatschappij der Wetenschappen (C.R.B.), and by the Stichting Fonds Doctor Catharine van Tussenbroek (C.R.B.).

REFERENCES

1. Hideshima, T. et al. The proteasome inhibitor PS-341 inhibits growth, induces apoptosis, and overcomes drug resistance in human multiple myeloma cells. *Cancer Res.* **61**, 3071–3076 (2001).
2. Rock, K.L. & Goldberg, A.L. Degradation of cell proteins and the generation of MHC class I–presented peptides. *Annu. Rev. Immunol.* **17**, 739–797 (1999).
3. Glickman, M.H. & Ciechanover, A. The ubiquitin-proteasome proteolytic pathway: destruction for the sake of construction. *Physiol. Rev.* **82**, 373–428 (2002).
4. Kisselev, A.F. & Goldberg, A.L. Proteasome inhibitors: from research tools to drug candidates. *Chem. Biol.* **8**, 739–758 (2001).
5. Adams, J. et al. Proteasome inhibitors: a novel class of potent and effective antitumor agents. *Cancer Res.* **59**, 2615–2622 (1999).
6. Dahlmann, B., Ruppert, T., Kloetzel, P.M. & Kuehn, L. Subtypes of 20S proteasomes from skeletal muscle. *Biochimie* **83**, 295–299 (2001).
7. Driscoll, J., Brown, M.G., Finley, D. & Monaco, J.J. MHC-linked LMP gene products specifically alter peptidase activities of the proteasome. *Nature* **365**, 262–264 (1993).
8. Gaczynska, M., Rock, K.L. & Goldberg, A.L. Gamma-interferon and expression of MHC genes regulate peptide hydrolysis by proteasomes. *Nature* **365**, 264–267 (1993).
9. Bogyo, M. et al. 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. USA* **94**, 6629–6634 (1997).

10. Greenbaum, D., Medzihradszky, K.F., Burlingame, A. & Bogyo, M. Epoxide electrophiles as activity-dependent cysteine protease profiling and discovery tools. *Chem. Biol.* **7**, 569–581 (2000).
11. Liu, Y., Patricelli, M.P. & Cravatt, B.F. Activity-based protein profiling: the serine hydrolases. *Proc. Natl Acad. Sci. USA* **96**, 14694–14699 (1999).
12. Adam, G.C., Cravatt, B.F. & Sorensen, E.J. Profiling the specific reactivity of the proteome with non-directed activity-based probes. *Chem. Biol.* **8**, 81–95 (2001).
13. Adam, G.C., Sorensen, E.J. & Cravatt, B.F. Proteomic profiling of mechanistically distinct enzyme classes using a common chemotype. *Nat. Biotechnol.* **20**, 805–809 (2002).
14. Kessler, B.M. et al. Extended peptide-based inhibitors efficiently target the proteasome and reveal overlapping specificities of the catalytic beta-subunits. *Chem. Biol.* **8**, 913–929 (2001).
15. Ovaa, H. et al. Chemistry in living cells: detection of active proteasomes by a twostep labeling strategy. *Angew. Chem. Int. Edn. Engl.* **42**, 3626–3629 (2003).
16. Dantuma, N.P., Lindsten, K., Glas, R., Jellne, M. & Masucci, M.G. Short-lived green fluorescent proteins for quantifying ubiquitin/proteasome-dependent proteolysis in living cells. *Nat. Biotechnol.* **18**, 538–543 (2000).
17. Hideshima, T. et al. The proteasome inhibitor PS-341 inhibits growth, induces apoptosis, and overcomes drug resistance in human multiple myeloma cells. *Cancer Res.* **61**, 3071–3076 (2001).
18. Bogyo, M., Shin, S., McMaster, J.S. & Ploegh, H.L. Substrate binding and sequence preference of the proteasome revealed by active-site-directed affinity probes. *Chem. Biol.* **5**, 307–320 (1998).
19. Garcia-Echeverria, C. Recent advances in the identification and development of 20S proteasome inhibitors. *Mini Rev. Med. Chem.* **2**, 247–259 (2002).
20. Lightcap, E.S. et al. Proteasome inhibition measurements: clinical application. *Clin. Chem.* **46**, 673–683 (2000).
21. Stein, R.L., Melandri, F. & Dick, L. Kinetic characterization of the chymotryptic activity of the 20S proteasome. *Biochemistry* **35**, 3899–3908 (1996).
22. Fiebiger, E., Story, C., Ploegh, H.L. & Tortorella, D. Visualization of the ER-to-cytosol dislocation reaction of a type I membrane protein. *EMBO J.* **21**, 1041–1053 (2002).
23. Lindsten, K., Menendez-Benito, V., Masucci, M.G. & Dantuma, N.P. A transgenic mouse model of the ubiquitin/proteasome system. *Nat. Biotechnol.* **21**, 897–902 (2003).
24. Luker, G.D., Pica, C.M., Song, J., Luker, K.E. & Piwnica-Worms, D. Imaging 26S proteasome activity and inhibition in living mice. *Nat. Med.* **9**, 969–973 (2003).
25. Elliott, P.J., Soucy, T.A., Pien, C.S., Adams, J. & Lightcap, E.S. Assays for proteasome inhibition. *Methods Mol. Med.* **85**, 163–172 (2003).

Supplementary information

SUPPLEMENTARY METHODS

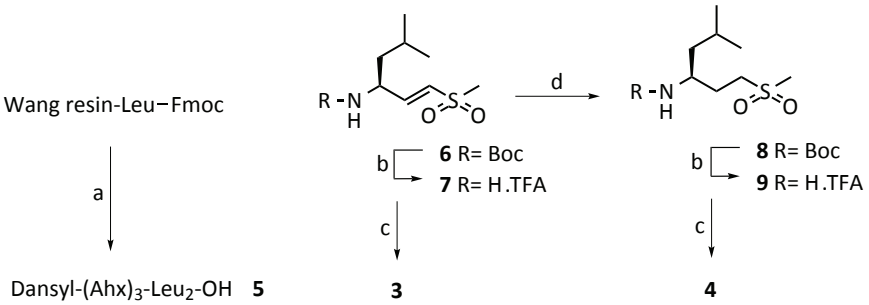
Inhibitors: Bortezomib was synthesized as previously described[1]. Dansyl-Ahx$_3$L$_2$-OH (**5**) was synthesized (Supplementary Scheme 1) on Wang resin using standard Fmoc-based solid-phase peptide synthesis protocols and purified by silica gel chromatography using a gradient of 10% to 20% of MeOH in CH_2Cl_2. **5** was coupled by *N*-(3-dimethylaminopropyl)-*N*'-ethylcarbodiimide hydrochloride (EDCI)-mediated condensation with (S,*E*)-5-methyl-1-(methylsulfonyl)hex-1-en-3-amine (leucine vinyl sulfone, LeuVS) **7** to afford **3**. MS(ESI): calculated 990.5 $(M+H)^+$ 1012.5 $(M+Na)^+$. Observed 990.5 $(M+H)^+$ 1012.5 $(M+Na)^+$. Hydrogenation and deprotection of **6** followed by condensation of the resulting (S)-5-methyl-1-(methylsulfonyl)hexan-3-amine **9** with **5** afforded the negative control **4**. MS(ESI): calculated 992.5 $(M+H)^+$ 1014.5 $(M+Na)^+$. Observed 992.5 $(M+H)^+$ 1014.5 $(M+Na)^+$.

SUPPLEMENTARY INFORMATION ONLINE

Detailed information on the characterization of compound **3** can be found in Supplementary Figures 2–4. This material is available via the Internet at http://www.nature.com/nmeth, as supplementary information of *Nature Methods* **2**, 357-362 (2005).

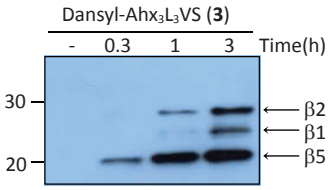
SUPPLEMENTARY REFERENCES

1. Adams, J. et al. Potent and selective inhibitors of the proteasome: dipeptidyl boronic acids. *Bioorg. Med. Chem. Lett.* **8**, 333-338 (1998).



Supplementary Scheme 1 | a) Fmoc based SPPS then TFA, (78%). b) TFA quant. c) EDCI, DIPEA, CH_2Cl_2/DMF (**3**: 64%, **4**: 84%). d) 10% Pd/C, H_2, MeOH (98%).

Supplementary figures



Supplementary Figure 1 | Incubation of HeLa cells with inhibitor 3.

Chapter 1.3

Profiling proteasome activity in tissue with fluorescent probes

Molecular Pharmaceutics
Volume 4, 2007, Pages 739-748

Profiling proteasome activity in tissue with fluorescent probes

Celia R. Berkers,[a,§] Fijs W. B. van Leeuwen,[a,§] Tom A. Groothuis,[b] Victor Peperzak,[c] Erica W. van Tilburg,[a] Jannie Borst,[c] Jacques J. Neefjes[b] & Huib Ovaa[a,*]

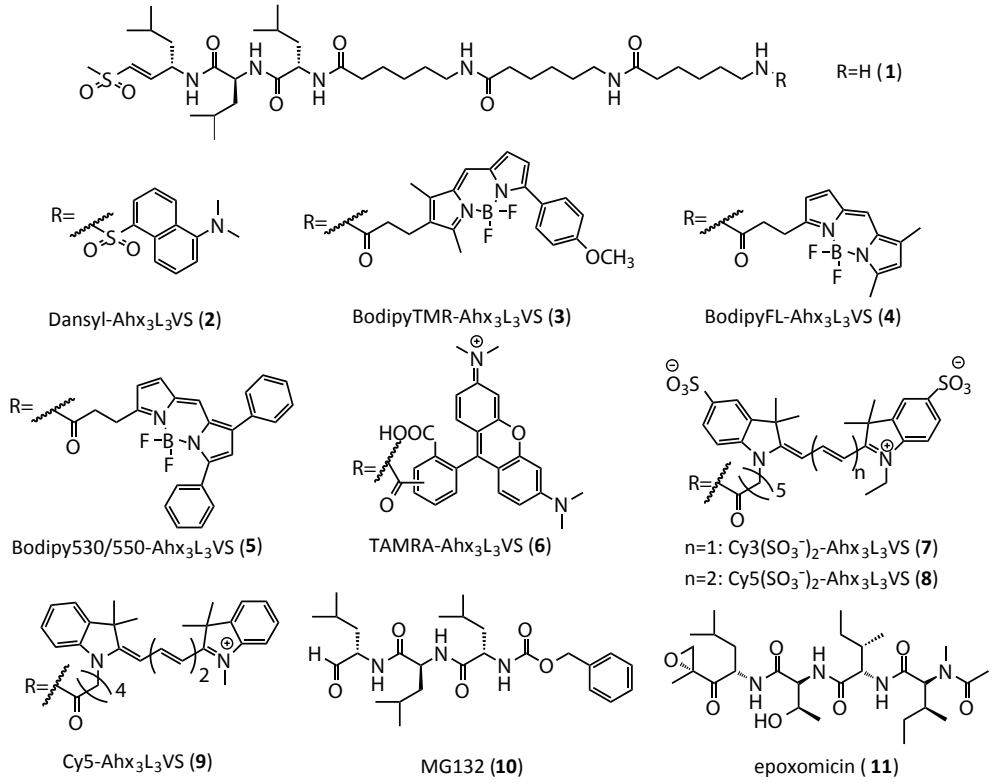
Divisions of [a]Cellular Biochemistry, [b]Tumor Biology, and [c]Immunology, The Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands.

With proteasome inhibitors in use in the clinic for the treatment of multiple myeloma and with clinical trials in progress investigating the treatment of a variety of hematologic and solid malignancies, accurate methods that allow profiling of proteasome inhibitor specificity and efficacy in patients are in demand. Here, we describe the development, full biochemical validation, and comparison of fluorescent proteasome activity reporters that can be used to profile proteasome activities in living cells with high sensitivity. Seven of the synthesized probes tested label proteasomes in lysates, although the fluorescent dye used affects their specificity. Two differentially labeled probes tested are suitable for studying proteasome activity in living cells by gel-based assays, by confocal laser scanning microscopy, and by flow cytometry. We established methods using these fluorescent reporters to profile proteasome activity in different mouse tissues, carefully avoiding postlysis artifacts, and we show that proteasome subunit activity is regulated in an organ-specific manner. The techniques described here could be used to study pharmacological properties of proteasome inhibitors.

INTRODUCTION

The clinical value of proteasome inhibition has recently become evident. The proteasome inhibitor bortezomib[1] is currently used to treat multiple myeloma, while a therapeutic effect on several other (solid) tumors has also been shown.[2] The proteasome is responsible for the degradation of misfolded and redundant proteins[3] as well as a variety of key proteins involved in the regulation of cell proliferation and survival.[4] Proteasome inhibition disrupts these pathways, leading to the suppression of NF-κB activity and stabilization of the tumor suppressor p53,[5] among others. Both suppression of NF-κB activity and p53 stability have been suggested to be at least in part responsible for the observed apoptotic effects.[2,5] Since side effects of bortezomib are substantial and since a response is observed only in a subset of patients,[2] there is an emerging demand for methods for accurately assessing the inhibitory action and efficacy of proteasome inhibitors in clinical subjects. Such a method may allow treatment to be adjusted to individual cases, minimizing side effects and strengthening treatment responses. Next to diagnostic applications,

[§]These authors contributed equally to this work. [*]Correspondence should be addressed to H.O. (h.ovaa@nki.nl).



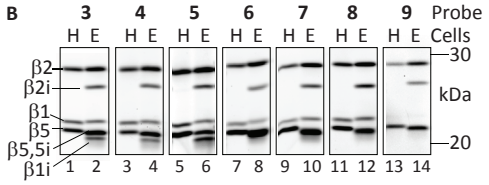


Figure 1 | Probes provide an in-gel readout of proteasome subunit availability in cell lysates. A) All probes are based on structure **1**, which contains a reactive vinylsulfone part (VS) coupled to three leucine residues (L_3) and a spacer consisting of three aminohexanoic acid moieties (Ahx_3). A dansyl moiety or different fluorophores can be coupled to this inhibitory segment, yielding the dansyl probe **2** and benchmark **3** described previously[14,19] as well as probes **4–9** used here. **B)** Representative gel images showing results of a labeling experiment with 500 nM probes **3–9** (1 h) in HeLa and EL4 cell lysates. Proteins were separated using a 12.5%, 20 cm SDS–PAGE gel, affording separation of most subunits in a one-dimensional SDS–PAGE format. Filter settings, exposure time, and image contrast settings were optimized for each individual probe. H) HeLa lysate; E) EL-4 lysate, and i) immunoproteasome subunit.

proteasome probes are also likely to advance basic research in proteasome action since this line of research heavily depends on the sensitivity of tools available.

The catalytic activity of the proteasome resides within the 20S core and is provided by three constitutive subunits, termed β1, β2, and β5, and/or their inducible immunoproteasome counterparts β1i, β2i, and β5i. It is widely accepted that immunoproteasome subunits are mainly expressed in lymphoid tissues,[6] e.g., spleen, lymph nodes, and thymus. In addition, it is rapidly becoming apparent that different organs express different ratios

of constitutive, immuno, and hybrid proteasomes.[7,8] These proteasome subtypes differ with regard to their activity against (fluorogenic) peptide and protein substrates.[7,8] Consequently, proteasome activity and specificity could be organ-specific.[6,7] This is, however, not experimentally validated and requires sensitive techniques to enable the detection of active subunit compositions.

Three different methods are currently used to monitor proteasome activity: fluorogenic assays,[9,10] small molecule–based activity assays,[11–14] and models based on recombinant reporter proteins.[15–18] Fluorogenic and most small molecule assays cannot be used to probe living cells or tissue, whereas the use of reporter proteins remains confined to genetically altered cells or organisms and thus does not allow profiling of patient material. Recently, we reported the development of the dansylated vinylsulfone–based proteasome inhibitor **2** (Figure 1), which contains an α,β-unsaturated sulfone part that reacts with the γ-hydroxyl of the N-terminal threonine residue of catalytic subunits of the proteasome,[19] while the use of antibodies against the dansyl moiety allows the detection of labeled subunits by SDS-PAGE and Western blot analysis. This probe is the first of its kind to profile the specificity of proteasome inhibitors in living cells.[19,20]

Since probe **2** requires Western blotting to produce an activity–based readout, we reasoned that this probe could be further optimized by replacing the environment sensitive and low-quantum yield dansyl fluorophore with high-quantum yield fluorophores. This allows direct scanning of the gel for fluorescence emission for visualization of labeled subunits. We recently showed[14] in an initial report that this concept affords a more reliable and sensitive readout in cell lysates, living cells, and tissue using inhibitor **3**. The use of fluorescent probes allows profiling of large numbers of samples, and the method should be applicable to patient material, allowing tumor profiling, which is likely to improve responses to proteasome inhibitor therapy. For these reasons, we embarked on the synthesis of a range of fluorescent inhibitors so we would be able to select for effective fluorescent proteasome inhibitors with good biochemical and biophysical properties. We furthermore describe the full validation of all probes described; we report the development of procedures that distinguish live cell labeling events from *in vitro* postlysis artifacts, and we have extended the live cell labeling procedures into flow cytometry-based assays. The latter approach could alter diagnostic means for (overall) proteasome activities and susceptibility to inhibition since it allows for significant throughputs using minimal amounts of material.

RESULTS AND DISCUSSION

Probe preparation and characterization in cell lysates

All probes were based on an H_2N-Ahx$_3$Leu$_3$-VS inhibitory segment (**1**) (Figure 1A) and functionalized with different types of fluorophores. Since different fluorophores possess different biophysical properties, such as membrane permeability, a library of fluorescent proteasome probes was synthesized (see the Supplementary Information and Figure 1 therein for experimental procedures). The following probes were selected for further investigation: BodipyFL-Ahx$_3$Leu$_3$VS (**4**), Bodipy530/550-Ahx$_3$Leu$_3$VS (**5**), TAMRA-Ahx$_3$Leu$_3$VS (**6**), Cy3(SO_3^-)$_2$-Ahx$_3$Leu$_3$VS (**7**), Cy5(SO_3^-)$_2$-Ahx$_3$Leu$_3$VS (**8**), and Cy5-Ahx$_3$-Leu$_3$VS (**9**). They were compared to the control probe BodipyTMR-Ahx$_3$Leu$_3$VS (**3**) (Figure 1A) about which we recently communicated our findings.[14] These probes span a large emission spectrum, ranging from green (**4**,

λ em ≈ 510 nm) to orange/red (**3** and **5–7**, λ em ≈ 590 nm) and far red (**8** and **9**, λ em ≈ 670 nm), which makes it possible to combine them with other fluorescent biomolecules (such as GFP and RFP) and with each other in pulse–chase experiments.

The ability of these probes to provide an in-gel readout of proteasome subunit availability was investigated (Figure 1B). Lysates from both HeLa cells (human, cervix carcinoma) and EL4 cells (murine thymoma) were incubated with probes **3–9**. Proteins were separated by SDS-PAGE, and the wet gel slab was directly scanned using a fluorescence scanner with appropriate filter settings. As shown in Figure 1B, probes **3–9** labeled a distinct set of proteins in both HeLa and EL4 cell lysates. These proteins could be specifically immunoprecipitated using anti-proteasome antibodies (data not shown) and identified as the indicated proteasome active subunits on the basis of the labeling patterns found with probe **2**, the identities of which we recently confirmed by mass spectrometry.[21] Nonspecific labeling in high-molecular weight regions was not observed (Supplementary Figure 2), suggesting that these probes are specific for proteasomes in cell lysates at the concentrations used.

The different fluorescent moieties were found to affect the fine specificities of individual probes. Both benchmark **3** and BodipyFL-labeled probe **4** afforded quite uniform labeling patterns (Figure 1B, lanes 1–4), suggesting that these probes have similar affinities for all catalytically active subunits. The other probes exhibited enhanced reactivity toward particular subunits (Figure 1B, lanes 5–14). In HeLa cells, probes **5–9** labeled the β5 subunit more efficiently than the β1 and β2 subunits. In addition, probe **6** and in particular probe **9** poorly labeled β1 subunits. Furthermore, Figure 1B shows that most subunits labeled with probes **3–5** can be resolved by one-dimensional SDS-PAGE. Subunits labeled with probes **6–9** are more difficult to separate with good resolution. This suggests that Bodipy-labeled probes allow easier interpretation of complex labeling patterns. Altogether, results indicate that all probes give an in-gel readout that correlates with proteasome activity. Nonetheless, probes **3** and **4** are preferred for in-gel readouts.

Proteasome labeling in living cells

To study which probes can be used to probe living cells, flow cytometry experiments were performed. MelJuSo cells (human melanoma) were incubated with individual probes, and fluorescence intensity was measured. Supplementary Figure 3 shows that benchmark **3** as well as probes **4–6** and **9** accumulated in living cells. Probes **7** and **8** on the other hand were not taken up by these cells and were therefore not subjected to further studies, but may be used in experiments on cell lysates.

Proteasome inhibition by **3–6** and **9** in living cells was studied, using MelJuSo cells stably expressing ubiquitin-R-GFP (Ub-R-GFP).[15] Following the N-end rule, Ub-R-GFP is rapidly degraded by the proteasome under steady state conditions. Upon inhibition of the proteasome, the GFP concentration in the cells will increase over time, which can be quantified by flow cytometry. Cells were incubated with probe or the general proteasome inhibitor MG132 (ZL_3CHO, **10**). The percentage of cells responding to MG132 was taken to be 100%, and GFP induction in other samples was normalized accordingly. Since the fluorescence emitted by **4** is comparable in wavelength to GFP fluorescence (λ ex/em = 480/530 nm), nontransfected MelJuSo cells incubated with this probe were used to normalize the signal (Supplementary Figure 4). Figure 2A shows that cellular accumulation did not necessarily result in proteasome binding and inhibition. Only probes **3** and **4** induced substantial GFP

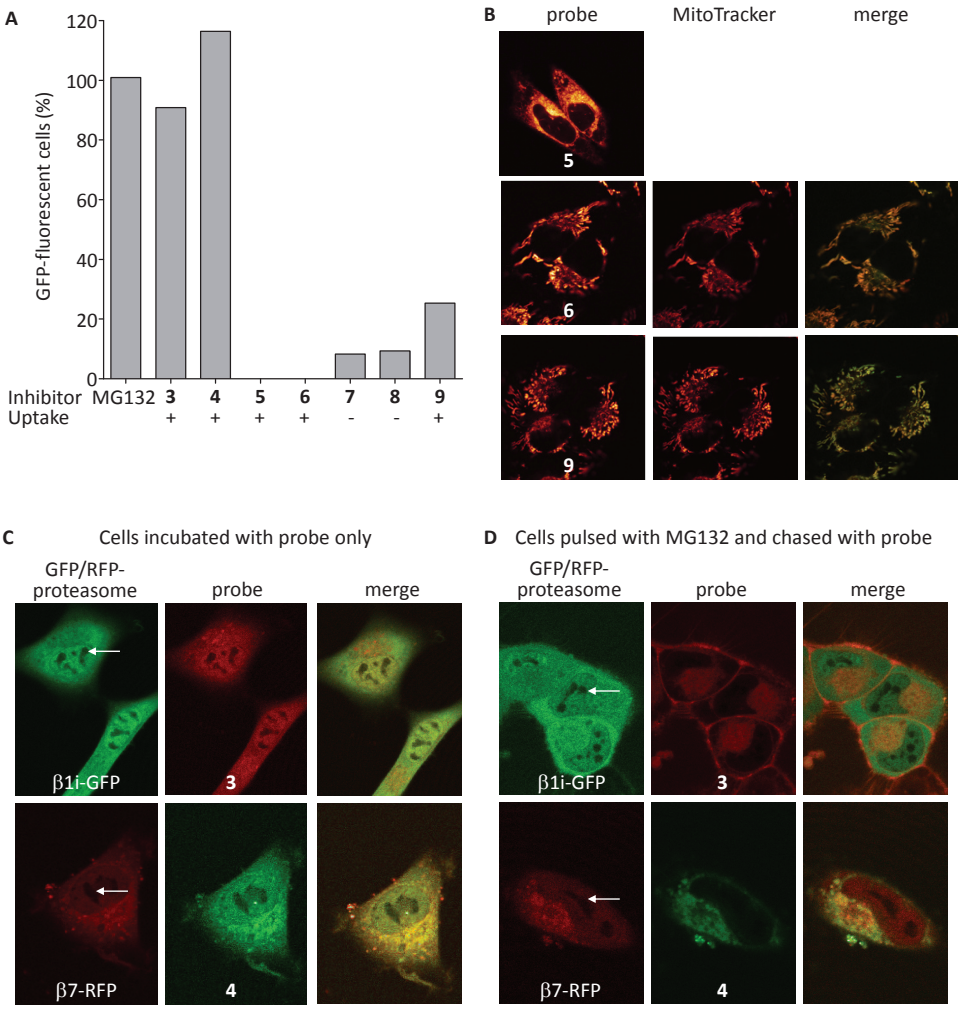


Figure 2 | Probes 3 and 4 can be used to study the distribution and activity of proteasomes in living cells. A) Flow cytometry analysis results of Ub-R-GFP-expressing MelJuSo cells upon incubation with probes **3–9** (500 nM) or MG132 (5 μM). The percentage of cells responding to MG132 was taken to be 100%. **B)** CLSM images showing MelJuSo cells, incubated with probe **5**, **6**, or **9** (500 nM, 1 h). Cells incubated with **6** or **9** were coincubated with MitoTracker (200 nM, 30 min). From left to right, distribution of probe, localization of MitoTracker, overlay. **C)** CLSM images showing MelJuSo cells expressing β7-RFP or β1i-GFP incubated with probe **3** or **4** (1 h). From left to right: localization of GFP- and RFP-labeled proteasome subunits, localization of probe, overlay showing colocalization of probe and proteasome. **D)** CLSM images showing MelJuSo cells expressing β7-RFP or β1i-GFP pulsed with MG132 (1 h) and chased with probe **3** or **4** (1 h). From left to right: localization of GFP- and RFP-labeled proteasome subunits after MG132 treatment, localization of probe after MG132 treatment, overlay. Arrows indicate unstained nucleoli.

fluorescence, confirming that only these two compounds efficiently inhibit proteasomes in cultured cells.

Confocal laser scanning microscopy (CLSM) allows us to study probe localization in living cells, as demonstrated previously.[14] To validate that probes **5**, **6**, and **9** accumulate in the cells but do not bind the proteasome, MelJuSo cells were incubated and imaged. Figure 2B shows that probe **5** was mainly visible in granular structures near the nucleus. Probes **6** and **9**, both lipophilic and cationic probes (Figure 1A), were found to accumulate in mitochondria, as visualized by colocalization with MitoTracker green (Figure 2B), and remained there for extended periods of time. Lipophilic cations are often found to accumulate in mitochondria due to the mitochondrial membrane potential.[22] Upon fixation with formaldehyde, **6** and **9** gave a distribution similar to that observed for GFP-labeled proteasomes,[23] indicating that these probes are not covalently bound to mitochondria but merely accumulate there (Supplementary Figure 5).

To study colocalization of proteasomes with benchmark **3** and probe **4**, both of which inhibited proteasomes in living cells, MelJuSo cells that stably expressed either β1i-GFP[23] or β7-red fluorescent protein (RFP) were incubated with individual probes, washed to remove excess probe, and imaged by CLSM. As shown in Figure 2C, GFP- and RFP-labeled proteasomes were distributed uniformly throughout the nucleus and cytoplasm, with the exception of nucleoli, the nuclear envelope, and the ER/Golgi, consistent with previous reports.[14,23] Compound **3** afforded a similar distribution, as reported,[14] as did compound **4**, and the overlays of proteasome and probe signals revealed substantial colocalization of both probes with GFP- and RFP-labeled proteasomes. Non-colocalizing, free probe could be observed in vesicular structures (Figure 2C), which are likely to be endosomes, since lipophilicity and high molecular weight are factors known to promote uptake via the endosomal pathway.[24] To verify the actual labeling of proteasomes in this experiment, cells were treated as described and lysed and proteins resolved by SDS-PAGE (Supplementary Figure 6). A clean proteasome labeling pattern was obtained, confirming that at least part of the colocalizing probe is proteasome-bound.

To further study whether all colocalizing probe is proteasome-bound, cells were pulsed with MG132, washed, and chased with either benchmark **3** or probe **4**. Figure 2D (merge) reveals that under these conditions, probes no longer colocalized with GFP- and RFP-labeled proteasomes. This effect could be visualized best in nuclei, which can be recognized from their unstained nucleoli. Instead, probe that was taken up by the cells now remained confined to the plasma membrane and the perinuclear area as free probe. MG132 treatment is reported to block all catalytic sites of the proteasome,[25] preventing probe binding. Since inhibition with MG132 can only alter the intracellular distribution of proteasome-bound probe, but not of free probe, it provides a reliable method for controlling for nonspecific interactions of probes **3** and **4**. Supplementary Figure 6 shows that MG132 treatment resulted in 85–94% block of proteasome labeling. Labeling of additional proteases was not observed, suggesting that only free probe was left in the cells after MG132 treatment. These results suggest that the signal observed in cytosol and nucleus in Figure 2C can fully be attributed to proteasome-bound probe and that probes **3** and **4** can be used to study proteasome distribution and profile the activity of proteasomes in living cells.

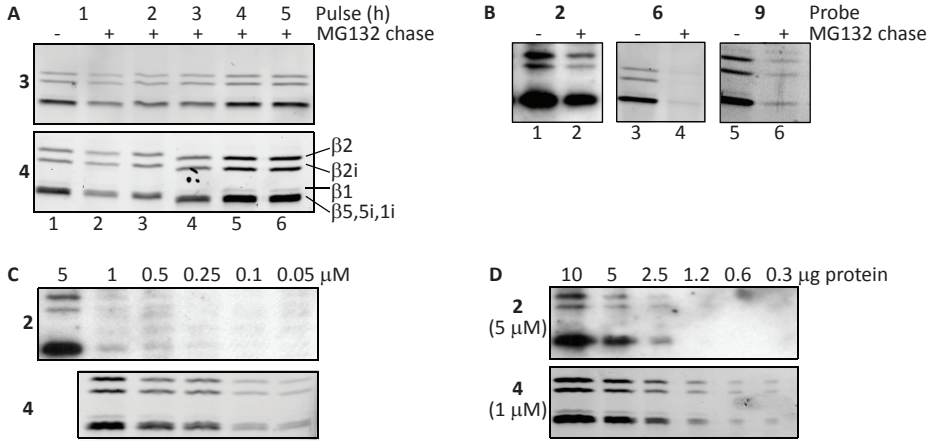


Figure 3 | Probes 3 and 4 function as highly sensitive proteasome subunit activity reporters in living cells. A,**B)** Gel images and immunoblot showing proteasome labeling in living EL4 cells that were pulsed with probe **3** or **4** (375 nM, indicated times) **(A)** or with probes **6** and **9** (375 nM, 1 h) or **2** (5 μM, 2 h) **(B)** and either harvested directly or chased with 5 μM MG132 (1 h). **C**,**D)** Gel image and immunoblot showing proteasome labeling in living EL4 cells, incubated with probe **2** or **4** (2 h), with either decreasing probe concentrations **(C)** or decreasing levels of protein loading on gel **(D)**. Proteins were separated using mini-gels (that do not separate the β1i subunit from the β5 subunits).

Optimization of in-gel activity assay in living cells

Next, we set out to optimize the gel assay described above. In CLSM experiments, some residual free probe **3** and **4** was observed in membrane structures (Figure 2C). Since cells have to be lysed for a gel-based readout, internal membrane structures are destroyed, which may release unbound probe. This potentially results in additional labeling, disturbing the accuracy of readouts. To prevent postlysis labeling, cells were pulsed with either probe **3** or **4** for various periods of time and chased with MG132 to block residual catalytic sites (Figure 3A, lanes 2–6). To show the contribution of postlysis labeling, cells were pulsed for 1 h and harvested directly (Figure 3A, lane 1). Proteins were separated using a mini-gel system, which does not separate the β1i, β5, and β5i subunits from each other, explaining the slightly different labeling patterns in Figure 3 compared to Figure 1B. A 1-h pulse with **3** or **4** followed by a MG132 chase revealed substantial proteasome labeling (Figure 3A, lane 2). Labeling intensities increased with incubation time and were saturated at an incubation time of 4 h. Since the relative labeling patterns did not change (Figure 3A, lanes 3–6), the 1-h labeling period (Figure 3A, lane 2) was sufficient to obtain a representative proteasome labeling profile. When samples were not chased, additional labeling could be observed (in Figure 3A, compare lanes 1 and 2), suggesting that postlysis artifacts are likely to occur. To verify that this procedure ensures a proper live cell readout, the experiment was repeated with probe **2**, which should afford a live cell readout,[19] and with probes **6** and **9**, which localize to the mitochondria and should therefore barely label proteasomes in living cells. Cells were pulsed with probes **2**, **6**, and **9** and either chased with MG132 or harvested directly. When probe **2** was used,[19] the results were similar to those observed with probes **3**

and **4** (Figure 3B, lanes 1 and 2). The MG132 chase completely abolished labeling using probes **6** and **9** (Figure 3B, lanes 4 and 6). Without an MG132 chase, on the other hand, both probes afforded proteasome labeling (Figure 3B, lanes 3 and 5). These observations confirm the validity of the method applied and are a warning for post-lysis labeling that can be mistaken for a genuine live cell labeling event. It is, therefore, important to chase with a non-fluorescent pan-proteasome inhibitor like MG132, to ensure a correct interpretation of the results.

Fluorescent probe 4 versus dansylated probe 2

Having optimized the in-gel activity assay as described above, we compared dansylated probe **2** to fluorescent probe **4**. EL4 cells were pulsed with various concentrations of probe **2** or **4**, followed by a MG132 chase. Figure 3C shows that for a clear and equal visualization of all subunits, cells needed to be incubated with 5 μM **2**, while 50 nM **4** was sufficient. To confirm these data, EL4 cells were pulsed with probe **2** or **4** and chased with MG132, after which different amounts of total lysate were analyzed by SDS-PAGE. As seen in Figure 3D, 300 ng of total lysate was sufficient for proteasome visualization when probe **4** was used. Although cells were incubated with a 5-fold molar excess of probe **2** (5 μM) compared to probe **4** (1 μM), 2.5 μg of total lysate was needed to generate sufficient signal using probe **2**. From these results, it can be concluded that when living cells are probed for proteasome activity using the pulse-chase procedure described here, fluorescent probe **4** affords an up to 100-fold improvement in sensitivity compared to the dansylated probe **2** described previously and properties comparable to that of **3**, in principle allowing bioorthogonal use of multiple probes.

Profiling mouse tissues

Several methods have been used to study the bioavailability of proteasome inhibitors (see above).[17,18] However, sensitive methods that can be used to study the *in vivo* subunit-specific activity of proteasome inhibitors are in great demand. In our previous setup, probe was injected into the mouse, after which organs could be imaged and analyzed.[14] This setup is, however, not suitable for studying proteasome activity in patient material or in tissues where these probes may not accumulate sufficiently. We therefore employed an *ex vivo* approach, using probe **4** to identify the tissue-specific activity of the different catalytic subunits of proteasomes in freshly isolated organs. To determine whether probe **4** was effectively taken up in different cell types, segments of different organs were incubated with **4** (Figure 4A). Due to the complex nature of tissues (e.g., muscle cells in the heart), nuclear staining (DAPI) was used as counter stain (Supplementary Figure 7). Images were obtained by sequential scanning, and autofluorescence was found to be minimal (Supplementary Figure 7). Figure 4A shows that probe **4** was taken up by all cell types.

Next, we profiled proteasome activity in this panel of mouse tissues (Figure 4B). To this end, single-cell suspensions were obtained and red blood cells lysed. Cells were pulsed with compound **4** and chased with MG132. Samples were normalized to either cell number or protein content without observing significant differences in the results that were obtained. Band intensities were quantified and plotted as a percentage of the total proteasome labeling intensity (Figure 4B, bottom panel). A great variation in subunit-specific labeling was observed, depending on the organ that was probed (Figure 4B). Substantial labeling of immunoproteasome subunits was found in lymphoid organs (Figure 4B, lanes 1, 2, and 5). Immunoproteasome subunits

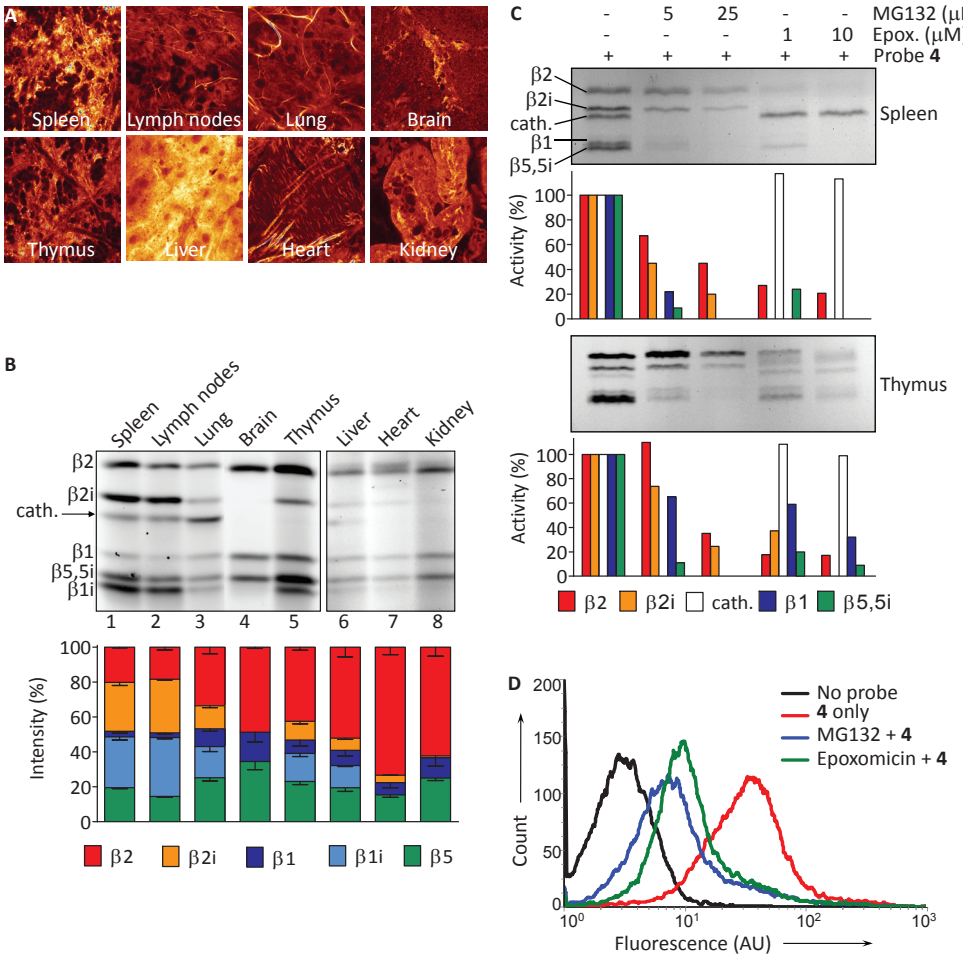


Figure 4 | Profiling the proteasome and the effects of proteasome inhibitor treatment in mouse tissues. A) CLSM images showing whole tissue segments from indicated organs incubated with 500 nM **4** (1 h). **B)** Gel image showing active subunit labeling in indicated tissues from mice, pulsed with 500 nM **4** (2 h) and chased with 5 μM MG132 (1 h). Proteins were separated using a 20 cm SDS–PAGE gel to allow separation of β1i and β5 subunits. Because of large differences in absolute labeling intensities between tissues, exposure times and image contrast settings were optimized for lanes 1–5 and 6–8 separately to allow clear visualization of labeled subunits. cath. represents cathepsin (top panel). Plot showing the quantification of band intensities, plotted as a percentage of total proteasome labeling intensity. Values are an average of two independent experiments ± the standard error of the mean (bottom panel). **C)** Gel images showing the active subunit labeling in whole tissue segments from spleen and thymus that were pulsed with the indicated concentrations of MG132 and epoxomicin (1 h) and chased with 500 nM probe **4** (1 h; top panels). Epox. represents epoxomicin. Plots showing quantification of band intensities, plotted as a percentage of labeling intensity in untreated tissue (bottom panels). **D)** Flow cytometry plots showing the fluorescence in mouse lymphocytes that were not incubated (black), incubated for 1 h with 50 nM **4** (red), or pulsed for 1 h with 1 μM epoxomicin (green) or 5 μM MG132 (blue) and chased with 50 nM **4** (1 h).

could also be detected in lung tissue (Figure 4B, lane 3), although at lower intensities. In brain, liver, heart, and kidney tissue (Figure 4B, lanes 4 and 6–8), active immunoproteasome subunits could not be detected. Immunoproteasome expression favors the production of antigenic peptides for MHC class I antigen presentation, and high-level immunoproteasome expression has been reported for lymphoid organs (spleen, lymph nodes, and thymus),[6] correlating with the immunoproteasome labeling by **4**. Thus, probe **4** seems well suited to visualization of immunoproteasome regulation.

Next to the presence of immunoproteasome subunits, an additional level of proteasomal activity regulation appeared visible. Although equal amounts of subunits β1, β2, and β5 are thought to be incorporated in functional proteasomes, different labeling intensities were observed in different tissues (Figure 4B), in line with the existence of different proteasome subtypes.[7,8] In lung and thymus, all subunits contributed equally to the total proteasome activity (Figure 4B, lanes 3 and 5). In brain, β2 and β5 subunits were labeled with the greatest intensity (Figure 4B, lane 4), while mainly β1 and β2 subunit labeling was observed in spleen and lymph nodes (Figure 4B, lanes 1 and 2). In the liver, kidney, and heart (Figure 4B, lanes 6–8), β2 subunits were most intensely labeled. These results indicate that different cell types display different ratios of proteasomal activities, suggesting that subunit activity is regulated in an organ-specific manner. This may influence substrate degradation, as well as the response to proteasome inhibitors. Probe **4** labeled an additional protein as well, which could not be immunoprecipitated with an antibody raised against the 20S proteasome (data not shown). Labeling in spleen tissue could be abolished by preincubation with MG132 (Figure 4C), which is known to inhibit cathepsins.[25] In H929 cells (human, multiple myeloma) and in spleen cell lysates, labeling was abolished by both the pan-cathepsin inhibitors LHVS[26,27] and JPM-565[28,29] (Supplementary Figure 8), suggesting that this band represents a cathepsin.

To validate whether an *ex vivo* approach would allow profiling of the pharmacodynamics of proteasome inhibitors in intact organs, intact tissue segments of spleen and thymus were incubated with different concentrations of MG132 and the broad-spectrum proteasome inhibitor epoxomicin (**11**) and chased with probe **4** (Figure 4C). Intact tissue segments were less sensitive to proteasome inhibitors, compared to single-cell suspensions. In the latter, 5 μM MG132 or 1 μM epoxomicin was sufficient to inhibit 90% of the proteasomal activity (Supplementary Figure 9). In intact tissue segments, on the other hand, 25 μM MG132 and 10 μM epoxomicin did not fully inhibit all proteasomal activities (Figure 4C). These results stress the importance of methods that can profile the *in vivo* inhibitory action of proteasome inhibitors. The specificity of both proteasome inhibitors in intact tissue segments could be profiled with ease. While MG132 exhibited a preference for the β5 and β1 subunits at low concentrations, as described previously,[19] epoxomicin inhibited all subunits with equal affinity.[30] Furthermore, MG132 inhibited not only the proteasomal subunits but also a cathepsin,[25] while epoxomicin was highly specific for the proteasome[30] (Figure 4C). To study whether these results could be confirmed by flow cytometry experiments, single-cell suspensions obtained from spleen tissue were pulsed with proteasome inhibitor and chased with **4**, or incubated with **4** alone. Cells were washed, and fluorescence was measured. As seen in Figure 4D, both inhibitors induced a substantial shift in intracellular fluorescence, indicating proteasome inhibition.

Together, these data indicate that two of the

fluorescent probes described here could be used to profile the *in vivo* pharmacological properties of proteasome inhibitors, both quantitative, using flow cytometry, and qualitative, using SDS-PAGE readouts to profile distinct proteasomal active subunits. Although our studies have been limited to mouse tissue, the method used should be equally applicable to any other sample, enabling proteasome profiling and competition assays in clinical settings.

CONCLUSION

We devised a range of fluorescent proteasome activity reporters. Two probes, Bodipy-TMR-labeled **3**,[14] and BodipyFL-labeled probe **4**, can efficiently bind to all catalytically active subunits of proteasomes in living cells. Other probes described herein may be used for the analysis of cell homogenates. In addition, when probe **4** is compared to previously developed methods for probing proteasome activity in living cells (probe **2**), the newly developed probe is up to 100-fold more sensitive. Using probe **4**, we set up gel-based and flow cytometry methods to study proteasome activity in living cells and to probe organ-specific proteasome activity. Our results indicate that different tissue types exhibit a different subunit labeling pattern, suggesting that subunit activity is regulated in an organ-specific manner. Furthermore, the tools and methods described here should be well-suited to the performance of pharmacological studies with proteasome inhibitors under clinical evaluation. In addition, the methodology described here should be applicable to monitoring proteasome activity in samples from patients that are treated with proteasome inhibitors.

MATERIALS AND METHODS

General methods

Bodipy and TAMRA dyes were purchased from Invitrogen or Molecular Probes. Sulfated cyanine (Cy3 and Cy5) dyes were purchased from Amersham, whereas the nonsulfated Cy5 analogue was synthesized according to a published procedure.[31] Peptide building blocks were purchased from Novabiochem and appropriately functionalized resins from Applied Biosystems. All solvents were purchased from Biosolve at the highest grade available. All other chemicals were purchased from Aldrich at the highest available purity. All solvents and chemicals were used as received. Liquid chromatography–mass spectroscopy (LC–MS) analyses were carried out on a WATERS LCT mass spectrometer in line with a WATERS 2795 HPLC system and a WATERS 2996 photodiode array detector. Analytical and preparative HPLC runs were performed on a Waters 1525 EF HPLC system coupled to a Waters 2487 dual λ absorbance detector.

Probe synthesis

See the Supplementary Information for detailed information on synthetic procedures and purity analysis.

Cell culture

EL4 (mouse, thymoma) and H929 (human, multiple myeloma) cells were cultured in RPMI 1640 (Invitrogen). Cell lines HeLa (human, cervix carcinoma) and MelJuSo (human, melanoma) were cultured in DMEM (Invitrogen). MelJuSo cells expressing β1i-GFP,[23] β7-RFP, or Ub-R-GFP[15] were cultured in DMEM in the presence of zeomycin (RFP) and geomycin (GFP) selection. Prior to CLSM imaging, cells were cultured overnight in DMEM lacking phenol red to prevent autofluorescence of the medium. All media were supplemented with 10% fetal calf serum (FCS), 100 units/mL penicillin, and 100 μg/mL streptomycin.

In-gel profiling of proteasomes in cell lysates

EL4 and HeLa cells were mechanically lysed using glass beads.[11] To this end, cells were washed with PBS, pelleted, and lysed by being vortexed at high speed at 4 °C for 45 min with 1 volume of glass beads (<106 μm, acid-washed, Sigma) and

1 volume of homogenization buffer [50 mM Tris (pH 7.4), 5 mM $MgCl_2$, 250 mM sucrose, 1 mM DTT, and 2 mM ATP]. Beads, membrane fractions, and cell debris were removed by centrifugation at 16000g for 5 min. Protein concentrations were determined using the Bradford assay (Bio-Rad), and 50 μg of protein was incubated with compounds **3**–**9** (500 nM) for 1 h at 37 °C. Proteins were denatured by being boiled in reducing sample buffer and analyzed by 12.5% SDS–PAGE using a PROTEAN II xi Cell system (Bio-Rad).

In-gel profiling of proteasomes in living cells

MelJuSo cells were incubated with 500 nM probe **3** or **4** or pulsed with 5 μM MG132 and chased with 500 nM **3** or **4**, both followed by a wash step. EL4 cells (1×10^6) in 1 mL of RPMI were incubated with 375 nM compound **3** or **4** for 1–5 h, 375 nM probe **6** or **9** for 1 h, or 5 μM **2** for 2 h, followed by a 1 h chase with 5 μM MG132. Control samples were incubated with 375 nM probe **3**, **4**, **6**, or **9** alone for 1 h or with 5 μM **2** alone for 2 h. Cells were harvested and lysed for 30 min in NP40 lysis buffer [50 mM Tris (pH 7.4), 150 mM NaCl, and 1% NP40] at 4 °C. The Bradford assay was used to measure protein content. For experiments using cathepsin inhibitors, H929 cells were incubated with 500 nM **4** (1 h) or pulsed with 10 or 50 μM LHVS or JPM-565 (1 h) and chased with 500 nM **4** (1 h). Cells were lysed in NP40 lysis buffer as described above.

Comparison between dansylated and fluorescent proteasome probes

EL4 cells (1×10^6/mL) were pulsed for 2 h with the indicated (Figure 4C) concentrations of probe **2** or **4**, followed by a 1 h chase with 5 μM MG132. Cells were harvested and lysed in NP40 lysis buffer. The Bradford assay was used to measure protein content, and equal amounts of protein were analyzed by 12.5% SDS-PAGE. Alternatively, EL4 cells (1×10^6/mL) were pulsed for 2 h with 5 μM **2** or 1 μM **4**, and chased with MG132 (5 μM, 1 h). Cells were lysed in NP40 lysis buffer, and protein content was determined as described above.

Mouse tissue analysis by SDS-PAGE

C57BL/6 mice were bred under pathogen-free conditions according to national and institutional guidelines and were sacrificed at 6–10 weeks of age. Subsequently, spleen, lymph nodes, lungs, liver, heart, kidneys, brain, and thymus were isolated. Single-cell suspensions were made using 70 μm nylon cell strainers (BD Biosciences). Heart tissue was pretreated with 1 mg/mL collagenase V (Sigma) and 1% DNase I (Ambion) for 30 min at 37 °C before single-cell suspensions were obtained. Red blood cells were lysed by hypotonic lysis, and cells were resuspended at a density of 5×10^6 cells/mL. Cells were incubated for 2 h with 500 nM **4**, followed by a 1 h chase with 5 μM MG132 to block postlysis labeling. Cells were harvested and lysed for 30 min at 4 °C using NP40 lysis buffer. Protein concentrations were determined using the Bradford assay. Lysate from 1×10^6 cells or equal amounts of protein were denatured by being boiled in reducing sample buffer, and polypeptides were analyzed by 12.5% SDS-PAGE using a PROTEAN II xi Cell system (Bio-Rad). To assess the effect of treatment with different proteasome inhibitors, whole tissue segments of spleen and thymus were placed in 1 mL of DMEM. In addition, a single-cell suspension was obtained from thymus (5×10^6 cells/mL). Tissue segments and single-cell suspensions were pulsed with MG132 (5 and 25 μM), epoxomicin (1 and 10 μM), or DMSO (1 h, 37 °C), chased with 500 nM **4** (1 h, 37 °C), and washed with fresh DMEM (1 h, 37 °C). Cells and tissue segments were lysed at 4 °C using NP40 lysis buffer. Protein concentrations were determined using the Bradford assay. Equal amounts of protein were denatured by being boiled in reducing sample buffer, and polypeptides were analyzed by 12.5% SDS-PAGE using a mini-gel system (Bio-Rad). For experiments using cathepsin inhibitors, spleen cells were lysed mechanically as described above. Equal amounts of protein were either incubated with 1 μM probe **4** (1 h) or pulsed with 10 or 50 μM LHVS or JPM-565 (1 h) and chased with 1 μM **4** (1 h).

In-gel fluorescence measurements

Proteins were denatured by being boiled in reducing sample buffer and analyzed by 12.5% SDS-PAGE. Wet gel slabs were imaged for 2 min, with a resolution of 100 μm, using a ProXPRESS 2D Proteomic imaging system (Perkin Elmer), using the following filter settings: probe **4**, λ ex/em = 480/530 nm; probes **3** and **5**–**7**, λ ex/em = 550/590 nm; probes **8** and **9**, λ ex/em = 625/680 nm. Images were analyzed using Totallab analysis software (Nonlinear Dynamics, Newcastle upon

Tyne, U.K.).

Assaying proteasome activity by immunoblotting (probe 2)

Proteins were denatured by being boiled in reducing sample buffer and separated by 12.5% SDS-PAGE gel followed by electro-transfer onto PVDF membranes. Immunoblotting was performed using an antidansyl-sulfonamidohexanoyl polyclonal Ab (1:1000, rabbit, Molecular Probes) and HRP-coupled swine anti-rabbit secondary antibody (DAKO) followed by enhanced chemiluminescence (SuperSignal West Dura Extended Duration, Pierce).

Flow cytometry experiments

Flow cytometry experiments were performed on a FACSCalibur 2 apparatus (BD Biosciences). For uptake experiments, MelJuSo cells were incubated with 500 nM probes **3–9** for 1 h. Cells were washed, trypsinized, and resuspended in medium, and intracellular fluorescence was measured. For inhibition experiments in cultured cells, both MelJuSo cells (control) and MelJuSo cells stably expressing a Ub-R-GFP construct were incubated with 500 nM probes **3–9** for 12 h. Cells were washed, trypsinized, and resuspended in medium. Subsequently, GFP fluorescence was measured. As a control for Ub-R-GFP expression, cells were incubated with 5 μM MG132 for 1.5 h, followed by analysis of GFP fluorescence. Since the fluorescence emitted by **4** is comparable in wavelength to GFP fluorescence (λ ex/em = 480/530 nm), nonengineered MelJuSo cells incubated with this probe were used to normalize the signal. For inhibition experiments in mouse tissue, a single-cell suspension from mouse spleen was obtained (see above). Samples were pulsed with either MG132 (5 μM) or epoxomicin (1 μM) for 1 h at 37 °C and chased with 50 nM **4** for 1 h at 37 °C. As a control, cells were incubated with **4** alone or not incubated. Cells were washed three times with fresh DMEM for 20 min at 37 °C. Subsequently, intracellular fluorescence was measured.

CLSM imaging of cultured cells

CLSM images were recorded using a Leica-SP2 system connected to an inverted microscope (Leica DM IR/E 2) and a tissue culture box at 37 °C. MelJuSo cells expressing β7-RFP or β1i-GFP were incubated for 1 h with 500 nM **3**, **4**, or **5**. Cells were washed with medium before being imaged. For the pulse–chase studies, MelJuSo cells expressing β1i-GFP or β7-RFP were incubated for 1 h with 5 μM MG132, washed with medium, and chased for 1 h with 500 nM **3** or **4**, respectively. To study probe **6** or **9**, MelJuSo cells were incubated for 1 h with either probe (500 nM) and washed. To visualize mitochondria, 200 nM MitoTracker green (Invitrogen) was added (30 min), and the cells were washed.

CLSM imaging of tissue slices

CLSM images were recorded using a Leica-SP2 AOBS system connected to an inverted microscope (Leica DM IR/E 2) and a tissue culture box at 37 °C with 5% CO_2. All images were obtained by sequential scanning and with identical settings. Spleen, lymph nodes, lungs, liver, heart, kidneys, brain, and thymus were isolated from C57BL/6 mice and sliced into small segments (~ 0.5 mm thick). Tissue segments were placed in 1 mL of DMEM and incubated with 500 nM **4** and 2×10^{-3} mg/mL DAPI. After incubation, the wet segment was placed on a glass coverslip with a drop of fresh medium, covered by a secondary coverslip, and imaged. To obtain autofluorescence images, tissue segments were not incubated, but directly imaged using identical settings.

ACKNOWLEDGMENTS

We thank Henk Hilkmann for assistance with automated peptide synthesis, Dr. Christoph Lippuner for technical assistance, and Dr. Danny Burg for in-depth discussions. This work was funded by the Dutch Cancer Society, Project Grant NKI 2005-3368.

REFERENCES

1. Hideshima, T. et al. The proteasome inhibitor PS-341 inhibits growth, induces apoptosis, and overcomes drug resistance in human multiple myeloma cells. *Cancer. Res.* **61**, 3071-3076 (2001).
2. Richardson, P.G., Mitsiades, C., Hideshima, T. & Anderson, K.C. Bortezomib: proteasome inhibition as an effective anticancer therapy. *Annu. Rev. Med.* **57**, 33-47 (2006).
3. Etlinger, J.D. & Goldberg, A.L. A soluble ATP-dependent proteolytic system responsible for the degra-

dation of abnormal proteins in reticulocytes. *Proc. Natl. Acad. Sci. U. S. A.* **74**, 54-58 (1977).
4. Glickman, M.H. & Ciechanover, A. The ubiquitin-proteasome proteolytic pathway: destruction for the sake of construction. *Physiol. Rev.* **82**, 373-428 (2002).
5. Concannon, C.G. et al. Apoptosis induced by proteasome inhibition in cancer cells: predominant role of the p53/PUMA pathway. *Oncogene.* **26**, 1681-1692 (2007).
6. Stohwasser, R., Standera, S., Peters, I., Kloetzel, P.M. & Groettrup, M. Molecular cloning of the mouse proteasome subunits MC14 and MECL-1: reciprocally regulated tissue expression of interferon-gamma-modulated proteasome subunits. *Eur. J. Immunol.* **27**, 1182-1187 (1997).
7. Dahlmann, B., Ruppert, T., Kuehn, L., Merforth, S. & Kloetzel, P.M. Different proteasome subtypes in a single tissue exhibit different enzymatic properties. *J. Mol. Biol.* **303**, 643-653 (2000).
8. Dahlmann, B., Ruppert, T., Kloetzel, P.M. & Kuehn, L. Subtypes of 20S proteasomes from skeletal muscle. *Biochimie.* **83**, 295-299 (2001).
9. Elliott, P.J., Soucy, T.A., Pien, C.S., Adams, J. & Lightcap, E.S. Assays for proteasome inhibition. *Methods Mol. Med.* **85**, 163-172 (2003).
10. Lightcap, E.S. et al. Proteasome inhibition measurements: clinical application. *Clin. Chem.* **46**, 673-683 (2000).
11. Bogyo, M., Shin, S., McMaster, J.S. & Ploegh, H.L. Substrate binding and sequence preference of the proteasome revealed by active-site-directed affinity probes. *Chem. Biol.* **5**, 307-320 (1998).
12. Kessler, B.M. et al. Extended peptide-based inhibitors efficiently target the proteasome and reveal overlapping specificities of the catalytic beta-subunits. *Chem. Biol.* **8**, 913-929 (2001).
13. Ovaa, H. et al. Chemistry in living cells: detection of active proteasomes by a two-step labeling strategy. *Angew. Chem. Int. Ed. Engl.* **42**, 3626-3629 (2003).
14. Verdoes, M. et al. A fluorescent broad-spectrum proteasome inhibitor for labeling proteasomes in vitro and in vivo. *Chem. Biol.* **13**, 1217-1226 (2006).
15. Dantuma, N.P., Lindsten, K., Glas, R., Jellne, M. & Masucci, M.G. Short-lived green fluorescent proteins for quantifying ubiquitin/proteasome-dependent proteolysis in living cells. *Nat. Biotechnol.* **18**, 538-543 (2000).
16. Fiebiger, E., Story, C., Ploegh, H.L. & Tortorella, D. Visualization of the ER-to-cytosol dislocation reaction of a type I membrane protein. *EMBO J.* **21**, 1041-1053 (2002).
17. Lindsten, K., Menendez-Benito, V., Masucci, M.G. & Dantuma, N.P. A transgenic mouse model of the ubiquitin/proteasome system. *Nat. Biotechnol.* **21**, 897-902 (2003).
18. Luker, G.D., Pica, C.M., Song, J., Luker, K.E. & Piwnica-Worms, D. Imaging 26S proteasome activity and inhibition in living mice. *Nat. Med.* **9**, 969-973 (2003).
19. Berkers, C.R. et al. Activity probe for in vivo profiling of the specificity of proteasome inhibitor bortezomib. *Nat. Methods.* **2**, 357-362 (2005).
20. Chauhan, D. et al. A novel orally active proteasome inhibitor induces apoptosis in multiple myeloma cells with mechanisms distinct from Bortezomib. Cancer Cell. **8**, 407-419 (2005).
21. Kraus, M. et al. Activity patterns of proteasome subunits reflect bortezomib sensitivity of hematologic malignancies and are variable in primary human leukemia cells. Leukemia. **21**, 84-92 (2007).
22. Murphy, M.P. & Smith, R.A. Drug delivery to mitochondria: the key to mitochondrial medicine. *Adv. Drug Deliv. Rev.* **41**, 235-250 (2000).
23. Reits, E.A., Benham, A.M., Plougastel, B., Neefjes, J. & Trowsdale, J. Dynamics of proteasome distribution in living cells. *EMBO J.* **16**, 6087-6094 (1997).
24. Lipinski, C.A., Lombardo, F., Dominy, B.W. & Feeney, P.J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* **46**, 3-26 (2001).
25. Kisselev, A.F. & Goldberg, A.L. Proteasome inhibitors: from research tools to drug candidates. *Chem. Biol.* **8**, 739-758 (2001).
26. Palmer, J.T., Rasnick, D., Klaus, J.L. & Bromme, D. Vinyl sulfones as mechanism-based cysteine protease inhibitors. *J. Med. Chem.* **38**, 3193-3196 (1995).
27. Riese, R.J. et al. Essential role for cathepsin S in MHC class II-associated invariant chain processing and peptide loading. *Immunity.* **4**, 357-366 (1996).
28. Meara, J.P. & Rich, D.H. Mechanistic studies on the inactivation of papain by epoxysuccinyl inhibitors. *J. Med. Chem.* **39**, 3357-3366 (1996).
29. Shi, G.P., Munger, J.S., Meara, J.P., Rich, D.H. & Chapman, H.A. Molecular cloning and expression of human alveolar macrophage cathepsin S, an elastinolytic cysteine protease. *J. Biol. Chem.* **267**, 7258-7262 (1992).
30. Meng, L. et al. Epoxomicin, a potent and selective proteasome inhibitor, exhibits in vivo antiinflammatory activity. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 10403-10408 (1999).
31. Korbel, G.A., Lalic, G. & Shair, M.D. Reaction microarrays: a method for rapidly determining the enantiomeric excess of thousands of samples. *J. Am. Chem. Soc.* **123**, 361-362 (2001).

Supplementary information

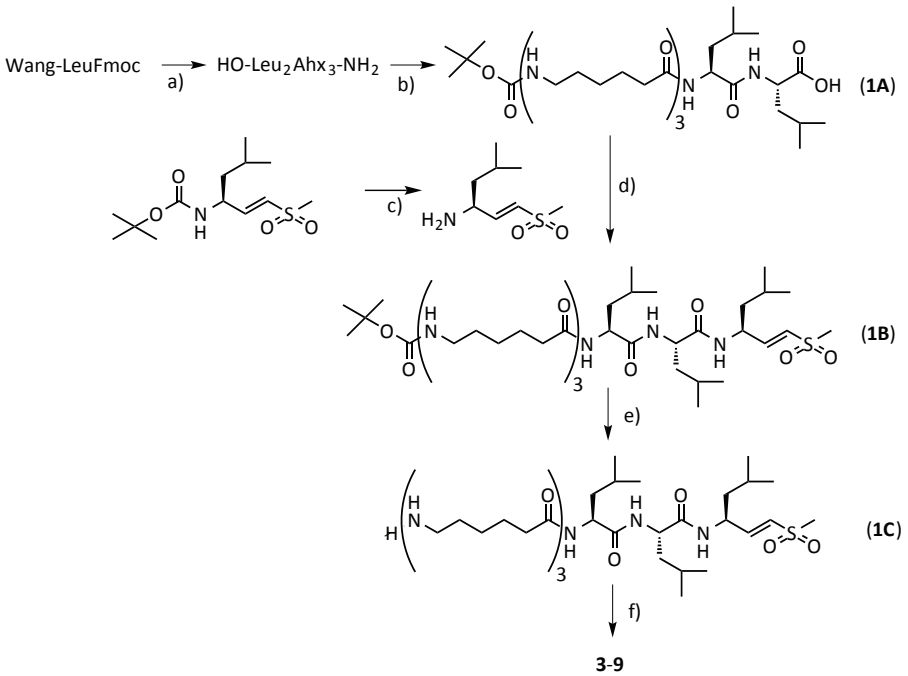
SUPPLEMENTARY METHODS

Boc-Ahx$_3$L$_2$-OH (1A)

Fmoc-Ahx$_3$L$_2$-Wang resin was synthesized using standard Fmoc-based solid-phase peptide synthesis protocols. Fmoc-L-Leu-PEG-polystyrene resin (Applied Biosystems) was subjected to subsequent coupling cycles, in which deprotection of the Fmoc-group with piperidine/NMP (1:4 v/v), was followed by coupling with Fmoc-Leu-OH or Fmoc-Ahx-OH (3 equivalents) and Di-isopropyl ethylamine (DIPEA, 3 equivalents). After the final coupling step, the Fmoc group was removed and subsequently the pentapeptide was released from the resin by treating the resin with TFA for 30 min to afford H$_2$N-Ahx$_3$L$_2$-OH, which was judged sufficiently pure by LCMS to be condensed directly with Di-tert-butyl-dicarbonate (1.2 equivalents) and DIPEA (5 equivalents) in DMF. The resulting suspension was stirred overnight to yield crude Boc-Ahx$_3$L$_2$OH (**1A**, Supplementary Figure 1) after concentration under reduced pressure. Boc-Ahx$_3$L$_2$OH was purified by silica column chromatography (10 % to 20% MeOH in CH_2Cl_2 containing 0.1 % HOAc). Alternatively, the crude could be used directly in the next step without further purification. MS(ESI): 684.46 $(M+H)^+$, calc: 684.48 $(M+H)^+$.

Boc-Ahx$_3$L$_3$-VS (1B)

Boc-Leu-VS was synthesized as described.[1] Briefly, deprotection of the Boc-group was achieved with TFA to yield LeuVS as a white powder upon concentration *in vacuo* and precipitation with diethylether. The resulting white powder (1.5 equivalents) was directly coupled to Boc-Ahx$_3$L$_2$OH by 1-Ethyl-3-(3-dimethylaminopropyl)-carbodiimide hydrochloride (EDC, 1.5 equivalents)-mediated condensation in the presence of DIPEA (3 equivalents) in DMF. The solution was stirred for 3



Supplementary Figure 1 | a) Fmoc-based solid phase peptide synthesis then TFA. b) Di-*tert*-butyl-dicarbonate, DIPEA, DMF. c) TFA. d) EDC, DIPEA, CH_2Cl_2/DMF. e) TFA. f) Activated ester of fluorescent dye, DMF, DIPEA, overnight, room temperature.

hours and concentrated under reduced pressure. Silica gel column chromatography (10 % MeOH in CH_2Cl_2) yielded homogeneous Boc-Ahx$_3$L$_3$-VS (**1B**, Supplementary Figure 1), as shown by HPLC (Supplementary Figure 10A) and LC-MS.

BodipyTMR-Ahx$_3$L$_3$VS (3)

BodipyTMR-Ahx$_3$L$_3$VS[2] (**3**) was prepared from Boc-Ahx$_3$L$_3$VS (**1B**) (10.0 mg, 11.6 μmol) and Bodipy TMR STP ester (5mg, 7.7 μmol) as described above. Purification was achieved using conditions identical to those described above to afford 8.0 mg of the title compound (**3**) (6.18 μmol, 80% yield two steps). R_f 0.65 (20% MeOH in CH_2Cl_2). HPLC revealed a single product (>95% purity) (Supplementary Figure 10B).

BodipyFL-Ahx$_3$L$_3$VS (4)

Boc-Ahx$_3$L$_3$VS (**1B**) (10.0 mg, 11.6 μmol) was dissolved in TFA (1 mL) and left at room temperature for 30 minutes. The solution was concentrated under reduced pressure and co-evaporated with toluene (2 mL 2x) and precipitated with ether (1 mL), which was carefully decanted and discarded. To the resulting white film was added a freshly prepared solution of Bodipy FL STP ester (5 mg, 9.2 μmol) in a mixture of DMF (0.5 mL) and DIPEA (50 μL). The resulting solution was left stirring overnight at room temperature under an argon atmosphere. The solution was concentrated under reduced pressure and purified by silica gel column chromatography (0 to 15% MeOH in EtOAc followed by 20% MeOH in CH_2Cl_2) fractions containing the title compound were pooled, concentrated, and further purified by reversed phase HPLC over a C-18 preparative column using a gradient of 75%-95% CH_3CN (0.05% TFA) in H_2O (0.05% TFA). Homogeneous fractions containing product (as judged by both TLC and LC-MS) were pooled and concentrated to a small volume and freeze-dried to afford 8.0 mg of the homogeneous (>99% pure) title compound (**4**) (2.7 μmol, 29% yield two steps). R_f 0.65 (20% MeOH in CH_2Cl_2). Analytical HPLC revealed a single product (>99% purity) (Supplementary Figure 10C).

General procedure preparation probes 5-9

Bodipy530/550-Ahx$_3$L$_3$VS (**5**), TAMRA-Ahx$_3$L$_3$VS (**6**), Cy3(SO_3^-)$_2$-Ahx$_3$L$_3$VS (**7**), Cy5(SO_3^-)$_2$-Ahx$_3$L$_3$VS (**8**), and Cy5-Ahx$_3$L$_3$VS (**9**) were prepared from Boc-Ahx$_3$L$_3$VS (**1B**) (2.0 mg) and from Bodipy530/550-, 5(6)TAMRA-, Cy3(SO_3^-)$_2$-, Cy5(SO_3^-)$_2$-, and Cy5-NHS ester (1 mg), respectively, as described above. Purification was achieved by HPLC as described, using the following gradients: 50%-80% CH_3CN (0.05% TFA), 50%-95% CH_3CN (0.05% TFA), 46%-56% CH_3CN (0.05% TFA), 30%-80% CH_3CN(0.05% TFA), and 30%-80% CH_3CN (0.05% TFA) in H_2O (0.05% TFA), to obtain **5**-**9**, respectively. HPLC revealed all products to be of at least 95% purity (Supplementary Figures 10D-H, for probes **5**-**9**, respectively).

Determination of probe concentrations

A dilution series of fluorophore-NHS ester in DMSO was made. 5 μL of each dilution in 500 μL total volume (ethanol) was excited at λ ex max and fluorescent emission at λ em max was measured using a Perkin Elmer LS50B luminescence spectrometer. Data obtained were fitted onto a standard curve and extrapolated. Purified dye-labeled probe was dissolved in DMF. A dilution series of dye labeled-probe was made in ethanol and fluorescence measurements were performed as described above with identical wavelength settings. Using the curve obtained from fluorophore-NHS ester measurements, the concentration fluorophore-labeled probe in DMF was calculated subsequently.

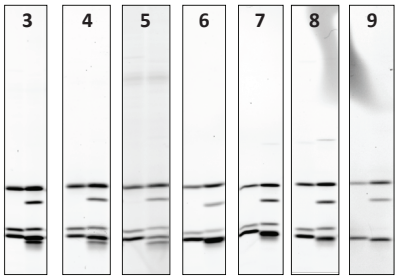
SUPPLEMENTARY INFORMATION ONLINE

HPLC profiles of compounds **1B** and **3**-**9** can be found in Supplementary Figure 10. This material is available via the Internet at http://pubs.acs.org, as supporting information of *Molecular Pharmaceutics* **4**, 739-748 (2007).

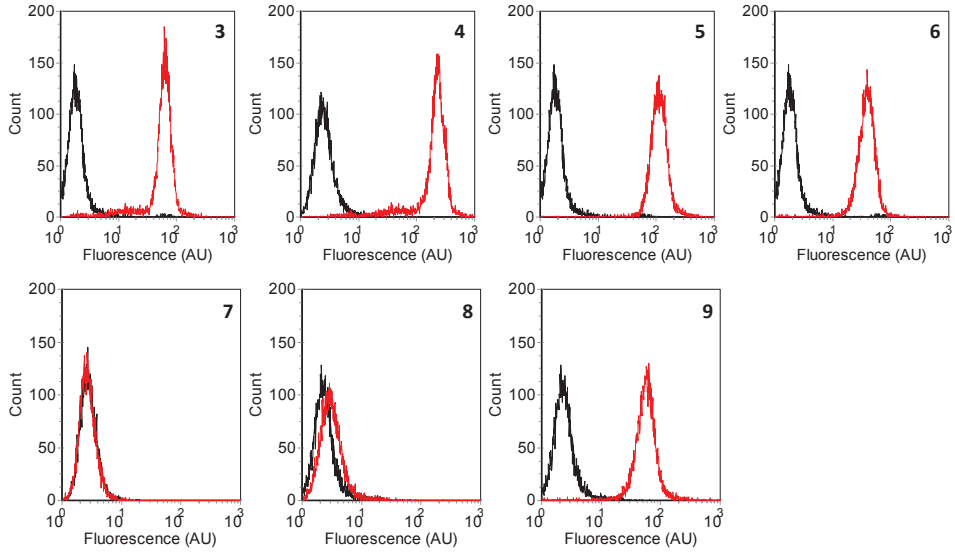
SUPPLEMENTARY REFERENCES

1. Rydzewski, R.M. et al. Optimization of subsite binding to the beta5 subunit of the human 20S proteasome using vinyl sulfones and 2-keto-1,3,4-oxadiazoles: syntheses and cellular properties of potent, selective proteasome inhibitors, *J. Med. Chem.* **49**, 2953-2968 (2006).
2. Verdoes, M. et al. A fluorescent broad-spectrum proteasome inhibitor for labeling proteasomes in vitro and in vivo, *Chem. Biol.* **13**, 1217-1226 (2006).

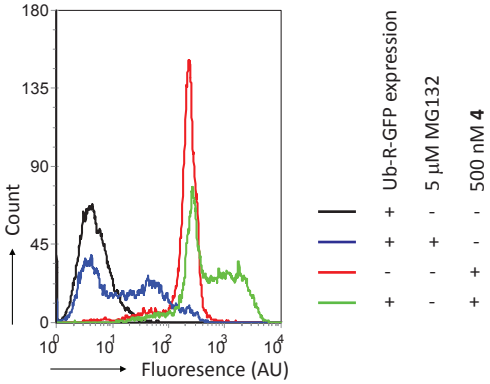
Supplementary figures



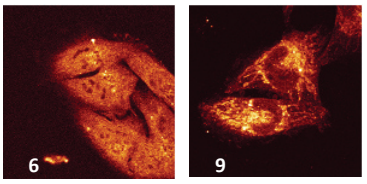
Supplementary Figure 2 | No unspecific labeling by the probes is observed in cell lysates. Gel images showing full scans of Figure 1B.



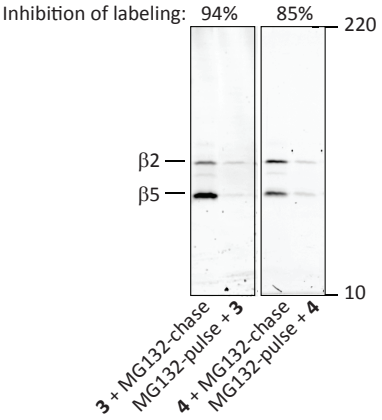
Supplementary Figure 3 | Probes 3, 4, 5, 6, and 9 enter cells. Flow cytometry plots showing the fluorescence of a population of non-incubated MelJuSo cells (black) vs. MelJuSo cells incubated with 500 nM probes **3**-**9** (1h) (red).



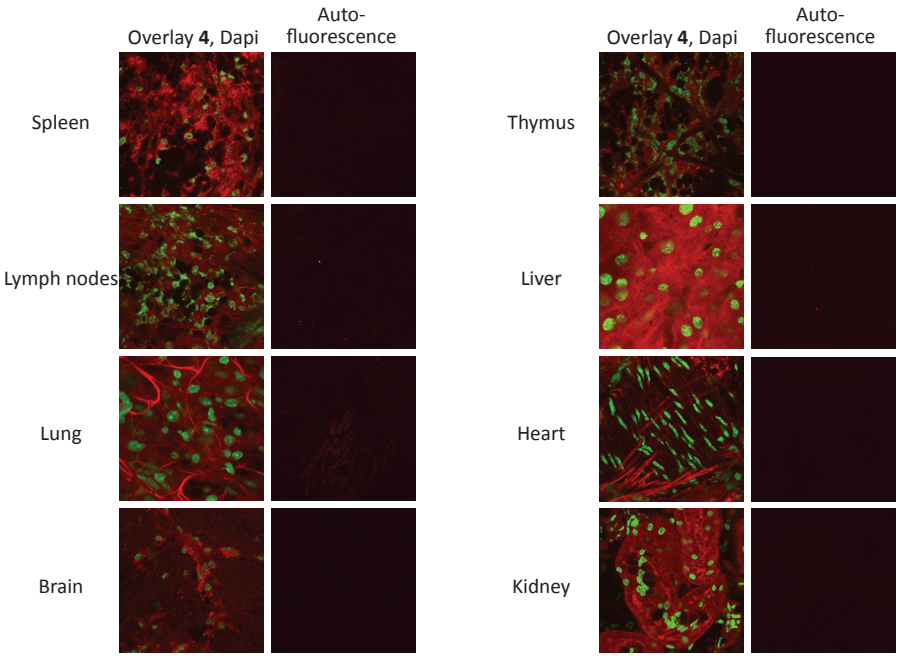
Supplementary Figure 4 | Flow cytometry plots showing normalization of Ub-R-GFP signal in probe 4 treated cells. Upon incubation with MG132, part of the Ub-R-GFP expressing cells show a shift in fluorescence (compare blue to black). Upon incubation with probe **4**, a similar shift is observed compared to non-transfected cells (compare green to red). We therefore concluded that probe **4** inhibits the proteasome in a similar percentage of cells as MG132 does.



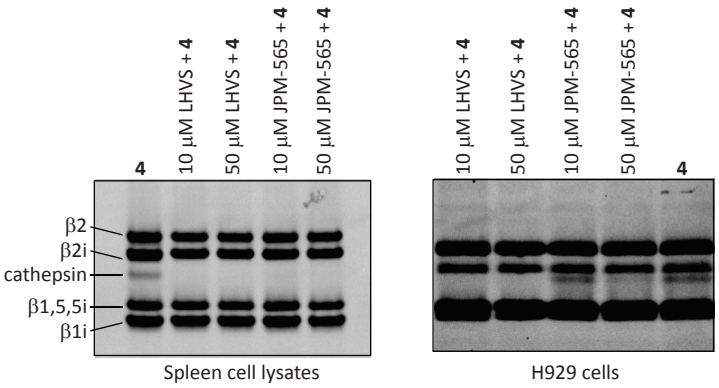
Supplementary Figure 5 | Localization of probes 6 and 9 upon fixation. CLSM images showing MelJuSo cells, incubated for 1 hour with 500 nM **6** or **9**, followed by fixation with formaldehyde.



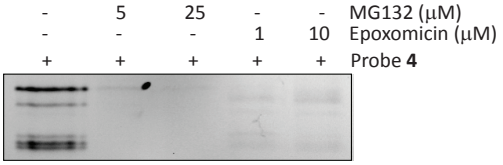
Supplementary Figure 6 | MG132 inhibits proteasome in MelJuSo cells. Gel images showing the result of a labeling experiment where cells were either pulsed with probe (1 h) and chased with MG132 (1 h) or pulsed with MG132 (1 h) and chased with probe (1 h).



Supplementary Figure 7 | Autofluorescence was found to be minimal when tissue segments were imaged by CLSM. CLSM images showing whole tissue segments from indicated organs incubated for one hour with 500 nM **4** (red) and 2×10^{-3} mg/mL DAPI (green), and the autofluorescence of different tissues.



Supplementary Figure 8 | Labeling of additional band by probe 4 in mouse tissue and H929 cells can be abolished by preincubation with cathepsin inhibitors. Gel image showing the active subunit labeling in H929 cells and cell lysates obtained from mouse spleen cells, pulsed with the indicated concentrations of LHVS and JPM-565 (1 h) and chased with probe **4** (1 h). Note that the image contrast settings were adjusted to visualize the cathepsin band, which was weakly labeled in these samples, causing relative intense proteasome band intensities.



Supplementary Figure 9 | Low concentrations of MG132 and epoxomicin are sufficient to fully inhibit the proteasome in single cell suspensions. Gel images showing the active subunit labeling in single cell suspensions obtained from mouse thymus, that were pulsed with the indicated concentrations of MG132 and epoxomicin (1 h) and chased with 500 nM probe **4** (1 h).

Chapter 2

Use of fluorescent proteasome activity probes

Chapter 2.1

Comparison of the specificity and activity profiles of the proteasome inhibitors bortezomib and CEP-18770

Submitted

Comparison of the specificity and activity profiles of the proteasome inhibitors bortezomib and CEP-18770

Celia R. Berkers,[a] Karianne G. Schuurman,[a] Yves Leestemaker,[a] Bruce Ruggeri,[b] Susan Jones-Bolin,[b] Michael Williams[b] & Huib Ovaa[a,*]

[a]Division of Cell Biology II, The Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands. [b]Cephalon Inc, Discovery Research, West Chester, PA, USA.

The ubiquitin proteasome system is an attractive pharmacological target for the treatment of cancer. The proteasome inhibitor bortezomib has been approved for the treatment of multiple myeloma (MM) and mantle cell lymphoma, but is associated with substantial adverse effects and the occurrence of resistance, underscoring the continued need for novel proteasome inhibitors. In this study, bortezomib and the novel proteasome inhibitor CEP-18770 were compared for their ability to inhibit proteasome activity using both fluorogenic substrates and a recently developed chemical proteasome activity probe. Bortezomib and CEP-18770 were equipotent in inhibiting distinct subunits of the proteasome in a panel of cell lines *in vitro*. In a preclinical MM model, both inhibitors inhibited the proteasome in normal tissues to a similar extent. Tumor proteasome activity was inhibited to a significantly higher extent by CEP-18770 (60%) compared to bortezomib (32%). In addition, CEP-18770 was able to overcome bortezomib resistance *in vitro*. The present findings demonstrate that proteasome activity probes can accurately monitor the effects of proteasome inhibitors on both normal and tumor tissues in preclinical models. Furthermore, the data presented here indicate that CEP-18770 has a favorable pharmacodynamic profile compared to bortezomib, and provide rationale for further clinical development of CEP-18770.

INTRODUCTION

The ubiquitin-proteasome system is responsible for the degradation of proteins involved in a wide variety of cellular processes, including cell proliferation and survival, cell-cycle control, transcriptional regulation and cellular stress responses, as well as for the destruction of abundant and misfolded proteins.[1,2] The catalytic activity of the 26S proteasome resides within its 20S core and is mediated by three catalytically active constitutive β-subunits termed β1, β2, and β5. These have three different catalytic activities, caspase-like, tryptic and chymotryptic activity, respectively. In lymphoid tissues or after induction with interferon γ, these subunits can be replaced by their immunoproteasomal counterparts, termed β1i, β2i, and β5i, to form the immunoproteasome, while mixed-type hybrid proteasomes have also been reported.[3-6] The increased sensitivity of malignant cells to proteasome inhibitors has resulted in proteasome inhibition emerging as a novel strategy

*Correspondence should be addressed to H.O. (h.ovaa@nki.nl).

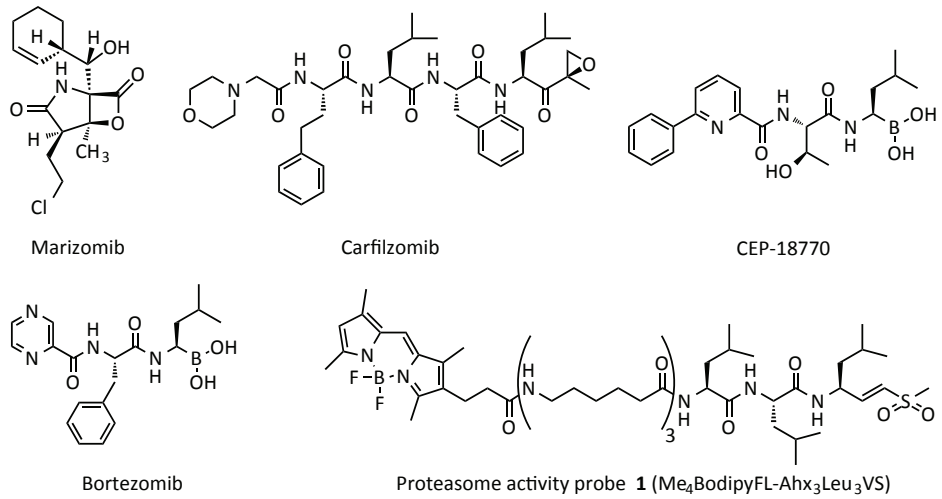


Figure 1 | Structures of bortezomib, CEP-18770, NPI-0052, carfilzomib and proteasome activity probe 1 (Me_4BodipyFL-Ahx_3Leu_3VS).

for cancer treatment.[2] In addition to eliciting direct apoptotic effects, proteasome inhibition may also enhance sensitivity and overcome drug resistance to standard chemo- and radiation therapy.[7] Bortezomib[8] (Figure 1), the active ingredient of velcade®, was the first proteasome inhibitor to be approved as a single agent for the treatment of relapsed or refractory multiple myeloma (MM)[9-11] and mantle cell lymphoma[12] and, in combination with PEGylated doxorubicin, for the treatment of relapsed MM.[13,14] Bortezomib induces apoptosis via stabilization of a number of proapoptotic and/or regulatory proteins. Both stabilization of p53 and inhibition of NF-κB (nuclear factor κB) mediated transcription via stabilization of its inhibitor IκBα, are thought to play important roles in bortezomib-induced apoptosis,[15,16] although additional mechanisms have been implicated to explain the anti-tumor efficacy of bortezomib.[17,18]

Bortezomib treatment is associated with manageable but substantial side effects e.g., thrombocytopenia and peripheral neuropathy[10] and the occurrence of both primary and acquired drug-resistance.[18] Consequently, there is an ongoing need for the development of novel proteasome inhibitors, and several second-generation proteasome inhibitors are now entering clinical trials. These include carfilzomib (PR-171),[19-21] an irreversible epoxyketone peptidyl inhibitor, and marizomib (NPI-0052),[22,23] an orally available proteasome inhibitor isolated from the marine actinomycete, *Salinispora spp* (Figure 1). Both these inhibitors demonstrate a favorable cytotoxicity profile as compared to bortezomib and can overcome bortezomib resistance *in vitro*.[19,20,22] Another second-generation proteasome inhibitor in clinical development is CEP-18770[24] (Figure 1). Like bortezomib, CEP-18770 is a reversible boronic acid-based inhibitor that suppresses NF-κB-mediated transcription, and induces apoptosis in human MM cell lines and patient derived cells to a similar extent.[25] It has a favorable cytotoxicity profile towards a variety of normal human cell lines as compared to bortezomib.[25] In a preclinical model of human MM, CEP-18770 evidenced significant and sustained proteasome inhibi-

tion in tumors relative to normal tissues and anti-tumor efficacy, resulting in complete tumor regression of MM xenografts.[25] Importantly, CEP-18770 is orally bioavailable, and oral administration of CEP-18770 resulted in significant reduction in tumor weight and dose-related incidence of complete tumor regression with minimal changes in animal body weight.[25,26]

In the present study, the proteasome inhibitory capacities of bortezomib and CEP-18770 were profiled side by side *in vitro* and in a variety of *in vivo* preclinical models, using both fluorogenic substrates and a recently developed chemical proteasome activity probe, that allows profiling of all constitutive and immunoproteasomal activities in live cells and *ex vivo*.[27] These studies demonstrate that both bortezomib and CEP-18770 inhibit β1 and β5 subunits to a similar extent in a panel of multiple myeloma cell lysates and cells. Furthermore, in a preclinical model of human MM, CEP-18770 inhibited tumor proteasome activity to a significantly higher extent (60% inhibition of total activity) compared to bortezomib (32% inhibition of total activity), while CEP-18770 inhibited the proteasome in normal tissues with equal potency compared to bortezomib. Finally, our data show that CEP-18770 is able to overcome bortezomib resistance *in vitro*.

RESULTS

Bortezomib and CEP-18770 show comparable inhibition profiles in a panel of cell lysates

CEP-18770 and bortezomib are chemically related (Figure 1) and have a comparable ability to induce apoptosis in MM cell lines and in purified primary human MM explants cultures, but show substantial differences in their cytotoxicity profile towards normal cells and in their antitumor-efficacy.[25] To assess whether these differences can be related to differences in proteasome inhibition, the proteasome inhibitory potential of CEP-18770 and bortezomib were first compared in cell lysates using fluorogenic substrates for the different proteasomal activities. Lysates from RPMI8226, H929, JJN3 (all MM) and HeLa (cervical carcinoma) cells were pre-incubated for 1 hour with increasing concentrations of CEP-18770 and bortezomib after which residual proteasome activity was measured (Figure 2A). IC_{50} values for the chymotryptic and caspase-like activities of the proteasome were calculated (Figure 2B). Bortezomib (black lines) and CEP-18770 (grey lines) showed almost identical inhibition profiles in all cell lines tested (Figure 2A) and inhibited both the chymotryptic activity (thin solid lines) and the caspase-like activity (dotted lines), but not the tryptic activity (thick solid lines), as has been described for bortezomib.[25,28] IC_{50} values for inhibition of the chymotryptic activity were 10–25 nM depending on the cell line used and did not differ significantly between bortezomib and CEP-18770 (Figure 2B). When IC_{50} values were compared between different tumor cell lines, RPMI8226 and H929 MM cells were more sensitive than JJN3 and HeLa cells to inhibition of chymotryptic activity. Whereas in RPMI8226 and H929 cells IC_{50} values between 9 and 13 nM were measured, HeLa and JJN3 cells required 20–24 nM of CEP-18770 or bortezomib for 50% inhibition (Figure 2B). Consistent with the finding that low concentrations of CEP-18770 inhibit the proteasome to a higher extent in H929 or RPMI8226 cells compared to JJN3 cells, a low concentration of CEP-18770 induced apoptosis in RPMI8226 and H929 cells, but not in JJN3 cells.[25] IC_{50} values for inhibition of the caspase-like activity were 40–75 nM depending on the cell line used. CEP-18770 was more specific for the chymotryptic activity as compared to bortezomib in both JJN3 and HeLa cells, but a significant difference between IC_{50} values ($p<0.05$) was only

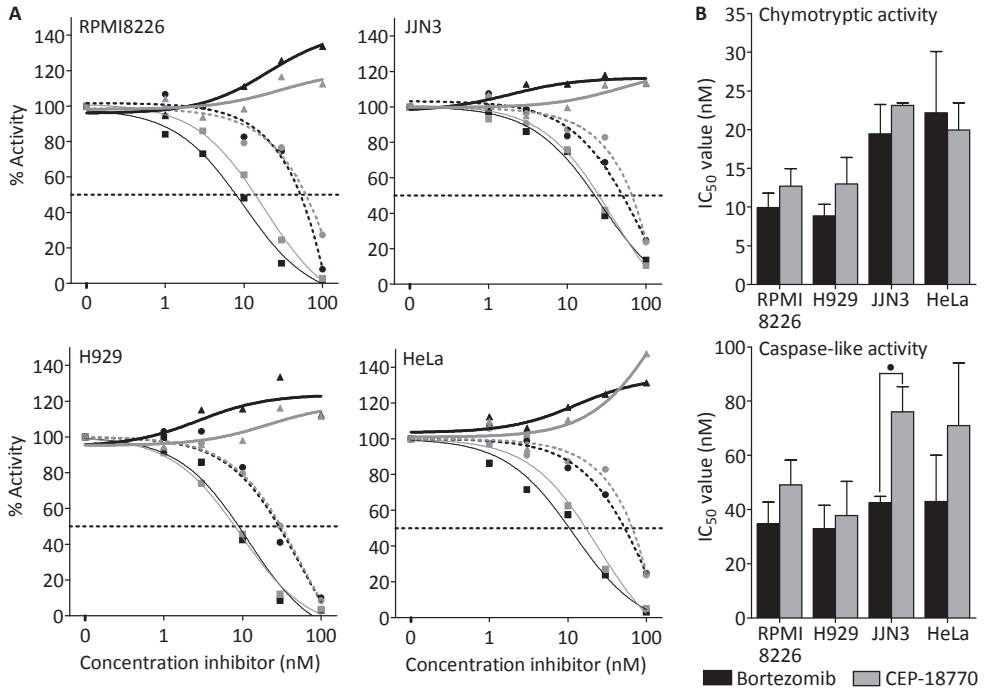


Figure 2 | Bortezomib and CEP-18770 have identical proteasome inhibition profiles in cell lysates. A) Representative proteasome inhibitory profiles in RPMI8226, H929, JJN3 and HeLa cell lysates after addition of increasing concentrations of bortezomib (black lines) or CEP-18770 (grey lines), obtained using fluorogenic substrates for the different proteasomal activities. Results were plotted as percentage activity compared to non-treated lysates. Thick solid lines: tryptic activity; thin solid lines: chymotryptic activity; dotted lines: caspase-like activity. **B)** IC_{50} values of CEP-18770 (grey bars) and bortezomib (black bars) in RPMI8226, H929, JJN3 and HeLa cell lysates determined using fluorogenic substrates for the different proteasomal activities. IC_{50} values were determined in three independent experiments and error bars indicate SEM. • $p<0.05$

observed in JJN3 cells (Figure 2B).

Bortezomib and CEP-18770 show comparable inhibition patterns in cells

The proteasome inhibitory effects of bortezomib and CEP-18770 in intact RPMI8226, H929, and HeLa cells were investigated. To this end, active proteasome subunits in cells were labeled using a fluorescent proteasome activity probe (Me_4BodipyFL-Ahx$_3$Leu$_3$VS (**1**); Figure 1) that contains a vinyl sulfone reactive group that covalently modifies all N-terminal threonine residues of catalytically active proteasomal subunits, and a fluorescent tag that allows in-gel visualization of the subunit-probe complex in an SDS-PAGE gel. When one or more proteasome active sites are targeted by CEP-18770 or bortezomib treatment, probe labeling is hampered, resulting in a decrease in in-gel fluorescence intensity and ultimately in the disappearance of specific bands.[27]

Cells were pulsed for 1 h with increasing concentrations of bortezomib and CEP-18770, followed by a 2-h chase with probe **1**, a close analog of the bodipyFL probe described previously.[27] Cells were then lysed, subunits

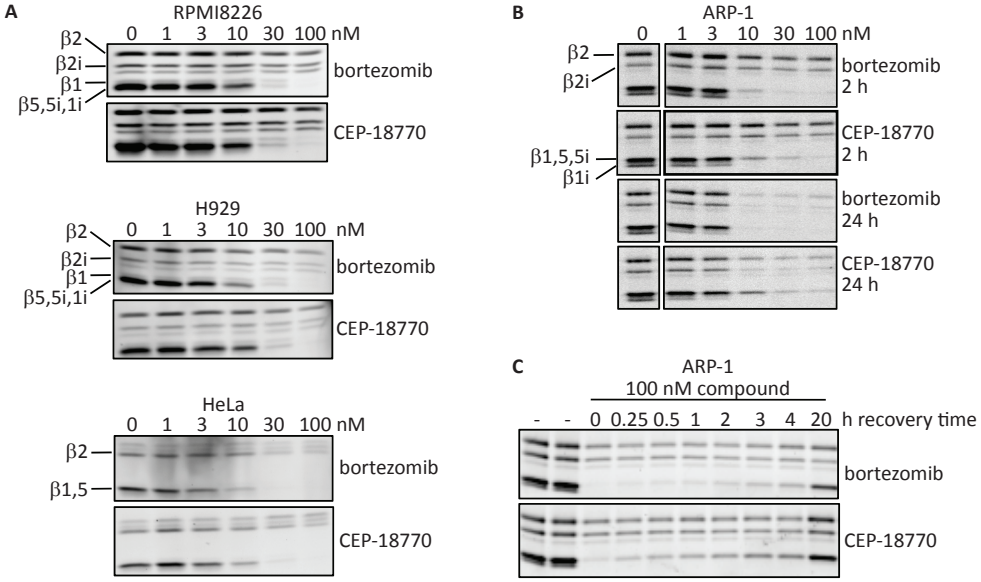


Figure 3 | Bortezomib and CEP-18770 induce near identical proteasome inhibition profiles and recovery patterns in live cells. A) In-gel fluorescence measurements showing representative proteasome activity profiles in RPMI8226, H929 and HeLa cells after a 1-hour incubation with increasing concentrations of bortezomib or CEP-18770. Results were obtained by incubating cells with inhibitor, followed by active proteasome labeling by probe **1**, and SDS-PAGE analysis. **B)** In-gel fluorescence measurements showing representative proteasome activity profiles in ARP-1 cells after a 2- or 24-hour incubation period with increasing concentrations of bortezomib and CEP-18770. Results were obtained by incubating cells with inhibitor, followed by active proteasome labeling by probe **1**, and SDS-PAGE analysis and fluorescence scanning. **C)** In-gel fluorescence measurements showing representative proteasome activity profiles in ARP-1 cells after a 1-hour incubation with 100 nM of bortezomib or CEP-18770, followed by the indicated recovery periods. Results were obtained by incubating cells with inhibitor, followed by active proteasome labeling by probe **1**, and SDS-PAGE analysis.

separated by SDS-PAGE and the gel scanned for fluorescence emission (Figure 3A). In untreated cells, the proteasome probe labeled all active subunits, and differences in subunit activity could be seen between different cell lines (Figure 3A, first lane). In H929 and RPMI8226 cells both constitutive and immunoproteasome subunits were active, as shown previously in RPMI8226 cells,[29] while in HeLa cells only constitutive subunits were labeled.[27,28] When cells were preincubated with either bortezomib or CEP-18770, the labeling of the β1 and β5 subunits disappeared in a concentration-dependent manner in all cell lines with estimated IC_{50} values in the low nanomolar range, while the labeling of the β2 subunits remained largely visible. These data confirm results in cell lysates and suggest that bortezomib and CEP-18770 are equipotent in inhibiting β1 and β5 subunits, but that they do not inhibit the β2 subunits in live cells.

Validating ARP-1 cells as a model system to compare bortezomib and CEP-18770 treatment *in vivo*

In a previous report, the *in vivo* efficacies of

CEP-18770 and bortezomib were compared in a RPMI8226 subcutaneous xenograft model of human MM, and in an ARP-1 systemic model of human MM.[25] In these models, differences were found between bortezomib and CEP-18770 in the extent and duration of tumor proteasome inhibition and antitumor efficacy. To assess whether the use of specific MM cell lines that respond differentially to CEP-18770 or bortezomib treatment were responsible for these observations, the effects of bortezomib and CEP-18770 were further studied in ARP-1 cells. ARP-1 cells were pulsed with increasing concentrations of bortezomib and CEP-18770 for 2 or 24 h, followed by a 2-h chase with proteasome probe **1**. Cells were then lysed, labeled subunits separated by SDS-PAGE and the resulting gel scanned for fluorescence emission. As can be seen in Figure 3B, both immuno- and constitutive proteasome subunits were labeled in untreated ARP-1 cells (Figure 3B, lane 1). In accord with results inform the RPMI8226, H929 and HeLa cells, bortezomib and CEP-18770 showed near identical inhibition profiles in ARP-1 cells (Figure 3B). After a 2-hour incubation period, both compounds dose-dependently inhibited β1 and β5 subunits with IC_{50} values between 3 and 10 nM, while β2 subunits remained relatively unaffected. Also after 24 h, both compounds inhibited β1 and β5 subunits, but in addition the labeling of β2 subunits was largely abolished at higher concentrations. As bortezomib and CEP-18870 do not interact with β2 subunits directly at the concentrations investigated (Figure 2A), this may be due to a secondary cellular response.

To investigate whether ARP-1 MM cells show differences in proteasome activity recovery after either bortezomib or CEP-18770 treatment, ARP-1 cells were pulsed for 1 h with 100 nM bortezomib or CEP-18770, washed and either probed directly (Figure 3C, lane 3), or allowed to recover for the indicated time periods before being analyzed. The patterns of activity recovery after withdrawal of either bortezomib or CEP-18770 were very similar (Figure 3C). After withdrawal of either inhibitor, proteasome activity recovered only slowly. After a 20-h recovery period, activity of the β5 subunits had not completely returned to baseline levels, while the activity of the β1i subunit was not detected (Figure 3C). These results indicate that although boronic-acid based proteasome inhibitors inhibit the proteasome in a reversible manner, proteasome synthesis, which occurs only after longer time periods, is the predominant factor contributing to recovery. Together, these data demonstrate that ARP-1 cells respond similarly to bortezomib and CEP-18770 treatment with respect to both proteasome inhibition and proteasome recovery patterns.

CEP-18770 shows an improved pharmacodynamic profile compared to bortezomib

To evaluate whether bortezomib and CEP-18770 show differences in proteasome inhibition *in vivo*, proteasome probe **1** was used to profile proteasome inhibition using an *ex vivo* approach.[27] NOD/SCID mice bearing subcutaneous ARP-1 tumor xenografts (n=5/group) were administered with the maximum tolerated dose (MTD) of bortezomib (1.2 mg/kg intravenous), CEP-18770 (4 mg/kg intravenous) or Solutol/PBS vehicle. At 2 h, 8 h or 24 h post dosing, mice were sacrificed and heart, lung, liver, spleen and tumor tissue obtained. Tissue cells were lysed, and lysates probed for proteasome activity, followed by SDS-PAGE. The resulting gels were scanned for fluorescence emission (Figure 4A) and the intensities of all bands were quantified using a sypro stain to control for protein loading. In all tissues, vehicle treatment had no effect on proteasomal subunit activity over a 24-h period (data not shown). Proteasome subunit activities in CEP-18770 and bortezomib treated samples were

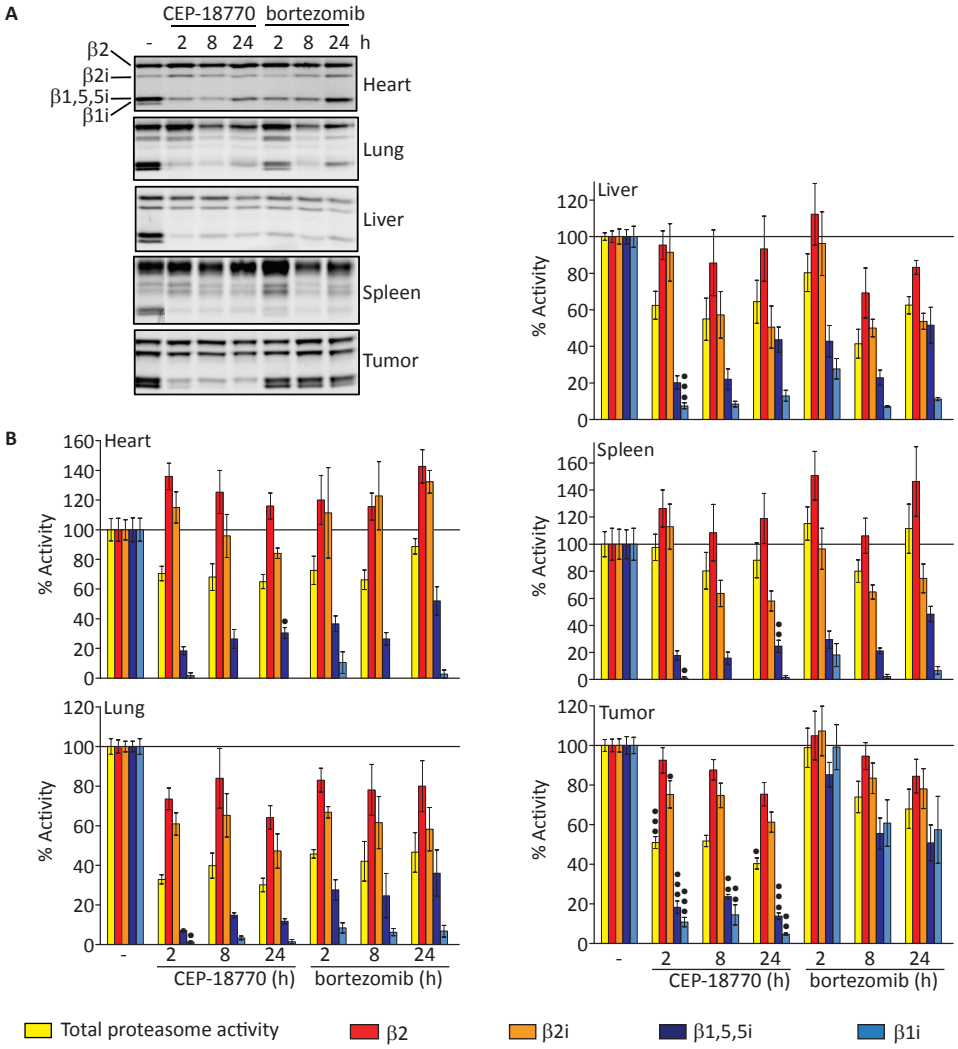


Figure 4 | CEP-18770 shows an improved pharmacodynamic profile compared to bortezomib in an ARP-1 MM tumor model. NOD/SCID mice (n=5/group) bearing ARP-1 tumor xenografts received a single injection of the maximum tolerated dose of bortezomib or CEP-18770 or vehicle, and tissues were removed at t = 2 h, 8 h, or 24 h post dosing. Tissues were lysed, followed by active proteasome labeling by probe **1** and SDS-PAGE analysis. **A)** In-gel fluorescence measurements showing representative proteasome activity profiles in mouse heart, lung, liver, spleen and tumor tissues after bortezomib or CEP-18770 treatment for the indicated time periods. **B)** Quantification of proteasome activity profiles in mouse heart, lung, liver, spleen and tumor tissues after bortezomib or CEP-18770 treatment for the indicated time periods. The proteasome subunit activity in treated samples was plotted as a percentage of the subunit activity in vehicle treated mice and total proteasome activity was defined as the sum of all individual activities. Error bars represent SEM. Statistical analysis was performed using GraphPad Prism software. CEP-18770 treated samples in which inhibition differed significantly from the corresponding bortezomib treated samples are indicated with dots. • $p<0.05$; •• $p<0.01$; ••• $p<0.001$

calculated as a percentage of the activities in vehicle-treated samples and averaged (Figure 4B). Total proteasome activity was defined as the sum of all individual activities.

In normal murine tissues, CEP-18770 and bortezomib showed comparable inhibition and recovery patterns, although subtle differences in inhibition patterns could be observed between the two inhibitors as well as between different tissue types. Bortezomib and CEP-18770 inhibited β1 and β5 subunits (Figure 4B, blue bars) to a similar degree in all tissues at all time points, with residual activities ranging between 20–50% for the β5,5i,1 subunits (dark blue bars) and 0–10% for the β1i subunit (light blue bars). The β2 subunit (red bars) was least affected by compound treatment. A small decrease in labeling was only observed in lung tissue. Labeling of the β2i subunit (orange bars) was decreased to a somewhat higher extent in lung, liver and spleen, predominantly at later time points. These results are in line with the results from ARP-1 cells (Figure 3B), suggesting that this was a secondary cellular response to proteasomal β1 and β5 subunit inhibition. Additional, little recovery of the activity of individual subunits was seen in this model over a period of 24 h, although in individual mice, proteasome activity appeared to recover in spleen, liver and heart after treatment with either compound (data not shown). Total proteasome activity (Figure 4B, yellow bars) was only partially inhibited in all normal tissues, predominantly because the β2,2i activities substantially contributed to total proteasome activity in all tissues examined (Figure 4A). The largest effects on total activity were observed in lung, where 40% residual proteasome activity was measured. In heart and liver, 60–70% of total proteasome activity remained after CEP-18770 and bortezomib treatment, while in spleen, total proteasome activity hardly decreased compared to vehicle (Figure 4B). No significant differences in total proteasome inhibition were observed between CEP-18770 and bortezomib in healthy murine tissue at all time points.

In ARP-1 tumor xenografts, differences between the inhibition produced by CEP-18770 and bortezomib were more pronounced. In primary tumor tissues β5 and β1 subunits were inhibited to a significantly higher extent by CEP-18770 treatment compared to bortezomib treatment at all time points, with residual activities ranging between 10–20% after CEP-18770 treatment and 60–80% after bortezomib treatment (Figure 4B). CEP-18770 induced maximal inhibition of these subunits within 2 h while maximal inhibition by bortezomib occurred only after 8–24 hours. Additionally, treatment with either compound resulted in small decreases in β2 and β2i labeling over time, as described above for normal tissues. As a result, the residual proteasome activity was significantly lower in CEP-18770-treated tumors as compared to bortezomib-treated tumors at 2 h and 24 h (Figure 4B). CEP-18770 reduced total proteasome activity to 40–50%, suggesting that this compound inhibited tumor proteasome to a similar or greater extent than proteasome in normal tissues. In contrast, the residual activity in bortezomib-treated tumor samples was 70–100%, suggesting that bortezomib inhibited proteasome activity in tumors to a lesser extent than in most normal tissues. Collectively, these data indicate that CEP-18770 induces a more favorable pharmacodynamic response in the ARP-1 MM tumor model as compared to bortezomib.

CEP-18770 displays activity against cells that are resistant to bortezomib

We next investigated whether CEP-18770 could overcome bortezomib resistance in the bortezomib resistant cell line THP1/BTZ_{100}.[30] THP1/BTZ_{100} cells acquired bortezomib resis-

tance by exposure to stepwise increasing concentrations of bortezomib (2.5 to 100 nM). Study of the molecular mechanism of bortezomib resistance in these cells revealed an Ala49Thr mutation in the β5 subunit protein.[30] This mutation has also been observed in Jurkat cells (human lymphoblastic T cells) with acquired bortezomib resistance.[31,32] THP1/BTZ_{100} cells were shown to be cross-resistant to other proteasome inhibitors that target the β5 subunit, including MG132, MG262 and 4A6.[30]

First, we compared the proteasome inhibitory potential of CEP-18770 and bortezomib in cell lysates of THP1/BTZ_{100} and THP1 wild type (THP1/WT) cells using fluorogenic substrates for the different proteasomal activities as described above. Lysates were pre-incubated for 1 hour with increasing concentrations of CEP-18770 and bortezomib after which residual proteasome activity was measured (Figure 5A) and IC_{50} values were calculated for the chymotryptic activity of the proteasome (Figure 5B). Bortezomib (black lines) and CEP-18770 (grey lines) showed near identical inhibition profiles in THP1/WT cells (dotted lines), in accordance with data obtained in other cell types (Figure 2A). Higher concentrations of bortezomib were required to inhibit the chymotryptic activity in THP1/BTZ_{100} cells as compared to THP1/WT cells, as expected (Figure 5A,B). Remarkably, CEP-18770 inhibited the chymoptryptic activity in THP1/BTZ_{100} cells with equal potency compared to the WT cell line (Figure 5A,B). In THP1/BTZ_{100} cells, no differences between bortezomib and CEP-18770 were observed in the degree of inhibition of the tryptic and caspase-like activities of the proteasome (Figure 5A).

Next, we set out to confirm that CEP-18770 was able to inhibit the β5 subunit in intact THP1/BTZ_{100} cells. To this end, THP1/WT and THP1/BTZ_{100} cells were pulsed with increasing concentrations of bortezomib and CEP-18770 for 3 or 24 h, followed by a 2-h chase with proteasome probe. Cells were then lysed, labeled subunits separated by SDS-PAGE and the resulting gel scanned for fluorescence emission. In line with results inform other cell lines tested, bortezomib and CEP-18770 showed near identical inhibition profiles in THP1/WT cells (Figure 5C, left panels). After a 3-hour incubation period, both compounds dose-dependently inhibited β1 and β5 subunits with IC_{50} values between 3 and 10 nM, while β2 subunits remained relatively unaffected. After 24 h, the labeling of β2 subunits by probe **1** was also hampered at higher concentrations of bortezomib and CEP-18770. In contrast, CEP-18770 and bortezomib showed different inhibition profiles in THP1/BTZ_{100} cells (Figure 5C, right panels). CEP-18770 inhibited the β5 subunits in THP1/BTZ_{100} cells with IC_{50} values between 3 and 10 nM, while higher concentrations of bortezomib were required to reach a similar extent of β5 inhibition. These differences are in accordance with results found in cell lysates using fluorogenic substrates and suggest that the β5 subunits in bortezomib resistant THP1 cells can still be inhibited by CEP-18770.

Finally, we investigated how the cell viability of THP1/BTZ_{100} and THP1/WT cells was influenced by bortezomib and CEP-18770 treatment. To this end, cells were exposed to increasing concentrations of bortezomib and CEP-18770 for 24 h or 48 h, after which cell viability was assessed using a cell titer blue assay. Values were normalized to control samples that were incubated with DMSO only and results were plotted as percentage compared to control values (Figure 5D). EC_{50} values, defined as the concentration of inhibitor causing 50% cell death, were determined at 48 h (Figure 5E). In accord with the proteasome inhibitory profiles found in these cells using probe **1**, THP1/WT cells responded similarly to bortezomib and CEP-18770 treatment

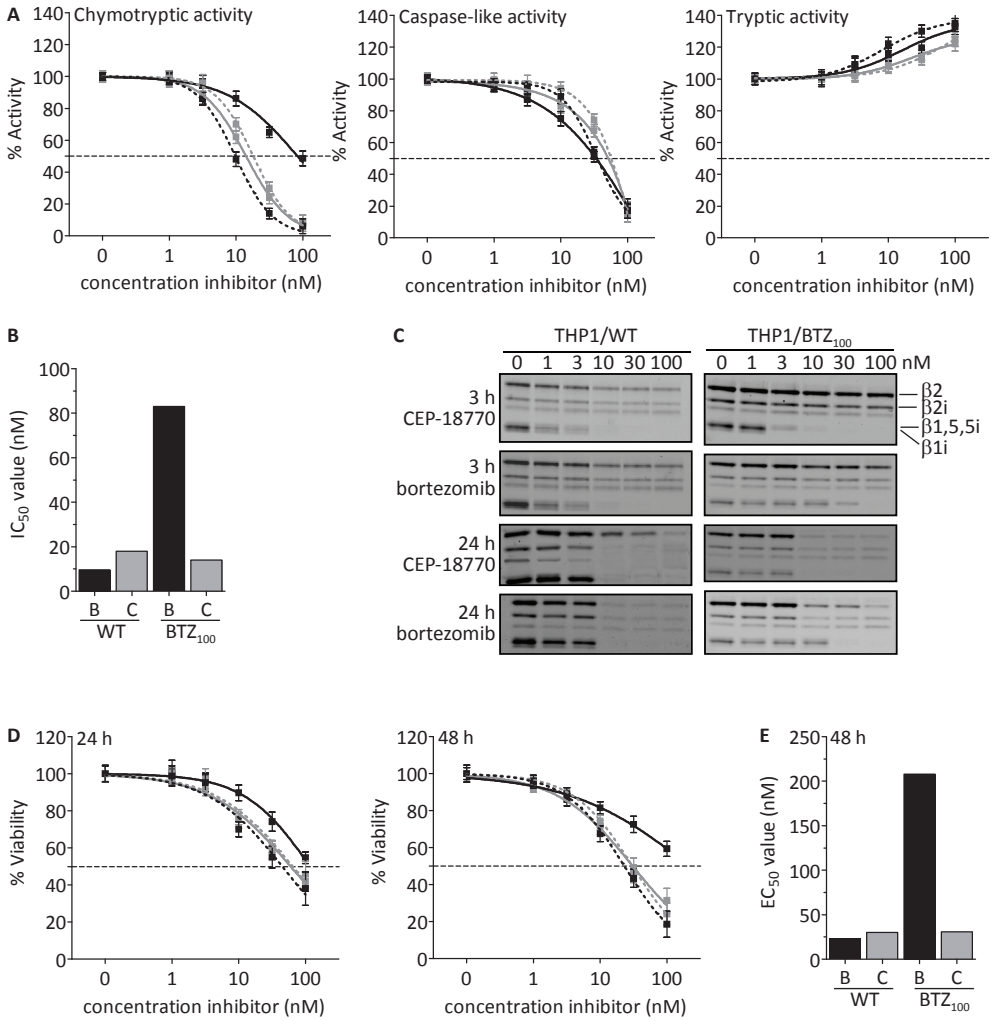


Figure 5 | CEP-18770 can overcome bortezomib resistance. A) Proteasome inhibitory profiles in THP1/WT and THP1/BTZ_{100} cell lysates after addition of increasing concentrations of bortezomib (black lines) or CEP-18770 (grey lines), obtained using fluorogenic substrates for the different proteasomal activities. Results were plotted as percentage activity compared to non-treated lysates. Solid lines: THP1/BTZ_{100} cells; dotted lines: THP1/WT cells. Error bars represent SEM of technical triplicates. **B)** IC_{50} values of CEP-18770 (grey bars) and bortezomib (black bars) in THP1/WT and THP1/BTZ_{100} cell lysates for inhibition of the chymotryptic activity. **C)** In-gel fluorescence measurements showing representative proteasome activity profiles in THP1/WT and THP1/BTZ_{100} cells after a 3 or 24-hour incubation with increasing concentrations of bortezomib or CEP-18770. Results were obtained by incubating cells with inhibitor, followed by active proteasome labeling by probe **1**, and SDS-PAGE analysis. **D)** Viability of THP1/WT and THP1/BTZ_{100} cells after incubation with increasing concentrations of bortezomib (black lines) or CEP-18770 (grey lines) for 24h or 48h. Results were plotted as percentage viability compared to a DMSO control. Solid lines: THP1/BTZ_{100} cells; dotted lines: THP1/WT cells. Error bars represent SEM of technical triplicates. **E)** EC_{50} values of CEP-18770 (grey bars) and bortezomib (black bars) in THP1/WT and THP1/BTZ_{100} cells.

(Figure 5D). In THP1/BTZ_{100} cells on the other hand, bortezomib showed higher EC_{50} values as compared to THP1/WT cells, whereas THP1/BTZ_{100} and THP1/WT cells were equally sensitive to CEP-18770 (Figure 5D,E). The EC_{50} value of CEP-18770 in THP1/BTZ_{100} cells was approximately 30 nM, whereas 100 nM bortezomib was not sufficient to kill 50% of all THP1/BTZ_{100} cells (Figure 5E). From the data obtained in THP1/BTZ_{100} and THP1/WT cells we conclude that THP1/BTZ_{100} cells are resistant to bortezomib due to a mutation in the β5 protein, which prevents bortezomib binding. Remarkably, THP1/BTZ_{100} cells are still equally sensitive to proteasome inhibition by CEP-18770 as compared to THP1/WT cells, and as a consequence, CEP-18770 displays similar EC_{50} values in THP1/BTZ_{100} and THP1/WT cells. Therefore, our data suggest that CEP-18770 can overcome acquired bortezomib resistance *in vitro*.

DISCUSSION

Proteasome inhibition is a validated strategy for the treatment of cancer, as evidenced by bortezomib, a boronic acid-based proteasome inhibitor, which is the active ingredient of velcade® and which is currently the only proteasome inhibitor approved for human use. Bortezomib treatment is however associated with adverse effects,[10] including thrombocytopenia[33] and peripheral neuropathy,[34] and with both primary and secondary resistance.[18] This has led to the search for new proteasome inhibitors with an improved cytotoxicity profile and the potential to overcome bortezomib resistance. CEP-18770, a novel proteasome inhibitor that is structurally similar to bortezomib, is one of several second-generation inhibitors in clinical trials.[24,25]

In the present study the proteasome inhibitory profiles of bortezomib and CEP-18770 were compared in cell lysates, whole cells and *ex vivo* in a mouse model of human MM, using both fluorogenic substrates and a chemical proteasome activity reporter that allows profiling of proteasome activity with high sensitivity in a wide range of cell types and tissues.[27] In addition, we investigated the proteasome inhibitory profiles of CEP-18770 and bortezomib in cells with acquired resistance to bortezomib, and assessed the effect of either CEP-18770 or bortezomib treatment on the cell viability of these cells.

CEP-18770 and bortezomib showed comparable inhibition patterns in a panel of human tumor cell lines. Both compounds were equipotent in inhibiting the β1 (caspase-like activity) and β5 (chymotryptic activity) subunits and had little effect on the β2 subunit activity (tryptic activity), in agreement with previous reports.[25,28] IC_{50} values in cell lysates were 40–75 nM and 10–25 nM for inhibition of the caspase-like and chymotryptic activity, respectively. In cell lines, IC_{50} values of both compounds for both the β1 and β5 subunits ranged between 3–10 nM. These data are in agreement with a previous report, in which bortezomib and CEP-18770 showed comparable effects in cells downstream of proteasome inhibition, including a comparable accumulation of polyubiquitinated proteins and the appearance of activated caspases-3, -7, and -9 with a comparable kinetic profile.[25] Data obtained using either fluorogenic substrates in cell lysates or using probe **1** in intact cells were similar, indicating the use of probe **1** provides a reliable method to assess proteasome activity.

Proteasome activity recovered similarly after withdrawal of either inhibitor from cells, with immediate but very slow recovery of β5(i)/1 subunit activity, and no recovery of the β1i activity within 20 hours. The slow rate of proteasome activity recovery after inhibition by either compound indicates that proteasome synthesis, and not the reversibility of boronic-

acid based proteasome inhibitors, is the predominant factor contributing to recovery, as suggested previously.[19] The factors contributing to a differential recovery of proteasomal activities are unknown, but differences in recovery between subunits have been reported.[22,35] After treatment with bortezomib, the recovery of caspase-like activity was delayed compared to that of chymotryptic activity.[22] Furthermore, continuous bortezomib treatment can abrogate immunoproteasome subunit expression and thus induce an altered subunit composition of newly synthesized proteasomes.[35]

In contrast to the finding that CEP-18770 and bortezomib have near identical effects on tumor cell proteasome activity *in vitro*, marked differences between CEP-18770 and bortezomib were found *in vivo* in an ARP-1 MM xenograft model. In tumor xenografts, CEP-18770 showed a significantly higher degree of inhibition of β1 and β5 subunits at all time points measured as compared to bortezomib (80–90% inhibition by CEP-18770 as compared to 40–60% inhibition by bortezomib), resulting in a significantly higher degree of overall proteasome inhibition (60% inhibition by CEP-18770 as compared to 32% inhibition by bortezomib), while the response of normal, healthy murine tissues to treatment with the maximum tolerated dose of bortezomib or CEP-18770 was near identical. CEP-18770 thus reduced proteasome activity in tumors to a greater extent than in most normal tissues, while bortezomib inhibited proteasome activity in tumors to a lesser extent than in most normal tissues, resulting in a favorable pharmacodynamic profile for CEP-18770 as compared to bortezomib. These results are in agreement with Piva et al., who have shown that in a RPMI8226 MM tumor xenograft model, CEP-18770 induced a similar degree of proteasome inhibition in normal mouse peripheral tissues compared to bortezomib, but a greater and more-sustained inhibition of tumor proteasome activity.[25] Consistent with these findings, CEP-18770 treatment has been shown to have a greater *in vivo* antitumor efficacy, and CEP-18770 administration resulted in a dose-related induction of complete tumor regression, whereas bortezomib treatment only delayed tumor growth.[25] Other studies have shown that bortezomib reduced proteasome activity in tumors to a lesser extent as compared to other tissues or whole blood samples[19,36,37] and similar observations have been reported for marizomib[38] and carfilzomib.[19] While the anti-tumor efficacy of bortezomib results in effective anti-tumor therapy, the improved response of CEP-18770 in tumors that can be correlated with its previously reported improved anti-tumor efficacy may result in further benefit in a clinical setting. In addition, the data obtained in mice *ex vivo* show that probe **1** can be successfully used for detailed studies of the effects of proteasome inhibitors on both normal and tumor tissues in preclinical models.

Finally, our data show that CEP-18770 displays activity against a bortezomib resistant cell line *in vitro*. The cell line THP1/BTZ_{100} is resistant to concentrations of bortezomib up to 100 nM. Study of the molecular mechanism of bortezomib resistance in these cells revealed an Ala49Thr mutation in the β5 subunit protein.[30] Ala49 is a highly conserved residue among many prokaryotic and mammalian species and very likely involved in tight interactions with the pyrazine-2-carboxy-phenylalanyl backbone of bortezomib, as can be reasoned from the bortezomib liganded yeast 20S proteasome structure.[39] Therefore, mutation of this residue to threonine is likely to have a profound effect on binding of bortezomib and of other proteasome inhibitors that target the β5 subunit. Our data show that CEP-18770 binding was not influenced by the Ala49Thr mutation. Bortezomib and

CEP-18770 are structurally similar, but differ in their P2 and P3 residues, suggesting that even small changes to the backbone structure of bortezomib can be sufficient to overcome bortezomib resistance. The data presented here suggest that the development of second-generation proteasome inhibitors will likely result in the generation of inhibitors that can overcome resistance and may ultimately contribute to better patient outcome and that this outcome can be predicted with the use of activity reporter reagents.

MATERIALS AND METHODS

Reagents

The THP1/BTZ_{100} cell line was kindly provided by Dr. G. Jansen (VU University Medical Centre, Amsterdam, The Netherlands). Fluorogenic substrates were purchased from Bachem (Weil am Rhein, Germany). All solvents were purchased from Biosolve (Valkenswaard, the Netherlands) at the highest grade available. Bortezomib and CEP-18770 were synthesized at Cephalon Inc for this study. All other chemicals were purchased from Sigma-Aldrich (Zwijndrecht, the Netherlands) at the highest available purity.

Probe synthesis and characterization

Detailed information on synthetic procedures and characterization can be found in the Supplementary Information and Supplementary Schemes 1 and 2 therein.

Cell culture

RPMI8226 (human, multiple myeloma), ARP-1 (human, multiple myeloma), THP1/WT (human monocytic/macrophage) and H929 (human, multiple myeloma) cells were cultured in RPMI 1640 medium (Invitrogen). THP1/BTZ_{100} cells were cultured in RPMI 1640 medium supplemented with 100 nM bortezomib. JJN3 cells were cultured in RPMI 1640 medium containing 10mM HEPES (Invitrogen). HeLa cells (human, cervix carcinoma) were cultured in DMEM medium (Invitrogen). All media were supplemented with 10% fetal calf serum (FCS), 100 units/mL penicillin and 100 μg/mL streptomycin.

Inhibition experiments in cell lysates

Cells were washed with phosphate-buffered saline (PBS), pelleted and lysed with one volume of glass beads (<106 microns, acid-washed; Sigma) and an equal volume of homogenization buffer (50 mM Tris pH 7.4, 1 mM dithiothreitol (DTT), 5 mM $MgCl_2$, 2 mM ATP and 250 mM sucrose) by vortexing at high speed for 45 minutes at 4 °C. Beads, membrane fractions, nuclei and cell debris were then removed from the supernatant by centrifugation at 14,000 g for 5 minutes. Protein contents were quantified using the Bradford method (BioRad, Veenendaal, the Netherlands) or with a nanodrop spectrophotometer. Equal amounts of protein were incubated with 1, 3, 10, 30 or 100 nM bortezomib or CEP-18770 for 1 h at 37 °C. Control samples were incubated with 1 μM epoxomicin. To assay proteasome activity, 10 μg of lysate was added to 100 μL of substrate buffer (20 mM Tris pH 7.4, 5 mM $MgCl_2$, 1 mM DTT, 1 mM ATP and 100 μM BzValGlyArg-AMC (tryptic activity), 100 μM SucLeuLeuValTyr-AMC (chymotryptic activity) or 50 μM ZLeuLeuGlu-AMC (caspase-like activity)). Fluorescence was measured every minute for 25 minutes at 37 °C using a fluostar optima 96 well plate reader (BMG labtechnologies) (λ ex/em = 355/450 nm) and the increase in fluorescence per minute was used to calculate specific activities of each sample. Non-specific activities were determined using 1 μM epoxomicin, which specifically inhibits all proteasomal activity at this concentration, and the background signal obtained was subtracted from each measurement. IC_{50} values were determined in three independent experiments and averaged. Data were analyzed by using GraphPad Prism software (GraphPad, La Jolla, CA, USA).

Probing proteasome inhibition in whole cells

H929, RPMI8226, ARP-1 or HeLa cells (0.5×10^6 cells/mL) were pulsed with 0, 1, 3, 10, 30 or 100 nM bortezomib or CEP-18770 for 1, 2 or 24 hours at 37 °C, followed by a 2 hour chase with 500 nM proteasome activity probe **1**. Alternatively, ARP-1 cells were pulsed with 100 nM bortezomib or CEP-18770 for 1 hour, washed, and either chased directly with 500 nM proteasome activity probe 1 or left to recover for 0.25, 0.5, 1, 2, 3, 4 or 20 hours before being chased with probe **1**. THP1/WT or THP1/BTZ_{100} cells (0.5×10^6 cells/mL) were pulsed with 0, 1, 3, 10, 30 or 100 nM bortezomib

or CEP-18770 for 1 or 22 hours at 37 °C, followed by a 2 hour chase with 200 nM proteasome activity probe **1**. Cells were harvested, washed with cold PBS and lysed for 30 min. in NP40 lysis buffer (50 mM Tris pH 7.4, 150 mM NaCl, 1% NP40) at 4 °C. The Bradford assay was used to measure protein content.

***Ex vivo* evaluations**

Nod/Scid mice with established human ARP-1 MM subcutaneous tumor xenografts (n=5/group) were dosed at Cephalon (West Chester, PA) with the maximum tolerated dose of bortezomib (1.2 mg/kg i.v.), CEP-18770 (4 mg/kg i.v.) or Solutol/PBS vehicle. Experiments were approved by the Institutional Animal Care and Use Committee of Cephalon. At 2 h, 8 h, and 24 h post dosing heart, liver, lung, spleen and tumor tissue were obtained, rinsed in cold PBS and flash frozen in a tissue cassette under liquid N_2. Organs were lysed in homogenization buffer using glass beads as described above. Protein concentrations were determined using the Bradford assay, and equal amounts of protein were incubated with 1 μM probe **1** for 1 h at 37 °C. Statistical analyses were performed using by using GraphPad Prism software (GraphPad, La Jolla, CA, USA).

In-gel fluorescence measurements

Equal amounts of protein were denatured by boiling in LDS (Lithium dodecyl sulfate) sample buffer (Invitrogen) containing 2.5% β-mercaptoethanol and polypeptides were resolved by 12% SDS-PAGE using the NuPAGE system from invitrogen. Wet gel slabs were imaged for 2 min, with a resolution of 100 μm, using the ProXPRESS 2D Proteomic imaging system (Perkin Elmer), using appropriate filter settings (λ ex/em = 480/530 nm). To verify protein loading, gels were stained with ruthenium II tris bathophenanthroline disulfonate, as described,[40,41] and imaged using appropriate filter settings (λ ex/em = 460/680 nm). Probe signals were then normalized to the sypro signal. Images were analyzed using Totallab analysis software (Nonlinear Dynamics, Newcastle upon Tyne, UK) to quantify the intensity of the bands detected.

Cell viability assay

THP1/WT or THP1/BTZ_{100} cells (2×10^5 cells/mL were incubated with 0, 1, 3, 10, 30 or 100 nM bortezomib or CEP-18770 for 24 or 48 hours, followed by incubation with 20 μg/mL resazurin. The increase in fluorescence was measured after 8 hours using a PE envision plate reader. All results were expressed as percentage relative to untreated THP1/WT cells (100%).

ACKNOWLEDGMENTS

The authors thank Dr. G. Jansen for providing the THP1/BTZ_{100} cell line. This study was supported by the Dutch Cancer Society (grant number NKI 2005-3368) and by research funding from Cephalon to H.O.

REFERENCES

1. Adams, J. The development of proteasome inhibitors as anticancer drugs. *Cancer Cell.* **5**, 417-421 (2004).
2. Nalepa, G., Rolfe, M. & Harper, J.W. Drug discovery in the ubiquitin-proteasome system. *Nat. Rev. Drug Discov.* **5**, 596-613 (2006).
3. Dahlmann, B., Ruppert, T., Kuehn, L., Merforth, S. & Kloetzel, P.M. Different proteasome subtypes in a single tissue exhibit different enzymatic properties. *J. Mol. Biol.* **303**, 643-653 (2000).
4. Dahlmann, B., Ruppert, T., Kloetzel, P.M. & Kuehn, L. Subtypes of 20S proteasomes from skeletal muscle. *Biochimie.* **83**, 295-299 (2001).
5. Pelletier, S. et al. Quantifying cross-tissue diversity in proteasome complexes by mass spectrometry. *Mol. Biosyst.* (2010).
6. Raijmakers, R. et al. Automated online sequential isotope labeling for protein quantitation applied to proteasome tissue specific diversity. *Mol. Cell. Proteomics.* **7**, 1755-1762 (2008).
7. Adams, J. The proteasome: a suitable antineoplastic target. *Nat. Rev. Cancer.* **4**, 349-360 (2004).
8. Hideshima, T. et al. The proteasome inhibitor PS-341 inhibits growth, induces apoptosis, and overcomes drug resistance in human multiple myeloma cells. *Cancer Res.* **61**, 3071-3076 (2001).
9. Jagannath, S. et al. A phase 2 study of two doses of bortezomib in relapsed or refractory myeloma. *Br. J. Haematol.* **127**, 165-172 (2004).
10. Richardson, P.G. et al. A phase 2 study of bortezomib in relapsed, refractory myeloma. *N. Engl. J. Med.* **348**, 2609-2617 (2003).
11. Richardson, P.G. et al. Bortezomib or high-dose dexamethasone for relapsed multiple myeloma. *N. Engl. J. Med.* **352**, 2487-2498 (2005).
12. Fisher, R.I. et al. Multicenter phase II study of bortezomib in patients with relapsed or refractory

mantle cell lymphoma. *J. Clin. Oncol.* **24**, 4867-4874 (2006).
13. Orlowski, R.Z. et al. Phase 1 trial of the proteasome inhibitor bortezomib and pegylated liposomal doxorubicin in patients with advanced hematologic malignancies. *Blood.* **105**, 3058-3065 (2005).
14. Orlowski, R.Z. et al. Randomized phase III study of pegylated liposomal doxorubicin plus bortezomib compared with bortezomib alone in relapsed or refractory multiple myeloma: combination therapy improves time to progression. *J. Clin. Oncol.* **25**, 3892-3901 (2007).
15. Concannon, C.G. et al. Apoptosis induced by proteasome inhibition in cancer cells: predominant role of the p53/PUMA pathway. *Oncogene.* **26**, 1681-1692 (2007).
16. Richardson, P.G., Mitsiades, C., Hideshima, T. & Anderson, K.C. Bortezomib: proteasome inhibition as an effective anticancer therapy. *Annu. Rev. Med.* **57**, 33-47 (2006).
17. Nencioni, A., Grunebach, F., Patrone, F., Ballestrero, A. & Brossart, P. Proteasome inhibitors: antitumor effects and beyond. *Leukemia.* **21**, 30-36 (2007).
18. Orlowski, R.Z. & Kuhn, D.J. Proteasome inhibitors in cancer therapy: lessons from the first decade. *Clin. Cancer Res.* **14**, 1649-1657 (2008).
19. Demo, S.D. et al. Antitumor activity of PR-171, a novel irreversible inhibitor of the proteasome. *Cancer Res.* **67**, 6383-6391 (2007).
20. Kuhn, D.J. et al. Potent activity of carfilzomib, a novel, irreversible inhibitor of the ubiquitin-proteasome pathway, against preclinical models of multiple myeloma. *Blood.* **110**, 3281-3290 (2007).
21. Stapnes, C. et al. The proteasome inhibitors bortezomib and PR-171 have antiproliferative and proapoptotic effects on primary human acute myeloid leukaemia cells. *Br. J. Haematol.* **136**, 814-828 (2007).
22. Chauhan, D. et al. A novel orally active proteasome inhibitor induces apoptosis in multiple myeloma cells with mechanisms distinct from Bortezomib. *Cancer Cell.* **8**, 407-419 (2005).
23. Fenical, W. et al. Discovery and development of the anticancer agent salinosporamide A (NPI-0052). *Bioorg. Med. Chem.* **17**, 2175-2180 (2009).
24. Dorsey, B.D. et al. Discovery of a potent, selective, and orally active proteasome inhibitor for the treatment of cancer. *J. Med. Chem.* **51**, 1068-1072 (2008).
25. Piva, R. et al. CEP-18770: A novel, orally active proteasome inhibitor with a tumor-selective pharmacologic profile competitive with bortezomib. *Blood.* **111**, 2765-2775 (2008).
26. Sanchez, E. et al. The Novel Proteasome Inhibitor CEP-18770 Inhibits Myeloma Tumor Growth In Vitro and In Vivo and Enhances the Anti-MM Effects of Melphalan. *ASH Annual Meeting Abstracts.* **112**, 843 (2008).
27. Berkers, C.R. et al. Profiling proteasome activity in tissue with fluorescent probes. *Mol. Pharm.* **4**, 739-748 (2007).
28. Berkers, C.R. et al. Activity probe for in vivo profiling of the specificity of proteasome inhibitor bortezomib. *Nat. Methods.* **2**, 357-362 (2005).
29. Altun, M. et al. Effects of PS-341 on the activity and composition of proteasomes in multiple myeloma cells. *Cancer Res.* **65**, 7896-7901 (2005).
30. Oerlemans, R. et al. Molecular basis of bortezomib resistance: proteasome subunit beta5 (PSMB5) gene mutation and overexpression of PSMB5 protein. *Blood.* **112**, 2489-2499 (2008).
31. Lu, S. et al. Point mutation of the proteasome beta5 subunit gene is an important mechanism of bortezomib resistance in bortezomib-selected variants of Jurkat T cell lymphoblastic lymphoma/leukemia line. *J. Pharmacol. Exp. Ther.* **326**, 423-431 (2008).
32. Lu, S. et al. Different mutants of PSMB5 confer varying bortezomib resistance in T lymphoblastic lymphoma/leukemia cells derived from the Jurkat cell line. *Exp. Hematol.* **37**, 831-837 (2009).
33. Lonial, S. et al. Risk factors and kinetics of thrombocytopenia associated with bortezomib for relapsed, refractory multiple myeloma. *Blood.* **106**, 3777-3784 (2005).
34. Argyriou, A.A., Iconomou, G. & Kalofonos, H.P. Bortezomib-induced peripheral neuropathy in multiple myeloma: a comprehensive review of the literature. *Blood.* **112**, 1593-1599 (2008).
35. Fuchs, D., Berges, C., Opelz, G., Daniel, V. & Naujokat, C. Increased expression and altered subunit composition of proteasomes induced by continuous proteasome inhibition establish apoptosis resistance and hyperproliferation of Burkitt lymphoma cells. *J. Cell. Biochem.* **103**, 270-283 (2008).
36. Adams, J. et al. Proteasome inhibitors: a novel class of potent and effective antitumor agents. *Cancer Res.* **59**, 2615-2622 (1999).
37. LeBlanc, R. et al. Proteasome inhibitor PS-341 inhibits human myeloma cell growth in vivo and prolongs survival in a murine model. *Cancer Res.* **62**, 4996-5000 (2002).
38. Williamson, M.J. et al. Comparison of biochemical and biological effects of ML858 (salinosporamide A) and bortezomib. *Mol. Cancer Ther.* **5**, 3052-3061 (2006).
39. Groll, M., Berkers, C.R., Ploegh, H.L. & Ovaa, H. Crystal structure of the boronic acid-based proteasome inhibitor bortezomib in complex with the yeast 20S proteasome. *Structure.* **14**, 451-456 (2006).
40. Lamanda, A., Zahn, A., Roder, D. & Langen, H. Improved Ruthenium II tris (bathophenantroline disulfonate) staining and destaining protocol for a better signal-to-background ratio and improved baseline resolution. *Proteomics.* **4**, 599-608

(2004).
41. Rabilloud, T., Strub, J.M., Luche, S., van Dorsselaer, A. & Lunardi, J. A comparison between Sypro Ruby and ruthenium II tris (bathophenanthroline disulfonate) as fluorescent stains for protein detection in gels. *Proteomics.* **1**, 699-704 (2001).

Supplementary information

SUPPLEMENTARY METHODS

Fmoc peptide building blocks were from Novabiochem. All solvents were purchased from Biosolve at the highest grade available. All other chemicals were purchased from Aldrich at the highest available purity. All solvents and chemicals were used as received. L-Leucinyl vinyl sulfone was prepared as reported.[1-4] Nuclear magnetic resonance (NMR) spectra were recorded in the indicated solvent using a Bruker Avance 300 (^{1}H: 300 MHz; ^{13}C: 75 MHz) spectrometer. LC-MS analyses were carried out on a WATERS LCT mass spectrometer in line with a WATERS 2795 HPLC system and a WATERS 2996 photodiode array detector. Reversed phase runs were performed on a 3 μm Atlantis T3, C18 RP, 2.1 x 100 mm column (Waters) using gradient elution with H_2O/0.1% formic acid as solvent A and acetonitrile/0.1% formic acid as solvent B at a flow rate of 0.4 mL/min. Reversed phase HPLC runs were performed on a Waters 1525 EF HPLC system in line with a Waters 2487 dual λ absorbance detector using gradient elution with H_2O/0.05% TFA as solvent A and acetonitrile/0.05% TFA as solvent B. Analytical runs were performed on a 10 μm 4.6x150 mm Atlantis dC18 column (Waters) at a flow rate of 1.4 ml/min and preparative runs were performed on a 10 μm 19x250 mm Atlantis dC18 column (Waters) at a flow rate of 18 mL/min.

Synthesis of Me_4BodipyFL-NHS ester (Supplementary Scheme 1)

I: (Z)-tert-butyl 2-(hydroxyimino)-3-oxobutanoate (1)

A solution of sodium nitrite (1.4 g, 20.3 mmol) in deionized water was added dropwise over 5 minutes to a stirred mixture of *tert*-butyl acetoacetate (2.80 g, 17.4 mmol) in acetic acid (4 mL), while keeping the reaction vessel in an ice-bath. The reaction mixture was stirred at 4 °C for an additional 16 hours.

II: tert-butyl 4-(3-methoxy-3-oxopropyl)-3,5-dimethyl-1H-pyrrole-2-carboxylate (2)

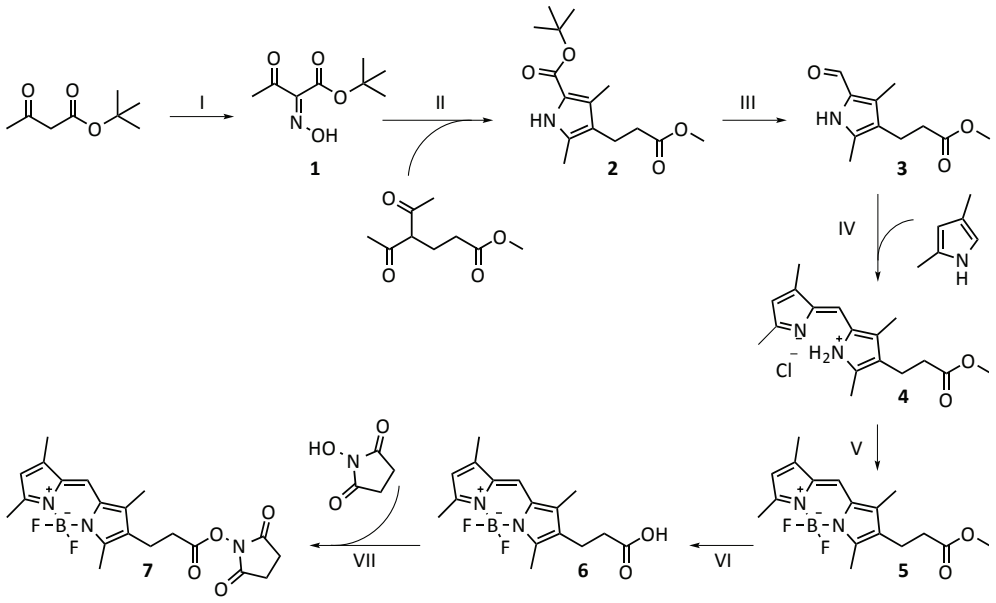
The resulting solution containing oxime **1** was slowly added in 5 portions to a solution of methyl 4-acetyl-5-oxohexanoate (5 mL, 28.6 mmol) in acetic acid (15 mL) kept at 65 °C. Simultaneously, a mixture of zinc dust (3 g, 45.9 mmol) and sodium acetate (3 g, 36.6 mmol) was added in portions such that zinc/sodium acetate was in excess of dione. After stirring the mixture for 2 more hours at 65 °C it was poured into ice water and left slowly stirring for 16 hours, allowing the mixture to warm up to room temperature. The precipitate, including product and zinc residues was filtered off using a glass filter. The residue was taken up in ethyl acetate (100 mL) and the organic layer was dried with magnesium sulfate ($MgSO_4$). The $MgSO_4$ was filtered off and the filtrate was concentrated *in vacuo* to dryness, yielding tetrasubstituted pyrrole **2** as a light brown solid (2.87 g, 10.2 mmol, 59 % yield), which was used without further purification in the next step.

III: Methyl 3-(5-formyl-2,4-dimethyl-1H-pyrrol-3-yl)propanoate (3)

Pyrrole **2** (1 g, 3.55 mmol) was dissolved in TFA (20 mL) and stirred for 20 minutes on ice. Triethylorthoformate (2 mL, 12 mmol) was added slowly and the solution was stirred gently at 0 °C for 15 minutes. Water (2 mL) was added and the solution was stirred for 10 minutes at 0 °C. The solution was coevaporated with toluene *in vacuo* to dryness at room temperature. The crude product was purified by flash column chromatography using ethyl acetate as the eluent to yield aldehyde **3** as a brown solid (469 mg, 2.25 mmol, 63 % yield).

IV: (Z)-2-((3,5-dimethyl-2H-pyrrol-2-ylidene) methyl)-4-(3-methoxy-3-oxopropyl)-3,5-dimethyl-1H-pyrrol-1-ium (4)

Aldehyde **3** (469 mg, 2.25 mmol) was dissolved in absolute ethanol (8 mL) and cooled to 0 °C. 2,4-dimethylpyrrole (256 mg, 2.69 mmol) was added under an argon atmosphere and 4 mL of cold 4N HCl in dioxane was added. After Stirring for 15 minutes at 0 °C, the solution was transferred to a 50 mL Falcon tube and centrifuged at 1500g for 10 minutes at 4 °C. The supernatant was discarded and the resulting orange solid was



Supplementary Scheme 1 | Synthesis of Me_4BodipyFL-NHS ester 7. Step numbers correspond to step numbers outlined in the text.

resuspended in cold diethylether (40 mL) and centrifuged again. The supernatant was decanted and the residue was dried under a stream of nitrogen, yielding crude dipyrrole **4** as an orange solid, which was used in the next step without further purification.

V: Me_4BodipyFL-methyl ester (5)

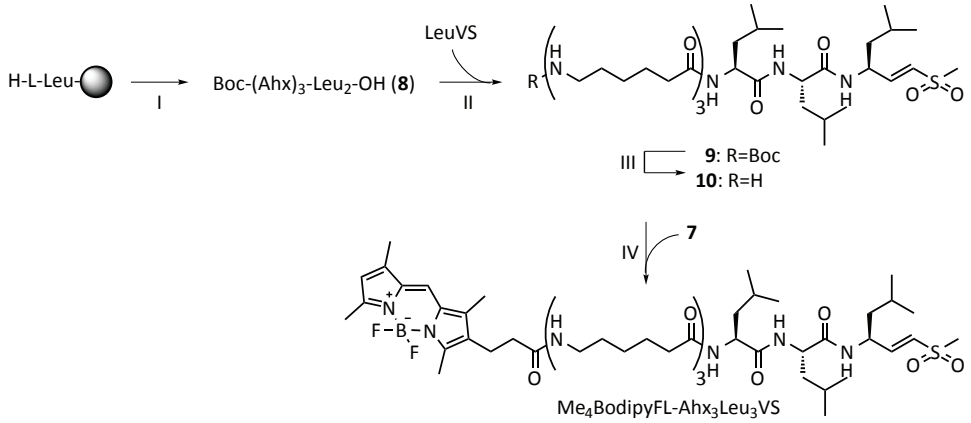
Dipyrrole **4** was suspended in 1,2-dichlorobenzene (20 mL) under an argon atmosphere and triethylamine (1 mL, 7.16 mmol) and boron trifluoride diethyl etherate (1 mL, 8.12 mmol) were added in concert. The resulting metallic solution was stirred for 30 minutes at 100 °C and allowed to cool to room temperature. The solution was diluted with dry hexanes (25 mL) and applied directly onto a silica gel column for purification by flash column chromatography using 25% EtOAc in toluene as the eluent to afford Me_4BodipyFL-methyl ester **5** as an orange solid (481 mg, 1.44 mmol, 64 % yield starting from aldehyde **3**).

VI: Me_4BodipyFL-OH (6)

Me_4BodipyFL-methyl ester **5** (400 mg, 1.24 mmol) was dissolved in ethanol (6 mL) and 1 M lithium hydroxide (LiOH) solution was added (1.3 mL). The solution was stirred until TLC (eluent: EtOAc containing 0.1% acetic acid) revealed complete conversion. Ethyl acetate (40 mL) and 0.1 M HCl (30 mL) were added and the two layers were separated. The organic layer was dried with $MgSO_4$ and concentrated *in vacuo* to dryness to obtain an orange solid. Purification by flash column chromatograpy using EtOAc containing 0.1% acetic acid as the eluent afforded Me_4BodipyFL-OH **6** as an orange solid (306 mg, 0.95 mmol, 77 % yield). ^{1}H NMR ($CDCl_3$): δ 7.09 (s, 1H), 6.10 (s, 1H), 2.66 (dd, *J*=8.2 Hz, *J*=7.1 Hz, 2H), 2.59 (s, 6H), 2.60-2.53 (m, 2H), 2.31 (s, 3H), 2.27 (s, 3H). ^{13}C NMR ($CDCl_3$): δ 178.2 (CO), 156.5 (C_q), 155.3 (C_q), 141.0 (C_q), 138.1 (C_q), 133.3 (C_q), 132.6 (C_q), 127.8 (C_q), 119.7 (CH), 118.9 (CH), 33.9 (CH_2), 19.2 (CH_2), 14.6 (CH_3), 12.7 (CH_3), 11.2 (CH_3), 9.6 (CH_3).

VII: Me_4BodipyFL-NHS ester (7)

Me_4BodipyFL-OH **6** (306 mg, 0.95 mmol) was dissolved in dichloromethane (60 mL) and N-hydroxy succinimide (120 mg, 1.05 mmol), EDCI (201 mg, 1.05 mmol) and DMAP (11 mg, 0.1 mmol) were added. The mixture was stirred for



Supplementary Scheme 2 | Synthesis of Me_4BodipyFL-Ahx_3Leu_3VS. Step numbers correspond to step numbers outlined in the text.

16 hours at room temperature, 0.05 M HCl (40 mL) was added the two layers separated. The organic layer was dried with $MgSO_4$ and concentrated *in vacuo* to dryness to obtain an orange solid, which was purified by flash column chromatography using EtOAc : hexanes (2:3 v/v) as the eluent to obtain Me_4BodipyFL-NHS ester **7** as a dark orange solid (242 mg, 0.85 mmol) in 61 % yield. ^{1}H NMR ($CDCl_3$): δ 7.02 (s, 1H), 6.02 (s, 1H), 2.85-2.78 (m, 6H), 2.75-2.70 (m, 2H), 2.50 (s, 6H), 2.22 (s, 3H), 2.19 (s, 3H). ^{13}C NMR ($CDCl_3$): δ 169.0 (2xCO), 167.7 (CO), 156.9 (C_q), 154.8 (C_q), 141.3 (C_q), 138.0 (C_q), 133.4 (C_q), 132.5 (C_q), 126.7 (C_q), 119.9 (CH), 119.0 (CH), 31.1 (CH_2), 25.6 (2xCH_2), 19.2 (CH_2), 14.6 (CH_3), 12.6 (CH_3), 11.2 (CH_3), 9.6 (CH_3).

Synthesis of Me_4BodipyFL-Ahx_3Leu_3VS (Supplementary Scheme 2)

The synthesis of H_2N-Ahx_3Leu_2-OH involving standard Fmoc-based solid-phase peptide synthesis (SPPS) protocols and the subsequent coupling of fluorescent dye and vinyl sulfone electrophile to yield the final proteasome probe are depicted in Scheme 2.

I: Boc-Ahx_3Leu_2OH (8)

Hyper acid-labile H-Leu-2-ClTrt resin (novabiochem) (1 g, 0.3 mmol equivalents) was subjected to four coupling cycles, in which deprotection of the Fmoc-group with piperidine/NMP (1:4 v/v; 10 mL/g of dry resin), was followed by a coupling cycle with 3 equivalents of Fmoc-protected amino acid, 3 equivalents of DIPEA and 3 equivalents of PyBop coupling reagent in NMP (10 mL/g of dry resin). Coupling steps were performed with Fmoc-Leu-OH (1x), Fmoc-6-aminohexanoic acid (Fmoc-Ahx-OH) (2x) and Boc-Ahx-OH (1x) sequentially. The peptide was cleaved from the resin using 2% TFA in dichloromethane and the compound was concentrated under reduced pressure. The product was purified by flash column chromatography using a gradient of 10 % to 20% MeOH in CH_2Cl_2 containing 0.1 % acetic acid as the eluent to obtain Boc-Ahx_3Leu_2OH **8** as a white solid (205 mg, 0.3 mmol) in > 80 % yield.

II: Boc-Ahx_3Leu_3VS (9)

L-Leucinyl vinyl sulfone (LeuVS, 50 mg, 0.26 mmol) was dissolved in DMF and Boc-Ahx_3Leu_2OH **8** (116 mg, 0.17 mmol), DIPEA (89 μL, 0.51 mmol) and EDCI (50 mg, 0.26 mmol) were added. The solution was stirred for 3 hours at room temperature and then concentrated *in vacuo* to dryness to yield an off-white solid. The product was purified by flash column chromatography using 10 % MeOH in CH_2Cl_2 as the eluent to obtain Boc-Ahx_3Leu_3VS **9** as a white solid (103 mg, 0.12 mmol) in 71 % yield.

III: H_2N-Ahx_3Leu_3VS (10)

Boc-Ahx_3Leu_3VS **9** (12 mg, 14 μmol) was dissolved in TFA (1 mL) and left at room temperature for 30 minutes. The product was precipitated by adding cold diethylether (40 mL) and isolated by centrifugation at 1000g (low brake speed) for 5 minutes at 4 ˚C to afford H_2N-Ahx_3Leu_3VS **10** as a white solid (10.6 mg, 14 μmol) in > 98 % yield.

IV: Me_4BodipyFL-Ahx_3Leu_3VS

H_2N-Ahx_3Leu_3VS **10** (10.6 mg, 14 μmol) was dissolved in DMF (1 mL) and Me_4BodipyFL-NHS ester **7** (6.4 mg, 15.5 μmol) and DIPEA (24 μL, 0.14 mmol) were added. The resulting solution was stirred at room temperature for 16 hours. The solution was concentrated *in vacuo* to dryness and purify by reversed phase HPLC to obtain the fluorescent proteasome probe Me_4BodipyFL-Ahx_3Leu_3VS (6 mg, 5.7 μmol) in 40% yield. MS (ESI): 1058.66 $(M+H)^+$; 520.27 $(M+H-F)^{2+}$; 1080.91 $(M+Na)^+$

SUPPLEMENTARY REFERENCES

1. Bogyo, M. et al. 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 (1997).
2. Palmer, J.T., Rasnick, D., Klaus, J.L. & Bromme, D. Vinyl sulfones as mechanism-based cysteine protease inhibitors. *J. Med. Chem.* **38**, 3193-3196 (1995).
3. Rydzewski, R. et al. Optimization of Subsite Binding to the â5 Subunit of the Human 20S Proteasome Using Vinyl Sulfones and 2-Keto-1,3,4-oxadiazoles: Syntheses and Cellular Properties of Potent, Selective Proteasome Inhibitors. *J. Med. Chem.* **49**, 2953-2968 (2006).
4. Wang, G., Mahesh, U., Chen, G.Y. & Yao, S.Q. Solid-phase synthesis of peptide vinyl sulfones as potential inhibitors and activity-based probes of cysteine proteases. *Org. Lett.* **5**, 737-740 (2003).

Chapter 2.2

Profiling patient response to marizomib therapy

Profiling patient response to marizomib therapy

Celia R. Berkers,[a] Karianne G. Schuurman,[a] Saskia K.C. Neuteboom,[b] Dharminder Chauhan,[c] Kenneth C. Anderson,[c] Michael A. Palladino[b] & Huib Ovaa[a,*]

[a]Division of Cell Biology II, The Netherlands Cancer Institute, Amsterdam, The Netherlands. [b]Nereus Pharmaceuticals, San Diego, CA, USA. [c]The Jerome Lipper Multiple Myeloma Center, Department of Medical Oncology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA, USA.

Marizomib (NPI-0052; Salinosporamide A) is a irreversible proteasome inhibitor isolated from the marine actinomycete *Salinispora tropica* that has demonstrated activity in a wide range of preclinical tumor models. Clinical trials with marizomib are currently ongoing, enabling a detailed study of patient response to marizomib treatment. In the present study, we used a fluorescent proteasome activity probe to profile proteasome inhibition in both packed whole blood (PWB) and peripheral blood mononuclear cells (PBMCs) from patients on marizomib therapy. We show that PWB, which contains >98% red blood cells in which immunoproteasome is not expressed, responds differently to proteasome inhibitors and recovers slower from proteasome inhibition compared to PBMCs. In addition, we demonstrate that proteasome activity levels prior to treatment differ between patients, and that this initial proteasome activity level is an important determinant of the sensitivity of cells to proteasome inhibitors. Furthermore, we show that PBMCs are able to rapidly change their proteasomal subunit composition in response to marizomib treatment, which seems to account for a quick recovery, suggesting that this ability may mediate sensitivity of tumor cells to proteasome inhibition. Together, these data indicate that fluorescent proteasome probes can be used to profile patient blood samples and that they may enable prediction of patient responses to proteasome inhibitors, which can contribute to a more personalized proteasome inhibition strategy to improve the therapeutic potential of this class of drugs and finally clinical outcome of proteasome targeted therapy.

INTRODUCTION

Proteasome inhibition has emerged as a powerful approach for the treatment of cancer, as evidenced by the clinical success of the proteasome inhibitor bortezomib (Figure 1).[1] Bortezomib has been approved for the treatment of relapsed and newly diagnosed multiple myeloma (MM)[2-4] and mantle cell lymphoma[5] and in combination with PEGylated doxorubicin for the treatment of relapsed MM.[6,7] Bortezomib treatment is associated with substantial side effects, and the occurrence of both primary and secondary resistance has been reported.[8] Consequently, there is an ongoing search for novel proteasome inhibitors that are able to overcome bortezomib resistance and that are better tolerated to increase their therapeutic potential. Several second-generation proteasome inhibitors have entered clinical trials, including marizomib (NPI-0052; Salinosporamide A)[9] (Figure 1), CEP-18770,[10]

*Correspondence should be adressed to H.O. (h.ovaa@nki.nl).

and Carfilzomib (PR-171).[11]

The proteasome is a large threonine protease complex responsible for the degradation of redundant and misfolded proteins and for the turnover of regulatory proteins that control a wide range of cellular processes, including cell cycle regulation, apoptosis and proliferation.[12,13] Mammalian 26S proteasomes consist of a 20S catalytic core, complexed at one or both ends with 19S regulatory caps. The 20S core is composed of four stacked rings of seven subunits each and has an overall architecture of $\alpha(1\text{-}7)\beta(1\text{-}7)\beta(1\text{-}7)\alpha(1\text{-}7)$. The two outer α-rings stabilize the complex and serve as docking stations for the binding of regulatory caps, while the catalytic activity resides within the two inner β-rings. Three constitutive β-subunits, termed $\beta1$, $\beta2$ and $\beta5$, provide the proteasomal caspase-like, tryptic and chymotryptic activity, respectively. In lymphoid tissues and in cells exposed to INF-γ and TNF-α, these subunits are replaced by their immunoproteasomal counterparts $\beta1i$, $\beta2i$ and $\beta5i$ to form the immunoproteasome,[14] which is thought to favor the production of antigenic peptides,[15] while the existence of mixed-type hybrid proteasomes has also been reported.[16,17]

Marizomib is an orally available proteasome inhibitor isolated from the marine actinomycete *Salinispora tropica*.[9] Marizomib irreversibly inhibits the proteasome via a two-step mechanism. In an initial step an ester is formed between the catalytic Thr1 O^{γ} of the proteasome and the carbonyl derived from the β-lactone ring. Subsequent chlorine substitution then gives rise to a cyclic ether.[18] In contrast to bortezomib which targets the $\beta5$ and $\beta1$ subunits, marizomib has been shown to inhibit all constitutive and immunoproteasome subunits.[9] Recent studies have shown that inhibition of multiple activities is a prerequisite for a significant inhibition of protein degradation, depending on the protein that is degraded,[19] suggesting that blocking all three activities may be therapeutically advantageous. Marizomib showed efficacy as a single agent or in combination with conventional chemotherapy in a variety of preclinical solid tumor and hematologic tumor models[20-22] and was able to overcome bortezomib resistance in tumor cells purified from MM patients that had relapsed from prior therapies including bortezomib treatment.[9] Interestingly, combination treatment with bortezomib and marizomib both *in vitro* and *in vivo* induced synergistic cytotoxicity in MM and Waldenstrom's macroglobulinemia.[23,24] Marizomib is currently being evaluated in phase I clinical trials for the treatment of both solid and hematologic tumors.

Proteasome activity is traditionally measured using fluorogenic substrates for the different catalytic activities.[25] The use of recently developed chemical proteasome activity reporters provides an alternative method for assaying proteasome activity in a wide range of tissues.[26-28] In this study, we used a fluorescent proteasome activity reporter[27] to compare the inhibitory profiles of marizomib and bortezomib in rat blood samples, and established that this probe can be used to accurately profile both packed whole blood (PWB) and peripheral blood mononuclear cell (PBMC) samples. In addition, we developed a protocol that enables the use of this fluorescent probe to profile blood samples from patients that are treated with proteasome inhibitors. The data presented here give for the first time insight into patient response to and recovery from marizomib treatment and show that both the initial proteasome activity in cells and the ability of cells to change the composition of newly synthesized proteasomes may mediate sensitivity to this novel class of drugs. In addition, we show that immunoproteasome is not expressed in PWB, which contains >98% red blood cells. As red

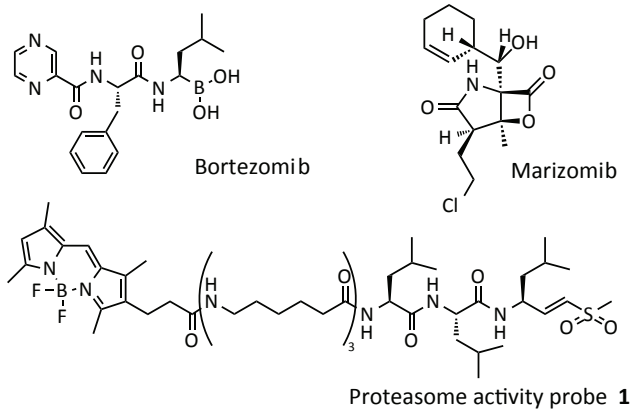


Figure 1 | Structures of bortezomib, marizomib, and proteasome activity probe 1.

blood cells have a relatively high, micromolar intracellular proteasome concentration,[25] they are thought to sequester intravenously administered proteasome inhibitors *in vivo*.[29] Consequently, proteasome inhibitors likely saturate proteasomes in red blood cells before peripheral sites can be reached. Specific immunoproteasome inhibitors, however, may saturate blood proteasome at relatively low concentrations, which may result in higher inhibitor concentrations in tumor cells. This suggests that immunoproteasome inhibitors may in future provide a good alternative for tumors that have high intrinsic immunoproteasome activity.

RESULTS

Profiling marizomib and bortezomib in rat blood samples *ex vivo*

To profile the effects of proteasome inhibitors on blood proteasome in both rats and patients *ex vivo*, fluorescent proteasome activity reporter **1** (Me_4BodipyFL-Ahx$_3$Leu$_3$VS, described in chapter 2.1) was used.[27] Probe **1** consists of a proteasome-targeting motif, a vinyl sulfone reactive group that covalently reacts with the N-terminal threonine of all active subunits and a fluorescent tag (Figure 1). This tag allows the use of an SDS-PAGE based assay to visualize proteasome (subunit) labeling by scanning the gel for fluorescence emission. Prior inhibition of proteasome subunits prevents probe binding, which results in the disappearance of fluorescent signal. Therefore, the measured fluorescence intensity directly correlates to the activity of a labeled β-subunit. To study whether proteasome inhibition could be reliably measured in blood samples using probe **1**, the inhibition and recovery patterns of bortezomib and marizomib were evaluated in rat blood samples. Rats (n=15/group) were treated with vehicle, 0.15 mg/kg bortezomib or 0.05 or 0.10 mg/kg marizomib (which will be referred to as marizomib low and marizomib high treatment, respectively). Blood samples were taken from 3 rats per dosing group per time point (1.5 h, 24 h, 48 h 72 h or 168 h after injection), packed whole blood (PWB) or PBMC lysates were obtained, and lysates were incubated with probe **1**. Proteins were then separated on SDS-PAGE, and the resulting gel was scanned for fluorescence emission. Figure 2A shows representative gel images of labeled PWB and PBMC samples.

Red blood cells (RBCs), which do not express immunoproteasome, form the majority of

PWB (>99%), and as a result mainly constitutive proteasome subunits were labeled in PWB samples (Figure 2A). In PBMCs, which consist mainly of white blood cells, both the constitutive β1, β2, and β5 subunits and the immunoproteasomal β2i and β1i subunits were clearly labeled (Figure 2A). In both PBMC and PWB samples, no substantial differences were found between untreated and vehicle treated mice. The contribution of each subunit to total activity was constant, indicating that vehicle treatment did not influence proteasome activity (Supplementary Figure 1). Total proteasome activity differed somewhat between individual mice within treatment groups (data not shown). Upon treatment with marizomib or bortezomib, both total proteasome activity (Figure 2A; lanes 7-15) and the relative subunit labeling (Supplementary Figure 1) changed, indicating inhibition.

Marizomib and bortezomib have differential effects on PBMCs but not on PWB

To evaluate the effects of marizomib and bortezomib treatment on total proteasome activity, total fluorescence intensity in each sample was quantified, averaged per group, and plotted as percentage of untreated samples ± SEM (Figure 2B). In PBMCs a single dose of 0.05 mg/kg (low) or 0.10 mg/kg (high) marizomib reduced total proteasome activity to 20-30% residual activity, whereas a single dose of bortezomib (0.15 mg/kg) reduced total proteasome activity to 60%. Differences between marizomib high and bortezomib treatment remained significant up to 72 h after dosing, while marizomib low and bortezomib treatment had similar effects at time points 24 h to 168 h after dosing. Total proteasome activity recovered fully within one week for all treatment schedules. In PWB, all treatment regimes inhibited total proteasome activity almost completely. Recovery was hardly evident at time points up to 48 h, and one week after dosing (168 h) only 50% of proteasome activity had recovered. Rats treated with a high dose of marizomib showed somewhat less recovery of total proteasome activity at 72 h and 168 h post dosing compared to bortezomib treated rats, but differences were not significant.

To study how individual subunits contributed to total proteasome activity, the fluorescence intensity of individual subunits was quantified, averaged per group, and plotted as the percentage of untreated samples ± SEM (Figure 2C). Since the β2 and β2i subunits responded similarly to all treatments (data not shown), β2 activity was defined as the sum of both bands. In PBMCs, the combined β5(i)/β1 activity was inhibited to a higher extent by marizomib high dose treatment compared to bortezomib treatment, while the β1i subunit was inhibited to a similar extent by both treatment regimes. The β2 subunits were inhibited by marizomib, but not by bortezomib. These data suggest that the more potent and sustained inhibition of proteasome activity in marizomib high dose-treated rats compared to bortezomib treated rats is mediated by marizomib's ability to inhibit all proteasome activities. When bortezomib treatment was compared to marizomib low dose treatment schedules, inhibition of the β5(i)/β1 subunits was comparable, but differences were observed in inhibition of both the β2 and β1i subunits. Inhibition of β2 activities was found only in marizomib low treated rats, but not in bortezomib treated rats, as also previously reported.[9] Inhibition of β1i activity was much more pronounced in bortezomib treated rats which recovered slowly, resulting in <50% recovery one week post-dosing. Together, these data suggest that the long-term effect of bortezomib and marizomib treatment on total proteasome activity was mediated via the inhibition of different sets of subunits. In PWB, differences between treatment regimes were

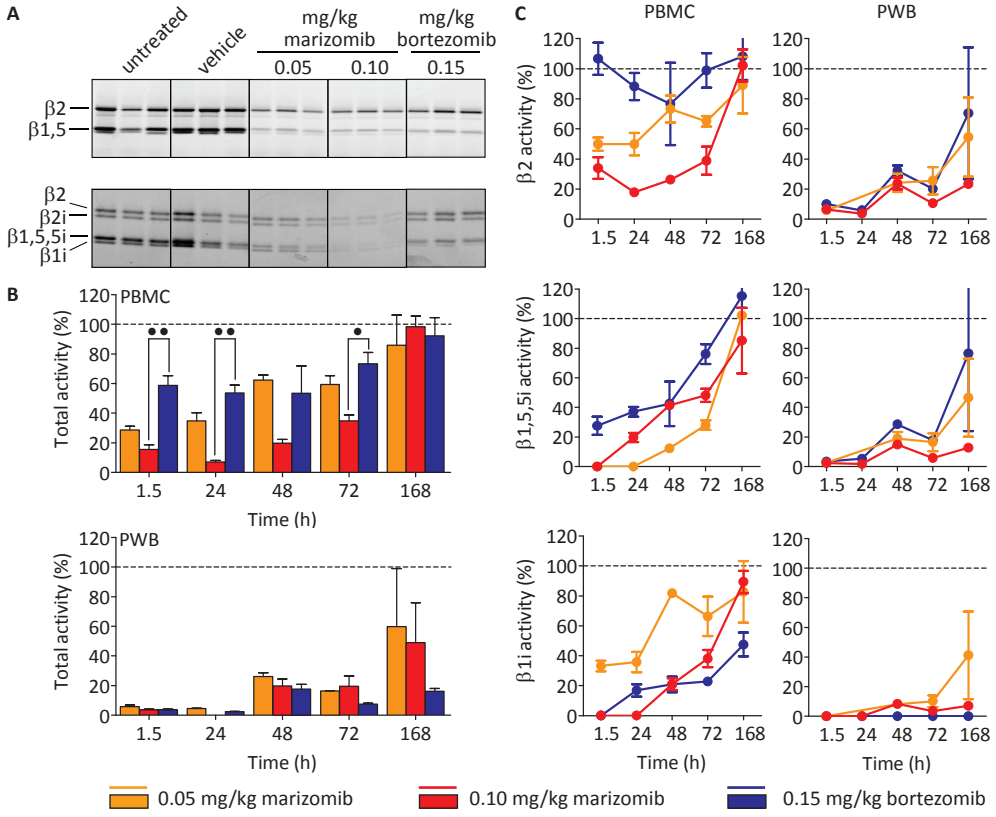


Figure 2 | Marizomib and bortezomib have different inhibition profiles in rat PBMC and PWB samples. Rats (n=15/group; 3/time point per dosing group) received a single injection of 0.05 or 0.10 mg/kg marizomib, 0.15 mg/kg bortezomib or vehicle. PWB and PBMC samples were obtained at 1.5 h, 24 h, 48 h, 72 h or 168 h post dosing. Samples were lysed, followed by proteasome labeling with probe **1** and SDS-PAGE analysis. **A)** In-gel fluorescence measurements showing proteasome activity profiles in rat PWB and PBMC samples 1.5 h after treatment with marizomib or bortezomib. **B)** Quantification of total proteasome activity in rat PWB and PBMC samples after marizomib or bortezomib treatment for the indicated periods. Total proteasome activity was defined as the sum of all individual activities and was plotted as a percentage of activity in untreated rats. Error bars represent SEM. Data were analyzed by GraphPad Prism software. • $p<0.05$; •• $p<0.01$. **C)** Quantification of proteasome subunit activity in rat PWB and PBMC samples after marizomib or bortezomib treatment for the indicated time periods. Subunit activities were plotted as percentage of subunit activity in untreated rats. Error bars represent SEM.

largely absent compared to PBMC samples. Both marizomib and bortezomib inhibited all subunits, and recovery was also largely comparable between treatment schedules. The main difference was found at the level of the β1i activity, which recovered from marizomib high or marizomib low treatment at a similar rate as the β2 and β5(i)/β1 activity, but did not recover from bortezomib treatment. Since β1i activity was hardly present in PWB samples (see above), the effect of this difference on total proteasome activity was negligible.

Table 1 | Treatment regimes of patients #1 to #5.

Patient	Malignancy	Cycle 1	Subsequent cycles
#1	Prostate carcinoma	0.9 mg/m^2	0.7 mg/m^2
#2	Anorectal squamous carcinoma	0.9 mg/m^2	0.7 mg/m^2
#3	Hodgkin's lymphoma	0.8 mg/m^2	0.8 mg/m^2
#4	Colon carcinoma	0.8 mg/m^2	0.7 mg/m^2
#5	Non-Hodgkin's lymphoma	0.8 mg/m^2	0.7 mg/m^2

Remarkably, both bortezomib and marizomib treatment resulted in the disappearance of β2 labeling by probe **1**, whereas bortezomib has been shown to somewhat enhance and not inhibit the activity of β2 subunits in PWB using fluorogenic substrates.[9] A similar disappearance of β2 labeling after bortezomib treatment has been observed using probe **1** in cell lines *in vitro* and *ex vivo* in mouse tissues (see Chapter 2.1).

Probe 1 can profile proteasome activity in marizomib treated patients *ex vivo*

Having established that probe **1** can be used to assay the effects of proteasome inhibitors in both PBMC and PWB samples *ex vivo*, we next used probe **1** to evaluate the effect of marizomib treatment on patients suffering from different types of cancer (Table 1). Marizomib was administered intravenously at Day 1 (D1), Day 8 and Day 15 (D15) during 28-day treatment cycles (indicated with C1-C2). In the first treatment cycle, patients were dosed with concentrations of marizomib ranging from 0.55 mg/m^2 to 0.9 mg/m^2. Due to the occurrence of dose-limiting toxicities, marizomib concentrations were adjusted to 0.7 mg/m^2 or 0.8 mg/m^2 in subsequent cycles (Table 1). Blood was taken pre- and post-marizomib treatment; PBMCs and PWB samples were obtained and prepared in cell pellets. To study proteasome inhibition in PWB samples, PWB samples were lysed and incubated with probe **1**. Labeled subunits were separated by SDS-PAGE, and the resulting gels were scanned for fluorescence emission. An example of such a gel is shown in Figure 3A. Analogous to results found in rat blood samples, only constitutive proteasome subunits were labeled in PWB. Marizomib treatment decreased the activity of both β2 and β5/β1 subunits in PWB in all patients. Although in most patients some proteasome activity recovery was observed before the next dose of marizomib was given, proteasome activity never recovered to baseline levels in any of the patients (Figure 3A, pre treatment samples). These data are in accord with a previous report which suggested that PWB proteasome activity could be titrated with repeated dosing.[29]

To validate that probe **1** could be used to accurately study the effects of proteasome inhibitors on patient blood samples, we compared probe **1** with the use of fluorogenic substrates for the different activities in PWB samples obtained at C1D1 (Cycle 1 Day 1) and C1D15 (Cycle 1 Day 15) post treatment (Figure 3B). For both assays, the residual activity of different subunits was plotted as a percentage of the baseline activity before the start of treatment. As the β1 and β5 subunits labeled by probe **1** were not completely resolved on gel, we averaged the β1 and β5 activity levels obtained using fluorogenic substrates to be able to compare the two methods. Results obtained with probe **1** were very similar to results found using fluorogenic substrates (Figure 3B). Both methods revealed a higher inhibition of the β1/β5 activity compared to the β2 activity. In addition, both methods showed

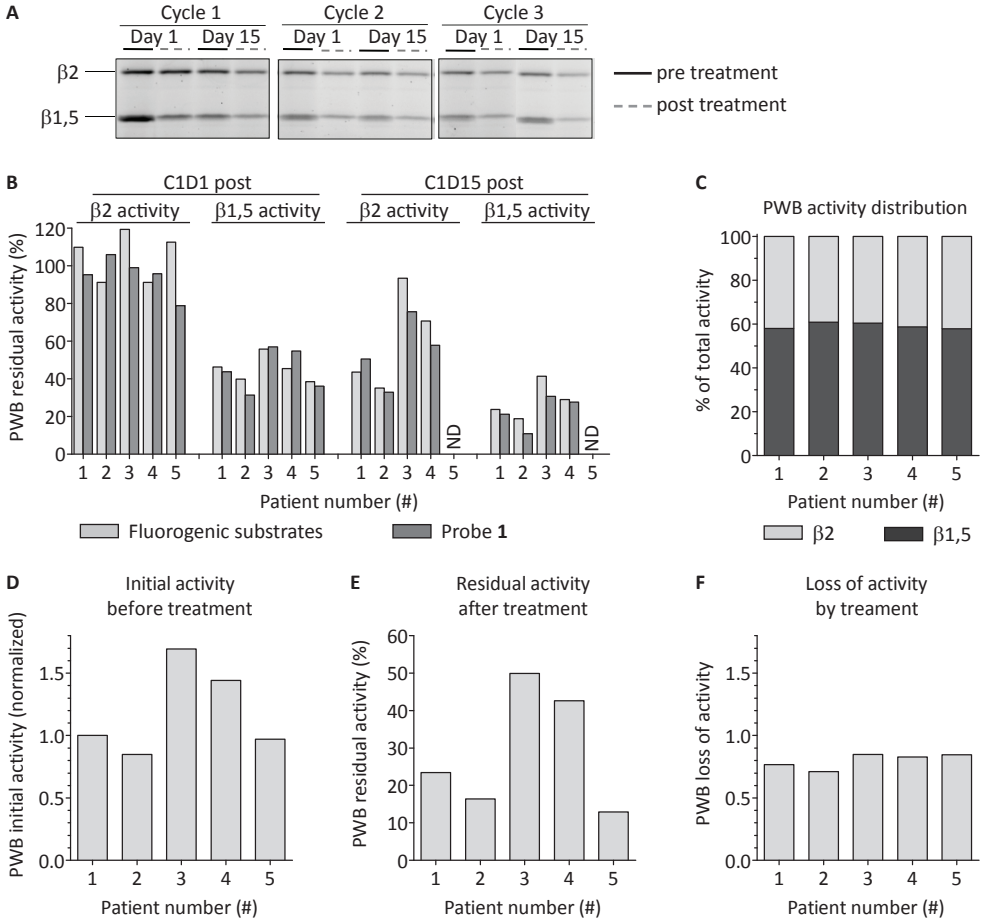


Figure 3 | *Ex vivo* profiling of proteasome activity in patient PWB samples. Patients with different malignancies (Table 1) were treated with varying doses of marizomib. Marizomib was administered at day 1, day 8, and day 15 in 28-day treatment cycles. Blood was taken pre and post dosing at days 1 and 15; PWB samples were prepared, lysed, and incubated with probe **1**. Labeled proteasome subunits were separated using SDS-PAGE, and the resulting gels were scanned for fluorescence emission. **A)** In-gel fluorescence measurements showing proteasome activity profiles in PWB samples of patient #1 that were obtained during marizomib treatment. **B)** Comparison of residual proteasome activities in PWB samples from patients #1 to #5 obtained at Cycle 1 Day 1 (C1D1) and Cycle 1 Day 15 (C1D15) post treatment. Residual activities compared to baseline values were calculated using probe **1** or fluorogenic substrates. ND: not determined. **C)** Quantification of the relative contribution of the β2 and β5/β1 subunits to total baseline proteasome activity in patient PWB samples prior to marizomib treatment. **D)** Quantification of baseline total proteasome activity levels in PWB samples of patients #1 to #5 before the start of marizomib treatment. Baseline activities were normalized to patient #1. **E)** Quantification of residual total proteasome activity levels at Cycle 2 Day 15 (C2D15) post treatment compared to baseline proteasome activity levels in PWB samples obtained from patients #1 to #5. **F)** Quantification of the absolute amount of inhibited proteasome activity at C2D15 post treatment compared to baseline values in PWB samples of patients #1 to #5. Samples were normalized to baseline proteasome activity in patient #1.

that repeated dosing resulted a higher extent of inhibition (compare C1D15 to C1D1 in Figure 3B). Both methods also revealed that the inhibition in PWB was dose-dependent, especially at C1D15 post treatment. Patients #1 and #2 were treated with 0.9 mg/m^2 in the first treatment cycle, while patients #3, #4 and #5 received 0.8 mg/m^2 during the first treatment cycle. In accord, patients #3 and #4 showed higher residual β2 and β5 activities at C1D15 post treatment compared to patients #1 and #2 (for patient #5 no PWB sample was available at C1D15 post treatment). Together, these results indicate that probe **1** can be reliably used to assay proteasome activity in patient blood samples *ex vivo*.

Proteasome inhibition by marizomib differs between individual patients

We quantified proteasome inhibition as profiled using probe **1** in PWB samples of patients #1 to #5 (obtained during C1 and C2). In PWB baseline samples of different patients that were obtained before the start of marizomib treatment, identical β2/β5,1 ratios (Figure 3C) with different total proteasome activity levels (Figure 3D) were observed. To compare inhibition profiles between patients, proteasome activity at C2D15 (Cycle 2 Day 15) post dosing was compared to baseline activity for patients #1 to #5. In accord with marizomib's inhibition profile, inhibition of the β5 and β1 subunits was approximately 2 fold higher than inhibition of the β2 subunit in all patients (Supplementary Figure 2A). Total proteasome inhibition, however, differed substantially between patients, with residual proteasome activities ranging from 13% to 50% (Figure 3E). The residual proteasome activity correlated well to the initial proteasome activity in all patients. Patients with high initial proteasome activity also displayed high residual activity, while patients with low initial activity also showed low residual proteasome activity (compare Figures 3D and 3E). This suggested that all patients lost a similar absolute amount of proteasome activity as a result of marizomib administration. To investigate this, we subtracted the total fluorescence intensity at C2D15 post dosing from the baseline fluorescence intensity in each patient to calculate the loss of fluorescence intensity over the course of two treatment cycles (Figure 3F). As the total fluorescence intensity directly correlates to total proteasome activity in these samples, the loss of fluorescence directly correlates to the loss of proteasome activity. As can be seen from Figure 3F, a similar amount of proteasome activity was inhibited in all patients in PWB at the end of the second treatment cycle. Therefore, the residual proteasome activity in PWB after marizomib treatment is mainly determined by the initial amount of proteasome activity.

Initial proteasome activity determines PBMC response to marizomib

Next we investigated proteasome inhibition by marizomib in patient PBMCs. To this end, PBMCs were lysed and incubated with probe **1**. Labeled subunits were separated by SDS-PAGE and the resulting gels were scanned for fluorescence emission. An example of such a gel is shown in Figure 4A. Analogous to results found in rat blood samples, both constitutive and immunoproteasome subunits were labeled in PBMCs (Figure 4A). In contrast to results found in PWB, the relative contribution of subunits to total proteasome activity differed between patients (Figure 4B). Diverse labeling patterns were observed, ranging from almost equal contribution of all subunits to total activity in patient #3, to subunits being responsible for over 50% (β2i and β1i subunits in patients #2 and #5, respectively) or less than 5% (β2 in patients #1 and #5) of total proteasome activity. In PBMCs baseline samples of different patients, total proteasome

activity levels were variable (Figure 4C) and resembled initial proteasome activity levels in the corresponding PWB samples. In both types of samples, high initial proteasome activity was found in patients #3 and #4, patient #1 had intermediate activity, while the lowest activity was found in patients #2 and #5 (Figures 3D and 4C). Administration of marizomib decreased the activity of all subunits in all patients. In contrast to results found in PWB, total proteasome activity was largely restored to baseline levels before a new dose of marizomib was administered (Figure 4A).

We quantified proteasome inhibition and recovery in PBMC samples of patients #1 to #5 (obtained during C1 and C2). Since proteasome activity recovers between doses in PBMCs, the residual proteasome activity at C2D15 post treatment compared to baseline levels depended on the extent of recovery. However, each dose resulted in a similar reduction in proteasome activity in post treatment samples compared to the preceding pre treatment sample, independent of the extent of recovery in this pre treatment sample. Therefore, we first normalized the inhibition in each post treatment sample to the corresponding pre treatment sample, and plotted the average residual activity ± SEM. In PBMCs, the average residual proteasome activity differed between patients and ranged from 33% to 52% (Figure 4D). In addition, the average residual proteasome activity correlated well to the initial proteasome activity in all patients, as observed in PWB. In accord with results found in PWB, inhibition of the β2,2i subunits in PBMCs was two fold lower compared to inhibition of the β5,5i and β1,1i subunits, with the exception of patient #5 in which all subunits were equally inhibited (Supplementary Figure 2B). Next, total fluorescence intensity in each post treatment sample was subtracted from the total fluorescence intensity in each corresponding pre treatment sample. On average, a similar absolute amount of proteasome activity was inhibited in PBMCs as a result of each marizomib administration in all patients (Figure 4E). This suggests that the initial amount of proteasome activity in PBMCs is one of the factors determining the residual amount of proteasome activity after marizomib treatment.

Patient PMBCs change their proteasome composition in response to marizomib

Next, we studied the influence of recovery on proteasome inhibition in patient PBMCs samples. To take recovery effects into account, we calculated the residual activity at C2D15 post treatment compared to baseline activity before the start of treatment (Figure 4F, shaded bars). We compared these values to the average residual activity (Figure 4F, filled bars) in which recovery effects were not taken into account (as plotted in Figure 4D). In patient #1, activity recovered almost completely between doses and therefore little difference was observed between the average residual activity and the residual activity at C2D15 post treatment. In patient #4, activity recovered >100%, resulting in an increase in proteasome levels over time. As a consequence, the residual activity at C2D15 post treatment was higher compared to the average residual activity. In patients #2, #3 and #5, proteasome activity did not completely recover to baseline levels between doses, and therefore the residual proteasome activity at C2D15 post treatment was lower compared to the average residual activity. Together, these data suggest that the extent of activity recovery between administrations influences the efficacy of marizomib treatment over multiple treatment cycles.

To study recovery in individual patients in more detail, we quantified the relative contribution of individual subunits to total proteasome activity in PBMC samples from patients

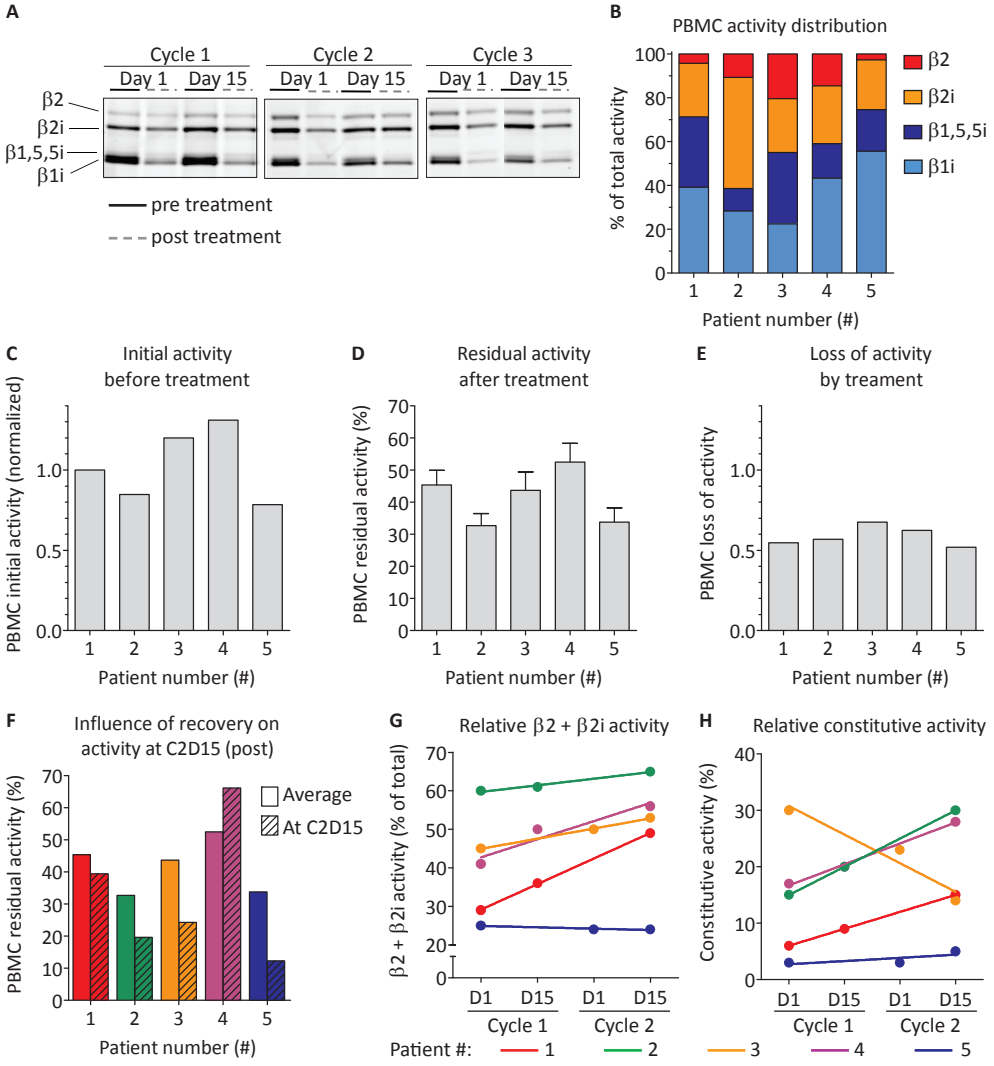


Figure 4 | *Ex vivo* profiling of proteasome activity in patient PBMC samples. Patients with different malignancies (Table 1) were treated with varying doses of marizomib. Marizomib was administered at day 1, day 8 and day 15 in 28-day treatment cycles. Blood was taken pre and post dosing at days 1 and 15 and PBMC samples were prepared, lysed and incubated with probe **1**. Labeled proteasome subunits were separated using SDS-PAGE and the resulting gels were scanned for fluorescence emission. **A)** In-gel fluorescence measurements showing proteasome activity profiles in PBMC samples of patient #1 that were obtained during marizomib treatment. **B)** Quantification of the relative contribution of the β2, β2i, β1/5/5i and β1i subunits to total baseline proteasome activity in patient PBMC samples prior to marizomib treatment. **C)** Quantification of baseline total proteasome activity levels in PBMC samples of patients #1 to #5 before the start of marizomib treatment. Baseline activities were normalized to patient #1. **D)** Quantification of residual total proteasome activity levels post marizomib treatment in PBMC samples obtained from patients #1 to #5. Proteasome activity in each post treatment sample was normalized to the preceding pre treatment sample and

averaged over all administrations. Error bars represent SEM. **E)** Quantification of the absolute amount of proteasome activity that disappeared on average in PBMC samples of patients #1 to #5 after each administration of marizomib. Samples were normalized to baseline proteasome activity in patient #1. **F)** Quantification of the average residual proteasome activity after each administration in PBMCs obtained from patients #1 to #5 (solid bars). Quantification of the residual proteasome activity at C2D15 (Cycle 2 Day 15) post treatment compared to baseline proteasome activity levels in PBMCs obtained from patients #1 to #5 (shaded bars). **G)** Quantification of the relative contribution of the β2 + β2i activities to total proteasome activity in patients #1 to #5 during treatment cycles 1 and 2 in PBMC samples in which activity has recovered (pre treatment samples). **H)** Quantification of the relative contribution of constitutive proteasome to total proteasome activity in patients #1 to #5 during treatment cycles 1 and 2 in PBMC samples in which activity has recovered (pre treatment samples).

#1 to #5 in which activity had recovered (pre treatment samples obtained during cycles 1 and 2). Both the contribution of the β2 activity to total proteasome activity (defined as the sum of the β2 and β2i activities, which are least targeted by marizomib; Figure 4G), and the contribution of constitutive activities to total activity (Figure 4H) were calculated. In patients #1 and #4, in which proteasome activity fully recovered between marizomib administrations, the relative β2 activity increased steadily over time in response to marizomib treatment, independent of the initial active subunit composition. In patients #2 and #3 in which proteasome activity did not fully recover between doses, the relative β2 activity increased less in response to marizomib treatment, while in patient #5 in which least proteasome activity recovery was observed the relative β2 activity somewhat decreased over time. In addition, the relative constitutive activity increased in all patients with the exception of patient #3. A similar shift from immunoproteasome to constitutive proteasome was observed *in vitro* in cells with acquired resistance to bortezomib. In these cells, this shift was accompanied by the upregulation of constitutive activities that could no longer be inhibited by bortezomib and of apoptosis resistance.[30] Together, these data suggest that the rapid modification of proteasome composition and thereby cleavage specificity allow PBMCs to adapt and survive marizomib treatment. The ability to upregulate β2 activities, which are less sensitive to inhibition by marizomib, seems to correlate to the ability of PBMCs to fully recover proteasome activity between administrations. Whether upregulation of β2 activities results in a decreased sensitivity and thereby less cell death in response to marizomib treatment remains to be determined. These data indicate that PBMCs modify their active subunit composition in response to marizomib treatment and that this may be a factor contributing to marizomib response.

DISCUSSION

Marizomib is a second-generation proteasome inhibitor isolated from the marine actinomycete *Salinispora tropica* that binds all active proteasome subunits irreversibly.[9] Marizomib has been evaluated in a wide range of preclinical models and has demonstrated clinical activity against both hematologic and solid tumors.[21] Clinical trials with marizomib are currently ongoing in both hematologic and solid tumors, enabling for the first time a thorough evaluation of patient response to marizomib treatment. Proteasome probe **1** is a recently developed reagent that enables profiling of all proteasome activities in a single measurement in a wide range of

tissue types, both *in vitro* and *ex vivo*.[27] The present study shows that probe **1** can be used to profile the effects of proteasome inhibitors on blood samples, derived both from animals and, importantly, from patients that are treated with this type of drug. Inhibition profiles of marizomib and bortezomib in rat PBMCs were in accord with previous reports[9,26] and revealed that marizomib inhibited all subunits with the highest affinity for β5 subunits, while bortezomib inhibited the β5 and β1 but not the β2 subunits. Parallel measurement of the proteasomal activities using both probe **1** and fluorogenic substrates in patient PWB samples gave similar results, further validating the use of probe **1** for assaying proteasome activity in blood samples. It has been shown that measuring only the chymotryptic activity may not accurately reflect the total proteasome activity, and that both the caspase-like and tryptic activities may considerably contribute to overall proteasome activity.[19] The use of chemical probes to assay proteasome activity has the advantage that the activity of all constitutive and immunoproteasome subunits can be profiled in a single measurement. In addition, probe **1** has equal affinity for all proteasome active sites and can therefore be used to establish the relative contribution of each subunit to total proteasome activity and to reliably measure total proteasome activity. Recovery patterns in PWB, which consists predominantly of erythrocytes (>99%) that are devoid of a nucleus, differed substantially from results found in PBMCs. Whereas in rat PBMCs total proteasome activity recovered within 72 h to 168 h post dosing, proteasome activity in PWB samples was only recovered by 20 to 50% after 168 h. In patients, PBMC proteasome activity recovered almost completely between doses, while recovery was near absent in PWB. These differences in recovery can be explained by the inability of erythrocytes to synthesize proteins, which makes proteasome activity recovery completely dependent on the slow turnover of erythrocytes ($t_{1/2}$ = 15-17 weeks).[25] It has been suggested previously that blood proteasome inhibition does not correlate with tumor proteasome inhibition or inhibitor efficacy.[11,29] The data presented here are a further indication that erythrocytes respond to and recover from proteasome inhibitors uniquely and caution against the use of PWB as a model system to monitor the effects of proteasome inhibitors in animals or patients, as is routinely done.

Probe **1** enabled a detailed study of patient response to marizomib treatment. Although proteasome activity was monitored in a small group of patients, factors that may mediate sensitivity to proteasome inhibition could be derived from these data. The data presented here indicate for the first time that marizomib treatment results in inhibition of similar absolute amounts of proteasome activity in both PWB and PBMCs. This suggests that both the relative and absolute residual proteasome activities after marizomib treatment are determined by the initial level of total proteasome activity. Therefore, total proteasome activity is likely one of the factors determining the sensitivity of cells to marizomib and possibly other proteasome inhibitors, as has been suggested earlier.[8,31] In accord with this, cells with acquired bortezomib resistance have shown increased expression and activity of the proteasome *in vitro*.[30,32,33] Initial proteasome activity in PWB seemed to correlate to initial activity in PBMCs, suggesting that proteasome activity levels may be regulated similarly in different cell types. Therefore, the initial proteasome activity in blood may be indicative of the initial activity in hematologic malignancies. Measuring the initial proteasome activity in PWB compared to a benchmark sample may thus be an easy approach to determine whether a patient is likely to respond to proteasome inhibitor therapy and could be a first

step towards personalized proteasome inhibitor therapy.

The ability of cells to shift their active subunit composition in response to proteasome inhibitor therapy may be another factor that mediates sensitivity to this class of compounds. Patient PBMCs responded to marizomib therapy by shifting their proteasome activity profile towards a higher degree of β2 activity and a higher degree of activity of constitutive subunits, in line with observations made *in vitro*.[30] These shifts probably enable PBMCs to better survive proteasome inhibition. Therefore, the (partial) absence of these shifts in tumor cells likely renders cells more vulnerable to proteasome inhibition. It is tempting to speculate that certain tumor types lack the ability to shift their active subunit composition in response to proteasome inhibitor therapy, suggesting that tumor types that are especially sensitive to proteasome inhibitors could be identified. Alternatively, it would be interesting to define the cellular mechanisms that mediate these shifts, as pharmacological interference with such pathways may overcome resistance to and enhance the therapeutic potential of proteasome inhibitors.

Finally, differences in initial subunit activity distribution are likely to influence the response of tumors to different proteasome inhibitors, as was suggested previously by Krauss et al.[32] Both red and white blood cells, which have a relatively high, micromolar intracellular proteasome concentration,[25] are thought to sequester proteasome inhibitors *in vivo*,[29] as blood cells are the first cell type to come in contact with intravenously administered proteasome inhibitors. Consequently, proteasome inhibitors likely saturate proteasomes in red blood cells, which constitute 99% of all blood cells, before peripheral sites can be reached,[29] which may limit inhibitor concentrations in tumor cells. PR-957 is a novel β5i-specific proteasome inhibitor that showed activity against immunoproteasomes in mice.[34] As red blood cells do not express immunoproteasome, specific immunoproteasome inhibitors may saturate blood proteasome at relatively low concentrations, resulting in higher inhibitor concentrations in tumor cells. In addition, immunoproteasome inhibitors are suggested to cause fewer side effects in tissues that express only low levels of immunoproteasome subunits.[8] Therefore, PR-957 or other immunoproteasome inhibitors may in future provide a good alternative for tumors that originate from immune tissues that have intrinsic immunoproteasome activity, such as spleen, lymph nodes and thymus.[27] The data presented here also suggest that the relative contribution of subunits to total proteasome activity may differ between patients and similar observations were made in primary human leukemia cells.[32] In addition, analysis of rat PBMCs suggest that proteasome inhibitor profiles may differ substantially, especially at low concentrations. Whereas the β1i subunit was especially susceptible to bortezomib inhibition, the β2 subunits could only be inhibited by marizomib. Probing tumor cells with probe **1** before treatment will establish which subunits are most active. This will enable matching of tumor proteasome profile to a complementary inhibitor or multiple inhibitors to personalize proteasome inhibitor therapy and maximize therapy response.

MATERIALS AND METHODS

Reagents

All solvents were purchased from Biosolve (Valkenswaard, the Netherlands) at the highest grade available**.** All other chemicals were purchased from Sigma-Aldrich (Zwijndrecht, the Netherlands) at the highest available purity. Bortezomib (obtained commercially) and marizomib synthesized at Nereus (San Diego, CA). Probe **1** was synthesized as described in chapter 2.1.

***Ex vivo* profiling rat blood samples using fluores-**

cent probe 1

Male Sprague-Dawley rats (280-325 grams) were treated once intravenously with 40% Propylene glycol/50% citrate buffer/10% ethanol based vehicle with or without marizomib (0.05 or 0.1 mg/kg) or bortezomib (0.15 mg/kg in saline). Fifteen rats served as untreated controls. At 90 minutes, 24, 48, 72, and 168 hours post dose, 3 rats per time point per dosing group were anesthetized and approximately 6-7 mL of whole blood was collected via the abdominal aorta into Na-Heparin vacutainer tubes. To obtain a PWB pellet, 2 mL of whole blood was centrifuged for 10 minutes at 800g; the blood pellet was then washed once with cold 1X PBS (supernatant was discarded) and frozen at -80 °C. PBMCs were isolated using Lympholyte®-Mammal (Cedarlane Laboratories, Ontario, Canada): 4 mL of whole blood was diluted 1:1 with 4 mL of PBS and the tube was inverted 2-3 times to mix; 4 mL of the diluted blood was then layered onto 3 mL of Lympholyte®-Mammal in a 15 ml centrifuge tube and centrifuged at room temperature for 20 minutes at 800g. The PBMC layer at the interface was transferred to a new centrifuge tube, diluted 1:1 with PBS, and pelleted for 10 minutes at 800g. The supernatant was discarded; the PBMCs were washed once in PBS and frozen at -80 °C.

PBMC and PWB pellets were thawed on ice. 50 μL of ice-cold 5 mM EDTA pH8 lysis buffer was added to the PBMC pellets, and 2 mL lysis buffer to the PWB pellets. Pellets were resuspended and lysed on ice for 1 h. Lysates were then centrifuged for 10 min at 16,000 x g at 4 °C, and the supernatants were aliquoted and stored at -80°C. Protein concentration was determined using a Bradford assay (Biorad). Protein concentrations were adjusted to 2 μg/μL with HR-buffer (50 mM Tris-HCl pH 7.4, 250 mM sucrose, 5 mM $MgCl_2$, 1 mM dithiothreitol (DTT), 2 mM ATP), and lysates were incubated with 1 μM probe **1** for 1 hour at 37 °C. Data were analyzed by using GraphPad Prism software (GraphPad, La Jolla, CA, USA).

***Ex vivo* profiling patient blood samples**

At indicated time points according the clinical protocol, 5 mL of blood was drawn in heparinized vacutainers. Blood samples were obtained pre dosing or within 2.5 h to 5 h post dosing for all patients. 1 mL of the heparinized whole blood was centrifuged at 2000g for 10 min at room temperature (RT) and the PWB pellet was washed once in 5 times the pellet volume of ice-cold 1X DPBS (Dulbecco's Phosphate Buffered Saline) with Ca^{+2}/Mg^{+2}. After centrifugation (2000g, 10 min RT) and discarding of supernatants, PWB pellets were frozen and stored at -70°C. The remaining 4 mL of heparinized whole blood was diluted by adding an equal volume of Hanks Balanced Salt Solution and layered over 6 mL of Ficoll-Paque™ PLUS. After centrifugation for 40 min at 400g at room temperature (RT), PBMCs located at the interphase were collected, transferred to a new tube, and washed with at least three volumes of Hanks Balanced Salt Solution. To lyse any co-isolated RBCs, PBMC pellets were resuspended in 1x BD PharmLyse™ buffer and incubated at RT for 5 minutes. 9 mL of Hanks Balanced Salt Solution was then added and cells were centrifuged for 10 min at 1000g (RT). PBMCs were resuspended in Hanks Balanced Salt Solution, an aliquot was taken for a viability cell count by a Flow based 7-Amino-actinomycin D (7-AAD) staining procedure and the remaining PBMC solution was centrifuged for 10min at 1000g (RT). Supernatant was discarded, and PBMC cell pellets were flash frozen and stored at -70 °C.

For *ex vivo* profiling of proteasome activities using probe **1**, cell pellets were lysed in HR lysis buffer (50 mM Tris pH 7.4, 1 mM dithiothreitol (DTT), 5 mM $MgCl_2$, 2 mM ATP and 250 mM sucrose) by sonifcation using the Bioruptor sonicator (Diagenode, Liège, Belgium). Lysates were centrifuged at 16,000g to remove cell debris and protein concentration was determined using the Bradford assay (Biorad). Protein concentrations were adjusted to 1 μg/μL (PBMC) or 5 μg/μL (PWB), and lysates were incubated with 1 μM probe **1** for 1 h at 37 °C. To measure proteasome activities *ex vivo* using fluorogenic substrates, cells were lysed in ice-cold 5 mM EDTA pH8 lysis buffer and protein concentrations were determined using a BCA (bicinchoninic acid) assay. The 20S proteasome activities in cell lysates were measured as described previously.[25,35] In brief 20 μg of protein was used in a total volume of 200μL consisting of assay buffer (20 mM HEPES, 0.5 mM EDTA, pH 8.0) with a final SDS concentration of 0.035% and 20 μM of the peptide substrates Suc-LLVY-AMC (Chymotryptic activity), Z-ARR-AMC (Tryptic activity) or Z-LLE-AMC (caspase-like activity) (Calbiochem). Fluorescence (λ ex/em = 390/460 nm)

was measured every 5 min for 2 hr using a Fluoroskan Ascent 96-well microplate reader (Thermo Electron, Waltham, MA).

In-gel fluorescence measurements

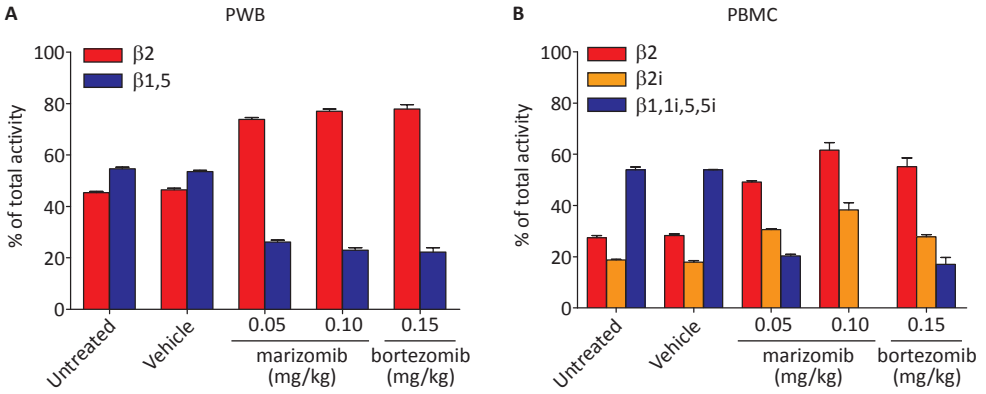
Incubated samples were denatured by boiling in LDS sample buffer (Invitrogen) containing 2.5% β-mercaptoethanol and polypeptides were resolved by 12% Bis-Tris gels using the NuPAGE system from invitrogen. Wet gel slabs were imaged for 10s to 120s, with a resolution of 100 μm, using the ProXPRESS 2D Proteomic imaging system (Perkin Elmer) with appropriate filter settings (λ ex/em = 480/530 nm). To verify protein loading, gels were stained with coomassie and probe signals were normalized accordingly. Images were analyzed using TotalLab analysis software (Nonlinear Dynamics, Newcastle upon Tyne, UK) to quantify the intensity of the bands detected. To compare baseline proteasome activities between patients, samples were rerun on a single gel.

REFERENCES

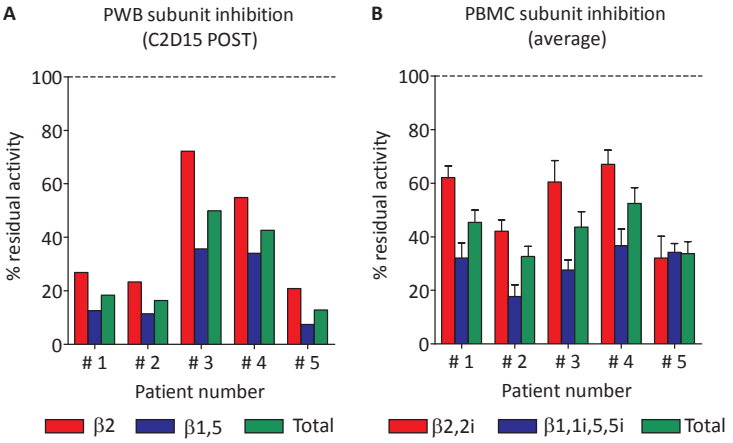
1. Hideshima, T. et al. The proteasome inhibitor PS-341 inhibits growth, induces apoptosis, and overcomes drug resistance in human multiple myeloma cells. *Cancer Res.* **61**, 3071-3076 (2001).
2. Jagannath, S. et al. A phase 2 study of two doses of bortezomib in relapsed or refractory myeloma. *Br. J. Haematol.* **127**, 165-172 (2004).
3. Richardson, P.G. et al. A phase 2 study of bortezomib in relapsed, refractory myeloma. *N. Engl. J. Med.* **348**, 2609-2617 (2003).
4. Richardson, P.G. et al. Bortezomib or high-dose dexamethasone for relapsed multiple myeloma. *N. Engl. J. Med.* **352**, 2487-2498 (2005).
5. Fisher, R.I. et al. Multicenter phase II study of bortezomib in patients with relapsed or refractory mantle cell lymphoma. *J. Clin. Oncol.* **24**, 4867-4874 (2006).
6. Orlowski, R.Z. et al. Phase 1 trial of the proteasome inhibitor bortezomib and pegylated liposomal doxorubicin in patients with advanced hematologic malignancies. *Blood* **105**, 3058-3065 (2005).
7. Orlowski, R.Z. et al. Randomized phase III study of pegylated liposomal doxorubicin plus bortezomib compared with bortezomib alone in relapsed or refractory multiple myeloma: combination therapy improves time to progression. *J. Clin. Oncol.* **25**, 3892-3901 (2007).
8. Orlowski, R.Z. & Kuhn, D.J. Proteasome inhibitors in cancer therapy: lessons from the first decade. *Clin. Cancer Res.* **14**, 1649-1657 (2008).
9. Chauhan, D. et al. A novel orally active proteasome inhibitor induces apoptosis in multiple myeloma cells with mechanisms distinct from Bortezomib. *Cancer Cell* **8**, 407-419 (2005).
10. Piva, R. et al. CEP-18770: A novel, orally active proteasome inhibitor with a tumor-selective pharmacologic profile competitive with bortezomib. *Blood* **111**, 2765-2775 (2008).
11. Demo, S.D. et al. Antitumor activity of PR-171, a novel irreversible inhibitor of the proteasome. *Cancer Res.* **67**, 6383-6391 (2007).
12. Adams, J. The development of proteasome inhibitors as anticancer drugs. *Cancer Cell* **5**, 417-421 (2004).
13. Nalepa, G., Rolfe, M. & Harper, J.W. Drug discovery in the ubiquitin-proteasome system. *Nat. Rev. Drug Discov.* **5**, 596-613 (2006).
14. Rock, K.L. & Goldberg, A.L. Degradation of cell proteins and the generation of MHC class I-presented peptides. *Annu. Rev. Immunol.* **17**, 739-779 (1999).
15. Strehl, B. et al. Interferon-gamma, the functional plasticity of the ubiquitin-proteasome system, and MHC class I antigen processing. *Immunol. Rev.* **207**, 19-30 (2005).
16. Dahlmann, B., Ruppert, T., Kuehn, L., Merforth, S. & Kloetzel, P.M. Different proteasome subtypes in a single tissue exhibit different enzymatic properties. *J. Mol. Biol.* **303**, 643-653 (2000).
17. Dahlmann, B., Ruppert, T., Kloetzel, P.M. & Kuehn, L. Subtypes of 20S proteasomes from skeletal muscle. *Biochimie* **83**, 295-299 (2001).
18. Groll, M., Huber, R. & Potts, B.C. Crystal structures of Salinosporamide A (NPI-0052) and B (NPI-0047) in complex with the 20S proteasome reveal important consequences of beta-lactone ring opening and a mechanism for irreversible binding. *J. Am. Chem. Soc.* **128**, 5136-5141 (2006).
19. Kisselev, A.F., Callard, A. & Goldberg, A.L. Importance of the different proteolytic sites of the proteasome and the efficacy of inhibitors varies with the protein substrate. *J. Biol. Chem.* **281**, 8582-8590 (2006).
20. Cusack, J.C., Jr. et al. NPI-0052 enhances tumoricidal response to conventional cancer therapy in a colon cancer model. *Clin. Cancer Res.* **12**, 6758-6764 (2006).
21. Fenical, W. et al. Discovery and development of the anticancer agent salinosporamide A (NPI-0052). *Bioorg. Med. Chem.* **17**, 2175-2180 (2009).
22. Sloss, C.M. et al. Proteasome inhibition activates epidermal growth factor receptor (EGFR) and EGFR-independent mitogenic kinase signaling pathways in pancreatic cancer cells. *Clin. Cancer Res.* **14**, 5116-5123 (2008).
23. Chauhan, D. et al. Combination of proteasome inhibitors bortezomib and NPI-0052 trigger in vivo

synergistic cytotoxicity in multiple myeloma. *Blood* **111**, 1654-1664 (2008).
24. Roccaro, A.M. et al. Dual targeting of the proteasome regulates survival and homing in Waldenstrom macroglobulinemia. *Blood* **111**, 4752-4763 (2008).
25. Lightcap, E.S. et al. Proteasome inhibition measurements: clinical application. *Clin. Chem.* **46**, 673-683 (2000).
26. Berkers, C.R. et al. Activity probe for in vivo profiling of the specificity of proteasome inhibitor bortezomib. *Nat. Methods* **2**, 357-362 (2005).
27. Berkers, C.R. et al. Profiling proteasome activity in tissue with fluorescent probes. *Mol. Pharm.* **4**, 739-748 (2007).
28. Verdoes, M. et al. A fluorescent broad-spectrum proteasome inhibitor for labeling proteasomes in vitro and in vivo. *Chem. Biol.* **13**, 1217-1226 (2006).
29. Williamson, M.J. et al. Comparison of biochemical and biological effects of ML858 (salinosporamide A) and bortezomib. *Mol. Cancer Ther.* **5**, 3052-3061 (2006).
30. Fuchs, D., Berges, C., Opelz, G., Daniel, V. & Naujokat, C. Increased expression and altered subunit composition of proteasomes induced by continuous proteasome inhibition establish apoptosis resistance and hyperproliferation of Burkitt lymphoma cells. *J. Cell Biochem.* **103**, 270-283 (2008).
31. Berkers, C.R. & Ovaa, H. Drug discovery and assay development in the ubiquitin-proteasome system. *Biochem. Soc. Trans* **38**, 14-20 (2010).
32. Kraus, M. et al. Activity patterns of proteasome subunits reflect bortezomib sensitivity of hematologic malignancies and are variable in primary human leukemia cells. *Leukemia* **21**, 84-92 (2007).
33. Ruckrich, T. et al. Characterization of the ubiquitin-proteasome system in bortezomib-adapted cells. *Leukemia* **23**, 1098-1105 (2009).
34. Muchamuel, T. et al. A selective inhibitor of the immunoproteasome subunit LMP7 blocks cytokine production and attenuates progression of experimental arthritis. *Nat. Med.* **15**, 781-787 (2009).
35. Stein, R.L., Melandri, F. & Dick, L. Kinetic characterization of the chymotryptic activity of the 20S proteasome. *Biochemistry* **35**, 3899-3908 (1996).

Supplementary figures



Supplementary Figure 1 | Proteasome activity profiles are identical in untreated and vehicle treated rat blood samples. A) Quantification of the relative contribution of the β2 and β5/1 subunits to total proteasome activity in rat PWB samples 1.5 h post treatment. **B)** Quantification of the relative contribution of the β2, β2i and β5/5i/1/1i subunits to total proteasome activity in rat PBMC samples 1.5 h post treatment.



Supplementary Figure 2 | Residual proteasome activities in marizomib treated patients. A) Quantification of the residual activity of individual subunits after marizomib treatment in patient PWB samples normalized to baseline activity levels. **B)** Quantification of the residual activity of individual subunits after marizomib treatment in patient PBMC samples. Activities in post treatment samples were normalized to the preceding pre treatment sample and averaged over all administrations. Error bars represent SEM.

Chapter 2.3

Automated online sequential isotope labeling for protein quantitation applied to proteasome tissue-specific diversity

Molecular & Cellular Proteomics
Volume 7, 2008, Pages 1755-1762

Automated online sequential isotope labeling for protein quantitation applied to proteasome tissue-specific diversity

Reinout Raijmakers,[a] Celia R. Berkers,[b] Annemieke de Jong,[b] Huib Ovaa,[b] Albert J. R. Heck[a,*] & Shabaz Mohammed[a,*]

[a]Biomolecular Mass Spectrometry and Proteomics Group, Bijvoet Center for Biomolecular Research and Utrecht Institute for Pharmaceutical Sciences, Utrecht University, Sorbonnelaan 16, 3584 CA Utrecht, the Netherlands. [b]Division of Cellular Biochemistry, The Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands.

Quantitation of protein abundance is a vital component in the proteomic analysis of biological systems, which can be achieved by differential stable isotopic labeling. To analyze tissue-derived samples, the isotopic labeling can be performed using chemical labeling of the peptides post-digestion. Standard chemical labeling procedures often require many manual sample handling steps, reducing the accuracy of measurements. Here, we describe a fully automated, online (in nanoLC columns), labeling procedure, which allows protein quantitation using differential isotopic dimethyl labeling of peptide N termini and lysine residues. We show that the method allows reliable quantitation over a wide dynamic range and can be used to quantify differential protein abundances in lysates and, more targeted, differences in composition between purified protein complexes. We apply the method to determine the differences in composition between bovine liver and spleen 20 S core proteasome complexes. We find that although all catalytically active immunoproteasome subunits were up-regulated in spleen (compared with liver), only one of the normal catalytic subunits was down-regulated, suggesting that the tissue-specific immunoproteasome assembly is more diverse than previously assumed.

INTRODUCTION

Comparative quantitation of proteins between biological samples is an important endeavor in the identification of proteins that play specific roles in biological pathways or diseases. In the past, the most common way to analyze and compare whole proteomes was by two-dimensional polyacrylamide gel electrophoresis (PAGE), with quantitative information being gleaned from the intensity of the spots observed after staining the separated proteins in the gel. Perhaps the most common quantitation methods in proteomics in use today are based on the application of differential stable isotopic labeling of protein samples (using for example ^{2}H, ^{13}C, ^{15}N, or ^{18}O), combined with detection by mass spectrometry, allowing both identification and quantitation of the sample components.[1,2] There are two main distinguishable methods commonly used for stable isotopic labeling of samples.

*Correspondence may be adressed to A.J.R.H or S.M. (a.j.r.heck@uu.nl or s.mohammed@uu.nl).

The first is the incorporation of stable isotopes in proteins by supplying these isotopes to the growth media consumed and metabolized by cells or smaller organisms, generally termed metabolic labeling. The isotope label is then either incorporated as the single carbon or nitrogen source[3] or incorporated via specific auxotrophic amino acids that contain isotopes, called SILAC (stable-isotope labeled amino acids in cell culture).[4] The second method involves isotope labeling of isolated proteins or peptides with chemically identical tags that are isotopomers. The latter method is particularly advantageous for human or animal tissue samples where metabolic based incorporation cannot easily be achieved.[1]

Many approaches have been described for the chemical labeling of proteins and peptides to allow quantitation.[1] These include the labeling of free cysteines in proteins by ICAT (isotope-coded affinity tagging)[5] and the labeling of free amines in the peptides obtained after protein digestion using N-hydroxysuccinimide esters like those used in the iTRAQ (isobaric Tagging for Relative and Absolute Quantification) approach.[6] One of the most efficient reactions for chemical labeling of peptides is the specific dimethylation of free amines (peptide N termini and Lysine residues) by reductive amination using formaldehyde and cyanoborohydride.[7] In this reaction, which is performed in near neutral conditions (between pH 6 and pH 8.5), the primary amines react with formaldehyde to create a Schiff base, which is then reduced by the cyanoborohydride. The label causes a mass increase of 28 Da per primary amine for regular formaldehyde and a mass increase of 32 Da when deuterated formaldehyde is used. The reaction is fast, does not create spurious side products, and does not adversely affect the identification of the peptides from the MS/MS spectra.

As all chemical labeling procedures are sensitive to small deviations in the reaction conditions and to human error in the handling of samples prior to and during labeling events, accuracy and reproducibility of quantitative experiments can benefit from automated labeling procedures. Here, we present a fully automated, online, and on-column sequential stable isotope labeling procedure based on the dimethylation of primary amines. We show that on-column double labeling, which we use in combination with high resolution MS/MS analysis,[8,9] is as sensitive as offline labeling, whereas the overall on-column procedure is more efficient and open to automation. The method is validated using model samples of varying complexity and can be used both to analyze protein levels in whole lysates as well as in specific proteins or in protein complexes. Finally, we apply the method to determine quantitatively differences in composition between proteasomes purified from different tissues.

RESULTS

Online sequential labeling

To test the feasibility of performing the dimethylation reaction on-column, we designed an online reaction protocol consisting of 5 main steps, all of which were performed on our regular nanoLC system, consisting of a six-port switching valve, a 100-μm i.d. trapping column, and a 50 μm i.d. analytical column (Figure 1A).[10] During sample loading and online labeling (steps 1 to 4), the restrictor is closed and the flow speed is 5 μL/min. The first step is the loading of the first peptide sample onto the trapping column, which is then chemically labeled by flushing the trapping column with 40 μl of light labeling reagent (CH_2O/$NaBH_3CN$) in the second step. After a short wash with 5% formic acid, the second peptide sample is loaded onto the trapping column in step three followed by chemical labeling

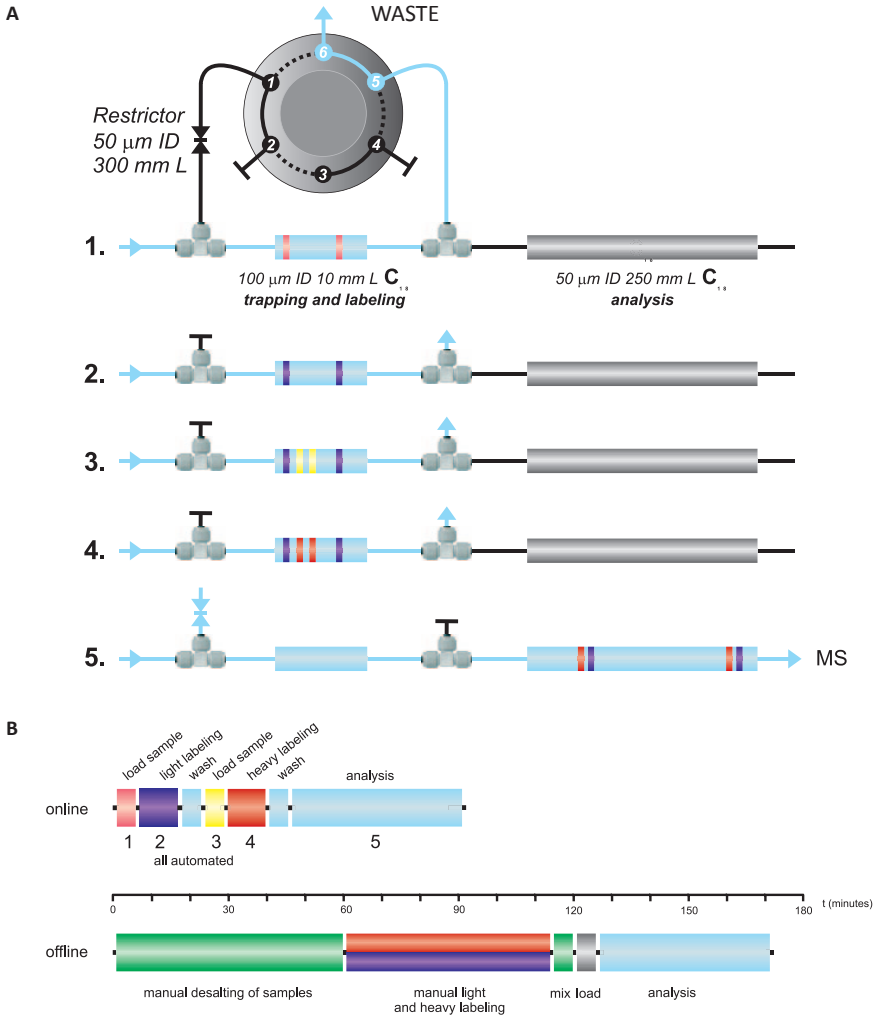


Figure 1 | A) Schematic showing the LC-MS setup and the main steps involved in the sequential online stable isotope labeling of two samples and subsequent LC-MS/MS analysis. Initially, the six-port valve closes the restrictor, directing the flow via the trapping column to the waste. In step 1, the first peptide sample (pink) is loaded onto the trapping column. The peptides now retained by the trapping column are chemically labeled in step 2 (purple) by flushing the trapping column with labeling reagent. Following a short wash, the second peptide mixture (orange) is loaded on the trapping column in step 3. During step 4, the trapping column is flushed with the second labeling reagent, containing the heavy isotope, labeling all peptides from the second sample (purple). After this step, the valve switches (opening the restrictor) and the flow speed increases, thus allowing flow over the analytical column to the mass spectrometer to perform a regular LC-MS analysis of all peptides. **B)** the estimated time involved in all of the steps, both automated online (top) and manual offline (bottom) procedures, with color coding taken from A. In green are the steps that are not required in the online method. Indicated in the middle is a time-scale providing an indication of the time involved.

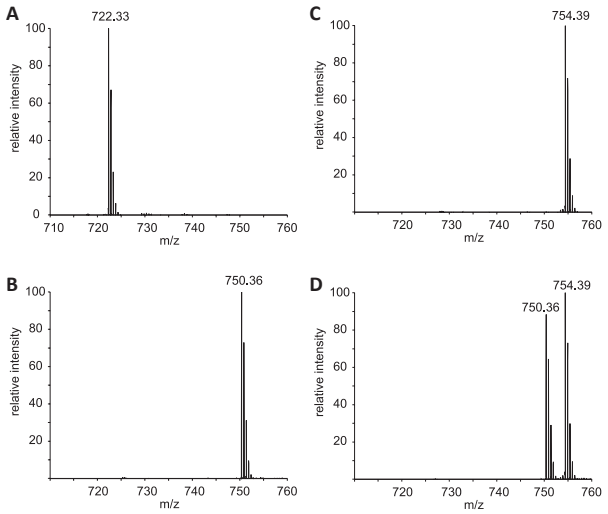


Figure 2 | Mass spectra of the doubly charged BSA peptide YICDNQDTISSK (amino acids 286–297) in its original form (panel A, calculated mono-isotopic mass 722.33) and after online labeling in either the light labeling step (panel B, calculated mono-isotopic mass 750.36), the heavy labeling step (panel C, calculated mono-isotopic mass 754.38), or both (panel D). All mass spectra are zoomed in on m/z range 710 to 760 and are averaged over the range in the chromatogram that includes the elution times of all three m/z values.

of that sample in step four when the column is flushed with the heavy labeling reagent ($CD_2O/NaBH_3CN$). Finally, again after a short wash with 5% FA, in the fifth step the valve switches, and the flow speed is increased to ~400 μl/min. This results in a pressure of ~150 bar and an effective flow of ~100 nl/min for a regular LC-MS analytical gradient to analyze all labeled peptides. The minimal time for the protocol is around 90 min, compared with about 170 min for an offline labeling (including desalting of the sample) and LC-MS analysis (Figure 1B). As the procedure is fully automated, no handling of the samples is required after digestion and loading the samples in the autosampler. In the offline method, several additional manual steps are required, including drying and mixing the samples, steps that are sensitive to sample loss and human error, respectively.

When performing a sequential labeling online it is of vital importance that the reaction goes to completion as otherwise the second label will also partially label the first sample, causing inaccurate quantitation. To test this, we first performed the online double labeling as described above, but although only injecting actual digested protein sample (bovine serum albumin (BSA), 100 fmol) in either the first (light) or the second (heavy) labeling step, though still performing the labeling in both steps. Figure 2 shows the averaged mass spectrum of one of the trypic peptides of BSA, YICDNQDTISSK (amino acids 286–297), which is detected as a doubly charged peptide ion at an m/z value of 722.33 when cysteine residues are carbamidomethylated. These spectra are averaged over the complete retention time span of the unmethylated and the fully methylated version of the peptide, to be able to see any intermediates. In Figure 2A, the unmodified peptide is seen from a regular LC-MS analysis. After dimethylation, the m/z of this peptide is expected to be 750 for the light label and 754 for the heavy label, as the peptide is dimethylated at both the N terminus as well as the C-terminal lysine residue. As can be seen for the light and heavy label in Figure 2, B and C, respectively, the online reaction goes to completion because no input peptide or any intermediates are visible after labeling. Figure 2D shows the same mass spectrum but now for a double labeling (two times 100 fmol) in which both the light and heavy la-

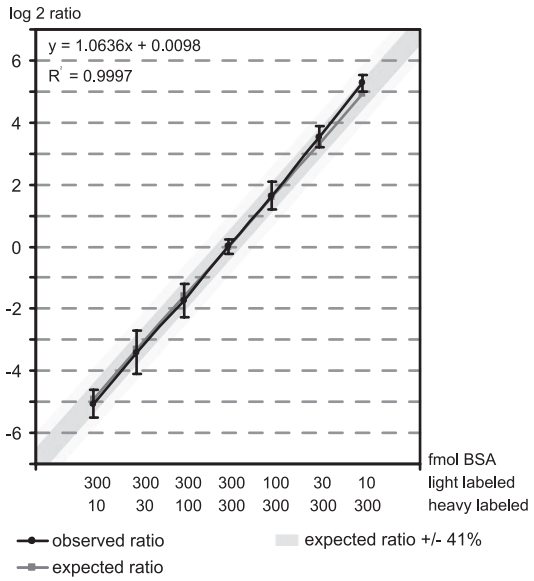


Figure 3 | Linearity of the online double stable isotope labeling procedure. Plotted is the correlation between the expected (x axis) and determined (y axis) ratio for several online labeling experiments involving different amounts of BSA, as indicated below the graph. The ratios are calculated from the 10 most intense peptide peak pairs of each analysis. Plotted are the log 2 values of the actual ratios, with the pink line (squares) indicating the expected ratios and the blue line (circles) the observed ratios. The dark blue shading indicates the spread of 0.5 log units from the expected value and the light blue shading a spread of 1 log unit. Bars indicate the standard deviation between the ten quantified peptides for each sample.

beled peptide are visible, in a ratio very close to 1:1. More examples of BSA peptides before and after double labeling can be found in supplementary Figure S1. We also analyzed this sample with a Mascot search with all dimethylation events set as variable modifications to determine the extent of identification of partially dimethylated peptides. The result of that search can be found in Supplementary Figure S2 and showed that at a Mascot ion score cutoff of 25, 98% of all identified spectra were dimethylated. Of the 157 identified spectra, four were not fully methylated and only one contained mixed light and heavy dimethylation. Because of the 30 s dynamic exclusion set in the mass spectrometer, these results strongly over-represent low intensity peptides, meaning the actual labeling percentage is much higher, as indicated in Figure 2. To check for recovery of peptides after the extended washing steps, we compared an online double labeling run of two times 20 fmol tryptic BSA with a subsequent normal LC-MS run of 20 fmol tryptic BSA. The observed signal intensities were similar, showing a near 100% recovery of the peptides following isotopic labeling. This behavior was typical and observed for most tryptic peptides analyzed in this way.

To test the dynamic range of the online labeling, different amounts of tryptic BSA peptides were loaded in the first and second sample loading step, varying between ratios of 30:1 to 1:30. Following the LC-MS analysis of the labeling, the ratio of BSA between the light and heavy label was determined from the average of the ten most intense peptide pairs using MSQuant. As can be seen in Figure 3, the method is linear in the base 2 log of the ratios and accurate over a wide dynamic range (the correlation between expected and observed ratios has an R^2 of 0.9997). The average ratios for all combinations tested fell within a 0.5 deviation on a base 2 log scale, which equates to a variation of ±41%. The standard deviations calculated from the log 2 ratios of each set of 10 peptides tend to be larger when the second (heavy) sample contains less peptide compared with the first (light) sample. This shows that with this method ratios can be reli-

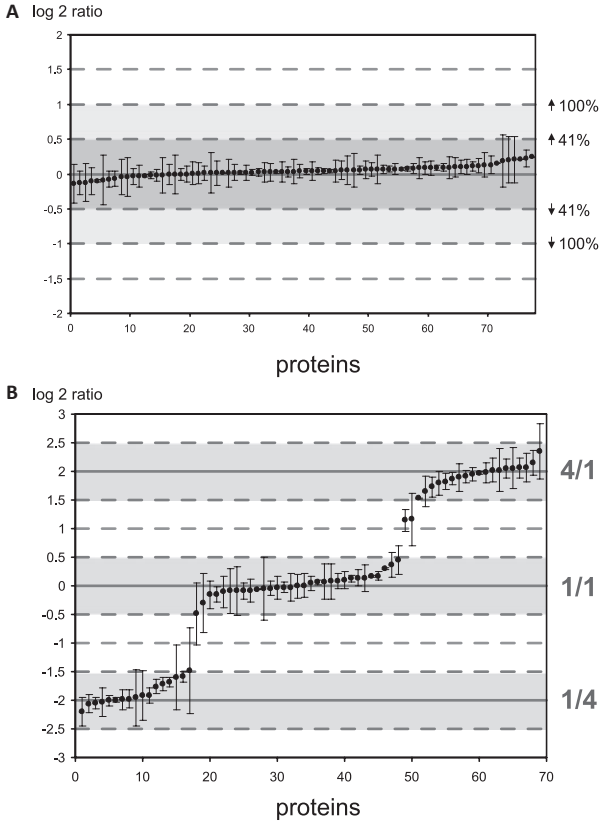


Figure 4 | Quantified ratios from mouse lung tissue proteins. Indicated on the x axis are the number of proteins that were identified from several gel bands that were mixed either all in a 1:1 ratio **(A)** or with several gel bands mixed in different ratios **(B)**. Ratios are calculated for all proteins from which at least two peptides were found that could be quantified and are plotted as the log 2 of the determined ratio. Shaded in blue, red, and green are the spread of 0.5 log units from the expected 1:1, 4:1, and 1:4 ratios, respectively. Bars indicate the standard deviation between the quantified peptides for each protein.

able calculated as low as for 1.5-fold changes, although usually a 2-fold change (100% variation) would be used as a cut-off in these types of experiments.

Of course, regular experiments would involve a large number of different proteins, with varying numbers of peptides that can be used for quantitation. To examine the behavior of the method under more complex conditions, a mouse lung tissue extract was run on a SDS-PAGE gel, cut into several bands, and the proteins present in the gel bands were subjected to in-gel tryptic digestion. These tryptic peptides were then first applied in equal amounts for heavy and light labeling, allowing us to determine the normal variation in the determined ratios introduced by the method. As can be seen in Figure 4A, we quantified just under 80 proteins, from three gel bands, with at least two peptides per protein. Similar to the quantitation of tryptic BSA, the variation of the average ratios did not exceed ±41%, and the standard deviation of the ratios observed for different peptides of the same protein was small. A “rough check” for the efficiency of labeling was performed via a Mascot search with all dimethylation events set as variable modifications. In this experiment 98% of all identified peptides were fully labeled, 1.6% were partially labeled, and 0.4% had mixed light and heavy labeling at a Mascot ion score cutoff of 25. The result of the search can be found in supplementary Figure S3 as well as in PRIDE (see “Materials and Methods” for details). Next, we mixed the gel bands in three different ratios (4:1, 1:1, and 1:4) and repeat-

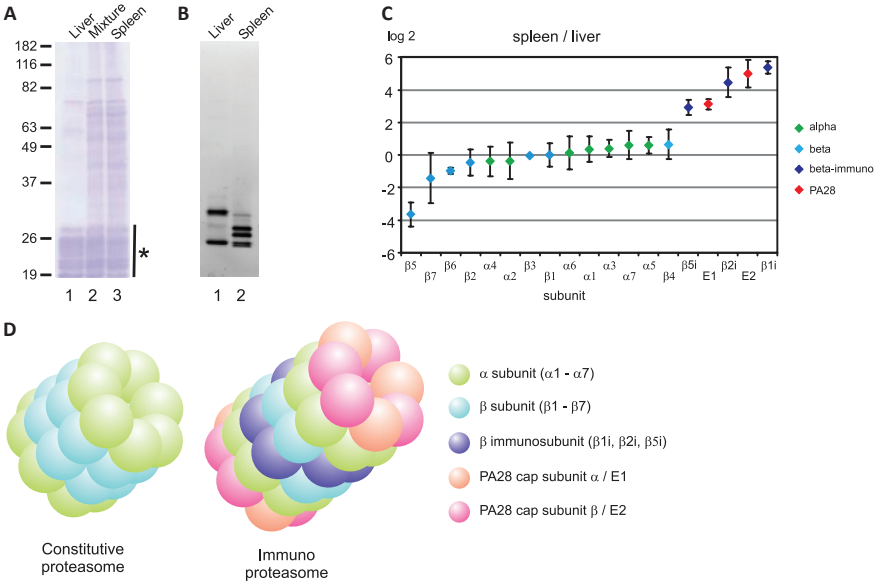


Figure 5 | Quantitation of differences between purified bovine liver and spleen proteasome complexes. A) the purified proteasome samples are shown after SDS-PAGE, stained using Coomassie Brilliant Blue. Lane 1 shows the purified liver proteasome, lane 2 the mixture of both proteasome samples, and lane 3 the purified spleen proteasome. The location of the proteasome proteins on the gel is indicated with an asterisk. **B)** the same proteasome purifications of liver (lane 1) and spleen (lane 2), but now stained using a fluorescent probe for catalytically active subunits. **C)** the relative quantitation of all Mascot identified proteasome subunits, with α-subunits indicated in green, β-subunits indicated in light blue (normal subunits), or dark blue (immunosubunits) and PA28 subunits indicated in orange. Bars indicate the standard deviation between the quantified peptides for each protein. **D)** the structure of the constitutive proteasome core and of the immunoproteasome core with the PA28 cap complex attached.

ed the procedure. Figure 4B shows that the quantified proteins nicely grouped into three ratios. With three exceptions, most likely caused by proteins being present in multiple slices of the gel (see "Discussion"), all proteins were again quantified within ±0.5 of the expected log 2 ratio (corresponding to ±41% of the actual ratios).

Compositional analysis of protein complexes

To test whether the method could be applied to analyze the differential composition of protein complexes, we purified the 20S core of the proteasome from bovine liver and spleen samples, using ammonium sulfate precipitation, sucrose gradient centrifugation, and DEAE separation (see "Materials and Methods"). Following this procedure, relatively pure proteasome samples were obtained from both tissues, as evidenced by SDS-PAGE and Coomassie Brilliant Blue staining of the purification (Figure 5A). To determine the presence and activity of catalytically active (β) subunits, the purified proteasomes were first incubated with a fluorescent probe followed by SDS-PAGE separation and fluorescence imaging. This revealed that both proteasome

purifications contained a number of active (β) subunits, but also that there were significant differences in the amount of constitutive proteasome and immunoproteasome between bovine liver and spleen samples, in agreement with previously reported data for mouse spleen and liver proteasomes.[11] After in-solution tryptic digestion, we used the online dimethylation method to further characterize and quantify these differences, with the liver proteasome being light labeled and the spleen proteasome being heavy labeled. All Mascot identified proteasome components in the sample were quantified using MS-Quant, and the data set was normalized on the median of the average of all α proteasome subunits, as these were expected not to change between the two proteasome samples. In Figure 5C, the base 2 logarithms of the measured ratios (intensity of the extracted ion chromatogram for the heavy labeled peptide divided by that of the light peptide) are plotted for all proteasome subunits. When using a cutoff value of 100% up- or down-regulation (which equals 1 unit on the log 2 scale), two proteins of the 20S proteasome core were more abundant in the proteasome purification from liver (subunits β5 and β7). Three core proteins were more abundant in the spleen sample (β5i, β2i, and β1i), corresponding exactly to the three catalytically active immunoproteasome subunits. In addition, two other 20S proteasome-associated proteins (PSME1 and PSME2) were found more abundantly in the spleen proteasome, which together can form an alternative cap of the proteasome that facilitates degradation of peptides for major histocompatibility complex (MHC) presentation,[12] called the PA28 (proteasome activator) or REG complex. Figure 5D displays a structural model of the differences between the regular constitutive proteasome and the immunoproteasome associated with the PA28 cap complex. To verify these results, the experiment was repeated with this time, the liver proteasome being heavy labeled and the spleen proteasome being light labeled. Although small differences were observed in the ratios calculated for the subunits having a log 2 ratio around 0, the same proteins were identified as being up- or down-regulated in the spleen sample, as compared with the liver proteasome (data not shown).

DISCUSSION

Quantitative proteomics by using stable isotope labeling in combination with mass spectrometry is an extensively used technique for the comparison of protein abundance in complex biological samples. To analyze tissue-derived samples, chemical labeling after digestion is the most often used method to determine relative quantities of peptides and proteins in these samples. As the efficiency of chemical labeling of samples is often sensitive to small (manual) errors and differences in sample preparation and handling, automation of the labeling process can reduce these errors and allow more reproducible results. However, care should still be taken in sample preparation to avoid human error prior to isotopic labeling as is the case with any other derivatization method.

Online labeling method

Here, we present a method for online, on-column, sequential derivatization of peptides, allowing relative quantitation of the peptides, by dimethylation using cyanoborohydride and either regular or deuterated formaldehyde.[7] The online quantitation procedure provided reliable quantitation over a wide dynamic range and allowed us to quantify differences between proteasome complexes purified from different tissue samples.

Online derivatization of peptides has been demonstrated before to enhance the frag-

mentation of the peptides in tandem mass spectrometry,[13] but for online sequential labeling, it is very important that the labeling reaction goes to completion to prevent bleeding of the samples. In our method, no intermediates or unmodified peptides were present at detectable levels, as demonstrated in Figure 2. Also, to prevent preferred binding of the first sample, the amount of peptide loaded on the trap column must be well below the capacity of the column. Therefore the method is especially suitable for analizing differences between single proteins or purified protein complexes, where the total amount of protein is unlikely to exceed the capacity of the column. When analyzing a mixture of ratios generated from bands cut from a gel containing mouse lung extract, we noticed a few proteins with unexpected ratios. Because some proteins might be present in multiple isoforms, with varying molecular weights, they might be present in more than one of the bands that were used for this experiments, causing the ratio to shift away from the expected value. This will not be a significant problem, however, as in most applications either in-solution digests or similar gel bands, but from different samples, will be compared. Evidently, in-solution digests will not resolve isoforms of the proteins,[14] but this would lead to inter-protein differences of quantified peptides rather than shifted ratios.

When automating the online labeling method, the stability of the reagent is another important issue. Using a colorimetric assay to determine cyanoborohydride concentrations[15] and by performing identical runs over a prolonged period of time, we found that the mixture of formaldehyde and cyanoborohydride was stable over at least 24 h when stored at 4 °C (supplemental Figure S4), whereas many other reagents used for chemical labeling (like iTRAQ) often required the reagent to be mixed immediately prior to the reaction. This stability makes the dimethylation reaction very suitable for the online labeling procedure, as it allows automated labeling and analyses of the samples to take place overnight or even longer period of time. Although the throughput speed of our method is higher than offline labeling when two samples are compared (Figure 1), offline labeling gets more favorable when larger numbers of samples have to be analyzed, as they can then all be labeled at the same time, rather than sequentially.

Compositional analysis of tissue-specific proteasomes

After comparing the 20S core proteasomes purified from bovine liver and spleen we observed that proteasome subunit β5 was significantly less abundant in spleen and to a lesser extent also β7 (although the standard deviation for the latter was high). The immunoproteasome subunits β1i, β2i, and β5i were all more abundantly present in the spleen-derived proteasome preparation. These differences between the proteasome complexes purified from bovine liver and spleen can be partially explained by the high abundance of immunoproteasome complexes in spleen, compared with liver, in full agreement with previous observations.[11] Still, it is remarkable that a single catalytic constitutive proteasome subunit is less abundant in the spleen-derived proteasome sample. This observation suggests that not only immunoproteasome complexes exist in which all constitutive active subunits are replaced by their immunoproteasome counterparts, but that "hybrid" proteasome complexes are present as well, in which only one of the β-subunits has been replaced by an immunosubunit. Such hybrid proteasome complexes have been suggested before, but their exact composition remains so far unclear.[16] Also, because only β5 was

relatively less abundant, and all three immunosubunits were more abundant, it cannot be excluded that β5 might be replaced by either β1i, β2i, or β5i in such a hybrid proteasome species.

Next to immunoproteasome subunits, we also found PSME1 and PSME2 to be almost uniquely present in the bovine spleen purified proteasomes. Together, these two highly homologous proteins form the alternative PA28 cap (also called 11S regulator) of the 20S proteasome, a heptameric ring containing most likely 4 subunits of PSME2 and 3 of PSME1.[12,17] The PA28 cap can replace the 19S regulator and is, like the catalytically active immunosubunits, up-regulated by interferon-γ.[18] The fact that we found the 11S complex in the proteasome purified from spleen, but did not find any subunits of the 19S complex, might suggest that the 11S regulator is more tightly bound to the proteasome core as the 19S cap is known to dissociate under the isolation conditions.

CONCLUSION

Here, we presented an online, sequential labeling procedure using differential stable isotopic dimethlylation of N termini and lysine residues. The method allows automated sequential labeling and LC-MS/MS analysis of complex biological samples. We validated the method on model proteins and a whole cell lysate and identified the differences in composition between proteasome complexes purified from bovine liver and spleen. The method is efficient, performs similar to other chemical labeling methods, and can be generally applied. The method is especially suitable for the quantitative analysis of gel bands or purified protein complexes, and its full automation permits easier analysis of large sample numbers with a greatly reduced chance of inaccuracy due to human error.

MATERIALS AND METHODS

NanoLC-MS/MS

All analyses were performed on nanoLC-LTQ-Orbitrap at a resolution of 60,000 or nanoLC-LTQ-FTICR at a resolution of 100,000. (Thermo, San Jose, CA) mass spectrometers. For nanoLC, Agilent 1100 series LC systems were equipped with 20-mm Aqua C18 (Phenomenex, Torrance, CA) trapping columns (packed in-house, i.d., 100 μm; resin, 5 μm) and 250 mm ReproSil-Pur C18-AQ (Dr. Maisch GmbH, Ammerbuch) analytical columns (packed in-house, i.d., 50 μm; resin, 3 μm). Solvents used where 0.6% HAc (buffer A) and 0.6% HAc/80% acetonitrile (ACN) (buffer B). Trapping was performed at 5 μL/min for 10 min, and elution was achieved with a gradient of 0–45% B in 45 min, 45–100% B in 1 min, 100% B for 4 min. The flow rate was passively split from 0.36 mL/min to 100 nL/min. Nanospray was achieved using a distally coated fused silica emitter (New Objective, Cambridge, MA) (outer diameter, 360 μm; i.d., 20 μm, tip i.d. 10 m) biased to 1.8 kV. The mass spectrometer was operated in the data-dependent mode to automatically switch between MS and MS/MS. Survey full scan MS spectra were acquired from m/z 350 to m/z 1500, and the three most intense ions were fragmented in the linear ion trap using collisionally induced dissociation. The target ion setting was 5e5 for the Orbitrap, with a maximum fill-time of 250 ms and 1e6 for the FTICR, with a maximum fill-time of 250 ms. Fragment ion spectra were acquired in the LTQ with a target ion setting of 3e4 and a maximum fill-time of 500 ms. Dynamic exclusion for selected precursor ions was set at 30 s.

Online labeling

All online labeling steps were performed at a flow speed of 5 μL/min with a continuous flow of buffer A. All reagents were injected with the autosampler using short injection programs of 5–20 min, depending on the injected volume. After loading the trap column with the first sample in 5% formic acid (FA), light isotope labeling was performed by flushing the trap column with 40 μL of 0.04% formaldehyde and 6 mM cyanoborohydride in 50 mM sodium phosphate buffer, pH 7.5. After washing the trap column with 10 μL of 5% FA, the second sample was loaded. Heavy

isotope labeling was then performed using 40 μL of 0.04% deuterated formaldehyde and 6 mM cyanoborohydride in 50 mM sodium phosphate buffer, pH 7.5. Finally, after washing with 40 μL of 5% FA, regular LC/MS analysis was performed as described above.

Proteasome purification

Proteasome samples were purified from bovine liver and spleen. The purification was monitored in each step using fluorescent probes for proteasome activity, followed by SDS-PAGE as has been described before.[11] After homogenization of the tissues in phosphate buffered saline, the extracts were clarified by centrifugation (27000 g, 4 °C, 50 min.). After precipitating unwanted proteins using 40% saturated ammonium sulfate and centrifugation, the proteasome-containing fraction was precipitated by increasing the ammonium sulfate concentration to 60% saturation. After centrifugation, precipitated proteins were dissolved in 10 mM Tris-HCl (pH 7.4) and dialyzed against the same buffer to remove all ammonium sulfate. The dialyzed solutions were further purified using a 10–40% sucrose gradient at 28,000 rpm at 4 °C for 16 h. Fractions containing liver and spleen proteasome activity were pooled for DEAE (diethylaminoethyl) separation (DEAE-Sephadex A25 and DEAE-Sephacel resin, respectively). After incubating the proteins with the DEAE resin for 30–60 min (80 mM potassium acetate buffer containing 20 mM Tris-HCl, pH 7.2, and 5 mM magnesium acetate for the Sephadex resin and 10 mM Tris-HCl, pH 8.0 containing 1 mM EDTA for the Sephacel resin), the resin was washed, and proteins were eluted using and increasing NaCl concentrations. All proteasome-containing fractions were pooled and concentrated by means of a nitrogen filter and protein concentrations determined using the Bradford assay (Bio-Rad). Proteasome preparation were stored at -80 °C. To check the purity of the proteasome preparations, samples were freeze-dried, boiled in reducing sample buffer, and analyzed by 12.5% SDS-PAGE using a PROTEAN II xi Cell system (Bio-Rad) followed by Coomassie Brilliant Blue staining. To visualize the active subunit composition of liver and spleen proteasome samples, 1 μg of proteasome was incubated with 500 nM of Me_4BodipyFL probe (a close analog of the bodipyFL-probe described previously) for 1 h at 37 °C. Proteins were denatured by being boiled in reducing sample buffer and analyzed by 12% SDS-PAGE using the NuPAGE pre-cast gel system (Invitrogen). The gel was then scanned for fluorescence emission using a ProXPRESS two-dimensional Proteomic imaging system (Perkin Elmer).

Trypsin digestion

Whole tissue lysates were separated by 12% SDS-PAGE followed by staining with Coomassie Brilliant Blue. Gel lanes were cut into slices, which were washed with MilliQ and 0.6% HAc/80% acetonitrile (ACN). Prior to in-gel digestion, proteins were reduced with 1,4-dithiothreitol (6.5 mM) and alkylated with iodoacetamide reagent (54 mM). After thorough washing, pieces were rehydrated in trypsin solution (10 ng/l) on ice. After addition of 30 ml of NH_4HCO_3 (50 mM, pH 8.5), samples were digested for 16 h at 37 °C. Supernatant of the digest was collected. The gel pieces were washed for 30 min in 5% formic acid at room temperature, after which the supernatant of this washing step was combined with the earlier fraction and stored at 30 °C until the analysis. Purified proteins and proteins complexes were digested in-solution after reduction and alkylation. Digestion was performed using 50 mM NH_4HCO_3 at 37 °C for 16 h, with a protein/trypsin ratio of 1:100 by weight. Prior to injection, samples were diluted in MilliQ water containing 5% formic acid.

Data processing and analysis

Raw MS data were converted to peak lists using Bioworks Browser software, version 3.1.1. For protein identification, MS/MS data were compared with the International Protein Index murine (v3.36, 51424 entries searched) or bovine (v3.22, 32915 entries searched) data bases (depending on the sample analyzed) using Mascot Version 2.1 (Matrix Science) with trypsin at the enzyme, allowing 1 missed cleavage. Precursor and fragment mass tolerances were set at 6 ppm and 0.9 Da deviation, respectively. As fixed modifications, carbamidomethylation (Cys) and dimethylation (Lys, N-terminal) were set as well as the following variable modifications: dimethylation: 2H(4) (Lys, N-terminal) and Oxidation (Met). Proteasome proteins were identified with a minimal Mascot protein score of 48 using at least two identified peptides with a Mascot ion score of >35 and an

expect value of <0.005. At these settings, the false discovery rate was 0.68% as determined using a decoy data base. The only exception was the protein PSME2, which was identified with only one peptide (Mascot peptide score 48) fulfilling those criteria. The Mascot MS/MS annotation of that peptide is supplied in Figure S5. Relative peptide abundance was calculated from extracted ion chromatograms of the different isotopic variants using MS-Quant software.[19] Very low intensity peptides (extracted ion chromatogram intensity < 3000) as well as peptides with a poor extracted ion chromatogram or MS/MS spectrum were excluded after manual inspection. The program StatQuant[20] was used to normalize the quantitation and calculate standard deviations of the log 2 of the ratios of all quantified peptides per protein. Where appropriate, quotation experiments were repeated with the light and heavy labels switched.

ACKNOWLEDGMENTS

We thank B. Rodenko and J. Neefjes (Netherlands Cancer Institute, Amsterdam, The Netherlands) and T. van de Goor and K. Kraiczek (Agilent Technologies, Waldbronn, Germany) for advice. This work was supported by the Netherlands Proteomics Centre, Dutch Cancer Society (Koningin Wilhelmina Fonds) Grant 2005-3368 and the Agilent Technologies Foundation. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

SUPPLEMENTARY INFORMATION ONLINE

More examples of BSA peptides before and after double labeling can be found in supplementary Figure S1. All quantitation results shown are listed in tables S1–S4. All identification results have been uploaded to PRIDE (under project title "Sequential Labeling for Protein Quantitation") and identification details for the proteasome proteins are supplied in table S5. This material is available via the Internet at http://www.mcponline.org, as supporting information of *Molecular & Cellular Proteomics* **7**, 1755-1762.

REFERENCES

1. Julka, S. & Regnier, F.E. Recent advancements in differential proteomics based on stable isotope coding. *Brief. Funct. Genomic. Proteomic.* **4**, 158–177 (2005).
2. Heck, A.J. & Krijgsveld, J. Mass spectrometry-based quantitative proteomics. *Expert Rev. Proteomics* **1**, 317–326 (2004).
3. Krijgsveld, J. et al. Metabolic labeling of C. elegans and D. melanogaster for quantitative proteomics. *Nat. Biotechnol.* **21**, 927–931 (2003).
4. Ong, S.E. et al. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell. Proteomics* **1**, 376–386 (2002).
5. Shiio, Y. & Aebersold, R. Quantitative proteome analysis using isotope-coded affinity tags and mass spectrometry. *Nat. Protoc.* **1**, 139–145 (2006).
6. Ross, P.L. et al. Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. *Mol. Cell. Proteomics* **3**, 1154–1169 (2004).
7. Hsu, J.L., Huang, S.Y., Chow, N.H., & Chen, S.H. Stable-isotope dimethyl labeling for quantitative proteomics. *Anal. Chem.* **75**, 6843–6852 (2003).
8. Makarov, A. et al. Performance evaluation of a hybrid linear ion trap/orbitrap mass spectrometer. *Anal. Chem.* **78**, 2113–2120 (2006).
9. Peterman, S.M., Dufresne, C.P. & Horning, S. The use of a hybrid linear trap/FT-ICR mass spectrometer for on-line high resolution/high mass accuracy bottom-up sequencing. *J. Biomol. Tech.* **16**, 112–124 (2005).
10. Meiring, H.D., van der Heeft, E., ten Hove, G.J. & de Jong, A.P.J.M. Nanoscale LC-MS(n): technical design and applications to peptide and protein analysis. *J. Sep. Sci.* **25**, 557–568 (2002).
11. Berkers, C.R. et al. Profiling proteasome activity in tissue with fluorescent probes. *Mol. Pharm.* **4**, 739–748 (2007)
12. Zhang, Z. et al. Proteasome activator 11S REG or PA28: recombinant REG alpha/REG beta hetero-oligomers are heptamers. *Biochemistry* **38**, 5651–5658 (1999).
13. Cardenas, M.S., van der Heeft, E. & de Jong, A.P. On-line derivatization of peptides for improved sequence analysis by microcolumn liquid chromatography coupled with electrospray ionization-tandem mass spectrometry. *Rapid Commun. Mass Spectrom.* **11**, 1271–1278 (1997).
14. Kolkman, A., Dirksen, E.H., Slijper, M. & Heck, A.J. Double standards in quantitative proteomics: direct comparative assessment of difference in gel

electrophoresis and metabolic stable isotope labeling. *Mol. Cell. Proteomics* **4**, 255–266 (2005).
15. Sorensen, K. Coomassie protein assay reagent used for quantitative determination of sodium cyanoborohydride (NaCNBH3). *Anal. Biochem.* **218**, 231–233 (1994).
16. Dahlmann, B., Ruppert, T., Kloetzel, P.M. & Kuehn, L. Subtypes of 20S proteasomes from skeletal muscle. *Biochimie* **83**, 295–299 (2001).
17. Knowlton, J.R. et al. Structure of the proteasome activator REGalpha (PA28alpha). *Nature* **390**, 639–643 (1997).
18. Tanahashi, N. et al. Molecular properties of the proteasome activator PA28 family proteins and gamma-interferon regulation. *Genes Cells* **2**, 195–211 (1997).
19. Schulze, W.X. & Mann, M. A novel proteomic screen for peptide-protein interactions. *J. Biol. Chem.* **279**, 10756–10764 (2004).
20. Van Breukelen, B., van den Toorm. H., Drugan, M. & Heck, A.J.R. In preparation.

Chapter 2.4

Summary and prospects: development and use of chemical proteasome probes

Summary and prospects: development and use of chemical proteasome probes

The ubiquitin-proteasome system is responsible for the degradation of proteins involved in a wide variety of cellular processes, including cell proliferation and survival, cell-cycle control, transcriptional regulation and cellular stress responses, as well as for the destruction of redundant and misfolded proteins. The ubiquitin proteasome system is an attractive pharmacological target for the treatment of cancer, because of the higher sensitivity of malignant cells to proteasome inhibition compared to normal cells, and the proteasome inhibitor bortezomib has been approved for the treatment of multiple myeloma (MM) and mantle cell lymphoma. Bortezomib treatment is however associated with substantial adverse effects, and the occurrence of both primary and secondary resistance has been reported. Consequently, there is an ongoing search for novel proteasome inhibitors that are able to overcome bortezomib resistance and that are better tolerated to increase their therapeutic potential. Several second-generation proteasome inhibitors, which differ in their mode of inhibition and subunit specificity, have entered clinical trials, including marizomib and CEP-18770.

Techniques that are commonly used to monitor proteasome activity include the application of fluorogenic substrates, small molecule-based activity assays and models based on recombinant reporter proteins. Traditionally, fluorogenic substrates are used to measure the activity of the different proteasome active sites, but most fluorogenic substrates cannot be used in cells, and prior cell lysis is required before activity measurements can be performed. Reporter proteins can be used in living cells, but their use is limited to genetically altered cells or organisms and therefore does not allow profiling of patient material. In addition, their activity readout depends on the balance between synthesis and degradation of fusion proteins, which involves many cellular factors other than the proteasome, including the rate of fusion-protein synthesis. With proteasome inhibitors in use in the clinic and with clinical trials in progress investigating the treatment of several proteasome inhibitors for a variety of hematologic and solid malignancies, accurate methods that allow profiling of proteasome inhibitor specificity and efficacy in patients are in demand.

In **chapter 1**, the development of chemical proteasome probes that provide such methods is described. In **chapter 1.2**, we describe the first small molecule-based activity assay in its class that could be used to profile the specificity of proteasome inhibitors in living cells. This dansylated vinylsulfone based proteasome probe contains an α,β-unsaturated sulfone part that reacts through a Michael addition reaction with the γ-hydroxy moiety of the N-terminal threonine residue of catalytic β-subunits of the proteasome, resulting in the formation of a β-sulfonyl ether link-

age. Antibodies against the dansyl moiety are used for detection of labeled active subunits by SDS-PAGE and Western blot analysis. Using this probe, we investigated the *in vivo* subunit specificities of bortezomib and another inhibitor, MG132. When we compared the specificities in cultured cells and in crude cell extracts, we observed substantial differences in activity patterns, stressing the importance for bioassays compatible with live cells to ensure accuracy of such measurements. In **chapter 1.3**, we describe the further optimization of this probe by replacement of the environment sensitive and low quantum yield dansyl fluorophore by high quantum yield fluorophores that allowed for direct scanning of the SDS-PAGE gel for fluorescence emission of fluorescently labeled subunits. In order to be able to select for effective fluorescent proteasome inhibitors with good biochemical and biophysical properties, we embarked on the synthesis of a range of fluorescent inhibitors. Two differentially labeled probes tested proved suitable to study proteasome activity in cell extracts, living cells and murine tissues using a range of techniques including SDS-PAGE, confocal laser scanning microscopy and flow cytometry.

In **chapter 2**, these fluorescent proteasome probes were used to study to the proteasome inhibitory profiles of two second-generation proteasome inhibitors, CEP-18770 and marizomib, in a preclinical model of MM and in patients in clinical trials, respectively. In **chapter 2.1**, we compared bortezomib and CEP-18770 in a variety of *in vitro* and preclinical *ex vivo* models. In a panel of cell lysates and cell lines, both proteasome inhibitors were equipotent in inhibiting distinct subunits of the proteasome, with proteasome activity recovering similarly after withdrawal of either inhibitor. In contrast, in a preclinical MM model, CEP-18770 inhibited tumor proteasome activity to a significantly higher extent compared to bortezomib, while the two inhibitors had an equal capacity to inhibit the proteasome in normal tissues. In addition, we showed that CEP-18770 was able to inhibit proteasomes and overcome bortezomib resistance in cells with acquired bortezomib resistance *in vitro*. Marizomib is currently being evaluated in phase I clinical trials for the treatment of both solid and hematologic tumors, enabling a detailed study of patient response to marizomib treatment. In **chapter 2.2**, we validated the use of a fluorescent proteasome activity probe for profiling proteasome inhibition in both packed whole blood (PWB) and peripheral blood mononuclear cells (PBMCs) using marizomib- or bortezomib-treated rats. In addition, we developed a protocol that enables the use of this fluorescent probe to profile blood samples from patients that are treated with proteasome inhibitors. We showed that PWB, which contains >98% red blood cells, responded differently to proteasome inhibitors and recovered slower from proteasome inhibition compared to PBMCs. These data caution against the use of PWB as a model system to monitor the effects of proteasome inhibitors in animals or patients, as is routinely done. In addition, we used the fluorescent activity-based probe to demonstrate that both initial proteasome activity level in cells and the ability of cells to change their proteasomal subunit composition in response to marizomib treatment may be important determinants of sensitivity to proteasome inhibitors.

Collectively, the findings presented in **chapter 2** demonstrate that chemical proteasome activity probes can reliably assay proteasome activity *in vitro* and *ex vivo* in a broad range of tissue types, both in preclinical models and during clinical trials. Furthermore, these

probes may enable prediction of patient responses to proteasome inhibitors, which can contribute to a more personalized proteasome inhibition strategy to improve the therapeutic potential of this class of drugs.

We expect that the development of second-generation proteasome inhibitors will likely result in the generation of inhibitors that can overcome resistance and may ultimately contribute to better patient outcome. Intravenously administered proteasome inhibitors likely saturate proteasomes in red blood cells, which predominantly display constitutive activity, before peripheral sites are reached. Specific immunoproteasome inhibitors may saturate blood proteasome at relatively low concentrations, resulting in higher inhibitor concentrations in tumor cells. This suggests that immunoproteasome inhibitors may in future provide a good alternative for tumors that have high intrinsic immunoproteasome activity. An interesting strategy to overcome resistance to the current generation of proteasome inhibitors may be the development of inhibitors that do not interfere with the 20S core, but with other elements of the UPS. Experiments setting up a high-throughput flow cytometry-based assay that can be used to measure both proteasome activation and proteasome inhibition in both cultured and primary cells are currently ongoing. This assay allows screening of large compound or siRNA libraries to identify either direct regulators of proteasome activity, such as inhibitors of the 20S core, or modulators that interfere elsewhere in the ubiquitin-proteasome pathway. This assay is currently optimized for the flow cytometry assisted measurement of overall proteasome inhibition in tumor cells originating from complex patient samples. We expect that our flow-cytometry based proteasome activity assay will not only identify proteasome inhibitors, but also activators of the ubiquitin proteasome system. These activators may find use in the treatment of neurodegenerative diseases, which are often characterized by a decrease in intracellular proteasome activity. In addition, the development of proteasome inhibitors that are specific for proteasome subtypes or bacterial proteasome may broaden the clinical application of UPS-modulating compounds in cancer, infectious- and autoimmune diseases.

Chapter 3

Proteasome splicing

Chapter 3.1

Transpeptidation and reverse proteolysis and their consequences for immunity

International Journal of Biochemistry and Cell Biology
Volume 41, 2009, Pages 66-71

Transpeptidation and reverse proteolysis and their consequences for immunity

Celia R. Berkers,[§] Annemieke de Jong,[§] Huib Ovaa[*] & Boris Rodenko

Division of Cellular Biochemistry, The Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands.

Reverse proteolysis and transpeptidation lead to the generation of polypeptide sequences that cannot be inferred directly from genome sequences as they are post-translational phenomena. These phenomena have so far received little attention although the physiological consequences may reach far. The protease mediated synthesis of several immunodominant MHC class I antigens was recently reported, underscoring its importance to immunity. Reverse proteolytic and transpeptidation mechanisms as well as conditions that favor successful protease-catalyzed synthetic events are discussed here.

INTRODUCTION

A broad repertoire of antigenic peptides is presented at the cell surface to the immune system by major histocompatibility complexes (MHCs) class I and class II, which are known in humans as human leukocyte antigens (HLAs) class I and class II. Intracellular proteins form the source of antigens that are presented by MHC class I, enabling the immune system to monitor changes in intracellular protein content and sense malignant transformation or the presence of viral proteins. Recognition of these foreign or abnormal peptides by CD8+ cytotoxic T cells ultimately results in the elimination of the antigen presenting cell.[1,2] MHC class I antigens are primarily generated in the cytosol, where proteins are initially degraded by the proteasome into smaller fragments. In order to fit the MHC class I binding groove, these peptides undergo further N-terminal trimming by various aminopeptidases,[3,4] that include tripeptidylpeptidase II (TPPII), bleomycin hydrolase, thymet oligopeptidase, puromycin sensitive aminopeptidase, neurolysin, and leucine aminopeptidase, although many other proteases may be involved in this trimming process. After processing, these peptides are transported into the endoplasmic reticulum (ER), where some can be loaded onto MHC class I directly, while others may require additional N-terminal trimming by ER-associated aminopeptidase (ERAAP).[5]

Antigens generated from exogenous pathogens are presented via the MHC class II pathway, which leads to the activation of CD4+ T helper cells. Following endocytosis by professional antigen presenting cells, exogenous proteins are degraded in early endosomes, late endosomes and lysosomes in a pH-dependent manner. Association of thus formed peptides with MHC class II can take place in any of these compartments,[6] but occurs mainly in late endocytic structures called MIIC compartments. Cysteine proteases, including cathepsins S, L, B, F, H and V as well as aspar-

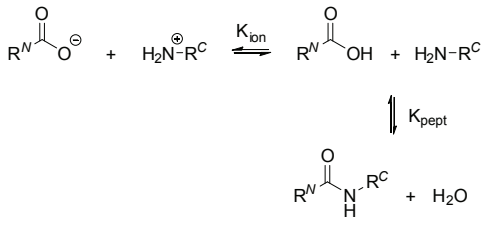
[§]These authors contributed equally to this work. [*]Correspondence should be addressed to H.O. (h.ovaa@nki.nl).

agine endopeptidase (AEP), have mainly been implicated in antigen processing, but various metallo-, aspartic acid- and serine-proteases have been shown to be involved in MHC class II antigen processing as well.[6]

Recently, it has become apparent that not only contiguous antigens, but also non-contiguous antigens, that consist of two post-translationally fused peptides, can be presented to the immune system.[1,7,8] The ligation that produces these epitopes is thought to be catalyzed by proteases *in vivo*. Here, we summarize how and under which conditions proteases catalyze peptide ligation, as well as the consequences of this process for immunity.

PROTEASE-CATALYZED PEPTIDE LIGATION

Based on their proteolytic mechanism, proteases can be divided into five classes,[9,10] and broadly form two groups: cysteine, serine, and threonine proteases, that hydrolyze peptide bonds via the formation of an acyl-enzyme intermediate (Figure 1), and aspartic- and metalloproteases, that act via the direct activation of a water molecule that acts as the primary hydrolyzing species. Since proteases are catalysts, they alter the rate at which the thermodynamic equilibrium of the protein hydrolysis reaction is reached, but they do not change the overall thermodynamic equilibrium between starting materials and proteolytic products.[11,12] Consequently, all proteases are also able to catalyze the reverse reaction, *i.e.* peptide bond synthesis, provided that the reaction conditions favor this process, which is referred to as reverse proteolysis (Scheme 1). Furthermore, proteases that act via an acyl- enzyme intermediate can catalyze peptide bond formation via a transpeptidation mechanism, which competes with hydrolysis (Figure 1).



Scheme 1 | Mechanism of reverse proteolysis. R^N and R^C refer to the N-terminal and C-terminal ligation partners, respectively. K_{ion} and K_{pept} refer to the ionization and peptidation equilibrium constants, respectively.

REVERSE PROTEOLYSIS

Peptide ligation by reverse proteolysis was first reported in the 1930s, when papain and chymotrypsin, a cysteine and serine protease, respectively,were used as catalysts for enzymatic synthesis.[13-16] Reverse proteolysis involves the direct reversal of catalytic hydrolysis by proteases, as shown in Scheme 1, in which K_{ion} is the equilibrium constant of ionization and K_{pept} is the equilibrium constant of peptidation.[11,12,17] In an aqueous environment, the reaction will be driven in the direction of hydrolysis.[11] Furthermore, the ligation of peptide fragments is an endergonic process (entropy is lost) and is therefore energetically unfavorable.[17] One strategy to direct the equilibrium towards ligation is based on the law of mass action.[12,17,18] *In vitro*, product removal by precipitation,[19-22] the use of biphasic systems,[23] the use of a large molar excess of one of the substrates,[11,17] or a conformational trap[22,24,25] have all been shown to promote ligation. Alternatively, the equilibrium of ionization can be manipulated to enhance ligation.[12,17,18] Peptide bond formation is dependent on the concentration of uncharged substrates formed in the ionization equilibrium, which is dependent on their pK_a values. Therefore, the pH optimum for ligation lies in the range between the pK_a of the

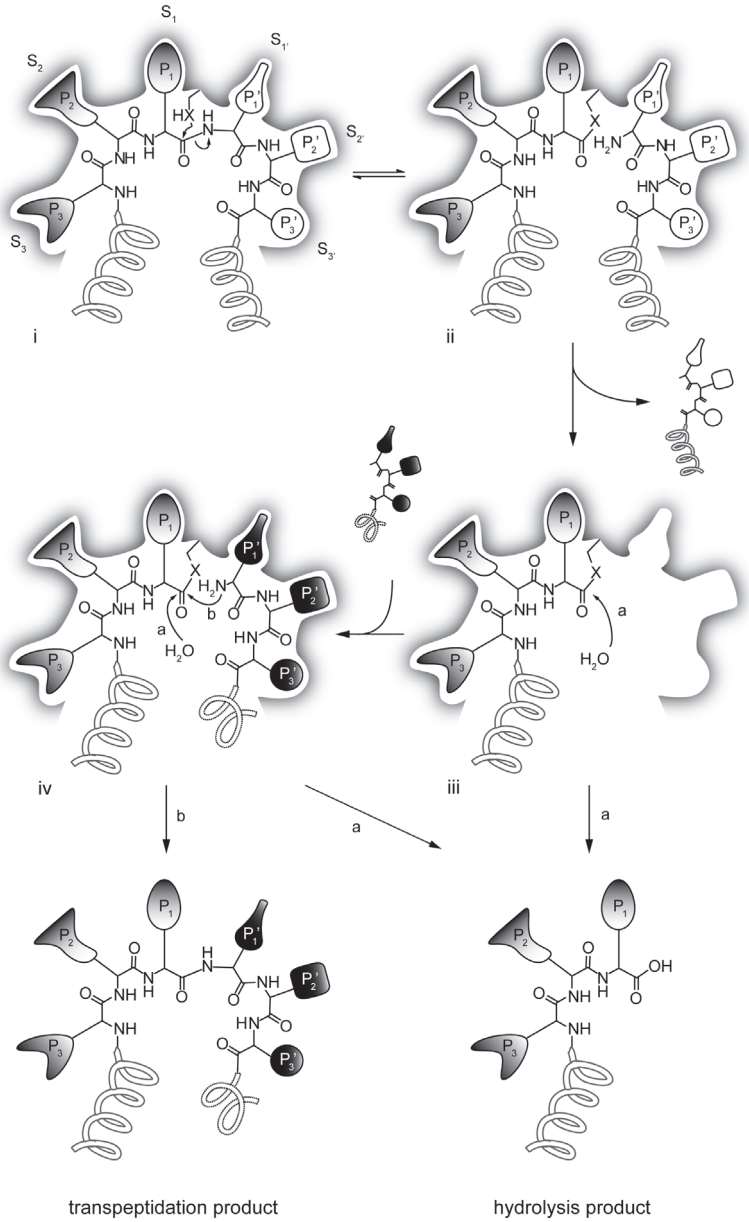


Figure 1 | General mechanism of protease mediated hydrolysis and transpeptidation via an acyl-enzyme intermediate. (i) The protein or peptide substrate with its P_1, P_2, P_3 and P_1', P_2' and P_3' residues docks into the complementary S_1, S_2, S_3 and S_1', S_2' and S_3' pockets of the protease. (ii) Nucleophilic attack of the reactive residue on the enzyme results in formation of the acyl-enzyme intermediate. (iii) The cleaved C-terminal substrate fragment leaves the active site and hydrolysis of the acyl-enzyme intermediate can take place - path a - under release of the N-terminal peptide or protein fragment. (iv) Alternatively, the vacated S_1', S_2' and S_3' binding sites are subsequently occupied by a protein or peptide fragment containing complementary P_1', P_2' and P_3' residues. Aminolysis by the free N-terminus - path b - results in net transpeptidation, and competes with hydrolysis - path a.

α-carboxylic group of the N-terminal component and the amino group of the C-terminal substrate, which is normally found between pH 6 and 7.[17] Several water-miscible organic co-solvents can reduce the acidity of the α-carboxy group of the N-terminal component, which shifts K_{ion} in favor of ligation, and have therefore been used *in vitro* to influence the ionization equilibrium.[18] Furthermore, organic co-solvents reduce the activity of water by lowering its concentration, thereby hampering the hydrolysis reaction.[11]

Under physiological conditions, the equilibrium as shown in Scheme 1 favors hydrolysis and ligation is thought by many to be negligible *in vivo*.[11] It has been postulated, however, that macromolecular crowding, defined by the high total concentration of background molecules, such as lipids, proteins, nucleic acids and carbohydrates that are present in cells, shifts the equilibrium towards ligation *in vivo*.[26,27] Due to non-specific steric repulsions, the presence of a significant volume fraction of macromolecules places constraints on the placement of additional macromolecules.[28] Macromolecules therefore exclude volume to each other. This excluded volume effect reduces the configurational entropy,[29] which is energetically unfavorable. Therefore, processes that are accompanied by a reduction in excluded volume, such as macromolecular compaction and association, are facilitated in a macromolecularly crowded environment.[26-29] Consequently, molecular crowding has been shown to promote peptide ligation[24,25,30,31] as well as the synthesis of proteins from partly structured large polypeptides,[32,33] in cases where ligation led to a large reduction in excluded volume. This suggests that *in vivo*, proteases might also catalyze peptide bond synthesis in addition to their ability to function as hydrolases.

TRANSPEPTIDATION

Transpeptidation occurs when the rate of aminolysis is faster than or comparable to the rate of hydrolysis, which is dependent on the competition between nucleophilic components and water for the nucleophilic attack on the acyl-enzyme intermediate (Figure 1). *In vitro*, the use of water-miscible organic co-solvents or biphasic systems can help to reduce the water content and therefore favor aminolysis over hydrolysis.[34-37] High reactant concentrations[36,38,39] and a high pH,[39,40] preferably higher than the pK_a of the attacking nucleophile to ensure its unprotonated state,[11,41] have also been shown to promote transpeptidation. Evidently, the specificity of the protease for the reactants, is also a crucial factor[17] as its active site has to be able to accommodate reactants for productive reactions. Both *in vitro* studies, in which different amino acids or peptides were used as nucleophiles,[38,39,42-45] and an *in vivo* study[46] (see below) indicate that ligation efficiencies are highly dependent on the fit of the amino acids at the P_1', P_2' and/or P_3' positions in the S' binding sites (Figure 1).

Examples of protease-catalyzed transpeptidation occurring *in vivo* are scarce, but few examples have been reported. Asparaginyl endopeptidases (AEPs) have been implicated in the catalysis of two distinct transpeptidation events inplants: (1) the post-translational processing of the lectin concanavalin A (con A) in maturing jack beans[47,48] and (2) the biosynthesis of the backbone of plant-produced cyclotides,[46,49] the largest family of circular proteins that are notoriously stable, due to the combination of a cyclic backbone and a cysteine knot of three disulfide bonds.

During cleavage of pro-con A into smaller fragments, two of these fragments are ligated in the reverse order by AEP in a transpeptidation reaction to form mature con A.[47] Both the

increased stability of mature con A compared to its unligated precursor[48] and the occurrence of a proximity effect may explain why transpeptidation is favored over hydrolysis *in vivo*. If protein complexation or conformation positions the amine-donating component involved in transpeptidation in close proximity to the acyl ester component, this will facilitate its nucleophilic attack on the acyl-enzyme intermediate and hence favor aminolysis. Although such an effect has never been implicated in con A transpeptidation, the proposed structure of pro-con A[50] indicates this is likely the case.

Plant-produced cyclotides exhibit a broad range of bioactivities, amongst which are antimicrobial-, hemolytic-, anti-HIV- and insecticidal activities.[46,49] Although the mechanism is not fully understood, cyclization of these proteins *in vivo* is thought to be the result of an AEP-catalyzed transpeptidation reaction.[46,49] The cysteine knots are believed to keep the amine-donating and carboxy-donating components involved in transpeptidation in close proximity to each other, indicating that a proximity effect facilitates ligation *in vivo*. Furthermore, a tripeptide binding motif at the N-terminus of the AEP bound propeptide binds, after release of the cleaved C terminal propeptide fragment, to the vacated S_1', S_2' and S_3' subsites, which facilitates cyclization.[46] This suggests that substrate specificity further enhances transpeptidation.

The proteasome, a threonine protease, has been shown to catalyze transpeptidation during the production of antigens for presentation by MHC class I. To date, three non-contiguous immunodominant antigens, formed by the ligation of two peptides from the same parental protein during proteasomal degradation, have been identified. The first non-contiguous antigen, which resulted from the fusion of residues 172–176 and 217–220 from fibroblast growth factor-5 was identified from a patient with renal cell carcinoma.[8] Vigneron et al. identified a second non-contiguous tumor antigen, composed of residues 40–42 and 47–52 of the melanocytic glycoprotein gp100, in a melanoma patient.[7] Subsequently, an antigen was identified in a recipient of MHC-matched hematopoietic cell transplantation, consisting of two peptides from the SP110 nuclear body protein, ligated in the reverse order.[1] Proteasomal involvement in the ligation events was suggested,[1,7,8] and *in vitro* digestion experiments with purified 20S proteasome confirmed this notion.[1,7] Further experiments showed that both cleavage of the carboxy component and a free N-terminus at the amine-donating reaction component were required for ligation, indicating that splicing in the proteasome occurs via a transpeptidation mechanism.[1,7] In most cytosolic proteases the reaction products diffuse away too quickly under physiological conditions to promote ligation. However, in barrel-shaped, self-compartmentalizing proteases such as the proteasome,[51] reaction partners are generated *in situ* and remain confined to a small volume, resulting in their high local concentrations and a high degree of substrate specificity, which likely facilitates peptide ligation *in vivo*.

CONSEQUENCES FOR IMMUNITY

Several reports showthat transpeptidation occurs readily under physiological conditions *in vitro*, during digestion of proteins for structural analysis. Trypsin,[52-56] Staphylococcus aureus V8 protease,[52,56] and Lys-C endoprotease[56] were all shown to catalyze transpeptidation. Based on their observations, Fodor and Zhang stated that transpeptidation occurs more often than was known previously.[56] Schaefer et al. emphasized that available database search algorithms fail to detect transpeptidation products, as they cannot be inferred from the

genome.[55] *In vivo*, the extent to which proteases catalyze transpeptidation rather than hydrolysis is fully dependent on specific local conditions. γ-Glutamyltranspeptidases, for example, which bind glutathione and function predominantly as hydrolases *in vitro*,[57] are assumed by many to mainly catalyze transpeptidation reactions *in vivo* due to the presence of high concentrations of amino acids in kidney as amine-donating reaction components.[58,59]

The conditions promoting ligation described above, including molecular crowding, proximity effects, a high degree of substrate specificity or confinement can certainly be met by proteases other than AEPs and the proteasome, suggesting that protease-catalyzed peptide ligation represents a more widespread and common mechanism than currently assumed. Macromolecular crowding in the cytosol may facilitate ligation of non-contiguous polypeptides to form novel proteins. During proteasomal degradation of cytosolic proteins into peptides, transpeptidation produces spliced and reordered antigens.[1,7,8] Notably, both bleomycin hydrolase[60.61] and TPPII,[62,63] peptidases that process these antigens downstream of the proteasome, are thought to assemble into self-compartmentalizing proteases, which may facilitate transpeptidation. Bleomycin hydrolase has been shown to be able to convert itself into a peptide ligase *in vitro*.[64] This suggests that transpeptidation is likely to occur during MHC class I epitope trimming downstream of the proteasome. Furthermore, functional homologs of mammalian AEP, which is actively involved in MHC class II antigen production, have been shown to catalyze transpeptidation in plants.[46,49]

In our current understanding of the cellular antigen presenting pathways, proteins are cleaved in contiguous sequences,[65] creating antigens that can traditionally be inferred from genomic sequences, and in which reverse proteolysis and transpeptidation have not been taken into account. The repertoire and diversity of both foreign and self antigens in the MHC routes may thus be significantly broader than currently thought. Amplifying the diversity of epitopes increases the chance that one or more of these epitopes are recognized by CD4+ and CD8+ T cells, which ultimately results in the elimination of infected or malignant cells. It is thus likely that the expansion of the epitope repertoire as a result of transpeptidation or reverse proteolysis events, results in enhanced anti-tumor immunity as well as in better protection from pathogens. Increasing diversity by V(D)J recombination is essential for protection from a constantly changing bacterial and viral repertoire.[65] V(D)J recombination results in a large variety of antibodies and T cell receptors, while differential splicing of RNA in B lymphocytes results in the production of either membrane bound or secreted antibody molecules. The protease-catalyzed ligation of peptides or even proteins could be yet another way of the immune system to increase diversity and enhance protection.

The increased diversity of self-antigens on the other hand demands additional negative selection to prevent autoimmunity. During T cell development, the thymus is responsible for the negative selection of T cells, resulting in a T cell repertoire that only recognizes foreign antigens.[66] Autoimmunity against spliced antigens may occur if self-antigens are spliced differently in antigen presenting cells compared to cells in the thymus, thereby evading negative selection. Differential splicing in different tissue types will occur if splicing is a random process. However, both transpeptidation and reverse proteolysis reactions only take place under specific circumstances, requiring, for instance, a high degree of substrate specificity. Therefore, it is likely that splicing is a controlled process that is performed identically in different cell types. Little data is available

to confirm this notion, but notably, Vigneron et al. have found identical results with proteasomes purified from either a human renal cell carcinoma line or human erythrocytes.[7] In splicing experiments performed by Warren et al. also constitutive and immunoproteasomes seem to behave in a similar fashion.[1] These findings may indicate that proteasomes from different cell types all behave similarly and point to a tightly controlled splicing process, which is carried out identically in different cell types or even with different proteasome compositions. The spliced antigens produced by both constitutive and immunoproteasomes in the thymus are therefore not likely to be different from antigens produced in other tissues. It should be noted that Murata et al. have discovered a thymus-specific proteasome type, which contains a novel catalytic subunit (β5t) and displays different catalytic properties.[67] However, these thymoproteasomes are not involved in negative T cell selection,[67] and the antigens produced by this type of proteasome are therefore not interfering with autoimmunity control. Together, these reports suggest that transpeptidation and reverse proteolysis reactions in the cell are not likely to lead to an increase in the recognition of self-antigens.

It is imperative to know the identity of the antigens presented by the MHC pathways, as they mediate T cell responses to cancer and transplants, as well as bacterial and viral infection. The characterization of novel antigens is therefore important for vaccine development, T cell-based therapeutic strategies as well as pharmacological immunosuppression. Further studies need to show the contribution of different proteases to the production of noncontiguous antigens and to elucidate their role in immunity.

REFERENCES

1. Warren, E.H. et al. An antigen produced by splicing of noncontiguous peptides in the reverse order. *Science* **313**, 1444-1447 (2006).
2. Shastri, N., Schwab, S. & Serwold, T. Producing nature's gene-chips: the generation of peptides for display by MHC class I molecules. *Annu. Rev. Immunol.* **20**, 463-493 (2002).
3. Saveanu, L., Fruci, D. & van Endert, P. Beyond the proteasome: trimming, degradation and generation of MHC class I ligands by auxiliary proteases. *Mol. Immunol.* **39**, 203-215 (2002).
4. Gromme, M. & Neefjes, J. Antigen degradation or presentation by MHC class I molecules via classical and non-classical pathways. *Mol. Immunol.* **39**, 181-202 (2002).
5. Groothuis, T.A., Griekspoor, A.C., Neijssen, J.J., Herberts, C.A. & Neefjes, J.J. MHC class I alleles and their exploration of the antigen-processing machinery. *Immunol. Rev.* **207**, 60-76 (2005).
6. Watts, C. The exogenous pathway for antigen presentation on major histocompatibility complex class II and CD1 molecules. *Nat. Immunol.* **5**, 685-692 (2004).
7. Vigneron, N. et al. An antigenic peptide produced by peptide splicing in the proteasome. *Science* **304**, 587-590 (2004).
8. Hanada, K., Yewdell, J.W. & Yang, J.C. Immune recognition of a human renal cancer antigen through post-translational protein splicing. *Nature* **427**, 252-256 (2004).
9. Cawston, T.E. & Wilson, A.J. Understanding the role of tissue degrading enzymes and their inhibitors in development and disease. *Best. Pract. Res. Clin. Rheumatol.* **20**, 983-1002 (2006).
10. Hedstrom, L. Introduction: Proteases. *Chemical Reviews* **102**, 4429 (2002).
11. Bordusa, F. Proteases in organic synthesis. *Chemical Reviews* **102**, 4817-4868 (2002).
12. Guzman, F., Barberis, S. & Illanes, A. Peptide synthesis: chemical or enzymatic. *Electronic Journal of Biotechnology* **10**, 279-314 (2007).
13. Bergmann, M. & Fraenkel-Conrat, H. The enzymatic synthesis of peptide bonds. *J. Biol. Chem.* **124**, 1-6 (1938).
14. Bergmann, M. & Behrens, O. On the asymmetric course of the enzymatic synthesis of peptide bonds. *J. Biol. Chem.* **124**, 7-10 (1938).
15. Bergmann, M. & Fruton, J. Some synthetic and hydrolytic experiments with chymotrypsin. *J. Biol. Chem.* **124**, 321-329 (1938).
16. Bergmann, M. & Fraenkel-Conrat, H. The role of specificity in the enzymatic synthesis of proteins. Syntheses with intracellular enzymes. *J. Biol. Chem.* **119**, 707-720 (1937).

17. Jakubke, H.D., Kuhl, P. & Konnecke, A. Basic Principles of Protease-Catalyzed Peptide-Bond Formation. *Angewandte Chemie-International Edition in English* **24**, 85-93 (1985).
18. Homandberg, G.A., Mattis, J.A. & Laskowski, M., Jr. Synthesis of peptide bonds by proteinases. Addition of organic cosolvents shifts peptide bond equilibria toward synthesis. *Biochemistry* **17**, 5220-5227 (1978).
19. Isowa, Y. et al. Thermolysin-Catalyzed Condensation-Reactions of N-Substituted Aspartic and Glutamic Acids with Phenylalanine Alkyl Esters. *Tetrahedron Letters* 2611-2612 (1979).
20. Oka, T. & Morihara, K. Peptide bond synthesis catalyzed by alpha-chymotrypsin. *J. Biochem.* **84**, 1277-1283 (1978).
21. Oka, T. & Morihara, K. Peptide bond synthesis catalyzed by thermolysin. *J. Biochem.* **88**, 807-813 (1980).
22. Stoineva, I.B., Galunsky, B.P., Lozanov, V.S., Ivanov, I.P. & Petkov, D.D. Enzymatic-Synthesis Design and Enzymatic-Synthesis of Aspartame. *Tetrahedron* **48**, 1115-1122 (1992).
23. Murakami, Y., Yoshida, T. & Hirata, A. Enzymatic synthesis of N-formyl-L-aspartyl-L-phenylalanine methyl ester (aspartame precursor) utilizing an extractive reaction in aqueous/organic biphasic medium. *Biotechnology Letters* **20**, 767-769 (1998).
24. Kumaran, S., Datta, D. & Roy, R.P. Conformationally driven protease-catalyzed splicing of peptide segments: V8 protease-mediated synthesis of fragments derived from thermolysin and ribonuclease A. *Protein Science* **6**, 2233-2241 (1997).
25. Roy, R.P., Khandke, K.M., Manjula, B.N. & Acharya, A.S. Helix formation in enzymically ligated peptides as a driving force for the synthetic reaction: example of alpha-globin semisynthetic reaction. *Biochemistry* **31**, 7249-7255 (1992).
26. Somalinga, B.R. & Roy, R.P. Volume exclusion effect as a driving force for reverse proteolysis. Implications for polypeptide assemblage in a macromolecular crowded milieu. *J. Biol. Chem.* **277**, 43253-43261 (2002).
27. Zimmerman, S.B. & Minton, A.P. Macromolecular crowding: biochemical, biophysical, and physiological consequences. *Annu. Rev. Biophys. Biomol. Struct.* **22**, 27-65 (1993).
28. Minton, A.P. The influence of macromolecular crowding and macromolecular confinement on biochemical reactions in physiological media. *J. Biol. Chem.* **276**, 10577-10580 (2001).
29. Minton, A.P. Implications of macromolecular crowding for protein assembly. *Curr. Opin. Struct. Biol.* **10**, 34-39 (2000).
30. Srinivasulu, S. & Acharya, A.S. Product-conformation-driven ligation of peptides by V8 protease. *Protein Science* **11**, 1384-1392 (2002).
31. Kumaran, S., Datta, D. & Roy, R.P. An enigmatic peptide ligation reaction: Protease-catalyzed oligomerization of a native protein segment in neat aqueous solution. *Protein Science* **9**, 734-741 (2000).
32. Ray, S.S., Balaram, H. & Balaram, P. Unusual stability of a multiply nicked form of Plasmodium falciparum triosephosphate isomerase. *Chem. Biol.* **6**, 625-637 (1999).
33. Vogel, K. & Chmielewski, J. Rapid and Efficient Resynthesis of Proteolyzed Triose Phosphate Isomerase. *Journal of the American Chemical Society* **116**, 11163-11164 (1994).
34. Chen, S., Wu, S., Chen, S. & Wang, K. Alcalase catalyzed peptide bond formation between Asp and D-Ala in anhydrous t-butanol. *Biotechnology Letters* **15**, 373-376 (1993).
35. Hou, R. et al. Synthesis of tripeptide RGD amide by a combination of chemical and enzymatic methods. *Journal of Molecular Catalysis B: Enzymatic* **37**, 9-15 (2005).
36. Kuhl, P., Sauberlich, S. & Jakubke, H.D. Model Studies on Protease-Catalyzed Peptide-Synthesis Using 9-Fluorenylmethoxycarbonyl Protected Amino-Acid Derivatives. *Monatshefte fur Chemie* **123**, 1015-1022 (1992).
37. Quiroga, E., Priolo, N., Obregon, D., Marchese, J. & Barberis, S. Peptide synthesis in aqueous-organic media catalyzed by proteases from latex of Araujia hortorum (Asclepiadaceae) fruits. *Biochemical Engineering Journal* **39**, 115-120 (2008).
38. Morihara, K. & Oka, T. alpha-Chymotrypsin as the catalyst for peptide synthesis. *Biochem. J.* **163**, 531-542 (1977).
39. Oka, T. & Morihara, K. Trypsin as a catalyst for peptide synthesis. *J Biochem.* **82**, 1055-1062 (1977).
40. Rose, K., Herrero, C., Proudfoot, A.E., Offord, R.E. & Wallace, C.J. Enzyme-assisted semisynthesis of polypeptide active esters and their use. *Biochem. J.* **249**, 83-88 (1988).
41. Bongers, J. & Heimer, E.P. Recent applications of enzymatic peptide synthesis. *Peptides* **15**, 183-193 (1994).
42. Schellenberger, V., Schellenberger, U., Mitin, Y.V. & Jakubke, H.D. Characterization of the S'-subsite specificity of porcine pancreatic elastase. *Eur. J. Biochem.* **179**, 161-163 (1989).
43. Schellenberger, V., Schellenberger, U., Mitin, Y.V. & Jakubke, H.D. Characterization of the S'-subsite specificity of bovine pancreatic alpha-chymotrypsin via acyl transfer to added nucleophiles. *Eur. J. Biochem.* **187**, 163-167 (1990).
44. Schellenberger, V., Jakubke, H.D. & Kasche, V. Electrostatic effects in the alpha-chymotrypsin-catalyzed acyl transfer. II. Efficiency of nucleophiles bearing charged groups in various locations. *Biochim. Biophys. Acta* **1078**, 8-11 (1991).
45. Stein, R.L. & Strimpler, A.M. Catalysis by human

leukocyte elastase. Aminolysis of acyl-enzymes by amino acid amides and peptides. *Biochemistry* **26**, 2238-2242 (1987).
46. Gillon, A.D. et al. Biosynthesis of circular proteins in plants. *Plant. J.* **53**, 505-515 (2008).
47. Yamauchi, D. & Minamikawa, T. Structure of the gene encoding concanavalin A from Canavalia gladiata and its expression in Escherichia coli cells. *FEBS Lett.* **260**, 127-130 (1990).
48. Min, W. & Jones, D.H. In vitro splicing of concanavalin A is catalyzed by asparaginyl endopeptidase. *Nat. Struct. Biol.* **1**, 502-504 (1994).
49. Saska, I. et al. An asparaginyl endopeptidase mediates in vivo protein backbone cyclization. *J. Biol. Chem.* **282**, 29721-29728 (2007).
50. Bowles, D.J. et al. Posttranslational processing of concanavalin A precursors in jackbean cotyledons. *J. Cell. Biol.* **102**, 1284-1297 (1986).
51. Baumeister, W., Walz, J., Zuhl, F. & Seemuller, E. The proteasome: paradigm of a self-compartmentalizing protease. *Cell* **92**, 367-380 (1998).
52. Canova-Davis, E., Kessler, T.J. & Ling, V.T. Transpeptidation during the analytical proteolysis of proteins. *Anal. Biochem.* **196**, 39-45 (1991).
53. Lippincott, J., Hess, E. & Apostol, I. Mapping of recombinant hemoglobin using immobilized trypsin cartridges. *Anal. Biochem.* **252**, 314-325 (1997).
54. Goepfert, A., Lorenzen, P.C. & Schlimme, E. Peptide synthesis during in vitro proteolysis--transpeptidation or condensation? *Nahrung* **43**, 211-212 (1999).
55. Schaefer, H. et al. Tryptic transpeptidation products observed in proteome analysis by liquid chromatography-tandem mass spectrometry. *Proteomics.* **5**, 846-852 (2005).
56. Fodor, S. & Zhang, Z. Rearrangement of terminal amino acid residues in peptides by protease-catalyzed intramolecular transpeptidation. *Anal. Biochem.* **356**, 282-290 (2006).
57. McIntyre, T.M. & Curthoys, N.P. Comparison of the hydrolytic and transfer activities of rat renal gamma-glutamyltranspeptidase. *J. Biol. Chem.* **254**, 6499-6504 (1979).
58. Allison, R.D. & Meister, A. Evidence that transpeptidation is a significant function of gamma-glutamyl transpeptidase. *J. Biol. Chem.* **256**, 2988-2992 (1981).
59. Castonguay, R., Lherbet, C. & Keillor, J.W. Kinetic studies of rat kidney gamma-glutamyltranspeptidase deacylation reveal a general base-catalyzed mechanism. *Biochemistry* **42**, 11504-11513 (2003).
60. Joshua-Tor, L., Xu, H.E., Johnston, S.A. & Rees, D.C. Crystal structure of a conserved protease that binds DNA: the bleomycin hydrolase, Gal6. *Science* **269**, 945-950 (1995).
61. O'Farrell, P.A., Gonzalez, F., Zheng, W., Johnston, S.A. & Joshua-Tor, L. Crystal structure of human bleomycin hydrolase, a self-compartmentalizing cysteine protease. *Structure* **7**, 619-627 (1999).
62. Tomkinson, B. & Lindas, A.C. Tripeptidyl-peptidase II: a multi-purpose peptidase. *Int. J. Biochem. Cell. Biol.* **37**, 1933-1937 (2005).
63. Tomkinson, B., Ni Laoi, B. & Wellington, K. The insert within the catalytic domain of tripeptidyl-peptidase II is important for the formation of the active complex. *Eur. J. Biochem.* **269**, 1438-1443 (2002).
64. Zheng, W., Johnston, S.A. & Joshua-Tor, L. The unusual active site of Gal6/bleomycin hydrolase can act as a carboxypeptidase, aminopeptidase, and peptide ligase. *Cell* **93**, 103-109 (1998).
65. Shastri, N. Cell biology. Peptides, scrambled and stitched. *Science* **313**, 1398-1399 (2006).
66. Goldrath, A.W. & Bevan, M.J. Selecting and maintaining a diverse T-cell repertoire. *Nature* **402**, 255-262 (1999).
67. Murata, S. et al. Regulation of CD8+ T cell development by thymus-specific proteasomes. *Science* **316**, 1349-1353 (2007).

Chapter 3.2

How the proteasome cleaves and religates peptides to expand the antigen repertoire

How the proteasome cleaves and religates peptides to expand the antigen repertoire

Celia R. Berkers,[a] Annemieke de Jong,[a] Karianne G. Schuurman,[a] Carsten Linnemann,[b] Hugo Meiring,[c] Ton Schumacher,[b] Boris Rodenko[a] & Huib Ovaa[a,*]

Divisions of [a]Cell Biology II and [b]Immunology, The Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands. [c]The Netherlands Vaccine Institute, Unit Research and Development, P.O. Box 457, 3720 AL Bilthoven, The Netherlands.

Recently, three HLA class I epitopes have been reported, in which two distant parts of a protein are excised and ligated together to form a novel peptide. Such non-contiguous antigens can be produced by the proteasome via a transpeptidation mechanism, suggesting that splicing rules are likely to exist. In the present study we use small libraries of short peptides to identify amino acid sequences that are likely to promote this transpeptidation process and to deduce splicing rules for both the primed and non-primed proteasomal sites. Subsequently, we use a long viral polypeptide from the Influenza A neuraminidase 1 protein as a model to validate these rules. We quantify peptide ligation using a model peptide and demonstrate that it can occur with up to 30% ligation efficiency if structural requirements for ligation are met by both ligating partners. In addition, we show that protein splicing preferentially occurs via intrastrand recombination (recombination of peptides from the same protein) and that many spliced products may be formed from a single protein, some of which are likely to be immunogenic. Our data suggest that protein splicing occurs frequently and may occur at least 1,000 times more efficient as assumed previously. Therefore, protein splicing may play a much more significant role in immunity than previously assumed. The splicing rules presented here will facilitate the prediction and detection of spliced antigens to study their roles in immunity.

INTRODUCTION

Following malignant transformation or viral infection, a broad repertoire of antigenic peptides is presented on the cell surface by MHC class I complexes, enabling CD8+ T cells to sense the change in intracellular protein content, which eventually leads to a kill of the antigen presenting cell.[1] The 26S proteasome, a large threonine protease complex, is largely responsible for the generation of these antigenic peptides *in vivo*. The 26S proteasome consists of a 20S core, in which the catalytic activity resides, and 19S regulatory caps, which regulate the entry of substrates into the 20S core. The 20S core is formed by four stacked rings consisting of 7 subunits each, with an overall architecture of α(1-7)β(1-7)β(1-7)α(1-7). Catalytic activity is provided by three β subunits, termed β1, β2 and β5, which are responsible for the caspase-like, tryptic and chymotryptic activ-

*Correspondence should be addressed to H.O. (h.ovaa@nki.nl).

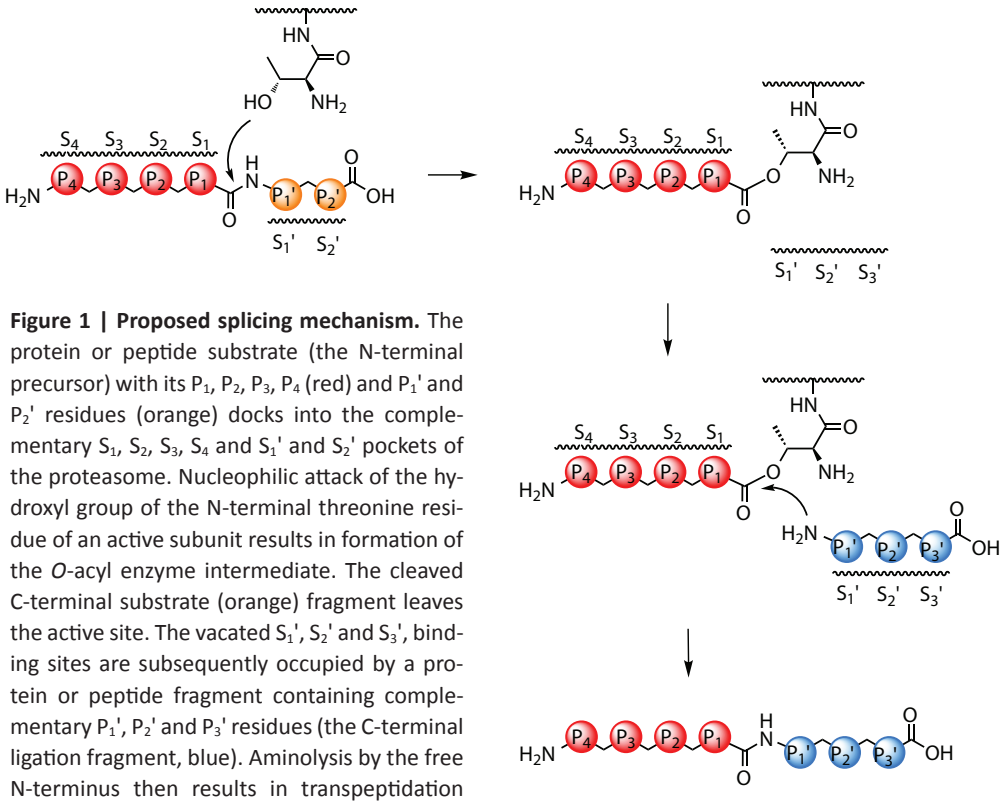


Figure 1 | Proposed splicing mechanism. The protein or peptide substrate (the N-terminal precursor) with its P_1, P_2, P_3, P_4 (red) and P_1' and P_2' residues (orange) docks into the complementary S_1, S_2, S_3, S_4 and S_1' and S_2' pockets of the proteasome. Nucleophilic attack of the hydroxyl group of the N-terminal threonine residue of an active subunit results in formation of the *O*-acyl enzyme intermediate. The cleaved C-terminal substrate (orange) fragment leaves the active site. The vacated S_1', S_2' and S_3', binding sites are subsequently occupied by a protein or peptide fragment containing complementary P_1', P_2' and P_3' residues (the C-terminal ligation fragment, blue). Aminolysis by the free N-terminus then results in transpeptidation and the formation of a novel peptide.

ity of the proteasome, respectively, and all have different cleavage preferences. The β1 subunit cleaves behind acidic amino acid residues, the β2 subunit prefers to cleave after basic residues, and the β5 subunit cleaves behind hydrophobic or aromatic residues. In immune tissues or under influence of the cytokine interferon-γ, these subunits are replaced by their immunoproteasomal counterparts, termed β1i, β2i and β5i to form the immunoproteasome,[2] while the existence of mixed-type hybrid proteasomes has also been reported.[3-5]

Recently, it has become apparent that not only contiguous antigens, but also non-contiguous antigens, that consist of two post-translationally fused peptides, can be presented to the immune system.[6-8] To date, three immunogenic, spliced antigens have been described, RTK-QLYPEW,[7] NTYAS-PRFK[6] and SLPRGT-STPK[8] (where the dash indicates the site of ligation). To form RTK-QLYPEW and NTYAS-PRFK, ligation occurred in-line (*i.e.* ligation of peptide fragments in the order in which they occurred in the parent protein), whereas SLPRGT-STPK was formed by religation of peptide fragments in the reverse order to that in which they occur in the parent protein. The ligation that produces these epitopes is thought to be catalyzed by the proteasome *in vivo* via a transpeptidation mechanism.[7-9] During hydrolysis, the N-terminal threonine residue of a catalytically active subunit reacts with the scissile peptide bond to form an *O*-acyl enzyme intermediate, resulting in the release of the C-terminal part of the peptide. In a second

step, water reacts with the intermediate ester resulting in the release of the N-terminal part of the peptide. During a transpeptidation reaction, the amino group of a second peptide competes with water for reaction with the intermediate ester, resulting in the formation of a new peptide bond (Figure 1).[10]

Proteasomal splicing results in a genuine post-translational modification, which can increase the diversity of the antigenic peptides presented, but which may also serve other yet unknown functions. A greater diversity of epitopes increases the chance that one or more of these epitopes are recognized by CD8+ T cells, which ultimately may result in more efficient elimination of infected or malignant cells.[10-12] It is currently not known whether proteasome splicing is a random event, or governed by explicit rules. As transpeptidation will only be able to compete with the hydrolysis reaction when both the N-terminal and C-terminal ligation fragments fit the active site pockets such that splicing is favored (Figure 1), it is not unlikely that splicing rules exist.[10] In addition, in barrel-shaped proteases such as the proteasome, reaction partners are generated *in situ* and remain confined to a small volume, which likely facilitates peptide ligation.[10] Because the various proteasomal active sites have different architectures, such rules, if existent, are likely site-specific.

In this study, we use short peptides to identify amino acid sequences that are likely to promote transpeptidation reactions, enabling the deduction of splicing rules. We validate these rules using a long viral polypeptide to mimic protein degradation in cells. In addition, we quantify the absolute amount of peptide ligation and demonstrate that it can occur efficiently with up to 30% ligation efficiency. Finally, we show that protein splicing preferentially occurs via intrastrand recombination (recombination of peptides from the same protein) and not via random recombination of peptides derived from different proteins. The elucidation of these splicing rules will facilitate the prediction and detection of spliced antigens to study their roles in immunity.

RESULTS

Rules for proteasomal antigen splicing exist

Two types of fragments participate in proteasomal ligation reactions: 1) N-terminal precursors, that fit into the proteasomal substrate binding non-primed pockets, so called S-pockets (S_1, S_2, S_3…S_n, depending on their proximity to the scissile peptide bond) and 2) C-terminal ligation partners, that fit into the proteasomal primed pockets (S_1', S_2', S_3'…S_n').[11] Residues in the substrate that interact with these proteasomal specificity pockets are referred to as P_1, P_2, P_3…P_n and P_1', P_2', P_3'…P_n', accordingly (Figure 1).[11] For both types of ligation partners, rules predicting proteasomal peptide ligation were deduced by comparing small libraries of potential ligation partners in digestion assays. To investigate which C-terminal ligation partners were likely to participate in ligation reactions, the N-terminal precursor YLDW|SY (where the vertical line indicates the site of cleavage) was incubated with proteasome in the presence of a 5-fold molar excess of the C-terminal ligation partner XXXSV in separate reactions (in which X is any natural amino acid). Transpeptidation led to the accumulation of YLDW-XXXSV, which has affinity for MHC class I (where the dash indicates the newly formed peptide bond). The extent of formation of MHC binders through transpeptidation in each sample could be measured using a fluorescence polarization (FP) assay[13,14] scoring for ligation efficiency (See Supplementary Information and Supplementary Figure 1 for details). By incubating YLDW|SY with proteasome in the presence of XLPSV (where

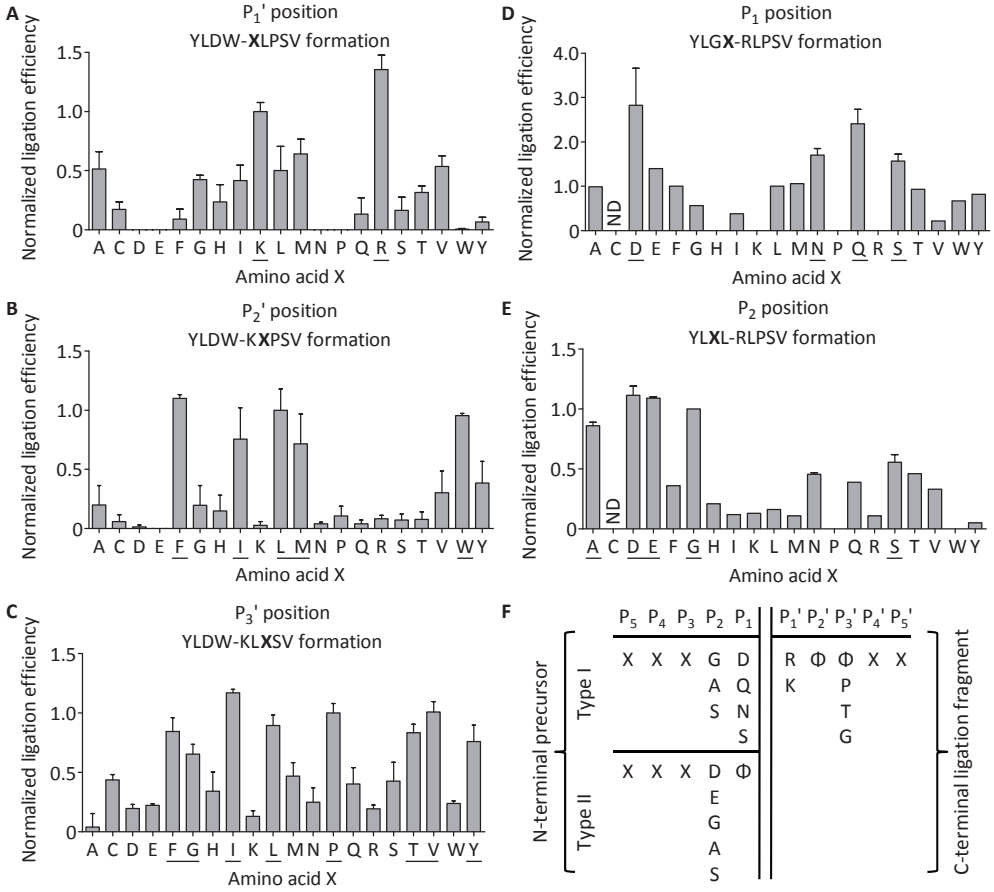


Figure 2 | Splicing rules to predict peptide ligation in the proteasome. A-C) Relative ligation efficiencies of different C-terminal ligation fragments measured in fluorescence polarization assays, indicating splicing preferences for the P_1' **(A)**, P_2' **(B)** and P_3' **(C)** positions. Efficiencies were normalized to the efficiency of KLPSV, which was included in all experiments and set to 1. Values are averaged over at least two independent experiments; error bars represent SEM. Residues that favored ligation on a particular position are underlined. **D,E)** Relative ligation efficiencies of different N-terminal ligation precursors measured in LC-MS assays, indicating splicing preferences for the P_1 **(D)** and P_2 **(E)** positions. Efficiencies were normalized to the efficiency of YLGL|SY, which was included in all experiments and set to 1. Residues that favored ligation on a particular position are underlined. Error bars represent SEM. ND: Not determined. **F)** Summary of the antigen splicing rules determined in these assays. Φ = hydrophobic residue; X = any natural amino acid.

X is any natural amino acid) in 20 separate reactions, we established which amino acid resides on P_1' promoted ligation (Figure 2A). In subsequent experiments, YLDW|SY was incubated with proteasome in the presence of the peptides KXPSV (Figure 2B) or KLXSV (Figure 2C) *in vitro* to define transpeptidation rules for the P_2' and P_3' position. All ligation efficiencies were related to the efficiency of YLDW-KLPSV formation, which was included in all experimental runs. To investigate which N-terminal precursors were likely to promote

transpeptidation reactions, the N-terminal precursor YLXX|SY was incubated with proteasome in the presence of one equivalent of the C-terminal ligation partner RLPSV for different time periods in separate reactions, in which transpeptidation led to the formation of YLXX-RLPSV. In contrast to experiments described above, hydrolysis rates differed between precursors (Supplementary Figure 2), hampering analysis of ligation efficiencies using the FP assay, which requires a constant hydrolysis rate for quantitative comparison of transpeptidation efficiencies. Therefore, we switched to LC-MS analysis of digestion mixtures. Ligation efficiency was defined as the percentage of YLXX|SY cleavages that resulted in ligation. Although MS is not generally used to quantify peptide levels, these peptides differ by a single amino acid per series and we compared efficiencies semi-quantitatively. By incubating YLGX|SY or YLXL|SY (where the vertical line is the site of cleavage and X is any natural amino acid) with proteasome in the presence of RLPSV in separate reactions, splicing rules for the P_1 and P_2 positions were determined (Figures 2D and 2E). Ligation efficiencies were determined for each time point and were normalized to the ligation efficiency of YLGL|SY, which was included in all experimental runs. Maximum ligation efficiencies were plotted for each residue position in Figures 2D and 2E.

From these data, distinct peptide motifs that are likely to either participate in or to abolish antigen splicing could be deduced for both the primed and the non-primed sites (Figure 2F). In general, rules seemed most strict for positions close to the actual site of transpeptidation (P_1, P_1', P_2 and P_2'), whereas multiple residues on P_3' favor transpeptidation. N-terminal precursors that favored ligation fell into the following categories: Type I precursors, which contained aspartate or a polar uncharged residue on P_1 combined with a small or polar (but not charged) residue on P_2 showed high ligation efficiencies but low cleavage rates (Figure 2F and Supplementary Figure 2A). Type II precursors, which combined a hydrophobic residue on P_1 with a basic or, to a lesser extent, polar residue on P_2, showed lower ligation efficiencies, but high cleavage rates (Figure 2F and Supplementary Figure 2A). The higher ligation efficiency of Type I precursors and the higher turnover rate of Type II precursors suggest that in proteins substantial amounts of ligation products may be produced from both types of precursor. C-terminal ligation partners that promoted ligation mostly contained a basic residue (K/R) on P_1' combined with hydrophobic residue on the P_2'. An additional hydrophobic residue or P/G on P_3' enhanced ligation further. Peptides containing a second charged residue on P_2' or P_3' in addition to K or R or P_1' did not serve as C-terminal ligation partners (Figure 2F).

We conclude from Supplementary Figure 2 that hydrolysis rates of N-terminal precursors were highest when a hydrophobic residue was present on P_1, suggesting that cleavage and thus ligation reactions were performed by the β5 subunit, for which the non-primed sites are well characterized.[15,16] Apart from the preference for a basic residue on P_1', the primed sites on the other hand have so far been only poorly characterized.[11,15] The amino acid preferences for ligation described above suggest that ligation efficiencies are highest with C-terminal ligation partners that have a good fit in the primed sites, judging from the requirement of K or R on P_1' for successful ligation. In addition, they further define the architecture of the β5 primed sites and suggest that hydrophobic residues are preferred on P_2' and to a lesser extent on P_3'. N-terminal ligation precursors on the other hand seemed more likely to be involved in ligation if the fit of the side chains of residues on the P_1 and P_2 positions in the corresponding S pockets was

sub-optimal, as inferred from the lower turnover rates for non hydrophobic residues (type I precursors).[15,16]

Splicing of optimized peptide sequences occurs with high efficiency

Next, we set out to characterize the efficiency and mechanism of splicing using an optimized model peptide. We combined the optimized N-terminal precursor YLGD (**a**) with the optimized C-terminal ligation fragment RLPSV (**c**), separated by a two amino acid sequence (SY; **b**) and we incubated this peptide YLGD|SY|RLPSV (**a**|**b**|**c**) with proteasome (where the vertical lines indicate proteasomal cleavage sites). The resulting reaction mixtures were analyzed by HPLC by monitoring absorption at 280 nm at various time points. At 280 nm, only the tyrosine residues in this peptide give rise to absorption, enabling quantification of hydrolysis and transpeptidation events. After a 2-h incubation period, the precursor peptide **abc** was still the main component in the digestion mixture (Supplementary Figure 3, red trace), while smaller amounts of the hydrolysis products **a**, **ab** and **bc** were also visible. In addition, various ligation products were formed (Peaks I to IV; Supplementary Figure 3) and identified by LC-MS as **a**-**a**, **a**-**c**, **a**-**ab** and **a**-**abc** (where the dash indicates the site of ligation). This suggests that peptide **a** is a very good precursor for ligation, as cleavage of the peptide bond between **a** and **b** (**ab** cleavage) resulted in the recombination of peptide **a** with other hydrolysis products or the complete precursor peptide to form a variety of ligation products. When the reaction was allowed to continue, the peak ratio changed. Precursor peptide **abc** almost completely disappeared, while hydrolysis product **a** became the main component in the digestion mixture (Supplementary Figure 3, black trace).

Ligation efficiency was defined as the percentage of **ab** cleavages that resulted in a ligation event. At 2h, the total ligation efficiency was 14% (Figure 3A), suggesting that one ligation product was formed every 7 times the **ab** bond was processed by the proteasome. Total ligation efficiency decreased gradually over time, while the ratio of different ligation products shifted from **a**-**abc** to shorter products. It is likely that the decrease in the amount of precursor **abc** over time led to increased processing of other peptides, including ligation products. Therefore, these shorter ligation products were probably predominantly formed by secondary processing of **a**-**abc**, the most abundant ligation product. When the concentration of precursor peptide **abc** was increased, hydrolysis rates increased, but the relative ratio of hydrolysis products did not shift (data not shown), indicating that the proteasomal cleavage preference was not altered. However, a 4-fold increase in peptide concentration doubled the ligation efficiency at 2 h from 15% to 30% (Figure 3B), suggesting that *in vitro* as many as 1 in 3 **ab** cleavages can result in a ligation event. Especially **a**-**abc**, resulting from the recombination of peptide **a** with the precursor **abc** was formed more efficiently (Figure 3B), suggesting that not only the sequence but also the concentration of C-terminal ligation partner in the proteasome is an important determinant for splicing. When reactions were allowed to continue, total ligation efficiencies decreased (data not shown), in accordance with data shown in Figure 3A.

The ligation efficiencies of splicing reactions for two reported spliced epitopes, RTK-QLYPEW[7] and NTYAS-PRFK[6] have been estimated. One spliced molecule of RTK-QLYPEW was estimated to be formed from ~ 1 x 10^4 precursor molecules RTK|AWNR|QLEPW.[7] The estimated efficiency of NTYAS-PRFK formation was even lower, with roughly one spliced peptide formed from ~ 5 x 10^5 precursor molecules.[9] Notably, these low splicing ef-

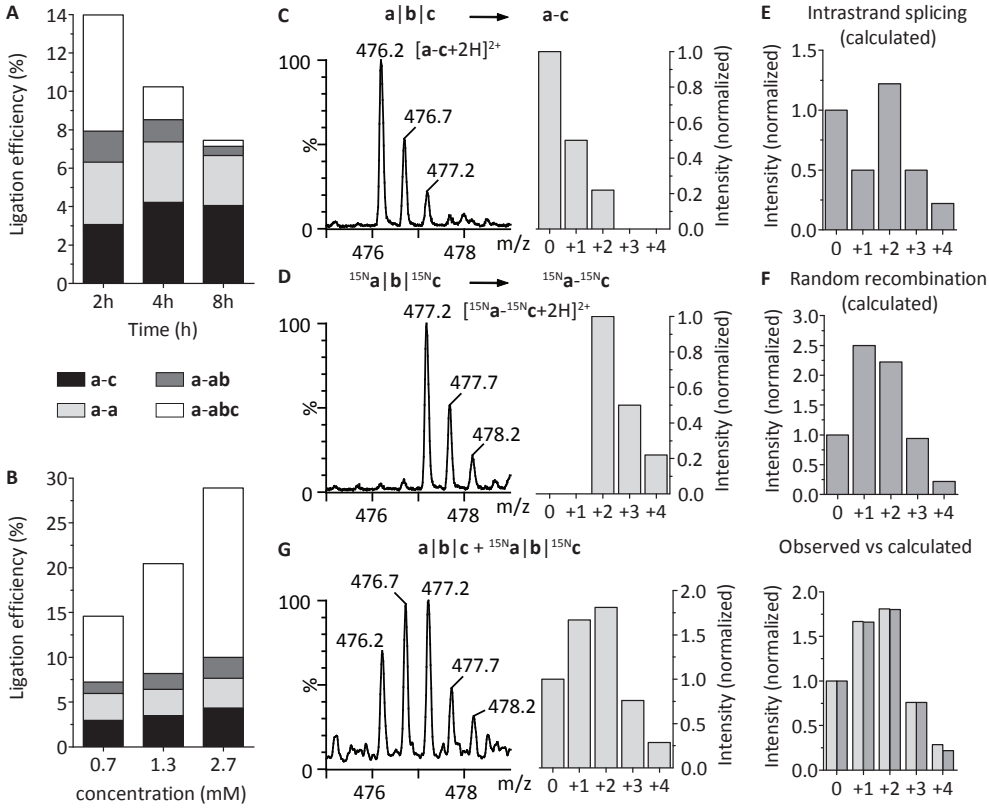


Figure 3 | Splicing of optimized peptide sequences occurs with high efficiency and predominantly via intrastrand splicing. **A)** Ligation efficiencies of the formation of different ligation products resulting from **a**|**b**|**c** digestion, measured at various time points. **B)** Ligation efficiencies of the formation of different ligation products resulting from digestion of increasing concentrations of **a**|**b**|**c**. **C)** Mass spectrum of **a**-**c**, formed during digestion of **a**|**b**|**c** (left panel). Isotope pattern derived from the mass spectrum of **a**-**c** by integration of mass peaks (right panel). **D)** Mass spectrum of ^{15N}**a**-^{15N}**c**, formed during digestion of ^{15N}**a**|**b**|^{15N}**c** (left panel). Isotope pattern derived from the mass spectrum of ^{15N}**a**-^{15N}**c** by integration of mass peaks (right panel). **E)** Calculated isotope distribution pattern for intrastrand splicing of a 1:1 digestion mixture of **a**|**b**|**c** and ^{15N}**a**|**b**|^{15N}**c**. **F)** Calculated isotope distribution pattern for random recombination of a 1:1 digestion mixture of **a**|**b**|**c** and ^{15N}**a**|**b**|^{15N}**c**. **G)** Mass spectrum resulting from the digestion of a 1:1 mixture of **a**|**b**|**c** and ^{15N}**a**|**b**|^{15N}**c** (left panel), showing a mixture of **a**-**c** and ^{15N}**a**-^{15N}**c**. Observed isotope pattern of a 1:1 digestion mixture of **a**|**b**|**c** and ^{15N}**a**|**b**|^{15N}**c** (middle panel). Comparison of the observed isotope distribution pattern with the calculated isotope pattern of 56% random recombination and 44% intrastrand splicing (right panel).

ficiencies of 0.0002% – 0.01% were sufficient to produce immunogenic epitopes that were presented at the cell surface by MHC class I to evoke an immune response. The splicing reactions creating these epitopes require two cleavage steps and for the model peptide described above the equivalent reaction would be the formation of **a**-**c** from **abc**. Splicing efficiencies for **a**-**c** formation were between 1.5% and 3% depending on concentration and reaction time. This indicates that splicing efficiencies may be several orders of magni-

tude higher than previously described if optimized N- and C-terminal ligation fragments are recombined. Therefore, spliced epitopes may be presented on the surface of cells to a larger extent than previously assumed.

Splicing of a model peptide occurs mainly via intrastrand splicing

Next, the mechanism of **abc** splicing was studied in more detail. In theory, **a**-**c** can be formed from **abc** via random recombination or via intrastrand splicing of **a** and **c**. In the latter case, **abc** is first be processed to **ab** and **c**. The liberated **ab** fragment then serves as precursor for ligation of the liberated **c** fragment derived from the same polypeptide chain. In contrast, a random **ab** or **abc** peptide can serve as precursor for **c** ligation, in which case recombination is said to occur randomly. To distinguish between random recombination and intrastrand splicing, a precursor peptide was used in which a ^{15}N-labeled amino acid was incorporated in part **a** and part **c** (^{15N}**a**|**b**|^{15N}**c**). First, either **abc** or ^{15N}**ab**^{15N}**c** were incubated with proteasome separately, and the resulting digestion mixtures were analyzed by LC-MS for the presence of **a**-**c** or ^{15N}**a**-^{15N}**c** (Figures 3C,D). In the digestion mixture of **abc**, a peak with an m/z value of 476.2 (M_0), corresponding to [**a**-**c**+2H]$^{2+}$ could be detected (Figure 3C, left panel), which showed a characteristic isotope pattern (Figure 3C, right panel). In the digestion mixture of ^{15N}**ab**^{15N}**c**, a peak with an m/z value of 477.2 and thus with a mass of +2 Dalton (M_{+2}), corresponding to [^{15N}**a**-^{15N}**c**+2H]$^{2+}$, was detected (Figure 3D, left panel), displaying a similar isotope pattern as **a**-**c** (Compare Figures 3C and 3D, right panels). Next, a 1:1 mixture of **abc** and ^{15N}**ab**^{15N}**c** was incubated with proteasome. In this experiment, intrastrand splicing would yield equal amounts of two products (**a**-**c**, and ^{15N}**a**-^{15N}**c**) with masses M_0 and M_{+2}. The resulting isotope pattern would therefore resemble the sum of Figures 3C and 3D, as shown in Figure 3E. When recombination of **a** and **c** in the mixture would be completely random on the other hand, indicating that intrastrand splicing is not preferred, four different peptides would be formed (**a**-**c**, ^{15N}**a**-**c**, **a**-^{15N}**c**, and ^{15N}**a**-^{15N}**c**) with masses M_0, M_{+1} and M_{+2} in a 1:2:1 ratio, resulting in a isotope distribution pattern as shown in Figure 3F. The actual MS spectrum and isotope distribution pattern of the 1:1 incubation mixture of **abc** and ^{15N}**ab**^{15N}**c** are shown in Figure 3G. The observed isotope pattern is an intermediate between Figures 3E and 3F, suggesting that both random and intrastrand recombination occur. The ratio between these two processes could be estimated by calculating the isotope distribution patterns of different ratios of random recombination and intrastrand splicing. The observed isotope pattern resembled an experiment in which 56% random recombination and 44% intrastrand splicing occurred (Figure 3G, right panel).

To interpret these results, the hydrolysis pattern of **abc** has to be taken into account. HPLC spectra indicate that **abc** is predominantly cleaved between **a** and **b** (66%), and not between **b** and **c** (34%) (data not shown). As intrastrand splicing can only occur when the **bc** bond is cleaved first, it can only occur in 34% of **abc** cleavages. Therefore, we expected that the isotope pattern would be predominantly determined by random recombination and that some enrichment in the M_0 and M_{+2} signals would occur as a consequence of possible intrastrand splicing of **a**-**c**. Remarkably, we observed 44% enrichment of M_0 and M_{+2} signals as a consequence of intrastrand splicing (see above). Such a high degree of enrichment is only possible if intrastrand splicing is highly preferred over random recombination. The proteasome creates N-terminal and C-terminal ligation partners almost simultaneously and as the two fragments are kept

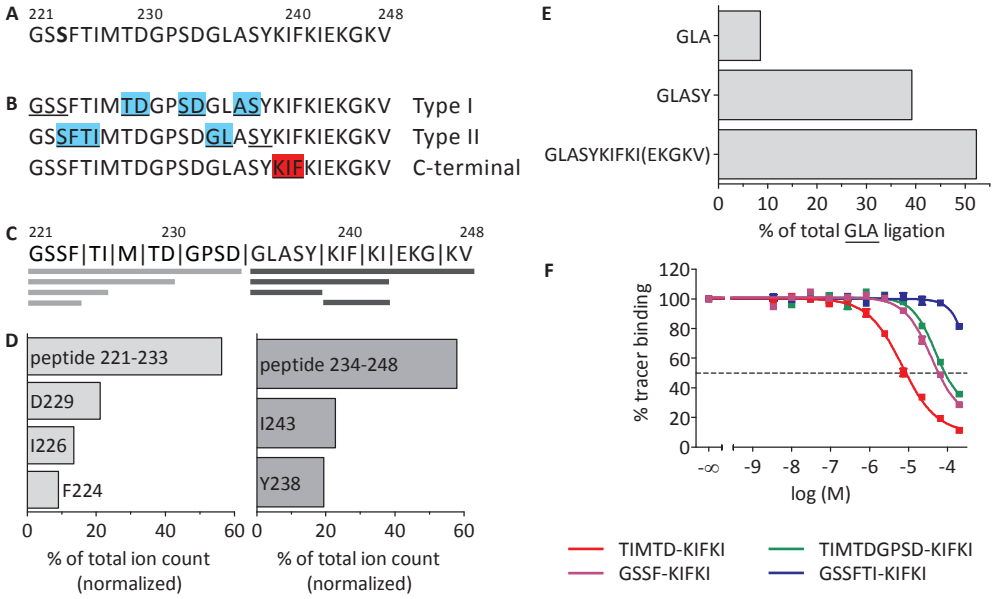


Figure 4 | Validating splicing rules: investigating influenza N1. A) Peptide 221-248/C223S (N1-1) derived from the neuraminidase 1 protein of Influenza A virus A/Puerto Rico/8/34 **B)** N-terminal (blue and underlined) and C-terminal ligation motifs (red and underlined) predicted by the splicing rules. Sequences that can participate in ligation but that are not likely to do so in this particular peptide are underlined only. **C)** Cleavage sites in peptide 221-248 (indicated by vertical lines) and the most predominant hydrolysis products formed (indicated by horizontal lines), as measured by LC-MS. **D)** Relative occurrence of secondary cleavages in peptides 221-233 (GSSFTIMTDGPSD) and 234-248 (GLASYKIFKIEKGKV). **E)** Relative contribution of different GLA-containing C-terminal fragments to total GLA ligation. **F)** HLA-A2.1 binding curves of peptides spliced from peptide 221-248/C223S (N1-1).

together inside the confined space of the catalytic chamber, ligation is promoted. This likely explains why intrastrand splicing occurs preferentially and more efficiently than random recombination.

Validating splicing rules: investigating influenza N1

Next, we validated the determined splicing rules by application to a polypeptide sequence of the neuraminidase 1 (N1) protein of Influenza A virus A/Puerto Rico/8/34 to mimic degradation of a real protein *in vivo*. *In vitro*, only peptides of limited lengths can be digested by 20S proteasome and therefore the 28-mer GSCFTIMTDGPSDGLASYKIFKIEKGKV (positions 221-248), which contained several N- and C-terminal splicing motifs, was selected for further analysis. Peptide 221-248 contained a cysteine residue on position 223, which complicates LC-MS analysis due to poor ionization efficiencies and/or the formation of disulfide bonds. Therefore, C223 was replaced by a serine residue (Figure 4A). Peptide 221-248/C223S (N1-1) contained several potential N- and C-terminal splicing precursors in accordance with our splicing rules. The sequences GS, SS, TD, SD and AS are all type I N-terminal transpeptidation motifs, while SF, TI, GL and SY are in line with Type II rules (Figure 4B, underlined). Only KIF (Figure 4B, underlined) was identified as a suitable

Table 1 | Ligation products identified in the digestion mixture of peptide 221-248.

Ligation product	Ion count	IC_{50} value
GSSF-KIFKI	1304	39 μM
GSSF-GLASY	477	> 1 mM
GSSF-GLASYKIFKIEKGKV	400	ND
GSSFTI-KIFKI	334	1 mM
GSSFTI-KIF	333	ND
GSSFTIMTDGPSD-KIFKI	328	ND
GSSF-KIF	319	ND
GSSF-KIFKIEKGKV	193	ND
GPSD-KIFKI	183	ND
GSSFTIMTDGPSD-KIFKIEKGKV	164	ND
GSSFTIMTD-GLASYKIFKIEKGKV	124	ND
GSSFTI-GLA	112	ND
GSSFTIMTD-KIFKI	109	ND
GSSF-GLASYKIFKI	104	ND
GSSFTI-KIFKIEKGKV	102	ND
GSSF-KIFKIEKG	79	ND
TIMTD-KIFKI	64	6.8 μM
TIMTDGPSD-KIFKI	56	49 μM
GSSFTIMTD-GLASY	40	ND
GSSFTI-GLASYKIFKI	38	ND
GSSFTIMTD-GLASYKIFKI	23	ND

ND: Not determined; a dash indicates the site of ligation.

C-terminal ligation fragment, (KIE contains a second charged residue, which was shown to be detrimental for ligation in experiments described above). In all our digestions, we have not observed cleavage within the first four residues from the N-terminus, which excludes GS and SS as suitable splicing precursors. Also SY must be rejected as splicing precursor, as SY directly precedes KIF and can therefore not participate in splicing with this C-terminal fragment. These restrictions leave six possible N-terminal precursors (indicated in blue) and one C-terminal ligation partner (indicated in red) in peptide N1-1 (Figure 4B) as predicted splicing sites.

Peptide N1-1 was synthesized and incubated with proteasome and the resulting digestion mixtures were analyzed by LC-MS. N1-1 contained nine main cleavage sites as indicated by the vertical lines (Figure 4C). Upon entering the proteasome, 90% of peptides were cleaved after D233, resulting in the accumulation of GSSFTIMTDGPSD (221-233) and GLASYKIFKIEKGKV (234-248) (Figures 4C,D). After this initial cleavage step, peptides 221-233 and 234-248 were both further processed. Fragment 234-248 was cleaved with equal preference after Y238 or I243, and to a lesser extent after F241 or G246. This yielded a variety of secondary cleavage products, of which GLASYKIFKI, GLASY and KIFKI were the most prominent. Fragment 221-233 was further processed after D229, I226 or F224, resulting in the accumulation of GSSFTIMTD, GSSFTI and GSSF (Figures 4C,D).

Table 1 shows all ligation products that could be identified in the digestion mixture. The identity of ligation products was confirmed by comparison with the corresponding synthetic peptide, as exemplified in Supplementary

Figure 4. Ligation was observed onto 4 different precursors of both type I (GSSFTIMTD and GSSFTIMTDGPSD) and type II (GSSF and GSSFTI). Peptides starting with either KIF or GLA served as C-terminal ligation partners, whereas only KIF was predicted by the model. GLA ligations mainly occurred with longer fragments (GLASYKIFKI and GLASYKIFKIEKGKV) (Figure 4E), which are rapidly formed during the initial cleavage step. In addition, GLA ligations were particularly favored over KIF ligations on position D229, which is separated from GLA by only four amino acids and is often the site where the next cleavage occurs. This suggests that GLA ligations occurred more efficiently than predicted because fragments were highly abundant and their generation directly preceded N-terminal precursor processing.

The analysis of ligation products indicates that the N-terminal ligation partner is the dominant factor determining ligation efficiency. Ligation will only take place if the N-terminal ligation precursor has a sequence that is suitable for ligation. Importantly, our splicing model accurately predicted the ability of peptides to serve as N-terminal ligation precursor for all cleavage sites. Ligation will preferentially occur with a suitable C-terminal ligation partner, but seems more dependent on the availability of peptides than on precise sequences. Whereas a hydrophobic amino acid on P_2' seemed required, the P_1' position was not limited to basic residues, but likely also allows for hydrophobic or small residues. Length was not a determining factor, as both short and long peptides were equally suited to serve as C-terminal ligation partners.

Splicing rules predict the formation of potential antigens from influenza N1

To investigate the affinity of spliced products for HLA-A2.1, binding curves of all products that consisted of 14 residues or less and contained at least one anchor residue (L/M/I on P2 and/or V/L/I on P9) were measured using an FP assay (Figure 4F, see Supplementary Information for details). Affinities (expressed as IC_{50} values) were defined as the concentration of peptide that inhibited 50% of tracer peptide binding and were determined for all binders (Table 1). Although N1-1 is a small protein fragment containing only 27 residues, N1-1 processing resulted in the splicing of four peptides that displayed affinity for HLA-A2.1. The peptide TIMTD-KIFKI, which contains both anchor residues and has the proposed optimal length for binding to HLA-A2, had the highest affinity for HLA-A2.1 (Figure 4F), with an IC_{50} value in the low micromolar range (Table 1). The splicing products GSSF-KIFKI and TIMTDGPSD-KIFKI, which had either the optimal length or both anchor residues, displayed intermediate affinity and showed IC_{50} values in the high micromolar range. The peptide GSSFTI-KIFKI had only low affinity for HLA-A2.1, while the other splicing products that were studied showed no affinity (Table 1). These data suggest that many potential HLA-A binding epitopes may be spliced from an average protein that contains 10 to 20 times as many residues as N1-1. Peptides that are spliced by the proteasome may be subjected to further processing by peptidases that reside in the cytosol and the endoplasmic reticulum. Therefore, peptides that do not display the right binding characteristics directly after proteasomal splicing, may be further processed to fit HLA-A molecules and be presented at the cell surface.

Validating splicing mechanism: investigating influenza N1

Next, we set out to compare the efficiency and mechanism of ligation of the influenza A derived N1-1 peptide to the **abc** model peptide described above. Ligation efficiency was defined as the percentage of precursor clea-

Table 2 | Ligation efficiencies of N-terminal precursors identified in peptide 221-248.

Ligation precusor	Efficiency (%)
GSSFTIMTDGPSD-	5.7
GSSFTIMTD-	7.3
GSSFTI-	23.9
GSSF-	59.5
Total	10.9

vages that resulted in a ligation event, and calculated from MS peak intensities. Total ligation efficiencies ranged from 5.7% to 59.5% for different N-terminal precursors (Table 2). The overall splicing efficiency was 10.9%, suggesting that one transpeptidation product was formed every 9 times peptide 221-248 was processed by the proteasome. These efficiencies are in the same range as ligation efficiencies determined for the **abc** peptide, suggesting that protein-derived peptides can be spliced efficiently by the proteasome.

To determine whether ligation would occur randomly or via intrastrand splicing, we replaced G221 and G234 in peptide N1-1 with ^{15}N Glycine (Figure 5A, indicated in red). N1-1 and ^{15N}N1-1 were incubated with proteasome and the formation of $^{(15N)}$GSSFTIMTD-$^{(15N)}$GLASYKIFKIEKGKV (**x**-**z** and ^{15N}**x**-^{15N}**z**; Figure 5A) by the proteasome was studied by LC-MS. In the digestion mixture of N1-1, a peak with an m/z value of 655.9 (M_0), corresponding to [**x**-**z**+4H]$^{4+}$ could be detected (Figure 5B), whereas in the digestion mixture of ^{15N}N1-1, a peak with a mass of +2 Dalton, corresponding to [^{15N}**x**-^{15N}**z**+4H]$^{4+}$, was detected (Figure 5C). A 1:1 mixture of these digestion samples, in which **x**-**z** and ^{15N}**x**-^{15N}**z** are present in equal amounts, resulted in the MS pattern shown in Figure 5D. The observed isotope pattern (Figure 5D, grey bars) resembled the sum of Figures 5B and 5C (Figure 5D, black bars) and is the expected isotope pattern for complete intrastrand splicing.

Next, a 1:1 mixture of N1-1 and ^{15N}N1-1 was incubated with proteasome (Figure 5E). The observed isotope pattern (grey bars) is an intermediate between complete intrastrand splicing (black bars) and complete random recombination (white bars), suggesting that both random and intrastrand recombination occur. A high peptide concentration within the proteasomal catalytic chamber likely increases the chance that two polypeptide strands are degraded simultaneously, favoring random recombination. We hypothesized that decreasing the precursor peptide concentration within the proteasome would shift the ratio between random recombination and intrastrand splicing towards intrastrand splicing. We expect that this shift is accompanied with only a minor decrease in ligation efficiency, as intrastrand splicing should be less effected by the peptide concentration as compared to random recombination. In addition, reducing the peptide concentration within the proteolytic chamber provides a better model for protein ligation *in vivo*, as the catalytic cavity is able to accommodate polypeptides between 25 and 30 kDa in size[17] and will therefore fit only one copy of most proteins. We incubated a 1:1 mixture of N1-1 and ^{15N}N1-1 with proteasome for longer time periods or with lower peptide concentrations to decrease the concentration of N1-1 within the proteasomal catalytic chamber (Figure 5F). Upon increasing the incubation time from 4 h to 16 h, the MS patterns shifted towards more intrastrand splicing. When the m/z peaks of M_0 and M_{+1} were integrated, the ratio shifted towards less M_{+1}, as less random recombination occurred and therefore fewer products with a mass of M_{+1} were formed. After a 16 h incubation period, the resulting MS

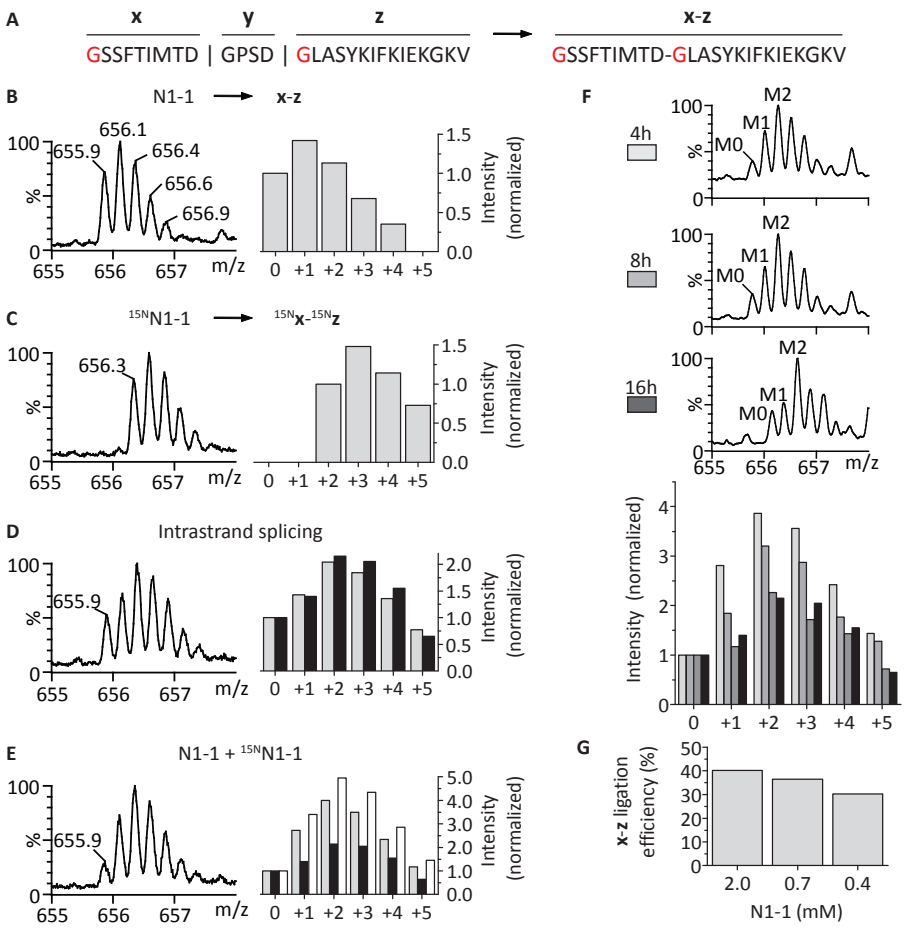


Figure 5 | Validating splicing mechanism: investigating influenza N1. A) Formation of $^{(15N)}$**x**-$^{(15N)}$**z** from $^{(15N)}$N1-1 (peptide 221-248/C223S from the neuraminidase 1 protein of Influenza A virus A/Puerto Rico/8/34) by the proteasome. ^{15}N glycine residues are indicated in red. **B)** Mass spectrum of **x**-**z**, formed during digestion of N1-1 (left panel). Isotope pattern derived from the mass spectrum of **x**-**z** by integration of mass peaks (right panel). **C)** Mass spectrum of ^{15N}**x**-^{15N}**z**, formed during digestion of ^{15N}N1-1 (left panel). Isotope pattern derived from the mass spectrum of ^{15N}**x**-^{15N}**z** by integration of mass peaks (right panel). **D)** Mass spectrum of a 1:1 mixture of digestion samples, in which **x**-**z** and ^{15N}**x**-^{15N}**z** are present in equal amounts (left panel). Observed and calculated isotope distribution patterns for intrastrand splicing (right panel). Grey bars represent the observed pattern; black bars represent the calculated isotope distribution pattern for intrastrand splicing of a 1:1 digestion mixture of N1-1 and ^{15N}N1-1. **E)** Mass spectrum resulting from the digestion of a 1:1 mixture of N1-1 and ^{15N}N1-1 (left panel). Observed isotope pattern of a 1:1 digestion mixture of N1-1 and ^{15N}N1-1 (right panel, grey bars), compared to the calculated isotope distribution patterns for random recombination (white bars) and intrastrand splicing (black bars). **F)** Mass spectra of samples resulting from the digestion of a 1:1 mixture of N1-1 and ^{15N}N1-1 measured at different time points during the digestion (top panel). Observed isotope patterns of a 1:1 digestion mixture of N1-1 and ^{15N}N1-1 measured at different time points during the digestion (4h, 8h, and 16h), compared to the calculated isotope distribution pattern for intrastrand splicing (black bars) (bottom panel). **G)** Ligation efficiencies of **x**-**z** formation from N1-1 at the indicated concentrations of N1-1.

spectrum closely resembled the calculated MS isotope pattern for complete intrastrand splicing (Figure 5F). Upon decreasing the concentration of N1-1 precursor in the sample, we observed only a minor decrease in ligation efficiencies for the formation of **x**-**z** (Figure 5G). Using **abc** as a model peptide, the efficiency of **a**-**abc** formation, which can only be formed via random recombination of two separate peptides, decreased from 19% to 7% (>60% decrease) when the concentration of **abc** was decreased with a factor 4 (Figure 3B). The efficiency of **x**-**z** formation on the other hand only decreased with 25% upon lowering the N1-1 concentration 5-fold (Figure 5G). Collectively, the observed shift in isotope pattern and the relative unchanged ligation efficiencies suggest that protein splicing occurs predominantly via intrastrand recombination.

DISCUSSION

Recently, it has become apparent that non-contiguous antigens, that consist of two post-translationally fused peptides are formed in cells and are presented to the immune system by MHC I.[6-8] These epitopes are thought to be produced by the proteasome via a transpeptidation mechanism.[7-9] As spliced antigens may mediate T cell responses to cancer, transplants and viral infection, knowledge of their identity is important for vaccine development, T cell based therapeutic strategies as well as pharmacological immunosuppression.[10] Novel epitopes are usually identified by cell surface elution, MS analysis and matching against protein databases.[18,19] Alternatively, epitopes may be predicted from protein sequences[20,21] and tested for T cell receptor recognition, e.g. by incorporation into fluorescent MHC class I tetramers, which are subsequently used to stain T cells in fluorescence assisted cell sorting (FACS) assays.[22,23] However, as spliced antigens are not contiguous and their formation requires post-translational splicing, they neither match existing databases nor can they be predicted from contiguous protein sequences. In the present study, we aimed at deciphering splicing rules that govern the production of non-contiguous antigens. Such rules will enable prediction and detection of spliced antigens to elucidate their role in immunity.

Splicing rules were determined for both the N- and C-terminal ligation fragments by analyzing proteasomal incubation samples from small libraries of both ligation partners using fluorescence polarization assays and LC-MS assays. Subsequently, these splicing rules were experimentally confirmed by application to a 27-mer polypeptide derived from the neuraminidase 1 protein from Influenza A virus A/Puerto Rico/8/34. Previous reports hypothesized that the transpeptidation mechanism would implicate that splicing is not determined by a particular sequence motif, but could occur at any major cleavage site.[7,9] Our data are however in strong disagreement with this notion and indicate that splicing rules exist, with ligation efficiencies being primarily determined by the N-terminal ligation partner. In general, N-terminal precursors promoted ligation when fitting sub-optimally in the non-primed S_1 and S_2 pockets. Two types of N-terminal precursors showed highest ligation efficiencies: precursors that I) contained aspartate or a polar uncharged residue on P_1 combined with a small or polar uncharged residue on P_2 or II) combined a hydrophobic residue on P_1 with a basic or small or, to a lesser extent, a polar uncharged residue on P_2. Although the highest ligation efficiency was observed with C-terminal fragments that had a good fit in the primed S_1', S_2' and S_3' pockets, structural requirements on the C-terminal ligation partner are less stringent, and ligation efficiency is rather determined by the presence than by the sequence of C-terminal ligation partners.

Ligation can only take place if the C-terminal ligation partner is able to compete with water molecules that are constantly present in the active site.[10,11] It has been hypothesized that splicing is most likely to occur with substrates that form intermediate esters with relatively long half-lives.[11]. Borrisenko and Groll suggested that this relatively long half-life could result from a high affinity of the N-terminal precursor for the non-primed sites.[11] Our data disagree with this hypothesis and suggest that the half-life of the *O*-acyl enzyme intermediate is prolonged if the fit of the N-terminal precursor in the non-primed pockets is suboptimal. Therefore, we propose a model in which a perfect fit may orientate the *O*-acyl enzyme intermediate such that the attack of a conserved water molecule on the intermediate ester is facilitated. A suboptimal fit of the N-terminal precursor, on the other hand, may prolong the half-life of the intermediate ester by influencing the orientation of the *O*-acyl enzyme intermediate, resulting in an intermediate complex that cannot be hydrolyzed easily and is more prone to ligation.

Our data indicate that a spliced product may be formed as often as once every 3 times a precursor peptide is processed (30% ligation efficiency). In previous studies, the ligation efficiencies of the splicing reactions for two described spliced epitopes were estimated to be 0.0002% to 0.01%.[7,9] Notably, these low efficiencies were sufficient to produce epitopes that were presented at the cell surface by MHC class I leading to an immune response. As our data suggest that splicing can occur with high efficiency *in vivo,* many spliced epitopes are likely to be present on the cell surface of all antigen presenting cells and may contribute to immunity. This splicing mechanism would ensure that a more diverse peptide repertoire reaches the cell surface, which increases the chance of recognition by CD8+ T cells and ultimately of elimination of the malignant or infected cell by the immune system.

The high splicing efficiencies observed in the present study imply that splicing rules must exist. If splicing would be a random process, high splicing efficiencies would result in the formation of high amounts of random (self)-antigens. Such random formation of self-antigens would hamper negative selection by cells in the thymus and cause auto-immunity.[10] Similarly, random recombination of different proteins would result in the formation of randomly formed spliced products, which would also severely hamper negative selection. We therefore expected protein ligation in the proteasome to occur via intrastrand splicing. Our data suggest that intrastrand splicing of peptides is strongly preferred over random ligation. In addition, only intrastrand splicing was observed when the peptide concentration within the proteasomal catalytic chamber was lowered. A reduced peptide concentration within the proteasomal chamber would provide a better model system for protein ligation *in vivo*. Only one copy of most proteins can physically fit into the confined space of the proteasomal catalytic cavity.[17] Experiments with ClpXP complexes, a proteasome-like protease complex from *E. coli,* have shown that two substrates could bind ClpXP simultaneously, but that simultaneous translocation from both ends only rarely occurred.[24] The 20S proteasome from *Thermoplasma acidophilum* can bind two substrates. However, the apparent rate-limiting step in protein degradation is the threading of the substrate through the pore followed by translocation, but not proteolytic cleavage.[25] It is likely that similar mechanisms underlie translocation in mammalian 26S proteasomes, which suggests that simultaneous proteolysis of two different proteins is a rare event. In accord with this, Dalet et al. have shown in a recent study that random recombination occurs readily *in vitro*, but that in cells random re-

combination occurs with much lower efficacy compared to intrastrand splicing.[9] Together, these data suggest that protein ligation occurs via intrastrand splicing *in vivo*, whereas the random recombination of proteins will rarely occur *in vivo* and is predominantly an *in vitro* artifact.

In conclusion, we determined splicing rules for both the primed and the non-primed sites. Such rules will enable the prediction and detection of spliced antigens to elucidate their role in immunity. In addition, we show that splicing occurs via intrastrand recombination with high efficiency and that many spliced products may be formed from a single protein, some of which are likely to be immunogenic. Therefore, the data presented here suggest that protein splicing may play a much more significant role in immunity than previously assumed.

MATERIALS AND METHODS

Peptide building blocks were purchased from Novabiochem and appropriately functionalized resins from Applied Biosystems. ^{15}N Glycine-N-Fmoc was purchased from Cambridge Isotope Laboratories. All solvents were purchased from Biosolve at the highest grade available. All other chemicals were purchased from Aldrich at the highest available purity. All solvents and chemicals were used as received. LC-MS analyses were carried out on a WATERS LCT mass spectrometer in line with a WATERS 2795 HPLC system and a WATERS 2996 photodiode array detector. HPLC purifications and analyses were performed on a Shimadzu LC-20AT prominence liquid chromatography system, coupled to a Shimadzu SPD-20A prominence UV/Vis detector and a Shimadzu CTO-20A prominence column oven.

Peptide synthesis

Peptides were synthesized using standard Fmoc-based solid-phase peptide synthesis protocols and appropriately functionalized PEG-polystyrene Wang resins. Functionalized resins were subjected to coupling cycles, in which deprotection of the Fmoc-group with piperidine/NMP (1/4 v/v), was followed by coupling with 4 equivalents each of Fmoc protected amino acid, di-isopropyl ethylamine (dipea) and (Benzotriazol-1-yloxy) tripyrrolidinophosphonium hexafluorophosphate (PyBop). Reactions were carried out in NMP at a volume of 1 mL/0.1 g of resin. After the final coupling step, the Fmoc group was removed and peptides were released from the resin and fully deprotected by treating the resin with TFA/H_2O/triisopropylsilane (93/5/2 v/v) for 2.5 h. Peptides were precipitated with cold diethylether/pentane (3/1 v/v) and lyophilized from H_2O/ACN/HOAc (65/25/10 v/v). Peptides were purified by HPLC over an Atlantis® Preparative T3 column (5 μm; 20 x 150 mm; Waters) to > 99.5% purity using a linear gradient (5% to 50% acetonitrile in H_2O containing 0.05% TFA). Peptide-containg fractions were pooled and lyophilized.

Proteasome purification

Proteasome was purified from bovine liver as described.[5] After each step, proteasome purity was monitored by incubating fractions with a fluorescent proteasome activity probe, followed by SDS-PAGE and scanning of the resulting gel for fluorescence emission, as described.[26] Briefly, bovine liver was homogenized in phosphate buffered saline (PBS), followed by precipitation using 40% saturated ammonium sulphate to remove unwanted proteins. The proteasome was subsequently precipitated by increasing the concentration of saturated ammonium sulphate to 60%. Following dialysis, the proteasome was further purified using a 10-40% sucrose gradient and anion exchange column chromatography, using DEAE Sephadex A25 resin. Proteasome containing fractions were pooled, concentrated and protein concentrations were determined using the Bradford assay (Biorad).

FP ligation assay to determine transpeptidation rules for the primed sites

50 nmol of the peptide YLDWSY was incubated with 5 μg proteasome in the presence of 250 nmol of XLPSV, KXPSV or KLXSV in ligation buffer (50 mM Tris-HCl, pH 8.5) in a total volume of 75 μL for 16 h at 37°C. As a control, 50 nmol of the peptide YLDWSY was incubated with 5 μg proteasome only. The DMSO concentration in all incubation mixtures was 8%. Following incuba-

tion, ligation mixtures were snap-frozen in liquid N_2, lyophilized, dissolved in 50 μL DMSO and the volume was adjusted to 1 mL with H_2O. Samples were purified over Strata™-X 33 μm polymeric reversed phase disposable columns (Phenomenex). To ensure equal input in all samples, 250 nmol of XLPSV, KXPSV or KLXSV was added to control samples prior to purification. Ligation products were eluted with ACN/H_2O (2/3 v/v), lyophilized and dissolved in FP buffer (20 mM Tris pH 7.4, 150 mM NaCl, 0.5 mg/ml BGG). Fluorescence polarization (FP) HLA-A2.1 binding assays with UV-mediated peptide exchange were performed as described previously[14] with minor changes. The peptide KILGFVFJ_1V (where J_1 is UV-cleavable 3-amino-3-(2-nitrophenyl)propionic acid) was used as conditional ligand and fluorescently labeled FLPSDC^{TMR}FPSV (where TMR is maleimido-linked tetramethylrhodamine) was used as tracer peptide. The ligation samples were used as competitor samples and binding curves of serial 2-fold dilutions of the expected ligation products were determined as reference. Assays were performed in 384-well low-volume black nonbinding surface assay plates (Corning 3820). Wells were loaded with 30μL total volume containing 0.5 μM HLA-A2.1-peptide complex, 1 nM tracer peptide and either 25% of each ligation or control sample or serial dilutions of the expected ligation products in FP buffer. The plate was spun for 1 min at 1000g at room temperature to ensure proper mixing of all components. To start UV-mediated cleavage of the conditional ligand and peptide exchange, the plate was placed under a 365-nm UV lamp at 10 cm distance (366 nm UV lamp, 2 × 15 W blacklight blue tubes, l×w×h 505 × 140 × 117 mm, Uvitec, UK) located in a cold room (4 °C). After 30 min irradiation, the plate was sealed with thermowell sealing tape (Corning) and incubated at room temperature for 24 hours, allowing cleaved peptide fragments to dissociate and to be exchanged for competitor and/or tracer peptide. Subsequently, fluorescence polarization measurements were performed, as described.[13] The relative ligation efficiency was calculated by comparing each sample to the corresponding control sample and ligation efficiencies were related to the efficiency of YLDW-KLPSV formation, which was included in all experimental runs. Data were analyzed using GraphPad Prism software (GraphPad).

LC-MS assay to determine transpeptidation rules for the non-primed sites

50 nmol of the peptide YLGXSY or YLXLSY was incubated with 5 μg proteasome in the presence of 50 nmol RLPSV in ligation buffer (50 mM Tris-HCl, pH 8.5) in a total volume of 75 μL for 2h, 4h, 8h, 16h or 24h at 37°C. The DMSO concentration in all incubation mixtures was 2.6%. Following incubation, ligation mixtures were snap-frozen in liquid N_2, lyophilized and dissolved in ACN/H_2O (1/1 v/v). Samples were filtered over Strata™-X 33 μm polymeric reversed phase disposable columns (Phenomenex) to remove proteasomal proteins and eluted directly. Eluted fractions were analyzed by LC-MS using an XBridge BEH300 C18 column (3.5 μm; 2.1 x 100 mm; Waters) using a linear gradient (5% to 50% acetonitrile in H_2O containing 0.1% Formic acid). Ligation efficiencies were calculated using the following formula: Efficiency = $I_L/(I_L+I_H)$ × 100%, where I_L is the peak intensity of the ligation product YLXX-RLPSV and I_H is the peak intensity of the hydrolysis product YLXX. Ligation efficiencies were related to the efficiency of YLGL-RLPSV formation, which was included in all experimental runs.

Investigating splicing efficiencies using a model peptide

The peptide YLGD|SY|RLPSV (**a**|**b**|**c**) (50 nmol, 100 nmol or 200 nmol) was incubated with proteasome (5 μg/ 50 nmol of peptide) in ligation buffer in a total volume of 75 μL for 2h, 4h or 8h. The DMSO concentration was 5.3% in all incubation mixtures. Incubation mixtures were snap-frozen, lyophilized and dissolved in ACN/H_2O (1/1 v/v). Samples were filtered over Strata™-X 33 μm polymeric reversed phase disposable columns (Phenomenex) to remove proteasomal proteins and eluted directly. Eluted fractions were diluted 1:1 with H_2O and analyzed by HPLC using a Symmetry® C18 column (5 μm, 4.6 x 150 mm, Waters) using a linear gradient (5% to 50% acetonitrile in H_2O containing 0.05% TFA). Spectra were analyzed using LCsolution software (Shimadzu). Integrated peaks were normalized to the number of tyrosine residues the corresponding peptide contained. Ligation efficiencies for the formation of distinct ligation products were calculated using the following formula: Efficiency = $I_L/(I_{LT}+I_H)$ × 100%, where I_L is the integral of ligation product (**a**-**a**, **a**-**c**, **a**-**ab** or **a**-**abc**), I_{LT} is the sum of the in-

tegrals of all ligation products (**a**-**a** + **a**-**c** + **a**-**ab** + **a**-**abc**) and I_H is the integral of the hydrolysis product **a**. Overall splicing efficiencies were calculated using the following formula: Efficiency = $I_{a\text{-}c}/(I_T\text{-}I_{abc}) \times 100\%$, where $I_{a\text{-}c}$ is the integral of the ligation product **a**-**c**, and I_T is the sum of all integrals and I_{abc} is the integral of the precursor peptide **abc**).

Investigating splicing mechanism using a model peptide

50 nmol YLGDSYRLGSV, 50 nmol YL^{15N}GDSYR-L^{15N}GSV, or 50 nmol of a 1:1 mixture of both peptides was incubated with 5 μg proteasome in ligation buffer in a total volume of 75 μL for 2 h. The DMSO concentration was 1.3% in all incubation mixtures. Incubation mixtures were snap-frozen, lyophilized and dissolved in ACN/H_2O (1/1 v/v). Samples were filtered over Strata™-X 33 μm polymeric reversed phase disposable columns (Phenomenex) to remove proteasomal proteins and eluted directly with ACN/H_2O (1/1 v/v). Eluted fractions were analyzed by LC-MS using an XBridge BEH300 C18 column (3.5 μm; 2.1 x 100 mm; Waters) using a linear gradient (5% to 50% acetonitrile in H_2O containing 0.1% Formic acid).

Investigating influenza N1

Synthetic GSSFTIMTDGPSDGLASYKIFKIEKGKV (150 nmol) was incubated with 5 μg proteasome in ligation buffer in a total volume of 75 μL for 16 h. To investigate the splicing mechanism, different amounts of GSSFTIMTDGPSDGLASYKIFKIEKGKV, ^{15N}GSSFTIMTDGPSD^{15N}GLASYKIFKIEKGKV, or a 1:1 mixture of both peptides were incubated with 5 μg proteasome in ligation buffer in a total volume of 75 μL for 4 h, 8 h or 16 h. Incubation mixtures were snap-frozen, lyophilized and dissolved in ACN/H_2O (1/1 v/v). Samples were filtered over Strata™-X 33 μm polymeric reversed phase disposable columns (Phenomenex) to remove proteasomal proteins and eluted directly with ACN/H_2O (1/1 v/v). Eluted fractions were analyzed by LC-MS using an XBridge BEH300 C18 column (3.5 μm; 2.1 x 100 mm; Waters) using a linear gradient (5% to 50% acetonitrile in H_2O containing 0.1% Formic acid). To determine the affinities of spliced products for HLA-A2.1, an FP assay was performed as described above, with minor changes. 3-fold serial dilutions of synthetic products were used as competitor peptide and fluorescence polarization measurements were performed after a 4 h incubated period at room temperature. The binding affinity (IC_{50} value) of each spliced product was defined as the concentration that inhibited 50% of tracer peptide binding. Data were analyzed using GraphPad Prism software (GraphPad).

ACKNOWLEDGMENTS

The authors thank Henk Hilkmann for peptide synthesis and Jacques Neefjes for help with proteasome purification. This work was funded by the Dutch Cancer Society (grant NKI 2005-3368) and NWO/ALW.

REFERENCES

1. Shastri, N., Schwab, S. & Serwold, T. Producing nature's gene-chips: the generation of peptides for display by MHC class I molecules. *Annu. Rev. Immunol.* **20**, 463-493 (2002).
2. Rock, K.L. & Goldberg, A.L. Degradation of cell proteins and the generation of MHC class I-presented peptides. *Annu. Rev. Immunol.* **17**, 739-779 (1999).
3. Dahlmann, B., Ruppert, T., Kuehn, L., Merforth, S. & Kloetzel, P.M. Different proteasome subtypes in a single tissue exhibit different enzymatic properties. *J Mol. Biol.* **303**, 643-653 (2000).
4. Dahlmann, B., Ruppert, T., Kloetzel, P.M. & Kuehn, L. Subtypes of 20S proteasomes from skeletal muscle. *Biochimie* **83**, 295-299 (2001).
5. Raijmakers, R. et al. Automated online sequential isotope labeling for protein quantitation applied to proteasome tissue-specific diversity. *Mol. Cell. Proteomics* **7**, 1755-1762 (2008).
6. Hanada, K., Yewdell, J.W. & Yang, J.C. Immune recognition of a human renal cancer antigen through post-translational protein splicing. *Nature* **427**, 252-256 (2004).
7. Vigneron, N. et al. An antigenic peptide produced by peptide splicing in the proteasome. *Science* **304**, 587-590 (2004).
8. Warren, E.H. et al. An antigen produced by splicing of noncontiguous peptides in the reverse order. *Science* **313**, 1444-1447 (2006).
9. Dalet, A., Vigneron, N., Stroobant, V., Hanada, K. & Van den Eynde, B.J. Splicing of distant Peptide fragments occurs in the proteasome by transpeptidation and produces the spliced antigenic peptide derived from fibroblast growth factor-5. *J. Immunol.* **184**, 3016-3024 (2010).
10. Berkers, C.R., de Jong, A., Ovaa, H. & Rodenko, B.

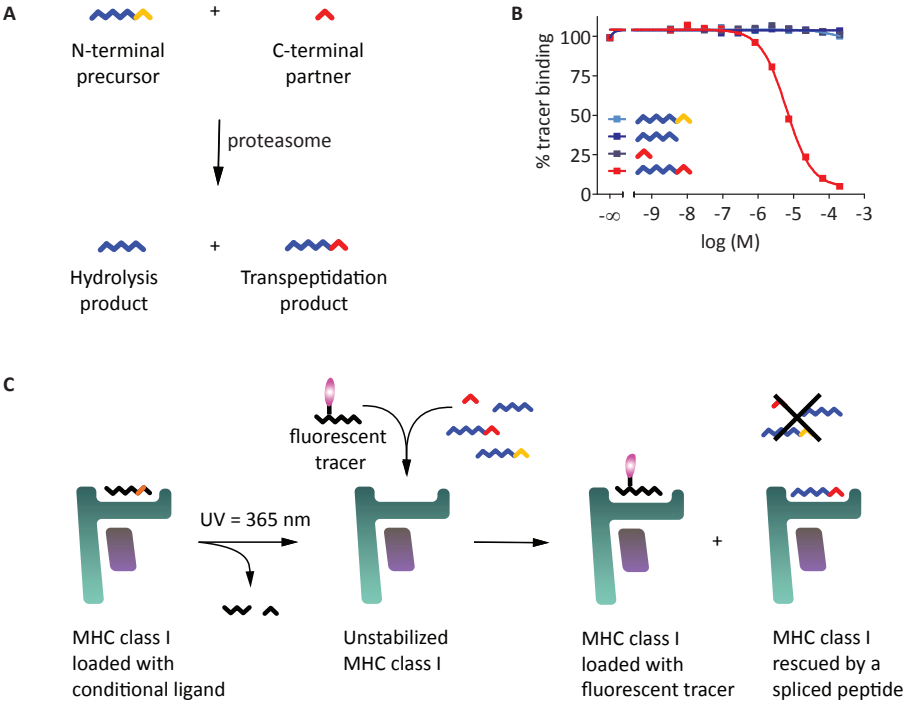
Transpeptidation and reverse proteolysis and their consequences for immunity. *Int J. Biochem. Cell Biol.* **41**, 66-71 (2009).
11. Borissenko, L. & Groll, M. Diversity of proteasomal missions: fine tuning of the immune response. *Biol. Chem.* **388**, 947-955 (2007).
12. Shastri, N. Cell biology. Peptides, scrambled and stitched. *Science* **313**, 1398-1399 (2006).
13. Bakker, A. H. et al. Conditional MHC class I ligands and peptide exchange technology for the human MHC gene products HLA-A1, -A3, -A11, and -B7. *Proc. Natl. Acad. Sci. U S A* **105**, 3825-3830 (2008).
14. Rodenko, B. et al. Class I major histocompatibility complexes loaded by a periodate trigger. *J. Am. Chem. Soc.* **131**, 12305-12313 (2009).
15. Nussbaum, A.K. et al. Cleavage motifs of the yeast 20S proteasome beta subunits deduced from digests of enolase 1. *Proc. Natl. Acad. Sci. U S A* **95**, 12504-12509 (1998).
16. Toes, R.E. et al. Discrete cleavage motifs of constitutive and immunoproteasomes revealed by quantitative analysis of cleavage products. *J. Exp. Med.* **194**, 1-12 (2001).
17. Sharon, M. et al. 20S proteasomes have the potential to keep substrates in store for continual degradation. *J. Biol. Chem.* **281**, 9569-9575 (2006).
18. Henderson, R.A. et al. HLA-A2.1-associated peptides from a mutant cell line: a second pathway of antigen presentation. *Science* **255**, 1264-1266 (1992).
19. Hunt, D.F. et al. Characterization of peptides bound to the class I MHC molecule HLA-A2.1 by mass spectrometry. *Science* **255**, 1261-1263 (1992).
20. Rammensee, H., Bachmann, J., Emmerich, N.P., Bachor, O.A. & Stevanovic, S. SYFPEITHI: database for MHC ligands and peptide motifs. *Immunogenetics* **50**, 213-219 (1999).
21. Nielsen, M. et al. Reliable prediction of T-cell epitopes using neural networks with novel sequence representations. *Protein Sci.* **12**, 1007-1017 (2003).
22. Toebes, M. et al. Design and use of conditional MHC class I ligands. *Nat. Med.* **12**, 246-251 (2006).
23. Altman, J.D. et al. Phenotypic analysis of antigen-specific T lymphocytes. *Science* **274**, 94-96 (1996).
24. Ortega, J., Lee, H.S., Maurizi, M.R. & Steven, A.C. Alternating translocation of protein substrates from both ends of ClpXP protease. *EMBO J.* **21**, 4938-4949 (2002).
25. Hutschenreiter, S., Tinazli, A., Model, K. & Tampe, R. Two-substrate association with the 20S proteasome at single-molecule level. *EMBO J.* **23**, 2488-2497 (2004).
26. Berkers, C.R. et al. Profiling proteasome activity in tissue with fluorescent probes. *Mol. Pharm.* **4**, 739-748 (2007).

Supplementary information

SUPPLEMENTARY RESULTS

Development of a fluorescence polarization assay to measure peptide ligation.

In order to screen small peptide libraries for their ability to participate in proteasomal transpeptidation reactions, we used a UV-mediated MHC exchange fluorescence polarization (FP) assay[1,2] scoring for ligation efficiency. In this assay, HLA-A complexes loaded with a UV-cleavable peptide ligand are subjected to UV light, which cleaves the ligand, resulting in the generation of empty HLA-A.[3] These empty HLA-A complexes are unstable, unless they are rescued by the addition of a 'rescue' peptide. When rescue is performed by adding both a fluorescent tracer peptide and increasing concentrations of a 'competitor' peptide, the affinity of this competitor peptide can be determined.[1,2] This assay was adapted so it could be used to determine the extent to which peptide ligation had taken place in a digestion mixture, resulting in potential HLA-A2.1 epitopes. First, proteasome was incubated with an N-terminal ligation precursor and a C-terminal ligation partner, resulting in the formation of both hydrolysis and transpeptidation products (Supplementary Figure



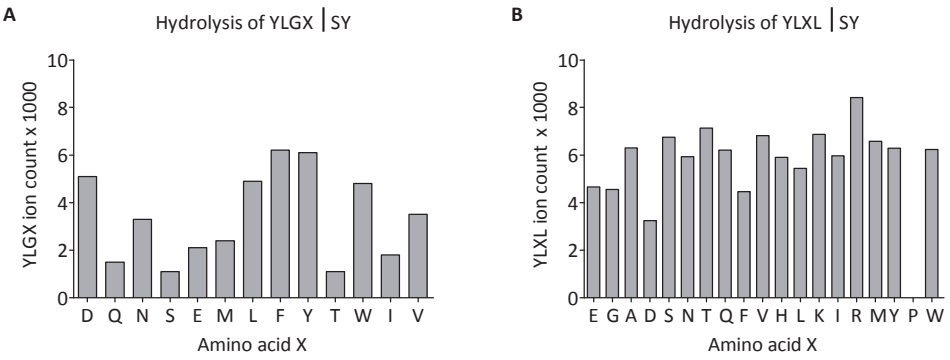
Supplementary Figure 1 | Development of a fluorescence polarization ligation assay. A) Schematic overview of hydrolysis and transpeptidation reactions in the digestion mixture. **B)** Affinities of components in the digestion mixture for MHC cass I. **C)** Schematic overview showing the fluorescence polarization assay used to determine ligation efficiencies in proteasomal digestion mixtures.

1A). This digestion mixture was used to rescue MHC class I. Peptides require two anchor residues to bind HLA-A2.1, one on the second position and one on the last position. Both N- and C-terminal precursors were designed such that they contained only one anchor residue and did not bind MHC class I (Supplementary Figure 1B). Upon ligation, peptides with both anchor residues were formed, resulting in a much-increased affinity for MHC class I (Supplementary Figure 1B). Therefore, competition with fluorescent tracer occurred only when transpeptidation products had been formed (Supplementary Figure 1C) and the relative amount of bound tracer inversely correlated to the efficiency of ligation in the corresponding digestion mixture. Samples were normalized to their respective control samples, in which ligation products could not be formed.

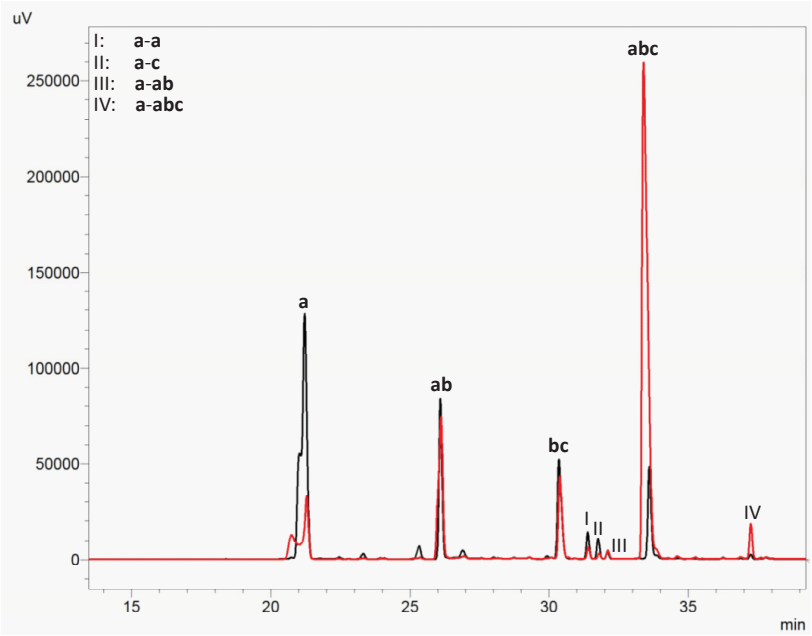
SUPPLEMENTARY REFERENCES

1. Bakker, A. H. et al. Conditional MHC class I ligands and peptide exchange technology for the human MHC gene products HLA-A1, -A3, -A11, and -B7. *Proc. Natl. Acad. Sci. U S A* **105**, 3825-3830 (2008).
2. Rodenko, B. et al. Class I major histocompatibility complexes loaded by a periodate trigger. *J. Am. Chem. Soc.* **131**, 12305-12313 (2009).
3. Toebes, M. et al. Design and use of conditional MHC class I ligands. *Nat. Med.* **12**, 246-251 (2006).

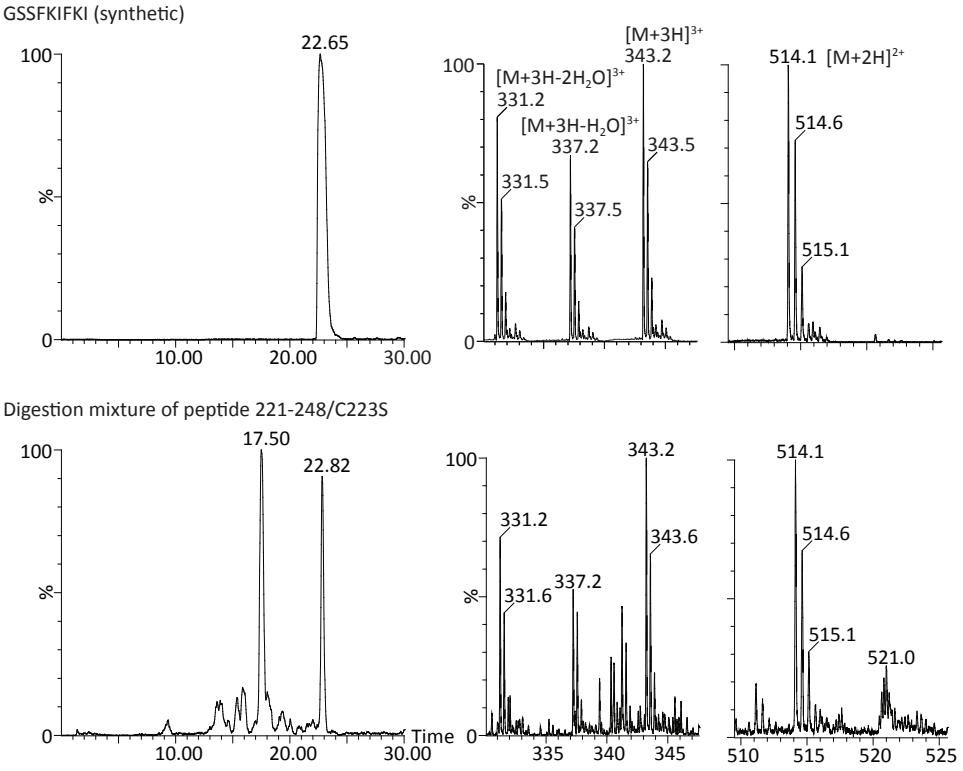
Supplementary figures



Supplementary Figure 2 | Hydrolysis rates of different N-terminal precursors. A) Amount of the hydrolysis product YLGX in proteasomal digestion mixtures of YLGX|SY after a 2h-incubation as measured by LC-MS. **B)** Amount of the hydrolysis product YLXL in proteasomal digestion mixtures of YLXL|SY after a 2h-incubation as measured by LC-MS. Cleavage sites are indicated with a vertical line.



Supplementary Figure 3 | Quantification of transpeptidation and splicing efficiencies using HPLC. HPLC spectra of proteasomal digestion mixtures of **abc** recorded at 280 nm after 2 h (red) and 8 h (black) incubations. Peaks were assigned to hydrolysis and ligation products using LC-MS. A dash indicates the site of ligation.



Supplementary Figure 4 | Identification of splicing products in peptide 221-248/C223S. Comparison of synthetic GSSFKIFKI and the digestion mixture of peptide 221-248/C223S. Ion traces detected at 342.7 m/z are compared (left panels) as well as in-line MS spectra of the peaks eluting at 22 to 23 minutes. Spectra taken from these peaks were identical, confirming that GSSF-KIFKI was formed in the digestion mixture by proteasomal splicing.

Chapter 3.3

Proteasomal splicing creates a novel type of antigen containing an isopeptide linkage

Proteasomal splicing creates a novel type of antigen containing an isopeptide linkage

Celia R. Berkers,[a] Annemieke de Jong,[a] Karianne G. Schuurman,[a] Jan A. J. Geenevasen,[b] Boris Rodenko[a,*] & Huib Ovaa[a,*]

[a]Department of Cell Biology II, The Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands. [b]Synthetic Organic Chemistry Group, Van 't Hoff Institute For Molecular Sciences, University of Amsterdam, Nieuwe Achtergracht 129, 1018 WS Amsterdam, The Netherlands.

Proteasomal transpeptidation reactions can occur particularly efficient if the C-terminal ligation partner has lysine or arginine at the site of ligation. As lysine has two amino groups that can theoretically both react with the *O*-acyl enzyme intermediate, this implies that the proteasome may be able to from isopeptide linkages. In the present study we use NMR to show for the first time that the proteasome can use both the α-amino group and the ε-amino group of lysine in transpeptidation reactions. The overall efficiency of $^{\varepsilon}$K-ligation reactions is 10-fold lower as compared to α-ligation and should produce sufficient amounts of peptide to evoke a T cell response, suggesting that the proteasome can create a novel type of antigen that may play a role in immunity in vivo. In addition, we show that isopeptides have unique properties that discern them from normal epitopes. Isopeptides of various lengths can bind to HLA-A2.1 and HLA-A3 proteins with high affinity. In addition, isopeptides are more stable towards further proteasomal processing as compared to normal peptides. These properties are likely to increase the fraction of ε-ligated peptides that enters the endoplasmic reticulum, and that is loaded onto MHC class I and transported to the cell surface for CD8+ T cell surveillance. We postulate that the formation of epsilon linkages is a genuine post-translational modification resulting from transpeptidation mechanisms.

INTRODUCTION

The eukaryotic 26S proteasome is responsible for the degradation of redundant and misfolded proteins and for the turnover of regulatory proteins involved in a wide range of cellular processes, including cell proliferation and survival, cell-cycle control and cellular stress responses.[1,2] In addition, the proteasome is responsible for the generation of peptides that are presented on the cell surface by Major Histocompatability Complex (MHC) class I proteins, known in humans as human leukocyte antigen (HLA) class I.[3] The proteasome generates peptides both from self and foreign proteins, which are subsequently transported into the endoplasmic reticulum, loaded onto MHC class I and transported to the cell surface. CD8+ T cells continuously scan the MHC I-peptide complexes on the cell surface, thereby receiving a blueprint of the intracellular protein content. This enables CD8+ T cells

*Correspondence may be addressed to H.O. or B.R. (h.ovaa@nki.nl or b.rodenko@nki.nl).

to sense viral infection or malignant transformation, which ultimately results in killing of the antigen presenting cell.

The 26S proteasome consists of a barrel-shaped 20S core, complexed at one or both ends with 19S regulatory particles. The 19S caps deubiquitinate and unfold protein substrates and regulate the entrance of substrates into the 20S core, where the proteolytic activity resides. The 20S catalytic chamber consists of four stacked heptameric rings and has an overall architecture of α(1-7) β(1-7)β(1-7)α(1-7). Whereas the two outer α-rings regulate entry into the complex and provide anchor points for the attachment of 19S regulatory caps, the actual catalytic activity resides within the inner β-rings. Once entered into the 20S core, a protein is degraded via the action of three catalytically active subunits, termed β1, β2 and β5, which all have different cleavage specificities and which are responsible for the proteasomal caspase-like, tryptic and chymotryptic activities, respectively. In lymphoid tissues these subunits are largely replaced by their immunoproteasome counterparts, termed β1i, β2i and β5i, to form the immunoproteasome.[4] The latter has been hypothesized to favor the production of antigenic peptides,[5] while the existence of hybrid 'mixed-type' proteasomes is now also well established.[6,7]

Recently, it has become apparent that the proteasome does not only produce contiguous peptides for presentation on MHC class I. Non-contiguous antigens, in which two distant parts of a protein are excised and ligated together to form a novel peptide, are presented on the cell surface and can evoke immune responses.[8-10] These non-contiguous antigens can be produced by the proteasome via a transpeptidation mechanism.[9-11] During the transpeptidation event, an active site N-terminal threonine hydroxyl group attacks the scissile peptide bond, resulting in the formation of an *O*-acyl enzyme intermediate and the release of the C-terminal part of the peptide. In a second step, this intermediate ester is captured by an amino group of a second peptide (the 'C-terminal ligation fragment'). This leads to the formation of a novel peptide bond and a spliced peptide in which two separate peptide fragments are combined (See Chapter 3.1, Figure 1; Chapter 3.2, Figure 1).[12] Transpeptidation competes with normal hydrolysis, in which the intermediate ester is hydrolyzed by water molecules present in excess. Hydrolysis results in the release of the N-terminal part of the peptide, which can be further processed in subsequent degradation rounds.

All peptides produced in the proteasomal catalytic core have a free amino terminus and can therefore theoretically participate in ligation reactions as C-terminal ligation partner. In a previous study (Chapter 3.2), we observed the highest ligation efficiencies with peptides that had a basic amino acid (arginine or lysine) as their N-terminal residue. As lysine has a free amino group on the C^{ε} position in addition to its free amino terminus, both the α-amine and the ε-amine are theoretically able to participate in ligation reactions. Whereas ligation with the α-amine results in the formation of a normal peptide bond, ligation with the ε-amine results in the formation of an isopeptide. Isopeptide linkage-containing epitopes may have unique properties and may form a novel class of post-translationally modified peptides that can be presented on MHC class I. Isopeptide bonds are notoriously difficult to identify and have not been observed in epitopes so far. In this study, we show that both the α- and ε-amino groups of lysine can participate in ligation reactions and we show how the unique properties of these isopeptides may contribute to an increased MHC class I surface expression, strongly suggesting that ε-ligation may contribute to immunity *in*

vivo, even though sensitive methods to detect such linkages directly do not exist.

RESULTS

The proteasome can form both normal and isopeptide bonds during splicing

To investigate whether both the α- and the ε-amino group of lysine could react with the *O*-acyl enzyme intermediate during splicing reactions, unprotected and mono-acetylated lysine were compared in splicing experiments. Only the unprotected α-amino group of N^{ε}-acetylated lysine ($K^{\varepsilon(Ac)}$) can participate in ligation reactions, resulting in the formation of normal peptide bonds only. On the other hand, N^{α}-acetylated lysine ($K^{\alpha(Ac)}$) can only react with the acyl-enzym intermediate through its unprotected ε-amino group, resulting in the formation of isopeptide bond containing peptides. The peptide C^{TMR}SLPRGTASSR is an N-terminally extended and fluorescently labeled version of the known splicing precursor peptide SLPRGTASSR.[10] This precursor was subjected to proteasomal degradation in the presence of proteasome inhibitors or lysine, followed by MALDI analysis. Peak heights in MALDI spectra could be semi-quantitatively compared, as all C-terminal hydrolysis and ligation products formed during these digestions contained a positively charged TMR (tetramethylrhodamine) label, which functioned as an ionization enhancer acting as the main determinant of signal intensities. The precursor C^{TMR}SLPRGTASSR (m/z 1616, Figure 1A, top panel) disappeared upon incubation with purified 20S proteasome, while two peaks appeared at m/z 1372 and 1214, which could be assigned to the hydrolysis products C^{TMR}SLPRGTAS and C^{TMR}SLPRGT, respectively (Figure 1A, second panel). Hydrolysis was completely abolished by the addition of the pan-proteasome inhibitor MG132, while the addition of the β5/β1 inhibitor bortezomib[13] only hampered the formation of C^{TMR}SLPRGTAS (m/z 1372), but not of C^{TMR}SLPRGT (m/z 1214) (Figure 1A, third and fourth panel). This suggests that C^{TMR}SLPRGTAS is formed by the β5 subunit, while C^{TMR}SLPRGT is a β2 hydrolysis product. When excess lysine was added to the reaction mixture, direct ligation of lysine onto the precursor peptide occurred, as evidenced by the appearance of additional peaks at m/z 1500 (Figure 1B, second panel) and m/z 1342 (data not shown). These peaks could be assigned to the β5 ligation product C^{TMR}SLPRGTAS-K and the β2 ligation product C^{TMR}SLPRGT-K (where a dash indicates the newly formed (iso)peptide bond). Ligation efficiencies, defined as the percentage of cleavages that resulted in ligation as opposed to hydrolysis, were 26% and 13% for the β5 and β2 active sites, respectively (Figure 1C). When either $K^{\alpha(Ac)}$ or $K^{\varepsilon(Ac)}$ was added to the digestion mixture, both MALDI spectra showed an additional peak at m/z 1542 (Figure 1B, third and fourth panel), which could be assigned to the β5 ligation product C^{TMR}SLPRGTAS-K^{Ac}. Ligation efficiencies were estimated at 29% for peptide bond formation with $K^{\varepsilon(Ac)}$ and 33% for isopeptide bond formation with $K^{\alpha(Ac)}$ (Figure 1C). Formation of the β2 ligation product C^{TMR}SLPRGT-K^{Ac} was only observed upon addition of $K^{\varepsilon(Ac)}$, with a ligation efficiency of 7% (Figure 1C). These data indicate that the α- and ε-amino groups of lysine are equally capable of participating in proteasomal splicing reactions in the β5 active site, whereas the β2 active site only accepts α-amino groups. This strongly suggests that both peptide and isopeptide bonds can be formed during proteasomal splicing reactions.

The β5 site accepts α-acetylated peptides as C-terminal ligation partner

Having established that isopeptide bonds can be formed by the proteasome when single

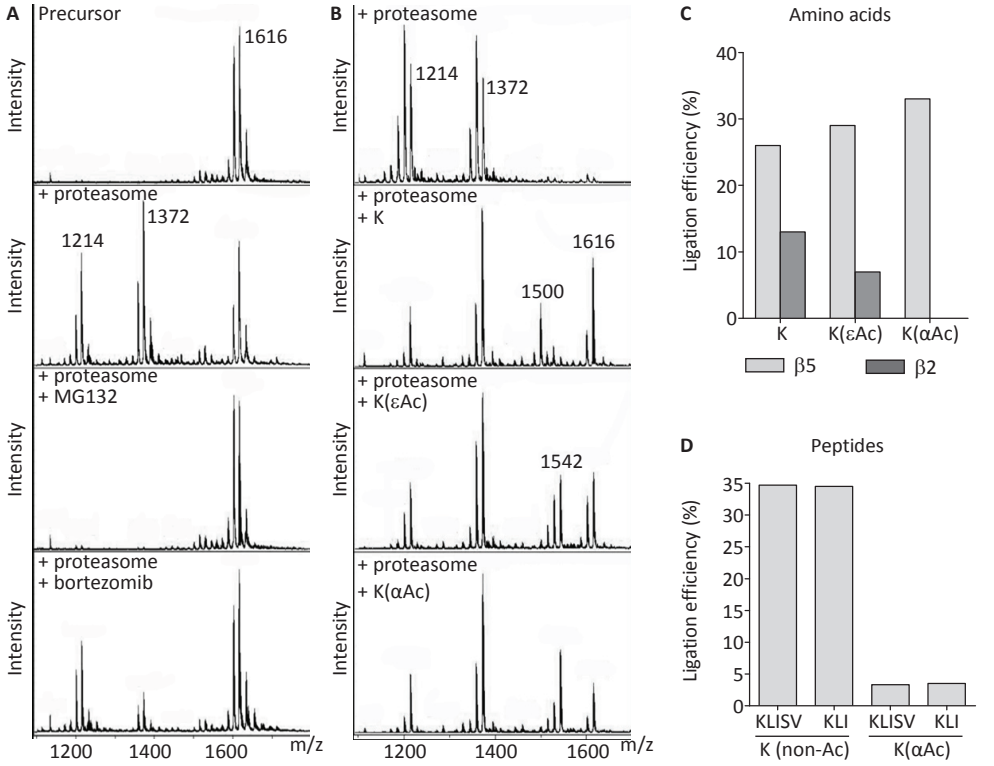


Figure 1 | The proteasome accepts α-acetylated amino acids and peptides as C-terminal ligation partner during splicing reactions. A) MALDI spectrum of C^{TMR}SLPRGTASSR (top panel). MALDI spectra of proteasomal digestion mixtures of C^{TMR}SLPRGTASSR only (second panel), C^{TMR}SLPRGTASSR and MG132 (third panel), and C^{TMR}SLPRGTASSR and bortezomib (bottom panel). **B)** MALDI spectra of proteasomal digestion mixtures of C^{TMR}SLPRGTASSR only (top panel), C^{TMR}SLPRGTASSR and lysine (second panel), C^{TMR}SLPRGTASSR and N^{ε}-acetylated lysine (third panel), and C^{TMR}SLPRGTASSR and N^{α}-acetylated lysine (bottom panel). **C)** Ligation efficiencies for the formation of β2 and β5 ligation products by lysine and acetylated lysine, calculated from MALDI experiments. **D)** Ligation efficiencies for the formation of YLGD-KLISV and YLGD-KLI using non-acetylated or N^{α}-acetylated C-terminal ligation fragments. Ligation efficiencies were calculated from LC-MS experiments. Ligation efficiency was defined as the percentage of cleavage that resulted in ligation as opposed to hydrolysis.

amino acids are added to digestion mixtures, we investigated whether N^{α}-acetylated peptides could serve as C-terminal ligation partners. The N-terminal precursor YLGD|SY has been established as a β5 splicing precursor (where the vertical line indicates the site of cleavage; Chapter 3.2). YLGD|SY was incubated with proteasome and a 5-fold excess of either unprotected or N^{α}-acetylated KLI or KLISV to form the hydrolysis product YLGD and the transpeptidation products YLGD-KLI/YLGD-$K^{\alpha(Ac)}$LI or YLGD-KLISV/YLGD-$K^{\alpha(Ac)}$LISV (where a dash indicates the newly formed (iso)peptide bond). The resulting digestion mixtures were analyzed by LC-MS. Ligation with non-acetylated KLI and KLISV occurred with equal efficiency (Figure 1D), suggesting that peptide bond formation is independent

of the length of the C-terminal ligation fragment, as has been observed before (Chapter 3.2). Ligation with N^{ε}-acetylated peptides gave similar results compared to reactions using non-acetylated peptides (data not shown). Ligation of N^{α}-acetylated peptides on the other hand occurred with 10-fold lower efficiency compared to non-acetylated peptides (Figure 1D), in contrast with results found with single amino acids. The length of the N^{α}-acetylated C-terminal ligation fragment did not influence ligation efficiency (Figure 1D), as was also observed with non-acetylated peptides. Together these data suggest that the proteasome can form isopeptide bonds between two peptide fragments, but that it favors α-ligation over ε-ligation.

ε-ligation can compete with α-ligation during proteasomal splicing events

Next, we studied whether the ε-amino group of an N-terminal lysine residue can compete with the α-amino group for ligation when both amines are unprotected. To this end, a ligation experiment was performed using a C-terminal ligation fragment with a ^{15}N-enriched ε-amine. A ligation product containing a ^{15}N-enriched amide will result from ε-ligation, whereas α-ligation will result in a ligation product containing a ^{15}N-enriched amino group (Figure 2A). As amides and amines are distinguishable by NMR spectroscopy, this enabled us to study to which extent ε-ligation occurred. ^{15}N-ε-Fmoc-Lys(Boc)-OH was synthesized and incorporated into $K^{\varepsilon 15N}$LI by standard solid phase peptide synthesis. Subsequently, a 5-fold excess of the N-terminal precursor YLGD|SY was incubated with proteasome and $K^{\varepsilon 15N}$LI to form YLGD-$K^{\varepsilon 15N}$LI, where the dash indicates the newly formed (iso)peptide linkage. The resulting digestion mixture was analyzed using LC-MS. Figure 2B shows absorption spectra recorded at 280 nm of the reaction mixture before incubation (left panel) and after incubation (second panel). In addition to the N-terminal precursor YLGD|SY (eluting at 11.9 min), the hydrolysis product YLGD (eluting at 10.0 min) and different ligation products could be detected in the digestion mixture after incubation (Figure 2B, second panel). MS analysis of the peak eluting at 13.7 min showed that this peak contained two peptides with masses of 821 Da (M_1) and 913.8 Da (M_2) (Figure 2B, third panel), which corresponded to the ligation products YLGD-$K^{\varepsilon 15N}$LI and YLGD-YLGD. The peak eluting at 14.6 min. contained a third ligation product with a mass of 1163.6 Da (M_3) (Figure 2B, fourth panel), corresponding to YLGD-YLGDSY.

Next, the ligation mixture was purified using HPLC and as YLGD-$K^{\varepsilon 15N}$LI and YLGD-YLGD were co-eluting and thus not separated using HPLC, NMR spectra were recorded of the fraction containing both peptides. First, we investigated whether both α-linked and ε-linked YLGD-$K^{\varepsilon 15N}$LI were formed during the ligation reaction. If both types of ligation would occur, a ^{15}N NMR spectrum should show two distinct peaks, one corresponding to the amine, and one to the amide (Figure 2A). As direct detection of ^{15}N atoms suffers from low sensitivity and long relaxation times,[14] ^{15}N signals are usually inverse-detected via the proton signal in a ^{1}H-^{15}N heteronuclear single quantum coherence (HSQC) experiment using a fixed $J_{H,N}$ coupling constant. Whereas amide groups can be readily detected in an HSQC experiment via the $^{1}J_{H,N}$ coupling, NMR characterization of lysine NH_3 groups via the $^{1}J_{H,N}$ coupling is challenging, due to rapid hydrogen exchange with water (Supplementary Information and Supplementary Figure 1).[15] In accordance, the ^{15}N amine signal could not be detected in HSQC experiments using $J_{H,N}$ coupling constant ranging from 55 to 90 Hz (data not shown), which covered the whole range of reported amine $^{1}J_{H,N}$ coupling constants (55-80 Hz).[14] When

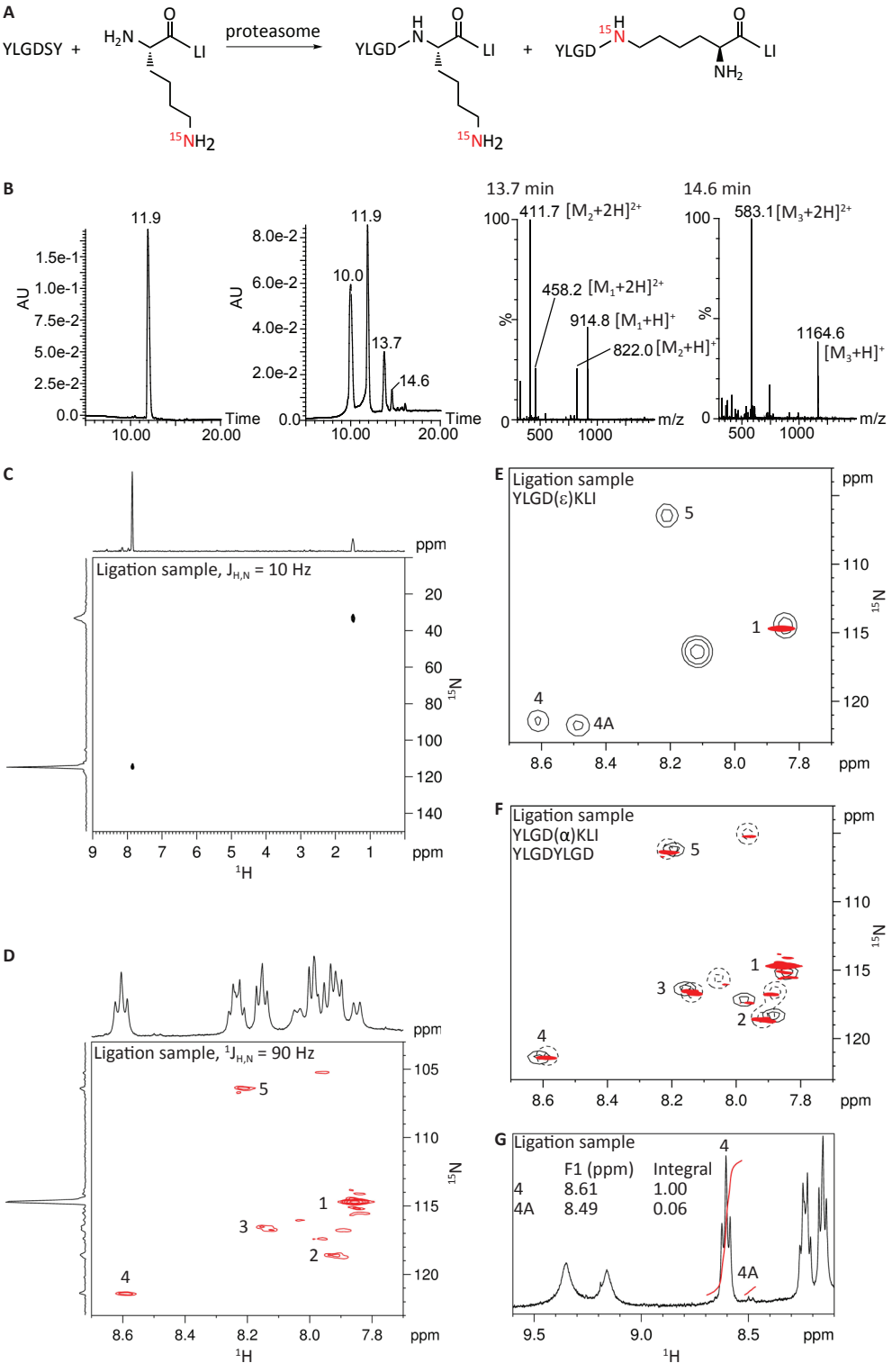

A
YLGDSY +
proteasome
YLGD
LI
B
11.9
10.0
13.7
14.6
AU
Time
13.7 min
411.7 [M2+2H]2+
458.2 [M1+2H]2+
914.8 [M1+H]+
822.0 [M2+H]+
14.6 min
583.1 [M3+2H]2+
1164.6
[M3+H]+
m/z
C
Ligation sample, JH,N = 10 Hz
D
Ligation sample, 1JH,N = 90 Hz
E
Ligation sample
YLGD(ε)KLI
F
Ligation sample
YLGD(α)KLI
YLGDYLGD
G
Ligation sample
F1 (ppm) Integral
4 8.61 1.00
4A 8.49 0.06
ppm
15N
1H

Figure 2 | ε-ligation can compete with α-ligation during proteasomal splicing reactions. **A)** Scheme showing the proteasomal ligation reaction of YLGDSY with $K^{\varepsilon 15N}$LI. **B)** HPLC profile (UV detection at 280 nm) of the ligation mixture before incubation (left panel) and after incubation (second panel). In-line MS spectra of the peaks eluting at 13.7 min (third panel) and 14.6 min (right panel). M_1: YLGD-KLI; M_2: YLGD-YLGD; M_3: YLGD-YLGDSY. **C)** ^{1}H-^{15}N HMQC spectrum of the ligation sample, recorded using $J_{H,N}$ = 10 Hz. 1D projections of the ^{1}H and ^{15}N signals are shown on top and left axes, respectively. **D)** ^{1}H-^{15}N HSQC spectrum of the ligation sample, recorded using $^1J_{H,N}$ = 90 Hz. The 1D projection of the ^{15}N signal is shown on the left axis, the ^{1}H spectrum of the ligation sample is shown on the top axis. Peaks were assigned as follows, 1: $^{\varepsilon}$K5, 2: not assigned, 3: L2/D4, 4: L2/D4, 4A: not assigned, 5: G3 (See Supplementary Information and Supplementary Figure 2 for details). **E)** Overlay of the ^{1}H-^{15}N HSQC spectra of the $^{15}N^{\varepsilon}$-enriched ligation sample (red) and synthetic, non ^{15}N enriched YLGD$^{\varepsilon}$KLI (black) **F)** Overlay of the ^{1}H-^{15}N HSQC spectra of the $^{15}N^{\varepsilon}$ enriched ligation sample (red), synthetic non ^{15}N-enriched YLGD$^{\varepsilon}$KLI (black, solid lines), and synthetic non ^{15}N enriched YLGDYLGD (black, dotted lines). **G)** ^{1}H spectrum of the ligation sample and integrals of peaks 4 and 4A. Peak 4 is attributed to L2 or D4 and is present in the ^{1}H spectra all products, whereas peak 4A is present in the ^{1}H spectrum of YLGD$^{\varepsilon}$KLI only (Supplementary Figures 2 and 3)

^{15}N signals are on the other hand inverse-detected via long-range $^2J_{H,N}$ and $^3J_{H,N}$ couplings ($J_{H,N}$ = 0.3 to 16 Hz),[14] both amide and amine resonances can be measured simultaneously in a single experiment (Supplementary Information and Supplementary Figure 1). Therefore, we performed a ^{1}H-^{15}N HMQC measurement, using a $J_{H,N}$ coupling constant of 10 Hz. The resulting HMQC spectrum showed two peaks, with shifts of 7.85; 114.7 and 1.45; 33.2 ppm (Figure 2C). Amines display a chemical shift of 0 to 70 ppm, while amides have a chemical shift of 80 to 170 ppm, compared to a $^{15}NH_3$ reference signal.[14] This indicates that during proteasomal digestion of YLGDSY in the presence of $K^{\varepsilon 15N}$LI, ligation products containing ^{15}N-enriched amines and ^{15}N-enriched amides were both formed. From these data we conclude that during proteasomal splicing reactions ε-ligation competes with α-ligation.

ε-ligation occurs in 1 out of 10 ligation events

Next, we set out to quantify the relative amount of ε-ligation product formed. It is however not possible to directly compare the ^{15}N amide and ^{15}N amine signals in a single HSQC measurement. ^{1}H-^{15}N correlations of the lysine NH_3 group are often difficult to observe in HSQC experiments optimized for backbone amides, even when the water exchange rate is slow enough to permit their detection by ^{1}H NMR.[15] This is due to two factors. First, the ^{15}N chemical shift of lysine NH_3 groups resonates ~ 90 ppm upfield from the backbone amide signals. As the rf strength that can be used for ^{15}N pulses is limited, application of a ^{15}N 180° pulse at 115 ppm (close to the amide resonance frequency) leads to an imperfect pulse at 30 ppm (close to the lysine NH_3 resonance frequency) and therefore to loss of signal intensity, and vice versa.[15] With standard HSQC pulse programs it is therefore not possible to obtain maximal signal intensities for ^{15}N amide and ^{15}N amine groups simultaneously. Second, the ^{15}N transverse relaxation of lysine NH_3 groups has been shown to be highly affected by water exchange, resulting in broadening of ^{15}N line shapes and a decrease in resolution and sensitivity,[15] as visible in Figure 2C. Therefore, the relative amount of ε-ligation product compared to α-ligation product was estimated by quantifying the signal intensity of the newly formed ^{15}N enriched amide bond

using amide resonances within the peptide due to naturally abundant ^{15}N as an internal reference. To ensure maximum signal intensities of amide resonances, the YLGD-K$^{\varepsilon 15N}$LI containing fraction was analyzed again in an HSQC experiment using a $^{1}J_{H,N}$ coupling constant of 90 Hz.[16] In the resulting spectrum, one intense cross peak was visible at 7.86; 114.7 ppm, which originated from the ^{15}N enriched isopeptide amide in YLGD-$^{\varepsilon}$KLI (Figure 2D, peak 1). In addition, several weaker signals were detected, which originate from naturally abundant ^{15}N amide signals of different ligation products.

The ligation sample contained the following ligation products: YLGD-$^{\alpha}$KLI, YLGD-$^{\varepsilon}$KLI and YLGD-YLGD as indicated by LC-MS (where a dash indicates the newly formed pepide bond and α/ε the type of linkage) (Figures 2B and 2C). We synthesized these three peptides and recorded ^{1}H and ^{1}H-^{15}N HSQC spectra as a reference to assign ^{1}H-^{15}N cross-peaks to one or more peptides (Supplementary Information and Supplementary Figure 2). An overlay of the ^{1}H-^{15}N HSQC spectra of the ^{15}N enriched ligation sample and synthetic YLGD$^{\varepsilon}$KLI confirmed that peak 1 resulted from the ^{15}N enriched $^{\varepsilon}$K5 resonance (Figure 2E). In addition, the overlay of HSQC spectra of the ligation sample, synthetic YLGD$^{\alpha}$KLI, and synthetic YLGDYLGD showed that all weaker signals could be attributed to naturally abundant ^{15}N amide resonances in one or more ligation products (Figure 2F). In the HSQC spectrum of the ligation sample, 56% of the signal of overlapping cross-peaks originated from YLGD-YLGD, while 44% originated from YLGD-$^{\alpha/\varepsilon}$KLI (Supplementary Information and Supplementary Figure 3). We integrated the signals of the $^{\varepsilon}$K5 cross peak (peak 1) and the 4 most intense naturally abundant cross-peaks (peaks 2 to 5) and calculated the signal intensity originating from YLGD-$^{\alpha/\varepsilon}$KLI resonances only (Table 1; 44% of total integrals). Taking the abundance of ^{15}N signals into account, the relative percentage of ε-ligation (YLGD-$^{\varepsilon}$KLI) versus α-ligation (YLGD-$^{\alpha}$KLI) was estimated to be 10.3 ± 1.7% (average ± SD of peaks 2 to 5). This relative percentage of ε-ligation could also be derived from the ^{1}H spectra of the ligation sample (Figure 2G). In this spectrum, the triplet at 8.61 ppm originates from naturally abundant peak 4, which is present in all peptides, while the doublet at 8.49 ppm originates from naturally abundant peak 4A, which is only present in YLGD$^{\varepsilon}$KLI (Supplementary Figure 2D). From the integral of these resonances, we derived that YLGD-$^{\varepsilon}$KLI constituted 6% of total ligation products (YLGD-YLGD + YLGD-$^{\alpha}$KLI + YLGD-$^{\varepsilon}$KLI) and estimated the relative percentage of ε-ligation versus α-ligation to be 14%. Both percentages also correlate with the LC-MS experiments described above, in which relative ε-ligation efficiencies of 10% were observed. Together, these data indicate that in one out of 8 to 10 ligation reactions ε-ligation occurred as opposed to α-ligation.

Table 1 | Integrals of ^{15}N enriched and naturally abundant amide cross-peaks.

Peak	Residue	f2 (ppm)	f1 (ppm)	Total Integral	YLGD-KLI only	^{15}N abundance
1	$^{\varepsilon}$K5	7.8552	114.6679	209654420	209654420	98%
2	-	7.9136	118.6336	20745315	9127938	0.37%
3	L2/D4	8.1374	116.6248	20142896	8862874	0.37%
4	L2/D4	8.5903	121.4045	15826456	6963640	0.37%
5	G3	8.2106	106.3924	14876321	6545581	0.37%

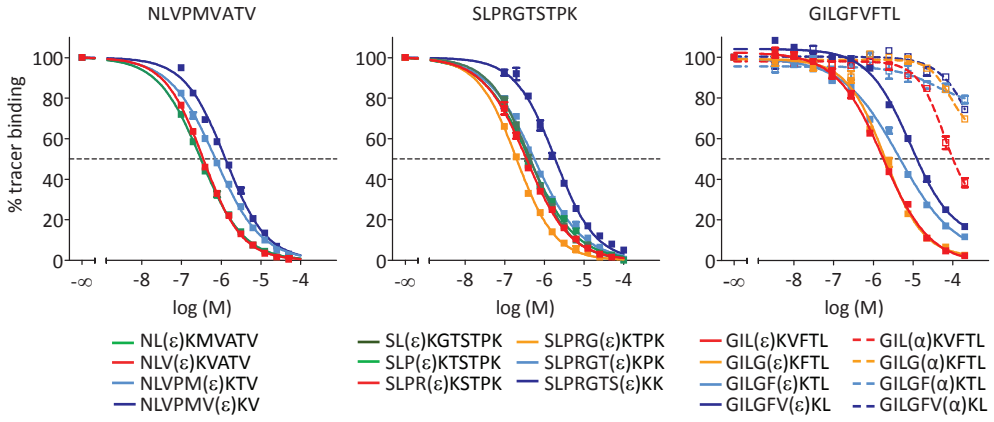


Figure 3 | Isopeptides have increased affinity for HLA-A complexes. Binding curves of (iso)peptides based on the epitopes NLVPMVATV (specific for HLA-A2.1), SLPRGTSTPK (specific for HLA-A3) and GILGFVFTL (specific for HLA-A2.1). (ε)K indicates the location of the isopeptide linkage. (α)K indicates a normally linked lysine residue.

Isopeptides have increased affinity for HLA-A complexes

It is likely that epitopes containing an isopeptide linkage have binding affinities for HLA-A molecules that differ from their non-isopeptide counterparts. To test whether epitopes containing an isopeptide bond could still bind HLA-A complexes, we measured the affinity of a set of isopeptide bond-containing peptides for HLA-A2.1 and HLA-A3. Isopeptides were based on three known epitopes, GILGFVFTL (influenza matrix protein 1 58-66),[17] NLVPMVATV (human cytomegalovirus pp65 495-503),[18] which are both presented on the cell surface by HLA-A2.1, and SLPRGTSTPK (spliced together from SP110 nuclear body protein 296-301 and 286-289),[10] which is presented by HLA-A3. As the length of an N^{ε}-linked lysine residue in a peptide backbone is twice that of an N^{α}-linked lysine residue (See Figure 2A), we replaced two adjacent amino acid residues at different positions in the epitope by one N^{ε}-linked K (Table 2). To ensure binding, the anchor residues on P2 and P9 (P10 in case of SLPRGTSTPK) were not modified.

To investigate binding to HLA-A2.1 or HLA-A3, a UV-mediated MHC exchange fluorescence polarization (FP) assay was used.[19,20] In this assay, HLA-A2.1/HLA-A3 complexes loaded with a UV-cleavable peptide ligand are subjected to UV light, which cleaves the ligand, resulting in the generation of "empty" HLA-A2.1/HLA-A3.[21] These "empty" HLA-A2.1/HLA-A3 complexes quickly disintegrate, unless they are rescued by the addition of a 'rescue' peptide. When rescue is performed by adding both a fluorescent tracer peptide and increasing concentrations of a 'competitor' peptide, the affinity of this competitor peptide can be measured.[19,20] Binding curves of the NLVPMVATV-, SLPRGTSTPK- and GILGFVFTL-based isopeptides were measured (Figure 3, solid lines) and the affinities (expressed as IC_{50} values), defined as the concentration of peptide that inhibited 50% of tracer peptide binding, were determined (Table 2). For all potential epitopes, the presence of an isopeptide linkage did not abrogate HLA-A2.1/HLA-A3 binding, as all peptides were able to bind to HLA-A2.1 or HLA-A3 complexes with similar

Table 2 | IC_{50} values of different (iso)peptides for binding to HLA-A2.1 and HLA-A3.

Parent peptide	Replaced positions	Isopeptide	IC_{50} (μM)	Peptide	IC_{50} (μM)
GILGFVFTL	-	-	1.7	-	-
	3-4	GI(ε)KFVFTL	ND	GI(α)KFVFTL	ND
	4-5	GIL(ε)KVFTL	1.8	GIL(α)KVFTL	111
	5-6	GILG(ε)KFTL	2.0	GILG(α)KFTL	509
	6-7	GILGF(ε)KTL	4.8	GILGF(α)KTL	577
	7-8	GILGFV(ε)KL	14.5	GILGFV(α)KL	601
NLVPMVATV	-	-	4.7		
	3-4	NL(ε)KMVATV	0.31		
	4-5	NLV(ε)KVATV	0.35		
	5-6	NLVP(ε)KATV	ND		
	6-7	NLVPM(ε)KTV	0.71		
	7-8	NLVPMV(ε)KV	1.3		
SLPRGTSTPK	-	-	0.19		
	3-4	SL(ε)KGTSTPK	0.44		
	4-5	SLP(ε)KTSTPK	0.42		
	5-6	SLPR(ε)KSTPK	0.35		
	6-7	SLPRG(ε)KTPK	0.20		
	7-8	SLPRGT(ε)KPK	0.55		
	8-9	SLPRGTS(ε)KK	1.9		

(ε)K indicates the location of the isopeptide linkage. (α)K indicates a normally linked lysine residue. ND: not determined

affinities as the parent peptides (Table 2). In general, the isotope linkage was tolerated in the middle of the epitope (Figure 3, red and orange lines) or towards the N-terminus of the epitope (Figure 3, green lines) without loss of HLA class I affinity. When the isopeptide linkage was shifted towards the C-terminal side of the epitope on the other hand, affinity decreased. For comparison, the affinities of GILGFVFTL-based epitopes with identical sequences but a normal peptide bond were determined. As these peptides are 8 amino acids in length, they are expected to have low affinity to HLA-A2.1, which only binds peptides of 9-10 amino acids in length with high affinity. As can be judged from Figure 3 (dotted lines) and Table 2, the binding affinity of these GILGFVFTL-based peptides was decreased by a factor 40 to 200 compared to their isopeptide counterparts. Together, these data suggest that isopeptides are not only formed by the proteasome, but isopeptides of different lengths can also be presented at the cell surface by different HLA class I alleles. In addition, the presence of an isopeptide linkage enables 8-mer isopeptides to bind HLA-A2.1/HLA-A3 complexes with much increased affinity as compared to normal 8-mer peptides that do not contain an isopeptide linkage.

Isopeptides have enhanced stability towards proteasomal degradation

Intramolecular isopeptide bonds have been shown to provide proteolytic stability to pili on the surface of Gram-positive bacteria,[22-24] suggesting that epitopes containing an isopeptide linkage may be less susceptible to processing by proteases. Therefore, we investigated whether isopeptides showed enhanced stability to proteasomal processing

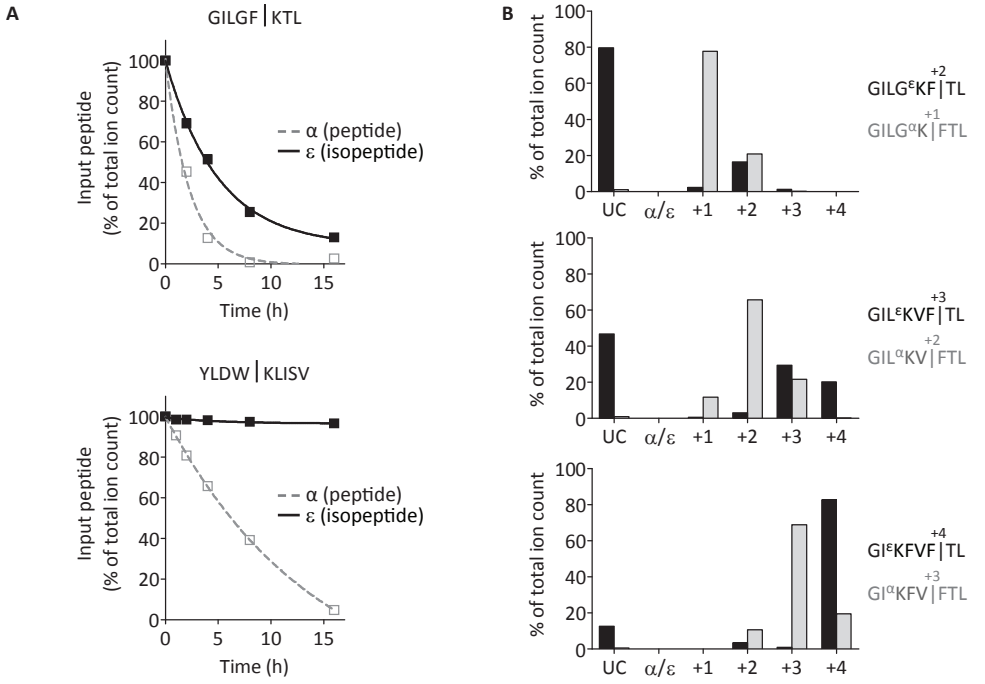


Figure 4 | Isopeptides show enhanced stability towards proteasomal degradation. A) Degradation of GILGF$^{\alpha/\varepsilon}$KTL and YLDW$^{\alpha/\varepsilon}$KLISV by the proteasome as a function of time. The amounts of input peptides and degradation products were quantified using LC-MS. The site of cleavage and the location of the lysine N$^{\alpha}$- or N$^{\varepsilon}$-linkage is indicated with a vertical line. **B)** Proteasomal degradation patterns of the indicated peptides. The amounts of input peptides and degradation products were quantified using LC-MS after incubation with proteasome for 16 h. Black bars indicate isopeptides and corresponding degradation products, grey bars indicate normal peptides and corresponding degradation products. The main site of cleavage is indicated with a vertical line. The location of the lysine N$^{\alpha}$- or N$^{\varepsilon}$-linkage is indicated with α or ε. Cleavage products are indicated by the site of cleavage relative to the lysine N$^{\alpha}$- or N$^{\varepsilon}$-linkage. UC: uncleaved input peptide.

over normal peptides. First, we investigated peptides in which cleavage of the (iso)peptide bond was the only proteasomal cleavage occurring. To this end, YLDW$^{\alpha/\varepsilon}$KLISV and GILGF$^{\alpha/\varepsilon}$KTL were incubated with proteasome for time periods up to 16 h, followed by LC-MS analysis of the degradation mixtures (where α/ε indicates the location of the lysine N$^{\alpha}$- or N$^{\varepsilon}$-linkage). The amounts of the different input peptides were plotted as the percentage of total ion count versus time (Figure 4A). GILGF$^{\alpha}$KTL (dotted lines) was cleaved readily ($t_{1/2}$ ~ 2 h), while GILGF$^{\varepsilon}$KTL (solid lines) was cleaved at a slower rate ($t_{1/2}$ ~ 4 h). The difference in hydrolysis rate between N$^{\alpha}$-linked and N$^{\varepsilon}$-linked lysine residues was even more apparent in YLDWKLISV. Whereas YLDW$^{\alpha}$KLISV was cleaved with a half-life of 6 h, less then 5% of YLDW$^{\varepsilon}$KLISV was cleaved within 16 h.

Next, we questioned whether isopeptides that were cleaved at positions other then at the isopeptide bond were also more stable to proteasomal processing compared to normal peptides. We incubated GI$^{\alpha/\varepsilon}$KFVFTL, GIL$^{\alpha/}$

$^{\varepsilon}$KVFTL, GILG$^{\alpha/\varepsilon}$KFTL and GILGFV$^{\alpha/\varepsilon}$KL for 16 h with proteasome and analyzed the resulting digestion mixtures with LC-MS to assess the extent as well as the site of peptide cleavage. The amounts of input peptide and cleavage products were quantified by LC-MS and plotted as the percentage of total ion count, which is a measure of total peptide content (Figure 4B). Cleavage products were indicated by the site of cleavage relative to the lysine N$^{\alpha}$- or N$^{\varepsilon}$-linkage (e.g. +1 cleavage products result from cleavage 1 residue towards the C-terminus from $^{\alpha/\varepsilon}$K). For all peptides in which the isopeptide linkage was located towards the N-terminus from the site of cleavage, differences were found both in the rate and the site of hydrolysis. The difference between peptide and isopeptide processing was most apparent when the isopeptide linkage was located within 1 to 2 residues of the site of cleavage (Figure 4B, top panel). The peptide GILG$^{\alpha}$KFTL was processed almost completely within 16 h and was cleaved predominantly at the KF bond, which is located one residue towards the C-terminus from $^{\alpha}$K (+1 cleavage products). Only 20% of GILG$^{\varepsilon}$KFTL was processed within 16 h on the other hand, and the main cleavage site in GILG$^{\varepsilon}$KFTL was the FT bond located two residues (corresponding to a length of three normal residues) from the isopeptide linkage (Figure 4B, top panel). Similar patterns were found when GIL$^{\alpha/\varepsilon}$KVFTL (Figure 4B, middle panel) and GI$^{\alpha/\varepsilon}$KFVFTL (Figure 4B, lower panel) were compared. Both GIL$^{\alpha}$KVFTL and GI$^{\alpha}$KFVFTL were completely processed within 16 h, while only 50% and 90% of GIL$^{\varepsilon}$KVFTL and GI$^{\varepsilon}$KFVFTL, respectively, were cleaved within the same time frame (Figure 4B). In line with results described above, the site of cleavage in isopeptides was located further away from the lysine residue compared to the normal peptides. Whereas GIL$^{\alpha}$KVFTL and GI$^{\alpha}$KFVFTL were cleaved preferably 2 and 3 residues towards the C-terminus from the N$^{\alpha}$-linked lysine, GIL$^{\varepsilon}$KVFTL and GI$^{\varepsilon}$KFVFTL were cleaved 3 and 4 residues away from the N$^{\varepsilon}$-linked lysine residue (Figure 4B). No differences in peptide composition were found between digestion mixtures of GILGFV$^{\alpha}$KL and GILGFV$^{\varepsilon}$KL, which are the only peptides in this series in which the $^{\alpha/\varepsilon}$K was located towards the C-terminus from the site of cleavage (data not shown). Together, these data suggest that peptides containing an isopeptide linkage are less susceptible to proteasomal degradation. This holds true both if the isopeptide linkage is the preferred site of cleavage and if the isopeptide linkage is located close to the preferred cleavage site. In addition, when proteasomal processing of isopeptides did occur, the actual cleavage site was located further away from the $^{\varepsilon}$K residue compared to an $^{\alpha}$K resiude.

DISCUSSION

Recent studies have shown that proteasomal splicing reactions can create a novel type of antigen via a transpeptidation mechanism.[8-11] During proteasomal transpeptidation reactions, a protein amide backbone linkage is attacked by one of the proteasomes catalytic N-terminal threonine residues, resulting in the formation of an *O*-acyl enzyme intermediate and the release of the C-terminal part of the protein. Subsequently, this intermediate ester reacts with an amino group, usually the amino terminus, from a second peptide (the 'C-terminal ligation partner'), resulting in the formation of a novel peptide linkage and a spliced product.[9,10] This transpeptidation reaction can apparently compete with normal hydrolysis, in which water reacts with the intermediate ester,[12] in such a way that sufficiently new peptide can be formed to invoke an immune response against transpeptidation products. Transpeptidation reactions can efficiently compete with hydrolysis if specific

structural requirements on the two ligating partners are met, as is discussed in Chapter 3.2. Ligation occurs particularly efficient if the C-terminal ligation partner has a basic amino acid residue (lysine or arginine) at the N-terminus (Chapter 3.2). As lysine has two amino groups that can theoretically both react with the *O*-acyl enzyme intermediate, this implies that the proteasome may be able to from isopeptide linkages.

In the present study we show for the first time that the proteasome can use both the α-amino group and the ε-amino group of lysine as a C-terminal ligation partner, suggesting that also *in vivo* the proteasome can create an isopeptide linkage and form a novel type of antigen. When single amino acids are ligated onto a precursor peptide, both amino groups were equally reactive and capable of participating in ligation reactions. However, when peptide ligation was monitored by NMR, ε-ligation was found to be 10-fold less efficient as compared to α-ligation. Likely, a single lysine amino acid is capable of fitting into the S_1' pocket in multiple orientations, thereby facilitating both α- and ε-ligation. A peptide on the other hand will adopt an orientation whereby its P_2' and P_3' side chains fit in the S_2' and S_3' pockets. This may favor orientation of the lysine residue in the S_1' pocket in such a way that the α-amino group is positioned correctly for a nucleophilic attack on the intermediate ester bond, resulting in normal peptide bond formation, disfavoring alternative ε-ligation and isopeptide bond formation.

The data presented here also show that isopeptides have unique properties that discern them from normal epitopes. First, isopeptides of various lengths can bind different HLA alleles with high affinity. The isopeptide linkage is tolerated in the middle and towards the N-terminus of different epitopes without loss of affinity for HLA-A2/A3 complexes. Only location of the isopeptide linkage towards the C-terminus of an epitope leads to loss of HLA-A2/A3 affinity. In addition, the presence of an isopeptide linkage allows peptides that are comprised of only 8 amino acids to bind HLA-A2/A3, whereas normally only peptides of 9 to 11 residues in length bind HLA-A2/A3 complexes with high affinity. Second, isopeptides are more stable towards further proteasomal processing compared to normal peptides. Degradation of isopeptides is slow compared to common peptides, both when the isopeptide bond itself is the preferred site of cleavage, and when the preferred site of cleavage is located close to the isopeptide linkage. In addition, the proteasome cleaves most isopeptides at a site that is located further away from the lysine residue compared to normal peptides. This indicates that proteasomal processing of isopeptides results in the generation of a peptide fragment in which the isopeptide linkage is still present and located towards the N-terminus of the peptide, favoring affinity for HLA-A2/A3. Together, these properties are likely to increase the fraction of ε-ligated peptides that enters the endoplasmic reticulum, and that is loaded onto MHC class I and transported to the cell surface for CD8+ T cell surveillance. As isopeptides are more stable towards proteasomal degradation, we consider it likely that the isopeptide linkage also hampers hydrolysis by other (cytosolic) proteases. Importantly, isopeptides that are formed and subsequently processed by the proteasome are likely to retain their isopeptide linkage at a location (towards the N-terminus) that does not affect HLA-A2/A3 affinity. Therefore, we anticipate an increased chance that an isopeptide containing epitope is efficiently loaded onto MHC allowing it to reach the cell surface once it is formed.

It has been proposed that proteasomal antigen splicing increases the diversity of the peptide repertoire available for presenta-

tion on MHC class I.[12,25] A greater diversity of epitopes increases the potential epitope repertoire that may be recognized by CD8+ T cells, which ultimately results in the elimination of infected or malignant cells.[12] Increasing diversity of T cell receptors by V(D)J recombination is essential for protection from a constantly changing bacterial and viral repertoire.[25] In addition, proteasomal splicing may increase diversity on the protein level. The data presented here suggest that also proteasomal $^{\varepsilon}$K-ligation reactions may increase immune diversity. To date, three spliced epitopes have been described *in vivo*,[8-10] suggesting that proteasomal splicing does occur in cells and with sufficient efficiency to evoke an immune response against spliced epitopes. The splicing efficiencies of two of the three described epitopes have been calculated and one spliced epitope was estimated to be produced from $\sim 5 \times 10^5$ to 1×10^4 molecules of precursor peptide (0.0002% to 0.01% efficiency).[9,11] This indicates that even very low splicing efficiencies lead to sufficient amounts of peptide for detection by CD8+ T cells. In a previous study, we have shown that splicing efficiencies can be several orders of magnitude higher if specific structural requirements for the two ligating partners are met (Chapter 4.2). We estimated that ligations reaction involving C-terminal fragments with a lysine residue at the site of ligation occurred with efficiencies of 1 to 10% (Chapter 4.2). This indicates that the overall efficiency of $^{\varepsilon}$K-ligation reactions, which is 10-fold lower as compared to α-ligation, can be as high as 0.1 to 1%, depending on the sequence of N- and C-terminal ligation partners, suggesting that one isopeptide may be formed from 100 to 1000 precursor molecules. Since a large fraction of spliced isopeptide epitopes may be able to reach the cell surface as a result of their increased proteolytic stability, ε-ligation should produce sufficient amounts of peptide to evoke a T cell response, suggesting that isopeptides may play a role in immunity *in vivo*. Thus although the epsilon linkage has not been evidenced before in MHC epitopes yet, which is easily explained by its difficult (and so far unanticipated) detection, we consider it more than likely that they will be detected, based on our *in vitro* experiments. With evidence of proteolytic stability and therefore accumulation of this modification we postulate it as a genuine post-translational modification resulting from transpeptidation mechanisms. We consider it likely that our *in vitro* findings will be validated *in vivo* soon.

MATERIALS AND METHODS

Peptide building blocks were purchased from Novabiochem and appropriately functionalized resins from Applied Biosystems. ^{15}N-ε-L-Lysine HCl was purchased from C/D/N Isotopes Inc (Pointe-Claire, QC, Canada). All solvents were purchased from Biosolve at the highest grade available. All other chemicals were purchased from Aldrich at the highest available purity. All solvents and chemicals were used as received. LC-MS analyses were carried out on a WATERS LCT mass spectrometer in line with a WATERS 2795 HPLC system and a WATERS 2996 photodiode array detector. All ^{1}H and ^{1}H-^{15}N HSQC (Heteronuclear Single Quantum Coherence) experiments were carried out in DMSO-d6 using a Bruker Avance 300 (^{1}H: 300 MHz, ^{15}N: 30.4 MHz) spectrometer. HMQC (Heteronuclear Multiple Quantum Coherence) experiments were performed on a Bruker ARX 400 (^{1}H: 400 MHz, ^{15}N: 40.5 MHz) spectrometer. NMR data were processed and analyzed using Bruker v2.1 Topspin software.

Peptide synthesis

Peptides were synthesized using standard Fmoc-based solid-phase peptide synthesis protocols and appropriately functionalized PEG-polystyrene Wang resins. Functionalized resins were subjected to coupling cycles, in which deprotection of the Fmoc-group with piperidine/NMP (1/4 v/v), was followed by coupling with 4 equivalents each of Fmoc protected amino acid, di-isopropyl

ethylamine (dipea) and (Benzotriazol-1-yloxy) tripyrrolidinophosphonium hexafluorophosphate (PyBop). Reactions were carried out in NMP at a volume of 1 mL/0.1 g of resin. After the final coupling step, the Fmoc group was removed and peptides were either directly fully deprotected and released from the resin directly by treating the resin with TFA/H_2O/triisopropylsilane (93/5/2 v/v) for 2.5 h or peptides were treated with 4 equivalents each of acetic anhydride and dipea for 40 minutes to acetylate the N-terminus prior to deprotection and release from the resin. Peptides containing N^{ε}-linked lysine residues were synthesized using Boc-L-lysine(Fmoc)-OH. Incorporation of this building blocked followed by removal of the Fmoc group ensures that peptide synthesis continues at the Lysine ε-amine group, resulting in isopeptide linkage containing peptides. Peptides were precipitated with cold diethylether/pentane (3/1 v/v) and lyophilized from H_2O/ACN/HOAc (65/25/10 v/v).

Synthesis of ^{15}N-ε-Fmoc-Lys(Boc)-OH

N^{α}-(9-Fluorenyl)methoxycarbonyl-*$^{15}N^{\varepsilon}$-tert*-butyloxycarbonyl-L-lysine was synthesized as described.[26] ^{1}H NMR (400 MHz, d_6-dmso) δ = 12.53 (bs, COOH), 7.91 (d, *J*=7.5 Hz, 2H, H_{Ar}), 7.73 (d, *J*=7.5 Hz, 2H, H_{Ar}), 7.62 (d, *J*=7.6 Hz, 1H, NH), 7.42 (t, *J*=7.5 Hz, 2H, H_{Ar}), 7.33 (t, *J*=7.5 Hz, 2H, H_{Ar}), 6.79 (dt, J_{NH}=92 Hz, J_{HH}=6.8 Hz, 1H, $^{15}N^{\varepsilon}H$), 4.29-4.26 (m, 2H, OCH_2), 4.24-4.20 (m, 1H, $OCH_2C\underline{H}$), 3.93-3.88 (m, 1H, H^{α}), 2.91-2.78 (m, 2H, CH_2), 1.70-1.52 (m, 2H, CH_2), 1.36 (s, 9H, $C(CH_3)_3$), 1.33-1.27 (m, 4H, 2xCH_2). ^{13}C NMR APT (300 MHz, $CDCl_3$) δ = 175.51 (COOH), 156.29 (2x CO), 143.74 (2x$C_{q\text{-}Ar}$), 141.30 (2x$C_{q\text{-}Ar}$), 127.69 (2xCH_{Ar}), 127.08 (2xCH_{Ar}), 125.14 (2xCH_{Ar}), 119.95 (2xCH_{Ar}), 79.57 ($\underline{C}CH_3$), 67.04 (CH_2O), 53.71 ($C^{\alpha}H$) 47.16 ($\underline{C}HCH_2O$), 40.15 ($C^{\varepsilon}H_2$, $^1J_{CN}$=11 Hz), 31.76 ($C^{\beta}H_2$), 29.54 ($C^{\delta}H_2$), 28.41 ((CH_3)$_3$), 22.31 ($C^{\gamma}H_2$). $^{15}N^1$H-HMQC (6.79 ppm, 83.7 ppm). MS (ESI): $[M+H]^+_{calc}$=470.2, $[M+H]^+_{found}$=470.1.

Proteasome purification

Proteasome was purified from bovine liver as described.[4] After each step, proteasome purity was monitored by incubating fractions with a fluorescent proteasome activity probe, followed by SDS-PAGE and scanning of the resulting gel for fluorescence emission, as described.[27] Briefly, bovine liver was homogenized in phosphate buffered saline (PBS), followed by precipitation using 40% saturated ammonium sulphate to remove unwanted proteins. The proteasome was subsequently precipitated by increasing the concentration of saturated ammonium sulphate to 60%. Following dialysis, the proteasome was further purified using a 10-40% sucrose gradient and anion exchange column chromatography, using DEAE Sephadex A25 resin. Proteasome containing fractions were pooled, concentrated and protein concentrations were determined using the Bradford assay (Biorad).

MALDI experiments

78 pmol C^{TMR}SLPRGTASSR was incubated with 10 μg proteasome in ligation buffer (50 mM Tris-HCl, pH 8.5) containing 6% DMSO for 16 h at 45°C. In different experiments, 100 μM MG132, 100 nM bortezomib, or 0.5 M lysine, 0.5 M N^{α}-acetylated lysine or 0.5 M N^{ε}-acetylated lysine were added to the reaction mixture. Following incubation, reaction mixtures were freeze-dried and dissolved in 50 μL acetonitrile/H_2O (5:95 v/v) containing 0.1 % TFA. Of this solution 10 μL was used for purification and desalting using reversed-phase ZipTip®C18 tips (Millipore, C18, spherical silica, 15 um, 200 A pore size). Peptide samples were bound to the ZipTip pipette tip by aspirating and dispensing 10 μL sample for approximately 20 cycles until the solution appeared nearly colorless and the ZipTip pipette tip had turned pink due to retained labeled peptides. The ZipTip pipette tip was washed twice with 10 μL of 0.1% TFA. Retained peptides were eluted by aspirating and dispensing 4 μL acetonitrile through the ZipTip pipette tip at least five times (the ZipTip pipette tip turned colorless). Eluates were analyzed by MALDI mass spectrometry. MALDI-TOF experiments were carried out on an Autoflex, linear MALDI-TOF-MS (Bruker Daltonik GmbH, Bremen, Germany). Droplets of 0.5 μL 10 mg/ml 2,5-dihydroxybenzoic acid (DHB, Bruker Daltonik) matrix solution in 0.1% TFA were spotted onto a MALDI target plate and 0.5 μL of the purified and desalted samples were pipetted into the DHB matrix droplets and left drying. For calibration, Peptide Calibration Standard (M 1046-3147 Da, Bruker Daltonik) was used. Spectra were analyzed with Bruker Daltonics FlexAnalysis Software. Ligation efficiencies were calculated using the following formula: Efficiency = $I_L/(I_L+I_H)$ × 100%, where I_L is

the peak intensity of the ligation product and I_H is the peak intensity of the hydrolysis product.

Peptide ligation experiments

To compare acetylated and non-acetylated peptides, 50 nmol of the peptide YLGDSY was incubated with 250 nmol of (acetylated) C-terminal ligation fragment and 5 μg proteasome in a total volume of 75 μL ligation buffer (50 mM Tris-HCl, pH 8.5) containing 2% DMSO for 16 h at 37 °C. Incubation mixtures were snap-frozen, lyophilized, dissolved in DMSO/H_2O/ACN (1/2/1 v/v) and filtered over Strata™-X 33 μm polymeric reversed phase disposable columns (Phenomenex). Eluted fractions were analyzed by LC-MS.

For NMR experiments, 16 μmol $^{\varepsilon(15N)}$KLI was incubated with 80 μmol YLGDSY and 2.8 mg proteasome in 23.2 mL ligation buffer containing 4.8% DMSO for 24 h at 37°C. The reaction mixture was snap-frozen, lyophilized, dissolved in 12 mL acetonitrile/H_2O (1:1) and filtered over a Strata™-X 33 μm polymeric reversed phase disposable column (Phenomenex) to remove proteasomal proteins. The filtrate was lyophilized again, dissolved in 7.5 mL DMSO/H_2O (1:4) and diluted and purified by HPLC over an Atlantis® Preparative T3 column (5 μm; 20 x 150 mm; Waters) using a linear gradient (5% to 50% acetonitrile in H_2O containing 0.05% TFA). HPLC purifications were performed on a Shimadzu LC-20AT prominence liquid chromatography system, coupled to a Shimadzu SPD-20A prominence UV/Vis detector and a Shimadzu CTO-20A prominence column oven. The YLGD-$^{\varepsilon(15N)}$KLI containing fraction was lyophilized and dissolved in 0.4 mL DMSO-d6 for NMR spectroscopy analysis. Reference spectra of synthetic YLGD$^{\alpha}$KLI, YLGD$^{\varepsilon}$KLI and YLGDYLGD were recorded using a 50 mM peptide solution in DMSO-d6. HSQC experiments were performed using a Bruker preprogrammed gradient enhanced echo-antiecho HSQC pulse program[28,29] (hsqcetgp, avance-version, 2D H-1/X correlation via double inept transfer, phase sensitive using Echo/Antiecho-TPPI gradient selection with decoupling during acquisition, using trim pulses in inept transfer) optimized for the detection of ^{15}N signals and ^{15}N signals were compared to a $^{15}NH_3$ reference signal. The relative percentage of ε-ligation versus α-ligation was calculated using the following formula: Efficiency of ε-ligation = I_1/(0.44×I_N) × 0.37/98 × 100%, where I_1 is the integral of the $^{15}N^{\varepsilon}$-enriched peak 1, I_N is the integral of peak N, resulting from a naturally abundant amide resonance, 0.44 is the fraction YLGD-$^{\alpha/\varepsilon}$KLI in the ligation sample as compared to total peptide content, 0.37 is the percentage of naturally abundant ^{15}N and 98 is the percentage of ^{15}N enrichment of $^{15}N^{\varepsilon}$-enriched peak 1. The efficiency of ε-ligation was calculated for peaks 2 to 5 and averaged.

Fluorescence polarization HLA-A2.1 and HLA-A3 binding assay

Fluorescence polarization (FP) HLA-A2.1 and HLA-A3 binding assays with UV-mediated peptide exchange were performed as described previously[19,20] with minor changes. For the HLA-A2.1 binding assay, KILGFVFJ_1V (where J_1 is UV-cleavable 3-amino-3-(2-nitrophenyl)propionic acid) was used as conditional ligand and fluorescently labeled FLPSDC^{TMR}FPSV (where TMR is maleimido-linked tetramethylrhodamine) was used as tracer peptide. For the HLA-A3 binding assay, RIYRJ_1GATR was used as conditional ligand and KVPC^{TMR}ALINK was used as tracer peptide. Serial 2- or 3-fold dilutions of different (iso)peptides were used as competitor peptides. Assays were performed in 384-well low-volume black nonbinding surface assay plates (Corning 3820). Wells were loaded with 15 or 30μL total volume containing 0.5 μM HLA-A-peptide complex, 1 nM tracer peptide and serial dilutions of competitor peptide in PBS containing 0.5 mg/mL bovine γ-globulin. The plate was spun for 1 min at 1000g at room temperature to ensure proper mixing of all components. To start UV-mediated cleavage of the conditional ligand and peptide exchange, the plate was placed under a 365-nm UV lamp at 10 cm distance (366 nm UV lamp, 2 × 15 W blacklight blue tubes, l×w×h 505 × 140 × 117 mm, Uvitec, UK) located in a cold room (4 °C). After 30 min irradiation, the plate was sealed with thermowell sealing tape (Corning) and incubated at room temperature for 4 hours, allowing cleaved peptide fragments to dissociate and to be exchanged for competitor and/or tracer peptide. Subsequently, fluorescence polarization measurements were performed, as described.[19] The binding affinity (IC_{50} value) of each competitor peptide was defined as the concentration that inhibited 50% of tracer peptide binding. Data were analyzed using GraphPad Prism software

(GraphPad).

(Iso)peptide stability assays

50 nmol YLDW$^{\alpha}$KLISV or YLDW$^{\varepsilon}$KLISV was incubated with 5 μg proteasome in ligation buffer containing 1-2% DMSO at 37 °C for different time periods up to 26 h. Incubation mixtures were snap-frozen, lyophilized and filtered over Strata™-X 33 μm polymeric reversed phase disposable columns (Phenomenex). Filtered samples were analyzed by LC-MS.

ACKNOWLEDGMENTS

The authors thank Olaf van Tellingen and Dorothe Linders for help with MALDI measurements, Henk Hilkmann for peptide synthesis, Jacques Neefjes for help with proteasome purification and Ton Schumacher for advice. This work was funded by the Dutch Cancer Society (grant NKI 2005-3368) and NWO/ALW.

REFERENCES

1. Adams, J. The development of proteasome inhibitors as anticancer drugs. *Cancer Cell* **5**, 417-421 (2004).
2. Nalepa, G., Rolfe, M. & Harper, J.W. Drug discovery in the ubiquitin-proteasome system. *Nat. Rev. Drug Discov.* **5**, 596-613 (2006).
3. Shastri, N., Schwab, S. & Serwold, T. Producing nature's gene-chips: the generation of peptides for display by MHC class I molecules. *Annu. Rev. Immunol.* **20**, 463-493 (2002).
4. Raijmakers, R. et al. Automated online sequential isotope labeling for protein quantitation applied to proteasome tissue-specific diversity. *Mol. Cell Proteomics* **7**, 1755-1762 (2008).
5. Strehl, B. et al. Interferon-gamma, the functional plasticity of the ubiquitin-proteasome system, and MHC class I antigen processing. *Immunol. Rev.* **207**, 19-30 (2005).
6. Dahlmann, B., Ruppert, T., Kuehn, L., Merforth, S. & Kloetzel, P.M. Different proteasome subtypes in a single tissue exhibit different enzymatic properties. *J. Mol. Biol.* **303**, 643-653 (2000).
7. Dahlmann, B., Ruppert, T., Kloetzel, P.M. & Kuehn, L. Subtypes of 20S proteasomes from skeletal muscle. *Biochimie* **83**, 295-299 (2001).
8. Hanada, K., Yewdell, J.W. & Yang, J.C. Immune recognition of a human renal cancer antigen through post-translational protein splicing. *Nature* **427**, 252-256 (2004).
9. Vigneron, N. et al. An antigenic peptide produced by peptide splicing in the proteasome. *Science* **304**, 587-590 (2004).
10. Warren, E.H. et al. An antigen produced by splicing of noncontiguous peptides in the reverse order. *Science* **313**, 1444-1447 (2006).
11. Dalet, A., Vigneron, N., Stroobant, V., Hanada, K. & Van den Eynde, B.J. Splicing of distant Peptide fragments occurs in the proteasome by transpeptidation and produces the spliced antigenic peptide derived from fibroblast growth factor-5. *J. Immunol.* **184**, 3016-3024 (2010).
12. Berkers, C.R., de Jong, A., Ovaa, H. & Rodenko, B. Transpeptidation and reverse proteolysis and their consequences for immunity. *Int J. Biochem. Cell Biol.* **41**, 66-71 (2009).
13. Berkers, C.R. et al. Activity probe for in vivo profiling of the specificity of proteasome inhibitor bortezomib. *Nat. Methods* **2**, 357-362 (2005).
14. Marek, R. & Lycka, A. N-15 NMR spectroscopy in structural analysis. *Current Organic Chemistry* **6**, 35-66 (2002).
15. Iwahara, J., Jung, Y.S. & Clore, G.M. Heteronuclear NMR spectroscopy for lysine NH(3) groups in proteins: unique effect of water exchange on (15)N transverse relaxation. *J. Am. Chem. Soc.* **129**, 2971-2980 (2007).
16. Tjandra, N., Grzesiek, S. & Bax, A. Magnetic field dependence of nitrogen-proton J splittings in N-15-enriched human ubiquitin resulting from relaxation interference and residual dipolar coupling. *J. Am. Chem. Soc.* **118**, 6264-6272 (1996).
17. Falk, K., Rotzschke, O., Stevanovic, S., Jung, G. & Rammensee, H.G. Allele-specific motifs revealed by sequencing of self-peptides eluted from MHC molecules. *Nature* **351**, 290-296 (1991).
18. Solache, A. et al. Identification of three HLA-A*0201-restricted cytotoxic T cell epitopes in the cytomegalovirus protein pp65 that are conserved between eight strains of the virus. *J. Immunol.* **163**, 5512-5518 (1999).
19. Bakker, A.H. et al. Conditional MHC class I ligands and peptide exchange technology for the human MHC gene products HLA-A1, -A3, -A11, and -B7. *Proc. Natl. Acad. Sci. U S A* **105**, 3825-3830 (2008).
20. Rodenko, B. et al. Class I major histocompatibility complexes loaded by a periodate trigger. *J. Am. Chem. Soc.* **131**, 12305-12313 (2009).
21. Toebes, M. et al. Design and use of conditional MHC class I ligands. *Nat. Med.* **12**, 246-251 (2006).
22. Budzik, J.M. et al. Intramolecular amide bonds stabilize pili on the surface of bacilli. *Proc. Natl. Acad. Sci. U S A* **106**, 19992-19997 (2009).
23. Kang, H.J., Coulibaly, F., Clow, F., Proft, T. & Baker, E.N. Stabilizing isopeptide bonds revealed in gram-positive bacterial pilus structure. *Science* **318**, 1625-1628 (2007).

24. Kang, H.J. & Baker, E.N. Intramolecular isopeptide bonds give thermodynamic and proteolytic stability to the major pilin protein of Streptococcus pyogenes. *J. Biol. Chem.* **284**, 20729-20737 (2009).
25. Shastri, N. Cell biology. Peptides, scrambled and stitched. *Science* **313**, 1398-1399 (2006).
26. Wiejak, S., Masiukiewicz, E. & Rzeszotarska, B. A large scale synthesis of mono- and di-urethane derivatives of lysine. *Chemical & Pharmaceutical Bulletin* **47**, 1489-1490 (1999).
27. Berkers, C.R. et al. Profiling proteasome activity in tissue with fluorescent probes. *Mol. Pharm.* **4**, 739-748 (2007).
28. Boyd, J., Soffe, N., John, B., Plant, D. & Hurd, R. The Generation of Phase-Sensitive 2D N-15-H-1 Spectra Using Gradient Pulses for Coherence-Transfer-Pathway Selection. *Journal of Magnetic Resonance* **98**, 660-664 (1992).
29. Davis, A.L., Keeler, J., Laue, E.D. & Moskau, D. Experiments for Recording Pure-Absorption Heteronuclear Correlation Spectra Using Pulsed Field Gradients. *Journal of Magnetic Resonance* **98**, 207-216 (1992).

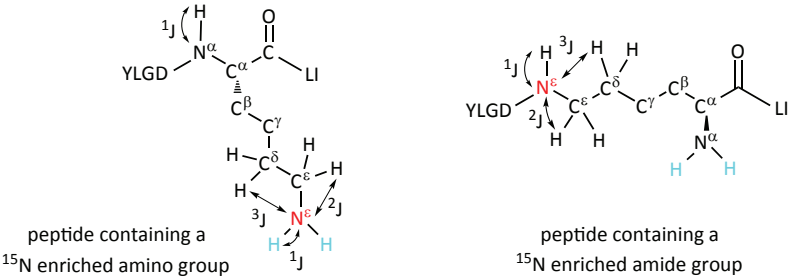
Supplementary information

SUPPLEMENTARY RESULTS

^{1}H-^{15}N HSQC and HMQC experiments

As direct detection of ^{15}N by NMR is difficult due to low naturally abundance of ^{15}N, low relative sensitivity and low relaxation times, the ^{15}N signal is frequently inverse-detected in ^{1}H-^{15}N HSQC (Heteronuclear Single Quantum Correlation) or HMQC (Heteronuclear Multiple Quantum Correlation) experiments. The correlation data provided by either HSQC or HMQC experiments are essentially equivalent and provide the same information. In inverse-detected experiments, ^{1}H nuclei that couple to ^{15}N nuclei are used for both the excitation and the detection of the ^{15}N signal. The ^{15}N signal is therefore detected via the ^{1}H signal. The resulting 2D spectrum contains one cross-speak for each proton that couples to a specific ^{15}N nucleus. Depending on the coupling constant J that is used in the experiment, ^{1}H-^{15}N coupling through one bond ($^{1}J_{H,N}$ ~ 90 Hz)[1] or through bonds (two or three bonds; $J_{H,N}$ = 0.3 to 16 Hz)[1] is utilized (Supplementary Figure 1). When coupling through one bond is utilized, the resulting 2D spectrum contains one cross-peak for each proton that is directly attached to a nitrogen. When coupling over multiple-bonds (^{2}J or ^{3}J) is utilized on the other hand, cross peaks are between protons and nitrogen atoms that can be either two or three bonds away.

In protein NMR, ^{1}H-^{15}N HSQC experiments are routinely performed to detect backbone amide resonances. Each residue in a protein, with the exception of proline, has an amide proton that is directly attached to a nitrogen and can therefore be used to inversely detect the ^{15}N signal. We performed ^{1}H-^{15}N HSQC experiments (J = 90 Hz) to detect backbone amide resonances in small peptides. In these experiments, both naturally abundant back-



Supplementary Figure 1 | Detection of amine and amide groups in HSQC experiments. Scheme showing ^{1}H-^{15}N couplings (J) that can be used to detect lysine's amide and amine groups in ^{1}H-^{15}N HSQC and HMQC experiments. Protons that exchange with the solvent are indicated in blue, other relevant protons are indicated in black, ^{15}N enriched nitrogen is indicated in red. Naturally abundant backbone amide groups (indicated in black) can be detected via the ^{1}J coupling (via the directly attached proton), while the amide group of a N$^{\varepsilon}$-linked lysine residue can be detected via ^{1}H-^{15}N coupling through one bond (^{1}J) or through multiple bonds (^{2}J or ^{3}J). As amino groups readily exchange protons with the solvent, the ^{15}N amine resonance cannot be inversely detected via the directly attached protons (^{1}J coupling). Lysine's free amino ^{15}N$^{\varepsilon}$ signal can be detected by utilizing ^{1}H-^{15}N coupling through multiple bonds.

bone amide resonances and ^{15}N enriched N^{ε}-linked lysine amide resonances can be detected (Supplementary Figure 1). By comparing the signal intensities of the naturally abundant backbone amide resonances to the intensity of the ^{15}N enriched N^{ε}-linked lysine signal, the relative amounts of different peptides in the sample can be determined.

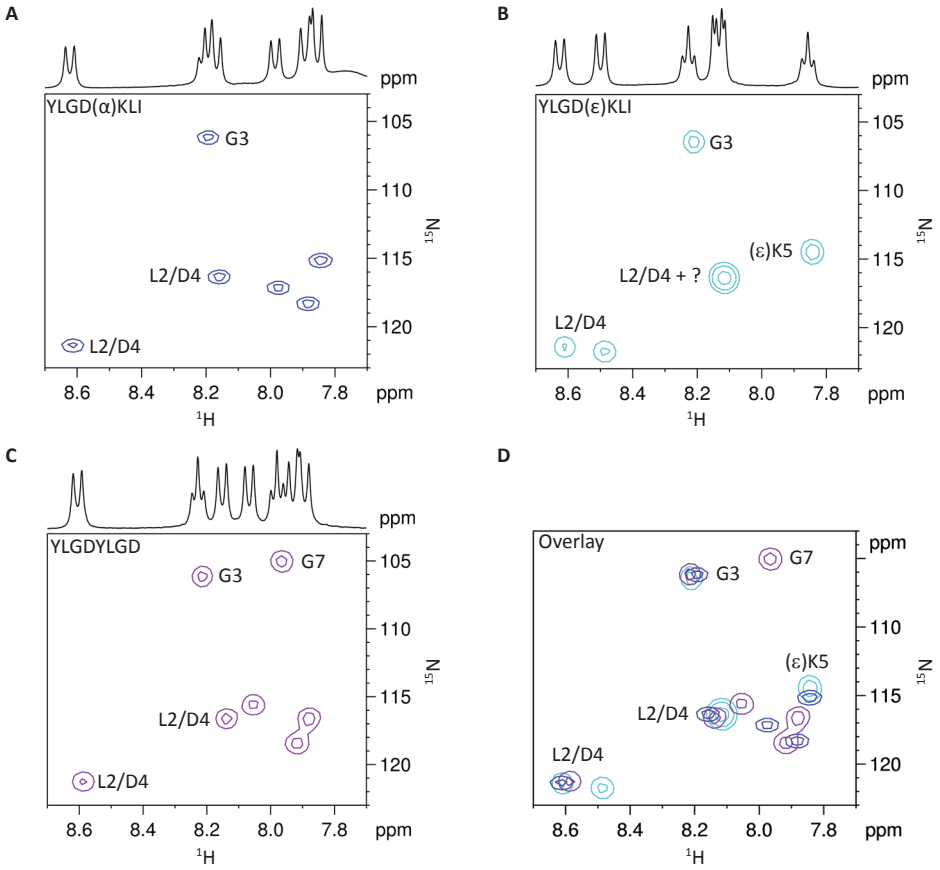
Whereas amide groups can be readily detected in an HSQC experiment via the $^{1}J_{H,N}$ coupling (J = 90 Hz), NMR characterization of lysine's free amino groups via the $^{1}J_{H,N}$ coupling is challenging. In aqueous solution, amino groups rapidly exchange protons with the solvent (Supplementary Figure 1), which hampers the detection of the amino protons in ^{1}H experiments and therefore the detection of amino $^{15}N^{\varepsilon}$ signals via these protons. Lysine's free amino $^{15}N^{\varepsilon}$ signal can be detected via protons that do not exchange with the solvent, e.g. the protons attached to lysine's C^{δ} or C^{ε}. Also the amide resonance of $^{15}N^{\varepsilon}$-linked lysine can be detected via these protons (Supplementary Figure 1). To visualize the ^{15}N enriched amide and amine resonances simultaneously, we therefore performed an HMQC experiment with a coupling constant of 10 Hz. In this experiment, the ^{15}N signal is detected via long range $^{2}J_{H,N}$ and $^{3}J_{H,N}$ couplings and the cross peaks are between the ^{15}N enriched lysine N^{ε} and the protons attached to the lysine C^{ε} and/or C^{δ}. Therefore, the resulting 2D spectrum of this experiment will show both the ^{15}N enriched amide and amine resonances. The amide and amine signal intensities measured in this experiment can however not be compared quantitatively, as explained in the main text.

Assignment of ^{1}H-^{15}N cross-peaks

To assign ^{1}H-^{15}N cross-peaks to one or more peptides, we synthesized YLGD$^{\alpha}$KLI, YLGD$^{\varepsilon}$KLI and YLGDYLGD and recorded ^{1}H and ^{1}H-^{15}N HSQC spectra as a reference (Supplementary Figure 2). In the HSQC reference spectra of both YLGD$^{\alpha}$KLI and YLGDYLGD all backbone amides are visible as separate cross-peaks, while in the spectrum of YLGD$^{\varepsilon}$KLI two backbone amide peaks overlap (Supplementary Figure 2). In all spectra several ^{1}H-^{15}N cross-correlations could be attributed to specific residues. Three cross-correlations were visible in all three spectra (8.2; 106, 8.1; 116, and 8.6; 121 ppm, Supplementary Figure 2D). Glycine ^{15}N signals resonate somewhat upfield (at 106 ppm) compared to other amino acids.[2] In addition, glycine amide protons are expected to show triplets in ^{1}H spectra due to the presence of two protons on C^{α}. The cross-peaks at 8.2; 106 ppm showed this upfield shift and characteristic splitting pattern (see ^{1}H spectra in Supplementary Figures 2B and 2C) and we therefore attributed this signal to G3, which is present in all peptides. As all peptides have YLGD as their first four residues, the other overlapping signals likely also originated from this part of the peptides (L2 and D4). The ^{1}H spectrum of YLGDYLGD (Supplementary Figure 2C) showed a second triplet at 8.0 ppm, and therefore the corresponding cross-peak in the ^{1}H-^{15}N HSQC spectrum (8.0; 105 ppm) was attributed to G7. The ^{1}H spectrum of YLGD$^{\varepsilon}$KLI (Supplementary Figure 2B) showed a second triplet at 7.86 ppm. As this splitting is expected for ε-linked K amide protons (due to the presence of two protons on the adjacent C^{δ}), the corresponding cross-peak in the ^{1}H-^{15}N HSQC spectrum (7.8; 114.6 ppm) was attributed to $^{\varepsilon}$K5.

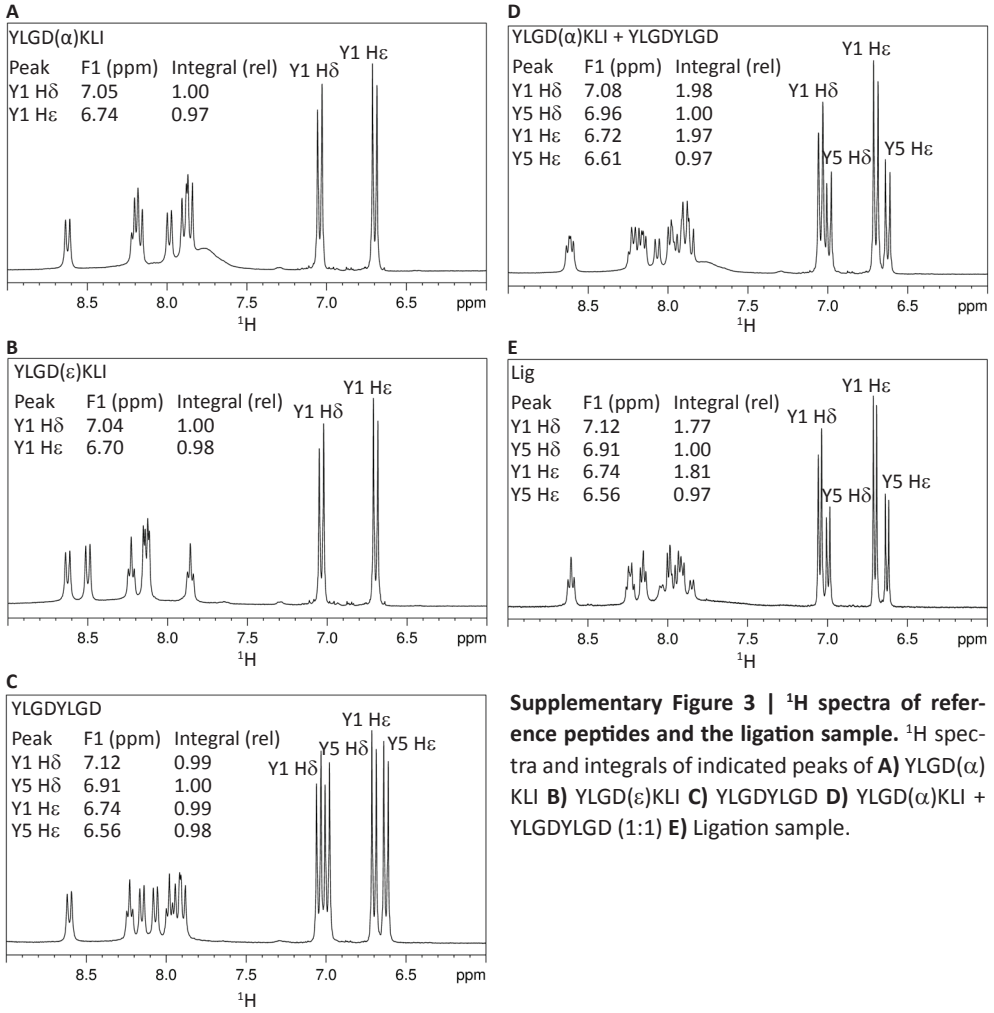
Estimation of the relative abundance of different ligation products in the ligation sample

As the most intense cross peaks due to naturally abundant amide resonances originate from all three peptides present in the ligation sample (Figure 2F), an estimation of the relative amounts of products in this sample is needed



Supplementary Figure 2 | HSQC spectra of reference peptides. A-C) ^{1}H-^{15}N HSQC spectra of YLGD(α)KLI **(A)** YLGD(ε)KLI (**B)** YLGDYLGD **(C)**. ^{1}H spectra of the corresponding peptides are shown on the top axis. **D)** Overlay of ^{1}H-^{15}N HSQC spectra of YLGD(α)KLI, YLGD(ε)KLI, and YLGDYLGD.

to estimate the efficiency of ε-ligation. To this end, the ^{1}H reference spectra of YLGD$^{\alpha}$KLI, YLGD$^{\varepsilon}$KLI and YLGDYLGD were compared to the ^{1}H spectrum of the ligation mixture (Supplementary Figure 3). All peptides contained one or two tyrosine residues, of which the Hδ and Hε protons are resonating at 7.0 and 6.6 ppm, respectively. Whereas YLGDKLI and YLGD$^{\varepsilon}$KLI show near identical resonances for Y1 Hδ and Hε (Supplementary Figures 3A and 3B), the ^{1}H spectrum of YLGDYLGD shows additional resonances for Y5, which are shifted somewhat downfield compared to Y1 and have equal intensities (Supplementary Figure 3C). The sum of the ^{1}H spectra of YLGDYLGD and YLGD$^{\alpha}$KLI (Supplementary Figure 3D) or YLGD$^{\varepsilon}$KLI (data not shown) indicates that the integrals of the Y1 signals are twice as high as the integrals of the Y5 signals if equal amounts of both peptides are used, as expected. In the ^{1}H spectrum of the ligation sample, the integrals of Y1 are only 1.8 times as intense as the integrals from Y5 (Supplementary Figure 3E). Therefore, we concluded that in the HSQC spectrum of the ligation sample, 56% of the signal of overlapping cross-peaks originated from YLGD-YLGD, while 44% originated from YLGD-$^{\alpha/\varepsilon}$KLI. In addition to the amount of pep-



Supplementary Figure 3 | ^{1}H spectra of reference peptides and the ligation sample. ^{1}H spectra and integrals of indicated peaks of **A)** YLGD(α)KLI **B)** YLGD(ε)KLI **C)** YLGDYLGD **D)** YLGD(α)KLI + YLGDYLGD (1:1) **E)** Ligation sample.

tide, the intensity of any ^{1}H-^{15}N cross-peak depends on the abundance of ^{15}N nuclei in these peptides. The naturally abundance of ^{15}N and thus of any ^{15}N amide is 0.37%, whereas the newly formed Lys N$^{\varepsilon}$-linkage is 98% ^{15}N enriched.

SUPPLEMENTARY REFERENCES

1. Marek, R. & Lycka, A. N-15 NMR spectroscopy in structural analysis. *Current Organic Chemistry* **6**, 35-66 (2002).
2. Wishart, D. S., Bigam, C. G., Holm, A., Hodges, R. S. & Sykes, B. D. H-1, C-13 and N-15 Random Coil Nmr Chemical-Shifts of the Common Amino-Acids .1. Investigations of Nearest-Neighbor Effects. *Journal of Biomolecular Nmr* **5**, 67-81 (1995).

Chapter 3.4

Summary and prospects: proteasome splicing

Summary and prospects: proteasome splicing

Following malignant transformation or viral infection, a broad repertoire of antigenic peptides is presented on the cell surface by MHC class I, enabling CD8+ T cells to sense the change in intracellular protein content, which eventually leads to killing of the antigen presenting cell. The 26S proteasome is responsible for the generation of these antigenic peptides *in vivo*. Recently, it has become apparent that not only contiguous antigens, but also non-contiguous antigens, that consist of two post-translationally fused peptides, can be presented to the immune system. The ligation that produces these epitopes is thought to be catalyzed by the proteasome *in vivo* via a transpeptidation mechanism, which makes it likely that splicing is not a random event, but governed by distinct splicing rules. The elucidation of such splicing rules will facilitate the prediction and detection of spliced antigens, which will enable studying the roles of such spliced epitopes in immunity.

To be able to study proteasome splicing on a molecular level, large amounts of purified proteasome are required. In **chapter 2.3** we described how the fluorescent proteasome activity probes described in chapter 1 can be used to monitor proteasome purity during large scale proteasome isolation from bovine liver and spleen. Using mass spectrometry, we subsequently showed that highly purified preparations of either constitutive or immunoproteasome could be obtained via this method. In **chapter 3.2**, we used these proteasome preparations to decipher splicing rules. By analyzing proteasome digestion mixtures of small libraries of potential ligation partners, we were able to show that splicing motifs exist. Subsequently, we confirmed and refined these motifs by application to a polypeptide derived from the Influenza A neuraminidase 1 protein. In addition, we showed that proteasomal splicing reactions can occur with high efficiencies of up to 30% *in vitro* if structural requirements for proteasome splicing are met by both ligation partners and that many spliced epitopes may be formed from a single protein. Finally, we examined the mechanism of splicing reactions, and showed that *in vivo,* intrastrand splicing of peptides derived from the same protein is likely preferred over random ligation of peptides derived from different proteins. In **chapter 3.3**, we investigated in more detail a specific type of proteasomal ligation reaction, in which the C-terminal ligation partner has lysine at the site of ligation. As lysine has two amino groups that can theoretically both participate in proteasomal transpeptidation reactions, this implies that the proteasome may be able to from isopeptide linkages. Using NMR, we showed that both the α- and the ε-amino groups of lysine were able to participate in ligation reactions, and we could determine that ε-ligation was favored over α-ligation in 10% of all splicing reactions involving lysine at the site of ligation. This suggested that the proteasome can create isopeptide linkages and may therefore form a novel type of antigen that may play a role in immunity *in vivo*. In addition, we showed that isopeptide linkage containing epitopes have unique properties that discern them from normal epitopes. First, isopeptides

of various lengths can bind to HLA-A2.1 and HLA-A3 proteins with high affinity. Second, isopeptides are more stable towards further proteasomal processing as compared to normal peptides. The combination of these properties suggests that a large fraction of isopeptide epitopes may reach the cell surface to evoke an immune response.

The data described in **chapters 3.2 and 3.3** collectively indicate that different types of proteasomal ligation reactions readily occur *in vitro*. The efficiencies of these reactions are high compared to the estimated ligation efficiencies of splicing reactions described *in vivo*, suggesting that proteasomal splicing reactions may occur more frequently than assumed thus far and may significantly contribute to immunity *in vivo*. To date, only three spliced epitopes have been described *in vivo*, as identifying spliced antigens *de novo* is not possible with the methodologies currently used to identify novel antigens. Novel epitopes can be identified by cell surface elution of all self- and foreign epitopes, followed by MS analysis and matching against protein databases. Alternatively, epitopes may be predicted from protein sequences and tested for T cell receptor recognition, e.g. by incorporation into fluorescent MHC class I tetramers, which are subsequently used to stain T cells in Fluorescence assisted cell sorting (FACS) assays. However, as spliced antigens are not contiguous and their formation requires post-translational splicing, they neither match existing databases nor can they be predicted from contiguous protein sequences, which hampers their identification. The splicing rules deduced in chapter 3.2 should facilitate the identification of more spliced epitopes, and several types of experiments that aim at identifying both normal and $^{\varepsilon}$K linked spliced antigens are currently ongoing.

To show that our splicing rules can predict splicing *in vivo*, we have expressed optimized precursor peptides in cells. If the predicted proteasomal splicing reactions occur, spliced epitopes are formed, loaded onto MHC class I and transported to the cell surface, where they can be identified using two strategies. The first strategy is based on elution of all epitopes from the cell surface, followed by LC-MS analysis. The successful identification of spliced epitopes requires the ability to detect and sequence low femtomolar amounts of individual peptides in highly complex mixtures eluted from the cell surface. Typically, the number of copies of a particular peptide presented by MHC molecules on a single cell ranges from 1 to 10.000, and therefore, MS identification of spliced epitopes is very challenging. Very low copy numbers of MHC-presented epitopes can already activate CD8+ T cells, and T cell activation assays can therefore also be used to identify spliced epitopes. To perform these assays, mice transgenic for HLA-A2 (A2-Kb) are first vaccinated with the anticipated splicing product to obtain specific CD8+ T cells. These T cells are subsequently added to the cells that are transfected with an optimized precursor peptide *in vitro*. The presence of the expected splicing product on the cell surface will activate the T cells, which can be monitored in interferon-γ release assays. Although T cell activation assays are very sensitive, the expected spliced product has to be immunogenic to evoke a T cell response, which may complicate the identification of predicted spliced products. Finally, vaccination experiments are currently ongoing to show that splicing occurs readily *in vivo*. In these experiments, mice are first vaccinated with a DNA vaccine encoding splicing precursors. At the top of the immune response, blood is analyzed for the presence of T cells that are reactive against the spliced product.

Identifying isopeptide linkage containing epitopes *in vivo* is even more challenging, as mass spectrometry cannot distinguish between N^{ε}- and N^{α}-linked lysine residues. Isopeptide epitopes have unique properties that can be used to distinguish them from normal peptides, but the amounts of epitope that are eluted from the surface of antigen presenting cells is generally not sufficient to perform secondary assays. Therefore, only T cell activation assays can be used to identify this novel type of antigen and we are currently in the process of peptide vaccinating HLA-A2 transgenic mice with both normal epitopes and isopeptide epitopes to generate T cells that are specific for either normal or isopeptide epitopes. By using these T cells in combination with cells that are transfected with an optimized splicing precursor peptide, we hope to identify isopeptide epitopes *in vivo*

Finally, studies are currently ongoing that will further refine the splicing rules that are postulated in this thesis. Application of the methods developed in chapter 3.2 to immunoproteasomal digestion mixtures will likely generate a set of splicing rules for the immunoproteasome. In addition, knowledge of proteasomal cleavage sites will allow a more accurate prediction of both C- and N-terminal ligation partners. Algorithms that predict proteasomal cleavage are available, but they did not accurately predict hydrolysis of our model peptides. By analyzing a larger set of splicing precursors, rules that predict proteasomal cleavage can likely be deduced. We expect that the studies presented in this thesis will aid the discovery of novel spliced epitopes and that they will contribute to unraveling the role of spliced epitopes and isopeptide antigens *in vivo*.

Nederlandse samenvatting

Curriculum Vitae

List of publications

Nederlandse samenvatting

Proteasoom activiteit in de breedste zin van het woord is het onderwerp van dit proefschrift. In de cel worden voortduren eiwitten gevormd en weer afgebroken om processen die van belang zijn voor het normaal functioneren van de cel te sturen. De afbraak van eiwitten in een cel wordt strak gereguleerd door het ubiquitine proteasoom systeem, waarbij de daadwerkelijke eiwitafbraak plaatsvindt in en door het proteasoom. Het proteasoom is een groot eiwitcomplex, dat verantwoordelijk is voor het overgrote deel van de eiwitafbraak in een cel. In het proteasoom worden eiwitten afgebroken tot kleinere stukjes, genaamd peptides, die vervolgens door andere proteases afgebroken worden tot aminozuren, de bouwstenen van eiwitten. De precieze activiteit van het proteasoom wordt bepaald door de activiteit van zes verschillende actieve 'subunits' die ieder na andere aminozuursequenties knippen. Naast gewoon (constitutief) proteasoom bestaat er ook een vorm die vooral aanwezig is in weefsels die een rol spelen in het immuunsysteem: het immunoproteasoom.

Het proteasoom is niet alleen verantwoordelijk voor de afbraak van eiwitten die een sleutelrol spelen in processen als de cel cyclus, cel differentiatie, signaal transductie en apoptose, het proteasoom speelt ook een rol in immuniteit. Een breed repertoire aan antigene peptiden wordt gepresenteerd op het celoppervlakte aan het immuunsysteem door Major Histocompatibility Complex (MHC) klasse I eiwit complexen. Deze peptiden worden gegenereerd door het proteasoom en kunnen afkomstig zijn van zowel lichaamseigen als lichaamsvreemde eiwitten. Zo wordt als het ware een blauwdruk van de celinhoud getoond aan het immuunsysteem. Dit stelt CD8+ cytotoxische T cellen in staat veranderingen in de cel als gevolg van bijvoorbeeld maligne transformatie of een virale infectie op te merken. De herkenning van een lichaamsvreemd of abnormaal peptide door CD8+ cellen leidt uiteindelijk tot de eliminatie van een maligne of geïnfecteerde cel.

Het remmen van het proteasoom verstoort vele regulatoire processen in de cel, wat uiteindelijk tot celdood kan leiden. Dit maakt het proteasoom tot een aantrekkelijk farmacologisch target. De observatie dat kanker cellen gevoeliger zijn voor proteasoomremming dan normale cellen heeft geleid tot de ontwikkeling van verscheidene proteasoomremmers voor de behandeling van kanker. Het medicijn velcade®, waarin de proteasoomremmer bortezomib het actieve ingrediënt is, wordt gebruikt voor de behandeling van de ziekte van Kahler (multipel myeloom) en van mantelcel lymfoom, een agressieve en ongewone vorm van non-Hodgkin lymfoom. Behandeling met bortezomib kan echter leiden tot substantiële bijwerkingen en ook treedt resistentie tegen bortezomib regelmatig op. Daarom is het belangrijk te zoeken naar nieuwe proteasoomremmers die beter getolereerd worden en/of werkzaam zijn in cellen die resistent zijn tegen bortezomib. Verscheidene tweede-generatie proteasoomremmers, die onder andere verschillen in hun remmingsmechanisme en subunit specificiteit, worden momenteel getest

in klinische trials. Voorbeelden van deze remmers zijn marizomib en CEP-18770. Nu verschillende proteasoomremmers beschikbaar komen voor klinisch gebruik, is het van groot belang om optimale behandelmethodes te definiëren en de meest geschikte remmer voor een bepaalde patiënt te bepalen. Hiervoor zijn accurate methodes nodig die gebruikt kunnen worden om de specificiteit en effectiviteit van deze verschillende remmers in patiënten te monitoren. Voor de meest gangbare methode die momenteel wordt gebruikt om proteasoomactiviteit te meten is het echter noodzakelijk om cellen te lyseren (open te breken) voordat activiteitsmetingen uitgevoerd kunnen worden. Het lyseren van cellen leidt ertoe dat het proteasoom uit zijn regulatoire omgeving gehaald wordt, wat de proteasoomactiviteit kan beïnvloeden. Andere methodes kunnen alleen gebruikt worden in genetisch gemodificeerde cellen of organismen, wat het monitoren van patiënten die behandeld worden met proteasoomremmers uitsluit.

In **hoofdstuk 1** beschrijf ik de ontwikkeling en karakterisatie van verschillende proteasoom probes, synthetische verbindingen die covalent binden aan actieve proteasoomsubunits en een label bevatten waardoor ze gevisualiseerd kunnen worden. Deze probes kunnen gebruikt worden om proteasoomactiviteit te bestuderen in vele soorten weefsels en met een breed scala aan technieken.
Hoofdstuk 1.2 beschrijft de ontwikkeling van een gedansyleerde proteasoom probe die in combinatie met op SDS-PAGE gebaseerde assays gebruikt kan worden om proteasoomactiviteit te monitoren. Deze probe is de eerste in zijn soort die gebruikt kon worden om de activiteit van bortezomib in intacte cellen te bestuderen zonder de noodzaak om bewerkelijke radiolabeling procedures uit te voeren. De gelabelde subunits kunnen eenvoudig zichtbaar gemaakt worden door te immunoblotten met een antilichaam tegen de dansyl groep. We vonden substantiële verschillen tussen de specificiteit van bortezomib in intacte cellen en die in gelyseerde cellen. Dit benadrukt de noodzaak om bioassays te ontwikkelen die uitgevoerd kunnen worden in levende cellen om de accuraatheid van zulke me-tingen te kunnen waarborgen.
In **hoofdstuk 1.3** beschrijf ik de verdere optimalisatie van deze gedansyleerde proteasoom probe, door vervanging van de dansyl groep door fluorophoren met een hogere quantum yield. Hierdoor kan de resulterende SDS-PAGE gel direct gescand worden om fluorescente emissie van de gelabelde proteasoomsubunits te meten. Om effectieve proteasoom probes met goede biochemische en biofysische eigenschappen te kunnen selecteren, hebben we een serie probes met verschillende fluorescente groepen gesynthetiseerd. Twee verschillende probes bleken geschikt om proteasoomactiviteit te bestuderen in celextracten, intacte cellen en weefsels afkomstig van muizen. Bovendien kunnen deze probes ook gebruikt worden in confocale microscopy assays en flow cytometrie assays om proteasoomactiviteit direct in cellen te kunnen detecteren.

Hoofdstuk 2 concentreert zich op de toepassing van deze fluorescente proteasoom probes. We hebben de effecten van verschillende proteasoom remmers in preklinische studies en klinische trials bestudeerd. Daarnaast hebben we deze probes gebruikt om proteasoom te zuiveren en proteasoom activiteit en compositie te analyseren in verschillende weefsels.
In **hoofdstuk 2.1** vergelijken we de effecten van bortezomib en de nieuwe proteasoom remmer CEP-18770 in een aantal *in vitro* en preklinische *ex vivo* modellen. CEP-18770 doet hetzelfde als en lijkt zelfs op bortezomib.

In een panel van cellysaten en cellijnen waren beide proteasoomremmers equipotent in het remmen van verschillende proteasoomactiviteiten. In een preklinisch model van de ziekte van Kahler werd proteasoomactiviteit in normale, gezonde weefsels door beide remmers evenveel geremd. Echter, de remming van tumor proteasoomactiviteit door CEP-18770 was significant hoger dan de remming die bortezomib bewerkstelligde. Daarnaast laten we zien dat CEP-18770 behandeling in staat is om proteasoom te remmen in een bortezomib resistente multiple myeloom cellijn. Deze resistentie is verworven door langdurige blootstelling aan bortezomib *in vitro* en wordt veroorzaakt door een specifieke mutatie in een van de actieve subunits van het proteasoom. CEP-18770 kon bovendien apoptose veroorzaken in deze cellen.

Marizomib is een proteasoomremmer die momenteel geëvalueerd wordt in fase I klinische trials. Daardoor is het mogelijk om de respons van patiënten op behandeling met marizomib gedetailleerd te bestuderen. In **hoofdstuk 2.2** laten we zien dat fluorescente proteasoom probes gebruikt kunnen worden om de proteasoom remming te meten in zowel 'packed whole blood' (PWB, bloed zonder plasma dat voornamelijk uit rode bloedcellen bestaat) als in 'peripheral blood mononuclear cells' (PBMCs, witte bloedcellen), afkomstig van ratten die behandeld zijn met bortezomib of marizomib. Daarnaast hebben we een methode ontwikkeld waarmee fluorescente probes gebruikt kunnen worden om bloed samples van patiënten die behandeld worden met marizomib te bestuderen. We laten zien dat PWB, dat voor meer dan 98% uit rode bloed cellen bestaat, anders reageert op proteasoomremmers en minder snel herstelt van proteasoomremming in vergelijking met PBMCs. Deze data waarschuwen tegen het gebruik van PWB als een modelsysteem om de effecten van proteasoomremmers in proefdieren of patiënten te monitoren, zoals nu routinematig gedaan wordt. Daarnaast hebben we fluorescente proteasoom probes gebruikt om aan te tonen dat de initiële proteasoomactiviteit in patiënten voor de start van de behandeling een factor is die de gevoeligheid van patiënten voor proteasoomremmers bepaalt. Daarnaast bepaalt het vermogen van cellen om hun proteasoom activiteitsprofiel te veranderen in reactie op behandeling met marizomib de gevoeligheid van cellen voor marizomib. Proteasoomremmers die intraveneus worden toegediend verzadigen waarschijnlijk eerst het constitutieve proteasoom in rode bloedcellen voordat tumorcellen bereikt worden. Specifieke immunoproteasoomremmers verzadigen rode bloedcellen waarschijnlijk bij relatief lage concentraties en zouden daarom in de toekomst een goed alternatief kunnen bieden voor de behandeling van tumoren met een hoge intrinsieke immunoproteasoomactiviteit.

In **hoofdstuk 2.3** ten slotte beschrijf ik hoe fluorescente proteasoom probes gebruikt kunnen worden om tijdens het isoleren van proteasoom uit runderlever en –milt, de zuiverheid van dit proteasoom te kunnen monitoren. Met behulp van massaspectrometrie hebben we vervolgens laten zien dat zuiver constitutief of immunoproteasoom via deze methode verkregen kan worden. Daarnaast hebben we de relatieve subunitsamenstelling van gepurificeerd runderlever en -milt proteasoom met massaspectrometrie bepaald.

De vindingen beschreven in **hoofdstuk 2** laten zien dat chemische proteasoom probes gebruikt kunnen worden om proteasoomactiviteit op een betrouwbare manier te meten in een grote variëteit aan samples, zowel in preklinische diermodellen als in patiënten tijdens klinische trials. Daarnaast kunnen deze probes mogelijk de respons van patiënten op verschillende proteasoomremmers voor-

spellen, wat kan bijdragen aan therapie op maat om het therapeutische potentieel van dit type medicijn te verbeteren en betere behandelresultaten voor patiënten te bewerkstelligen.

In **hoofdstuk 3** wordt proteasoomactiviteit op moleculair niveau bestudeerd en wordt gekeken naar de rol van het proteasoom in het immuunsysteem. Recente studies hebben nieuwe epitopen geïdentificeerd die bestaan uit twee aparte delen van een eiwit die eerst losgeknipt zijn en vervolgens weer aan elkaar geplakt *gesplicet* om een nieuw peptide te vormen. De ligatiereactie die deze verknipte *gesplicete* epitopen produceert vindt plaats in en wordt gekatalyseerd door het proteasoom via een zogeheten transpeptidatiemechanisme. Dat mechanisme maakt het aannemelijk dat peptide *splicing* geen willekeurig proces is, maar gedicteerd wordt door specifieke regels. Het ophelderen van deze regels zal het voorspellen en identificeren van *gesplicete* antigenen vergemakkelijken. Pas als de identiteit van meer *gesplicete* epitopen bekend is, kan hun rol in immuniteit nauwkeurig bestudeerd worden.

In **hoofdstuk 3.2** wordt het moleculaire mechanisme van proteasomale antigen *splicing* bestudeerd, en worden regels opgesteld die deze antigen *splicing* kunnen voorspellen. Door kleine peptide bibliotheken met mogelijke *splicing* partners te incuberen met proteasoom, en deze reactiemengsels vervolgens te analyseren, hebben we laten zien dat splicing motieven bestaan. Specifieke aminozuursequenties bevorderden *splicing*, terwijl andere juist verhinderden dat *splicing* reacties plaatsvonden. Daarna hebben we deze *splicing* motieven bevestigd en verfijnd door de regels toe te passen op een polypeptide afgeleid van een viraal eiwit van het Influenza A/PR/8/34 virus (een H1N1 griepvirus, net als de mexicaanse griep). Daarnaast hebben we laten zien dat *splicing* reacties in het proteasoom heel efficiënt kunnen verlopen, met efficiënties tot 30% *in vitro*, als aan de structurele eisen voor proteasoom *splicing* voldaan wordt door beide *splicing* partners. Ook hebben we aangetoond dat een groot aantal *gesplicete* epitopen gevormd kan worden uit een enkel eiwit. Tot slot hebben we het mechanisme van *splicing* reacties bestudeerd en hebben laten zien dat splicing bij voorkeur gebeurt met twee peptiden die afkomstig zijn van hetzelfde eiwit. Willekeurige recombinatie van peptiden die afkomstig zijn van verschillende eiwitten vindt dus waarschijnlijk nauwelijks plaats *in vivo*.

In **hoofdstuk 3.3** laten we zien dat als een lysine residu betrokken is bij de ligatiereactie, proteasoom splicing een nieuw type antigen, dat een isopeptide binding bevat, kan creëren. Lysine heeft twee amino groepen (een α- en een ε-amino groep) die theoretisch allebei kunnen participeren in ligatiereacties. Dit impliceert dat het proteasoom naast normale peptide bindingen, waarin de α-amino groep betrokken is bij ligatie, ook antigenen kan vormen met een isopeptide binding (als de ε-amino groep participeert in de ligatiereactie). Met behulp van NMR hebben we laten zien dat zowel de α- als de ε-amino groep van lysine betrokken kan zijn bij ligatiereacties in het proteasoom. Ook konden we bepalen dat isopeptide vorming geprefereerd wordt over α-ligatie in 10% van alle *splicing* reacties waarin een lysine betrokken was op de plek van ligatie. Dit suggereert dat het proteasoom een nieuw type antigen vormt dat een rol zou kunnen spelen in immuniteit. Daarnaast hebben we laten zien dat epitopen die een isopeptide binding bevatten unieke eigenschappen hebben, die hen onderscheidt van normale epitopen. Zo kunnen isopeptiden met verschillende lengtes met hoge affiniteit binden aan MHC klasse I. Ook worden isopeptiden minder snel afgebroken door het

proteasoom in vergelijking met normale peptiden. De combinatie van deze eigenschappen suggereert dat een groot deel van de isopeptide epitopen die gevormd worden ook het celoppervlakte kan bereiken om daar een immuunrespons op te wekken.

Gezamenlijk geven de data die beschreven worden in **hoofdstuk 3.2** en **3.3**. aan dat verschillende typen ligatiereacties in het proteasoom kunnen plaatsvinden *in vitro*. De efficiënties van deze reacties zijn hoog (tot 30%) vergeleken bij de geschatte efficiënties voor de formatie van de *gesplicete* epitopen die tot nu toe beschreven zijn (0.0002-0.01%). Dit suggereert dat proteasomale *splicing* reacties veel vaker plaatsvinden dan tot nu toe werd aangenomen en dus een significante bijdrage zouden kunnen leveren aan immuniteit *in vivo*. Tot nu toe zijn pas drie *gesplicete* epitopen beschreven *in vivo*. Dit is niet verwonderlijk als je bedenkt dat identificatie van deze epitopen niet mogelijk is met de huidige methodes die gebruikt worden om nieuwe antigenen te identificeren. Nieuwe epitopen kunnen worden geïdentificeerd door elutie van het celoppervlakte, gevolgd door analyse met massaspectrometrie, waarbij pepitdes gematchet worden tegen databases met eiwitsequenties. Een alternatieve methode is het voorspellen van epitopen uitgaande van sequenties van bijvoorbeeld virale eiwitten. Vervolgens wordt dan getest of reactieve CD8+ T cellen tegen deze epitopen gevonden kunnen worden in muizen die gevaccineerd zijn met het desbetreffende virale eiwit. In het geval van *gesplicete* epitopen echter, bestaan de gevormde antigenen niet uit een aaneengesloten stuk eiwit, maar is een post-translationele *splicing* reactie noodzakelijk om deze epitopen te vormen. Daardoor kunnen deze peptiden niet gematchet worden tegen databases die alleen aaneengesloten eiwitsequenties bevatten. Bovendien kunnen ze niet voorspeld worden op basis van aaneengesloten sequenties alleen. De *splicing* regels die beschreven staan in hoofdstuk 3.2 maken het voorspellen van *gesplicete* epitopen mogelijk. Dit zou de identificatie van meer *gesplicete* epitopen moeten vergemakkelijken. Ik verwacht dat de studies die beschreven staan in dit proefschrift zullen bijdragen aan de ontdekking van nieuwe *gesplicete* epitopen en aan het ontrafelen van mogelijke functies van *gesplicete* epitopen en isopeptide antigenen *in vivo*.

Curriculum Vitae

Celia Rosita Berkers werd geboren op 1 mei 1980 te Amsterdam. In 1998 behaalde zij haar VWO diploma aan het Vossius Gymnasium in Amsterdam. In datzelfde jaar begon zij aan de studie scheikunde aan de Universiteit Utrecht. Tijdens haar studie liep zij stage bij de vakgroep biochemie van membranen aan de Universiteit Utrecht, waar zij de interactie tussen lipid II en het lantibioticum nisine bestudeerde. Een tweede stage liep zij bij de vakgroep Bio-organische chemie aan de Universiteit Utrecht. Hier deed zij onderzoek naar de rol van lipoxygenase in neurodegeneratie en neuroprotectie. In november 2003 behaalde zij *cum laude* het doctoraal examen scheikunde. In januari 2004 startte zij haar promotieonderzoek in de groep van Prof. dr. Hidde Ploegh aan de Harvard Medical School te Boston onder leiding van Dr. Huib Ovaa. Voor dit onderzoek ontving zij onderzoeksbeurzen van de Koninklijke Hollandse Maatschappij der Wetenschappen en de Stichting Fonds Doctor Catharine van Tussenbroek. Vanaf juli 2004 zette zij dit promotieonderzoek voort in de groep van Dr. Huib Ovaa aan het Nederlands Kanker Instituut te Amsterdam. Zij presenteerde de resultaten van haar promotieonderzoek op verschillende nationale en internationale congressen. Voor haar lezing op het jaarlijkse congres van de NWO-CW studiegroep eiwitonderzoek won zij in 2006 de *PhD-Student Lecture Prize*. Ook won zij in 2006 de *Best Lecture Prize Fundamental Science* voor haar lezing tijdens de jaarlijkse PhD retraite van de Oncology Graduate School Amsterdam.

List of publications

C.R. Berkers, K.G. Schuurman, Y. Leestemaker, B. Ruggeri, S. Jones-Bolin, M. Williams & H. Ovaa. Comparison of the specificity and activity profiles of the proteasome inhibitors bortezomib and CEP-18770. *Manuscript in preparation.*

C.R. Berkers, K.G. Schuurman, S.K.C. Neuteboom, D. Chauhan, K.C. Anderson, M.A. Palladino & H. Ovaa. Profiling patient response to marizomib therapy. *Manuscript in preparation.*

C.R. Berkers, A. de Jong, K.G. Schuurman, J.A.J. Geenevasen, B. Rodenko & H. Ovaa. Proteasomal splicing creates a novel type of antigen containing an isopeptide linkage. *Manuscript in preparation.*

C.R. Berkers, A. de Jong, K.G. Schuurman, C. Linnemann, H.D. Meiring, T.N. Schumacher, B. Rodenko & H. Ovaa. How the proteasome cleaves and religates peptides to expand the antigen repertoire. *Manuscript in preparation.*

A. de Jong, K.G. Schuurman, B. Rodenko, H. Ovaa & **C.R. Berkers**. Fluorescent-based proteasome activity profiling. *Methods in Molecular Biology. Accepted for publication.*

S. Pelletier, K.G. Schuurman, **C.R. Berkers**, H. Ovaa, A.J. Heck & R. Raijmakers. Quantifying cross-tissue diversity in proteasome complexes by mass spectrometry. *Molecular BioSystems. 2010, **DOI:** 10.1039/c004989a.*

C.R. Berkers & H. Ovaa. Drug discovery and assay development in the ubiquitin-proteasome system. *Biochemical Society Transactions. 2010, 38 (Part 1): 14-20.*

J. Middeldorp, W. Kamphuis, J.A. Sluijs, D. Achoui, C.H. Leenaars, M.G. Feenstra, P. van Tijn, D.F. Fischer, **C.R. Berkers**, H. Ovaa, R.A. Quinlan & E.M. Hol. Intermediate filament transcription in astrocytes is repressed by proteasome inhibition. *FASEB Journal. 2009, 23 (8): 2710-2726.*

C.R. Berkers, A. de Jong, H. Ovaa & B. Rodenko. Transpeptidation and reverse proteolysis and their consequences for immunity. *International Journal of Biochemistry and Cell Biology. 2009, 41 (1): 66-71.*

continued on the next page...

R. Oerlemans, N.E. Franke, Y.G. Assaraf, J. Cloos, I. van Zantwijk, **C.R. Berkers**, G.L. Scheffer, K. Debipersad, K. Vojtekova, C. Lemos, J.W. van der Heijden, B. Ylstra, G.J. Peters, G.L. Kaspers, B.A. Dijkmans, R.J. Scheper & G. Jansen. Molecular basis of bortezomib resistance: proteasome subunit beta5 (PSMB5) gene mutation and overexpression of PSMB5 protein. *Blood. 2008, 112 (6): 2489-2499*

R. Raijmakers, **C.R. Berkers**, A. de Jong, H. Ovaa, A.J. Heck & S. Mohammed. Automated online sequential isotope labeling for protein quantitation applied to proteasome tissue-specific diversity. *Molecular & Cellular proteomics. 2008, 7 (9): 1755-1762.*

A.H. Bakker, R. Hoppes, C. Linnemann, M. Toebes, B. Rodenko, **C.R. Berkers**, S.R. Hadrup, W.J. van Esch, M.H. Heemskerk, H. Ovaa & T.N. Schumacher. Conditional MHC class I ligands and peptide exchange technology for the human MHC gene products HLA-A1, -A3, -A11, and -B7. *Proceedings of the National Acadamy of Sciences of the United States of America. 2008, 105 (10): 3825-3830.*

R. Piva, B. Ruggeri, M. Williams, G. Costa, I. Tamagno, D. Ferrero, V. Giai, M. Coscia, S. Peola, M. Massaia, G. Pezzoni, C. Allievi, N. Pescalli, M. Cassin, S. di Giovine, P. Nicoli, P. de Feudis, I. Strepponi, I. Roato, R. Ferracini, B. Bussolati, G. Camussi, S. Jones-Bolin, K. Hunter, H. Zhao, A. Neri, A. Palumbo, **C.R. Berkers**, H. Ovaa, A. Bernareggi & G. Inghirami. CEP-18770: A novel, orally active proteasome inhibitor with a tumor-selective pharmacologic profile competitive with bortezomib. *Blood. 2008, 111 (5): 2765-2775.*

C.R. Berkers, F.W. van Leeuwen, T.A. Groothuis, V. Peperzak, E.W. van Tilburg, J. Borst, J.J. Neefjes & H. Ovaa. Profiling proteasome activity in tissue with fluorescent probes. *Molecular Pharmaceutics. 2007, 4 (5): 739-748.*

M. Kraus, T. Rückrich, M. Reich, J. Gogel, A. Beck, W. Kammer, **C.R. Berkers**, D. Burg, H. Overkleeft, H. Ovaa & C. Driessen. Activity patterns of proteasome subunits reflect bortezomib sensitivity of hematologic malignancies and are variable in primary human leukemia cells. *Leukemia. 2007, 21 (1): 84-92.*

M. Verdoes, B.I. Florea, V. Menendez-Benito, C.J. Maynard, M.D. Witte, W.A. van der Linden, A.M. van den Nieuwendijk, T. Hofmann, **C.R. Berkers**, F.W. van Leeuwen, T.A. Groothuis, M.A. Leeuwenburgh, H. Ovaa, J.J. Neefjes, D.V. Filippov, G.A. van der Marel, N.P. Dantuma & H.S. Overkleeft. A fluorescent broad-spectrum proteasome inhibitor for labeling proteasomes in vitro and in vivo. *Chemistry & Biology. 2006, 13 (11): 1217-1226.*

MDP Risseeuw, **C.R. Berkers**, H.L. Ploegh & H. Ovaa. Synthesis of tritium labeled KRN7000. *Tetrahedron Letters. 2006, 47 (22): 3677-3679.*

M. Groll, **C.R. Berkers**, H.L. Ploegh & H. Ovaa. Crystal structure of the boronic acid-based proteasome inhibitor bortezomib in complex with the yeast 20S proteasome. *Structure. 2006, 14 (3): 451-456.*

M. Verdoes, **C.R. Berkers**, B.I. Florea, P.F. van Swieten, H.S. Overkleeft & H. Ovaa. Chemical proteomics profiling of proteasome activity. *Methods in Molecular Biology. 2006, 328: 51-69.*

D. Chauhan, L. Catley, G. Li, K. Podar, T. Hideshima, M. Velankar, C. Mitsiades, N. Mitsiades, H. Yasui, A. Letai, H. Ovaa, **C.R. Berkers**, B. Nicholson, T.H. Chao, S.T. Neuteboom, P. Richardson, M.A. Palladino & K.C. Anderson. A novel orally active proteasome inhibitor induces apoptosis in multiple myeloma cells with mechanisms distinct from Bortezomib. *Cancer Cell. 2005, 8 (5): 407-419*

C.R. Berkers, M. Verdoes, E. Lichtman, E. Fiebiger, B.M. Kessler, K.C. Anderson, H.L. Ploegh, H. Ovaa & P.J. Galardy. Activity probe for in vivo profiling of the specificity of proteasome inhibitor bortezomib. *Nature Methods. 2005, 2 (5): 357-362*

C.R. Berkers & H. Ovaa. Immunotherapeutic potential for ceramide-based activators of iNKT cells. *Trends in Pharmacological Sciences. 2005, 26 (5): 252-257*