BJ www.biochemj.org

Biochem. J. (2010) **430**, 461–476 (Printed in Great Britain) doi:10.1042/BJ20100383

BJ ChemBio

online data

Characterization of a new series of non-covalent proteasome inhibitors with exquisite potency and selectivity for the 20S β5-subunit

Christopher BLACKBURN, Kenneth M. GIGSTAD, Paul HALES, Khristofer GARCIA, Matthew JONES, Frank J. BRUZZESE, Cynthia BARRETT, Jane X. LIU, Teresa A. SOUCY, Darshan S. SAPPAL, Nancy BUMP, Edward J. OLHAVA[1], Paul FLEMING, Lawrence R. DICK, Christopher TSU, Michael D. SINTCHAK and Jonathan L. BLANK[2]

Discovery, Millennium Pharmaceuticals, Inc., 40 Landsdowne Street, Cambridge, MA 02139, U.S.A.

The mammalian 26S proteasome is a 2500 kDa multi-catalytic complex involved in intracellular protein degradation. We describe the synthesis and properties of a novel series of non-covalent di-peptide inhibitors of the proteasome used on a capped tri-peptide that was first identified by high-throughput screening of a library of approx. 350000 compounds for inhibitors of the ubiquitin–proteasome system in cells. We show that these compounds are entirely selective for the β5 (chymotrypsin-like) site over the β1 (caspase-like) and β2 (trypsin-like) sites of the 20S core particle of the proteasome, and over a panel of less closely related proteases. Compound optimization, guided by X-ray crystallography of the liganded 20S core particle, confirmed their non-covalent binding mode and provided a structural basis for their enhanced *in vitro* and cellular potencies. We demonstrate that such compounds show low nanomolar IC_{50} values for the human 20S β5 site *in vitro*, and that pharmacological inhibition of this site in cells is sufficient to potently inhibit the degradation of a tetra-ubiquitin–luciferase reporter, activation of NFκB (nuclear factor κB) in response to TNF-α (tumour necrosis factor-α) and the proliferation of cancer cells. Finally, we identified capped di-peptides that show differential selectivity for the β5 site of the constitutively expressed proteasome and immunoproteasome *in vitro* and in B-cell lymphomas. Collectively, these studies describe the synthesis, activity and binding mode of a new series of non-covalent proteasome inhibitors with unprecedented potency and selectivity for the β5 site, and which can discriminate between the constitutive proteasome and immunoproteasome *in vitro* and in cells.

Key words: chymotrypsin-like, immunoproteasome, 26S proteasome, proteasome inhibitor, β5-subunit, ubiquitin–proteasome system (UPS).

INTRODUCTION

The 26S proteasome is a very large (2500 kDa) multi-subunit proteolytic complex found in high abundance in all eukaryotic cells that is responsible for the regulated degradation of the majority of intracellular proteins, including those involved in signal transduction, cell-cycle control, apoptosis and inflammatory responses [1–6]. It is composed of the 20S catalytic core particle that is capped at one or both ends by the 19S regulatory complex, also called PA700 [1–6]. The 20S core particle is comprises four stacked heteroheptameric rings that form a cylindrical structure with 2-fold axial symmetry. The two outer α rings gate substrate entry through their interaction with the 19S regulatory complex, which itself serves to bind, unfold, de-ubiquitinate and translocate into the catalytic core substrate proteins that have been conjugated to poly-ubiquitin [1–6]. The two inner β rings form the central proteolytic chamber and each contain three active sites with distinct substrate specificities. These are located on the β1, β2 and β5 subunits and are referred to as having caspase-like, trypsin-like and chymotrypsin-like activities on the basis of their preference for cleaving peptides after acidic, basic or hydrophobic amino acid residues respectively [1–6]. A second form of the 20S proteasome termed the immunoproteasome exists in cells of lymphoid origin and can be induced in most, if not all, cells in response to the pro-inflammatory cytokine interferon-γ [5–7]. This enzyme contains distinct catalytic subunits, designated β1i, β2i and β5i, and is primarily involved in the generation of antigenic peptides for MHC class I presentation, although a function for the immunoproteasome in cytokine production has also been described [8].

For each of the 20S catalytic β-subunits, the hydroxy group of the N-terminal threonine residue ($Thr^{1O\gamma}$) serves as the nucleophile that initiates cleavage of the peptide bond [4,5,9]. Various natural and synthetic compounds that contain electrophilic centres or 'warheads' inhibit the proteasome by forming covalent adducts with these active-site threonine residues, including peptide aldehydes, vinyl sulfones, boronic acids, α′β′-epoxyketones, 2-keto-1,3,4-oxadiazoles and β-lactones (Figure 1A) [4–6,9–11]. The di-peptide boronic acid bortezomib (PS-341 or VELCADE®, Millennium Pharmaceuticals, Inc.) is among the most potent, stable and selective of these inhibitors [12–15], and shows nanomolar potency with respect to cytotoxicity across a broad range of human tumour cell types *in vitro* [13,14]. It is in clinical use for the treatment of multiple myeloma [16–19] and refractory mantle cell lymphoma [20],

Abbreviations used: Ac, acetyl; AMC, 7-amino-4-methylcoumarin; Boc, t-butoxycarbonyl; HBTU, *O*-benzotriazole-*N*,*N*,*N*′,*N*′-tetramethyluronium hexafluorophosphate; HEK, human embryonic kidney; LC_{50}, half-maximal lethal concentration; MPD, 2-methyl-2,4-pentanediol; NFκB, nuclear factor κB; IκB, inhibitory protein of NFκB; NFκB-Luc, NFκB–luciferase; PA, proteasomal activator; PDL, poly-D-lysine; RNAi, RNA interference; siRNA, small interfering RNA; Suc, succinyl; TEV, tobacco etch virus; TNF-α, tumour necrosis factor-α; 4xUb-Luc, tetra-ubiquitin–luciferase; UPS, ubiquitin–proteasome system; Z, benzyloxycarbonyl.

[1] Present address: Epizyme, Inc., 840 Memorial Drive, Cambridge, MA 02139, U.S.A.

[2] To whom correspondence should be addressed (e-mail jonathan.blank@mpi.com).

The structural co-ordinates of the yeast 20S proteasome with the indicated ligand bound reported will appear in the PDB under accession codes: 3MG0 (bortezomib); 3MG4 (compound 1); 3MG6 (compound 6); 3MG7 (compound 8); 3MG8 (compound 16).

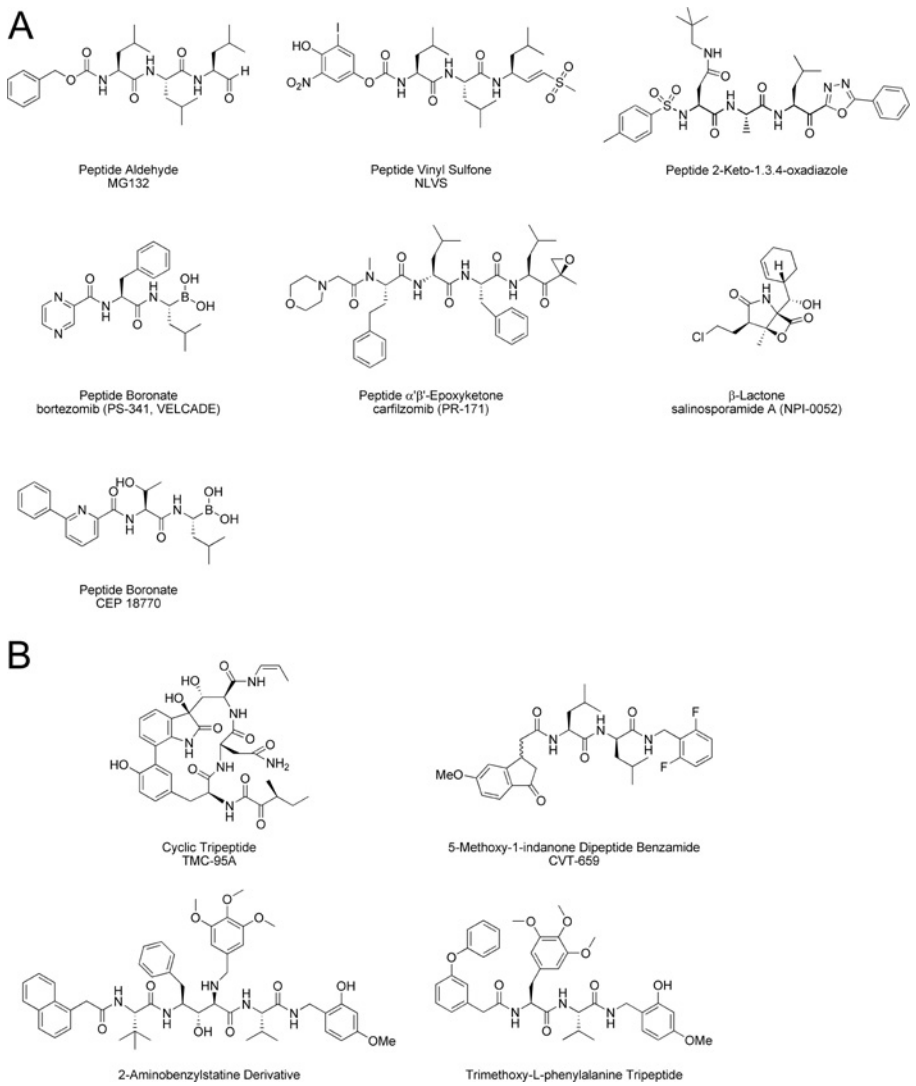


Figure 1 Examples of covalent (A) and non-covalent (B) proteasome inhibitors

and is being evaluated for the treatment of other malignancies [21–23]. Bortezomib induces cell death through a variety of transcriptional, translational and post-translational mechanisms, and may be preferentially cytotoxic to cancer cells by enhancing endoplasmic reticulum stress, increasing the expression of pro-apoptotic factors and/or inhibiting pro-survival or DNA-damage repair pathways [4–6,21–23]. More recently, two further closely related di-peptide boronic acids, CEP-18870 and MLN9708, have been described that inhibit cancer cell proliferation *in vitro* and show anti-tumour activity in solid and haematological preclinical tumour models [24,25].

Bortezomib binds with very high affinity to the $\beta 5$ site of the proteasome, and to a lesser extent the $\beta 1$ and $\beta 2$ sites [15], and behaves as a slowly reversible inhibitor ($t_{1/2} \sim 110$ min for dissociation from the $\beta 5$ site [25]). Analysis of the crystal structure of the yeast 20S in complex with bortezomib indicates that the high affinity of the drug is partly mediated by the covalent interaction of the boron atom with the nucleophilic oxygen of $Thr^{1O\gamma}$, and by the hydrogen bonding between the two acidic boronate hydroxy groups and the amine groups of Gly^{47N} and Thr^{1N} within each active site [5,9,26]. These properties are likely to contribute to the effectiveness of bortezomib as an anti-cancer agent for the treatment of haematological malignancies, but may limit its distribution and therefore broader utility in solid tumours due to the high tissue abundance of the target (the proteasome has been estimated to comprise 2 % of total cellular protein [6]). More recently, the β-lactone salinosporamide A (NPI-0052) [27] and the epoxyketone carfilzomib (PR-171) (Figure 1A) [28], which are analogues of lactacystin and epoxomicin respectively, have demonstrated activity in preclinical models of multiple myeloma and chronic lymphocytic leukaemia, and have entered clinical trials for advanced solid and haematological malignancies [23]. These covalent inhibitors have comparable potency to bortezomib, but differ in that they form essentially irreversible adducts with the active site $Thr^{1O\gamma}$ residues of the proteasome [4,5,9,10,15,27–30].

Although the reactive 'warheads' of covalent proteasome inhibitors can contribute to enzyme potency, they can also lack specificity and be excessively reactive and unstable [10,11,15], properties which may limit their efficacy *in vivo*. Non-covalent inhibitors that are readily reversible in their interactions with the catalytic 20S β-subunits provide an alternative mechanism for proteasome inhibition. Such inhibitors may offer therapeutic

advantages by enabling more widespread tissue distribution through their more rapid binding and dissociation kinetics. TMC-95A is a complex natural cyclic tri-peptide (Figure 1B) that competitively and non-covalently inhibits the chymotrypsin-like activity of the proteasome with nanomolar potency, and also inhibits the caspase-like and trypsin-like activities at higher concentrations [4,5,9,10,31]. The high affinity of TMC-95A is due to its rigid heterocyclic ring system, which provides specific hydrogen-bonding interactions within each 20S active site without covalent modification of the catalytic threonine residues [4,5,9,32]. In addition, synthetic analogues of TMC-95A have been described, including endocyclic biphenyl-ether macrocyles and non-constrained linear tri-peptides that bind non-covalently with altered affinities for the trypsin- and caspase-like sites relative to the chymotrypsin-like site [5,9,33,34]. TMC-95A shows cytotoxicity in the low micromolar range in cancer cells, although the mechanism-based effects of TMC-95A or its derivatives have not been demonstrated [4]. Other non-covalent proteasome inhibitors include a series of 5-methoxy-1-indanone di-peptide benzamides (e.g. CVT-659) [35,36] and 2-aminobenzylstatine derivatives (Figure 1B) [37] that selectively inhibit the chymotrypsin-like site with submicromolar potencies, although these compounds also have poor cellular activity. However, homology modelling of a 2-aminobenzylstatine compound in the chymotrypsin-like binding pocket of the proteasome has led to the design a non-covalent tri-methoxy-L-phenylalanine-containing tri-peptide with maintained selectivity and greatly increased cell potency (Figure 1B) [38]. This compound is the most potent non-covalent proteasome inhibitor described to date, inhibiting the chymotrypsin-like activity in MDA-MB-435 cells and their proliferation with IC_{50} values of 20 nM and 60 nM respectively [38].

In an effort to identify cell-active inhibitors of the UPS (ubiquitin–proteasome system) with novel chemical scaffolds, we have screened a library of approx. 350 000 compounds using a cell-based assay that monitors accumulation of a 4xUb-Luc (tetra-ubiquitin–luciferase) reporter in response to proteasome inhibition [15,39]. Using this approach, we report the identification of a C- and N-terminally capped tri-peptide derived from the unnatural amino acid *S*-homo-phenylalanine that potently and selectively inhibits the chymotrypsin-like activity of the mammalian and yeast 20S proteasomes. Further optimization, guided by X-ray crystallography of compounds in complex with purified yeast 20S, has yielded a series of non-covalent di-peptide inhibitors of the proteasome with unprecedented *in vitro* and cellular potencies. The synthesis, binding mode and cellular activity of these compounds are described in the present study.

EXPERIMENTAL

Cell culture

Cells were from the A.T.C.C. (Manassas, VA, U.S.A.), with the exception of the diffuse large B-cell lymphoma lines which were obtained from the following sources: Karpas-1106P, Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSMZ; Braunschweig, Germany); WSU-DLCL2, Asterand (Detroit, MI, U.S.A.); and OCI-Ly10, provided by Dr Louis M. Staudt (National Cancer Institute, National Institutes of Health, Bethesda, MD, U.S.A.). Cells were cultured at 37 °C in a humidified air/6 % CO_2 atmosphere in medium supplemented with 10 % fetal bovine serum, except for the medium for Karpas-1106P and OCI-Ly10 cells which contained 20 % fetal bovine serum, and 100 units/ml penicillin/100 μg/ml streptomycin (all from Invitrogen), as specified: Calu6 cells, minimum essential medium; H460, WSU-DLCL2 and Karpas-1106P cells, RPMI 1640 medium; HCT116 and HT29 cells, McCoy's 5a medium; and OCI-Ly-10, Iscove's modified Dulbecco's medium. Clonally-derived stable MDA-MB-231 cells expressing four tandem copies of ubiquitin fused to firefly luciferase (4xUb-Luc) and HEK (human embryonic kidney)-293 cells expressing NFκB-Luc [NFκB (nuclear factor κB)–luciferase] were generated and maintained as described previously [15].

Reporter assays

Cells were seeded at 10 000 cells per well in white BioCoat™ PDL (poly-D-lysine)-coated 384-well plates (BD Biosciences) at 16–24 h prior to compound treatment. For the 4xUb-Luc assays, MDA-MB-231 cells were incubated with compound for 8 h. For NFκB-Luc assays, HEK-293 cells were pre-treated for 1 h with proteasome inhibitor and then stimulated with 10 ng/ml recombinant human TNF-α (tumour necrosis factor-α) (R&D Systems) for a further 3 h in the continued presence of the compound. Firefly luciferase activity was measured using Bright-Glo™ reagents according to the manufacturer's instructions (Promega) in a LEADseeker™ plate reader (GE Healthcare Life Sciences). Inhibition of NFκB-Luc activity was calculated relative to a no-compound (DMSO) control, whereas 4xUb-Luc reporter accumulation was expressed as a fold increase in luciferase activity over the DMSO control.

Cell viability assay

Calu6, HT29, MDA-MB-231 cells (each at 2000 cells/well), H460 cells (1000 cells/well) and HCT116 cells (1500 cells/well) were plated in black clear-bottomed BioCoat™ PDL-coated 384-well plates (BD Biosciences). Cells were incubated with compound for 72 h, after which the medium was removed to leave 25 μl per well. An equal volume of ATPlite™ reagent (PerkinElmer) was then added and luminescence was measured using a LEADseeker™ instrument.

siRNA (small interfering RNA) transfection and assay

MDA-MB-231 4xUb-Luc cells were transfected in a 384-well format with 10 nM siRNAs (siGENOME SMARTpool, Dharmacon) using DharmaFECT 1 (DH1) reagent (Dharmacon) as follows. For the preparation of the RNAi (RNA interference) transfection mixture for each time point, 40 μl of OptiMEM I (Invitrogen) was dispensed into duplicate wells of a 384-well plate each containing 9 μl/well of 0.5 μM siRNA in siRNA buffer (Dharmacon). OptiMEM (50 μl) containing 0.53 μl of DH2 transfection reagent (Dharmacon) was then added to each well and the plates were incubated at room temperature (22 °C) for 20 min. The transfection mixture (14 μl) from duplicate wells was then transferred into six replicate wells of two separate BioCoat™ PDL-coated 384-well cell plates (BD Biosciences), one for 4xUb-Luc assay and one for ATPlite assay (white and black clear-bottomed respectively). Reverse transfection was performed by the addition of 50 μl of MDA-MB-231 4xUb-Luc cells to each well to give 1200 cells/well. At 48, 72 or 96 h after transfection, one set of duplicate plates was assayed for 4xUb-Luc reporter activity and cell viability. For each time point, six replicate wells of cells transfected with buffer only and the non-targeting GL2 luciferase control duplex were each treated with either DMSO or 1 μM bortezomib for 8 h prior to performing the 4xUb-Luc assay.

Cell-based Proteasome-Glo™ β5 assays

Calu6 cells were plated as for the reporter assays and incubated with compound for 1 h at 37 °C. B-cell lymphoma subtypes Karpas-1106P, WSU-DLCL2 and OCI-Ly10 were plated at 20 000 cells per well in 384-well plates and treated with compounds identically. The activity of the 26S proteasome was measured *in situ* after compound removal by monitoring hydrolysis of the β5 (chymotrypsin-like) substrate Suc-LLVY-aminoluciferin (where Suc is succinyl) in the presence of luciferase using the Proteasome-Glo™ assay reagents according to the manufacturer's instructions (Promega). Luminescence was measured using a LEADseeker™ instrument.

In vitro assays of purified 20S

The peptidase activities of purified human erythrocyte and peripheral blood monocyte 20S proteasomes (Boston Biochem™) were assayed using fluorogenic tri-and tetra-peptide substrates coupled to AMC (7-amino-4-methylcoumarin) [40] (AnaSpec) in the presence of recombinant PA28α (proteasomal activator 28α; Boston Biochem™). The following selective substrates were obtained from AnaSpec, with the exception of Z-LLE-AMC (where Z is benzyloxycarbonyl), which was from Boston Biochem™, and were each used at a final concentration of 15 μM to assay the constitutive (c) and immunoproteasome (i) sub-sites: β1c, Z-LLE-AMC; β2c, Ac-KQL-AMC (where Ac is acetyl); β5c, Ac-WLA-AMC; β1i, Ac-PAL-AMC; β2i, Ac-KQL-AMC; and β5i, Ac-ANW-AMC. Reactions were performed at 37 °C in 384-well black microtitre plates (Corning) using 0.25 nM 20S and 12 nM PA28α in a final volume of 50 μl of buffer containing 20 mM Hepes, pH 7.4, 0.5 mM EDTA and 0.01 % BSA. Peptidase activity was measured by monitoring the AMC liberation over time with a Polarstar Galaxy fluorimeter (BMG Labtechnologies) using excitation and emission wavelengths of 340 nm and 460 nm respectively. Percentage inhibition of 20S activity was calculated relative to the controls with DMSO and 10 μM ML599698, an analogue of bortezomib that contains a phenyl group in place of the N-terminal pyrazine cap.

In vitro assays of 20S proteasomes in cell extracts

B-cell lymphoma extracts were prepared by hypotonic lysis and assayed for proteasome activity as described previously [15] using 10 μg of protein extract per reaction in the presence of 12 nM PA28α and the following substrates, each at 100 μM, to monitor the catalytic activity of the c or i 20S sub-sites, as indicated: β1i, Ac-PAL-AMC; β5i, Ac-ANW-AMC; β5c, Ac-WLA-AMC; and β5 Suc-LLVY-AMC.

Kinetic analysis of purified 20S inhibition

Inactivation of purified human erythrocyte 20S was determined by monitoring the hydrolysis of the β5-specific fluorigenic peptide substrate Suc-LLVY-AMC in the presence of different concentrations of inhibitor. Substrate cleavage was continuously monitored as a change in the fluorescence emission at 460 nm (excitation wavelength, 360 nm) and plotted as a function of time. Assays were performed at 37 °C in cuvettes in a final volume of 2 ml in a reaction buffer containing 20 mM Hepes, pH 7.5, 0.5 mM EDTA, 0.01 % SDS, 0.25 nM 20S and 10 μM Suc-LLVY-AMC, with continuous stirring. The apparent dissociation constant, K_i^{app}, was determined by non-linear least-fit of the fractional velocity, v_s/v_0, as a function of [I], where v_s is the steady-state residual activity of the enzyme in the presence of inhibitor (I) and v_0 is the initial velocity in the absence of inhibitor:

$$\frac{v_s}{v_0} = \frac{1}{1 + \frac{[I]}{K_i^{app}}}$$

The dissociation constant, K_i, was calculated from the apparent dissociation constant, K_i^{app}, using the following expression, where [S] is the substrate concentration and K_m is the substrate binding constant (for 20 μM Suc-LLVY-AMC):

$$K_i = \frac{K_i^{app}}{1 + \frac{[S]}{K_m}}$$

Protein turnover assay

Bulk protein turnover was measured in HCT116 cells according to previously published procedures [41], as described in the Supplementary material at http://www.BiochemJ.org/bj/430/bj4300461add.htm.

Synthesis of compound 1, the capped tri-peptide 5-methyl-*N*-((*S*)-1-((*S*)-1-((*S*)-1-(4-methylbenzylamino)-1-oxo-4-phenylbutan-2-ylamino)-1-oxo-4-phenylbutan-2-ylamino)-1-oxo-4-phenylbutan-2-yl)pyrazine-2-carboxamide

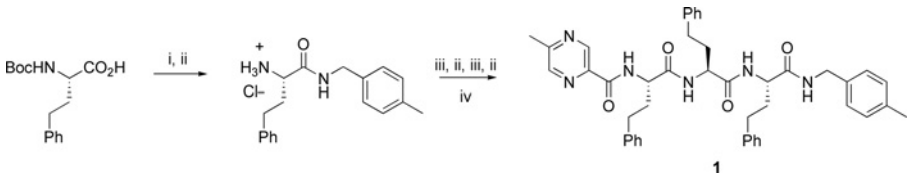
The trimeric compound 1 was identified by LC-MS as an impurity in an active well identified during the high-throughput screen [42] and its structure was confirmed by full characterization of material prepared according to Scheme 1, as detailed in the Supplementary Experimental section at http://www.BiochemJ.org/bj/430/bj4300461add.htm.

Synthesis of capped di-peptide proteasome inhibitors

The synthesis of di-peptide analogues was carried out using standard peptide-coupling conditions, typically using HBTU (*O*-benzotriazole-*N,N,N′,N′*-tetramethyluronium hexafluorophosphate) to activate the carboxylic acids according to Scheme 2. The P1 and P2 residues (R_1 and R_2 respectively) were installed as described above for the preparation of compound 1 by coupling the appropriate amine to a Boc (t-butoxycarbonyl)-protected α-amino acid, followed by de-protection to give intermediates II as hydrochloride salts. The three-step elaboration of II to the di-peptide analogues of compound 1 (IV) was carried out using automated solution-phase parallel synthesis techniques in deep-well plates without purification of intermediates, but with rigorous purification and full characterization of final products. A representative procedure for the parallel synthesis of compound arrays of IV can be found in the Supplementary Experimental section.

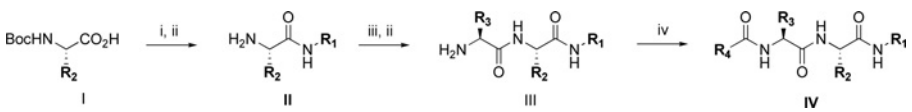
Purification of yeast 20S

Yeast 20S wild-type and open-gate (an N-terminal tail deletion from the α3 and α7 subunits, designated α3/7ΔN) mutant containing a TEV (tobacco etch virus)-protease-cleavable Protein A-derived tag in the core particle subunit Pre1 (β4) were affinity-purified from the SDL135 and yMS159 strains respectively (gifts from Dr M. Schmidt and Dr D. Finley, Harvard University Medical School, Boston, MA, U.S.A.) [43,44] on IgG–Sepharose 6 Fast Flow (GE Healthcare), according to the procedure of Leggett et al. [45], except that ATP was omitted from the buffers and the enzyme was eluted from the IgG resin by incubation with the His_6–TEV protease (Invitrogen) overnight at 4 °C. The resultant



Scheme 1 Synthesis of compound 1

(i) $ArCH_2$-NH_2, HBTU, NMM (*N*-methylmorpholine), DMF (*N*,*N*-dimethylformamide), 25 °C, 24 h; (ii) 4 M HCl, dioxane; (iii) *N*-Boc-homo-phenylalanine, TBTU [*O*-(benzotriazol-1-yl)-*N*,*N*,*N*′,*N*′-tetramethyluronium-tetrafluoroborate], collidine, DMF, 25 °C, 3 h; (iv) 4-methylpyrazinecarboxylic acid, TBTU, collidine, DMF, 25 °C, 3 h.



Scheme 2 Synthesis of compound IV

(i) R_1-NH_2, HBTU, NMM (*N*-methylmorpholine), DMF (*N*,*N*-dimethylformamide), 25 °C, 24 h; (ii) 4 M HCl, dioxane; (iii) P3 Boc-amino acid, HBTU, NMM, DMF, 25 °C, 24 h; (iv) R_4COOH, HBTU, NMM, DMF, 25 °C, 24 h.

yeast 20S was purified further by gel filtration on a 24 ml Superose 6 10/30 column (GE Healthcare) in 10 mM Tris/HCl, pH 7.5, containing 1 mM EDTA, concentrated to 30 mg/ml using an Amicon Centriplus YM-100 centrifugal filter device (Millipore) and stored at −80 °C.

X-ray crystallography

Crystals of purified wild-type and the open-gate mutant yeast 20S proteasome were grown in hanging drops at room temperature, as described previously [46], using a drop volume of 1.5 μl of 20S (20 mg/ml in 10 mM Tris/HCl, pH 7.5, and 1 mM EDTA) and 0.5 μl of reservoir solution containing 100 mM Mes, pH 7.0, 40 mM $Mg(CH_3COO)_2$, 15 % MPD (2-methyl-2,4-pentanediol) and 10 mM EDTA. Proteasome–inhibitor complexes were generated by soaking crystals overnight in reservoir buffer containing 1 mM compound, 10 % DMSO and 20 % MPD followed by an additional 5 h in reservoir buffer containing 1 mM compound, 10 % DMSO and 25 % MPD before being flash-cooled in liquid nitrogen. Crystal data and refinement statistics are given in Supplementary Table S1 at http://www.BiochemJ.org/bj/430/bj4300461add.htm. Data were collected using the Structural GenomiX (SGX)-CAT beamline at the Advanced Photon Source (APS) synchrotron of the Argonne National Laboratory (U.S. Department of Energy Chicago, IL, U.S.A.) and processed using the programs iMOSFLM [47] and SCALA [48]. Starting co-ordinates for each of the proteasome–inhibitor structures were taken from PDB entry 1G0U [43]. SigmaA-weighted $F_o - F_c$ difference electron density for the β5 active site was used to model inhibitor co-ordinates, starting with a conformation generated using the small-molecule topology generator PRODRG [49]. Model building was performed using the program Coot [50] and refinement was carried out using the CCP4i graphical user interface [51] to the REFMAC program [52].

RESULTS

Cell-based screen for inhibitors of the UPS

To identify novel cell-active small-molecule inhibitors of the UPS, a high-throughput screen of the Millennium Pharmaceuticals, Inc., library was performed for compounds that stabilized 4xUb-Luc in MDA-MB-231 cells. In this system, bortezomib and other catalytic 20S inhibitors cause a concentration-dependent accumulation of the constitutively expressed reporter by blocking its intracellular degradation, as reported previously [15]. To further validate the dependence of this reporter on the 26S proteasome, siRNAs were used to deplete selected components of the 20S core particle and 19S regulatory complex in MDA-MB-231-4xUb-Luc cells. As shown in Figure 2, depletion of subunits of either the 20S (α1 and β6) or 19S (rpn8 and rpn11) complex resulted in a time-dependent increase in the 4xUb-Luc signal

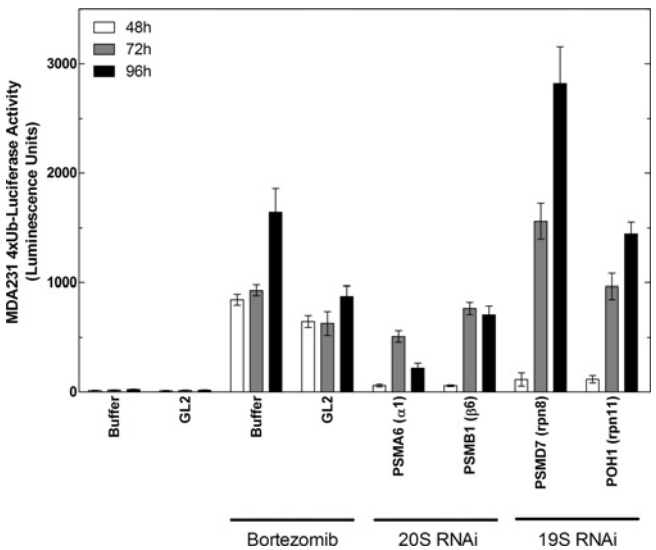


Figure 2 Effects of bortezomib and depletion of 26S subunits by RNAi on accumulation of 4xUb-Luc in cells

MDA-MB-231 cells stably expressing a 4xUb-Luc reporter were reverse-transfected with 10 nM siRNA SMARTPools to deplete components of the 19S regulatory cap or 20S catalytic core particle of the 26S proteasome, as indicated. At 48, 72 or 96 h after transfection, the accumulation of the 4xUb-Luc was measured as luminescence using firefly luciferase reagents, as described in the Experimental section. Mock transfections with buffer or with the GL2 luciferase control duplex were also performed to monitor basal expression of the 4xUb-Luc reporter and its stabilization by 1 μM bortezomib for the final 8 h of transfection. Results are the means ± S.E.M. for six assay points and are representative of four independent experiments. PSMA, proteasome subunit α-type; PSMB, proteasome subunit β-type; PSMD, proteasome 26S subunit non-ATPase; rpn, regulatory particle non-ATPase.

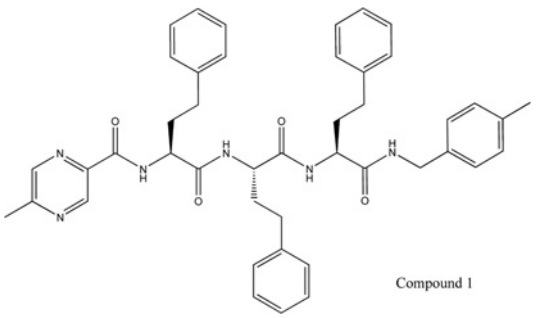


Figure 3 Structure of the capped tri-peptide screening hit, 5-methyl-*N*-[(*S*)-1-{(*S*)-1-[(*S*)-1-(4-methylbenzylamino)-1-oxo-4-phenylbutan-2-ylamino]-1-oxo-4-phenylbutan-2-ylamino}-1-oxo-4-phenylbutan-2-yl]pyrazine-2-carboxamide, compound 1

comparable with that observed by an 8 h treatment of the cells with a near-saturating concentration of bortezomib of 1 μM.

High-throughput screening of approx. 352 500 compounds at 50 μM gave 3015 hits that stabilized 4xUb-Luc by 2-fold or greater, representing a primary hit rate of 0.86 %, and 611 of these positives were confirmed when re-tested in triplicate at the same concentration and cut-off, representing a re-test rate of 20.3 %. The majority of these compounds were either peptide boronic acid inhibitors of the proteasome or ubiquitin-activating enzyme (E1) inhibitors that were part of the library generated by proprietary medicinal chemistry and served as screening controls. There were 191 compounds which were previously uncharacterized as inhibitors of the UPS, for which cell-based potency determinations were performed. Of these, 52 gave modest 2–3-fold stabilization, whereas 102 compounds gave a 3–10-fold stabilization and 37 compounds gave >10-fold stabilization of the reporter at concentrations up to 50 μM.

Identification of compound 1 as a potent selective 20S β5 inhibitor

The cell-active compounds were assessed for inhibition of purified rat brain 20S *in vitro* by measuring hydrolysis of the β5 (chymotrypsin-like) substrate Ac-WLA-AMC in the presence of PA28α. Only a small minority of compounds inhibited 20S, showing low and sub-micromolar potencies (results not shown). Among the most potent of these were four closely related peptidic benzyl amides synthesized by combinatorial chemistry, of which one was subsequently identified by MS and re-synthesis as a C- and N-terminally capped tri-peptide derived from the unnatural amino acid *S*-homo-phenylalanine, and designated as compound 1 (Figure 3) *In vitro*, this compound inhibited the β5 activity of the human constitutive and immunoproteasome with low nanomolar potencies, comparable with those of bortezomib, but was essentially inactive against the β1 (caspase-like) and β2 (trypsin-like) sites of either proteasome isoform (Table 1). β5 selectivity is in contrast with that of bortezomib, which also inhibits the β1 and, to a lesser extent, the β2 sites of the constitutive and immunoproteasome (Table 1). Assays against a diverse panel of ten proteases indicated that compound 1 was highly selective for the proteasome, showing only partial (<50 %) inhibitory activity against cathepsin L, coagulation Factor β-XIIa, thrombin, tissue plaminogen activator and trypsin at concentrations approx. 10000-fold higher than its IC_{50} value for the 20S β5 site (Supplementary Table S2 at http://www.BiochemJ.org/bj/430/bj4300461add.htm). Kinetic analysis of human 20S β5 inhibition by compound 1 showed that it behaved as a rapid-equilibrium inhibitor with a K_i value of 0.50 $\pm$ 0.12 nM (mean $\pm$ S.E.M., $n = 3$), consistent with reversible non-covalent binding to the active site of the enzyme (Figure 4). This behaviour is again in contrast with that of bortezomib, which acts as a time-dependent slow–tight binding inhibitor that interacts covalently with the active site Thr[1] residue, but which displays an equivalent K_i value for the β5 site of 0.56 $\pm$ 0.072 nM (mean $\pm$ S.E.M., $n = 3$) (Figure 4 and [25]).

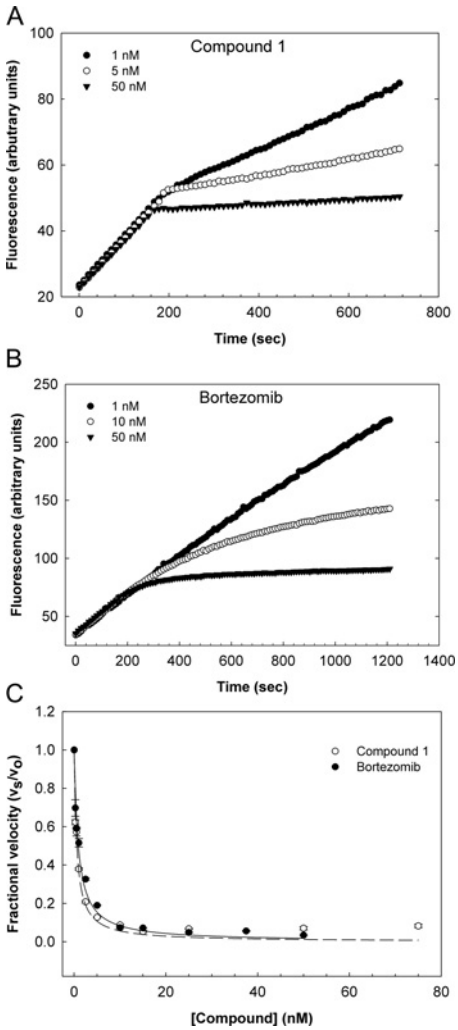


Figure 4 Kinetics of proteasome inhibition by compound 1 and bortezomib

(**A** and **B**) The peptidase activity of human erythrocyte 20S (0.25 nM) was monitored over time using the fluorigenic substrate Suc-LLVY-AMC (10 μM) in the presence of 0.01 % SDS and following the addition of various concentrations of either compound 1 or bortezomib, as indicated. (**C**) A non-linear least-fit of the fractional velocity of the reactions, v_s/v_0, as a function of inhibitor concentration, [I], was used to determine a K_i of 0.50 $\pm$ 0.12 nM for compound 1 and 0.56 $\pm$ 0.072 nM for bortezomib (means $\pm$ S.E.M., $n = 3$).

Cellular inhibition of proliferation, 4xUb-Luc degradation, NFκB-Luc activation by compound 1

On the basis of the high affinity and selectivity of compound 1 for the 20S β5 site *in vitro*, its activity was characterized further in cells as compared with that of bortezomib. Compound 1 inhibited the proliferation of MDA-MB-231 cells in a concentration-dependent manner with an LC_{50} (half-maximal lethal concentration) value of 150 $\pm$ 23 nM (mean $\pm$ S.E.M., $n = 3$), as determined by an ATPlite assay at 72 h (see Supplementary

Table 1 Active-site-selectivity of the peptide boronic acid, bortezomib, the capped tri-peptide screening hit, compound 1, and a series of fifteen capped di-peptides *in vitro*

Purified human erythrocyte 20S and peripheral blood monocyte 20S (0.25 nM each) were used as a source of constitutive (c) and immunoproteasome (i) respectively to measure the peptidase activity of the β1 (caspase-like), β2 (trypsin-like) and β5 (chymotrypsin-like) sub-sites of each proteasome isoform. Assays were performed in the presence of 15 μM of the indicated fluorigenic peptide-AMC substrate, 12 nM PA28α and a titration of each inhibitor. As shown, and unlike bortezomib, compounds 1–16 were essentially inactive against the β1 and β2 sites of the constitutive proteasome and immunoproteasome. The IC_{50} values of the compounds against the β5c and β5i sites are mean values for at least three independent determinations.

	Human 20S–PA28 IC_{50} (nM)					
Compound	β1c Z-LLE-AMC	β1i Ac-PAL-AMC	β2c Ac-KQL-AMC	β2i Ac-KQL-AMC	β5c Ac-WLA-AMC	β5i Ac-ANW-AMC
Bortezomib	140	5.5	1500	940	8.2	3.3
1	28000	>100000	57000	>100000	16	8.9
2	83000	>100000	>100000	>100000	1100	230
3	43000	65000	54000	47000	130	37
4	48000	>100000	>100000	>100000	470	41
5	39000	>100000	>100000	70000	340	27
6	>100000	>100000	>100000	44000	25	10
7	70000	>100000	>100000	91000	9100	1300
8	31000	83000	63000	47000	13	4.1
9	68000	>100000	>100000	>100000	6.8	6.7
10	90000	>100000	>100000	>100000	4.6	3.2
11	>100000	>100000	>100000	>100000	2.4	1.1
12	89000	>100000	95000	>100000	4.4	3.5
13	84000	>100000	99000	>100000	3.9	3.4
14	66000	>100000	>100000	>100000	11	200
15	64000	>100000	>100000	>100000	12	44
16	>100000	>100000	>100000	>100000	1.2	1.1

Figure S1A at http://www.BiochemJ.org/bj/430/bj4300461add.htm). Indeed, viability assays performed across a panel of five cancer cell lines, comprising MDA-MB-231, Calu6, H460, HCT116 and HT29 cells, indicated that compound 1 has broad cytotoxic activity, with LC_{50} values ranging from 50 nM to 380 nM and being approx. 10–20-fold less potent than bortezomib (Table 2 and Supplementary Table S3 at http://www.BiochemJ.org/bj/430/bj4300461add.htm). In the screening assay, compound 1 caused up to an approx. 100-fold accumulation of the 4xUb-Luc reporter at 8 h [maximal effect of 52 ± 16-fold, with an EC_{50} (half-maximally effective concentration) of 690 ± 22 nM (mean $\pm$ S.E.M., $n = 6$)], showing an approx. 3-fold higher EC_{50} and one-seventh of the efficacy of bortezomib [maximal effect, 340 ± 12-fold; $EC_{50} = 240 \pm 21$ nM (mean $\pm$ S.E.M., $n = 102$)] (Supplementary Figure S1B and Table 2). The activity of compound 1 was also assessed in the Proteasome-Glo™ assay that measures directly the peptidase activity of the β5 site of the 26S proteasome in cells using the luminogenic substrate Suc-LLVY-aminoluciferin. Treatment of Calu6 cells for 1 h with compound 1 caused a concentration-dependent inhibition of β5 activity in this assay with an IC_{50} (half-maximal inhibitory concentration) value of 53 ± 16 nM (mean $\pm$ S.E.M., $n = 3$). Bortezomib was approx. 10-fold more potent, inhibiting hydrolysis of the chymotrypsin-like substrate with an IC_{50} of 3.9 ± 0.49 nM (mean $\pm$ S.E.M., $n = 8$) (Supplementary Figure S1C and Table 2). Finally, we assessed the effect of compound 1 on activation of the transcription factor NFκB by TNF-α, which is dependent on proteasomal degradation of IκBα (inhibitory protein of NFκB α) following its rapid phosphorylation by the IκB kinase complex and ubiquitination by the $SCF^{\beta TrCP}$ (E3 ubiquitin-protein ligase complex Skp1–Cullin–F-box) [53]. Compound 1 inhibited TNFα-induced activation of NFκB-Luc in HEK-293 cells with an IC_{50} of 47 ± 7.7 nM (mean $\pm$ S.E.M., $n = 7$), consistent with potent proteasome inhibition in these cells. Interestingly, it did not fully inhibit NFκB-Luc, even at saturating concentrations [maximal inhibition, 79 ± 3.2 % (mean $\pm$ S.E.M., $n = 7$)], whereas bortezomib inhibited NFκB-Luc activity by 100 % with an IC_{50} of 9.7 ± 0.73 nM (mean $\pm$ S.E.M., $n = 14$) (Supplementary Figure S1D and Table 2). Collectively, these data indicate that selective inhibition of the β5 site of the proteasome by compound 1 in cells is sufficient to inhibit the degradation of the 4xUb-Luc reporter, TNF-α-dependent NFκB activity and the proliferation of cancer cells with potencies that are approximately one order of magnitude lower than those of bortezomib.

X-ray crystallography of compound 1 in complex with the yeast 20S proteasome

To determine its binding mode, compound 1 was soaked into crystals of the yeast 20S core particle. X-ray diffraction data obtained at 3.1 Å resolution indicated that it only occupied the β5 sites of the proteasome (Figure 5A), consistent with selective inhibition of this active site *in vitro*. The C-terminal 4-methylbenzyl group (P1) was seen to occupy a well-defined S1 pocket, with the first homo-phenylalanine residue (P2) pointing toward solvent, the second (P3) occupying a well-defined S3 pocket and the third (P4) projecting into a much larger and ill-defined S4 pocket (Figure 5B). No electron density was observed corresponding to the N-terminal cap, presumably because this region of the molecule is disordered. The structure of compound 1 overlaid with that of bortezomib, as determined previously [26], suggests that, although the side chains of the two compounds differ in their orientations, most of the backbone hydrogen-bond interactions within the β5 site are the same (Figure 5C). Of note, in this structure as in others, residual electron density corresponding to Mes, the buffer in which the crystals were prepared, was observed near Thr^1 in the β5 active site (Figure 5A). The main interaction between Mes and compound 1 involves hydrophobic contacts between the morpholino group of Mes and the homo-phenylalanine residue in the P2 position. For structures without ligand in the active site, no density for the morpholino group was observed (M.D. Sintchak, unpublished work), rather electron density at the position of the sulfur atom of Mes was seen. This sits

Table 2 Cell-based activity of bortezomib, compound 1 and fifteen capped di-peptide proteasome inhibitors

The cell-based potencies of bortezomib and compounds 1–16 were assessed in a diverse panel of assays that monitor the short- and longer-term effects of proteasome inhibition in cells, as indicated. In particular, inhibition of the $\beta5$ activity of the proteasome was measured directly with the Proteasome-Glo™ assay using Suc-LLVY-aminoluciferin as substrate following incubation of Calu6 cells (10000 per well) with compound for 1 h. Inhibition of NFκB-luciferase activity in HEK-293 cells was determined by pre-incubation of the cells (10000 per well) with compound for 1 h followed by stimulation with 10 ng/ml TNF-α in the continued presence of compound for 3 h. The IC_{50} and percentage inhibition of NFκB-Luc activity are given. Accumulation of the 4xUb-Luc reporter in MDA-MB-231 cells was determined by incubation of the cells (10000 per well) with compound for 8 h and expressed as the fold accumulation of luciferase activity together with an EC_{50} value for this effect. The effects of the compounds on the viability of Calu6 cells (2000 per well) and H460 cells (1000 per well) were assessed by the ATPlite assay following incubation of the cell with compound for 72 h. The results are mean values for at least three independent experiments.

	Calu6 Proteasome-Glo™ $\beta5$	HEK-293 NFκB-Luc		MDA231 4xUb-Luc		Cell viability	
Compound	IC_{50} (nM)	IC_{50} (nM)	Maximal inhibition (%)	EC_{50} (nM)	Maximal induction (fold)	Calu6 LC_{50} (nM)	H460 LC_{50} (nM)
Bortezomib	3.9	9.7	100	240	340	5.2	13
1	53	47	79	690	52	130	380
2	48000	>50000	28	1800	3	>25000	>25000
3	610	3100	84	2100	39	1700	3100
4	13000	24000	30	6900	12	13000	18000
5	12000	>50000	16	4700	28	2000	2600
6	310	320	89	2100	76	1200	4300
7	5400	4400	86	10000	65	6600	17000
8	43	29	87	930	86	240	470
9	10	22	78	130	77	10	34
10	16	6.6	72	47	64	55	82
11	5.7	9.8	93	112	65	21	110
12	36	10	74	71	64	36	110
13	24	9.2	77	64	76	61	240
14	520	64	77	650	64	210	650
15	230	39	80	450	82	355	1000
16	9.5	12	92	31	75	9.6	44

at the same position as the nucleophilic water molecule referred to as NUK in a previous publication [54]. Importantly, although Mes was observed in the active site following crystallization, this buffer molecule neither affects the specific activity of 20S nor the affinity of compound 1 for the active site on the basis of K_i value determination (F.J. Bruzzese, unpublished work).

Optimization of capped di-peptide 20S inhibitors

The partial crystal structure of compound 1 (Figure 5) together with the report of a potent related tri-methoxy-phenylalanine-containing tri-peptide 20S inhibitor (Figure 1 and [38]) suggested that one homo-phenylalanine residue could be deleted, provided that a compensating hydrophobic N-terminal capping group could be incorporated in place of the pyrazine. A series of di-peptide-like libraries was therefore prepared by solution-phase parallel synthesis as outlined in the Experimental section (Scheme 2), denoting residues P1–P4 (Figure 6) according to the binding mode indicated by the structure of compound 1 (Figure 5). Given the close correlation in IC_{50} values of proteasome inhibitors from various chemical classes for human and yeast 20S (C. Tsu, unpublished work), compounds were optimized for *in vitro* $\beta5$ potency against human 20S, whereas crystal structures of representative examples were determined with the yeast 20S open-gate mutant in order to guide further library design.

The first library synthesized retained 4-methylbenzylamine as the C-terminal capping group at P1, contained varied amino acid residues at P2 and P3 and probed two alternative hydrophobic N-terminal caps at P4 with 348 compounds in total. All of these di-peptide analogues were less potent than the trimer compound 1, although several inhibited $\beta5$ activity with low and sub-micromolar potencies. For example, the truncated analogue of compound 1 with an N-terminal cap substitution, compound 2 (Figure 7), had a $\beta5$ IC_{50} value of 1.1 μM *in vitro* (Table 1). Substitution of the homo-phenylalanine residues of compound 2 by a tryptophan at P1 and by the di-methoxy-phenylalanine residue [38] at P2 gave compound 3 (Figure 7) with a $\beta5$ IC_{50} value of 130 nM (Table 1). In the next library, smaller replacements for the P3 homo-phenylalanine residue were sought in combination with alternative N-terminal P4 capping groups that provided compensating hydrophobic interactions within the large S4 binding pocket. Numerous compounds, such as compound 4 and the 4-benzyloxybenzoyl derivative compound 5, with small polar P3 residues, in this case a threonine residue (Figure 7), retained reasonable potency *in vitro* ($\beta5$ IC_{50} values of approx. 500 nM). Significant further improvements were effected by varying the capping group, with the indanone in compound 6 being among the optimal residues at P4 with this compound showing a $\beta5$ IC_{50} of 25 nM (Figure 7 and Table 1) and a K_i value of 13 ± 1.4 nM (mean $\pm$ S.E.M., $n = 3$).

The crystal structure of compound 6 bound to the $\beta5$ site of yeast 20S was determined at 2.6 Å resolution (Figure 8) and compared with that of bortezomib determined at 2.7 Å, a resolution comparable with that determined previously [26]. In the structure of bortezomib, in addition to the covalent and hydrogen-bonding interactions involving the boronate, backbone hydrogen-bonding interactions with the $\beta5$ site can be seen (Figure 8A). Interestingly, the same hydrogen-bonding interactions between the backbone of compound 6 and the $\beta5$ site are observed, including the presence of the bridging water molecule between the P3 carbonyl group and Ala^{49} and Ala^{50} (Figure 8A). Figures 8(B) and 8(C) show the occupancy of the binding pockets by compound 6, whereas Supplementary Figure S2 (at http://www.BiochemJ.org/bj/430/bj4300461add.htm) shows an alternative view to include the electron density of the water molecule. The 4′-methylbenzyl P1 residue occupies S1, with the methyl group occupying a small sub-pocket at the bottom of the cavity. The P2 homo-phenylalanine residue occupies S2 and the P3 threonine residue projects, to a small degree, towards the deep S3 binding pocket. The P2 homo-phenylalanine residue does not fill the S2 binding pocket, which is quite shallow, but likely confers potency by contributing a hydrophobic

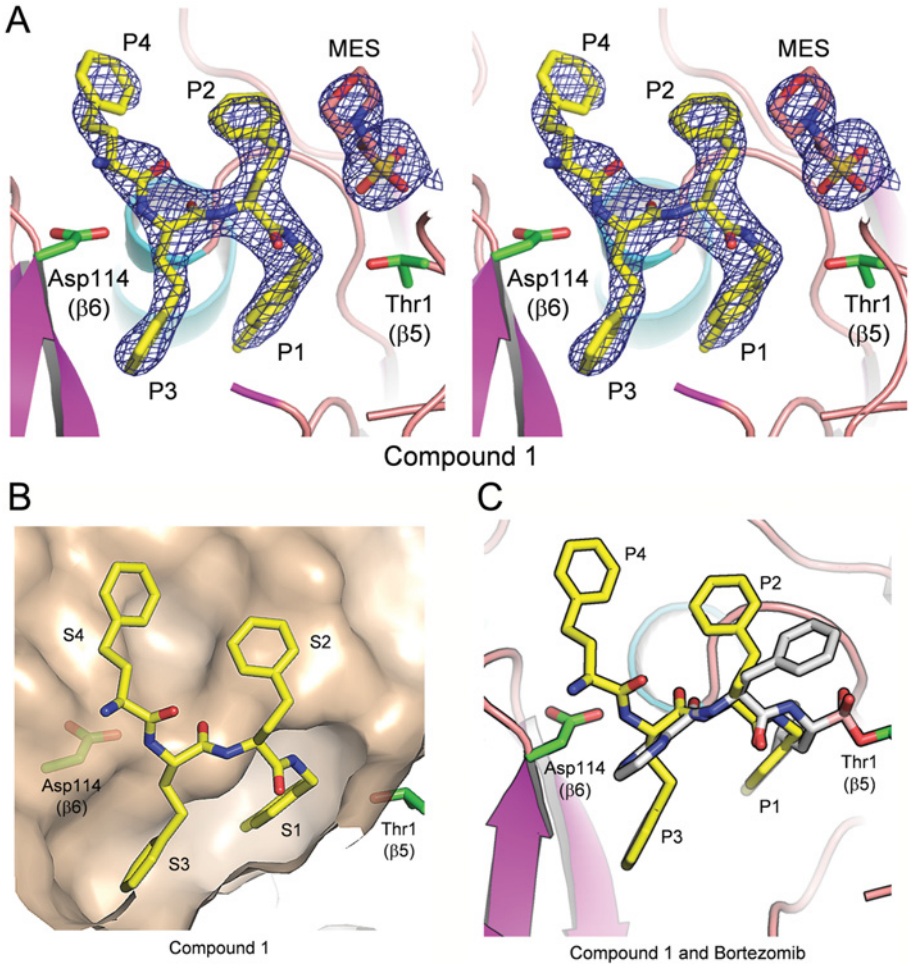


Figure 5 Electron-density map of compound 1 and occupancy of the specificity pockets of the 20S β5 subunit

(**A**) A 3.1 Å sigmaA-weighted $2F_o$-F_c electron-density map contoured at 1.0σ covering compound 1 (PDB entry 3MG4) and Mes in the $\beta5/\beta6$ active site of wild-type 20S is shown in stereo view, with residue positions P1–P4 indicated. Atoms are in ball-and-stick representation with carbon in yellow (compound 1) or salmon (Mes), nitrogen in blue and oxygen in red. The side chain atoms for Thr^1 ($\beta5$) and Asp^{114} ($\beta6$) are shown in ball-and-stick representation, coloured as above, except for carbon in green. Protein atoms from 20S are shown in cartoon representation coloured by secondary structure with helices in cyan, β-strands in magenta and loops in salmon. The Figure was made using the PyMOL Molecular Graphics System (DeLano Scientific, Palo Alto, CA, U.S.A.). (**B**) The $\beta5/\beta6$ active site with compound 1 bound (coloured as in **A**), showing the molecular surface of 20S. Specificity pockets S1–S4 are indicated. (**C**) Superposition of compound 1 and bortezomib bound to the $\beta5/\beta6$ active site of 20S is shown, with positions P1–P4 indicated. Colour scheme is as described above, except bortezomib carbon atoms are coloured grey.

interaction at a 'ledge' formed from Gly^{47}-Gly^{48} (Figure 8C). Of note, the P4 indanone in compound 6 is closely related to inhibitors reported previously [35,36], for example, CVT-659 (Figure 1B). Whereas the indanone carbonyl in these molecules was originally designed as a 'putative electrophilic head group' [35,36], the crystal structure of compound 6 places this moiety in the S4 binding pocket remote from the $\beta5$ catalytic Thr^1 residue, thereby precluding such an interaction. In addition, the water molecule appears to be present in all higher-resolution structures (<2.6 Å), including those without ligand bound, such as that reported originally [54]. This water molecule was also observed in structures with inhibitors having IC_{50} values greater than 1 μM (M.D. Sintchak, unpublished work). Therefore it is unlikely that this single water-mediated hydrogen bond contributes significantly to potency.

In the third library, optimization of the C-terminal P1 residue was explored. The requirement of the hydrogen-bonding interaction with Gly^{47} can be seen from N-methylation of P1 to give compound 7 (Figure 7), which diminished activity greatly ($\beta5$ $IC_{50} = 9.1$ μM; Table 1). However, comparable potency with compound 6 was observed for the 2′-chlorobenzyl derivative compound 8 ($\beta5$ $IC_{50} = 13$ nM; Table 1 and Figure 7). A crystal of compound 8 bound to 20S indicated an additional hydrophobic interaction between the ortho-halogen and a second 'sub-pocket' within the S1 site that cannot be accessed by the less active meta-substituted analogue ([42] and see Supplementary Figure S3 at http://www.BiochemJ.org/bj/430/bj4300461add.htm).

In the final library, the influence of a hydrophobic P3 residue was assessed, with emphasis on asparagine derivatives. As with the potency-enhancing N-terminal P4 capping groups described above, several low nanomolar inhibitors were identified, such as compounds 9 ($\beta5$ $IC_{50} = 6.8$ nM, Table 2) and 10 ($\beta5$ $IC_{50} = 4.6$ nM, Table 2) (Figure 7). Potent compounds such as 11 ($\beta5$ $IC_{50} = 2.4$ nM, Table 1) and 12 ($\beta5$ $IC_{50} = 4.4$ nM, Table 1) in which the homo-phenylalanine residue was replaced by an alanine residue and compound 13 ($\beta5$ $IC_{50} = 3.9$ nM, Table 1), in which 3-pyridyl was incorporated, were also identified (Figure 7). Similarly, the neopentyl-asparagine residue was able to compensate for sub-optimal interactions in S1, with compounds 14 and 15 showing good inhibitory activity ($\beta5$ $IC_{50} = 11$ and

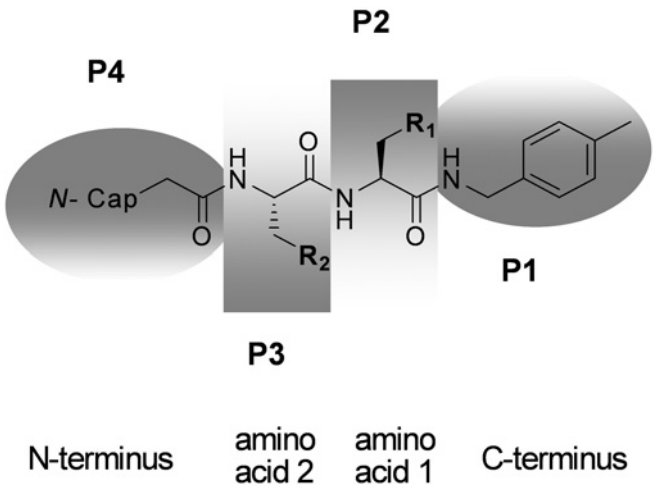


Figure 6 Residue assignments of di-peptide inhibitors

Optimization of compound 1 involved contracting the chain to a di-peptide of the general structure shown and systematically modifying the central amino acids and capping groups, as denoted by residues P1–P4 in a series of compound libraries prepared by solution-phase parallel synthesis (see the Experimental section).

12 nM respectively; Table 1 and Figure 7). Like compound 1, these di-peptides at 100 μM had negligible to modest activity against ten unrelated proteases (Supplementary Table S2), indicating that the most potent compounds are selective for the chymotrypsin-like site of the proteasome by 4 or 5 orders of magnitude.

The most potent proteasome inhibitor identified was compound 16 (Figure 7) with a $\beta5$ $IC_{50} = 1.2 \pm 0.35$ nM for the human enzyme (Table 1) and a K_i below the 20S concentration of 250 pM in the assay, indicating that it behaved as a non-covalent active-site titrant (F.J. Bruzzese, unpublished work). The X-ray structure of this compound bound to yeast 20S resolved at 2.6 Å is shown in Figure 9. The potentially electrophilic α-ketoamide N-terminal cap does not form any covalent interactions; it is positioned in the S4 binding pocket which is located away from the catalytic Thr[1] and it also does not appear to interact with any residues in S4 (Figure 9A). Importantly, the high potency of compound 16 and all compounds with a P3 neopentyl-asparagine group can be accounted for by the near-optimal fit of this residue in the S3 binding pocket (Figure 9B). As with compound 8 (Supplementary Figure S3), the P1 ortho-chlorine of compound 16 is also likely to contribute to the potency by accessing the S1 sub-pocket (Figure 9A). Of note, however, the aromatic P2 contains only one methylene and therefore the ledge interaction cannot take place with the pyridyl projecting toward solvent (compare with Figure 8C).

Relationship between 20S $\beta5$ potency and cellular activity of capped di-peptides

The effects of the di-peptides on 26S activity in cells were measured (Table 2) and compared with their 20S $\beta5$ IC_{50} values determined *in vitro* (Table 1 and Supplementary Figure S4 at http://www.BiochemJ.org/bj/430/bj4300461add.htm). Inhibition of the $\beta5$ site of the 26S proteasome was measured directly in

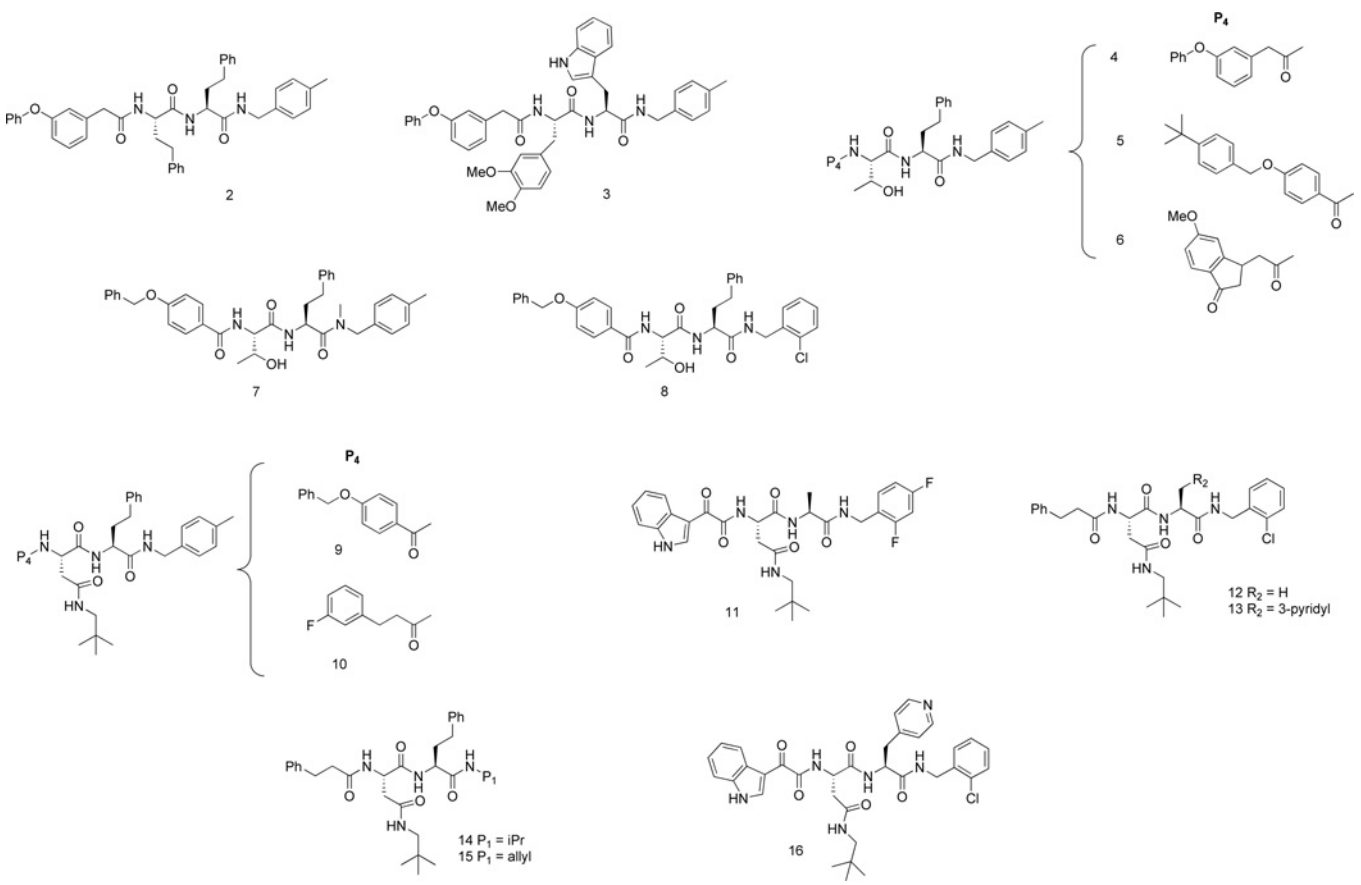


Figure 7 Chemical structures of capped di-peptide 20S $\beta5$ proteasome inhibitors

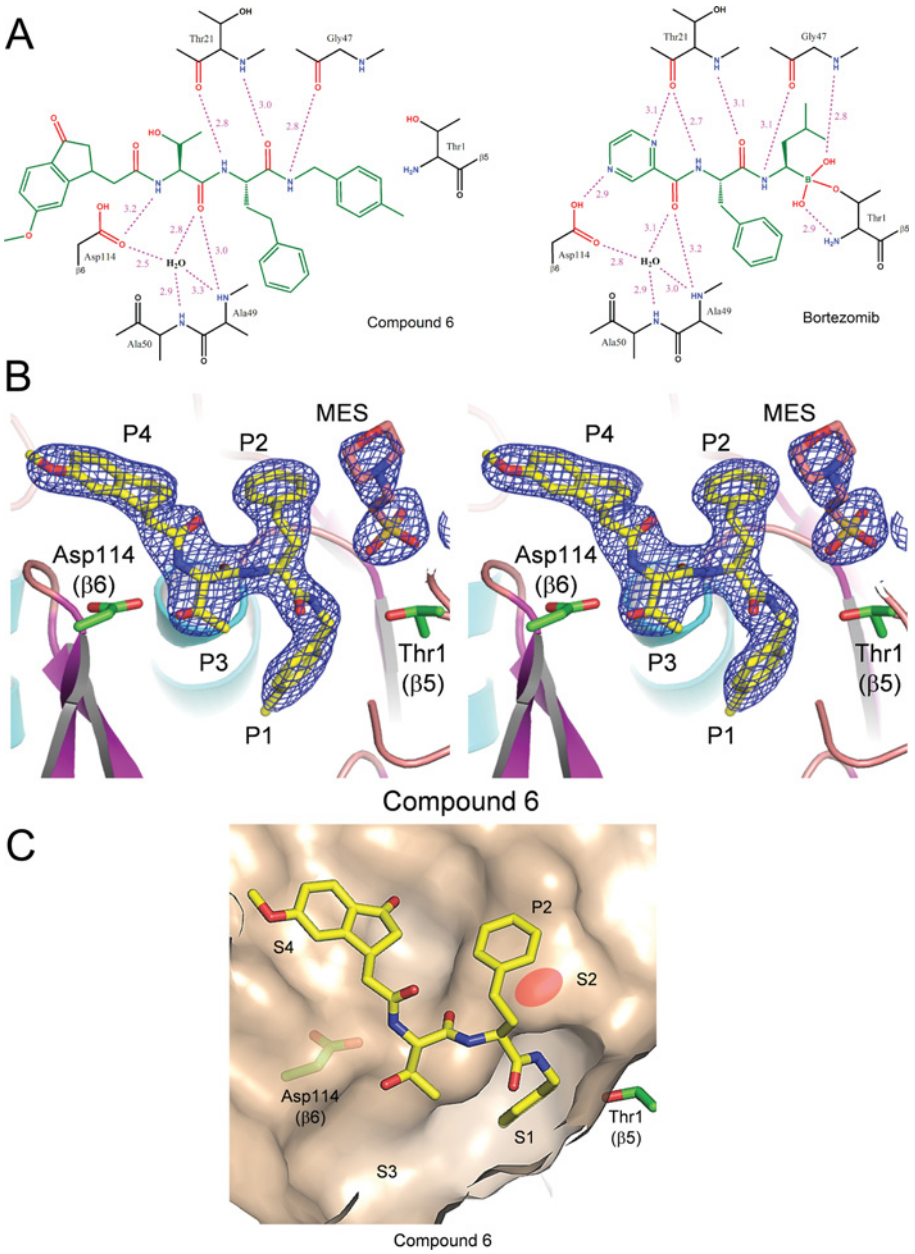


Figure 8 Hydrogen-bonding interactions and crystal structure of compound 6 bound to the chymotrypsin-like site of the 20S proteasome with reference to bortezomib

(**A**) Schematic representation of the $\beta5/\beta6$ active site of 20S with compound 6 (left-hand panel) (PDB entry 3MG6) and bortezomib (right-hand panel) (PDB entry 3MG0) bound. Key hydrogen bonds between the inhibitors and the protein are shown as dashed lines coloured magenta, with distances indicated in Å. Thus in the structure of bortezomib, the NH of the P1 leucine residue forms a hydrogen bond with the carbonyl of Gly^{47}, whereas the carbonyl and amino groups of the P2 residue form hydrogen-bond interactions with Thr^{21}. The carbonyl of the pyrazinoyl cap is involved in a network of hydrogen-bond interactions with Ala^{49}, Ala^{50} and Asp^{114} that includes a bridging water molecule. Similar hydrogen-bonding interactions between the backbone of compound 6 and the $\beta5$ site are shown, including the presence of the bridging water molecule between the P3 carbonyl and Ala^{49} and Ala^{50}; the NH of P3 makes the hydrogen-bond interaction with Asp^{114} that is formed by the pyrazine N in bortezomib. (**B**) 2.6 Å sigmaA-weighted $2F_o$-F_c electron-density map contoured at 1.0σ covering compound 6 (PDB entry 3MG6) and MES is shown in the $\beta5/\beta6$ active site of yeast 20S open-gated mutant in stereo view, with positions P1–P4 indicated. The colour scheme is as described in Figure 5. (**C**) $\beta5/\beta6$ active site with compound 6 bound (coloured as described above), showing the molecular surface of 20S. Specificity pockets S1–S4 are indicated. The P2 binding 'ledge' is shown by shading of the molecular surface.

Calu6 cells using the Proteasome-Glo™ assay and corresponded well with the inhibitory potencies determined *in vitro* with purified 20S. There was also a striking correlation between the IC_{50} values of the compounds for inhibition of NFκB-Luc activity in cells and 20S $\beta5$ activity *in vitro*, with five of the analogues, specifically compounds 10–13 and 16, showing inhibitory NFκB potencies in the range 6.6–12 nM, equivalent to or greater than that of bortezomib [NFκB-Luc $IC_{50} = 13 \pm 2.0$ nM (mean $\pm$ S.E.M., $n = 6$)]. Interestingly, the activity of NFκB-Luc was not fully inhibited even at saturating compound concentrations (Table 2). The di-peptides also stabilized the 4xUb-Luc reporter in MDA-MB-231 cells with EC_{50} values that tracked closely with their *in vitro* 20S $\beta5$ potencies (Table 2), although the maximal stabilization effects of approx. 100-fold were lower than those observed for bortezomib (Table 2), other boronates (J.L. Blank, unpublished work) and covalent inhibitors, such as

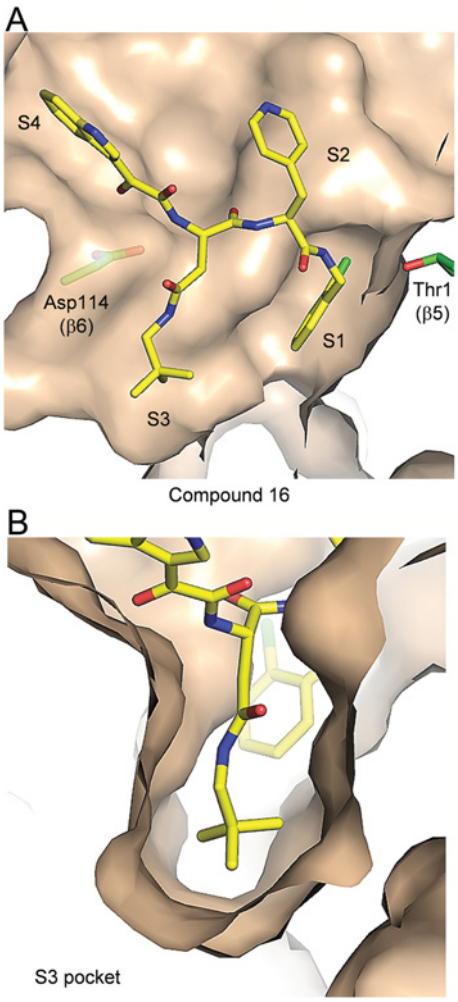


Figure 9 Crystal structure of compound 16 bound to the chymotrypsin-like site of the 20S proteasome

(**A**) $\beta5/\beta6$ active site with compound 16 (PDB entry 3MG8) bound to yeast open-gated 20S, showing the molecular surface of 20S (coloured as described in Figure 5). Selectivity pockets S1–S4 are indicated. (**B**) Close-up view of the S3 specificity pocket to illustrate occupancy by the neopentyl-asparagine residue of compound 16 in the $\beta5/\beta6$ active site of the 20S core particle (coloured as described above).

salinosporamide A [15]. Finally, the cytotoxic LC_{50} values of these compounds also correlated well between cell lines (Table 2) and with their 20S $\beta5$ potencies (Table 1). Interestingly, none of the compounds were as potent as bortezomib in this assay, although several gave LC_{50} values that were approx. 100 nM or less in at least one cell line, specifically compounds 9–13 and 16 (Table 2). Significantly, each of these contained the neopentyl-asparagine substituent in P3 (Figure 8), indicating that this residue contributes to both *in vitro* and cell-based potency. Furthermore, since these compounds were also the most potent inhibitors of NFκB-Luc activity, these data strongly suggest that their effects on cell viability and NFκB activity are mediated by proteasome inhibition.

Effect of 20S $\beta5$ inhibition on bulk protein turnover

The majority of intracellular proteins are degraded by the proteasome, as demonstrated using inhibitors that irreversibly inactivate each of the 20S active sites [4]. The effect of selectively inhibiting the $\beta5$ site alone on the degradation of short-lived protein was assessed using compounds 10 and 16, two of the most potent cell-active di-peptides. As shown previously [41], bortezomib inhibits approx. 50 % of bulk protein turnover in HCT116 cells (see Supplementary Figure S5 at http://www.BiochemJ.org/bj/430/bj4300461add.htm), an effect that is comparable with that observed with ML858, a synthetic version of salinosporamide A that irreversibly inhibits each of the 20S active sites [15] (T.A. Soucy, unpublished work). Under identical conditions, the di-peptides reduced protein turnover by approx. 20 % (Supplementary Figure S5), indicating that inhibition of the $\beta5$ site alone results in partial inhibition of protein degradation.

Constitutive and immunoproteasome selectivity of di-peptide proteasome inhibitors

In addition to monitoring the $\beta5$ inhibitory potency of the capped di-peptides using purified human erythrocyte constitutive 20S proteasomes, IC_{50} values were also obtained for the $\beta5$i site of the immunoproteasome from human peripheral blood monocytes using the selective substrates Ac-WLA-AMC and Ac-ANW-AMC respectively (Table 1). Bortezomib, compound 1 and the majority of the capped di-peptides were non-selective in these assays, inhibiting the constitutive and immunoproteasome with similar IC_{50} values, although they tended to inhibit $\beta5$i with slightly greater potencies (Table 1). In common with compound 1, the di-peptides were essentially inactive with respect to inhibition of the $\beta1$i or $\beta2$i sites of the immunoproteasome. However, a limited number of compounds showed significantly greater selectivity for the immunoproteasome, with compounds 4 and 5 being the best examples, showing $\sim$10-fold $\beta5$i selectivity [compound 4, $\beta5$i $IC_{50} = 41 \pm 3.3$ nM and $\beta5$c $IC_{50} = 470 \pm 96$ nM; compound 5, $\beta5$i $IC_{50} = 27 \pm 2.7$ nM and $\beta5$c $IC_{50} = 340 \pm 31$ nM (means $\pm$ S.E.M., $n = 3$)]. Conversely, compounds 14 and 15 showed the reverse selectivity, the former of these two compounds being 18-fold more selective for the constitutive proteasome [$\beta5$c $IC_{50} = 11 \pm 2.3$ nM; $\beta5$i $IC_{50} = 200 \pm 62$ nM (means $\pm$ S.E.M., $n = 3$)] and the latter 3.6-fold more selective [$\beta5$c $IC_{50} = 12 \pm 2.8$ nM (mean $\pm$ S.E.M., n = 4); $\beta5$i $IC_{50} = 44 \pm 9.2$ nM (mean $\pm$ S.E.M., $n = 3$)] (Table 1).

To determine whether this selectivity was maintained in cells, three B-cell lymphoma subtypes were used that differ in their proteasome isoform expression. As shown in Figure 10(A), OCI-Ly10 cells do not express the immunoproteasome, as demonstrated by the negligible hydrolysis of the $\beta1$i- and $\beta5$i-specific substrates Ac-PAL-AMC and Ac-ANW-AMC respectively, whereas Karpas-1106P express substantial $\beta1$i and $\beta5$i activities, indicating considerable enrichment of the immunoproteasome. The hydrolysis of the $\beta5$c substrate Ac-WLA-AMC supported this expression profile, although it is not entirely selective for $\beta5$c over $\beta5$i and therefore over-represents the amount of constitutive proteasome in Karpas-1106P cells. The chymotrypsin-like substrate Suc-LLVY-AMC is hydrolysed approximately equally well by the $\beta5$i and $\beta5$c sites *in vitro*, although cells enriched in the immunoproteasome can show approx. 2–3-fold higher activity with this substrate. On the basis of these data, Karpas-1106P cells are highly enriched for the immunoproteasome, whereas OCI-Ly10 cells only express the constitutive proteasome. WSU-DLCL2 cells are intermediate in this activity profile, indicating that they express a mixed population of the proteasome isoforms.

We assessed the potency of the $\beta5$i- and $\beta5$c-selective compounds 5 and 14 respectively in the three B-cell

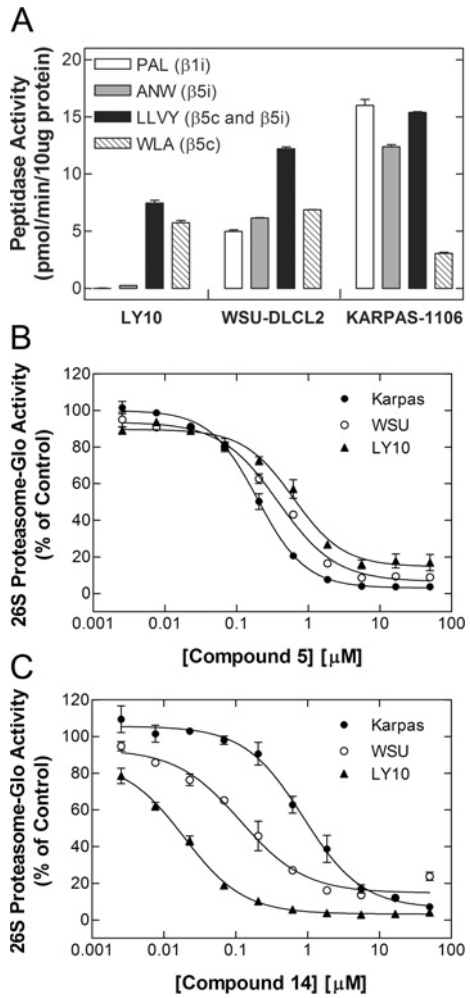


Figure 10 Proteasome-subtype selectivity of compounds 5 and 14 in B-cell lymphomas expressing different levels of the immunoproteasome

(**A**) The activities of the β1 and β5 subunits of the constitutive (c) and immunoproteasome (i) were assessed in diffuse large B-cell lymphoma extracts (10 μg of protein per reaction) using recombinant PA28α (12 nM) and subtype-selective fluorigenic peptide substrates (100 μM), as indicated. As shown, the rank order of immunoproteasome expression in these cell lines was Karpas-1106 > WSU-DLCL2 $\gg$ LY10. (**B** and **C**) The B-cell lymphoma lines (20000 cells per well) were treated with the indicated concentrations of the immunoproteasome-selective compound 5 (**B**) or constitutive-selective compound 14 (**C**) for 1 h under cell culture conditions. Inhibition of the β5 site of the constitutive and immunoproteasome was then assessed *in situ* using the Proteasome-Glo^{TM} assay with the chymotrypsin-like luminogenic substrate Suc-LLVY-aminoluciferin, as described in the Experimental section. As shown, the IC_{50} values of compound 5 in Karpas-1106, WSU-DLCL2 and LY10 cells were 200, 380 and 620 nM respectively; the IC50 values of compound 14 in Karpas-1106, WSU-DLCL2 and LY10 cells were 870, 120 and 19 nM respectively. Results are means $\pm$ S.E.M. for triplicate determinations and are representative of three independent experiments, the results of which are summarized in Table 3.

lymphomas using the non-subtype-selective β5 substrate Suc-LLVY-aminoluciferin in the Proteasome-Glo^{TM} assay. As shown in Figure 10(B) and Table 3, the IC_{50} values of the β5i-selective compound 5 correlated with immunoproteasome expression (i.e. the rank order of potency being Karpas-1106P > WSU-DLCL2 > OCI-Ly10), although the differences were relatively modest (i.e. approx. $\leqslant$ 4-fold). However, the β5c-selective inhibitor compound 14, which was the most isoform-selective compound *in vitro*, displayed substantial selectivity for the constitutive proteasome in cells (Figure 10C). The rank order of effect observed (i.e. OCI-Ly10 > WSU-DLCL2 > Karpas-1106P) indicated an approx. 22-fold increase in inhibitory potency

Table 3 Proteasome inhibition in B-cell lymphoma subtypes expressing differing levels of immunoproteasome by β5i- and β5c-selective compounds 5 and 14 respectively compared with bortezomib

Inhibition of the β5 sites of the constitutive and immunoproteasome was assessed in the indicated B-cell lymphoma lines (20000 cells per well) *in situ* using the Proteasome-Glo^{TM} assay. Cells were incubated with compound for 1 h and assays were performed with the chymotrypsin-like substrate Suc-LLVY-aminoluciferin as described in the Experimental section. The results are the means $\pm$ S.E.M. for triplicate determinations.

	Proteasome-Glo^{TM} β5 IC_{50} (nM)		
Compound	OCI-Ly10	WSU-DLCL2	Karpas-1106P
5	1300 $\pm$ 400	320 $\pm$ 120	310 $\pm$ 150
14	39 $\pm$ 7.6	190 $\pm$ 76	850 $\pm$ 240
Bortezomib	2.9 $\pm$ 0.25	2.7 $\pm$ 0.26	3.3 $\pm$ 0.40

of compound 14 in cells that exclusively express the constitutive proteasome as compared with those that predominantly express the immunoproteasome. Importantly, bortezomib, which does not discriminate substantially between β5c and β5i sites *in vitro* (Table 1), gave IC_{50} values that were within 2-fold of each other in the three B-cell lymphomas (Table 3). These data therefore provide the first description of non-covalent inhibitors of the proteasome that can discriminate between the immunoproteasome and constitutive proteasome *in vitro* and in cells.

DISCUSSION

The present study has identified a new series of potent non-covalent proteasome inhibitors on the basis of a capped trimeric peptide derived from the unnatural amino acid *S*-homophenylalanine, first identified by a large-scale high-throughput cell-based screen for small-molecule inhibitors of the UPS. These capped di-peptides, prepared by high-throughput liquid-phase peptide synthesis methods, are entirely selective for the β5 site over the β1 and β2 sites of the 20S core particle, and over a panel of less closely related proteases. The most potent compounds show IC_{50} values in the single-digit nanomolar range for the human 20S β5 site *in vitro*, with the best example, compound 16, displaying an IC_{50} of 1.2 nM and a K_i below the enzyme concentration in the assay (0.25 nM 20S), indicating that it behaves as an active-site titrant. In this respect, compound 16 is remarkable in having a greater affinity for the β5 site of the proteasome than that of the covalent inhibitor bortezomib, which displays a K_i of 0.56 nM for this site.

Compound optimization was guided by X-ray crystallography of representative examples bound to the yeast 20S core particle, which established their non-covalent binding mode and provided a structural basis for potency-enhancing modifications. The crystal structures of compounds 1, 6, 8 and 16 bound to the 20S β5 active site are shown in Figures 5, 8, S2, and 9 respectively, and are overlaid in Supplementary Figure S6 (at http://www.BiochemJ.org/bj/430/bj4300461add.htm). The overlay confirms a common binding mode in which the P1 benzyl group occupies a well-defined S1 binding pocket, whereas the P2 residue points out towards the solvent, unless it is the homophenylalanine residue in which case a hydrophobic 'ledge' interaction can be accessed. The S3 binding pocket is also well-defined, whereas the S4 pocket is very large and can tolerate a range of capping groups that are likely to contribute to potency by picking up further hydrophobic interactions. This binding mode broadly corresponds to that proposed on the basis of a homology model [38], except that the crystal structures obtained in the present study place the N-terminal

cap in a large S4 binding pocket rather than in small 'accessory hydrophobic pockets', designated previously AS1 and AS2 [38]. Importantly, the most potent compounds identified each contained a neopentyl-asparagine substituent in P3, which appears to provide a near-optimal fit for the S3 binding pocket of the $\beta5$ site (Figure 9B). This residue can also be found in the tri-peptide 2-keto-1,3,4-oxadiazoles (Figure 1A) that are slowly reversible and presumably covalent in nature [11].

Characterization of these compounds in cells demonstrated that they can functionally inhibit the 26S proteasome with potencies that correlated well with those determined *in vitro* using the purified enzyme (Table 2 and Supplementary Figure S4). Furthermore, since the analogues are essentially inactive against the $\beta1$ (caspase-like) and $\beta2$ (trypsin-like) sites of the proteasome (Table 1), these data indicate that inhibition of the $\beta5$ (chymotrypsin-like) site of the proteasome alone is sufficient to potently inhibit the degradation of the 4xUb-Luc reporter, activation of NFκB in response to TNF-α and the proliferation of cancer cells. Interestingly, despite the fact that the cellular IC_{50} values of these compounds spanned three orders of magnitude, with the most potent being in the single- or double-digit nanomolar range, neither the stabilization of 4xUb-Luc nor the inhibition of the NFκB pathway was complete, even at saturating compound concentrations (Table 2 and Supplementary Figure S1). These observations suggest that non-covalent inhibition of 20S $\beta5$ activity alone is not sufficient to fully inhibit the proteasomal degradation of either 4xUb-Luc or IκBα, at least over the timescale of these cell-based assays. This is in contrast with the covalent inhibitors bortezomib and salinosporamide A, which completely block the degradation of these proteins in cells and can minimally inhibit both the $\beta1$ and $\beta5$ sites of the proteasome. Furthermore, estimates of bulk protein turnover in cells suggest that the overall catalytic rate of proteasome is reduced, but not abolished, by non-covalent inhibition of the $\beta5$ site (Supplementary Figure S5), although it remains a possibility such inhibition completely blocks the degradation of a subset of proteins that are only substrates of the $\beta5$ site and not the $\beta1$ and/or $\beta2$ sites of the proteasome. Importantly, however, prolonged inhibition of the $\beta5$ site alone appears to sufficient to inhibit cancer cell proliferation.

Proteasome inhibitors that are currently in clinical use do not discriminate between the chymotrypsin-like activities of the constitutive or immunoproteasome, and the potential therapeutic benefit of selective inhibition of either of these activities has not been demonstrated. However, recent evidence from preclinical studies with PR-957, a $\beta5$i-selective peptide epoxyketone analogue of carfilzomib [28], suggests that the anti-inflammatory effects of proteasome inhibitors, such as bortezomib, may be mediated by their inhibition of the chymotrypsin-like site of the immunoproteasome [8]. Our studies have identified compounds that differentially inhibit the $\beta5$ sites of the constitutive and immunoproteasome *in vitro* and in B-cell lymphomas. Of note, compound 5 that preferentially inhibited the $\beta5$i site contained the small threonine residue in P3 and the larger 4-methylbenzylamine aromatic cap in P1. Conversely, the two $\beta5$c-selective compounds 14 and 15 each contained a smaller residue at P1 and the bulkier potency-enhancing neopentyl-asparagine in P3 (Figure 8). This structure–activity relationship, albeit limited, is entirely consistent with selectivity data obtained from an approx. 6000 tri-peptide library screen for substrates that are preferentially hydrolysed by each active site of the constitutive and immunoproteasome ([55] and C. Tsu, unpublished work). For example, the tri-peptide substrate Ac-ANW-AMC, where P1 is a tryptophan residue and P3 is an alanine residue, is preferentially cleaved by the $\beta5$i subunit of the immunoproteasome, whereas Ac-WLA-AMC, where the P1 and P3 residues are reversed, is preferentially hydrolysed by the constitutive $\beta5$c subunit. The crystal structure of the immunoproteasome has not yet been determined to provide an explanation for these selectivity differences. However, the primary sequences of the mature forms of the human $\beta5$c and $\beta5$i subunits are 60.7 % identical (see Supplementary Figure S7 at http://www.BiochemJ.org/bj/430/bj4300461add.htm). We have therefore mapped the sequence differences between the $\beta5$c and $\beta5$i subunits on to the bovine 20S crystal structure [56] and modelled compound 1 into the active site (see Supplementary Figure S8 at http://www.BiochemJ.org/bj/430/bj4300461add.htm). The majority of the sequence differences between $\beta5$c and $\beta5$i lie outside of the active site, whereas the residues that make contacts with compound 1 appear to be mainly the same, with some relatively minor exceptions, such as the replacement of Ala^{46} and Ser^{53} with Ser^{46} and Gln^{53} within the S1 binding pocket, Gly^{48} with Cys^{48} in the S2 'ledge', and the reversal of the Ala^{27}-Ser^{28} sequence with Ser^{27}-Ala^{28} in the S3 pocket. Since only subtle differences exist between the ligand-interacting residues that form the $\beta5$c and $\beta5$i active sites, it is not possible to rationalize selectivity differences with respect to substrate preference or inhibitor sensitivity based on such an analysis. Clearly, crystal structures of the constitutive and immunoproteasomes in complex with selective inhibitors will help identify the structural features of each sub-site that contribute to selectivity. Finally, although a covalent immunoproteasome-selective inhibitor has been described recently [8], the function of the immunoproteasome in cancer cells is unknown. The identification of a new class of proteasome inhibitors, such as compounds 14 and 5, that preferentially inhibit the 20S $\beta5$c and $\beta5$i sites respectively and can thus discriminate between the constitutive and immunoproteasome in cells may help address their functional roles in various cellular models of disease including cancer.

AUTHOR CONTRIBUTION

Christopher Tsu, Paul Hales and Frank Bruzzese and designed and performed the biochemical characterization of the compounds. Cynthia Barrett, Jane Liu and Khristofer Garcia performed the cell-based assays and Teresa Soucy performed the protein turnover assays. Darshan Sappal generated the MDA231-4xUb-Luc cell line and performed the high-throughput library screen. Edward Olhava performed the re-synthesis of compound 1, and Christopher Blackburn, Kenneth Gigstad and Matthew Jones designed and performed the chemical syntheses of the capped di-peptides. Nancy Bump purified yeast 20S core particles for crystallization and Michael Sintchak prepared the crystals for X-ray crystallography and analysed the data. Paul Fleming and Lawrence Dick provided project oversight, participated in experimental design and data interpretation, and reviewed the manuscript. Jonathan Blank performed the high-throughput screen and the initial characterization of compound 1 in cells, contributed to experimental design and data analysis, and wrote the manuscript with Christopher Blackburn.

ACKNOWLEDGEMENTS

We are grateful to Dr Marion Schmidt and Dr Dan Finley for providing the yeast 20S expression strains, and to Dr Juan Gutierrez and Zhi Li their assistance with preparation of the yeast 20S enzymes. We also thank Alejandra Raimondi for formatting the 4xUb-Luc cell-based assay for the high-throughput screening, Ben Knight for his assistance with data processing and management, and Maria Dawn-Linsley, Kim Lincoln, Charles Gauthier, Dan Onea, Pam Ward, Ted Peters, Christina Majer and Elena Musotto for their help with assay automation and compound management. We gratefully acknowledge the assistance of Marjorie Solomon, David Lok, Lenny Dang and Qing Lu with the identification of compound 1 by MS. We thank Mingkun Fu who performed the high-resolution MS on the synthetic compounds. We acknowledge Abe Achab and Alla Mishechkina for their synthetic contributions towards development of the series. We are also grateful to Dr Gary Luker (Department of Microbiology and Immunology, University of Michigan Medical School, Ann Arbor, MI, U.S.A.) and Dr David Piwnica-Worms (Division of Biology and Biomedical Sciences, University of Washington, St Louis, MO, U.S.A.) from whom the 4xUb-Luc expression plasmid was licenced. We are also grateful to Dr Friedrich Feuerhake and Dr Margaret Shipp (Lymphoma Program, Dana-Faber Cancer Institute, Boston, MA,

U.S.A.) for providing cell pellets derived from a panel of B-cell lymphoma subtypes to profile for immunoproteasome expression. Finally, we thank Mark Williamson and Dr Mark Rolfe for valuable discussions.

FUNDING

This work was supported by Millennium Pharmaceuticals, Inc. (Cambridge, MA, U.S.A.). All authors were employed by Millennium Pharmaceuticals at the time of their contribution to this work.

REFERENCES

1 Hershko, A. and Ciechanover, A. (1998) The ubiquitin system. Annu. Rev. Biochem. **67**, 425–479
2 Baumeister, W., Walz, J., Zuhl, F. and Seemuller, E. (1998) The proteasome: paradigm of a self-compartmentalizing protease. Cell **92**, 367–380
3 Pickart, C. M. and Cohen, B. (2004) Proteasomes and their kin: proteases in the machine age. Nat. Rev. Mol. Cell Biol. **5**, 177–187
4 Kisselev, A. F and Goldberg, A. L. (2001) Proteasome inhibitors: from research tools to drug candidates. Chem. Biol. **8**, 739–758
5 Borissenko, L. and Groll, M. (2007) 20S Proteasome and its inhibitors: crystallographic knowledge for drug development. Chem. Rev. **107**, 687–717
6 Goldberg, A. L. (2007) Functions of the proteasome: from protein degradation and immune surveillance to cancer chemotherapy. Biochem. Soc. Trans. **35**, 12–17
7 Kloetzel, P.-M. (2001) Antigen processing by the proteasome. Nat. Rev. Mol. Cell Biol. **2**, 179–187
8 Muchamuel, T., Basler, M., Aujay, M. A., Suzuki, E., Kalim, K. W., Lauer, C., Sylvain, C., Ring, E. R., Shields, J., Jiang, J. et al. (2009) A selective inhibitor of the immunoproteasome subunit LMP7 blocks cytokine production and attenuates progression of experimental arthritis. Nat. Med. **15**, 781–787
9 Groll, M., Huber, R. and Moroder, L. (2009) The persisting challenge of selective and specific proteasome inhibition. J. Pept. Sci. **15**, 58–66
10 García-Echeverria, C. (2006) Peptide and peptide-like modulators of 20S proteasome enzymatic activity in cancer cells. Int. J. Pept. Res. Ther. **12**, 49–64
11 Rydzewski, R. M., Burrill, M., Mendonca, R., Palmer, J. T., Rice, M., Tahilramani, R., Bass, K. E., Leung, L., Gjerstad, E., Janc, J. W. and Pan, L. (2006) Optimization of subsite binding to the $\beta 5$ 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
12 Adams, J., Behnke, M., Chen, S., Cruickshank, A. A., Dick, L. R., Grenier, L., Klunder, J. M., Ma, Y-T., Plamondon, L. and Stein, R. L. (1998) Potent and selective inhibitors of the proteasome: dipeptidyl boronic acids. Bioorg. Med. Chem. Lett. **8**, 333–338
13 Adams, J., Palombella, V. J., Sausville, E. A., Johnson, J., Destree, A., Lazarus, D. D., Maas, J., Pien, C. S., Prakash, S. and Elliott, P. J. (1999) Proteasome inhibitors: a novel class of potent and effective antitumor agents. Cancer Res. **59**, 2615–2622
14 Hideshima, T., Richardson, P., Chauhan, D., Palombella, V. J., Elliott, P. J., Adams, J. and Anderson, K. C. (2001) The proteasome inhibitor PS-341 inhibits growth, induces apoptosis, and overcomes drug resistance in human multiple myeloma cells. Cancer Res. **61**, 3071–3076
15 Williamson, M. J., Blank, J. L., Bruzzese, F. J., Cao, Y., Daniels, J. S., Dick, L. R., Labutti, J., Mazzola, A. M., Patil, A. D., Reimer, C. L. et al. (2006) Comparison of biochemical and biological effects of ML858 (salinosporamide A) and bortezomib. Mol. Cancer Ther. **5**, 3052–3061
16 Richardson, P. G., Barlogie, B., Berenson, J., Singhal, S., Jagannath, S., Irwin, D., Rajkumar, S. V., Srkalovic, G., Alsina, M., Alexanian, R. et al. (2003) A phase 2 study of bortezomib in relapsed, refractory myeloma. N. Eng. J. Med. **348**, 2609–2617
17 Richardson, P. G., Sonneveld, P., Schuster, M. W., Irwin, D., Stadtmauer, E. A., Facon, T., Harousseau, J.-L., Ben-Yehuda, D., Lonial, S., Goldschmidt, H. et al. (2005) Bortezomib or high-dose dexamethasone for relapsed multiple myeloma. N. Eng. J. Med, 352, 2487–2498
18 Richardson, P. G., Sonneveld, P., Schuster, M. W., Irwin, D., Stadtmauer, E. A., Facon, T., Harousseau, J.-L., Ben-Yehuda, D., Lonial, S., Goldschmidt, H. et al. (2007) Extended follow-up of a phase 3 trial in relapsed multiple myeloma: final time-to-event results of the APEX trial. Blood **110**, 3557–3560
19 San Miguel, J. F., Schlag, R., Khuageva, N. K., Dimopoulos, MA., Shpilberg, O., Kropff, M., Spicka, I., Petrucci, M. T., Palumbo, A., Samoilova, O. S. et al. (2008) Bortezomib plus melphalan and prednisone for initial treatment of multiple myeloma. N. Eng. J. Med. **359**, 906–917
20 Fisher, R. I., Bernstein, S. H., Kahl, B., Djulbegovic, B., Robertson, M. J., de Vos, S., Epner, E., Krishnan, A., Leonard, J. P., Lonial, S. et al. (2006) Multicenter phase II study of bortezomib in patients with relapsed or refractory mantle cell lymphoma. J. Clin. Oncol. **24**, 4867–4874
21 Voorhees, P. M. and Orlowski, R. Z. (2006) The proteasome and proteasome inhibitors in cancer therapy. Annu. Rev. Pharmacol. Toxicol. **46**, 189–213
22 Richardson, P. G., Mitsiades, C., Hideshima, T. and Anderson, K. C. (2006) Bortezomib: proteasome inhibition as an effective anticancer therapy. Annu. Rev. Med. **57**, 33–47
23 Orlowski, R. Z. and Kuhn, D. J. (2008) Proteasome inhibitors in cancer therapy: lessons from the first decade, Clin. Cancer Res.**14**, 1649–1657
24 Piva, R., Ruggeri, B., Williams, M., Costa, G., Tamagno, I., Ferrero, D., Giai, V., Coscia, M., Peola, S., Massaia, M. et al. (2008) CEP-18770: a novel, orally active proteasome inhibitor with a tumor-selective pharmacologic profile competitive with bortezomib. Blood **111**, 2765–2775
25 Kupperman, E., Lee, E. C., Cao, Y., Bannerman, B., Fitzgerald, M., Berger, A., Yu, J., Yang, Y., Bruzzese, F., Liu, J. et al. (2010) Evaluation of the proteasome inhibitor MLN9708 in preclinical models of human cancer. Cancer Res. **70**, 1970–1980
26 Groll, M., Berkers, C. R., Ploegh, H. L. and Ovaa, H. (2006) Crystal structure of the boronic acid-based proteasome inhibitor bortezomib in complex with the yeast 20S proteasome. Structure **14**, 451–456
27 Chauhan, D., Catley, L., Li, G., Podar, K., Hideshima, T., Velankar, M., Mitsiades, C., Mitsiades, N., Yasui, H., Letai, A. et al. (2005) A novel orally active proteasome inhibitor induces apoptosis in multiple myeloma cells with a mechanism distinct from bortezomib. Cancer Cell **8**, 407–419
28 Kuhn, D. J., Chen, Q., Voorhees, P. M., Strader, J. S., Shenk, K. D., Sun, C. M., Demo, S. D., Bennett, M. K., van Leeuwen, F. W. B., Chanan-Khan, A. A. and Orlowski, R. Z. (2007) Potent activity of carfilzomib, a novel, irreversible inhibitor of the ubuquitin–proteasome pathway, against preclinical models of multiple myeloma. Blood **110**, 3281–3290
29 Groll, M., Huber, R. and Potts, B. C. M. (2006) Crystal structures of salinoporamide A (NPI-0047) in complex with the 20S porteasome reveal important consequences of β-lactone ring opening and a mechanism for irreversible binding. J. Am. Chem. Soc. **128**, 5136–5141
30 Groll, M., Kim, K. B., Kairies, N., Huber, R. and Crews, C. M. (2000) Crystal structure of expoxomicin:20S proteasome reveals a molecular basis for selectivity of α',β'-epoxyketone proteasome inhibitors. J. Am. Chem. Soc. **122**, 1237–1238
31 Koguchi, Y., Kohno, J., Nishio, M., Takahashi, K., Okuda, T., Ohnuki, T. and Komatsubara, S. (2000) TMC-95A, B, C, and D, novel proteasome inhibitors produced by *Apiospora montagnei Sacc.* TC 1093. Taxonomy, production, isolation, and biological activities. J. Antibiot. **53**, 105–109
32 Groll, M., Koguchi, Y., Huber, R. and Kohno, J. (2002) Crystal structure of the 20 S proteasome:TMC-95A complex: a non-covalent proteasome inhibitor. J. Mol. Biol. **311**, 543–548
33 Groll, M., Gotz, M., Kaiser, M., Weyher, E. and Moroder, L. (2006) TMC-95-based inhibitor design provides evidence for the catalytic versatility of the proteasome. Chem. Biol. **13**, 607–614
34 Basse, N., Piguel, S., Papapostolou, D., Ferrier-Berthelot, A., Richey, N., Pagano, M., Sarthou, P., Sobczak-Thépot, J., Reboud-Ravaux, M. and Vidal, J. (2007) Linear TMC-95-based proteasome inhibitors. J. Med. Chem. **50**, 2842–2850
35 Lum, R. T., Kerwar, S. S., Meyer, S. M., Nelson, M. G., Schow, S. R., Shiffman, D., Wick, M. M. and Joy, A. (1998) A new structural class of proteasome inhibitors that prevent NF-κB activation. Biochem. Pharmacol. **55**, 1391–1397
36 Lum, R. T., Nelson, M. G., Joly, A., Horsma, A. G., Lee, G., Meyer, S. M., Wick, M. M. and Schow, S. R. (1998) Selective inhibition of the chymotrypsin-like activity of the 20S proteasome by 5-methoxy-1-indanone di-peptide benzamides. Bioorg. Med Chem. Lett. **8**, 209–214
37 García-Echeverría, C., Imbach, P., France, D., Furst, P., Lang, M., Noorani, M., Scholz, D., Zimmermann, J. and Furet, P. (2001) A new class of selective and non-covalent inhibitors of the chymotrypsin-like activity of the 20S proteasome. Bioorg. Med. Chem. Lett. **11**, 1317–1319
38 Furet, P., Imbach, P., Noorani, M., Koeppler, J., Laumen, K., Lang, M., Guagnano, V., Fuerst, P., Roesel, J. and García-Echeverría, C. (2004) Entry into a new class of potent proteasome inhibitors having high antiproliferative activity by structure-based design. J. Med. Chem. **47**, 4810–4813
39 Luker, G. D., Pica, C. M., Song, J., Luker, K. E. and Piwnica-Worms, D. (2003) Imaging 26S proteasome activity and inhibition in living mice. Nat. Med. **9**, 969–973
40 Lightcap, E. S., McCormack, T. A., Pien, C. S., Chau, V., Adams, J. and Elliott, P. J. (2000) Proteasome inhibition measurements: clinical applications. Clin. Chem. **46**, 673–683
41 Soucy, T. A., Smith, P. G., Milhollen, M. A., Berger, A. J., Gavin, J. M., Adhikari, S., Brownell, J. E., Burke, K. E., Cardin, D. P., Critchley, S. et al. (2009) An inhibitor of NEDD8-activating enzyme as a new approach to treat cancer. Nature **458**, 732–737
42 Blackburn, C., Achab, A., Blank, J., Bump, N., Bruzzese, F., Dick, L., Fleming, P., Garcia, K., Gigstad, K., Hales, P. et al. (2009) Identification and optimization of a series of non-covalent proteasome inhibitors guided by X-ray crystallography. 36th Northeast Regional Meeting of the American Chemical Society, Hartford, CT, U.S.A., October 7–10 2009, abstract NERM-055
43 Groll, M., Bajorek, M., Köhler, A., Moroder, L., Rubin, D. M., Huber, R., Glickman, M. H. and Finley, D. (2000) A gated channel into the proteasome core particle. Nat. Stuct. Biol. **7**, 1062–1067
44 Leggett, D. S., Hanna, J., Borodovsky, A., Crosas, B., Schmidt, M., Baker, R. T., Walz, T., Ploegh, H. and Finley, D. (2002) Multiple associated proteins regulate proteasome structure and function. Mol. Cell **10**, 495–507

45 Leggett, D. S., Glickman, M. H. and Finley, D. (2005) Purification of proteasomes, proteasome subcomplexes and proteasome-associated proteins from budding yeast. Methods Mol. Biol. **301**, 57–70
46 Groll, M. and Huber, R. (2005) Purification, crystallization, and X-ray analysis of yeast 20S proteasome. Methods Enzymol. **398**, 329–336
47 Leslie, A. G. W. (2006) The integration of macromolecular data. Acta Crystallogr. D Biol. Crystallogr. **62**, 48–57
48 Evans, P. (2005) Scaling and assessment of data quality. Acta Crystallogr. D Biol. Crystallogr. **62**, 72–82
49 Schüttelkopf, A. W. and van Aalten, D. M. F. (2004) PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. Acta Crystallogr. D Biol. Crystallogr. **60**, 1355–1363
50 Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. **60**, 2126–2132
51 Potterton, E., Briggs, P., Turkenburg, M. and Dodson, E. (2003) A graphical user interface to the CCP4 program suite. Acta Crystallogr. D Biol. Crystallogr. **59**, 1131–1137
52 Murshudov, G. N., Vagin, A. A. and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr. D Biol. Crystallogr. **53**, 240–255
53 Kroll, M., Margottin, F., Kohl, A., Renard, P., Durand, H., Concordet, J.-P., Bachelerie, F., Arenzana-Seisdedos, F. and Benarous, R. (1999) Inducible degradation of IκBα by the proteasome requires interaction with the F-box protein h-βTrCP. J. Biol. Chem. **274**, 7941–7945
54 Groll, M., Ditzel, L., Löwe, J., Stock, D., Bochtler, M., Bartunik, H. D. and Huber, R. (1997) Stucture of 20S proteasome from yeast at 2.4 Å resolution. Nature **386**, 463–471
55 Lin, G., Tsu, C., Dick, L., Zhou, X. K. and Nathan, C. (2008) Distinct specificities of *Mycobacterium tuberculosis* and mammalian proteasomes for *N*-acetyl tripeptide substrates J. Biol. Chem. **283**, 34423–34431
56 Unno, M., Mizushima, T., Morimoto, Y., Tomisugi, Y., Tanaka, K., Yasuoka, N. and Tsukihara, T. (2002) The structure of the mammalian 20S proteasome at 2.75 Å resolution. Structure **10**, 609–618

Received 15 March 2010/2 July 2010; accepted 15 July 2010
Published as BJ Immediate Publication 15 July 2010, doi:10.1042/BJ20100383

Biochem. J. (2010) **430**, 461–476 (Printed in Great Britain) doi:10.1042/BJ20100383

online data

SUPPLEMENTARY ONLINE DATA

Characterization of a new series of non-covalent proteasome inhibitors with exquisite potency and selectivity for the 20S β5-subunit

Christopher BLACKBURN, Kenneth M. GIGSTAD, Paul HALES, Khristofer GARCIA, Matthew JONES, Frank J. BRUZZESE, Cynthia BARRETT, Jane X. LIU, Teresa A. SOUCY, Darshan S. SAPPAL, Nancy BUMP, Edward J. OLHAVA[1], Paul FLEMING, Lawrence R. DICK, Christopher TSU, Michael D. SINTCHAK and Jonathan L. BLANK[2]

Discovery, Millennium Pharmaceuticals, Inc., 40 Landsdowne Street, Cambridge, MA 02139, U.S.A.

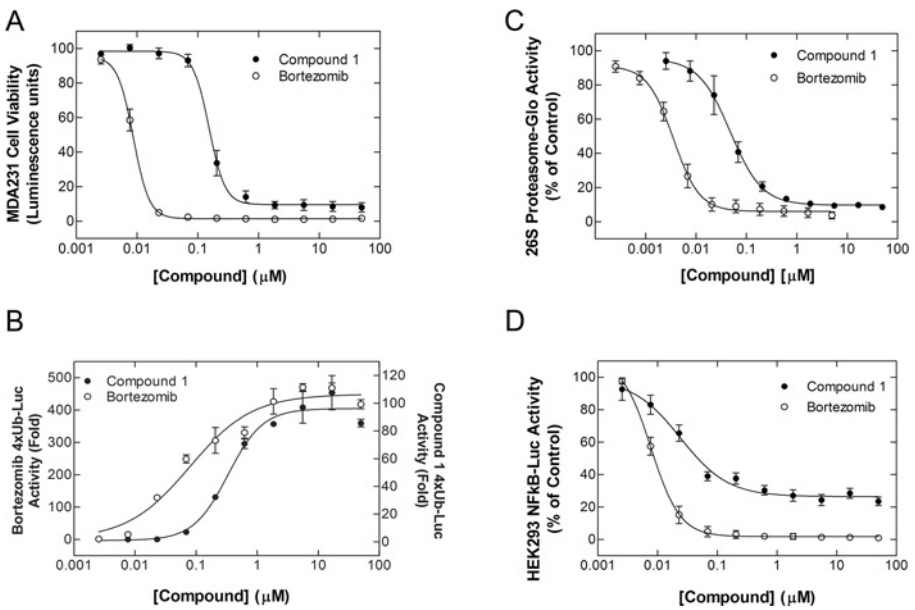


Figure S1 Activities of compound 1 and bortezomib in cells

(**A**) The concentration-dependent effects of compound 1 and bortezomib on the proliferation of MDA-MB-231-4xUb-Luc reporter cell line were determined by the ATPlite assay at 72 h. (**B**) Accumulation of the 4xUb-Luc reporter in the stable MDA-MB-231 cell line was determined by incubation with the indicated concentrations of compounds for 8 h followed by luciferase assay. (**C**) The β5 activity of the 26S proteasome was measured in Calu6 cells by the Proteasome-Glo™ assay using Suc-LLVY-aminoluciferin as a substrate following incubation of the cells with the indicated concentration of compounds for 1 h. (**D**) Inhibition of NFκB–luciferase activity in the stable HEK-293 cell line was determined by pre-incubation of the cells with the indicated concentration of compounds for 1 h followed by stimulation with 10 ng/ml TNF-α in the continued presence of compound for a further 3 h. Results are means $\pm$ S.E.M. from replicate determinations and are representative of several independent experiments.

SUPPLEMENTARY EXPERIMENTAL

Protease selectivity assays

Protease selectivity of the compounds was assessed using a panel of purified human enzymes (all from Calbiochem, EMD Chemicals) by monitoring hydrolysis rates of peptide-AMC substrates (all from Bachem) using fluorescence measurements as described for '*In vitro* assays of purified 20S' in the main paper, with modifications as follows. To measure the activities of coagulation Factor β-XIIa (200 nM), chymotrypsin (2 nM), elastase (4 nM), plasmin (4 nM), thrombin (4 nM), tissue plaminogen activator (25 nM) and trypsin (4 nM) (final concentrations), reactions were performed at 30 °C in 50 μl of buffer containing 50 mM Tris/HCl, pH 8.0, 100 mM NaCl, 5 mM $CaCl_2$, 0.01 % Tween 20 and their respective substrates Boc-QGR-AMC (50 μM), Suc-AAPF-AMC (40 μM), MeOSuc-AAPV-AMC (50 μM), H-AFK-AMC (65 μM), Bz-FVR-AMC (20 μM), Boc-E(OBzl)GR-AMC (50 μM) and Z-FR-AMC (12 μM). The activities of cathepsin B and L were measured using 2 nM enzyme in 50 μl of 100 mM sodium acetate, pH 5.5, containing 5 mM DTT (dithiothreitol), 1 mM EDTA, 0.01 % Brij 35 and the respective substrates Z-RR-AMC (60 μM) and Z-FR-AMC (37.5 μM). Calpain I activity was measured using 200 nM enzyme in 50 μl of 50 mM Tris/HCl, pH 7.4, containing 75 μM $CaCl_2$, 1.5 mM DTT and 100 μM Suc-LY-AMC. Percentage inhibition of the proteases was

[1] Present address: Epizyme, Inc., 840 Memorial Drive, Cambridge, MA 02139, U.S.A.
[2] To whom correspondence should be addressed (e-mail jonathan.blank@mpi.com).

The structural co-ordinates of the yeast 20S proteasome with the indicated ligand bound reported will appear in the PDB under accession codes: 3MG0 (bortezomib); 3MG4 (compound 1); 3MG6 (compound 6); 3MG7 (compound 8); 3MG8 (compound 16).

calculated relative to DMSO and either 2–20 μM E64 [*trans*-epoxysuccinyl-L-leucylamido(4-guanadino)butane; Sigma] for the cysteine proteases (cathepsins and calpain) or complete EDTA-free protease inhibitor cocktail (1 tablet/ml diluted 1:10 (v/v) in assay buffer, Roche) for the remaining serine proteases.

Protein turnover assay

Bulk protein turnover was measured in HCT-116 cells as described previously [1]. Briefly, HCT116 cells (100 000 cells/well in 12-well plates) were labelled for 20 min under cell culture incubation conditions in methionine- and cysteine-free DMEM (Dulbecco's modified Eagle's medium; Invitrogen), supplemented with 10 % dialysed FBS (fetal bovine serum), 2 mM L-glutamine and 50 μCi/well of EasyTag™ EXPESS ^{35}S-labelled methionine/cysteine mix (NEN Radiochemicals, PerkinElmer). The cells were then washed three times with the same medium containing the unlabelled 2 mM methionine and 0.2 mM cysteine, and then incubated with 3 μM compound in this medium for the indicated times. Medium (50 μl) was removed from the cells and subjected to liquid scintillation counting to monitor the release of ^{35}S-labelled peptides. At the end of the experiment, cells were washed in PBS and then solubilized in 1 ml of 0.2 M NaOH and subjected to liquid scintillation counting. Percentage protein turnover at each time point was calculated as: [(total counts released into medium)/(total counts released into medium + total counts in NaOH-solubilized cells)] × 100.

Synthesis of compound 1, the capped tri-peptide 5-methyl-*N*-((*S*)-1-((*S*)-1-((*S*)-1-(4-methylbenzylamino)-1-oxo-4-phenylbutan-2-ylamino)-1-oxo-4-phenylbutan-2-ylamino)-1-oxo-4-phenylbutan-2-yl)pyrazine-2-carboxamide

To a mixture of (*S*)-2-(tert-butoxycarbonylamino)-4-phenylbutanoic acid (0.634 g, 2.26 mmol) and 4-methylbenzylamine (0.306 ml, 2.40 mmol) in dichloromethane (10.0 ml) was added HBTU (*O*-benzotriazole-*N*,*N*,*N*′,*N*′-tetramethyluronium-hexafluorophosphate) (0.984g, 2.6 mmol) and *N*-methylmorpholine (0.50 ml, 4.54 mmol). The reaction mixture was stirred at room temperature overnight and then washed with saturated aqueous $NaHCO_3$. The organic layer was dried over anhydrous $MgSO_4$ and evaporated and the residue dissolved in tetrahydrofuran (20 ml), treated with 4.0 M HCl in 1,4-dioxane (20 ml) and stirred for 16 h at ambient temperature. The solvents were removed *in vacuo* to give (*S*)-2-amino-*N*-(4-methylbenzyl)-4-phenylbutanamide hydrochloride as an oil (0.70 g, 98 %). ESI (electrospray ionization)–MS +: calculated for $C_{18}H_{22}N_2O$ [M + H] 283; found 283.

A solution of (*S*)-2-amino-*N*-(4-methylbenzyl)-4-phenylbutanamide hydrochloride (0.50 g, 1.57 mmol) and (*S*)-2-(tert-butoxycarbonylamino)-4-phenylbutanoic acid (0.439 g, 1.57 mmol) and 2,4,6-collidine (0.571 mg, 4.71 mmol) in DMF (10 ml) was treated with TBTU [*O*-(benzotriazol-1-yl)-*N*,*N*,*N*′,*N*′-tetramethyluronium tetrafluoroborate] (0.556 g, 1.73 mmol) and stirred at ambient temperature for 3 h. The reaction mixture was diluted with water and extracted with dichloromethane. The extracts were dried over $MgSO_4$ and evaporated. The residue was re-dissolved in dichloromethane containing 1 % MeOH and hexane was added to precipitate tert-butyl (*S*)-1-[(*S*)-1-(4-methylbenzylamino)-1-oxo-4-phenylbutan-2-ylamino]-1-oxo-4-phenylbutan-2-ylcarbamate (0.546 g, 90 %).

A solution of tert-butyl (*S*)-1-[(*S*)-1-(4-methylbenzylamino)-1-oxo-4-phenylbutan-2-ylamino]-1-oxo-4-phenylbutan-2-ylcarbamate (0.754 g, 1.4 mmol) in dichloromethane (10 ml) was treated with 4 M HCl in dioxane (2.1 ml, 10.4 mmol) and stirred at room temperature for 4 h. The solvent was evaporated and the residue triturated with ether to give the corresponding amine hydrochloride (0.65 g, 100 %) which was used without further purification. The next *S*-homo-Phe residue was coupled as described for the previous step and the Boc protecting group removed as described above to give (*S*)-2-amino-*N*-{(*S*)-1-[(*S*)-1-(4-methylbenzylamino)-1-oxo-4-phenylbutan-2-ylamino]-1-oxo-4-phenylbutan-2-yl}-4-phenylbutanamide as the hydrochloride salt (81 % yield). The amine described above was coupled to 5-methylpyrazine-2-carboxylic acid using the coupling conditions described above and worked up in the usual way. The crude product was purified by flash chromatography

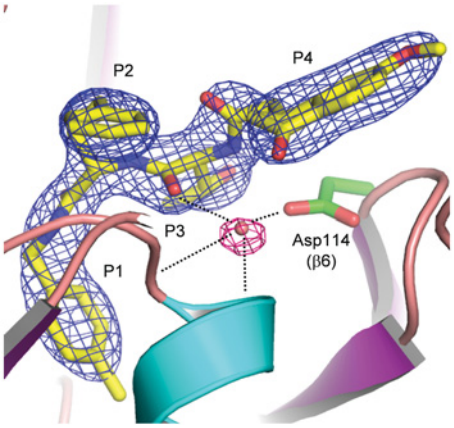


Figure S2 Electron density of compound 6 and the water molecule involved in peptide binding

A 2.6 Å sigmaA-weighted $F_o - F_c$ electron-density omit map contoured at 2.5σ covering compound 6 and an associated water molecule in the β5/β6 active site of 20S is shown, with positions P1–P4 indicated. After removal of all ligand atoms and water molecules from the model, random co-ordinate displacements were added to remove model bias prior to several cycles of refinement and omit map calculation. For clarity, the portion of the omit map covering the water molecule is coloured magenta and hydrogen bonds are indicated by dashed lines. Atoms are shown in ball-and-stick representation with carbon in yellow, nitrogen in blue, oxygen in red, and sulfur in orange. The side-chain atoms for Asp^{114} (β6) are shown in ball-and-stick representation coloured as above, except for carbon in green. Protein atoms from 20S are shown in cartoon representation coloured by secondary structure with helices in cyan, β-strands in magenta and loops in salmon. The figure was made using the PyMOL Molecular Graphics System (DeLano Scientific, Palo Alto, CA, U.S.A.).

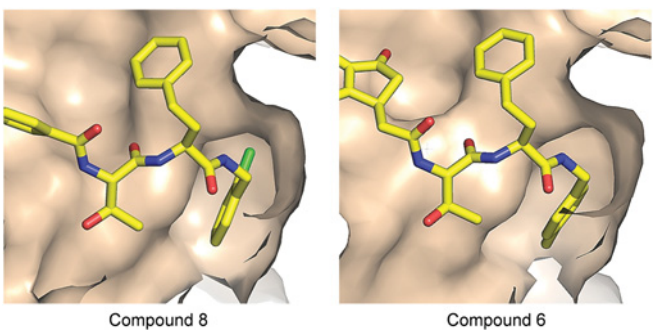


Figure S3 Comparison of the structures of compounds 8 and 6 bound to the chymotrypsin-like site of the 20S proteasome

Close-up view of the S1 sub-pocket being filled by the ortho-chloro substituent in P1 for compound 8 (left-hand panel), compared with compound 6 that lacks this substituent (right-handpanel). The colour scheme is as described in Figure S2.

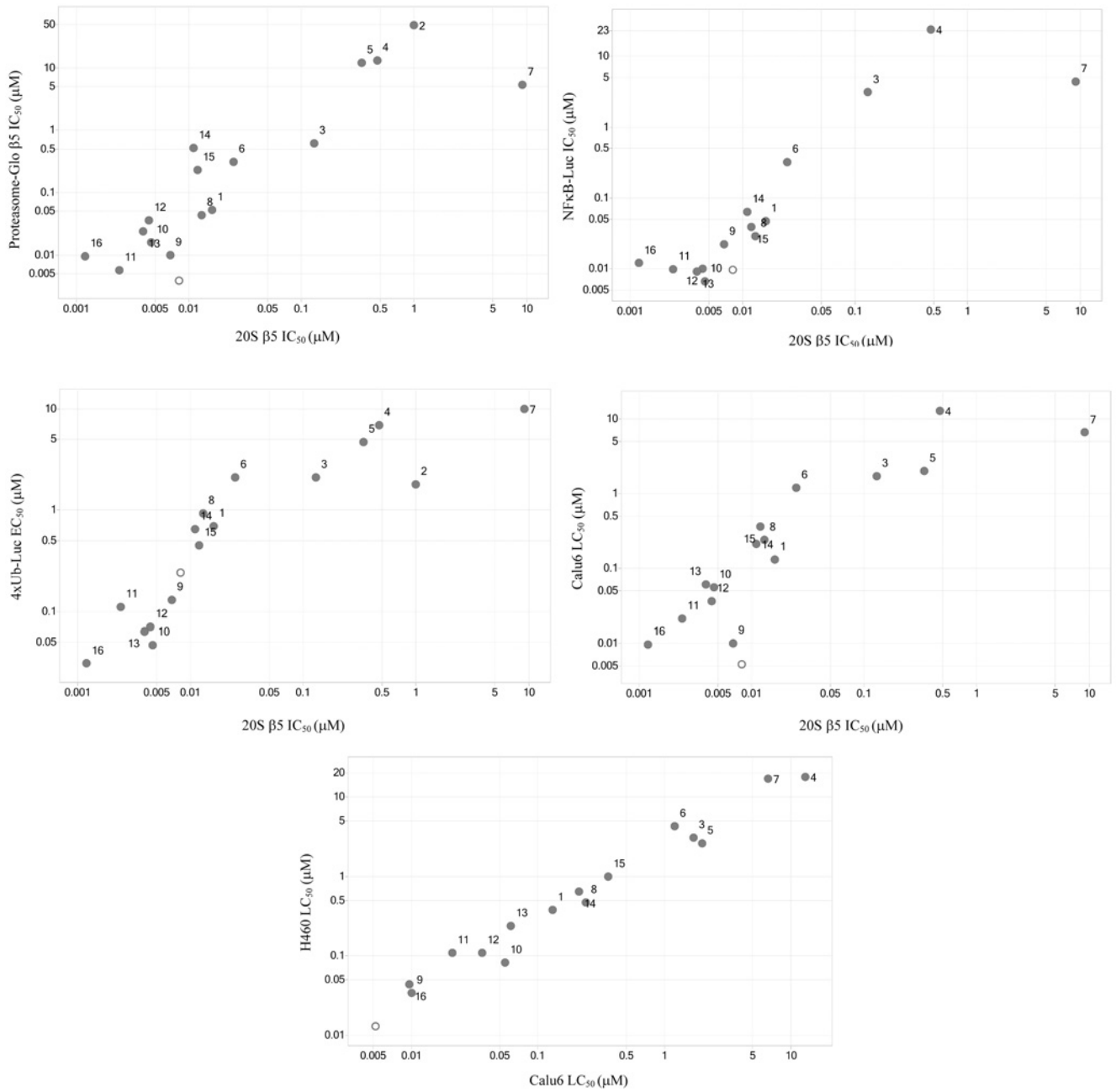


Figure S4 Correlation between *in vitro* and cellular potencies of bortezomib and compounds 1–16

Data presented in Tables 1 and 3 in the main paper are expressed in μM concentrations and compared to show the correlation between *in vitro* potency of the compounds against the β5 site of the constitutive 20S proteasome determined with Ac-WLA-AMC as substrate and Calu6 cell Proteasome-Glo™ β5 IC_{50}, HEK293-NFκB-Luc IC_{50}, MDA-MD-231-4xUb-Luc EC_{50} and Calu6 cell ATPlite viability LC_{50}. The viability effects of the compounds in Calu6 cells and H460 cells, as shown in Table 3 in the main paper, shown as LC_{50} values are expressed in μM concentrations and correlated. The open symbol represents bortezomib and the closed symbols are numbered to indicate compounds 1–16.

on silica gel eluting with a gradient of 0 to 5 % methanol in dichloromethane to give 5-methyl-*N*-[(*S*)-1-{(*S*)-1-[(*S*)-1-(4-methylbenzylamino)-1-oxo-4-phenylbutan-2-ylamino]-1-oxo-4-phenylbutan-2-ylamino}-1-oxo-4-phenylbutan-2-yl]pyrazine-2-carboxamide (compound 1) in 73 % yield.

Compound 1

^{1}H NMR (300 MHz, d6 DMSO) δ: 1.91 (m, 5H), 2.10 (m, 2H), 2.21 (s, 3H), 2.59 (m, 8H), 3.31 (s, 1H), 4.19 (m, 2H), 4.25 (m, 1H), 4.66 (m, 1H), 7.01 (d, 2H, $J = 7.8$ Hz), 7.17 (m, 17H), 8.21

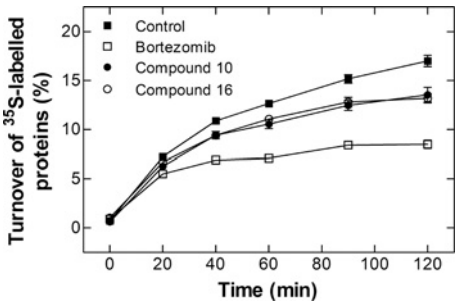


Figure S5 Effect on intracellular protein turnover of potent di-peptide proteasome inhibitors 11 and 16 compared with bortezomib

HCT116 cells were metabolically labelled for 20 min with [^{35}S]methionine/cysteine and then chased with excess unlabelled amino acids in the presence of compound 11, 16 or bortezomib (each at 3 μM) for the times indicated. Intracellular degradation of ^{35}S-labelled proteins was then determined as described in the Experimental section in the main paper and expressed as a percentage of total label incorporated into the proteins. The results are means $\pm$ S.E.M. ($n = 6$) and are representative of two similarly performed experiments.

(d, 1H, $J = 7.9$ Hz), 8.34 (d, 1H, $J = 7.1$ Hz), 8.45 (d, 1H, $J = 7.8$ Hz), 8.65 (m, 1H), 8.81 (d, 1H, $J = 8.3$ Hz), and 9.05 (d, 1H, $J = 11.6$ Hz). HRMS (high-resolution MS; ESI+): calculated for $C_{44}H_{48}N_6O_4$ [M + H] 725.3815; found 725.3799.

Synthesis of (*S*)-2-(tert-butoxycarbonylamino)-4-(neopentylamino)-4-oxobutanoic acid (compound 18)

Compound 18 was prepared according to Scheme S1. A solution of Boc-Asp-OBn (0.937g, 2.9 mmol), neopentylamine (0.267 g, 2.9 mmol), HBTU (*O*-benzotriazole-*N*,*N*,*N*′,*N*′-tetramethyl-uronium-hexafluorophosphate) (1.33 g, 3.5 mmol) and *N*-methylmorpholine (0.44 g, 4.4 mmol) in anhydrous methylene chloride (15.0 ml) was stirred at ambient temperature for 16 h. The solvent was evaporated and the residue obtained was partitioned between ethyl acetate (35 ml) and saturated aqueous $NaHCO_3$ (25 ml). The organic layer was dried over anhydrous $MgSO_4$, concentrated and purified by flash chromatography on silica gel to afford (*S*)-benzyl 2-(tert-butoxycarbonylamino)-4-(neopentylamino)-4-oxobutanoate (17) as a white solid (0.84 g, 96 %) which was dissolved in methanol (3 ml) and stirred under 20 psi (1 psi = 6.9 kPa) H_2 for 2 h in the presence of 10 % palladium on carbon (10 mol%). The Pd/C was removed via filtration through celite and the solvent evaporated to afford 18 in quantitative yield.

Compound 18

1H NMR (400 MHz, C^2HCl_3) δ 6.59 (s, 1H), 5.90 (d, $J = 5.2$ Hz, 1H), 4.41–4.27 (m, 1H), 3.48 (s, 1H), 3.22 (dd, $J = 13.3$, 7.1 Hz, 1H), 3.02–2.94 (m, 2H), 2.76 (dd, $J = 15.6$, 9.0 Hz, 1H), 1.44 (s, 9H), 0.91 (s, 9H). MS (ESI+): calculated for $C_{14}H_{26}N_2O_5$ [M + H] 303; found 303.

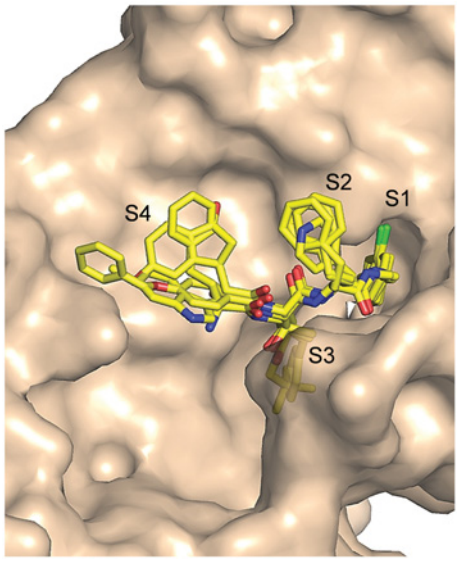
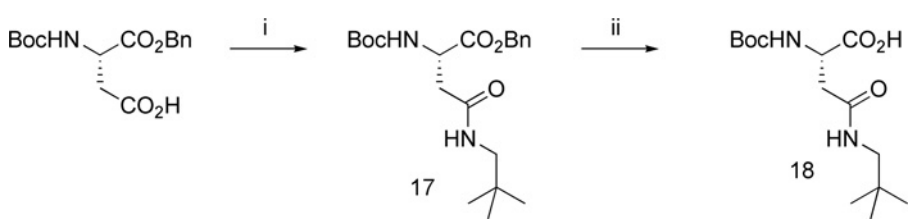


Figure S6 Common occupancy of the chymotrypsin-like site of the 20S proteasome by compounds 1, 6, 8 and 16

Superposition of the structures of compounds 1, 6, 8, and 16 bound to the $\beta5/\beta6$ active site of 20S, showing the molecular surface of 20S with positions P1–P4 indicated. The colour scheme is as described in Figure S2.

Synthesis of capped di-peptide proteasome inhibitors

A representative procedure for the parallel synthesis of compound arrays (denoted IV in Scheme 2 in the main paper) follows: a deep-well polypropylene synthesis plate was charged with the appropriate amine salts II (0.25 mmol) in one dimension and a solution of the amino acid (0.25 mmol) constituting the P3 coupling partner in DMF (2.0 ml) added in the second dimension by use of a Tecan liquid handler. HBTU (0.5 ml, 0.5 M solution in DMF) and *N*-methylmorpholine (0.9 mmol) were then added to each well by means of the liquid handler. The plate was shaken at room temperature for 16 h. The residue in each well obtained after evaporation of the solvent in a Genevac was partitioned between $CHCl_3$/THF (3:1) (3 ml per well) and saturated aqueous $NaHCO_3$ (2 ml) with each solution delivered by the liquid handler. The phases were separated by centrifugation and the aqueous layer was extracted with additional $CHCl_3$/THF (3:1) (3 ml) and the combined organic layers were evaporated. The residue thus obtained was treated with 4 M HCl in dioxane (2 ml), shaken at ambient temperature for 5 h and concentrated in a Genevac evaporator to remove solvent and residual HCl. The N-terminal cap was then coupled and the final products isolated as in the



Scheme S1 Synthesis of compound 18

(i) Neopentyl amine, HBTU, NMM, DCM, 25 °C, 24h; (ii) H_2, Pd/C, MeOH.

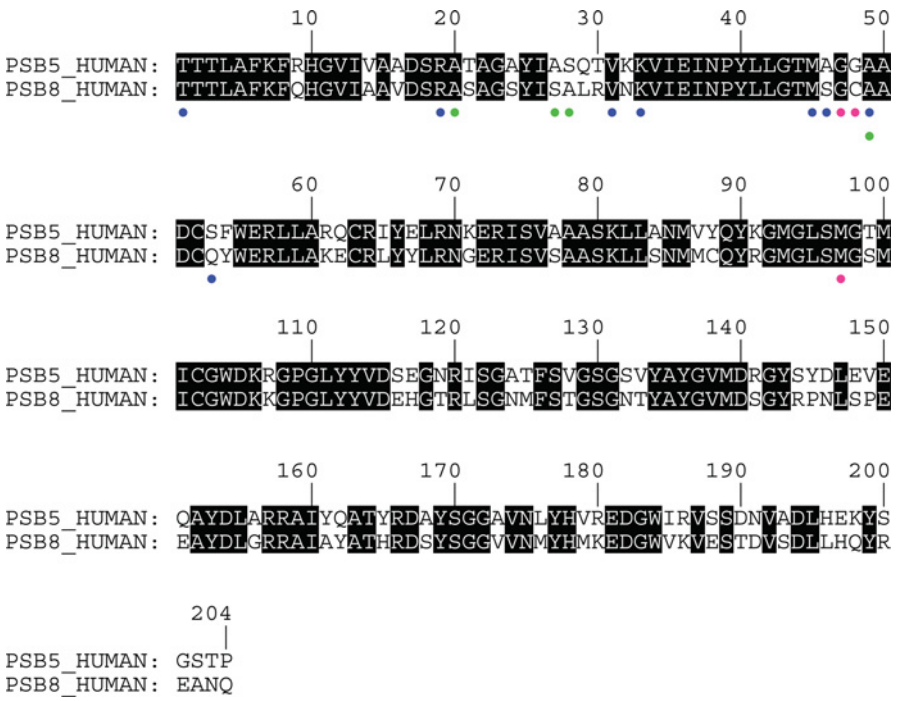


Figure S7 Sequence alignment of the human β5 subunits of the constitutive and immunoproteasome

PSB5_HUMAN (PSMB5, β5 constitutive subunit, upper sequence) is compared with PSB8_HUMAN (PSMB8, β5 immunoproteasome subunit, lower sequence), with regions of sequence identity shaded black. Residues forming the inhibitor-binding pockets are indicated by circles below the aligned sequences and coloured as follows: S1 pocket in blue, S2 pocket in magenta and S3 pocket in green. Note that Ala[49] sits between the S1 and S3 pockets and has been assigned to both. Residues from the β6 subunit that form the S4 pocket, as well as a portion of the S3 pocket, are not shown because this subunit does not differ between the constitutive and immunoproteasome.

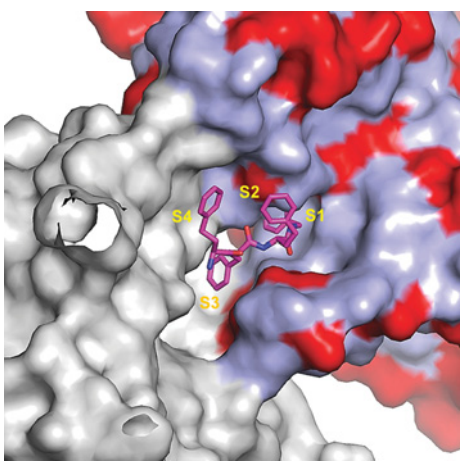


Figure S8 Residue conservation between the β5/β6 active site of the constitutive and immunoproteasome

Surface representation of the β5 (light blue) and β6 (grey) subunits of bovine 20S proteasome (PDB entry 1IRU [56]). Residues that are non-identical between constitutive and immunoproteasome are coloured red. The structure of compound 1 has been modelled into the active site on the basis of a superposition of the β5 subunit structure from yeast 20S, reported in the present study, with the corresponding subunit of the bovine 20S structure.

previous step. All final compounds IV (some 1500 in total) were purified by preparative reverse phase HPLC using mass-directed fraction collection and their purities and identities established by LC-MS. Preparative HPLC was conducted on an Agilent 1100 series LC/MSD instrument using a Waters SunFire C18 5 μm Prep OBD column (19 mm$\times$150 mm). The compounds were eluted with a water–MeCN gradient (0.1 % formic acid) optimized by the A2Prep Agilent software. Compounds 2–16 described herein were further characterized by 1H NMR spectroscopy and HRMS (see below).

HRMS

Molecular mass measurements were performed using a QSTAR® XL quadrupole TOF (time-of-flight) mass spectrometer (Applied Biosystems) coupled to an 1100 series HPLC system (Agilent Technologies). An isocratic flow of 50 % solvent A (20 % CH_3CN/80 % H_2O with 0.1 % CH_3COOH) and 50 % solvent B (100 % CH_3CN with 0.1 % CH_3COOH) at 200 μl/min was used to deliver each sample to the electrospray source of the mass spectrometer, tuned in ESI TOF-MS positive mode with a 3 min acquisition time for each analysis. The mass spectrometer was externally calibrated between m/z 400 and 800 by using API Calibration (NaCsI) Solution (Waters Corp). For each analysis, approx. 5 scans were summed and the centroid m/z value of the protonated monoisotopic molecular ion $[M+H]^+$ was recorded. The relative deviations from the calculated exact masses based on the expected formulae were also calculated. For each compound, the experimental and calculated $[M+H]$ values were within ± 4 ppm, supporting their expected elemental composition (see the Supplementary Results section).

SUPPLEMENTARY RESULTS

1H NMR and HRMS data for synthetic compounds

Compound 2

1H NMR (400 MHz, d6 DMSO) δ: 1.82 (m, 2H), 1.92 (m, 2H), 2.22 (s, 3H), 2.56 (m, 2H), 3.45 (d, 1H), 3.55 (d, 1H), 4.20 (d,

Table S1 Crystallographic data and refinement statistics

rmsd, root mean square deviation.

	Compound				
	Bortezomib	1	6	8	16
PDB ID	3MG0	3MG4	3MG6	3MG7	3MG8
Space group	$P2_1$	$P2_1$	$P2_1$	$P2_1$	$P2_1$
Unit cell dimensions (Å, °)	a=136.1, b=300.5, c=145.8, β=113.2	134.6, 300.8, 144.7, 112.7	137.3, 299.6, 145.8, 113.7	136.0, 300.4, 145.8, 113.7	136.3, 299.2, 145.7, 113.1
Wavelength (Å)	0.97947	0.97931	0.97958	0.97958	0.97969
Resolution (Å)	50 - 2.7	50 - 3.1	50 - 2.6	50 - 2.8	50 - 2.6
R_{sym}* (%)	9.8 (32.8)†	11.6 (30.5)	9.2 (35.3)	12.3 (44.2)	7.3 (32.7)
Total observations	1017555	723147	2255995	950351	1034821
Unique reflections	276539	177491	315914	257519	315407
Average redundancy	3.7	4.1	7.1	3.7	3.3
$<I/\sigma>$	9.7 (2.5)	7.1 (2.8)	11.1 (2.7)	8.0 (2.2)	11.1 (2.9)
Completeness (%)	92.9 (71.6)	94.1 (61.2)	96.3 (76.9)	97.2 (85.7)	95.6 (85.9)
Refinement resolution (Å)	50 - 2.7	50 - 3.1	50 - 2.6	50 - 2.8	50 - 2.6
Reflections (working/test)	272720/2818	173803/3565	310987/3189	251281/5191	308947/6368
R_{cryst}/R_{free} (%)‡	22.6/25.3	20.5/25.2	21.5/24.9	21.7/25.9	22.4/25.3
Protein atoms	49541	49541	49298	49298	49298
Ligand atoms	168	90	86	88	92
Mes atoms	0	24	24	24	24
Mg^{2+} atoms	0	10	10	10	10
Water atoms	1037	0	858	4	5
rmsd bond lengths (Å)	0.008	0.008	0.011	0.008	0.008
rmsd bond angles (°)	1.11	1.13	1.31	1.14	1.12
Ramachandran analysis					
Most-favoured (%/number)	88.5/4896	89.8/4969	90.4/4997	90.5/4993	91.3/5041
Additional allowable (%/number)	10.7/594	9.5/526	8.9/491	8.6/475	8.1/445
Generously allowed (%/number)	0.5/29	0.5/27	0.4/23	0.6/31	0.4/21
Disallowed (%/number)	0.3/14	0.2/13	0.3/16	0.3/16	0.2/13

*$R_{sym} = (\Sigma_{hkl}\Sigma_i|I_i(hkl) - <I(hkl)>|)/\Sigma_{hkl}\Sigma_i I_i(hkl)$ for n independent reflections and i observations of a given reflection. $<I(hkl)>$ is the average intensity of the i^{th} observation.
†Numbers in parenthesis are for highest resolution shell.
‡$R_{cryst} = \Sigma_h||F_o(h)| - |F_c(h)||/\Sigma_h|F_o(h)|$, where F_o and F_c are the observed and calculated structure factors respectively.

2H), 4.28 (m, 3H), 6.84 (m, 1H), 6.96 (m, 2H), 7.01 (m, 3H), 7.10 (m, 8H), 7.17 (m, 2H), 7.25 (m, 5H), 7.33 (m, 3H), 8.16 (d, 1H, $J = 6.7$ Hz), 8.32 (t, 1H, $J = 6.1$ Hz), and 8.40 (d, 1H, $J = 8$ Hz). HRMS (ESI+): calculated for $C_{42}H_{43}N_3O_4$ [M + H] 654.3332; found 654.3346.

Compound 3

1H NMR (300 MHz, d6 DMSO) δ: 2.25 (s, 3H), 2.67 (m, 1H), 2.90 (m, 1H), 3.01 (m, 1H), 3.12 (m, 1H), 3.42 (m, 2H), 3.65 (s, 3H), 3.67 (s, 3H), 4.18 (m, 2H), 4.51 (m, 1H), 4.57 (m, 1H), 6.65 (m, 1H), 6.71 (m, 1H), 6.81 (m, 4H), 6.97 (m, 5H), 7.05 (d, 3H, $J = 8.3$ Hz), 7.12 (m, 2H), 7.20 (t, 2H, $J = 7.8$ Hz), 7.35 (m, 4H), 7.59 (d, 1H, $J = 7.7$ Hz), 8.19 (d,d, 1H, $J = 14.0$, 7.9 Hz), and 8.33 (m, 1H). HRMS (ESI+): calculated for $C_{44}H_{44}N_4O_6$ [M + H] 725.3339; found 725.3336.

Compound 4

1H NMR (400 MHz, d6 DMSO) δ: 0.99 (d, 3H, $J = 6.3$ Hz), 1.83 (m, 1H), 1.98 (m, 1H), 2.25 (s, 3H), 2.57 (m, 2H), 3.56 (d, 1H, $J = 14.1$ Hz), 3.60 (d, 1H, $J = 13.8$ Hz), 4.00 (m, 1H), 4.22 (m, 4H), 5.05 (d, 1H, $J = 5.1$ Hz), 6.84 (m, 1H), 6.98 (m, 3H), 7.12 (m, 9H), 7.27 (m, 3H), 7.36 (m, 2H), 8.03 (d, 1H, $J = 8.1$ Hz), 8.10 (d, 1H, $J = 8.3$ Hz), and 8.36 (t, 1H, $J = 5.9$ Hz). HRMS (ESI+): calculated for $C_{33}H_{39}N_3O_5$ [M + H] 594.2968; found 594.2981.

Compound 5

1H NMR (400 MHz, DMSO) δ 8.32 (t, $J = 6.0$ Hz, 1H), 8.15 (d, $J = 8.0$ Hz, 1H), 7.98 (d, $J = 8.0$ Hz, 1H), 7.86 (d, $J = 8.9$ Hz, 2H), 7.43–7.36 (m, 4H), 7.24 (t, $J = 7.3$ Hz, 2H), 7.19–7.03 (m, 9H), 5.14 (s, 2H), 5.02 (d, $J = 6.5$ Hz, 1H), 4.42 (dd, $J = 8.0$, 4.9 Hz, 1H), 4.30 (td, $J = 8.7$, 4.8 Hz, 1H), 4.20 (dt, $J = 12.5$, 6.2 Hz, 2H), 4.11 (dd, $J = 11.5$, 6.3 Hz, 1H), 2.65–2.52 (m, 2H), 2.25 (s, 3H), 2.03–1.93 (m, 1H), 1.92–1.75 (m, 1H), 1.27 (s, 9H), 1.12 (d, $J = 6.3$ Hz, 3H). HRMS (ESI+): calculated for $C_{40}H_{47}N_3O_5$ [M + H] 650.3596; found 650.3594.

Compound 6

1H NMR (400 MHz, DMSO) δ 8.38 (dt, $J = 11.9$, 6.0 Hz, 1H), 8.03 (dd, $J = 8.1$, 3.8 Hz, 2H), 7.53 (dd, $J = 8.5$, 2.6 Hz, 1H), 7.26 (dd, $J = 10.1$, 4.8 Hz, 2H), 7.21–7.04 (m, 8H), 6.96 (dd, $J = 8.5$, 2.0 Hz, 1H), 5.03 (s, 1H), 4.34 (dd, $J = 8.1$, 4.5 Hz, 1H), 4.29 (td, $J = 8.6$, 4.7 Hz, 1H), 4.22 (dd, $J = 5.5$, 2.8 Hz, 2H), 4.03 (dd, $J = 6.3$, 4.7 Hz, 1H), 3.84 (d, $J = 2.2$ Hz, 3H), 3.69–3.59 (m, 1H), 2.85 (dd, $J = 14.3$, 5.1 Hz, 1H), 2.76–2.67 (m, 1H), 2.65–2.51 (m, 2H), 2.46–2.35 (m, 2H), 2.24 (s, 3H), 2.06–1.93 (m, 1H), 1.90–1.78 (m, 1H), 1.06 (d, $J = 6.3$ Hz, 2H), 1.00 (d, $J = 6.3$ Hz, 1H). HRMS (ESI+): calculated for $C_{34}H_{39}N_3O_6$ [M + H] 586.2917; found 586.2931.

Table S2 Protease selectivity of compounds

Protease selectivity of compounds was assessed using fluorigenic peptide-AMC substrates and a panel of purified human enzymes, as described in the Supplementary Experimental section. Percentage inhibition of the proteases was determined in the presence of 100 μM compound and calculated relative to DMSO and the inhibitor control E64 or complete EDTA-free protease inhibitor cocktail. Results are means obtained from two independent experiments, each performed in duplicate.

	Enzyme inhibition at 100 μM compound (%)									
Compound	Cathepsin B	Cathepsin L	Coagulation Factor β-XIIa	Chymotrypsin	Elastase	Plasmin	Thrombin	Tissue plasminogen activator	Trypsin	Calpain I
1	−15	12	45	−4	−8	−20	22	21	12	−9
2	15	5	37	−3	−6	−12	12	14	0	17
3	1	22	34	1	−7	−15	21	16	−1	−12
4	−3	16	38	1	−3	−9	17	19	3	29
5	20	4	36	−8	−11	−18	12	12	−2	5
6	3	3	26	84	1	−11	12	12	−6	8
7	20	−9	31	3	0	−12	21	14	2	−11
8	−3	33	36	9	−1	−11	29	18	1	55
9	−14	3	34	−4	−6	−13	13	13	3	25
10	−13	18	32	2	−5	−18	13	13	−1	22
11	17	−6	40	36	−4	5	30	26	15	−40
12	−7	21	38	2	12	−1	15	17	2	7
13	−6	28	33	0	1	−10	12	17	−2	50
14	−8	0	35	−10	−10	−18	9	8	−4	15
15	−6	7	42	−4	2	−15	15	12	1	32
16	10	−3	49	23	14	11	34	34	20	6

Table S3 Anti-proliferative effects of compound 1 and bortezomib in cancer cell lines

The indicated cell lines (1000–2000 cells per well) were incubated with a range of concentrations of compound 1 or bortezomib for 72 h under cell culture conditions. Cell viability was assessed my monitoring cellular ATP levels using the luminescence-based ATPlite assay as described in the Experimental section of the main paper. IC_{50} data are means $\pm$ S.E.M. for 3–11 determinations.

	ATPlite cell viability LC_{50} (nM)	
Cell line	Compound 1	Bortezomib
MDA-MB-231	150 $\pm$ 23	8.6 $\pm$ 1.0
Calu6	130 $\pm$ 10	5.2 $\pm$ 0.76
H460	380 $\pm$ 110	13 $\pm$ 2.0
HCT116	50 $\pm$ 8.1	4.6 $\pm$ 0.68
HT29	55 $\pm$ 8.9	5.4 $\pm$ 1.1

Compound 7

1H NMR (400 MHz, d6 DMSO) δ: 1.12 (d, 3H, $J = 6.3$ Hz), 1.90 (m, 2H), 2.60 (m, 2H), 2.75 (s, 1H), 2.84 (s, 2H), 4.09 (m, 1H), 4.48 (m, 3H), 4.78 (d, 1H, $J = 6.4$ Hz), 4.91 (m, 1H), 5.19 (s, 2H), 7.20 (m, 11H), 7.33 (t, 2H, $J = 7.4$ Hz), 7.40 (t, 2H, $J = 6.1$ Hz), 7.47 (d, 2H, $J = 7$ Hz), 7.89 (m, 2H), 7.97 (m, 1H), and 8.29 (m, 1H). HRMS (ESI+): calculated for $C_{36}H_{39}N_3O_5$ [M + H] 594.2968; found 594.2967.

Compound 8

1H NMR (300 MHz, d6 DMSO) δ: 1.12 (d, 3H, $J = 6.3$ Hz), 1.89 (m, 1H), 2.01 (m, 1H), 2.63 (m, 2H), 3.31 (s, 1H), 4.13 (m, 1H), 4.38 (m, 2H), 4.43 (m, 1H), 5.00 (d, 1H, $J = 6.5$ Hz), 5.19 (s, 2H), 7.10 (d, 2H, $J = 9.1$ Hz), 7.16 (m, 3H), 7.36 (m, 11H), 7.87 (d, 2H, $J = 8.98$ Hz), 7.99 (d, 1H, $J = 8$ Hz), 8.22 (d, 1H, $J = 8$ Hz), and 8.39 (t, 1H, $J = 5.9$ Hz). HRMS (ESI+): calculated for $C_{35}H_{36}ClN_3O_5$ [M + H] 614.2422; found 614.2404.

Compound 9

1H NMR (400 MHz, d6 DMSO) δ: 0.78 (s, 9H), 1.83 (m, 1H), 2.01 (m, 1H), 2.25 (s, 3H), 2.72 (m, 5H), 4.21 (m, 3H), 4.76 (d, 1H), 5.17 (s, 2H), 7.16 (m, 11H), 7.39 (m, 6H), 7.84 (m, 3H), 8.21 (d, 1H, $J = 8.1$ Hz), 8.42 (m, 1H), and 8.53 (d, 1H). HRMS (ESI+): calculated for $C_{41}H_{48}N_4O_5$ [M + H] 677.3703; found 677.3729.

Compound 10

1H NMR (400 MHz, d6 DMSO) δ: 0.78 (s, 9H), 1.82 (m, 1H), 2.03 (m, 1H), 2.24 (s, 3H), 2.44 (m, 3H), 2.63 (m, 3H), 2.80 (m, 4H), 4.14 (m, 2H), 4.25 (m, 1H), 4.60 (m, 1H), 6.98 (m, 3H), 7.11 (m, 7H), 7.25 (m, 3H), 7.85 (t, 1H, $J = 6.1$ Hz), 8.22 (d,d, 2H, $J = 5.4$, 7.7 Hz), and 8.45 (t, 1H, $J = 6.1$ Hz). HRMS (ESI+): calculated for $C_{36}H_{45}FN_4O_4$ [M + H] 617.3503; found 617.3527.

Compound 11

1H NMR (400 MHz, d6 DMSO) δ: 0.75 (s, 9H), 1.25 (d, 3H, $J = 7.4$ Hz), 2.71 (m, 2H), 2.80 (d, 2H, $J = 6.2$ Hz), 4.27 (m, 3H), 4.67 (m, 1H), 7.02 (m, 1H), 7.19 (m, 1H), 7.26 (m, 2H), 7.35 (m, 1H), 7.53 (m, 1H), 7.89 (t, 1H, $J = 6.3$ Hz), 8.22 (m, 1H), 8.32 (d, 1H, $J = 7.3$ Hz), 8.46 (t, 1H, $J = 6.1$ Hz), 8.71 (d, 1H, $J = 8.1$ Hz), 8.75 (s, 1H), and 12.25 (s, 1H). HRMS (ESI+): calculated for $C_{29}H_{33}F_2N_5O_5$ [M + H] 570.2528; found 570.2510.

Compound 12

1H NMR (400 MHz, d6 DMSO) δ: 0.75 (s, 9H), 1.27 (d, 3H, $J = 6.9$ Hz), 2.41 (m, 3H), 2.71 (m, 5H), 4.24 (t, 1H, $J = 7.4$ Hz), 4.32 (m, 2H), 4.58 (d, 1H, $J = 6.8$ Hz), 7.18 (m, 3H), 7.27 (m, 5H), 7.41 (m, 1H), 7.84 (t, 1H, $J = 5.7$ Hz), 8.14 (d, 1H, $J = 8.1$ Hz), 8.33 (d, 1H, $J = 6.9$ Hz), and 8.51 (t, 1H, $J = 5.3$ Hz). HRMS (ESI+): calculated for $C_{28}H_{37}ClN_4O_4$ [M + H] 529.2582; found 529.2602.

Compound 13

1H NMR (400 MHz, d6 DMSO) δ: 0.75 (s, 9H), 2.36 (m, 4H), 2.56 (m, 1H), 2.74 (m, 3H), 2.89 (m, 1H), 3.18 (m, 1H), 4.35 (m, 2H), 4.54 (m, 2H), 7.17 (d, 3H, $J = 6.5$ Hz), 7.28 (m, 6H), 7.42 (m, 1H), 7.72 (d, 1H, $J = 7.9$ Hz), 7.84 (t, 1H, $J = 3.9$ Hz), 8.08 (d, 1H, $J = 7.7$ Hz), 8.36 (d, 1H, $J = 7.9$ Hz), 8.44 (m, 2H), and 8.60 (t, 1H, $J = 5.4$ Hz). HRMS (ESI+): calculated for $C_{33}H_{40}ClN_5O_4$ [M + H] 606.2847; found 606.2866.

Compound 14

1H NMR (400 MHz, DMSO) δ 8.20 (d, $J = 7.8$ Hz, 1H), 8.08 (d, $J = 7.9$ Hz, 1H), 7.88 (dd, $J = 13.3$, 7.2 Hz, 1H), 7.77 (d, $J = 7.8$ Hz, 1H), 7.22 (dd, $J = 13.2$, 6.5 Hz, 4H), 7.13 (t, $J = 7.3$ Hz, 6H), 4.55 (dd, $J = 14.2$, 7.5 Hz, 1H), 4.05 (s, 1H), 3.80 (dd, $J = 13.9$, 6.5 Hz, 2H), 2.89–2.83 (m, 2H), 2.80–2.73 (m, 2H), 2.70 (t, $J = 2.7$ Hz, 2H), 2.44–2.36 (m, 3H), 1.97 (s, 1H), 1.74 (s, 1H), 1.02 (dd, $J = 18.7$, 6.6 Hz, 6H), 0.80 (s, 9H). HRMS (ESI+): calculated for $C_{31}H_{44}N_4O_4$ [M + H] 537.3441; found 537.3442.

Compound 15

1H NMR (400 MHz, DMSO) δ 8.20 (dd, $J = 7.9$, 3.2 Hz, 2H), 8.11 (t, $J = 5.7$ Hz, 1H), 7.87 (t, $J = 6.2$ Hz, 1H), 7.33–7.07 (m, 10H), 5.77 (ddt, $J = 17.2$, 10.3, 5.1 Hz, 1H), 5.10 (dq, $J = 17.2$, 1.7 Hz, 1H), 5.03–4.98 (m, 1H), 4.60 (dd, $J = 14.5$, 7.4 Hz, 1H), 4.15–4.07 (m, 1H), 3.67 (dd, $J = 12.1$, 5.2 Hz, 2H), 2.94–2.75 (m, 4H), 2.69–2.57 (m, 2H), 2.56–2.51 (m, 2H), 2.42 (dd, $J = 9.2$, 6.5 Hz, 2H), 2.08–1.96 (m, 1H), 1.90–1.72 (m, 1H), 0.81 (s, 9H). HRMS (ESI+): calculated for $C_{31}H_{42}N_4O_4$ [M + H] 535.2384; found 535.328.

Compound 16

1H NMR (300 MHz, d6 DMSO) δ: 0.74 (s, 9H), 2.66 (m, 2H), 2.79 (m, 2H), 2.91 (m, 1H), 3.10 (m, 1H), 4.34 (d, 2H, $J = 5.8$ Hz), 4.64 (m, 2H), 7.27 (m, 7H), 7.42 (m, 1H), 7.54 (m, 1H), 7.87 (t, 1H, $J = 6.2$ Hz), 8.23 (m, 1H), 8.37 (d, 2H, $J = 5.93$ Hz), 8.44 (d, 1H, $J = 8.24$ Hz), 8.62 (t, 1H, $J = 5.84$ Hz), 8.72 (m, 2H), and 12.26 (m, 1H). HRMS (ESI+): calculated for $C_{34}H_{37}ClN_6O_5$ [M + H] 645.2592; found 645.2586.

REFERENCE

1 Soucy, T. A., Smith, P. G., Milhollen, M. A., Berger, A. J., Gavin, J. M., Adhikari, S., Brownell, J. E., Burke, K. E., Cardin, D. P., Critchley, S. et al. (2009) An inhibitor of NEDD8-activating enzyme as a new approach to treat cancer. Nature **458**, 732–737

Received 15 March 2010/2 July 2010; accepted 15 July 2010
Published as BJ Immediate Publication 15 July 2010, doi:10.1042/BJ20100383