



Universiteit
Leiden
The Netherlands

Chemical tools to monitor and control human proteasome activities

Bruin, G. de

Citation

Bruin, G. de. (2016, June 1). *Chemical tools to monitor and control human proteasome activities*. Retrieved from <https://hdl.handle.net/1887/39834>

Version: Not Applicable (or Unknown)

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/39834>

Note: To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/39834> holds various files of this Leiden University dissertation.

Author: Bruin, G. de

Title: Chemical tools to monitor and control human proteasome activities

Issue Date: 2016-06-01

Systematic analyses of substrate preferences of 20S proteasomes using peptidic epoxyketone inhibitors*

Introduction

The majority of intracellular protein degradation in eukaryotes is mediated by the 20S proteasome core particle (CP). While yeast expresses only one type of proteasome (γ CP), divergent evolution endowed vertebrates with three distinct CPs that differ in their subunit compositions, substrate specificity and biological significance: the constitutive proteasome (cCP, active sites β 1c, β 2c, β 5c), the immunoproteasome (iCP: β 1i, β 2i, β 5i) and the thymoproteasome (tCP: β 1i, β 2i, β 5t).^{1, 2} With their three different proteolytic centers CPs can cleave polypeptides after virtually all amino acids.³ The trypsin-like (T-L) active sites, residing in the β 2 subunits, preferentially cut substrates after positively charged residues. This cleavage specificity is strictly conserved among eukaryotes and between CP types, thus making it highly challenging to design β 2c or β 2i specific inhibitors.⁴ The β 5 active sites exert the most important chymotrypsin-like (ChT-L) activity (processing after hydrophobic residues) by accommodating apolar amino acids in their S1 specificity pocket. In contrast, substrate cleavage preferences between β 1 subunits of γ CP/cCP and iCP significantly differ.^{5, 6} The hydrophilic nature of the $\gamma\beta$ 1 and β 1c active sites promotes the hydrolysis of peptide bonds C-terminally of acidic amino acids (peptidylglutamyl-peptide hydrolyzing (PGPH)⁷ or caspase-like (C-L)⁸ activity), whereas the more hydrophobic lining of the β 1i active site of iCPs is optimized to generate peptides with hydrophobic C-termini for immune surveillance.^{3, 6} These cleavage specificities, which have mostly been investigated by activity assays with natural or synthetic model substrates,^{5, 8-12} served as guidelines for the development of proteasome inhibitors. However, the design of subunit-selective compounds, which represent valuable tools to evaluate the contribution of the individual active sites to antigen processing and to diseases like autoimmune disorders and cancer, requires a more detailed understanding of

* Huber, E.M.; de Bruin, G. *et al.* Systematic analyses of substrate preferences of 20S proteasomes using peptidic epoxyketone inhibitors. *J. Am. Chem. Soc.* **137**, 7835-7842 (2015).

proteasome substrate specificities. This chapter describes the systematic analysis of substrate specificities using 18 synthesized tripeptide α',β' -epoxyketone inhibitors (Table 1),¹³ featuring Leu or Pro in P3, Ala or Leu in P2 and distinct side chains in P1. Binding preferences and IC₅₀ values for the human cCP and iCP subunits were determined by competitive activity based protein profiling (ABPP) in Raji cell lysates and for purified γ CPs by fluorogenic substrate assays. A separate series of five inhibitors was used to assess inhibition of β 1 depending on the presence or absence of a P4 site.⁸ X-ray crystallographic analyses of all compounds in complex with the γ CP together with yeast mutagenesis has been performed and will be discussed partly in this chapter to explain the substrate specificities of the synthesized epoxyketone proteasome inhibitors.

Compound	R	P4	P3	P2	P1	Compound	R	P4	P3	P2	P1
Ac-PAD-EK		-				Ac-LAL-EK		-			
Ac-LAD-EK		-				Ac-PAV-EK		-			
Ac-PAE-EK		-				Ac-LAV-EK		-			
Ac-LAE-EK		-				Ac-PAA-EK		-			
Ac-PAF-EK		-				Ac-LAA-EK		-			
Ac-LAF-EK		-				Ac-PLL-EK		-			
Ac-PAY-EK		-				Ac-LLL-EK		-			
Ac-LAY-EK		-				Ac-APLL-EK					
Ac-PAI-EK		-				N ₃ G-APAL-EK					
Ac-LAI-EK		-				N ₃ G-APnLL-EK					
Ac-PAL-EK		-				N ₃ G-A(4,4-F ₂ P)nLL-EK					

Table 1. Chemical structures of synthesized compounds.

Results and Discussion

Choice of employed organisms, proteasome inhibitor types and assay set-up.

The structures of the synthesized peptide epoxyketone inhibitors are shown in Table 1 (for synthetic details: see experimental section). Apparent IC_{50} values were determined for yeast and human proteasomes in separate experimental setups. Inhibitors were incubated for 1 h with human Raji cell lysate or purified yeast CP. Subsequently, blockage of human CPs was analyzed by determining residual proteasome activity with fluorescently labelled activity based probes^{14, 15} whereas inhibition of the yCP was determined by measuring the residual proteasome activity after the addition of subunit-specific 7-amino-4-methylcoumarin (AMC) substrates (see experimental section).

Compound	Raji lysates (human)						Purified yCP		
	$\beta 1c$	$\beta 1i$	$\beta 2c$	$\beta 2i$	$\beta 5c$	$\beta 5i$	$\gamma \beta 1$	$\gamma \beta 2$	$\gamma \beta 5$
Ac-PAD-EK	1.4	>1000	>1000	>1000	>1000	>1000	22.4	>200	>200
Ac-LAD-EK	1.6	>1000	115	226	21.7	12.3	24.5	>200	~100
Ac-PAE-EK	17.7	>1000	>1000	>1000	>1000	>1000	>200	>200	~200
Ac-LAE-EK	14.8	>1000	3.6	106	>1000	>1000	>200	>200	~200
Ac-PAF-EK	10.8	0.021	540	625	5.5	2.1	>200	>200	10.9
Ac-LAF-EK	61.7	0.23	0.55	0.78	0.02	0.11	>200	10.0	0.23
Ac-PAY-EK	10.3	0.38	69.0	65.1	4.9	1.5	>200	>200	7.9
Ac-LAY-EK	5.3	0.33	0.14	0.14	0.054	0.15	>200	6.0	0.27
Ac-PAI-EK	338	28.9	>1000	>1000	>1000	>1000	>200	>200	>200
Ac-LAI-EK	261	10.8	9.2	8.1	19.7	253	>200	>200	100.4
Ac-PAL-EK	0.11	0.028	326.7	446.6	9.5	10.7	43.3	>200	>200
Ac-LAL-EK	0.95	0.020	0.083	0.074	0.06	1.0	55.5	14.4	0.67
Ac-PAV-EK	>1000	36.0	>1000	>1000	~1000	>1000	>200	>200	>200
Ac-LAV-EK	>1000	101.5	7.2	15.6	25.9	~1000	>200	>200	>200
Ac-PAA-EK	>1000	99.5	457	432	248	>1000	>200	>200	>200
Ac-LAA-EK	>100	26.8	1.4	2.8	1.1	>1000	>200	65.8	89.3
Ac-PLL-EK	0.11	0.040	93.2	46.4	10.8	8.8	39.7	>200	~200
Ac-LLL-EK	0.14	0.015	0.041	0.017	0.1	1.0	40.9	10.4	0.72
H-APLL-EK	0.72	0.085	61.5	160	21.9	131	101	>200	>200
Ac-APLL-EK	0.029	0.035	75.1	37.3	26.6	52.4	10.3	>200	~200
N ₃ G-APAL-EK	0.20	0.043	>1000	>1000	43.8	46.2	7.6	>200	>200
N ₃ G-APnLL-EK	0.067	0.026	>100	>100	~100	~100	5.3	>200	~200
N ₃ G-A(4,4-F ₂ P)nLL-EK	2.7	0.031	>100	>100	6.2	17.1	11.4	>200	>200
Carfilzomib	0.075	0.017	0.031	0.017	0.00089	0.0013	~200	1.8	0.03
ONX 0914 ¹⁶	>10	0.46	1.1	0.59	0.054	0.0057	>200	5.5	0.42
Bortezomib	0.017	0.0065	0.51	0.57	0.0027	0.0027	1.1	9.3	0.05

Table 2. Apparent IC_{50} values [μM] of compounds as determined by ABPP with Raji cell lysates or by fluorogenic substrate hydrolysis assays for purified wt yCP

Despite recent progress,^{6, 17} mouse and human CPs are still difficult to crystallize and not suitable for extensive structural analyses with various ligands. Therefore the yCP is used for crystallographic and mutagenesis studies. The conserved fold of proteasome subunits and the identical binding mode of peptidic compounds to yeast and mammalian CPs justify usage of the model system yeast.⁶ Active yCP crystals were soaked with final ligand concentrations of 3.3 mM. At this concentration, even poor inhibitors bind to the active sites of the yCP and thus can be visualized by X-ray crystallography. All synthesized inhibitors share the α',β' -epoxyketone (EK/ep) warhead, which afforded us to use carfilzomib and ONX-0914 as reference compounds. Although distinct covalent inhibitor types¹⁸ and natural substrates may slightly vary in their binding mode to the S pockets of the proteasomal active site, the analyses provided here disclose significant tendencies for substrate preferences.

Epoxyketones with P1-valine or -isoleucine are weak proteasome inhibitors

Among the different P1 residues tested (Tables 1 and 2), Val and Ile are the most disfavored ones (IC_{50} for human CP $\geq 7.2 \mu\text{M}$, for yCP $\geq 100 \mu\text{M}$; Tables 2). For example, the compounds Ac-LAI-EK ($\geq 8.1 \mu\text{M}$ for human CP) and Ac-LAV-EK ($\geq 7.2 \mu\text{M}$ for human CP) are significantly less potent than Ac-LAL-EK ($\leq 1 \mu\text{M}$ for human CP). The crystal structures of yCP in complex with the inhibitors showed that the conformation of Val and Ile in the S1 pocket clearly differs from that observed for Leu. Val and Ile are sandwiched in the S1 pocket between the main chain heteroatoms 19O, 45O, 47O and 47N (Figure 1). Based on the interatomic distances, it is expected that also Thr is disfavored at the P1 position, but experimental support for this hypothesis is lacking due to the abortive synthesis of the respective epoxyketones.

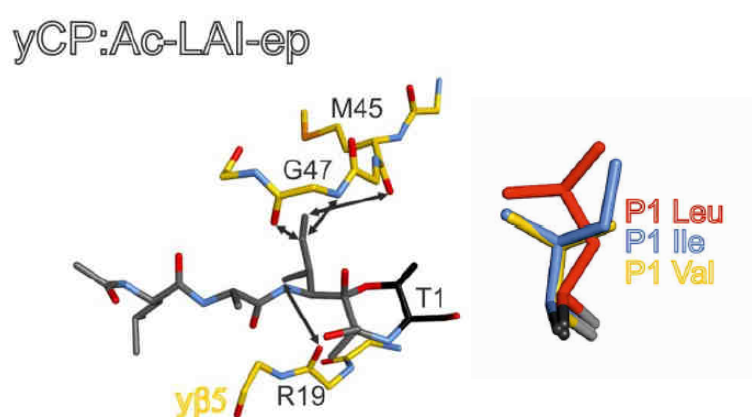


Figure 1. The P1 site: Ile and Val versus Leu. The crystal structure of Ac-LAI-EK at the $\gamma\beta 5$ active site clearly shows that main chain atoms 19O, 45O, 47O and 47N come close to the P1-Val/Ile residues. The overlay of the side chains of P1 Leu, Ile and Val residues show a different conformation for Ile and Val compared to Leu. The main chain atoms of residues 18-20 and 44-48 are shown as sticks. The side chains, except for Met45 forming the bottom of the S1 pocket, are removed for clarity.

Blockage of β 2 subunits

The marginal structural differences between the β 2c and β 2i subunits (Asp53 (β 2c) versus Glu53 ($\gamma\beta$ 2/ β 2i) and Thr48 ($\gamma\beta$ 2/ β 2c) versus Val48 (β 2i)) do not provide an obvious explanation for the biological need of their mutual exchange in mammalian CPs⁶ and complicate the design of β 2c- and β 2i-selective inhibitors.⁴ In order to gain additional insights into the β 2 substrate specificities, all compounds were tested for blockage of the T-L activities. Ligands bearing Pro at P3 poorly inhibit the β 2 active sites of yeast and human CPs, while P3-Leu analogues are much more potent (up to 6000 times). Among the P3-Leu series, inhibitors with Leu, Phe and Tyr at P1 display highest affinities and do not discriminate between the human c- and i-subunits (Table 2). These results confirm that the β 2 subunits accept nonpolar P1 side chains³ besides basic ones such as Arg.⁴ Small P1 residues like Ala, however, are less potent due to the large S1 pockets of β 2 subunits. The acidic amino acids Asp and Glu are also hardly effective towards yeast and human CPs, except for Ac-LAE-EK (IC_{50} : 3.6 μ M; Table 2) which selectively targets the human β 2c subunit. In contrast to β 2i, where Glu53 is oriented towards His35, Asp53O⁶ of β 2c may hydrogen-bond to the P1-Glu side chain of the ligand via a solvent molecule, thereby enhancing the affinity of Ac-LAE-EK for β 2c compared to β 2i (IC_{50} : 106 μ M; Table 2) by a factor of 30.^{6, 17}

Inhibition of subunit β 5

Characterization of the β 5 substrate binding channel by numerous natural and synthetic ligands revealed that the S3 pockets of yeast and mammalian β 5 subunits readily accept Leu residues.^{17, 19} Upon testing the impact of Leu versus Pro at the P3 site it was found that the latter is not suitable to target the ChT-L activity (up to 300 times decreased potency compared to P3-Leu inhibitors). Yeast β 5 and mammalian constitutive β 5c subunits lack a defined S2 pocket^{3, 6, 17} (Gly48 vs Cys48 for β 5i) and consequently, exchange of the P2-Ala by Leu does not significantly affect IC_{50} values. For instance, the compounds Ac-LAL-EK and Ac-LLL-EK as well as Ac-PAL-EK and Ac-PLL-EK are equipotent (Table 2). Nonetheless, potency and subunit selectivity may be affected by larger P2 side chains such as Phe, Tyr or Trp, which provide additional anchorage especially in subunit β 5i by interacting with Cys48.⁶

Proteasomal β 5 subunits select for distinct sizes of apolar P1-residues. In agreement with structural data,⁶ the fluorogenic substrate Ac-WLA-AMC, featuring a P1-Ala side chain can be used to monitor β 5c activity.⁵ We observed that the IC_{50} value of the P1-Ala compound Ac-LAA-EK for subunit β 5c (1.1 μ M) is reduced up to 55 fold compared to analogues bearing Leu, Phe or Tyr as P1 side chains (IC_{50} : < 0.06 μ M; Tables 2-3). Ac-LAA-EK, however, does not inhibit subunit β 5i and thus it represents a basic scaffold for developing β 5c selective compounds. Crystallographic data disclose that Ala undergoes hydrophobic contacts with Met45 of the

$\gamma\beta 5$ -S1 pocket (Ala C $^{\beta}$ to Met45 C $^{\gamma}$ /S: 4 Å) without changing its position (Figures S3d and S3g). The $\beta 5i$ -S1 site does not provide this stabilization because the peculiar conformation of Met45 enlarges the S1 pocket. Hence, subunit $\beta 5i$ is not targeted by Ac-LAA-EK (AlaC $^{\beta}$ to Met45: 5.7-7.4 Å; Figure 2).⁶

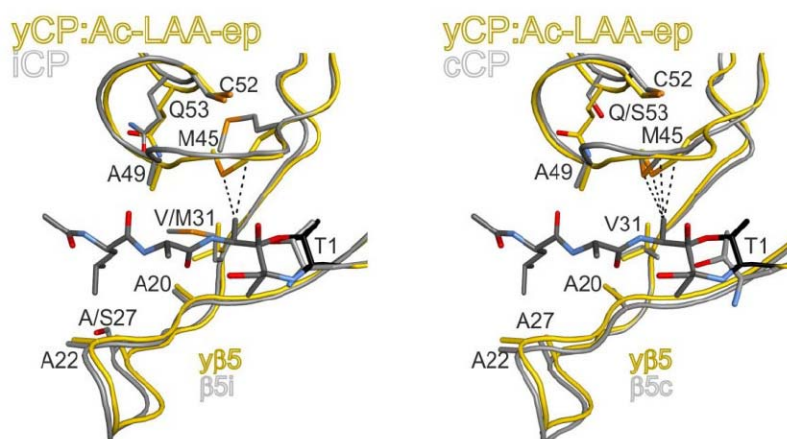


Figure 2. Superposition of $\gamma\beta 5$ in complex with Ac-LAA-EK onto the mouse $m\beta 5$ proteasome subunits. The peculiar conformation of Met45 in $\beta 5i$ (left) cannot stabilize the P1-Ala residue by hydrophobic interactions (black dashed lines; distances 5.7-7.4 Å), whereas the orientations of Met45 in the yeast $\beta 5$ subunit (distances 3.9-4.0 Å) and in $\beta 5c$ of the constitutive proteasome are more favorable (distances 4.2-4.8 Å).

Fluorogenic substrates for $\beta 5c$ and $\beta 5i$ feature either Leu or Tyr as P1 residues (carboxybenzyl-Gly-Gly-Leu-para-nitroanilide (Z-GGL-pNA), N-succinyl-Leu-Leu-Val-Tyr-AMC (Suc-LLVY-AMC)²⁰) and most inhibitors of the $\beta 5$ active site, including the FDA approved drugs bortezomib and carfilzomib, bear Leu in P1. Epoxyketones with Leu, Phe or Tyr in P1 as well as Leu in P3 are very potent inhibitors of both human $\beta 5c$ ($IC_{50} \leq 0.1 \mu M$) and $\beta 5i$ ($IC_{50} \leq 1.0 \mu M$; Table 2) as well as of the yeast counterpart ($IC_{50} \leq 0.72 \mu M$; Table 3). Consistent with previous suggestions,^{6, 21} the P1-Leu residue drives $\beta 5c$ selectivity: Ac-LAL-EK is 17 times more specific for $\beta 5c$ than $\beta 5i$ (Table 2). The tendency of Ac-LAF-EK to favor $\beta 5c$ over $\beta 5i$ (up to 6 times) probably results from the P3 and P2 residues (Table 2). The P3-Leu, which is accepted by both $\beta 5c$ and $\beta 5i$,^{6, 17, 19, 21, 22} and the P2-Ala residue, which fails to interact with Cys48 of $\beta 5i$, do not promote $\beta 5i$ but $\beta 5c$ selectivity. Interestingly, reported $\beta 5i$ inhibitors always have aromatic P2 residues, which interact with Cys48 and small P3 residues which are disfavored by $\beta 5c$.²¹ Altogether, these results imply that the P1 side chain is a major determinant of affinity for $\beta 5$ subunits, but does not act as the sole.

General aspects for targeting $\beta 1$

The yeast $\gamma\beta 1$ and the mammalian constitutive $\beta 1c$ subunits have been attributed to cleave peptide bonds after negatively charged amino acids.²³ To assay this activity of the proteasome, the fluorogenic substrate Z-Leu-Leu-Glu-AMC is frequently used. However, it has been found^{8,9} that $\gamma\beta 1$ and $\beta 1c$ prefer ligands featuring Asp and Leu at P1 over those bearing Glu or any other amino acid in this position (Table 2). For instance, the IC_{50} values of P1-Asp/Leu compounds for $\beta 1c$ range $\leq 1.6 \mu M$, whereas those for all other compounds are $\geq 5.3 \mu M$. In agreement, bortezomib (Leu in P1; $IC_{50} \leq 1 \mu M$) but not ONX-0914 (Phe in P1; $IC_{50} > 10 \mu M$) targets the subunits $\gamma\beta 1$ and $\beta 1c$ with high affinity (Tables 2). The wild-type $\gamma CP:Ac-PAE-EK$ crystal structure revealed that Glu fits well in the $\gamma\beta 1$ S1 pocket (Figure 3), but lacks any direct interaction with surrounding protein residues. In contrast to Glu, the P1-Asp is hydrogen-bonded to Thr200 γ (2.8 Å) and the P1-Leu undergoes favorable van-der-Waals interactions with the methyl group of Thr20 (3.7 Å). These contacts cause the Asp and Leu side chains to adopt distinct orientations in the S1 pocket. Repulsion of the positively charged Arg45 and the nonpolar P1-Leu side chain is prevented by a negatively charged counter ion (e.g. Cl⁻) that is hydrogen-bonded to Arg45NH1 and Arg45NH2 (Figure 3). The minor stabilization of the Glu side chain compared to Asp and Leu causes its disfavor at the P1 position. In conclusion, the $\gamma\beta 1$ and $\beta 1c$ active sites of the proteasome rather exert C-L and ChT-L than PGPH activities, and rather select for a certain size of side chain than for its charge. Regarding potency and selectivity for $\beta 1c$, Ac-PAD-EK and Ac-LAD-EK are the most outstanding compounds ($IC_{50} \leq 1.6 \mu M$; Table 2).

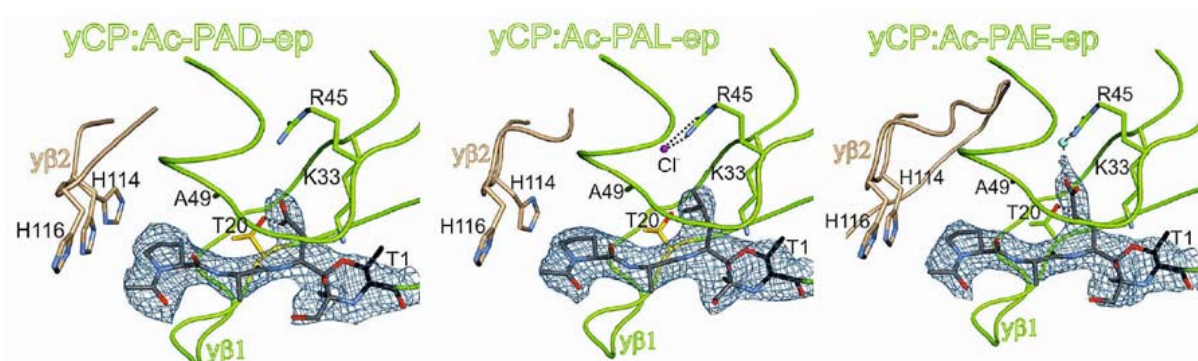


Figure 3. The $\gamma\beta 1$ active site favors Asp or Leu in its S1 pocket. $2F_o - F_c$ electron density map (blue mesh; contoured at 1σ) for Ac-PAD-EK (left), Ac-PAL-EK (middle) and Ac-PAE-EK (right) bound to the $\gamma\beta 1$ active site (green) of the wt yeast 20S proteasome (γCP). The active site Thr1 is marked in black, Thr20 in yellow. The chloride ion is depicted as a purple sphere. Hydrogen bonds are indicated by black dashed lines. Important residues of the neighboring subunit $\gamma\beta 2$ (brown), contributing to the S3 and S4 pockets of the $\gamma\beta 1/\gamma\beta 2$ substrate binding channel are illustrated. His114 adopts alternative conformations.

Consistent with structural data^{3, 6} and cleavage pattern analyses¹¹, the $\beta 1i$ substrate binding channel is targeted by compounds featuring hydrophobic residues (Table 2). The apolar $\beta 1i$ active site surroundings (e.g. Val20 and Leu45 in $\beta 1i$ versus Thr20 and Arg45 in $\beta 1c$) enhance

the IC₅₀ values of Ac-PAL-EK (IC₅₀: 0.11 μM (β1c)/0.028 μM (β1i)) and Ac-LAL-EK (IC₅₀: 0.95 μM (β1c)/0.020 μM (β1i)) at least four-times compared to β1c by providing favorable van-der-Waals stabilization. Besides Leu, also the aromatic amino acids Phe (IC₅₀: ≥10.8 μM (β1c)/≤0.23 μM (β1i)) and Tyr (IC₅₀: ≥5.3 μM (β1c)/≤0.38 μM (β1i)) represent appropriate P1 residues for β1i-ligands. In fact, the respective compounds are up to 500 times more selective for β1i than for β1c (Table 2).¹⁶ Previously it has been suggested that ONX-0914 targets β5i over β1i due to steric hindrance with Phe31 of the β1i S1 pocket.⁶ Nonetheless, Phe containing ligands can be used to block subunit β1i (see also chapter 5).¹⁶

Next, acetyl-capped tetrapeptides and tripeptides were evaluated and it was found that elongation of Ac-PLL-EK by a P4-Ala residue enhances the IC₅₀ value for γβ1 and β1c by a factor of four (Table 2). In addition, Ac-APLL-EK (IC₅₀: 0.029 μM) is 25 times more active for β1c than H-APLL-EK (IC₅₀: 0.72 μM; Table 2). The structural data show that a hydrogen bond between the carbonyl oxygen atom of the acetyl-cap and Thr22O^γ additionally stabilizes these elongated inhibitors in γβ1 (distance ~3 Å; data not shown). Thr22 is conserved in mammalian β1c subunits, but exchanged for Ala in β1i entities. Congruently, the IC₅₀ values of Ac-APLL-EK (IC₅₀: 0.035 μM) and H-APLL-EK (IC₅₀: 0.085 μM) for subunit β1i are almost identical. In summary, capped tetrapeptides may be useful to target β1c.

The S3 pocket – A peculiarity of the β1 active site.

P3-Pro ligands serve as potent inhibitors of C-L^{8, 16} or BrAAP (branched chain amino acid-preferring)^{24, 25} activities and as selective substrates (e.g. Ac-PAL-AMC) to monitor peptide bond hydrolysis by β1i⁵. The structural basis for the β1-preference of the P3-Pro residue, however, remained elusive. The analysis of the subunit selectivity profile of various compounds bearing either Leu or Pro residues at their P3 site for the yeast as well as mammalian CPs revealed that, in contrast to the β1 active sites, the β2 and β5 subunits disfavor inhibitors with Pro at P3 (Table 2). For instance, Ac-LAD-EK is selective for β1c (IC₅₀: 1.6 μM), but is also able to inhibit β5c and β5i quite efficiently (IC₅₀: 21.7 resp. 12.3 μM) and to a lesser extent also the β2 subunits are inhibited. In contrast, Ac-PAD-EK is not able to inhibit the β5 subunits even at 1 mM. Hence, it is the poor affinity of the Pro-ligands for the subunits β2 and β5 that causes their β1 selectivity.

The non-primed substrate binding channels of the proteasome are formed by two adjacent β subunits (β1/2, β2/3 and β5/6), with β2, β3 and β6 contributing to the S3 pockets of the β1, β2 and β5 active sites. For instance, Asp114O^δ of β3 and β6 hydrogen bonds to the amide nitrogen of P3-Leu ligands bound to β2 and β5, respectively (Figure 4c)^{3, 6, 17} and thereby stabilizes the ligand in the substrate binding channel. The backbone nitrogen of P3-Pro compounds, however, is not accessible for this interaction. Furthermore, in β2/3 and β5/6,

the P3-Pro inhibitors are shifted compared to the Leu-counterparts (Figures 4A-B), because Asp114 hinders placing of the Pro side chain in the S3 pocket (distance ≥ 3.0 Å). By contrast, P3-Pro and P3-Leu featuring inhibitors adopt the same conformation in β_1 , suggesting that the structure of the $\beta_1/2$ substrate binding channel tolerates P3-Pro-residues better than the β_2/β_3 and β_5/β_6 active sites. Indeed, the $\beta_1/2$ active site is ideally suited to accommodate P3-Pro ligands. Instead of Asp114, β_2 encodes either His114 (in $\gamma\beta_2/\beta_2i$ subunits) or Tyr114 (in β_2c subunits). According to the mouse cCP and iCP crystal structures, both Tyr114 in β_2c and His114 in β_2i and $\gamma\beta_2$ adopt the same side chain orientation. However, this conformation clearly differs from that observed for Asp114 in β_3 and β_6 and prevents hydrogen-bonding to the P3 nitrogen of ligands (Figures 4D-E). Consequently, at the $\beta_1/2$ active site, P3-Leu ligands are less stabilized and P3-Pro ones do not receive repulsion.

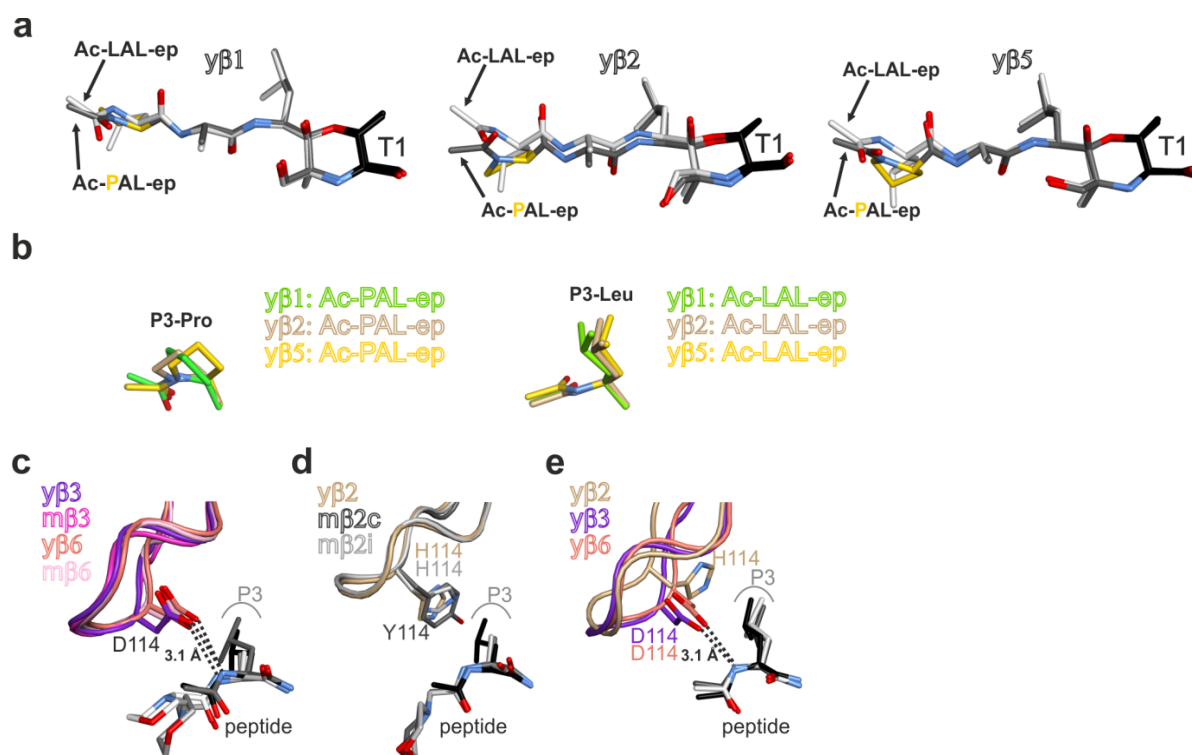


Figure 4. The impact of residue 114 of β_2 , β_3 and β_6 on ligand binding. A) Superposition of the compounds Ac-LAL-EK (light gray) and Ac-PAL-EK (dark gray) bound to the yeast β_1 , β_2 and β_5 subunits highlights differences at the P3 site. P3-Pro and P3-Leu featuring compounds adopt almost identical conformations in the $\gamma\beta_1$ substrate binding channel, while in the $\gamma\beta_2$ and $\gamma\beta_5$ counterpart profound changes are observed. These explain the disfavor of P3-Pro compounds by the latter two active sites. B) Superposition of the Ac-PAL-EK bound to $\gamma\beta_1$ (green), $\gamma\beta_2$ (brown) and $\gamma\beta_5$ (yellow) depicts subunit-specific differences in the orientation of the P3-Pro residue, whereas for the P3-Leu side chain of Ac-LAL-EK no significant alterations are observed. C) Superposition of yeast and mouse β_3 and β_6 subunits visualize that Asp114 occupies the same position in all subunits. Notably, Asp114 is hydrogen-bonded to the P3 amide nitrogen atom of the ligands peptide backbone (Ac-LAL-EK for yeast and ONX 0914 for mouse). The P3 sites of ligands are depicted as sticks. D) Superposition of yeast and mouse β_2 subunits indicates that His114 ($\gamma\beta_2$ and $m\beta_2i$) and Tyr114 ($m\beta_2c$) perfectly overlay and lack any interaction with the ligand's peptide backbone. E) Superposition of the yeast β_2 , β_3 and β_6 subunits depicts that only Asp114 in β_3 and β_6 hydrogen-bonds to the P3 amide nitrogen atom of the ligand. Notably, the position of the C $^\alpha$ atom of His114 in $\gamma\beta_2$ significantly deviates from that observed for Asp114 of $\gamma\beta_3$ and $\gamma\beta_6$.

Conclusion

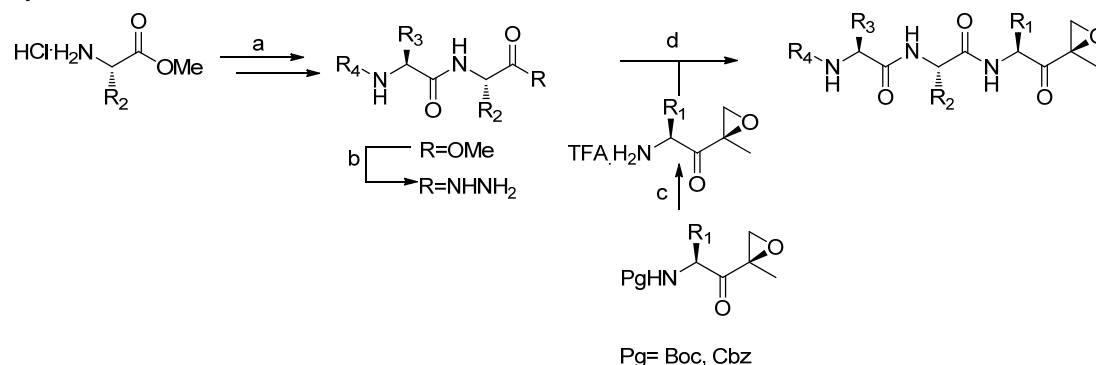
A series of 23 tri- and tetrapeptide epoxyketone proteasome inhibitors was synthesized. They feature varying P1 residues, including hydrophobic and acidic ones and bear either a P3-Leu or P3-Pro residue. All compounds were examined for their potency and subunit selectivity in human cell lysates and with purified yeast proteasome. The binding modes of all inhibitors were visualized by X-ray crystallography. The obtained data reveal the following key findings: 1) epoxyketones featuring Val or Ile as P1 residues are disfavored by yeast and human proteasomes; 2) two compounds were identified that favor $\beta 2c$ and $\beta 5c$, respectively, over the i-counterparts: Ac-LAE-EK and Ac-LAA-EK; 3) yeast and human constitutive $\beta 1$ active sites prefer Asp over Glu at the P1 position and thus exert rather caspase-like activities. Besides Asp, Leu is also well accepted by the $\gamma\beta 1/\beta 1c$ S1 pocket; 4) a structural explanation for the $\beta 1$ -preference of P3-Pro-featuring compounds is provided. Asp114 of $\beta 3$ and $\beta 6$ impairs binding of P3-Pro ligands to the $\beta 2$ and $\beta 5$ subunits, while His/Tyr114 of $\beta 2$ allows their accommodation in the $\beta 1/2$ substrate binding channel due to exceptional backbone and side chain orientations. Together, the structural and biochemical data presented here will support future efforts to improve existing proteasome inhibitors as well as to design proteasome-type selective and subunit-specific inhibitors, as will be described in the next chapters. Such compounds would serve to characterize in more detail the biological roles of the individual proteasomal active sites and might qualify for diverse medical applications including cancer and inflammatory diseases.

Experimental

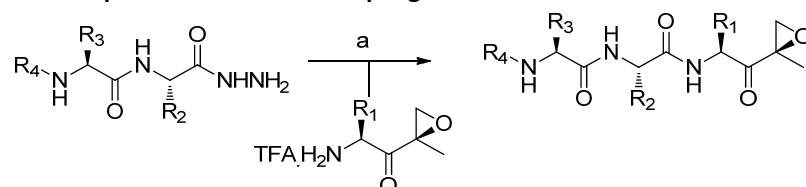
Synthetic procedures

General procedures

Acetonitrile (ACN), dichloromethane (DCM), N,N-dimethylformamide (DMF), methanol (MeOH), diisopropylethylamine (DiPEA) and trifluoroacetic acid (TFA) were of peptide synthesis grade, purchased at Biosolve, and used as received. All general chemicals (Fluka, Acros, Merck, Aldrich, Sigma, Iris Biotech) were used as received. Traces of water were removed from reagents used in reactions that require anhydrous conditions by co-evaporation with toluene. Diethylether was stored over 4 Å molecular sieves. Column chromatography was performed on Screening Devices b.v. Silica Gel, with a particle size of 40-63 µm and pore diameter of 60 Å. TLC analysis was conducted on Merck aluminium sheets (Silica gel 60 F254). Compounds were visualized by UV absorption (254 nm), by spraying with a solution of (NH₄)₆Mo₇O₂₄·4H₂O (25 g/L) and (NH₄)₄Ce(SO₄)₄·2H₂O (10 g/L) in 10% sulphuric acid, a solution of KMnO₄ (20 g/L) and K₂CO₃ (10 g/L) in water, or ninhydrin (0.75 g/L) and acetic acid (12.5 mL/L) in ethanol, where appropriate, followed by charring at ca. 150°C. ¹H and ¹³C-NMR spectra were recorded on a Bruker AV-400 (400 MHz) or AV-600 (600 MHz) spectrometer. Chemical shifts are given in ppm (δ) relative to tetramethylsilane, CD₃OD or CDCl₃ as internal standard. High resolution mass spectra were recorded by direct injection (2 µL of a 2 µM solution in water/acetonitrile 50/50 (v/v) and 0.1% formic acid) on a mass spectrometer (Thermo Finnigan LTQ Orbitrap) equipped with an electrospray ion source in positive mode (source voltage 3.5 kV, sheath gas flow 10, capillary temperature 250 °C) with resolution R = 60,000 at m/z 400 (mass range m/z = 150-2,000) and dioctylphthalate (m/z = 391.28428) as a "lock mass". The high resolution mass spectrometer was calibrated prior to measurements with a calibration mixture (Thermo Finnigan). Optical rotations were recorded on a Propol automatic polarimeter. LC-MS analysis was performed on a Finnigan Surveyor HPLC system with a Gemini C₁₈ 50 × 4.60 mm column (detection at 200–600 nm) coupled to a Finnigan LCQ Advantage Max mass spectrometer with ESI. The applied buffers were H₂O, MeCN and 1.0% TFA in H₂O (0.1% TFA end concentration). Methods used are: 10→90% MeCN, 13.5 min (0→0.5 min: 10% MeCN; 0.5→8.5 min: gradient time; 8.5→10.5 min: 90% MeCN; 10.5→13.5 min: 90% → 10% MeCN). 0→50% MeCN, 13.5 min (0→0.5 min: 0% MeCN; 0.5→8.5 min: gradient time; 8.5→10.5 min: 90% MeCN; 10.5→13.5 min: 90% → 10 (or 0)% MeCN). HPLC purification was performed on a Gilson HPLC system coupled to a Phenomenex Gemini 5µm 250×10 mm column, or on an Agilent HPLC/MS system coupled to a Phenomenex Gemini 5µm 250×10 mm column or on a Waters autopurifier HPCL/MS system coupled to a Phenomenex Gemini 5µm 150×21.2 mm column. All tested compounds are >95% pure on the basis of LC-MS and NMR, except Ac-PAD-EK, which is >60% pure on the basis of LC-MS and NMR. Boc-Leu-Leu-OMe²⁶, Boc-Leu-EK, Boc-Phe-EK²⁷ were synthesized according to literature procedures. Compound N₃G-A(4,4-F₂P)nLL-EK has been described elsewhere¹⁶.

Synthetic details

Scheme 1. General synthetic route towards peptide-epoxyketones. Reagents and conditions: (a). Sequential peptide coupling/Boc removal. Peptide coupling: HCTU, DiPEA, Boc-AA-OH, DCM. Boc-removal: TFA/DCM. (b) $\text{NH}_2\text{NH}_2 \cdot \text{H}_2\text{O}$, MeOH; (c) TFA or Pd/C, MeOH, TFA, $\text{H}_2(\text{g})$; (d) i) $t\text{BuONO}$, HCl, DMF, DCM, -30°C ; ii) amine (warhead), DiPEA, $-30^\circ\text{C} \rightarrow \text{RT}$.

General procedure for azide couplings.

Scheme 2. Azide coupling towards peptide epoxyketones. Reagents and conditions: (a) i) $t\text{BuONO}$, HCl, DMF, DCM, -30°C ; ii) amine (warhead), DiPEA, $-30^\circ\text{C} \rightarrow \text{RT}$.

All peptide epoxyketones were prepared via azide coupling of properly protected tripeptide hydrazide and properly deprotected epoxyketone amines (see scheme 2). The appropriate hydrazide was dissolved in 1:1 DMF:DCM (v/v) and cooled to -30°C . $t\text{BuONO}$ (1.1 equiv.) and HCl (4M solution in 1,4-dioxane, 2.8 equiv.) were added, and the mixture was stirred for 3h at -30°C after which TLC analysis (10% MeOH/DCM, v/v) showed complete consumption of the starting material. The epoxyketone a free amine was added to the reaction mixture as a solution in DMF. DiPEA (5 equiv.) was added to the reaction mixture, and this mixture was allowed to warm to RT slowly overnight. The mixture was diluted with EtOAc or DCM and extracted with H_2O (3×). The organic layer was dried over MgSO_4 concentrated and purified by flash column chromatography (1-5% MeOH in DCM) and HPLC-purification (if necessary).

Ac-PAD(OtBu)-EK

This compound was obtained by the general protocol for azide coupling on a 140 μmol scale. Purification by column chromatography (0→3% MeOH in DCM) provided the title compound (40 mg, 65%) Complex NMR due to presence of rotamers (8:1) and presence of by-product (reduced epoxide). Peaks of major rotamer of the product are reported. ^1H NMR (400 MHz, CDCl_3) δ 7.40 (d, J = 6.8 Hz, 1H), 7.09 (d, J = 7.4 Hz, 1H), 4.69 (dt, J = 7.4, 5.5 Hz, 1H), 4.53 (dd, J = 8.0, 2.1 Hz, 1H), 4.34 (p, J = 6.8 Hz, 1H), 3.63 – 3.53 (m, 1H), 3.43 (td, J = 9.5, 7.0 Hz, 1H), 3.18 (d, J = 4.8 Hz, 1H), 2.87 (d, J = 4.9 Hz, 1H), 2.84 – 2.66 (m, 2H), 2.35 (dtd, J = 14.6, 8.7, 7.6, 3.3 Hz, 1H), 2.10 (s, 3H), 2.04 – 1.80 (m, 3H), 1.52 (s, 3H), 1.41 (d, J = 3.7 Hz, 9H), 1.35 (t, J = 7.2 Hz, 3H). ^{13}C NMR (101 MHz, CDCl_3) δ 205.57, 172.08, 171.41, 171.16, 169.51, 81.99, 59.93, 59.79, 59.48, 52.67, 49.55, 49.43, 49.28, 48.47, 37.02, 28.11, 27.78, 25.20, 22.67, 17.79, 17.07. LC-MS (linear gradient 10 → 90% MeCN, 0.1% TFA, 12.5 min): R_t (min): 5.37 (ESI-MS (m/z): 440.07 ($\text{M}+\text{H}^+$)). HRMS: calculated for $\text{C}_{21}\text{H}_{33}\text{N}_3\text{O}_7$ 440.23913 [$\text{M}+\text{H}^+$] $^+$; found 440.23914.

Ac-PAD-EK

Ac-PAD(OtBu)-EK (40 mg) was dissolved in dry DCM (1 mL), followed by the addition of dry TFA (1 mL, ampule). After 15 min, 2 mL dry toluene was added and the reaction mixture was concentrated and co-evaporated with toluene. Purification by HPLC (C₁₈, Waters, 5-15% MeCN, 0.1% TFA, 9 min gradient), yielding the product as a white solid after lyophilisation. Analysis revealed partial hydrolysis of the product during HPLC-purification (about 30%). Complex NMR due to mixtures of product with hydrolysed product and rotamers. Peaks reported for product, major rotamer. ¹H NMR (600 MHz, MeOD) δ 4.79 (m, 1H), 4.44 – 4.38 (m, 2H), 3.65 – 3.56 (m, 2H), 3.28 – 3.23 (m, 1H), 2.95 (d, *J* = 5.0 Hz, 1H), 2.82 (dd, *J* = 16.6, 4.8 Hz, 1H), 2.73 (dd, *J* = 16.5, 7.9 Hz, 1H), 2.22 (d, *J* = 2.4 Hz, 1H), 2.13 – 2.10 (m, 3H), 2.10 – 1.96 (m, 3H), 1.51 (s, 3H), 1.37 (d, *J* = 7.2 Hz, 3H). ¹³C NMR (151 MHz, MeOD) δ 207.04, 174.72, 174.49, 173.28, 172.64, 61.37, 53.25, 50.51, 50.11, 49.48, 35.89, 31.04, 25.75, 22.29, 17.64, 17.15. LC-MS (linear gradient 0 → 50% MeCN, 0.1% TFA, 12.5 min): R_t (min): 5.40 (ESI-MS (*m/z*): 384.07 (M+H⁺)); 4.76 (ESI-MS (*m/z*): 402.07 (M+H⁺) (hydrolysed product)). HRMS: calculated for C₁₇H₂₅N₃O₇ 384.17653 [M+H]⁺; found 384.17653.

Ac-LAD(OtBu)-EK

This compound was obtained by the general protocol for azide coupling on a 100 μmol scale. Purification by column chromatography (0→3% MeOH in DCM) provided the title compound (24.5 mg, 54%) as a white powder after lyophilisation. Complex NMR due to presence of rotamers and presence of by-product (reduced epoxide). Peaks of major rotamer of the product are reported. ¹H NMR (400 MHz, CDCl₃) δ 7.42 (d, *J* = 7.7 Hz, 1H), 7.12 (d, *J* = 7.6 Hz, 1H), 6.61 (d, *J* = 8.5 Hz, 1H), 4.82 – 4.67 (m, 1H), 4.71 – 4.48 (m, 2H), 3.16 (d, *J* = 4.9 Hz, 1H), 2.88 (d, *J* = 4.8 Hz, 1H), 2.74 (qt, *J* = 12.4, 5.7 Hz, 2H), 2.02 (d, *J* = 3.4 Hz, 3H), 1.70 – 1.46 (m, 6H), 1.40 (d, *J* = 2.5 Hz, 9H), 1.34 (d, *J* = 7.0 Hz, 3H), 0.91 (d, *J* = 6.2 Hz, 6H). ¹³C NMR (101 MHz, CDCl₃) δ 205.70, 172.26, 172.14, 170.39, 169.43, 82.02, 59.43, 52.58, 51.63, 49.32, 48.82, 41.99, 37.17, 28.12, 24.87, 23.22, 23.06, 22.25, 18.95, 17.03. LC-MS (linear gradient 10 → 90% MeCN, 0.1% TFA, 12.5 min): R_t (min): 6.25 (ESI-MS (*m/z*): 456.07 (M+H⁺)).

Ac-LAD-EK

Ac-LAD(OtBu)-EK (8.7 mg) was dissolved in dry DCM (1 mL), followed by the addition of dry TFA (1 mL, ampule). After 15 min, 2 mL dry toluene was added and the reaction mixture was concentrated and co-evaporated with toluene. Purification by HPLC (C₁₈, Gemini, 21-24% MeCN, 0.1% TFA, 12 min gradient), yielding the product as a white solid after lyophilisation (3.31 mg, 43%). ¹H NMR (600 MHz, MeOD) δ 4.78 – 4.70 (m, 1H), 4.43 – 4.29 (m, 2H), 3.20 (s, 1H), 2.93 – 2.88 (m, 1H), 2.76 (dd, *J* = 16.4, 4.7 Hz, 1H), 2.68 (dd, *J* = 16.2, 7.6 Hz, 1H), 1.98 (s, 3H), 1.68 (dt, *J* = 13.4, 6.7 Hz, 1H), 1.57 (ddt, *J* = 9.2, 6.0, 3.2 Hz, 2H), 1.49 (s, 3H), 1.34 (d, *J* = 7.1 Hz, 3H), 0.95 (dd, *J* = 22.2, 6.6 Hz, 6H). ¹³C NMR (151 MHz, MeOD) δ 174.75, 174.55, 173.82, 173.56, 68.14, 53.31, 53.15, 50.61, 50.05, 41.76, 36.36, 25.92, 23.46, 22.38, 21.86, 18.05, 17.16. LC-MS (linear gradient 0 → 50% MeCN, 0.1% TFA, 12.5 min): R_t (min): 6.59 (ESI-MS (*m/z*): 400.07 (M+H⁺)). HRMS: calculated for C₁₈H₂₉N₃O₇ 400.20782 [M+H]⁺; found 400.20783.

Ac-PAE(OtBu)-EK

This compound was obtained by the general protocol for azide coupling on a 100 μmol scale. Purification by column chromatography (0→3% MeOH in DCM) provided the title compound (18.06 mg, 41%) as a white powder after lyophilisation. Complex NMR due to presence of rotamers (7:1) and presence of by-product (reduced epoxide). Peaks of major rotamer of the product are reported. ¹H NMR (400 MHz, CDCl₃) δ 7.34 (d, *J* = 6.9 Hz, 1H), 7.06 (d, *J* = 7.3 Hz, 1H), 4.56 – 4.42 (m, 2H), 4.38 – 4.27 (m, 1H), 3.63 – 3.53 (m, 1H), 3.50 – 3.38 (m, 1H), 3.30 (d, *J* = 4.9 Hz, 1H), 2.87 (d, *J* = 5.0 Hz, 1H), 2.38 – 2.22 (m, 3H), 2.10 (s, 3H), 2.09 – 1.85 (m, 4H), 1.76 (ddd, *J* = 15.8, 14.4, 7.2 Hz, 1H), 1.50 (d, *J* = 3.5 Hz, 3H), 1.43 (s, 9H), 1.34 (d, *J* = 7.1 Hz, 3H). ¹³C NMR (101 MHz, CDCl₃) δ 207.64, 172.45, 172.34, 171.45, 171.14, 81.03, 59.88, 59.34, 52.59, 51.40, 49.22, 48.48, 31.61, 28.19, 27.97, 25.97, 25.20, 22.67, 17.69, 16.68. LC-MS (linear gradient 10 → 90% MeCN, 0.1% TFA, 12.5 min): R_t (min): 5.61 (ESI-MS (*m/z*): 454.00 (M+H⁺)). HRMS: calculated for C₂₂H₃₅N₃O₇ 454.25478 [M+H]⁺; found 454.25461.

Ac-PAE-EK

Ac-PAE(OtBu)-EK (7.5 mg) was dissolved in dry DCM (1 mL), followed by the addition of dry TFA (1 mL, ampule). After 15 min, 2 mL dry toluene was added and the reaction mixture was concentrated and co-evaporated with toluene. Purification by HPLC (C₁₈, Gemini, 13-16% MeCN, 0.1% TFA, 12 min gradient), yielding the product as a white solid after lyophilisation (1.86 mg, 30%). Complex NMR due to presence of rotamers, peaks of major rotamer are reported. ¹H NMR (600 MHz, MeOD) δ 4.50 – 4.41 (m, 1H), 4.41 – 4.35 (m, 1H), 4.32 (q, *J* = 7.2 Hz, 1H), 3.69 – 3.62 (m, 1H), 3.62 – 3.52 (m, 1H), 3.25 (d, *J* = 5.1 Hz, 1H), 2.92 (d, *J* = 5.1 Hz, 1H), 2.44 – 2.27 (m, 3H), 2.26 – 2.18 (m, 1H), 2.09 (s, 3H), 2.08 – 2.00 (m, 2H), 2.00 – 1.92 (m, 2H), 1.82 – 1.72 (m, 1H), 1.48 (s, 3H), 1.35 (t, *J* = 6.5 Hz, 3H). ¹³C NMR (151 MHz, MeOD) δ 208.76, 175.09, 174.80, 174.48, 172.50, 61.21, 53.13, 53.11, 50.23, 49.48, 32.26, 31.03, 27.09, 25.74, 22.29, 17.70, 16.88. LC-MS (linear gradient 0 → 50% MeCN, 0.1% TFA, 12.5 min): R_t (min): 5.77 (ESI-MS (m/z): 398.07 (M+H⁺)). HRMS: calculated for C₁₈H₂₇N₃O₇ 398.19218 [M+H]⁺; found 398.19208.

Ac-LAE(OtBu)-EK

This compound was obtained by the general protocol for azide coupling on a 100 μmol scale. Purification by column chromatography (0→3% MeOH in DCM) provided the title compound (30.0 mg, 64%) as a white powder after lyophilisation. Complex NMR due to presence of rotamers and presence of by-product (reduced epoxide). Peaks of major rotamer of the product are reported. ¹H NMR (400 MHz, CDCl₃) δ 7.39 (d, *J* = 7.4 Hz, 1H), 7.08 (d, *J* = 7.6 Hz, 1H), 6.59 (d, *J* = 8.4 Hz, 1H), 4.67 – 4.43 (m, 3H), 3.28 (d, *J* = 5.0 Hz, 1H), 2.89 (d, *J* = 5.0 Hz, 1H), 2.36 – 2.22 (m, 2H), 2.13 – 1.96 (m, 4H), 1.78 (dd, *J* = 14.4, 7.8 Hz, 1H), 1.51 (s, 6H), 1.42 (s, 10H), 1.33 (d, *J* = 7.0 Hz, 3H), 0.95 – 0.84 (m, 6H). ¹³C NMR (101 MHz, CDCl₃) δ 207.72, 172.43, 172.40, 172.26, 170.39, 81.18, 59.34, 52.49, 51.75, 51.35, 48.79, 41.85, 31.66, 28.18, 26.20, 24.90, 23.21, 23.02, 22.27, 18.81, 16.65. LC-MS (linear gradient 10 → 90% MeCN, 0.1% TFA, 12.5 min): R_t (min): 6.42 (ESI-MS (m/z): 470.00 (M+H⁺)).

Ac-LAE-EK

Ac-LAE(OtBu)-EK (15 mg) was dissolved in dry DCM (1 mL), followed by the addition of dry TFA (1 mL, ampule). After 15 min, 2 mL dry toluene was added and the reaction mixture was concentrated and co-evaporated with toluene. Purification by HPLC (C₁₈, Gemini, 20-23% MeCN, 0.1% TFA, 12 min gradient), yielding the product as a white solid after lyophilisation (6.89 mg, 52%). ¹H NMR (600 MHz, MeOD) δ 4.47 (dd, *J* = 9.2, 4.0 Hz, 1H), 4.43 – 4.27 (m, 2H), 3.25 (d, *J* = 5.1 Hz, 1H), 2.93 (d, *J* = 5.1 Hz, 1H), 2.46 – 2.28 (m, 2H), 2.15 – 2.03 (m, 1H), 1.97 (s, 3H), 1.82 – 1.72 (m, 1H), 1.72 – 1.63 (m, 1H), 1.63 – 1.52 (m, 2H), 1.48 (s, 3H), 1.33 (d, *J* = 7.1 Hz, 3H), 0.95 (dd, *J* = 21.2, 6.6 Hz, 6H). ¹³C NMR (151 MHz, MeOD) δ 208.37, 176.16, 174.52, 174.36, 173.06, 59.80, 52.78, 52.74, 52.50, 49.68, 41.42, 31.11, 30.73, 26.57, 25.52, 23.07, 21.99, 21.50, 17.46, 16.46. LC-MS (linear gradient 0 → 50% MeCN, 0.1% TFA, 12.5 min): R_t (min): 6.85 and 6.61 (ESI-MS (m/z): 414.07 / 414.13 (M+H⁺)) (elutes in two peaks, most likely due to different salt-forms). HRMS: calculated for C₁₉H₃₁N₃O₇ 414.22342 [M+H]⁺; found 414.22348.

Ac-PAF-EK

This compound was obtained by the general protocol for azide coupling on a 100 μmol scale. Purification by column chromatography (0→3% MeOH in DCM) provided the title compound (19.40 mg, 47%) as a white powder after lyophilisation. Complex NMR due to presence of rotamers (8:1). Peaks of major rotamer are reported. ¹H NMR (400 MHz, CDCl₃) δ 7.35 (d, *J* = 7.0 Hz, 1H), 7.30 – 7.24 (m, 2H), 7.23 – 7.20 (m, 1H), 7.18 – 7.12 (m, 2H), 6.86 (d, *J* = 7.4 Hz, 1H), 4.77 (td, *J* = 8.2, 4.9 Hz, 1H), 4.41 (dd, *J* = 8.1, 2.0 Hz, 1H), 4.29 (p, *J* = 7.1 Hz, 1H), 3.54 (td, *J* = 9.1, 8.2, 3.1 Hz, 1H), 3.41 (td, *J* = 9.5, 7.1 Hz, 1H), 3.34 (d, *J* = 4.9 Hz, 1H), 3.11 (dd, *J* = 14.0, 4.9 Hz, 1H), 2.88 (d, *J* = 5.0 Hz, 1H), 2.76 (dd, *J* = 14.0, 8.4 Hz, 1H), 2.36 – 2.21 (m, 1H), 2.09 (s, 3H), 2.07 – 1.75 (m, 3H), 1.47 (s, 3H), 1.23 (d, *J* = 7.1 Hz, 3H). ¹³C NMR (101 MHz, CDCl₃) δ 207.62, 172.05, 171.45, 171.21, 136.01, 129.44, 128.76, 128.57, 127.09, 59.68, 59.40, 52.68, 52.67, 48.93, 48.48, 37.07, 27.61, 25.19, 22.65, 17.09, 16.67. LC-MS (linear gradient 10 → 90% MeCN, 0.1% TFA, 12.5 min): R_t (min): 5.48 (ESI-MS (m/z): 416.13 (M+H⁺)). HRMS: calculated for C₂₂H₂₉N₃O₅ 416.21800 [M+H]⁺; found 416.21783.

Ac-LAF-EK

This compound was obtained by the general protocol for azide coupling on a 100 μ mol scale. Purification by column chromatography (0 $\rightarrow$ 3% MeOH in DCM) provided the title compound (32.69 mg, 76%) as a white powder after lyophilisation. ^1H NMR (400 MHz, CDCl_3 , MeOD) δ 7.33 – 7.15 (m, 5H), 4.79 (dd, J = 8.2, 5.0 Hz, 1H), 4.44 – 4.25 (m, 2H), 3.31 (d, J = 4.9 Hz, 1H), 3.10 (dd, J = 13.9, 4.9 Hz, 1H), 2.92 (d, J = 4.9 Hz, 1H), 2.79 (dd, J = 13.9, 8.3 Hz, 1H), 1.99 (s, 3H), 1.69 – 1.53 (m, 1H), 1.52 – 1.40 (m, 5H), 1.28 (d, J = 7.1 Hz, 3H), 0.92 (dd, J = 11.8, 6.5 Hz, 6H). ^{13}C NMR (101 MHz, CDCl_3 , MeOD) δ 207.72, 172.75, 172.36, 171.32, 135.85, 129.18, 128.40, 126.95, 59.21, 52.79, 52.37, 51.42, 48.58, 40.84, 36.79, 24.59, 22.77, 22.42, 21.52, 17.44, 16.32. LC-MS (linear gradient 10 $\rightarrow$ 90% MeCN, 0.1% TFA, 12.5 min): R_t (min): 6.31 (ESI-MS (m/z): 432.07 ($\text{M}+\text{H}^+$)). HRMS: calculated for $\text{C}_{23}\text{H}_{33}\text{N}_3\text{O}_5$ 432.24930 [$\text{M}+\text{H}$] $^+$; found 432.24902.

Ac-PAY-EK

This compound was obtained by the general protocol for azide coupling on a 100 μ mol scale. Purification by column chromatography (0 $\rightarrow$ 5% MeOH in DCM) provided the title compound (13.22 mg, 31%) as a white powder after lyophilisation. Complex NMR due to presence of rotamers (7:1). Peaks of major rotamer are reported. ^1H NMR (400 MHz, CDCl_3) δ 7.30 (d, J = 7.1 Hz, 1H), 6.96 (t, J = 8.5 Hz, 3H), 6.72 (d, J = 8.5 Hz, 2H), 4.74 (td, J = 8.1, 4.5 Hz, 1H), 4.48 – 4.38 (m, 1H), 4.33 (t, J = 7.1 Hz, 1H), 3.63 – 3.48 (m, 1H), 3.43 (q, J = 9.3, 8.9 Hz, 1H), 3.30 (d, J = 4.9 Hz, 1H), 3.08 (dd, J = 14.1, 4.4 Hz, 1H), 2.91 (d, J = 4.9 Hz, 1H), 2.66 (dd, J = 14.1, 8.5 Hz, 1H), 2.23 – 2.13 (m, 1H), 2.08 (s, 4H), 2.08 – 1.84 (m, 5H), 1.51 (s, 3H), 1.21 (d, J = 7.1 Hz, 3H). ^{13}C NMR (101 MHz, CDCl_3) δ 207.70, 172.22, 171.78, 171.34, 155.64, 130.52, 127.20, 115.93, 115.73, 59.93, 59.41, 53.25, 52.73, 49.00, 48.54, 36.15, 28.29, 25.11, 22.56, 17.62, 16.80. LC-MS (linear gradient 10 $\rightarrow$ 90% MeCN, 0.1% TFA, 12.5 min): R_t (min): 4.70 (ESI-MS (m/z): 432.07 ($\text{M}+\text{H}^+$)). HRMS: calculated for $\text{C}_{22}\text{H}_{29}\text{N}_3\text{O}_6$ 432.21291 [$\text{M}+\text{H}$] $^+$; found 432.21283.

Ac-LAY-EK

This compound was obtained by the general protocol for azide coupling on a 100 μ mol scale. Purification by column chromatography (0 $\rightarrow$ 5% MeOH in DCM) provided the title compound (22.15 mg, 50.0%) as a white powder after lyophilisation. Complex NMR due to presence of rotamers (7:1). ^1H NMR (400 MHz, CDCl_3) δ 8.11–8.05 (m, 2H), 7.04 (d, J = 8.4 Hz, 1H), 7.04 (d, J = 8.4 Hz, 1H), 6.71–6.68 (m, 2H), 4.87–4.67 (m, 1H), 4.36–4.31 (m, 2H), 3.24 (d, J = 5.2 Hz, 1H), 2.98 (dd, J = 14.0, 4.4 Hz, 1H), 2.91 (d, J = 8.4 Hz, 1H), 1.96 (s, 3H), 1.66–1.63 (m, 1H), 1.53–1.50 (m, 2H), 1.42 (s, 3H), 1.28 (d, J = 7.2 Hz, 3H), 0.92 (dd, J = 16.0, 6.8 Hz, 6H). ^{13}C NMR (101 MHz, MeOD) δ 208.97, 174.71, 174.62, 173.46, 157.36, 131.31, 128.55, 116.22, 60.24, 54.95, 54.86, 53.22, 53.10, 41.81, 36.71, 25.89, 23.46, 22.44, 21.86, 18.11, 16.81. LC-MS (linear gradient 10 $\rightarrow$ 90% MeCN, 0.1% TFA, 12.5 min): R_t (min): 5.36 (ESI-MS (m/z): 448.13 ($\text{M}+\text{H}^+$)). HRMS: calculated for $\text{C}_{23}\text{H}_{33}\text{N}_3\text{O}_6$ 448.24421 [$\text{M}+\text{H}$] $^+$; found 448.24411.

Ac-PAI-EK

This compound was obtained by the general protocol for azide coupling on a 100 μ mol scale. Purification by column chromatography (0 $\rightarrow$ 3% MeOH in DCM) provided the title compound (27.50 mg, 72%) as a white powder after lyophilisation. Complex NMR due to presence of rotamers (8:1). Peaks of major rotamer are reported. ^1H NMR (400 MHz, CDCl_3) δ 7.40 (d, J = 6.9 Hz, 1H), 6.79 (d, J = 8.4 Hz, 1H), 4.55 – 4.42 (m, 2H), 4.42 – 4.20 (m, 1H), 3.57 (td, J = 8.9, 8.1, 3.1 Hz, 1H), 3.43 (td, J = 9.4, 7.0 Hz, 1H), 3.30 (d, J = 5.0 Hz, 1H), 2.84 (d, J = 5.0 Hz, 1H), 2.34 – 2.23 (m, 1H), 2.13 – 2.02 (m, 4H), 2.03 – 1.73 (m, 3H), 1.47 (s, 3H), 1.43 – 1.33 (m, 1H), 1.29 (d, J = 7.1 Hz, 3H), 1.15 – 0.99 (m, 1H), 0.92 (dd, J = 6.8, 3.5 Hz, 3H), 0.85 (t, J = 7.4 Hz, 3H). ^{13}C NMR (101 MHz, CDCl_3) δ 209.36, 172.23, 171.66, 171.05, 59.74, 59.58, 55.11, 52.04, 49.24, 48.44, 37.13, 27.92, 25.15, 24.43, 22.60, 17.28, 16.50, 15.95, 11.34. LC-MS (linear gradient 10 $\rightarrow$ 90% MeCN, 0.1% TFA, 12.5 min): R_t (min): 5.17 (ESI-MS (m/z): 382.13 ($\text{M}+\text{H}^+$)). HRMS: calculated for $\text{C}_{19}\text{H}_{31}\text{N}_3\text{O}_5$ 382.23365 [$\text{M}+\text{H}$] $^+$; found 382.23367.

Ac-LAI-EK

This compound was obtained by the general protocol for azide coupling on a 100 μ mol scale. Purification by column chromatography (0 $\rightarrow$ 3% MeOH in DCM) provided the title compound (27.50 mg, 72%) as a white powder after lyophilisation. ^1H NMR (400 MHz, MeOD) δ 4.46 (d, J = 7.0 Hz, 1H), 4.43 – 4.29 (m, 2H), 3.27 (d, J = 5.1 Hz, 1H), 2.91 (d, J = 5.1 Hz, 1H), 1.97 (s, 3H), 1.83 (dtq, J = 13.8, 6.9, 3.5 Hz, 1H), 1.66 (dq, J = 13.2, 6.6 Hz, 1H), 1.55 (t, J = 7.3 Hz, 2H), 1.52 – 1.39 (m, 4H), 1.29 (d, J = 7.1 Hz, 3H), 1.24 – 1.05 (m, 1H), 1.01 – 0.83 (m, 12H). ^{13}C NMR (101 MHz, MeOD) δ 210.59, 174.72, 174.68, 173.29, 60.42, 56.14, 53.04, 52.40, 50.02, 41.96, 38.19, 25.89, 25.71, 23.46, 22.39, 21.93, 17.82, 16.53, 16.05, 11.37. LC-MS (linear gradient 10 $\rightarrow$ 90% MeCN, 0.1% TFA, 12.5 min): R_t (min): 6.07 (ESI-MS (m/z): 398.20 ($\text{M}+\text{H}^+$)). HRMS: calculated for $\text{C}_{20}\text{H}_{35}\text{N}_3\text{O}_5$ 398.26495 [$\text{M}+\text{H}^+$] $^+$; found 398.26486.

Ac-PAL-EK

This compound was obtained by the general protocol for azide coupling on a 100 μ mol scale. Purification by column chromatography (0 $\rightarrow$ 3% MeOH in DCM) provided the title compound (14.35 mg, 38%) as a white powder after lyophilisation. Complex NMR due to presence of rotamers (8:1). Peaks of major rotamer are reported. ^1H NMR (400 MHz, CDCl_3) δ 7.34 (d, J = 7.0 Hz, 1H), 6.70 (d, J = 7.0 Hz, 1H), 4.53 (ddd, J = 14.8, 10.0, 5.0 Hz, 2H), 4.35 (t, J = 7.1 Hz, 1H), 3.58 (td, J = 9.1, 8.2, 3.2 Hz, 1H), 3.50 – 3.35 (m, 1H), 3.33 (d, J = 5.0 Hz, 1H), 2.87 (d, J = 5.0 Hz, 1H), 2.36 – 2.24 (m, 1H), 2.10 (s, 4H), 2.05 – 1.74 (m, 3H), 1.62 (dd, J = 6.2, 3.7 Hz, 1H), 1.50 (d, J = 5.1 Hz, 3H), 1.31 (dd, J = 9.9, 5.6 Hz, 4H), 1.32 – 1.21 (m, 3H), 0.92 (d, J = 6.5 Hz, 6H). ^{13}C NMR (101 MHz, CDCl_3) δ 208.55, 172.25, 171.51, 171.14, 59.82, 59.23, 52.60, 50.39, 49.14, 48.48, 40.00, 27.87, 27.85, 25.32, 25.20, 23.48, 22.65, 21.41, 17.36, 16.87. LC-MS (linear gradient 10 $\rightarrow$ 90% MeCN, 0.1% TFA, 12.5 min): R_t (min): 5.21 (ESI-MS (m/z): 382.13 ($\text{M}+\text{H}^+$)). HRMS: calculated for $\text{C}_{19}\text{H}_{31}\text{N}_3\text{O}_5$ 382.23365 [$\text{M}+\text{H}^+$] $^+$; found 382.23367.

Ac-LAL-EK

This compound was obtained by the general protocol for azide coupling on a 100 μ mol scale. Purification by column chromatography (0 $\rightarrow$ 3% MeOH in DCM) provided the title compound (15.49 mg, 81%) as a white powder after lyophilisation. ^1H NMR (400 MHz, CDCl_3) δ 7.66 (dd, J = 21.3, 8.1 Hz, 2H), 7.14 (d, J = 8.8 Hz, 1H), 4.77 (dp, J = 15.1, 8.1, 7.5 Hz, 2H), 4.69 – 4.53 (m, 1H), 3.29 (d, J = 5.0 Hz, 1H), 2.88 (d, J = 5.0 Hz, 1H), 2.02 (s, 3H), 1.69 – 1.43 (m, 8H), 1.33 (dd, J = 10.2, 4.2 Hz, 1H), 1.29 (d, J = 7.0 Hz, 3H), 0.88 (t, J = 6.4 Hz, 12H). ^{13}C NMR (101 MHz, CDCl_3) δ 208.99, 172.63, 172.44, 170.21, 59.14, 52.40, 51.39, 50.01, 48.51, 42.48, 40.00, 25.20, 24.85, 23.44, 23.14, 22.87, 22.54, 21.40, 18.54, 16.80. LC-MS (linear gradient 10 $\rightarrow$ 90% MeCN, 0.1% TFA, 12.5 min): R_t (min): 6.08 (ESI-MS (m/z): 398.13 ($\text{M}+\text{H}^+$)). HRMS: calculated for $\text{C}_{20}\text{H}_{35}\text{N}_3\text{O}_5$ 398.26495 [$\text{M}+\text{H}^+$] $^+$; found 398.26428.

Ac-PAV-EK

This compound was obtained by the general protocol for azide coupling on a 100 μ mol scale. Purification by column chromatography (0 $\rightarrow$ 3% MeOH in DCM) provided the title compound (6.95 mg, 19%) as a white powder after lyophilisation. Complex NMR due to presence of rotamers (8:1). Peaks of major rotamer are reported. ^1H NMR (400 MHz, CDCl_3) δ 7.50 (d, J = 6.8 Hz, 1H), 6.77 (d, J = 8.4 Hz, 1H), 4.53 (dd, J = 8.0, 2.4 Hz, 1H), 4.47 (dd, J = 8.5, 5.6 Hz, 1H), 4.32 (p, J = 7.1 Hz, 1H), 3.59 (ddd, J = 11.4, 8.5, 4.4 Hz, 1H), 3.44 (td, J = 9.4, 7.1 Hz, 1H), 3.30 (d, J = 4.9 Hz, 1H), 2.86 (d, J = 5.0 Hz, 1H), 2.37 – 2.29 (m, 1H), 2.16 – 2.04 (m, 4H), 2.03 – 1.82 (m, 3H), 1.49 (s, 3H), 1.34 (d, J = 7.1 Hz, 3H), 0.96 (d, J = 6.8 Hz, 3H), 0.83 (d, J = 6.9 Hz, 3H). ^{13}C NMR (101 MHz, CDCl_3) δ 209.20, 172.33, 171.69, 171.26, 59.82, 59.52, 55.90, 52.17, 49.49, 48.52, 30.43, 27.84, 25.23, 22.69, 19.82, 17.53, 17.18, 16.54. LC-MS (linear gradient 10 $\rightarrow$ 90% MeCN, 0.1% TFA, 12.5 min): R_t (min): 4.72 (ESI-MS (m/z): 368.13 ($\text{M}+\text{H}^+$)). HRMS: calculated for $\text{C}_{18}\text{H}_{29}\text{N}_3\text{O}_5$ 368.21800 [$\text{M}+\text{H}^+$] $^+$; found 368.21811.

Ac-LAV-EK

This compound was obtained by the general protocol for azide coupling on a 100 μ mol scale. Purification by column chromatography (0 $\rightarrow$ 3% MeOH in DCM) provided the title compound (25.68 mg, 67%) as a white powder after lyophilisation. ^1H NMR (400 MHz, CDCl_3 , MeOD) δ 4.39 (d, J = 5.5 Hz, 1H), 4.37 – 4.23 (m, 2H), 3.19 (d, J =

4.8 Hz, 1H), 2.83 (d, J = 4.9 Hz, 1H), 2.14 – 1.97 (m, 1H), 1.92 (s, 3H), 1.61 – 1.35 (m, 6H), 1.24 (d, J = 7.0 Hz, 3H), 0.95 – 0.81 (m, 9H), 0.78 (d, J = 6.8 Hz, 3H). ^{13}C NMR (101 MHz, CDCl_3 , MeOD) δ 208.81, 172.93, 172.62, 171.39, 59.25, 55.81, 51.74, 51.40, 48.60, 40.89, 30.03, 24.53, 22.56, 22.21, 21.46, 19.27, 17.03, 16.94, 16.08. LC-MS (linear gradient 10 $\rightarrow$ 90% MeCN, 0.1% TFA, 12.5 min): R_t (min): 5.57 (ESI-MS (m/z): 384.13 ($\text{M}+\text{H}^+$)). HRMS: calculated for $\text{C}_{19}\text{H}_{33}\text{N}_3\text{O}_5$ 384.24930 [$\text{M}+\text{H}$] $^+$; found 384.24939.

Ac-PAA-EK

This compound was obtained by the general protocol for azide coupling on a 100 μmol scale. Purification by column chromatography (0 $\rightarrow$ 5% MeOH in DCM) provided the title compound (10.43 mg, 31%) as a white powder after lyophilisation. Complex NMR due to presence of rotamers (9:1). Peaks of major rotamer are reported. ^1H NMR (400 MHz, CDCl_3) δ 7.33 (d, J = 7.0 Hz, 1H), 6.81 (d, J = 6.6 Hz, 1H), 4.50 (dtd, J = 14.3, 7.8, 7.2, 2.3 Hz, 2H), 4.43 – 4.23 (m, 1H), 3.58 (ddd, J = 11.2, 8.1, 3.2 Hz, 1H), 3.44 (td, J = 9.4, 7.0 Hz, 1H), 3.26 (d, J = 5.0 Hz, 1H), 2.87 (d, J = 5.0 Hz, 1H), 2.33 (ddd, J = 12.3, 6.3, 3.0 Hz, 1H), 2.10 (s, 3H), 2.11 – 1.82 (m, 3H), 1.50 (s, 3H), 1.40 – 1.22 (m, 6H). ^{13}C NMR (101 MHz, CDCl_3) δ 208.37, 171.81, 171.44, 171.15, 59.90, 59.86, 59.13, 52.63, 49.15, 49.09, 49.06, 48.50, 48.05, 47.95, 27.91, 25.20, 22.67, 17.60, 17.56, 17.12, 17.08, 16.94. LC-MS (linear gradient 10 $\rightarrow$ 50% MeCN, 0.1% TFA, 12.5 min): R_t (min): 4.55 (ESI-MS (m/z): 340.07 ($\text{M}+\text{H}^+$)). HRMS: calculated for $\text{C}_{16}\text{H}_{25}\text{N}_3\text{O}_5$ 340.18670 [$\text{M}+\text{H}$] $^+$; found 340.18677.

Ac-LAA-EK

This compound was obtained by the general protocol for azide coupling on a 100 μmol scale. Purification by column chromatography (0 $\rightarrow$ 5% MeOH in DCM) provided the title compound (22.46 mg, 63%) as a white powder after lyophilisation. ^1H NMR (400 MHz, MeOD) δ 4.48 – 4.26 (m, 3H), 3.25 (d, J = 5.1 Hz, 1H), 2.93 (d, J = 5.1 Hz, 1H), 1.97 (s, 3H), 1.71 – 1.59 (m, 1H), 1.58 – 1.53 (m, 2H), 1.47 (s, 3H), 1.33 (d, J = 7.1 Hz, 3H), 1.27 (d, J = 7.2 Hz, 3H), 0.94 (dd, J = 14.3, 6.5 Hz, 6H). ^{13}C NMR (101 MHz, MeOD) δ 209.49, 174.65, 174.50, 173.39, 60.07, 53.19, 53.16, 49.86, 49.35, 41.83, 25.90, 23.45, 22.40, 21.90, 18.07, 17.05, 16.29. LC-MS (linear gradient 10 $\rightarrow$ 90% MeCN, 0.1% TFA, 12.5 min): R_t (min): 4.77 (ESI-MS (m/z): 356.07 ($\text{M}+\text{H}^+$)). HRMS: calculated for $\text{C}_{17}\text{H}_{29}\text{N}_3\text{O}_5$ 356.21800 [$\text{M}+\text{H}$] $^+$; found 356.21805.

Ac-PLL-EK

This compound was obtained by the general protocol for azide coupling on a 100 μmol scale. Purification by column chromatography (0 $\rightarrow$ 2% MeOH/DCM) provided the title compound (24.4 mg, 58%) as a white powder after lyophilisation. Complex NMR due to presence of rotamers (7:1). Peaks of major rotamer are reported. ^1H NMR (400 MHz, CDCl_3) δ 7.24 (d, J = 7.6 Hz, 1H), 6.66 (d, J = 7.9 Hz, 1H), 4.54 (ddt, J = 7.8, 5.4, 3.3 Hz, 2H), 4.36 – 4.20 (m, 1H), 3.55 (ddd, J = 11.3, 8.3, 4.1 Hz, 1H), 3.43 (td, J = 9.4, 7.1 Hz, 1H), 3.31 (d, J = 5.0 Hz, 1H), 2.85 (d, J = 5.0 Hz, 1H), 2.31 (ddd, J = 12.2, 6.3, 3.1 Hz, 1H), 2.08 (s, 3H), 2.06 – 1.83 (m, 3H), 1.71 – 1.44 (m, 8H), 1.29 (ddd, J = 14.0, 10.4, 4.4 Hz, 1H), 0.94 – 0.82 (m, 12H). ^{13}C NMR (101 MHz, CDCl_3) δ 208.53, 172.09, 171.60, 171.10, 59.78, 59.19, 52.56, 52.19, 50.34, 48.41, 40.33, 40.05, 27.70, 25.29, 25.19, 25.00, 23.46, 23.01, 22.51, 21.99, 21.42, 16.86. LC-MS (linear gradient 10 $\rightarrow$ 90% MeCN, 0.1% TFA, 12.5 min): R_t (min): 6.30 (ESI-MS (m/z): 424.13 ($\text{M}+\text{H}^+$)). HRMS: calculated for $\text{C}_{22}\text{H}_{37}\text{N}_3\text{O}_5$ 424.28060 [$\text{M}+\text{H}$] $^+$; found 424.28040.

Ac-LLL-EK

This compound was obtained by the general protocol for azide coupling on a 100 μmol scale. Purification by column chromatography (0 $\rightarrow$ 2% MeOH/DCM) provided the title compound (28.55 mg, 65%) as a white powder after lyophilisation. ^1H NMR (400 MHz, CDCl_3 , MeOD) δ 4.49 (dd, J = 10.6, 3.0 Hz, 1H), 4.35 (dt, J = 8.7, 6.1 Hz, 2H), 3.25 (d, J = 5.0 Hz, 1H), 2.86 (d, J = 5.0 Hz, 1H), 1.93 (s, 3H), 1.68 – 1.35 (m, 11H), 1.34 – 1.14 (m, 1H), 0.97 – 0.76 (m, 18H). ^{13}C NMR (101 MHz, MeOD) δ 208.75, 173.09, 172.53, 171.43, 59.25, 52.52, 51.73, 51.53, 50.51, 41.10, 40.54, 39.62, 25.24, 24.84, 24.70, 23.33, 22.82, 22.76, 22.52, 21.90, 21.12, 16.77. LC-MS (linear gradient

10 → 90% MeCN, 0.1% TFA, 12.5 min): R_t (min): 7.05 (ESI-MS (m/z): 440.20 ($M+H^+$)). HRMS: calculated for $C_{23}H_{41}N_3O_5$ 440.31190 [$M+H$] $^+$; found 440.31171.

Boc-APLL-EK

This compound was obtained by the general protocol for azide coupling on a 100 μ mol scale. Purification by column chromatography (0→2% MeOH/DCM) provided the title compound (yield n.d., all material used in the next step). Complex NMR due to presence of rotamers (3.5:1). 1H NMR (400 MHz, $CDCl_3$) δ 7.62 (d, J = 6.9 Hz, 0.3H), 7.14 (d, J = 7.7 Hz, 0.7H), 6.88 (d, J = 9.2 Hz, 0.3H), 6.59 (d, J = 7.8 Hz, 0.7H), 5.31 (d, J = 7.6 Hz, 0.7H), 5.22 (d, J = 7.4 Hz, 0.3H), 4.70 – 4.09 (m, 4H), 3.70 – 3.60 (m, 1H), 3.57 – 3.46 (m, 1H), 3.30 (d, J = 5.0 Hz, 0.8H), 3.24 (d, J = 4.8 Hz, 0.2H), 2.85 (d, J = 5.1 Hz, 0.8H), 2.83 (d, J = 4.9 Hz, 0.2H), 2.55 (dd, J = 12.3, 6.3 Hz, 0.2), 2.34 – 1.83 (m, 4.8H), 1.74 – 1.44 (m, 8H), 1.41 (s, 8H), 1.37 – 1.22 (m, 4H), 0.96 – 0.79 (m, 12H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 209.43, 208.51, 173.36, 172.64, 172.05, 171.12, 155.31, 80.99, 79.97, 77.48, 77.16, 76.84, 60.87, 60.06, 59.19, 54.87, 52.55, 52.44, 51.87, 50.32, 49.41, 49.24, 48.22, 47.39, 46.76, 40.81, 40.55, 40.10, 39.11, 31.59, 28.45, 27.22, 25.36, 25.26, 24.76, 23.45, 23.35, 23.14, 22.96, 22.13, 21.79, 21.69, 21.41, 18.46, 16.84, 16.71, 16.54. LC-MS (linear gradient 10 → 90% MeCN, 0.1% TFA, 12.5 min): R_t (min): 7.65 (ESI-MS (m/z): 553.13 ($M+H^+$)). HRMS: calculated for $C_{28}H_{48}N_4O_7$ 553.35958 [$M+H$] $^+$; found 553.35925.

TFA-H-APLL-EK

Boc-AlaProLeuLeu-EK was treated with 1:1 TFA/DCM. After stirring for 30 min the reaction mixture was concentrated and co-evaporated with toluene (2x). HPLC-purification (C_{18} 15-50% MeCN, 0.1 % TFA, 10 min gradient) provided the title compound (6.90 mg, 12%) as a white powder after lyophilisation. Complex NMR due to presence of rotamers (7:1). Peaks of major rotamer are reported. 1H NMR (600 MHz, MeOD) δ 4.53 (dd, J = 10.9, 2.9 Hz, 1H), 4.50 (dd, J = 8.6, 4.9 Hz, 1H), 4.37 (dd, J = 9.4, 5.8 Hz, 1H), 4.22 (q, J = 6.9 Hz, 1H), 3.71 – 3.55 (m, 2H), 3.23 (d, J = 5.1 Hz, 1H), 2.94 (d, J = 5.0 Hz, 1H), 2.28 – 2.17 (m, 1H), 2.14 – 2.03 (m, 1H), 2.03 – 1.93 (m, 2H), 1.77 – 1.66 (m, 2H), 1.59 – 1.52 (m, 2H), 1.50 (d, J = 7.0 Hz, 3H), 1.47 (s, 3H), 1.41 – 1.31 (m, 2H), 1.02 – 0.86 (m, 12H). ^{13}C NMR (151 MHz, MeOD) δ 209.58, 174.77, 173.72, 169.51, 61.27, 61.15, 60.11, 53.34, 53.03, 53.01, 52.00, 51.88, 49.43, 49.28, 49.14, 49.00, 48.86, 48.72, 48.57, 42.01, 40.08, 30.46, 26.34, 26.32, 26.03, 25.73, 23.74, 23.45, 22.09, 21.43, 17.04, 16.13. LC-MS (linear gradient 10 → 90% MeCN, 0.1% TFA, 12.5 min): R_t (min): 5.29 (ESI-MS (m/z): 453.20 ($M+H^+$)). HRMS: calculated for $C_{23}H_{40}N_4O_5$ 453.30715 [$M+H$] $^+$; found 453.30704.

Ac-APLL-EK

This compound was obtained by the general protocol for azide coupling on a 100 μ mol scale. Purification by column chromatography (0→2% MeOH/DCM) provided the title compound (32.01 mg, 65%) as a white powder after lyophilisation. Complex NMR due to presence of rotamers (3:1). Peaks of major rotamer are reported for 1H , all peaks are reported for ^{13}C . 1H NMR (400 MHz, $CDCl_3$) δ 7.19 – 7.08 (m, 1H), 6.62 (s, 1H), 6.54 (s, 1H), 4.74 (p, J = 6.9 Hz, 1H), 4.59 – 4.51 (m, 2H), 4.42 – 4.31 (m, 1H), 3.68 (q, J = 8.7, 8.0 Hz, 1H), 3.55 (td, J = 10.1, 9.1, 4.1 Hz, 1H), 3.30 (d, J = 5.0 Hz, 1H), 2.86 (d, J = 5.0 Hz, 1H), 2.25 (d, J = 8.4 Hz, 1H), 2.19 – 2.00 (m, 3H), 1.98 (s, 3H), 1.91 (ddt, J = 12.4, 9.6, 5.9 Hz, 1H), 1.72 – 1.50 (m, 4H), 1.49 (d, J = 4.7 Hz, 3H), 1.43 (s, 1H), 1.34 – 1.29 (m, 2H), 1.29 – 1.21 (m, 1H), 0.96 – 0.81 (m, 12H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 209.55, 208.67, 172.98, 172.50, 172.10, 172.01, 171.03, 170.75, 169.67, 60.99, 59.96, 59.19, 54.68, 52.55, 52.42, 51.83, 50.34, 48.95, 48.86, 47.48, 46.91, 46.79, 40.81, 40.38, 40.23, 40.13, 31.82, 27.39, 25.39, 25.27, 25.24, 25.15, 24.76, 23.46, 23.21, 23.08, 22.85, 22.13, 21.99, 21.83, 21.42, 18.33, 16.83, 16.64, 16.30. LC-MS (linear gradient 10 → 90% MeCN, 0.1% TFA, 12.5 min): R_t (min): 6.01 (ESI-MS (m/z): 495.13 ($M+H^+$)). HRMS: calculated for $C_{25}H_{42}N_4O_6$ 495.31771 [$M+H$] $^+$; found 495.31741.

N_3 -GAPAL-EK

This compound was obtained by the general protocol for azide coupling on a 240 μ mol scale. Purification by column chromatography (0→4% MeOH in DCM) provided the title compound (39.0 mg, 76%) as a white powder

after lyophilisation. Complex NMR due to presence of rotamers (4:1). Peaks of major rotamer are reported. ^1H NMR (400 MHz, CDCl_3) δ 7.22 (d, J = 7.5 Hz, 1H), 7.12 (d, J = 7.4 Hz, 1H), 6.67 (d, J = 7.8 Hz, 1H), 4.79 – 4.70 (m, 1H), 4.60 – 4.49 (m, 2H), 4.46 – 4.38 (m, 1H), 3.95 (d, J = 1.9 Hz, 2H), 3.73 – 3.52 (m, 2H), 3.29 (d, J = 5.0 Hz, 1H), 2.88 (d, J = 5.0 Hz, 1H), 2.26 – 1.91 (m, 6H), 1.61 (dd, J = 6.5, 2.4 Hz, 1H), 1.49 (s, 3H), 1.38 (d, J = 6.9 Hz, 3H), 1.28 (d, J = 7.0 Hz, 3H), 0.92 (dd, J = 6.5, 3.1 Hz, 6H). ^{13}C NMR (101 MHz, CDCl_3) δ 208.74, 172.15, 172.03, 170.99, 166.31, 60.09, 59.20, 52.57, 52.52, 50.39, 48.83, 47.44, 46.82, 40.08, 27.95, 25.28, 25.20, 23.46, 21.35, 18.25, 18.01, 16.83. LC-MS (linear gradient 10 $\rightarrow$ 90% MeCN, 0.1% TFA, 12.5 min): Rt (min): 5.57 (ESI-18), MS (m/z): 494.13 $[\text{M}+\text{H}]^+$. HRMS: calculated for $\text{C}_{22}\text{H}_{35}\text{N}_7\text{O}_6$ 494.27216 $[\text{M}+\text{H}]^+$; found 494.27191.

Synthesis of warheads

Standard procedures amino acid epoxyketone synthesis

Boc-AA-C(CH_3)=CH₂

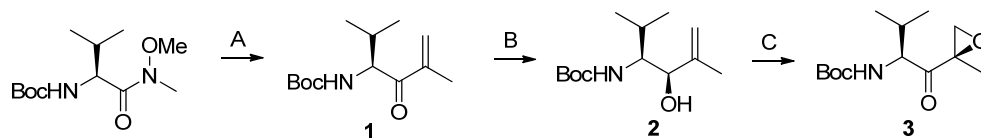
To a solution of 2-bromopropene (3 equiv.) in Et_2O at -78°C is added $t\text{BuLi}$ (4.5 equiv, from 1.7 M in pent) in 10 min. After stirring for 15 min. at -78°C , the Weinreb amide (1 equiv.) in Et_2O is added slowly in 10 min. The reaction mixture is stirred for 2–4 h, while warming up to max. -40°C . After TLC analysis revealed completion of the reaction, the reaction is quenched by the addition of sat. NH_4Cl and warmed to RT. The mixture is transferred to a separatory funnel and the water layer is extracted with EtOAc (3X). The combined organic layers are washed with brine, dried over Na_2SO_4 , filtered and concentrated. The crude product is purified by column chromatography (EtOAc /pent mixtures).

Boc-AA-OH-C(CH_3)=CH₂

To a solution of alkene **A** (1 equiv.) in MeOH is added $\text{CeCl}_3 \cdot 7\text{H}_2\text{O}$ (1.6 equiv.) and the mixture is stirred at RT. After the solution became clear, the mixture is cooled to 0°C and NaBH_4 (1.3 equiv.) is added in portion in 10 min. After TLC analysis showed completion of the reaction (about 30 min), the reaction is quenched by the addition of AcOH. The mixture is stirred for 15 min. followed by the addition of toluene and removal of the solvent. The residue is redissolved in a H_2O / EtOAc mixture, which is then transferred to a separatory funnel. The layers were separated and the aqueous layer was extracted with EtOAc (2X). The combined organic layers are washed with brine, dried over Na_2SO_4 , filtered and concentrated. The crude product was purified (if necessary) by column chromatography (EtOAc /pent mixtures).

Boc-AA-EK

To a solution of alcohol **B** in DCM at 0°C is added $\text{VO}(\text{acac})_2$ (0.1 equiv.) followed by the addition of $t\text{BuOOH}$ (5.5 M in decane, 3 equiv.). The reaction mixture is stirred at 0°C for 2–3 h. after which TLC analysis showed completion of the reaction. The reaction mixture is concentrated, redissolved in EtOAc and washed with 0.5 sat. NaHCO_3 (2x), H_2O and brine. The organic layer is dried over Na_2SO_4 , filtered and concentrated. The crude product is added as a solution in DCM to a solution of Dess-Martin-Periodane (1.5–3 equiv.) in DCM at 0°C . After TLC analysis revealed completion of the reaction, the reaction was quenched by the addition of sat. NaHCO_3 . The mixture was transferred to a separatory funnel and the layers were separated. The aqueous layer was extracted with DCM (1x) and the combined organics were washed with sat. NaHCO_3 (1x) and brine and dried over Na_2SO_4 , filtered and concentrated. The crude product was purified by column chromatography (EtOAc /pent mixtures).

Boc-Val-EK**(S)-tert-butyl (2,5-dimethyl-4-oxohex-5-en-3-yl)carbamate (1)**

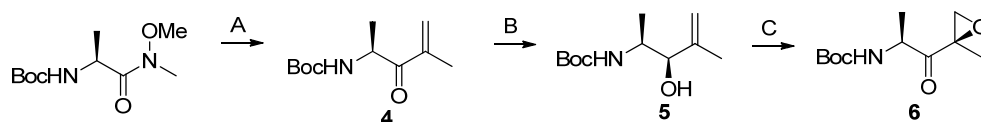
This compound was synthesized according to the general procedure **A** described above on a 10 mmol scale and was isolated after column chromatography (10% EtOAc:pent) (808 mg, 3.33 mmol, 33%). ^1H NMR (400 MHz, CDCl_3) δ 6.04 (s, 1H), 5.96 – 5.78 (m, 1H), 5.25 (d, J = 8.6 Hz, 1H), 4.90 (dd, J = 9.1, 4.7 Hz, 1H), 2.12 – 1.94 (m, 1H), 1.87 (s, 3H), 1.40 (s, 9H), 0.94 (d, J = 6.8 Hz, 3H), 0.74 (d, J = 6.8 Hz, 3H). ^{13}C NMR (101 MHz, CDCl_3) δ 201.41, 155.88, 143.05, 79.49, 58.63, 32.07, 28.37, 19.92, 17.76, 16.99.

tert-butyl ((3S,4R)-4-hydroxy-2,5-dimethylhex-5-en-3-yl)carbamate (2)

This compound was synthesized according to the general procedure **B** described above on a 3.33 mmol scale. The crude compound (quantitative yield) was not purified by column chromatography but used crude in procedure **C**. ^1H NMR (400 MHz, CDCl_3) δ 4.96 (s, 1H), 4.91 (s, 1H), 4.49 (d, J = 9.7 Hz, 1H), 4.04 (d, J = 6.7 Hz, 1H), 3.74 – 3.57 (m, 1H), 2.30 – 2.10 (m, 1H), 2.10 – 1.90 (m, 1H), 1.76 (s, 3H), 1.41 (s, 9H), 0.94 (d, J = 6.9 Hz, 3H), 0.86 (d, J = 6.8 Hz, 3H). ^{13}C NMR (101 MHz, CDCl_3) δ 156.37, 145.71, 112.99, 79.22, 77.20, 56.84, 28.39, 27.60, 20.91, 18.17, 16.73.

tert-butyl ((S)-3-methyl-1-((R)-2-methyloxiran-2-yl)-1-oxobutan-2-yl)carbamate (3)

This compound was synthesized according to the general procedure **D** described above on a 4.44 mmol scale and was isolated after column chromatography (5→20% EtOAc:pent) (485 mg, 1.89 mmol, 57%). ^1H NMR (400 MHz, CDCl_3) δ 4.96 (d, J = 9.2 Hz, 1H), 4.17 (dd, J = 9.4, 5.0 Hz, 1H), 3.18 (d, J = 4.9 Hz, 1H), 2.81 (d, J = 5.0 Hz, 1H), 2.01 (dq, J = 13.4, 6.8 Hz, 1H), 1.44 (s, 3H), 1.34 (s, 9H), 0.93 (d, J = 6.8 Hz, 3H), 0.76 (d, J = 6.9 Hz, 3H). ^{13}C NMR (101 MHz, CDCl_3) δ 209.73, 155.75, 79.64, 59.21, 56.99, 51.86, 30.15, 28.32, 19.78, 17.05, 16.47. $[\alpha]_D^{20}$ = 150.6 (C=1, CHCl_3)

Boc-Ala-EK**(S)-tert-butyl (4-methyl-3-oxopent-4-en-2-yl)carbamate (4)**

This compound was synthesized according to the general procedure **A** described above on a 4 mmol scale. The crude product was used in the next step (791 mg, 3.71 mmol, 93%).

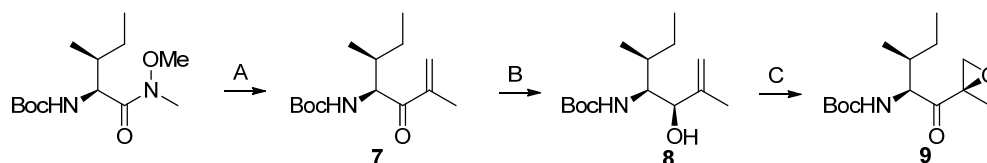
^1H NMR (400 MHz, CDCl_3) δ 6.01 (s, 1H), 5.92 – 5.78 (m, 1H), 5.45 – 5.31 (m, 1H), 4.99 (p, J = 7.1 Hz, 1H), 1.87 (s, 3H), 1.40 (s, 9H), 1.27 (d, J = 7.1 Hz, 3H). ^{13}C NMR (101 MHz, CDCl_3) δ 200.89, 155.00, 141.73, 126.08, 79.47, 50.04, 28.27, 20.04, 17.74.

tert-butyl ((2S,3R)-3-hydroxy-4-methylpent-4-en-2-yl)carbamate (5)

This compound was synthesized according to the general procedure **B** described above on a 3.71 mmol scale. The crude compound (quantitative yield) was not purified by column chromatography but used crude in procedure **C**. Equilibrating rotamers in NMR, peaks of major rotamer are reported. ^1H NMR (400 MHz, CDCl_3) δ 5.08 – 4.81 (m, 3H), 4.10 (s, 1H), 3.84 (d, J = 5.3 Hz, 1H), 1.72 (s, 3H), 1.42 (s, 9H), 1.00 (d, J = 6.8 Hz, 3H). ^{13}C NMR (101 MHz, CDCl_3) δ 156.65, 144.89, 111.17, 48.39, 28.39, 19.35, 13.80.

***tert*-butyl ((*S*)-1-((*R*)-2-methyloxiran-2-yl)-1-oxopropan-2-yl)carbamate (6)**

This compound was synthesized according to the general procedure **D** described above on a 3.71 mmol scale and was isolated after column chromatography (5→10% EtOAc:pent) (180 mg, 0.789 mmol, 21%). ¹H NMR (400 MHz, CDCl₃) δ 5.04 (d, *J* = 6.6 Hz, 1H), 4.27 (p, *J* = 7.2 Hz, 1H), 3.20 (d, *J* = 4.7 Hz, 1H), 2.86 (d, *J* = 5.0 Hz, 1H), 1.49 (s, 3H), 1.37 (s, 9H), 1.22 (d, *J* = 7.1 Hz, 3H). ¹³C NMR (101 MHz, CDCl₃) δ 209.20, 155.08, 79.65, 58.84, 52.27, 48.74, 28.24, 17.32, 16.73. [α]_D²⁰ = +111.0 (C=1, CHCl₃).

Boc-Ile-EK***tert*-butyl ((4*S*,5*S*)-2,5-dimethyl-3-oxohept-1-en-4-yl)carbamate (7)**

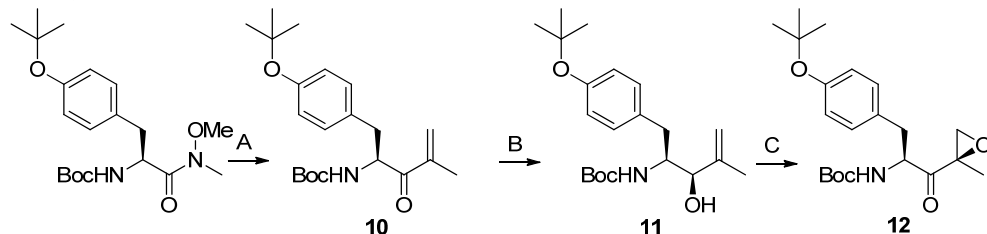
This compound was synthesized according to the general procedure **A** described above on a 9.7 mmol scale and was isolated after column chromatography (0→5% EtOAc:pent) (2.14 g, 8.4 mmol, 84%). ¹H NMR (400 MHz, CDCl₃) δ 6.01 (s, 1H), 5.81 (s, 1H), 5.20 (d, *J* = 8.8 Hz, 1H), 4.87 (dd, *J* = 9.0, 5.2 Hz, 1H), 1.83 (s, 3H), 1.76 – 1.62 (m, 1H), 1.36 (s, 9H), 1.29 – 1.18 (m, 1H), 1.05 – 0.90 (m, 1H), 0.86 (d, *J* = 6.8 Hz, 3H), 0.78 (t, *J* = 7.4 Hz, 3H). ¹³C NMR (101 MHz, CDCl₃) δ 201.76, 155.75, 143.29, 79.45, 58.19, 38.74, 28.34, 24.08, 17.75, 16.14, 11.53.

***tert*-butyl ((3*R*,4*S*,5*S*)-3-hydroxy-2,5-dimethylhept-1-en-4-yl)carbamate (8)**

This compound was synthesized according to the general procedure **B** described above on a 3.71 mmol scale. The crude compound (quantitative yield) was not purified by column chromatography but used crude in procedure **C**. ¹H NMR (300 MHz, CDCl₃) δ 4.91 (d, *J* = 13.3 Hz, 2H), 4.47 (d, *J* = 9.8 Hz, 1H), 4.08 (dd, *J* = 6.8, 2.3 Hz, 1H), 3.66 (dt, *J* = 10.7, 6.3 Hz, 1H), 2.45 (s, 1H), 1.74 (s, 4H), 1.66 – 1.48 (m, 1H), 1.39 (s, 9H), 1.06 – 0.72 (m, 6H). ¹³C NMR (75 MHz, CDCl₃) δ 156.45, 145.78, 79.27, 76.95, 57.26, 34.70, 28.43, 23.65, 18.27, 16.93, 11.82.

***tert*-butyl ((2*S*,3*S*)-3-methyl-1-((*R*)-2-methyloxiran-2-yl)-1-oxopentan-2-yl)carbamate (9)**

This compound was synthesized according to the general procedure **D** described above on a 8.4 mmol scale and was isolated after column chromatography (0→10% EtOAc:pent) (1.20 g, 4.4 mmol, 52%). ¹H NMR (300 MHz, CDCl₃) δ 4.92 (d, *J* = 9.0 Hz, 1H), 4.23 (dd, *J* = 9.4, 5.7 Hz, 1H), 3.24 (d, *J* = 4.8 Hz, 1H), 2.84 (d, *J* = 5.0 Hz, 1H), 1.85 – 1.63 (m, 1H), 1.48 (s, 3H), 1.38 (s, 9H), 1.12 – 0.74 (m, 2H), 0.94 (d, *J* = 6.8 Hz, 3H), 0.85 (t, *J* = 7.3 Hz, 3H). ¹³C NMR (75 MHz, CDCl₃) δ 210.29, 155.69, 79.72, 59.36, 56.41, 51.81, 37.21, 28.36, 24.17, 16.50, 16.01, 11.40. [α]_D²⁰ = 113.8 (C=1, CHCl₃)

Boc-Tyr(OtBu)-EK**(*S*)-*tert*-butyl (1-(4-(*tert*-butoxy)phenyl)-4-methyl-3-oxopent-4-en-2-yl)carbamate compound (10)**

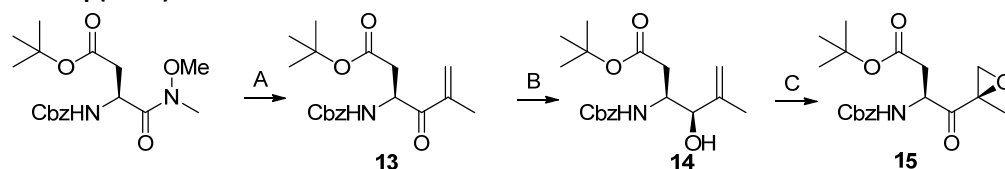
This compound was synthesized according to the general procedure **A** described above on a 10 mmol scale and was isolated after column chromatography (5→10% EtOAc:pent) (2.86 g, 7.9 mmol, 79%). ¹H NMR (400 MHz, CDCl₃) δ 6.96 (d, *J* = 8.3 Hz, 2H), 6.86 (d, *J* = 8.4 Hz, 2H), 5.93 (s, 1H), 5.78 (d, *J* = 1.4 Hz, 1H), 5.24 (dt, *J* = 20.8, 7.8 Hz, 2H), 2.94 (ddd, *J* = 46.6, 13.7, 6.2 Hz, 2H), 1.81 (s, 3H), 1.40 (s, 9H), 1.30 (s, 9H). ¹³C NMR (101 MHz, CDCl₃) δ 200.63, 155.15, 154.30, 142.72, 131.26, 129.91, 126.71, 124.28, 79.77, 78.46, 55.10, 39.61, 28.93, 28.46, 17.77.

***tert*-butyl ((2*S*,3*R*)-1-(4-(*tert*-butoxy)phenyl)-3-hydroxy-4-methylpent-4-en-2-yl)carbamate (11)**

This compound was synthesized according to the general procedure **B** described above on a 7.9 mmol scale and was isolated after column chromatography (10→20% EtOAc:pent) (1.90 g, 5.2 mmol, 66%). The product is formed as 5:1 mixture of diastereomers, but was isolated as 95:5 mixture after column chromatography. ¹H NMR (400 MHz, CDCl₃) δ 7.05 (d, *J* = 8.3 Hz, 2H), 6.87 (d, *J* = 8.3 Hz, 2H), 5.04 (s, 1H), 4.93 (s, 1H), 4.79 (d, *J* = 9.0 Hz, 1H), 4.14 (s, 1H), 3.95 (s, 1H), 3.04 – 2.73 (m, 2H), 2.70 – 2.53 (m, 1H), 1.76 (s, 3H), 1.31 (s, 9H), 1.29 (s, 9H).

***tert*-butyl ((*S*)-3-(4-(*tert*-butoxy)phenyl)-1-((*R*)-2-methyloxiran-2-yl)-1-oxopropan-2-yl)carbamate (12)**

This compound was synthesized according to the general procedure **D** described above on a 5.2 mmol scale and was isolated after column chromatography (5→10% EtOAc:pent) (0.54 g, 1.43 mmol, 27%). ¹H NMR (400 MHz, CDCl₃) δ 7.04 (d, *J* = 8.2 Hz, 2H), 6.90 (d, *J* = 8.4 Hz, 2H), 4.97 (d, *J* = 8.2 Hz, 1H), 4.55 (q, *J* = 8.0 Hz, 1H), 3.26 (d, *J* = 4.8 Hz, 1H), 3.03 (dd, *J* = 13.9, 5.0 Hz, 1H), 2.87 (d, *J* = 4.9 Hz, 1H), 2.67 (dd, *J* = 13.9, 8.0 Hz, 1H), 1.45 (s, 3H), 1.35 (s, 9H), 1.31 (s, 9H). ¹³C NMR (101 MHz, CDCl₃) δ 208.66, 155.24, 154.37, 130.76, 129.86, 124.22, 79.86, 78.45, 59.17, 53.70, 52.37, 37.05, 28.88, 28.31, 16.60. [α]_D²⁰ = 108.2 (C=1, CHCl₃)

Cbz-Asp(OtBu)-EK**(*S*)-*tert*-butyl 3-(((benzyloxy)carbonyl)amino)-5-methyl-4-oxohex-5-enoate (13)**

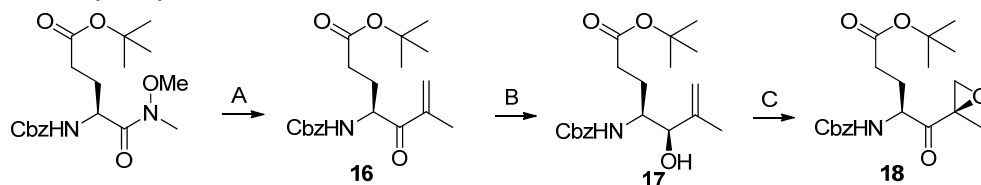
This compound was synthesized according to the general procedure **A** described above on a 8.8 mmol scale and was isolated after column chromatography (5→10% EtOAc:pent) (2.27 g, 5.6 mmol, 64%). ¹H NMR (400 MHz, CDCl₃) δ 7.39 – 7.28 (m, 5H), 6.07 (s, 1H), 5.87 (s, 1H), 5.81 (s, 1H), 5.33 – 5.21 (m, 1H), 5.09 (s, 2H), 2.78 (dd, *J* = 15.7, 5.6 Hz, 1H), 2.59 (dd, *J* = 15.7, 5.7 Hz, 1H), 1.90 (s, 3H), 1.39 (d, *J* = 1.9 Hz, 9H). ¹³C NMR (101 MHz, CDCl₃) δ 198.46, 169.34, 155.46, 141.81, 136.11, 128.38, 128.03, 127.92, 125.97, 81.42, 66.87, 51.47, 38.78, 27.83, 17.86.

(3*S*,4*R*)-*tert*-butyl 3-(((benzyloxy)carbonyl)amino)-4-hydroxy-5-methylhex-5-enoate compound (14)

This compound was synthesized according to the general procedure **B** described above on a 6.5 mmol scale and was isolated after column chromatography (10→30% EtOAc:pent) (quantitative yield). Complex NMR due to presence of rotamers or diastereomers (2:1). ¹H NMR (400 MHz, CDCl₃) δ 7.33 (q, *J* = 5.8, 5.1 Hz, 5H), 5.62 (d, *J* = 8.6 Hz, 0.65H), 5.43 (d, *J* = 8.3 Hz, 0.35H), 5.18 – 5.01 (m, 3H), 4.93 (d, *J* = 15.1 Hz, 1H), 4.20 (s, 0.7H), 4.11 (m, 1.3 H), 3.00 (s, 0.65H), 2.84 (s, 0.35H), 2.67 – 2.35 (m, 2H), 1.76 (s, 3H), 1.42 (s, 3H), 1.40 (s, 6H). ¹³C NMR (101 MHz, CDCl₃) δ 171.97, 156.01, 144.46, 128.55, 128.52, 128.15, 128.12, 128.05, 112.44, 81.32, 76.63, 76.04, 66.79, 50.34, 50.27, 38.28, 34.96, 28.07, 28.01, 19.03.

(*S*)-*tert*-butyl 3-(((benzyloxy)carbonyl)amino)-4-((*R*)-2-methyloxiran-2-yl)-4-oxobutanoate (15)

This compound was synthesized according to the general procedure **D** described above on a 6.5 mmol scale and was isolated after column chromatography (5→10% EtOAc:pent) (0.859 g, 1.43 mmol, 36%) as a single diastereomer. ¹H NMR (400 MHz, CDCl₃) δ 7.45 – 7.24 (m, 5H), 5.77 (d, *J* = 7.8 Hz, 1H), 5.24 – 4.89 (m, 2H), 4.53 (dt, *J* = 7.9, 5.4 Hz, 1H), 3.13 (d, *J* = 4.7 Hz, 1H), 2.88 (d, *J* = 4.7 Hz, 1H), 2.83 – 2.64 (m, 2H), 1.54 (s, 3H), 1.41 (s, 9H). ¹³C NMR (101 MHz, CDCl₃) δ 205.98, 169.32, 155.80, 136.21, 128.57, 128.22, 128.07, 81.95, 67.01, 59.40, 52.49, 51.27, 37.46, 28.03, 17.03. [α]_D²⁰ = 91.6 (C=1, CHCl₃)

Cbz-Glu(OtBu)-EK

(S)-tert-butyl 4-(((benzyloxy)carbonyl)amino)-6-methyl-5-oxohept-6-enoate (16)

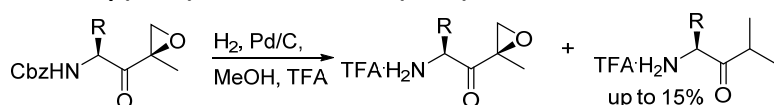
This compound was synthesized according to the general procedure **A** described above on a 10 mmol scale and was isolated after column chromatography (7.5% EtOAc:pent) (1.01 g, 2.8 mmol, 28% (conversion incomplete, but reaction was quenched due to appearance of by-products). ^1H NMR (400 MHz, CDCl_3) δ 7.37 – 7.17 (m, 5H), 6.19 (s, 1H), 5.92 (s, 1H), 5.74 (d, J = 8.2 Hz, 1H), 5.23 – 4.97 (m, 3H), 2.39 – 2.05 (m, 3H), 1.86 (s, 3H), 1.72 – 1.60 (m, 1H), 1.41 (s, 9H). ^{13}C NMR (101 MHz, CDCl_3) δ 199.96, 171.96, 156.00, 141.77, 136.31, 128.44, 128.06, 128.01, 127.21, 80.57, 66.80, 53.77, 30.81, 28.95, 28.02, 17.66.

(4S,5R)-tert-butyl 4-(((benzyloxy)carbonyl)amino)-5-hydroxy-6-methylhept-6-enoate (17)

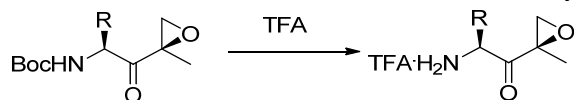
This compound was synthesized according to the general procedure **B** described above on a 2.8 mmol scale and was isolated after column chromatography (10→30% EtOAc:pent) (quantitative yield). Complex NMR due to presence of rotamers or diastereomers (8:1). ^1H NMR (400 MHz, CDCl_3) δ 7.39 – 7.28 (m, 5H), 5.24 (d, J = 9.2 Hz, 1H), 5.14 – 4.88 (m, 4H), 4.18 – 4.04 (m, 1H), 3.90 – 3.71 (m, 1H), 2.48 (bs, 1H), 2.28 (t, J = 7.4 Hz, 2H), 1.89 – 1.70 (m, 4H), 1.69 – 1.54 (m, 1H), 1.43 (s, 1H), 1.41 (s, 8H). ^{13}C NMR (101 MHz, CDCl_3) δ 173.26, 156.55, 144.62, 136.55, 128.60, 128.54, 128.20, 128.16, 111.86, 80.58, 77.27, 66.83, 52.90, 32.15, 28.15, 23.22, 19.35.

(S)-tert-butyl 4-(((benzyloxy)carbonyl)amino)-5-((R)-2-methyloxiran-2-yl)-5-oxopentanoate (18)

This compound was synthesized according to the general procedure **D** described above on a 2.8 mmol scale and was isolated after column chromatography (5→10% EtOAc:pent) (0.480 g, 1.27 mmol, 45%) as a single diastereomer. ^1H NMR (400 MHz, CDCl_3) δ 7.28 (dd, J = 8.9, 2.8 Hz, 5H), 5.63 (d, J = 8.4 Hz, 1H), 5.02 (q, J = 12.3 Hz, 2H), 4.32 (td, J = 8.7, 4.1 Hz, 1H), 3.21 (d, J = 4.9 Hz, 1H), 2.85 (d, J = 4.9 Hz, 1H), 2.29 (t, J = 7.3 Hz, 2H), 2.04 (dtd, J = 14.4, 7.3, 4.3 Hz, 1H), 1.84 – 1.58 (m, 1H), 1.49 (s, 3H), 1.40 (s, 9H). ^{13}C NMR (101 MHz, CDCl_3) δ 208.24, 171.99, 156.07, 136.13, 128.46, 128.10, 127.95, 80.75, 66.83, 59.07, 52.81, 52.15, 31.49, 28.00, 26.31, 16.48. $[\alpha]_D^{20}$ = +82.8 (C =1, CHCl_3)

TFA.H-Asp(OtBu)-EK and TFA.H-Glu(OtBu)-EK


Cbz protected epoxyketons are dissolved in MeOH (20 mg/mL), flushed with argon, followed by the addition of 1.2 equiv. TFA and 10% Pd/C (10 mg/100 mg epoxyketone). The reaction mixture is stirred under a H_2 atmosphere for 15-60 min and is filtered and concentrated upon completion. The product was directly used in the azide couplings to the corresponding hydrazides. Depending on the hydrogenation time, up to 15% reduced epoxide was formed, as was observed by NMR after azide coupling to the corresponding hydrazides. This impurity was removed by HPLC purification of the final compounds.

TFA·H-Leu-EK, TFA·H-Val-EK, TFA·H-Ala-EK, TFA·H-Tyr-EK, TFA·H-Ile-EK

Boc-protected epoxyketones were deprotected by treatment with TFA (neat) for 30 min, followed by concentrations and coevaporation with toluene. The product was directly used in the azide couplings to the corresponding hydrazides.

Synthesis of peptide hydrazides**Boc-Leu-Ala-OMe (19)**

To a solution of HCl-H-Ala-OMe (973 mg, 7 mmol, 1 equiv.) in DCM (70 mL) were added HCTU (3.19 g, 7.7 mmol, 1 equiv.), Boc-Leu-OH (1.78 g, 7.7 mmol, 1.1 equiv.) and DiPEA (4.3 mL, 24.5 mmol, 3.5 equiv.). After stirring overnight, the reaction mixture was concentrated, dissolved in EtOAc and washed with 1N HCl (2x), sat. NaHCO₃ (2x) and brine (1x). The organic layer was dried over NaSO₄, filtered and concentrated. Purification by column chromatography (10%→40% EtOAc/pent) provided the product (2.0 g, 6.3 mmol 90%). ¹H NMR (300 MHz, CDCl₃) δ 7.36 – 7.19 (m, 1H), 5.38 (d, *J* = 8.0 Hz, 1H), 4.41 (q, *J* = 6.9 Hz, 1H), 4.29 – 3.99 (m, 1H), 3.62 (s, 3H), 1.71 – 1.38 (m, 3H), 1.33 (s, 9H), 1.26 (d, *J* = 7.2 Hz, 3H), 0.96 – 0.68 (m, 6H). ¹³C NMR (75 MHz, CDCl₃) δ 173.14, 172.79, 155.84, 79.65, 52.86, 52.21, 47.89, 41.45, 28.31, 24.62, 22.95, 21.90, 17.70.

Ac-Leu-Ala-OMe (20)

Boc-Leu-Ala-OMe (**19**) (1.80g, 5.7 mmol) was dissolved in 1:1 TFA/DCM (20 mL) and stirred for 1 h, after which the mixture was concentrated and co-evaporated with toluene. TFA-H-Leu-Ala-OMe was dissolved in DCM (60 mL), followed by the addition of DiPEA (3.0 mL, 17 mmol, 3 equiv.) and Ac₂O (652 μL, 6.9 mmol, 1.2 equiv.). After stirring overnight, the reaction mixture was concentrated, dissolved in EtOAc and washed with 1N HCl (2x), sat. NaHCO₃ (2x) and brine (1x). The organic layer was dried over NaSO₄, filtered and concentrated. Purification by column chromatography (0%→2.5% MeOH/DCM) provided the product (1.03 g, 4 mmol, 70%). ¹H NMR (400 MHz, CDCl₃) δ 7.90 (d, *J* = 6.8 Hz, 1H), 7.68 (d, *J* = 8.1 Hz, 1H), 4.53 (q, *J* = 8.4 Hz, 1H), 4.33 (p, *J* = 7.1 Hz, 1H), 3.61 (s, 3H), 1.83 (s, 3H), 1.72 – 1.57 (m, 1H), 1.58 – 1.35 (m, 2H), 1.22 (d, *J* = 7.3 Hz, 3H), 0.80 (dd, *J* = 10.4, 6.6 Hz, 6H). ¹³C NMR (101 MHz, CDCl₃) δ 173.26, 173.04, 170.57, 52.11, 51.68, 48.11, 41.43, 24.59, 22.80, 22.64, 22.00, 17.19.

Ac-Leu-Ala-NHNH₂ (21)

To a solution of Ac-Leu-Ala-OMe (**20**) (1.03 g, 4 mmol) in MeOH (35 mL) was added hydrazine-hydrate (5.8 mL, 120 mmol, 30 equiv.). After stirring for 3 hours, the reaction was concentrated and co-evaporated with toluene (2x). The product was obtained in a quantitative yield. ¹H NMR (400 MHz, MeOD) δ 4.47 – 4.21 (m, 2H), 1.98 (s, 1H), 1.68 (dp, *J* = 13.8, 6.8 Hz, 1H), 1.62 – 1.46 (m, 2H), 1.34 (d, *J* = 6.9 Hz, 3H), 1.05 – 0.87 (m, 6H).

Boc-Pro-Ala-OMe (22)

To a solution of HCl-H-Ala-OMe (973 mg, 7 mmol, 1 equiv.) in DCM (70 mL) were added HCTU (3.19 g, 7.7 mmol, 1 equiv.), Boc-Pro-OH (1.67 g, 7.7 mmol, 1.1 equiv.) and DiPEA (4.3 mL, 24.5 mmol, 3.5 equiv.). After stirring overnight, the reaction mixture was concentrated, dissolved in EtOAc and washed with 1N HCl (2x), sat. NaHCO₃ (2x) and brine (1x). The organic layer was dried over NaSO₄, filtered and concentrated. Purification by column chromatography (20%→50% EtOAc/pent) provided the product in a quantitative yield. ¹H NMR (300 MHz, MeOD) δ 4.41 (q, *J* = 7.3 Hz, 1H), 4.28 – 4.14 (m, 1H), 3.71 (s, 3H), 3.56 – 3.36 (m, 2H), 2.36 – 2.10 (m, 1H), 2.05 – 1.77 (m, 3H), 1.51 – 1.30 (m, 12H).

Ac-Pro-Ala-OMe (23)

Boc-Pro-Ala-OMe (**22**) (2.0 g, 6.7 mmol) was dissolved in 1:1 TFA/DCM (20 mL) and stirred for 1 h, after which the mixture was concentrated and co-evaporated with toluene. TFA-H-Leu-Ala-OMe was dissolved in DCM (60 mL), followed by the addition of DiPEA (3.5 mL, 20 mmol, 3 equiv.) and Ac₂O (760 μ L, 8.0 mmol, 1.2 equiv.). After stirring overnight, the reaction mixture was concentrated, dissolved in EtOAc and washed with 1N HCl (2x), sat. NaHCO₃ (2x) and brine (1x). The organic layer was dried over NaSO₄, filtered and concentrated. Purification by column chromatography (0%→5% MeOH/DCM) provided the product (0.52 g, 2.14 mmol, 32%). Complex NMR due to presence of rotamers (1:3.5). ¹H NMR (400 MHz, CDCl₃) δ 7.45 (d, *J* = 6.4 Hz, 0.7H), 7.23 (d, 0.3H), 4.54 – 4.42 (m, 1H), 4.34 (p, *J* = 7.1 Hz, 0.75H), 4.23 (dd, *J* = 8.5, 2.5 Hz, 0.25H), 3.62 (s, 3H), 3.58 – 3.41 (m, 1.2H), 3.35 (q, *J* = 9.3 Hz, 0.8H), 2.29 – 1.71 (m, 4H), 1.98 (s, 1H), 2.00 (s, 2H), 1.32 (d, *J* = 7.3 Hz, 0.7H), 1.27 (d, *J* = 7.2 Hz, 2.3H). ¹³C NMR (101 MHz, CDCl₃) δ 173.13, 172.90, 171.81, 170.99, 170.68, 170.55, 61.88, 59.33, 52.26, 52.19, 48.19, 48.10, 47.88, 46.69, 31.98, 27.69, 24.82, 22.80, 22.39, 22.23, 17.67, 17.38.

Ac-Pro-Ala-NHNH₂ (24)

To a solution of Ac-Leu-Ala-OMe (**23**) (0.52 g, 2.14 mmol) in MeOH (20 mL) was added hydrazine-hydrate (3.1 mL, 64 mmol, 30 equiv.). After stirring for 3 h, the reaction was concentrated and co-evaporated with toluene (2x). The product was obtained in a quantitative yield. Complex NMR due to presence of rotamers. ¹H NMR (400 MHz, MeOD) δ 4.56 – 4.30 (m, 2H), 3.74 – 3.46 (m, 2H), 2.30 – 2.17 (m, 1H), 2.13 (s, 3H), 2.08 – 1.88 (m, 3H), 1.38 (d, *J* = 7.2 Hz, 3H). ¹³C NMR (101 MHz, MeOD) δ 174.35, 174.27, 173.97, 173.91, 172.55, 172.37, 62.22, 61.34, 61.30, 49.43, 49.10, 48.96, 47.96, 33.06, 30.92, 25.68, 23.81, 22.39, 22.23, 18.36, 17.98.

Boc-Ala-Pro-Ala-OMe (25)

Boc-L-Ala-OH (1.23 equiv., 0.464 g, 2.45 mmol) was dissolved in DCM (20 mL). HTCU (2 equiv., 0.992 g, 2.4 mmol) and DiPEA (3.53 equiv., 1.2 mL, 7.06 mmol) and TFA-Pro-Ala-OMe (**22b**) (1 equiv., 0.851 g, 2.0 mmol) were added to the mixture. After 1 h the solvent was evaporated and the crude was redissolved in EtOAc. The product was washed with 1M HCl (2x), sat. NaHCO₃ (2x) and brine (1x). The organic layer was dried over MgSO₄, filtered and concentrated. Column chromatography (0→2% MeOH/DCM) provided the product (0.232 g, 0.625 mmol, 31%). Complex NMR due to presence of rotamers. ¹H NMR (300 MHz, CDCl₃) δ 7.34 – 7.13 (m, 1H), 5.47 (d, *J* = 8.1 Hz, 1H), 4.60 – 4.14 (m, 3H), 3.73 – 3.44 (m, 5H), 2.19 – 1.84 (m, 4H), 1.38 – 1.32 (m, 9H), 1.31 – 1.24 (m, 6H). ¹³C NMR (75 MHz, CDCl₃, peaks of major rotamer) δ 173.12, 172.98, 171.01, 155.25, 79.66, 59.73, 52.33, 48.08, 47.75, 47.21, 28.27, 27.85, 24.94, 18.18, 17.78.

N₃Gly-Ala-Pro-Ala-OMe (26)

Boc-Ala-Pro-Ala-OMe (**26**) (140 mg, 0.376 mmol) was dissolved in DCM (2 mL) and TFA (2 mL) and the reaction was stirred for 30 min. After that the solution was evaporated and co-evaporated with toluene (3x), after which it was dissolved in DMF (5 mL), followed by the addition of DiPEA (0.2 mL, 1.1 mmol, 3 eq.) and (ClOAc)₂O (81 mg, 0.45 mmol (1.2 eq.) and stirred till completion (3h). NaN₃ (98 mg, 1.5 mmol, 4 eq.) was added and the reaction was stirred overnight. Reaction mixture was concentrated and redissolved in DCM and washed with 1 M HCl (1x) and brine (1x). Both water layers were back extracted several times with DCM. The organic layer was dried over MgSO₄, filtered and evaporated. Purification by column chromatography (0%→4% MeOH/DCM) yielded the desired compound (84 mg, 0.24 mmol, 63%). Complex NMR due to presence of rotamers. ¹H NMR (400 MHz, CDCl₃) δ 7.28 – 7.23 (m, 1H), 7.03 (d, *J* = 6.5 Hz, 1H), 4.84 – 4.66 (m, 1H), 4.51 (dt, *J* = 14.5, 5.8 Hz, 2H), 3.95 (s, 2H), 3.75 – 3.53 (m, 5H), 2.25 (d, *J* = 9.7 Hz, 1H), 2.20 – 2.05 (m, 1H), 2.05 – 1.90 (m, 2H), 1.45 – 1.29 (m, 6H). ¹³C NMR (101 MHz, CDCl₃, peaks of major rotamer) δ 173.33, 171.96, 170.54, 166.26, 59.99, 52.54, 48.20, 47.39, 46.79, 27.82, 25.14, 18.25.

N₃Gly-Ala-Pro-Ala-NHNH₂ (27)

N₃-Gly-Ala-Pro-Ala-OMe (**26**) (84 mg, 0.24 mmol) was dissolved in MeOH (2.5 mL). Hydrazine hydrate (0.35 mL, 7.1 mmol, 30 eq.) was added and the reaction mixture was stirred overnight. The reaction mixture was concentrated and co-evaporated with toluene (2x). The product was obtained in a quantitative yield. ¹H NMR (400 MHz, MeOD) δ 4.73 – 4.61 (m, 1H), 4.61 – 4.36 (m, 2H), 3.92 (s, 2H), 3.87 – 3.76 (m, 1H), 3.76 – 3.61 (m, 1H), 2.36 – 2.15 (m, 1H), 2.13 – 1.92 (m, 3H), 1.49 – 1.28 (m, 6H). ¹³C NMR (101 MHz, MeOD) δ 172.80, 172.64, 171.95, 171.78, 170.06, 168.46, 60.28, 60.02, 51.06, 47.82, 47.25, 29.17, 29.06, 24.64, 24.60, 16.69, 16.58, 15.49.

Ac-Pro-Leu-OMe (28)

To a solution of HCl-H-Leu-OMe (322 mg, 2 mmol, 1 equiv.) in DCM (20 mL) were added HCTU (913 mg, 2.2 mmol, 1.1 equiv.), Boc-Pro-OH (473 mg, 2.2 mmol, 1.1 equiv.) and DIPEA (1.2 mL, 7 mmol, 3.5 equiv.). After stirring overnight, the reaction mixture was concentrated, dissolved in EtOAc and washed with 1N HCl (2x), sat. NaHCO₃ (2x) and brine (1x). The organic layer was dried over NaSO₄, filtered and concentrated. Purification by column chromatography (20%→50% EtOAc/pent) provided the product (617 mg, 1.80 mmol, 90%) which was dissolved in 1:1 TFA/DCM (10 mL) and stirred for 30 min, after which the mixture was concentrated and co-evaporated with toluene. TFA-H-Pro-Leu-OMe (214 mg, 0.62 mmol) was dissolved in DCM (6 mL), followed by the addition of DIPEA (0.32 mL, 5.4 mmol, 3 equiv.) and Ac₂O (71 μL, 2.2 mmol, 1.2 equiv.). After stirring overnight, the reaction mixture was concentrated. Purification by column chromatography (50%→100% EtOAc/pent followed by 0%→2% MeOH/EtOAc) provided the product (136 mg, 0.47 mmol, 76%). Complex NMR due to presence of rotamers (1:2:3). ¹H NMR (400 MHz, CDCl₃) δ 7.36 (d, *J* = 7.3 Hz, 0.7H), 6.82 (d, *J* = 8.5 Hz, 0.3H), 4.62 – 4.49 (m, 1H), 4.48 – 4.33 (m, 7H), 4.27 (dd, *J* = 8.5, 2.6 Hz, 0.3H), 3.65 (s, 3H), 3.60 – 3.44 (m, 1H), 3.37 (td, *J* = 9.6, 7.1 Hz, 1H), 2.39 – 2.27 (m, 1H), 2.25 – 2.05 (m, 1H), 2.03 (s, 3H), 1.98 – 1.71 (m, 2H), 1.54 (ttd, *J* = 18.4, 8.8, 5.6 Hz, 3H), 0.90 – 0.79 (m, 6H). ¹³C NMR (101 MHz, CDCl₃) δ 173.18, 172.98, 172.07, 171.07, 170.93, 170.78, 62.15, 59.35, 52.28, 52.15, 51.07, 50.66, 48.24, 46.78, 41.06, 40.37, 32.12, 27.30, 25.10, 25.00, 24.91, 22.95, 22.69, 22.39, 22.35, 21.94, 21.36.

Ac-Pro-Leu-NHNH₂ (29)

To a solution of Ac-Pro-Leu-OMe (**28**) (136 mg, 0.47 mmol) in MeOH (5 mL) was added hydrazine-hydrate (0.69 mL, 64 mmol, 30 equiv.). After stirring for 3 hours, the reaction was concentrated and co-evaporated with toluene (2x). The product was obtained in a quantitative yield. Complex NMR due to presence of rotamers. ¹H NMR (400 MHz, MeOD) δ 4.48 – 4.42 (m, 0.4H), 4.42 – 4.36 (m, 1.6H), 3.70 – 3.47 (m, 2H), 2.38 – 2.27 (m, 0.2H), 2.27 – 2.14 (m, 0.8H), 2.10 (s, 2.5H), 2.07 – 1.83 (m, 3.5H), 1.75 – 1.44 (m, 3H), 1.01 – 0.87 (m, 6H). ¹³C NMR (101 MHz, MeOD) δ 174.53, 174.11, 173.65, 173.56, 172.71, 62.30, 61.38, 54.82, 51.79, 51.61, 48.00, 41.97, 41.67, 33.12, 30.94, 25.96, 25.78, 25.72, 23.82, 23.42, 23.32, 22.33, 22.14, 21.93, 21.86.

Ac-Leu-Leu-OMe (30)

A solution of Boc-Leu-Leu-OMe (358 mg, 1 mmol) in 1:1 TFA/DCM (10 mL) was stirred for 30 min, after which the mixture was concentrated and co-evaporated with toluene. TFA-H-Leu-Leu-OMe (214 mg, 0.62 mmol) was dissolved in DCM (6 mL), followed by the addition of DIPEA (0.52 mL, 3 mmol, 3 equiv.) and Ac₂O (113 μL, 1.2 mmol, 1.2 equiv.). After stirring overnight, the reaction mixture was concentrated. Purification by column chromatography (20%→50% EtOAc/pent) provided the product (238 mg, 0.79 mmol, 79%). ¹H NMR (400 MHz, CDCl₃) δ 7.44 (d, *J* = 7.6 Hz, 1H), 7.25 (d, *J* = 8.4 Hz, 1H), 4.57 (td, *J* = 8.5, 6.2 Hz, 1H), 4.43 (td, *J* = 8.5, 5.5 Hz, 1H), 3.64 (s, 3H), 1.88 (s, 3H), 1.76 – 1.39 (m, 6H), 0.83 (dt, *J* = 7.7, 6.1 Hz, 12H). ¹³C NMR (101 MHz, CDCl₃) δ 173.10, 172.98, 170.28, 52.08, 51.59, 50.94, 41.51, 40.78, 24.73, 24.65, 22.82, 22.76, 22.70, 22.39, 21.88.

Ac-Leu-Leu-NHNH₂ (31)

To a solution of Ac-Leu-Leu-OMe (238 mg, 0.79 mmol) in MeOH (8 mL) was added hydrazine-hydrate (1.12 mL, 24 mmol, 30 equiv.). After stirring for 3 hours, the reaction was concentrated and co-evaporated with toluene

(2x). The product was obtained in a quantitative yield. ^1H NMR (400 MHz, MeOD) δ 4.39 (dd, J = 8.7, 6.1 Hz, 2H), 1.98 (s, 3H), 1.73 – 1.45 (m, 6H), 0.93 (ddd, J = 16.3, 6.4, 4.4 Hz, 12H). ^{13}C NMR (101 MHz, MeOD) δ 174.74, 173.66, 173.39, 53.25, 51.62, 42.00, 41.73, 25.85, 25.76, 23.44, 23.35, 22.42, 22.10, 21.96.

Boc-Ala-Pro-Leu-OMe (32)

To a solution of TFA-H-Pro-Leu-OMe (400 mg, 1.17 mmol) in DCM (12 mL) were added HCTU (580 mg, 1.4 mmol, 1.2 equiv.), Boc-Ala-OH (226 mg, 1.4 mmol, 1.2 equiv.) and DiPEA (0.71 mL, 4.1 mmol, 3.5 equiv.). After stirring overnight, the reaction mixture was concentrated, dissolved in EtOAc and washed with 1N HCl (2x), sat. NaHCO_3 (2x) and brine (1x). The organic layer was dried over NaSO_4 , filtered and concentrated. Purification by column chromatography (50%→100% EtOAc/pent) provided the product (434 mg, 1.1 mmol, 89%). Complex NMR due to presence of rotamers (1:6), peaks of major rotamer reported. ^1H NMR (400 MHz, CDCl_3) δ 7.16 (d, J = 7.5 Hz, 1H), 5.36 (d, J = 8.1 Hz, 1H), 4.56 (dd, J = 8.0, 1.9 Hz, 1H), 4.40 (td, J = 8.0, 7.2, 4.0 Hz, 2H), 3.63 (s, 3H), 3.59 (d, J = 4.4 Hz, 1H), 3.47 (qd, J = 9.6, 8.7, 3.6 Hz, 1H), 2.35 – 2.23 (m, 1H), 2.14 – 1.98 (m, 1H), 1.98 – 1.88 (m, 1H), 1.88 – 1.73 (m, 1H), 1.57 – 1.42 (m, 3H), 1.35 (s, 9H), 1.23 (d, J = 7.0 Hz, 3H), 0.81 (dd, J = 10.3, 5.9 Hz, 6H). ^{13}C NMR (101 MHz, CDCl_3) δ 173.11, 173.09, 170.66, 155.10, 79.59, 59.52, 52.16, 50.85, 47.66, 47.16, 41.05, 28.31, 26.87, 25.08, 24.70, 22.81, 21.70, 18.48.

Boc-Ala-Pro-Leu-NHNH₂ (33)

To a solution of Boc-Ala-Pro-Leu-OMe (32) (215 mg, 0.5 mmol) in MeOH (5 mL) was added hydrazine-hydrate (0.71 mL, 15 mmol, 30 equiv.). After stirring for 3 hours, the reaction was concentrated and co-evaporated with toluene (2x). The product was obtained in a quantitative yield. ^1H NMR (400 MHz, MeOD) δ 4.45 (dd, J = 8.1, 4.9 Hz, 1H), 4.41 – 4.32 (m, 2H), 3.85 – 3.74 (m, 1H), 3.70 – 3.59 (m, 1H), 2.27 – 1.86 (m, 4H), 1.63 (dtd, J = 24.2, 10.4, 9.1, 5.9 Hz, 3H), 1.43 (s, 9H), 1.30 (d, J = 7.0 Hz, 3H), 0.93 (dd, J = 16.6, 6.4 Hz, 6H). ^{13}C NMR (101 MHz, MeOD) δ 174.44, 174.16, 173.71, 157.62, 80.46, 61.71, 51.84, 49.66, 41.77, 30.22, 28.79, 28.70, 26.04, 25.73, 23.40, 22.01, 17.13.

Ac-Ala-Pro-Leu-OMe (34)

Boc-Ala-Pro-Leu-OMe (32) (215 mg, 0.5 mmol) was treated with TFA (2 mL) for 30 min. The mixture was concentrated and co-evaporated with toluene (2x). To a solution TFA-H-Ala-Pro-Leu-OMe in DCM were added DiPEA (0.26 mL, 1.5 mmol, 3 equiv.) and Ac_2O (60 μL , 0.6 mmol, 1.2 equiv.). After stirring overnight, the reaction mixture was concentrated. Purification by column chromatography (0%→2% MeOH/DCM) provided the product (135 mg, 0.38 mmol, 76%). Complex NMR due to presence of rotamers, peaks of major rotamer reported ^1H NMR (400 MHz, MeOD) δ 4.58 (q, J = 7.0 Hz, 1H), 4.47 (dd, J = 8.6, 4.2 Hz, 1H), 4.41 (dd, J = 8.9, 6.1 Hz, 1H), 3.87 – 3.75 (m, 1H), 3.70 (s, 3H), 3.69 – 3.61 (m, 1H), 2.26 – 2.12 (m, 1H), 2.12 – 1.97 (m, 3H), 1.95 (s, 3H), 1.84 – 1.68 (m, 1H), 1.66 – 1.57 (m, 2H), 1.33 (d, J = 7.1 Hz, 3H), 0.94 (dd, J = 16.7, 6.6 Hz, 6H). ^{13}C NMR (101 MHz, MeOD) δ 174.58, 174.50, 173.55, 172.88, 61.11, 52.61, 52.23, 49.43, 49.21, 49.00, 48.79, 48.58, 48.46, 48.40, 41.40, 30.36, 25.90, 25.84, 23.31, 22.20, 21.89, 16.76.

Ac-Ala-Pro-Leu-NHNH₂ (35)

To a solution of Ac-Ala-Pro-Leu-OMe (34) (215 mg, 0.5 mmol) in MeOH (5 mL) was added hydrazine-hydrate (0.71 mL, 15 mmol, 30 equiv.). After stirring for 3 hours, the reaction was concentrated and co-evaporated with toluene (2x). The product was obtained in a quantitative yield. ^1H NMR (400 MHz, MeOD) δ 4.62 (q, J = 7.1 Hz, 1H), 4.46 (dd, J = 8.2, 4.6 Hz, 1H), 4.36 (dd, J = 9.5, 5.6 Hz, 1H), 3.83 (dt, J = 10.0, 6.7 Hz, 1H), 3.74 – 3.61 (m, 1H), 2.29 – 2.15 (m, 1H), 2.15 – 1.92 (m, 6H), 1.79 – 1.52 (m, 3H), 1.36 (d, J = 7.0 Hz, 3H), 0.96 (dd, J = 16.2, 6.4 Hz, 6H). ^{13}C NMR (101 MHz, MeOD) δ 172.86, 172.51, 172.47, 171.53, 60.25, 50.57, 47.23, 40.57, 28.99, 24.71, 24.44, 22.09, 20.92, 20.75, 15.49.

Biochemical methods

Proteasomes

Wild-type and mutant proteasomes were purified as previously described.²⁸ Purified human constitutive proteasome was bought from Boston Biochem (USA).

Proteasome substrates and inhibitors

AMC-substrates (Bachem) and inhibitors were stored as 50-100 mM solutions in DMSO at -20 °C. Bortezomib was purchased from Selleck Chemicals, carfilzomib from Active Biochemicals, and ONX-0914 from MedKoo Biosciences.

IC₅₀ Determination with purified yCP

Concentrations of purified yCP variants were determined spectrophotometrically at 280 nm. Initial point measurements were carried out with 200 μM of inhibitor. Only compounds that showed significant inhibition at this concentration were further evaluated. Purified yCPs (final concentration: 66 nM in 100 mM Tris-HCl, pH 7.5) were mixed with DMSO as a control or serial dilutions of inhibitor and incubated for 1 h at room temperature. After addition of the peptide substrate Z-LLE-AMC, Boc-LRR-AMC or Suc-LLVY-AMC (final concentration of 200 μM) and incubation for 1 h at room temperature, proteolysis was stopped by diluting the samples 1:10 in 20 mM Tris-HCl, pH 7.5. The AMC-molecules released by residual proteasomal activity were measured in triplicate with a Varian Cary Eclipse Fluorescence Spectrophotometer (Agilent Technologies) at λ_{ex} =360 nm and λ_{em} =460 nm. Relative fluorescence units were normalized to the DMSO treated control. The calculated residual activities were plotted against the logarithm of the applied inhibitor concentration and fitted with GraphPad Prism 5. IC₅₀ values were deduced from the fitted data. They depend on enzyme concentration and are comparable only within the same experimental settings.

Competition Assays in Raji Cell Lysate.

Lysates of Raji cells were prepared by addition of 4 volumes of lysis buffer containing 50 mM Tris pH 7.5, 2 mM DTT, 5 mM MgCl₂, 10% glycerol, 2 mM ATP, and 0.05% digitonin. Protein concentration was determined by the Bradford assay. Cell lysates (diluted to 10-15 μg total protein in buffer containing 50 mM Tris pH 7.5, 2 mM DTT, 5 mM MgCl₂, 10% glycerol, 2 mM ATP) were exposed to the inhibitors for 1 h at 37 °C prior to incubation with AzidoBODIPY-MeTyr-Phe-Leu-VS (BODIPY(TMR)-NC-005-VS; 0.1 μM; to probe β5), or a mixture of BODIPY(TMR)-EKoxomicin (0.5 μM; to probe all subunits, used for β2-profiling) and BODIPY(FL)-Ala-Pro-Nle-Leu-EK (BODIPY-NC001; 0.25 μM; to probe β1 activity, followed by 3 min boiling with a reducing gel-loading buffer and fractionation on 12.5% SDS-PAGE. In-gel detection of residual proteasome activity was performed in the wet gel slabs directly on a ChemiDoc™ MP System using Cy3 settings to detect BODIPY(TMR)-NC-005 and BODIPY(TMR)-EKoxomicin and Cy2 settings to detect BODIPY-NC-001-VS. Intensities of bands were measured by fluorescent densitometry and normalized to the intensity of bands in mock-treated extracts. Average values of three independent experiments were plotted against inhibitor concentrations. IC₅₀ values were calculated using GraphPad Prism software.

Crystallization and structure determination

The yeast 20S proteasome was crystallized by hanging drop vapour diffusion according to published procedures.²⁸ Inhibitor complex structures were obtained by incubating crystals in 5 μl cryobuffer (20 mM magnesium acetate, 100 mM MES (2-(N-morpholino)ethanesulfonic acid), pH 6.8 and 30% (v/v) MPD (2-methyl-2,4-pentanediol) supplemented with inhibitor at a final concentration of 3.3 mM for at least 12 h prior to vitrification in liquid nitrogen. Diffraction data were collected using synchrotron radiation of $\lambda = 1.0 \text{ \AA}$ at the beamline X06SA, Swiss Light Source (SLS), Villigen, Switzerland. Evaluation of reflection intensities and data reduction were performed with the program package XDS. Molecular replacement using the coordinates of the

yCP (PDB entry code: 1RYP13) was carried out by rigid body and anisotropic TLS refinements with REFMAC5; MAIN and COOT served as model building software. The coordinates finally yielded excellent R_{crys}, R_{free}, r.m.s.d. bond and angle values as well as good stereochemistry from the Ramachandran Plot and have been deposited in the RCSB Protein Data Bank (for X-ray data collection and refinement statistics, PDB codes and all structures: see²⁹).

References

1. Groettrup, M., Kirk, C.J. & Basler, M. Proteasomes in immune cells: more than peptide producers? *Nat. Rev. Immunol.* **10**, 73-8 (2010).
2. Murata, S., Takahama, Y. & Tanaka, K. Thymoproteasome: probable role in generating positively selecting peptides. *Curr. Opin. Immunol.* **20**, 192-6 (2008).
3. Groll, M. et al. Structure of 20S proteasome from yeast at 2.4 Å resolution. *Nature* **386**, 463-71 (1997).
4. Mirabella, A.C. et al. Specific cell-permeable inhibitor of proteasome trypsin-like sites selectively sensitizes myeloma cells to bortezomib and carfilzomib. *Chem. Biol.* **18**, 608-18 (2011).
5. Blackburn, C. et al. Characterization of a new series of non-covalent proteasome inhibitors with exquisite potency and selectivity for the 20S beta5-subunit. *Biochem. J.* **430**, 461-76 (2010).
6. Huber, E.M. et al. Immuno- and constitutive proteasome crystal structures reveal differences in substrate and inhibitor specificity. *Cell* **148**, 727-738 (2012).
7. Orłowski, M. The multicatalytic proteinase complex, a major extralysosomal proteolytic system. *Biochemistry* **29**, 10289-97 (1990).
8. Kisselev, A.F. et al. The caspase-like sites of proteasomes, their substrate specificity, new inhibitors and substrates, and allosteric interactions with the trypsin-like sites. *J. Biol. Chem.* **278**, 35869-77 (2003).
9. Nussbaum, A.K. et al. Cleavage motifs of the yeast 20S proteasome beta subunits deduced from digests of enolase 1. *Proc. Natl. Acad. Sci.* **95**, 12504-9 (1998).
10. Harris, J.L., Alper, P.B., Li, J., Rechsteiner, M. & Backes, B.J. Substrate specificity of the human proteasome. *Chem. Biol.* **8**, 1131-41 (2001).
11. Toes, R.E. et al. Discrete cleavage motifs of constitutive and immunoproteasomes revealed by quantitative analysis of cleavage products. *J. Exp. Med.* **194**, 1-12 (2001).
12. Nazif, T. & Bogoy, M. Global analysis of proteasomal substrate specificity using positional-scanning libraries of covalent inhibitors. *Proc. Natl. Acad. Sci.* **98**, 2967-72 (2001).
13. Groll, M., Kim, K.B., Kairies, N., Huber, R. & Crews, C.M. Crystal structure of epoxomicin: 20S proteasome reveals a molecular basis for selectivity of α' , β' -epoxyketone proteasome inhibitors. *J. Am. Chem. Soc.* **122**, 1237-1238 (2000).
14. Verdoes, M. et al. A panel of subunit-selective activity-based proteasome probes. *Org. Biomol. Chem.* **8**, 2719-27 (2010).
15. Li, N. et al. Relative quantification of proteasome activity by activity-based protein profiling and LC-MS/MS. *Nat. Prot.* **8**, 1155-68 (2013).
16. de Bruin, G. et al. Structure-based design of beta1i or beta5i specific inhibitors of human immunoproteasomes. *J. Med. Chem.* **57**, 6197-209 (2014).
17. Harshbarger, W., Miller, C., Diedrich, C. & Sacchettini, J. Crystal Structure of the Human 20S Proteasome in Complex with Carfilzomib. *Structure* **23**, 418-24 (2015).
18. Groll, M., Korotkov, V.S., Huber, E.M., de Meijere, A. & Ludwig, A. A Minimal β -Lactone Fragment for Selective β 5c or β 5i Proteasome Inhibitors. *Angew. Chem. Int. Ed.* **54**, 7810-7814 (2015).
19. Huber, E.M., Heinemeyer, W. & Groll, M. Bortezomib-Resistant Mutant Proteasomes: Structural and Biochemical Evaluation with Carfilzomib and ONX 0914. *Structure* **23**, 407-17 (2015).

20. Heinemeyer, W., Kleinschmidt, J.A., Saidowsky, J., Escher, C. & Wolf, D.H. Proteinase yscE, the yeast proteasome/multicatalytic-multifunctional proteinase: mutants unravel its function in stress induced proteolysis and uncover its necessity for cell survival. *EMBO J.* **10**, 555-62 (1991).
21. Huber, E.M. & Groll, M. Inhibitors for the immuno- and constitutive proteasome: current and future trends in drug development. *Angew. Chem. Int. Ed.* **51**, 8708-20 (2012).
22. Demo, S.D. et al. Antitumor activity of PR-171, a novel irreversible inhibitor of the proteasome. *Cancer Res.* **67**, 6383-91 (2007).
23. Orlowski, M. & Wilk, S. Catalytic activities of the 20 S proteasome, a multicatalytic proteinase complex. *Arch. Biochem. Biophys.* **383**, 1-16 (2000).
24. Cardozo, C., Vinitsky, A., Michaud, C. & Orlowski, M. Evidence that the nature of amino acid residues in the P3 position directs substrates to distinct catalytic sites of the pituitary multicatalytic proteinase complex (proteasome). *Biochemistry* **33**, 6483-9 (1994).
25. Orlowski, M., Cardozo, C. & Michaud, C. Evidence for the presence of five distinct proteolytic components in the pituitary multicatalytic proteinase complex. Properties of two components cleaving bonds on the carboxyl side of branched chain and small neutral amino acids. *Biochemistry* **32**, 1563-72 (1993).
26. Geurink, P.P. et al. Incorporation of non-natural amino acids improves cell permeability and potency of specific inhibitors of proteasome trypsin-like sites. *J. Med. Chem.* **56**, 1262-1275 (2013).
27. Zhou, H.-J. et al. Design and synthesis of an orally bioavailable and selective peptide epoxyketone proteasome inhibitor (PR-047). *J. Med. Chem.* **52**, 3028-3038 (2009).
28. Gallastegui, N. & Groll, M. Analysing properties of proteasome inhibitors using kinetic and X-ray crystallographic studies. *Methods Mol. Biol.* **832**, 373-90 (2012).
29. Huber, E.M. et al. Systematic analyses of substrate preferences of 20S proteasomes using peptidic epoxyketone inhibitors. *J. Am. Chem. Soc.* **137**, 7835-7842 (2015).