Bioorganic & Medicinal Chemistry Letters 23 (2013) 2696–2703

Contents lists available at SciVerse ScienceDirect

Bioorganic & Medicinal Chemistry Letters

journal homepage: www.elsevier.com/locate/bmcl

C^1 and N^5 derivatives of cerpegin: Synthesis of a new series based on structure–activity relationships to optimize their inhibitory effect on 20S proteasome

Anna Hovhannisyan [b,†], The Hien Pham [a,†], Dominique Bouvier [c], Lixian Qin [a], Gagik Melikyan [b], Michèle Reboud-Ravaux [a], Michelle Bouvier-Durand [a,*]

[a] Enzymologie Moléculaire et Fonctionnelle, UR4-Université de Paris 6, UPMC-Sorbonne Universités, Case Courrier 256, 7 Quai Saint Bernard, F 75252 Paris Cedex 05, France
[b] Department of Organic Chemistry, Yerevan State University, A. Manoogian Str. 1, 0025 Yerevan, Armenia
[c] Atelier de Bioinformatique, Université de Paris 6, UPMC-Sorbonne Universités, Case Courrier 1202, 4 Place Jussieu, F 75252 Paris Cedex 05, France

ARTICLE INFO

Article history:
Received 23 December 2012
Revised 15 February 2013
Accepted 18 February 2013
Available online 27 February 2013

Keywords:
Alkaloid cerpegin
Furo[3,4-*c*]pyridine-3,4-dione
Benzenesulfanilamide
Proteasome inhibitors
Selective inhibition of PA activity
In silico docking

ABSTRACT

Thirty-two new derivatives of cerpegin (1,1,5-trimethylfuro[3,4-*c*]pyridine-3,4-dione) were designed and synthesized in high yield by a new method, combining several C^1 and N^5 substituents. All compounds were tested for their inhibitory effect on the CT-L, T-L and PA proteolytic activities of a purified mammalian 20S proteasome. Only one molecule inhibited both CT-L and PA activities. Sixteen molecules specifically inhibited PA at the micromolar range, out of which fourteen had IC_{50} values around 5 μM and two had IC_{50} values closer to 2 μM. Except in one case, neither calpain I nor cathepsin B was inhibited. In silico docking suggests a unique mode of binding of the most efficient compounds to the β1 catalytic site (PA activity) in relation to the chemical nature of C^1 substituents.

A number of naturally occurring molecules and their derivatives have been shown to inhibit the proteolytic activity of proteasome (reviewed in[1–3]). The natural alkaloid cerpegin enters this case. Indeed, the cerpegin scaffold (1,1,5-trimethylfuro[3,4-*c*]pyridine-3,4-dione) has recently been used to develop series of N^5 derivatives with good performances as selective inhibitors of the post-acid activity (PA) of the 20S proteasome.[4] The 20S proteasome is the multiproteolytic core of the more complex 26S proteasome, working downstream the ubiquitinylation/deubiquitinylation steps necessary for the recognition of cell substrates.[1–3,5–9] The 20S complex comprises four stacked heptameric rings (α1-7, β1-7, β1-7 and α1-7) forming a cylindrical structure. The different proteolytic activities are confined to the two inner β rings: two PA active sites (post-acid or caspase-like) are found in the so-called β1 subunits (each β1 subunit belonging to a separate heptameric β ring), two chymotrypsin-like sites (CT-L, located on β5 subunits) and two trypsin-like sites (T-L, located on β2 subunits) The six catalytic subunits represent the main drug targets of the proteasome. Each type of catalytic site possesses a threonine residue at position one capable of nucleophilic attack on the substrates. However, the activities can be distinguished in vitro by their specificity towards amino acids adjacent to the scissile peptide bonds of specific synthetic substrates. In brief, cleavage occurs on the C-side of mainly acidic (PA site), basic (T-L site) or hydrophobic (CT-L site) residues.[10] Three similar types of catalytic subunits (β1i, β2i and β5i) with similar activities arise in eukaryotic cells after induction by biological stimuli. They are assembled in a 20S core particle in the so-called immunoproteasome in place of the constitutively expressed β1, β2 and β5 subunits. In this work, only the constitutive fraction was used.

Covalent inhibitors including peptide mimetics and non-peptidic molecules have been first described and represent the largest class of molecules known to target the 20S catalytic core.[2,11–13] Non-covalent inhibitors have been more recently developed.[13–19]

Dipeptide boronate bortezomib (Velcade or PS341) belongs to the class of covalent inhibitors. It was the first to be used as a

Abbreviations: DMF/DMA, dimethylformamide/dimethylacetal; DMSO, dimethyl sulfoxide; 20S, 20S proteasome catalytic core; CT-L, chymotrypsin-like; T-L, trypsin-like; PA, post-acid or caspase-like; AMC, 7-amino-4-methyl-coumarin; ®-NA, ®-naphthylamide; PDB, Protein Data Bank.

* Corresponding author. Tel.: +33 1 44 27 69 52; fax: +33 1 44 27 51 40.
E-mail addresses: annahovh@gmail.com (A. Hovhannisyan), bsthehien@yahoo.com (T.H. Pham), dbouvier@snv.jussieu.fr (D. Bouvier), lixian.qin@free.fr (L. Qin), gagikmelikyan@yahoo.com (G. Melikyan), Michele.Reboud@upmc.fr (M. Reboud-Ravaux), mbouvier@snv.jussieu.fr (M. Bouvier-Durand).
† These authors contributed equally to this work.

0960-894X/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.bmcl.2013.02.079

prescription drug to treat multiple myeloma and mantle lymphoma diseases.[20,21] Acquired resistance to, and adverse effects of the chemotherapy making use of bortezomib represent practical difficulties.[22] Thus, second-generation inhibitors have been developed: CEP-18770 and MLN9708, two other peptide-boronates, carfilzomib, a tetrapeptide epoxyketone, and salinosporamide A, a β-lactone.[23] All gained intensive clinical studies and carfilzomib (Kyprolis or PR-171) has been recently approved by FDA for the treatment of multiple myeloma (MM).

The natural product salinosporamide A (SalA) is a potent inhibitor of all three proteasome activities whose lactone is attacked by the nucleophilic Thr1 side chain, leading to formation of a highly stable acyl-enzyme.[24] The γ-lactone ring of cerpegin (Fig. 1A, left) is also susceptible to nucleophilic attack by Thr1 and a first series of N^5 derivatives with a dimethyl at C^1 has been developed in the laboratory.[4,25] A number of them specifically inhibit PA activity at the micromolar range (**7** molecules have an IC_{50} of about 5–6 μM).[4] Within this range, seven N^5 substituents were identified as good candidates to promote PA inhibition by the compounds to which they belong (Fig. 1A).

For practical reasons, out of the seven N^5 substituents associated with acceptable IC_{50} values, only four (Fig. 1A, **1′**–**4′**) were retained in the present study to serve as a basis for the synthesis of new cerpegin derivatives (Fig. 1B) with the aim to further optimize inhibition. A selection of R^1, R^2 substituents at C^1 (including spiro-compounds) was done (Fig. 1B). We report here two new series of cerpegin derivatives (here generically called compounds **III**) classified according to the nature of their N^5 moiety: purely aromatic or aliphatic displaying or not an aromatic moiety (substituted aliphatic).

The first series (**1**–**15**) mainly combined the aromatic benzenesulfonamides at N^5 with either aliphatic or aliphatic/aromatic R^1, or spiro-fused R^1/R^2 substituents (Fig. 1B). Our interest towards benzenesulfonamides comes from their interesting inhibitory effect on 20S proteasome as previously reported[4] (Fig. 1A) and from the published properties of sulfonamide-substituted heterocyclic systems.[26] Moreover, sulfone-containing heterocycles have successfully been tested for proteasome inhibition.[27] The second series of derivatives (**16**–**32**) combined one of the three other most efficient substituents at position N^5 (Fig. 1B) with the R^1 or R^1/R^2 substituents described for the first series (Fig. 1B).

The new furo[3,4-*c*]pyridine-3,4-dione derivatives substituted at N^5 and C^1 positions with aromatic and/or aliphatic substituents on both the pyridinone and furanone rings were prepared as outlined in Scheme 1.

The starting lactones **I** are mostly commercially available and could be synthesized by condensation of the corresponding ketoalcohols with diethyl malonate in the presence of catalytic amount of sodium ethylate.[28]

Special interest has also been paid to compounds with R^1/R^2 alicyclic (cyclohexane) and heterocyclic (tetrahydropyran) spiro-structures (Fig. 1B, bottom right). Such spiro-structures were obtained by the same method,[28] starting with 1-(1-hydroxy-cyclohexyl)-ethanone and 1-(4-hydroxy-2,2-dimethyl-tetrahydropyran-4-yl)-ethanone, respectively. In the case of 1-(4-hydroxy-2,2-dimethyl-tetrahydropyran-4-yl)-ethanone new spiro-lactone **Ih** was obtained (Supplementary data).

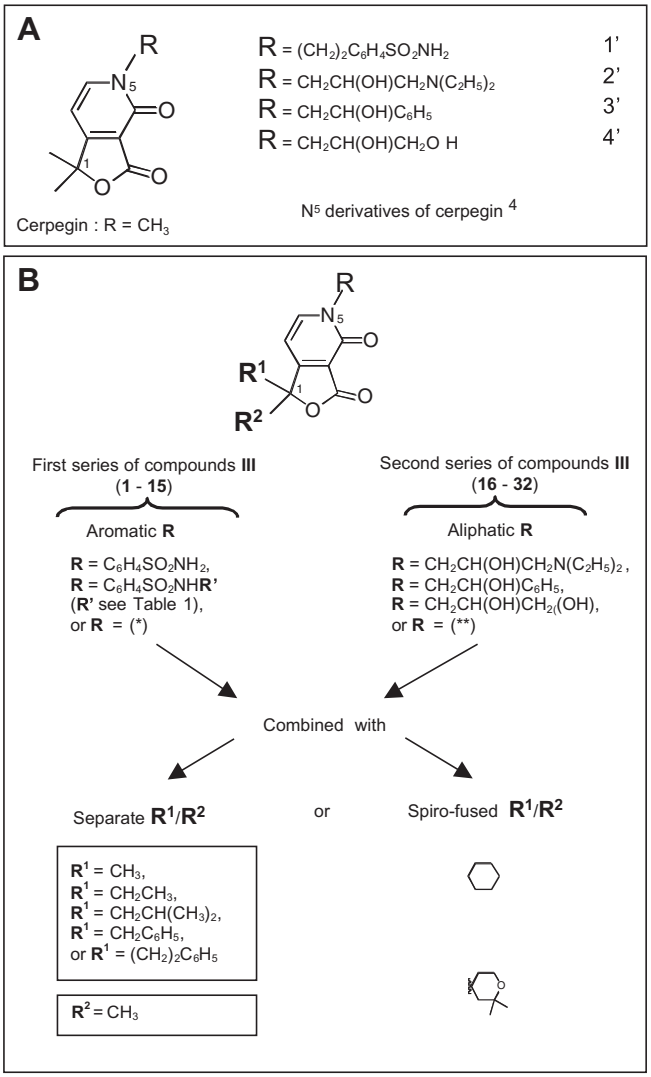
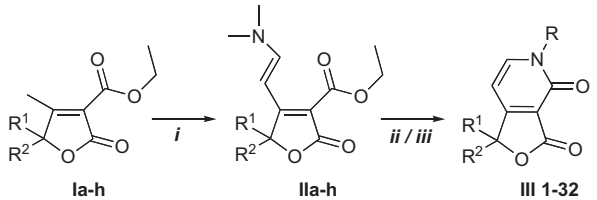


Figure 1. General presentation of the C^1 and N^5 derivatives of cerpegin synthesized in this study: (A) structure of cerpegin and formulae of the main N^5 substituents (R) already described in[4] for 1,1-dimethyl cerpegin derivatives. Primed numbers on the right are used in the text to refer to these 1,1-dimethyl derivatives. (B) New derivatives of cerpegin (compounds **III**) combining some of the most performant R substituents as defined in[4] at position N^5 and different R^1 or spiro-fused R^1/R^2 substituents at position C^1. By convention, in the case of separate R^1/R^2, R^1 is the variable substituent and R^2 is always a methyl group. In the case of spiro-fused R^1/R^2, R^1 and R^2 are forming a cyclohexyl ring or a tetrahydropyran ring. (*) Another N^5 substituent was included in this series: 5-amino-3-methyl-1-phenylpyrazole. (**) Other N^5 substituents were included in this series: 2-(4-chloro)-phenethyl (**23** and **24**), 3-morpholinopropyl (**25**–**29**), 2-hydroxy-3-morpholinopropyl (**30**), thiourea (**31**), as well as 1-phenyl-substituted cerpegin (**32**).



R^1, R^2 = Me, Me (a); Et, Me (b); i-Bu, Me (c); Ph, Me (d); Bz, Me (e); PhEt, Me (f)

R^1/R^2 = $(CH_2)_5$ (g); $(CH_2)_2O(CH_2)_2$ (h)

Scheme 1. Synthesis of new furo[3,4-*c*]pyridine-3,4-dione compounds **III**, **1**–**32**. Reagents and conditions: (*i*) for compound **I**, dimethylformamide dimethylacetal (DMF/DMA), xylene, reflux 3 h, 75–89% yield; (*ii*) for compounds **III**, **1**–**15**: **II**, aromatic amines (1:1 molar ratio), AcOH, reflux 15 h. (*iii*) for compounds **III**, **16**–**32**: **II**, NH_2R, xylene, reflux 15 h, typical yields: 75–90%.

Condensation of lactones **I** with DMF/DMA by refluxing in anhydrous xylene (Scheme 1, *i*) led to the corresponding dimethylaminovinyl derivatives **II**. Starting from **II**, two distinct pathways yielded two series of N^5-substituted furopyridine structures (compounds **III**) in one step by interaction with either aromatic (Scheme 1, *ii*) or aliphatic (Scheme 1, *iii*) amines. Experimental details for the synthesis of **II** are given in Supplementary data. The first series of compounds **III**, mainly comprised of N^5-benzenesulfonamide derivatives (**1**–**14**, see Table 1), was synthesized with high yield from **II** (Scheme 1, *ii*) and 4-aminobenzenesulfonamides in glacial acetic acid (Table 1 and Supplementary data).

Compounds **1**–**14** correspond to the first R entry in Figure 1B (left). Compounds **1**–**6** are derivatives of sulfanilamide with the same N^5-substituent. Compounds **1**–**5** have different R^1 substituents so as to make a decision about the influence of R^1 on the investigated activities. Compound **6** possesses unique spiro-fused R^1/R^2. Compounds **7**–**14** have two methyl groups as R^1, R^2 and can be viewed as N-substituted sulfonamide derivatives of compound **1**.

For comparison, compound **15** with N^5 aromatic heterocyclic substituent was obtained by condensation with 5-amino-3-methyl-1-phenylpyrazole in glacial acetic acid (Table 1). For the second series of compounds **III** (**16**–**32**, see Table 2), the cyclocondensation and insertion of an aliphatic-substituted group at N^5 of the pyridinone ring was performed in the xylene ambience following step *iii* (Scheme 1).

The most frequently used substituents were chosen on the basis of our previous results[4] and correspond to R entries in Figure 1B (right). In spite of their low efficiency when associated with a dimethyl at C^1,[4] other N^5 groups were also included in this second series to check for possible improvements with larger R^1 substitutions. These were: 2-(4-chloro)-phenethyl (**23** and **24**), 3-morpholinopropyl (**25**–**29**), 2-hydroxy-3-morpholinopropyl (**30**), thiourea (**31**), as well as 1-phenyl-substituted cerpegin (**32**).

The combined effects of R^1, R^2 and R substitutions of compounds **III** on the inhibition of the activities of the proteasomal 20S catalytic core (CT-L, T-L and PA) were investigated. For practical reasons, the three proteolytic activities were assayed using 20S proteasome fractions purified from rabbit erythrocytes. We have recently reported that the results remain quantitatively the same when using 20S proteasome fractions from human cells.[4] The detailed protocol used to measure the three peptidase activities has previously been described[29] and is summed up thereafter.[30] A systematic search for the selectivity of the inhibitions was performed using two other proteases: cathepsin B, purified from human hepatocytes and calpain I, purified from human erythrocytes, the activities of which were measured as described.[30] Results were expressed as IC_{50} values, which correspond to the inhibitor concentration causing 50% inhibition, relative to controls (without inhibitor).

Results obtained with the first series of compounds **III** (**1**–**15**) for the three proteolytic activities of the 20S proteasome are summed up in Table 1.

In this series designed after the 1,1-dimethylated compound **1′** (Fig. 1A), the N^5-substituents are mainly benzenesulfonamides (**1**–**6**) and N-derivatives thereof (**7**–**14**). As for compound **1′** from which they derive, these compounds **III** did not inhibit CT-L nor T-L activities. Only PA activity was inhibited at the micromolar range for a limited number of compounds (**1**–**9**). The best PA inhibition was obtained with compound **5** (IC_{50} = 2.8 μM), with a phenethyl as R^1. Compound **4** with a benzyl group as R^1 was a little less efficient (IC_{50} = 3.6 μM). The gradual impact of the chemical nature of R^1 on inhibition in the series from **1** to **6** can be summed up as follows: **5** (phenethyl) > **4** (benzyl) > **3** (isobutyl) ≈ **2** (ethyl) > **1** (methyl) ≈ **6** (cyclohexyl) and is discussed below in the context of docking results.

Compounds **7**–**9** with the respective N-4-acetyl, benzoyl, carbamidoyl benzenesulfonamide group in N^5 position were at least as efficient as **3**, **2** and **1** (IC_{50} around 5–6 μM). A dramatic loss of effect was observed for compounds **10**–**14** with the respective N-4-pyridinyl-, pyrimidinyl-, (1-phenyl-pyrazolyl-), (3,4-dimethylisoxazolyl-), thiazolyl-benzenesulfonamide, that is, heterocyclic-substituted benzenesulfonamides at N^5, with practically no inhibition with 100 μM compounds **12** and **14**. Compound **15** is not a sulfonamide but bears the aromatic 3-methyl-1-phenyl-1*H*-pyrazol-4-yl substituent and was not inhibitor on the three activities of the 20S proteasome. Neither calpain I nor cathepsin B was inhibited by these compounds (**1**–**15**).

Results obtained with the second series of compounds **III** (**16**–**32**) are summed up in Table 2.

Once again, none of these new molecules were inhibitors of CT-L or T-L activities, except for molecule **24** which caused mild CT-L inhibition (IC_{50} = 22.5 μM). A number of these molecules (**16**–**22**), designed after compounds **2′**, **3′** and **4′** (Fig. 1A), inhibited only PA within an acceptable micromolar range. Compared with 1,1-dimethyl derivatives of cerpegin, a significant optimization of inhibition (2- to 4-fold) was obtained only when R^1 was a benzyl or a phenethyl group. For example, molecule **19**, which combines the 3-diethylamino-2-hydroxypropyl substituent at N^5 with one phenethyl at C^1 (IC_{50} = 2.5 μM) was 2.5 times more efficient on PA than the 1,1-dimethyl homolog. The same was true of the comparison between compound **27** of this study (3-morpholinopropyl N^5-group) and its dimethyl homolog (optimization ≈4-fold as compared to compound **10** in[4]). In the same way, associating the 2-(4-chloro-phenethyl) N^5-group with a benzyl at C^1 yielded a molecule (**24**), which not only inhibited PA with higher efficiency (optimization ≈3-fold relative to the dimethyl counterpart) but also inhibited CT-L. No significant improvement of inhibition was associated with the other types of substituents at C^1, including cases where these substituents were replaced by a cyclohexyl ring (**16** and **21**), a tetrahydropyran ring (**17**) or when R^1 was simply a phenyl group (**18** and **26**).

As observed for compounds **1**–**15**, none of the compounds studied in this series (**16**–**32**) worked on purified fractions of cathepsin B and only one (**24**) worked on calpain I (IC_{50} = 29.6 μM).

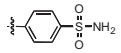
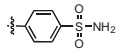
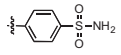
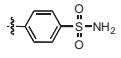
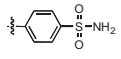
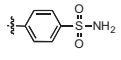
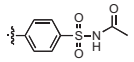
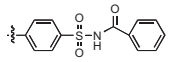
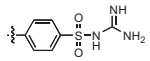
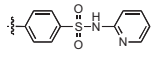
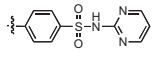
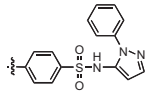
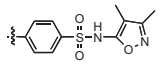
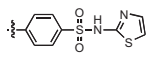
In summary, for these two series of compounds **III** (**1**–**15**) and (**16**–**32**), the inhibitory effects of N^5 derivatives of cerpegin upon PA activity can be optimized by providing them with a large, hydrophobic and flexible R^1 such as the phenethyl group. The spiro groups initially introduced to assess the effect of ligand stiffening did not improve PA inhibition relative to a methyl group.

In silico docking experiments were performed to illustrate the binding mechanisms of compounds **24**, **5** and **19** to the proteasome active sites (Fig. 2). The catalytic chains from the bovine 20S proteasome and their cognate partners were used as targets.[31,32] The choice of this model was made on the basis of the very strong amino acid sequence identity between human and bovine β-type subunits.[33] For example, the bovine β1 precursor protein is 95% identical to the human one and the identity reaches 98.5% when considering only the mature subunit, after precursor autolysis. Moreover, the few non-identical residues are located at the interfaces of β1 with its neighbors and are thus absent from the substrate binding sites. The AudoDock Vina program[34] was used for docking calculations,[35] with default parameters. In cases where asymmetric carbons were present, all corresponding configurations were docked.

Visual Molecular Dynamics (VMD)[36] was used to prepare molecular pictures.

We observed very similar poses of the inhibitors in the first ranks (1–3) of results for all the combinations of N^5 and C^1 substitutions subjected to docking analysis. As already noted with other derivatives of cerpegin,[4] the bicyclic furopyridinone ring fitted into

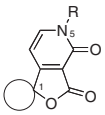
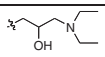
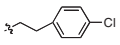
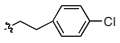
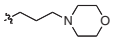
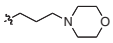
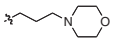
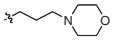
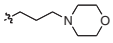
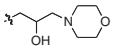
Table 1
Inhibition of 20S proteasome by C^1 and N^5 derivatives of cerpegin (compounds **III**, **1**–**15**): effects of diverse N^5 aromatic substituents (mainly benzenesulfonamides) and C^1 substituents

Compounds III				Yield (%)	Mp (°C)	IC_{50} (μM) or *(%)* inhibition at 100 μM		
	R	R^1	R^2	Spiro-Fused R^1/ R^2		CT-L	T-L	PA
1					84.0	>270	ni	ni	**6.1±0.3**
2					87.2	272	ni	ni	**4.9±0.2**
3					65.8	216-8	ni	ni	**4.6±0.7**
4					67.8	>260	ni	ni	**3.6±0.2**
5					74.4	193	ni	ni	**2.8±0.2**
6					95.5	>300	ni	ni	**6.6±0.6**
7					68,7	205	ni	ni	**5.7±0.3**
8					60.5	278	ni	ni	**4.5±0.3**
9					52.1	310	ni	ni	**5.9±0.4**
10					73.0	>260	ni	ni	*(64%)*
11					88.2	>260	ni	ni	*(51%)*
12					89.3	170	ni	ni	*(28%)*
13					44.7	147	ni	ni	*(51%)*
14					66.1	>250	ni	ni	*(22%)*
15					99.0	245	ni	ni	ni
Cerpegin[4]							ni	ni	10.4±0.5
1'							ni	ni	5.5±0.1
SalA[23]							0.0026	0.021	0.46

Yield (%) and melting points are indicated for each compound. Inhibitions of 20S proteasome from rabbit erythrocytes are expressed as IC_{50} values or as % inhibition at 100 μM (italics in parenthesis). ni = no inhibition. When R^1 and R^2 are different the synthesis led to racemates. Cerpegin and compound **1′** (from[4]) and salinosporamide A (SalA)[23] are given as reference inhibitors.

a slit in the active site between Thr residues 1 and 21 (Fig. 2A–F), which appeared as the main hydrogen bond donors in the interaction (dotted lines in Fig. 2B). This placed the inhibitors alongside the peptide chain of a sequence from residue 46 to residue 49 (Ser-Gly-Ser-Ala in PA). In all cases, the free end of the N^5 substituent was close to the aromatic ring of Tyr114 of the β2 subunit.

Table 2
Inhibition of 20S proteasome by C^1 and N^5 derivatives of cerpegin (compounds **III**, **16**–**32**): effects of diverse N^5 aliphatic substituents and C^1 substituents

Compounds III					Yield (%)	Mp (°C)	IC_{50} (μM) or % inhibition at 100 μM		
	R	R^1	R^2	Spiro-fused R^1/ R^2			CT-L	T-L	PA
16					89.5	183	ni	ni	**6.7±0.3**
17					72.1	175.6	ni	ni	**4.1±0.3**
18*			—		76.2	137	ni	ni	**4.4±0.2**
19*			—		71.1	95-7	ni	ni	**2.5±0.06**
20*			—		54.5	184	ni	ni	**4.8±0.5**
21					57.2	225	ni	ni	**11.9±0.4**
22					77.9	224	ni	ni	**8.1±0.2**
23			—		89.4	151	ni	ni	**12.7±0.5**
24			—		71.6	191	**22.5±1.2**	ni	**7±0.5**
25			—		83.9	143	ni	ni	*(31.8%)*
26			—		85.1	167	ni	ni	**56.4±1.8**
27			—		82.6	137	ni	ni	**19.6±1.6**
28					92.4	175	ni	ni	**54.7±2.3**
29					80.9	168	ni	ni	**46.9±3.4**
30					83.7	175	ni	ni	*(30.7%)*
31					65.5	270	*(23.8%)*	ni	ni
32	—		—		91.4	214	ni	ni	**36±4.2**
Cerpegin[4]							ni	ni	10.4±0.5
2'							ni	ni	6,.1±0.7
SalA[23]							0.0026	0.021	0.46

Yield (%) and melting points are indicated for each compound. Inhibitions of 20S proteasome from rabbit erythrocytes are expressed as IC_{50} values or as % inhibition at 100 μM (italics in parenthesis). ni = no inhibition. Cerpegin and compound **2'** (from[4]) and salinosporamide A (SalA)[23] are given as reference inhibitors.
[a] Compounds **18**–**20** were produced as mixtures of diastereoisomers (see Supplementary data).

When the N^5 substituent comprised one (Fig. 2C and D) or several (Fig 2E and F) polar group(s), the poses predicted the possibility of hydrogen bond formation with the carbonyl oxygen of Thr21 in β1 and/or the side chain alcohol of Ser118 in β2.

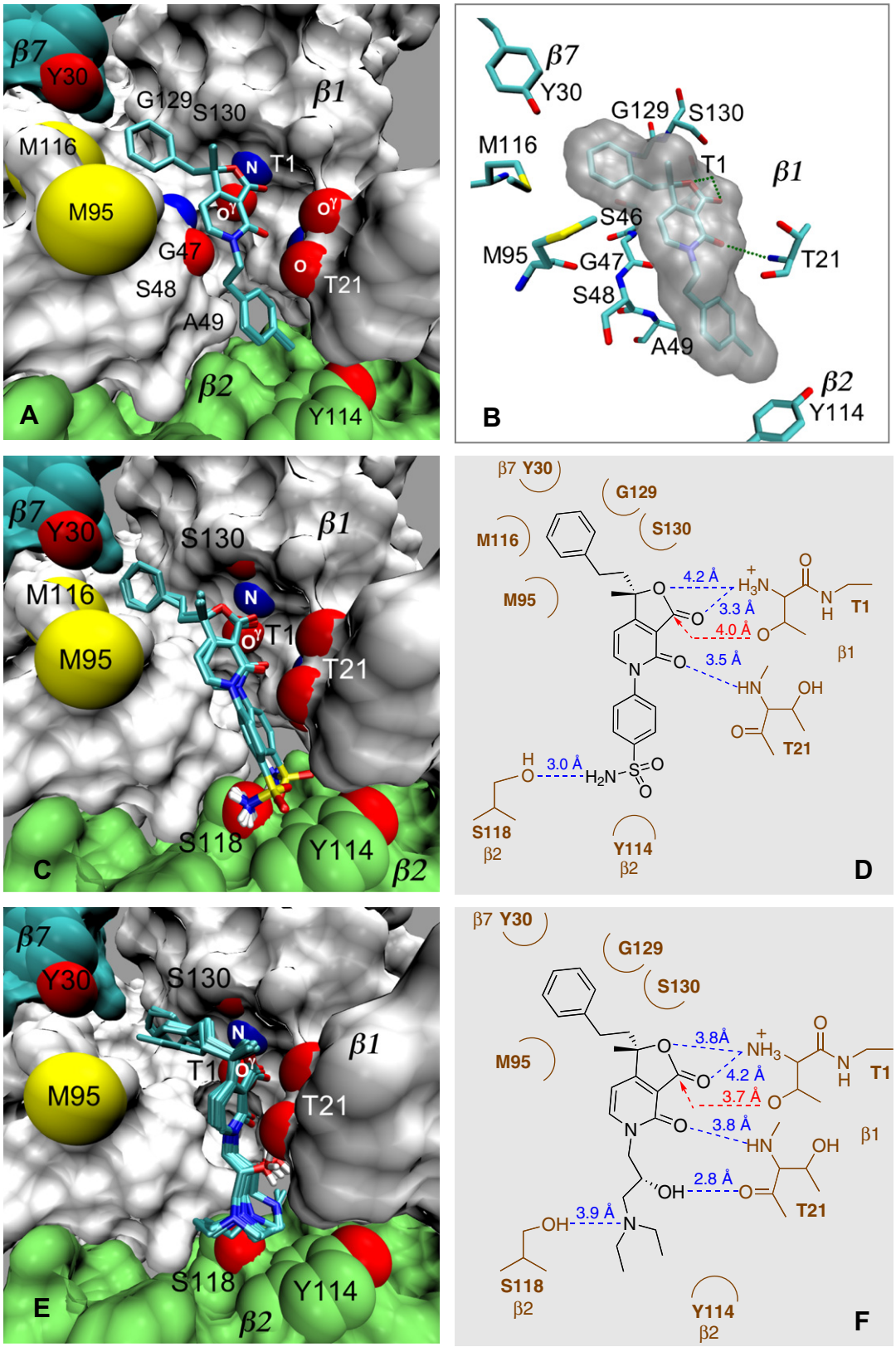


Figure 2. In silico docking of compounds **24** (A and B), **5** (C and D) and **19** (E and F) in the PA active site of bovine 20S proteasome. For all three compounds illustrated, the absolute configuration of the asymmetric carbon C^1 is *S* and that of the alcohol carbon in **19** is *R*. (A, C, E) Protein chains are shown as solvent accessible surfaces and the residues and atoms relevant to an enzyme activity are labelled. Inhibitors are displayed as stick models colored by atom type. Each docking mode is shown as six representative poses from the most populated and lowest-energy cluster. (B) Compound **24** is shown with transparent molecular surface and relevant amino-acids in the β1, β2 and β7 subunits are displayed as sticks. The dotted lines between the ligand and T1 and T21 represent putative hydrogen bonds. (D, F) Distances reported for the β1 subunit to atoms in **5** and **19**, respectively, were measured from the lowest energy pose. Putative hydrogen bonds are shown as blue dashed lines and the predicted nucleophilic attack on the lactone carbon, by a dashed arrow. Note: the crystal structure of the eukaryotic bovine proteasome (PDB ID: 1IRU)[31] was obtained from the Protein Data Bank (PDB).[32]

Table 3
Comparison of calculated K_D values for compounds **5**, **19** and **24**, with opposite configurations *S* and *R* at C^1 and for their respective 1,1-dimethylated analogs **1**, **2′** and **21**[4]

Calculated K_D for compounds **5**, **19** and **24**[a]				Calculated K_D for the 1,1-dimethylated analogs **1**, **2′**, **21**[4]	
Compound	K_D (μM)	Compound	K_D (μM)	Compound	K_D (μM)
5 *S*	6–7	**5** *R*	20–24	1	20–25
19 *S*[b]	20–25	**19** *R*[b]	200	**2′**	170
24 *S*	10	**24** *R*	10	**21**[4]	30

K_D values were computed from the binding energies given by AutoDock Vina.
[a] The substituents at C^1 are a methyl and a phenethyl for compounds **5** and **19**, a methyl and a benzyl for **24**.
[b] The asymmetric carbon of the hydroxyl group of the N^5 substituent is *R*.

The R^1 substituents lied in a rather hydrophobic groove running into Tyr30 in subunit β7 and lined by the peptide groups of Gly129 and Ser130 on one side and by the side chains of Met residues 95 and 116 of β1 on the other side (Fig. 2). This groove corresponds to the primed substrate binding or specificity channel of the PA activity.[37]

Substitutions at C^1 made this carbon asymmetric and two configurations had to be considered.

The absolute configuration at C^1 is *S* for all the molecules shown in Figure 2. This configuration corresponded to the best docking results in terms of rank and calculated K_D (Table 3) for compounds **5** (Fig. 2C and D) and **19** (Fig. 2E and F). The opposite configuration (*R*) gave rise to much higher K_D values due to the fact that the R^1 substituent was excluded from the binding site (illustrated in Fig. S1A and B). In contrast, the two stereoisomers of **24** adopted very similar poses (illustrated in Fig. S1C and D) and the computed K_D values were the same (Table 3). Table 3 also compares the binding affinities of 1-monosubstituted cerpegin derivatives **5**, **19** and **24** with those of their respective 1,1-dimethyl analogs, which adopted similar position in the β1 binding site (illustrated in Fig. S2A for **24**). K_D values were always higher for the dimethyl compounds than for the monomethyl ones and in some cases (see for examples **5** and **19**) this effect depended on the stereochemistry of the substitution. The consequences of this stereochemistry varied among compounds (Table 3), ranging from no effect for **24** to a 10-fold increase in K_D for **19** and an intermediate threefold increase in K_D for **5**. Thus, docking calculations reproduced the optimization associated with C^1 substitutions already revealed by in vitro assays (IC_{50}). In addition, the K_i values calculated from experimental IC_{50} using the Cheng–Prusoff equation[38] were of the same order of magnitude as the K_D reported in Table 3.

In the present study none of the tested compounds were predicted to interact with the S1 pocket. Therefore, the molecular basis for the PA specificity of these compounds should be searched elsewhere. We have already pointed out[4] the possible involvement of a tyrosine residue (Tyr114) in the contiguous β2 subunit, which most often makes favorable contacts with N^5 substituents and is absent from the catalytic site of the two other proteolytic activities. The present study also suggests that the R^1 substituent could play some part in PA specificity by binding to the primed specificity channel of the β1 active site, which is the most hydrophobic of the three β primed binding channels. The binding of **24** to the β5 active site (CT-L) is shown in Figure S3 for comparison.

In conclusion, the two series of cerpegin derivatives synthesized in this study have been characterized for their selectivity towards 20S proteasome. All molecules but one were inefficient in inhibiting the activity of both calpain I and cathepsin B. Sixteen molecules specifically inhibited PA activity within the micromolar range, out of which 14 had IC_{50} values around 5 μM and two IC_{50} values closer to 2 μM. Interestingly, the latter two molecules have a phenethyl-substituted C^1 and this substitution gives rise to inhibitors of higher potency relative to the 1,1-dimethylated analogs, depending on the steoreochemistry. These results open new ways for enhancing the inhibitory power of such molecules while preserving their specificity towards PA. Moreover, it will be necessary to develop protocols for the purification of the most efficient stereoisomers where asymmetric carbons are involved. Subunit-selective inhibitors of proteasome are still poorly documented. One example to date is a urea-containing peptide epoxyketone derived from syringolins. [39] The C^1 and N^5 derivatives of cerpegin could constitute a convenient set of tools to further analyze the respective roles of proteasome subunits in cell physiology and pathologies involving the proteasome.

Acknowledgments

A.H. and T.H.P. received financial supports from the Yerevan University (A.H.) and from the Vietnamian Ministery for Education (T.H.P.). This study was also supported by the Universities of Yerevan (G.M.), Pierre et Marie Curie (UPMC, Paris 6) (M.B.-D., M.R.-R., L.Q. and D.B.) and by a Grant from the 'French Association Against Myopathies'.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.bmcl.2013.02.079.

References and notes

1. Grawert, M. A.; Groll, M. *Chem. Commun. (Camb.)* **2012**, *48*, 1364.
2. Kisselev, A. F.; van der Linden, W. A.; Overkleeft, H. S. *Chem. Biol.* **2012**, *19*, 99.
3. Genin, E.; Reboud-Ravaux, M.; Vidal, J. *Curr. Top. Med. Chem.* **2010**, *10*, 232.
4. Pham, T. H.; Hovhannisyan, A.; Bouvier, D.; Tian, L.; Reboud-Ravaux, M.; Melikyan, G.; Bouvier-Durand, M. *Bioorg. Med. Chem. Lett.* **2012**, *22*, 3822.
5. Coux, O.; Tanaka, K.; Goldberg, A. L. *Annu. Rev. Biochem.* **1996**, *65*, 801.
6. Hershko, A.; Ciechanover, A. *Ann. Rev. Biochem.* **1998**, *67*, 425.
7. Borissenko, L.; Groll, M. *Chem. Rev.* **2007**, *107*, 687.
8. Marques, A. J.; Palanimurugan, R.; Matias, A. C.; Ramos, P. C.; Dohmen, R. J. *Chem. Rev.* **2009**, *109*, 1509.
9. Weissman, A. M.; Shabek, N.; Ciechanover, A. *Nat. Rev. Mol. Cell Biol.* **2011**, *12*, 605.
10. Groll, M.; Heinemeyer, W.; Jâger, S.; Ullrich, T.; Bochtler, M.; Wolf, D. H.; Huber, R. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 10976.
11. Kisselev, A. F.; Goldberg, A. L. *Chem. Biol.* **2001**, *8*, 739.
12. Zhang, S.; Shi, Y., et al *J. Mol. Model.* **2009**, *15*, 1481.
13. Beck, P.; Dubiella, C.; Groll, M. *Biol. Chem.* **2012**, *393*, 1101.
14. Koguchi, Y.; Kohno, J.; Nishio, K.; Takahashi, T.; Okuda, T.; Ohnuki, S.; Komatsubara *J. Antibiot.* **2000**, *53*, 105.
15. Groll, M.; Goetz, M.; Kaiser, M.; Weyher, E.; Moroder, L. *Chem. Biol.* **2006**, *13*, 607.
16. Basse, N.; Piguel, S.; Papapostolou, D.; Ferrier-Berthelot, A.; Richy, N.; Pagano, M.; Sarthou, P.; Sobczak-Thepot, J.; Vidal, J. *J. Med. Chem.* **2007**, *50*, 2842.
17. Groll, M.; Gallastegui, N.; Maréchal, X.; Le Ravalec, V.; Basse, N.; Richy, N.; Genin, E.; Huber, R.; Moroder, L.; Vidal, J.; Reboud-Ravaux, M. *ChemMedChem* **2010**, *1701*.
18. Blackburn, C.; Gigstad, K. M.; Hales, P.; Garcia, K. M.; Barrett, M.; Liu, J. X.; Soucy, T. A.; Sappal, D. S.; Bump, N.; Olhava, E. J.; Fleming, P.; Dick, L. R.; Tsu, C.; Sintchak, M. D.; Blank, J. L. *Biochem. J.* **2010**, *430*, 461.
19. Marechal, X.; Pujol, A.; Richy, N.; Genin, E.; Basse, N.; Reboud-Ravaux, M.; Vidal, J. *Eur. J. Med Chem.* **2012**, *52*, 322.
20. Adams, J. *Nat. Rev. Cancer* **2004**, *4*, 349.

21. Adams, J. *Cancer Cell* **2003**, *5*, 417.
22. Orlowski, R. Z.; Kuhn, D. *J. Clin. Cancer Res.* **2008**, *14*, 1649.
23. Macherla, V. R.; Mitchell, S. S.; Manam, R. R.; Reed, K. A.; Chao, T. H.; Nicholson, B.; Deyanat-Yadzi, G.; Mai, B.; Jensen, P. R.; Fenical, W. F.; Neuteboom, S. T. C.; Lam, K. S.; Palladino, M. A.; Potts, B. C. M. *J. Med. Chem.* **2005**, *48*, 3684.
24. Groll, M.; Huber, R.; Potts, B. C. *J. Am. Chem. Soc.* **2006**, *128*, 5136.
25. Communication by Melikyan, G.; Hovhannisyan, A.; Avetisyan, A. at the Congress of the French Chemical Society «La Chimie du Futur. Le Futur de la Chimie» SFC 07, Paris, July 16–18, 2007.
26. Haque, W.; Dinkur, J. U.S. Patent No. US 7,459,468 B2, 2008.
27. Basse, N.; Montes, M.; Marechal, X.; Qin, L.; Bouvier-Durand, M.; Genin, E.; Vidal, J.; Villoutreix, B. O.; Reboud-Ravaux, M. *J. Med. Chem.* **2010**, *53*, 509.
28. Avetisyan, A. A.; Tadevosyan, G. E.; Mangasaryan, C. A.; Matsoyan, S. G.; Dangyan, M. T. *J. Org. Chim. (Rus.)* **1970**, *5*, 962.
29. Formicola, L.; Maréchal, X.; Basse, N.; Bouvier-Durand, M.; Bonnet-Delpon, D.; Milcent, T.; Reboud-Ravaux, M.; Ongeri, S. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 83.
30. CT-L, PA and T-L activities were determined using the respective synthetic substrates Suc-LLVY-AMC, Z-LLE-βNA and Boc-LRR-AMC. All substrates and inhibitors were previously dissolved in dimethylsulfoxide (DMSO) whose final concentration was 2% (v/v). After a pre-incubation period of 15 min with the enzyme in the absence (control) or presence of test compounds (0.1–200 μM), reactions were followed for 45 min at 37 °C. The buffers (pH 7.5) were 20 mM Tris, 1 mM DTT, 10% glycerol, 0.02% (w/v) SDS for CT-L and PA activities, and 20 mM Tris, 1 mM DTT, 10% glycerol for T-L activity. The inhibitor concentrations giving 50% inhibition (IC_{50} values) were obtained by plotting the percent inhibition against inhibitor concentration to equation: $\%\ \text{inhibition} = 100[I]/(IC_{50}^{nH} + [I]^{nH})$, or equation $\%\ \text{inhibition} = 100[I]^{nH}/(IC_{50}^{nH} + [I]^{nH})$ where nH is the Hill number. The K_M values of the fluorogenic substrates in our experimental conditions were: 30 ± 5 μM (Suc-LLVY-AMC), 77 ± 4 μM (Z-LLE-βNA) and 26 ± 6 μM (Boc-LRR-AMC). For calpain I and cathepsin B activities the respective Suc-LLVY-AMC and Z-RR-AMC substrates were used. Buffers and temperatures used were as follows: 50 mM Tris, 100 μM $CaCl_2$ and 10 mM DTT (pH 7.2) at 25 °C for calpain I, and 352 mM KH_2PO_4, 48 mM Na_2HPO_4, 1 mM EDTA, 1 mM DTT (pH 6) at 37 °C for cathepsin B. In all assays, emitted fluorescence was measured using black 96-well microplates and a BMG Fluostar microplate reader.
31. Unno, M.; Mizushima, T.; Morimoto, Y.; Tomisugi, Y.; Tanaka, K.; Yasuoka, N.; Tsukihara, T. *Structure* **2002**, *10*, 609.
32. Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235.
33. UniProt Consortium, *Nucleic Acids Res.*, **2012**, *40*, D71.
34. Trott, O.; Olson, A. J. *J. Comput. Chem.* **2010**, *31*, 455.
35. The docking space was limited to a 20 × 20 × 20 Å box centered on the N terminal threonine of the catalytic subunits. Docking modes were selected on the basis of the distance (3–4 Å) between the carbonyl carbon of the compound lactone and the hydroxyl oxygen of the N terminal threonine of each subunit.
36. Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33.
37. Groll, M.; Larionov, O. V.; Huber, R.; de Meijere, A. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 4576.
38. Cheng, Y.; Prusoff, W. H. *Biochem. Pharmacol.* **1973**, *22*, 3099.
39. Van der Linden, W. A.; Willems, L. I.; Shabaneh, T. B.; Li, N.; Ruben, M.; Florea, I. B.; Van der Marel, G. A.; Kaiser, M.; Kisselev, A. F.; Overkleeft, H. S. *Org. Biomol. Chem.* **2012**, *10*, 181.