Pergamon

SCIENCE DIRECT®

Bioorganic & Medicinal Chemistry 11 (2003) 3723–3733

BIOORGANIC & MEDICINAL CHEMISTRY

C_2-Symmetric Inhibitors of *Plasmodium falciparum* Plasmepsin II: Synthesis and Theoretical Predictions

Karolina Ersmark,[a] Isabella Feierberg,[b] Sinisa Bjelic,[b] Johan Hultén,[a] Bertil Samuelsson,[c] Johan Åqvist[b] and Anders Hallberg[a,*]

[a]*Department of Medicinal Chemistry, Uppsala University, BMC, Box 574, SE-751 23 Uppsala, Sweden*
[b]*Department of Cell and Molecular Biology, Uppsala University, BMC, Box 596, SE-751 24 Uppsala, Sweden*
[c]*Medivir AB, Lunastigen 7, SE-141 44 Huddinge, Sweden*

Received 18 March 2003; accepted 14 May 2003

Abstract—A series of C_2-symmetric compounds with a mannitol-based scaffold has been investigated, both theoretically and experimentally, as Plm II inhibitors. Four different stereoisomers with either benzyloxy or allyloxy P1/P1′ side chains were studied. Computational ranking of the binding affinities of the eight compounds was carried out using the linear interaction energy (LIE) method relying on a complex previously determined by crystallography. Within both series of isomers the theoretical binding energies were in agreement with the enzymatic measurements, illustrating the power of the LIE method for the prediction of ligand affinities prior to synthesis. The structural models of the enzyme–inhibitor complexes obtained from the MD simulations provided a basis for interpretation of further structure–activity relationships. Hence, the affinity of a structurally similar ligand, but with a different P2/P2′ substituent was examined using the same procedure. The predicted improvement in binding constant agreed well with experimental results.

Introduction

Malaria, one of the major public health problems in the world, is caused by parasites of the genus Plasmodium. Of these, *Plasmodium falciparum* is most virulent and responsible for the majority of all malaria deaths. Hundreds of millions of people are afflicted and almost 3 million people die from the disease each year. Since resistance to current antimalarial drugs is spreading rapidly, these numbers are increasing, highlighting an urgent need for new drugs acting at new parasitic targets.[1] In the intraerythrocytic phase, the parasite uses host hemoglobin as its primary source of amino acids, and the enzymes involved in the hemoglobin degradation process have now emerged as promising targets. Among these enzymes, the aspartic proteases plasmepsin I (Plm I), plasmepsin II (Plm II), and the cystein protease falcipain are of particular interest to the medicinal chemists.[2]

Recently, Plm II inhibitors with good potencies in enzyme assays and with fair selectivity versus the human enzyme cathepsin D have been identified.[3–14] In addition, a 3-D structure of Plm II complexed with pepstatin A has been determined (1SME) and utilized for modeling studies.[8–10] Unfortunately, in cases where efforts are directed towards new classes of inhibitors, that is those comprising scaffolds that are structurally significantly different from those in 3-D determined ligand–enzyme complexes, the design process is frequently more troublesome.

We aimed at examining new linear C_2-symmetric compounds, with the symmetrical 1,2-dihydroxyethylene transition state mimetic, as potential Plm II inhibitors. We previously used this type of symmetric scaffold, easily available from mannitol, successfully for the inhibition of both Plm I, Plm II,[13] and the C_2-symmetric HIV-1 protease.[15] However, before embarking on a more extensive medicinal chemistry program, we were strongly encouraged to examine computational tools that could guide our selection of scaffolds with the optimal stereochemical configuration for efficient inhibition of Plm II.

*Corresponding author. Tel.: +46-18-471-4284; fax: +46-18-471-4474; e-mail: andersh@orgfarm.uu.se

0968-0896/03/$ - see front matter © 2003 Elsevier Ltd. All rights reserved.
doi:10.1016/S0968-0896(03)00339-0

It has been shown that the S1′ pocket is collapsed in the pepstatin A/Plm II crystal structure,[3] but still Ellman's group identified potent ligands with large P1′ side chains,[8] demonstrating the difficulties in predicting activity when one 3-D structure is available solely. For our purpose, a method that allows enzyme flexibility is needed, since it appeared that the ligands studied here do not fit readily into the available crystal structure if the protein is kept fixed. Some methods have been developed that consider receptor flexibility and this may be the most promising route to more accurate docking.[16] In this case, we choose to treat the protein flexibility by molecular dynamics (MD) simulations. Although this is not a general solution to flexible docking, it can be expected to yield good results if the initial model is reasonably reliable. Combined with simulated annealing the scope of the MD conformational searching may, however, be further expanded.[17–19] The docking procedure used here is based on the similarities between the present ligands and pepstatin A. Although there are significant differences between the chemical structures of pepstatin A and the ligands examined in this report, the similarities suggest a reasonable positioning of the ligand side chains and a large part of the backbone. Some guidance can also be obtained by considering structures of HIV-1 protease complexes.[20] Åqvist et al. have described an approach, in which absolute binding free energies are estimated by a linear response approximation from thermal averaging of the structure and interaction energies by MD simulations, referred to as the linear interaction energy (LIE) method.[21–23] The molecular dynamics conformational sampling allows relaxation, or induced fit, of the ligand–receptor complex and solvation is treated microscopically. The LIE method has in several cases, with different enzymes, proven to be a useful tool for the prediction of binding affinities.[23–26] We have now employed this procedure in an attempt to predict the activities of mannitol-based compounds with the 1,2-dihydroxyethylene scaffold as inhibitors of the flexible Plm II. The ability to predict the stereochemical requirements and the character of the P1/P1′ substituents prior to synthesis seemed particularly attractive. Since it has been demonstrated previously that a β-branched carbon is preferred in the P2 position, valine methylamide was selected as a suitable P2/P2′ residue.[5,27] The computational results were compared to experimental data from a standardized Plm II assay.

We herein report that the ranking of the C_2-symmetric stereoisomers obtained by our calculations agrees with the experimental results. Furthermore, the calculations provide useful models for the 3-D structures of Plm II complexed with 1,2-dihydroxyethylene-based inhibitors. One inhibitor, comprising the optimal core configuration (*RRRR*) and a (1*S*,2*R*)-1-amino-2-indanol P2/P2′ substituent, was employed as another test compound for the computational model. An increase of inhibitor activity of greater than 40-fold compared to the corresponding valine methylamide analogue was predicted which was found to be in agreement with the experimental data.

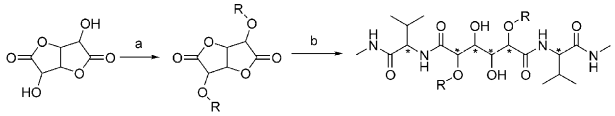
Results

Chemistry

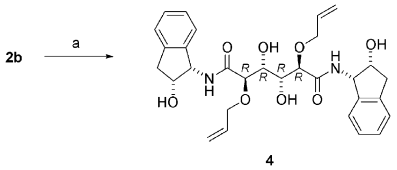
The synthetic route to the benzyloxy and allyloxy stereoisomers **3a–h** is outlined in Scheme 1. The compounds were prepared by a similar procedure to that previously reported for the synthesis of **3a**.[20] Thus, oxidation of L-mannonic γ-lactone or D-mannitol with aqueous nitric acid provided the bislactones **1a**[20] and **1b**.[28] Subsequent benzylation using commercially available benzyl trichloroacetimidate and a catalytic amount of trifluoromethanesulfonic acid gave the dibenzylated enantiomers **2a** and **2b**.[20] The allylation of the bislactones **1a** and **1b** was carried out using allyl trichloroacetimidate.[29] Trifluoromethanesulfonic acid was first employed as a catalyst and afforded the diallylated enantiomers **2c** and **2d**, although in moderate yields. A cleaner reaction occurred with boron trifluoride etherate as the catalyst, affording the diallylated bislactone **2c** in 76% yield. An excess of either D- or L-valine methylamide, which was used to open the bislactone rings, delivered the final stereoisomers **3a–h** in 47–86% yields (Scheme 1). Ring opening of bislactone **2b** with (1*S*,2*R*)-1-amino-2-indanol afforded the diamide **4** in 62% yield (Scheme 2).

Biological data

The inhibitory activities of the benzyloxy and allyloxy isomers **3a–h**, and the allyloxy indan derivative **4** were measured experimentally in a standardized Plm II enzyme assay[8] and the results are presented as K_i-values in Table 1. The role of stereochemistry proved to be crucial. Thus, only the *SRRRRS* isomers **3a** and **3c**, originating from L-mannonic γ-lactone and L-valine methylamide exhibited any inhibitory effect at concentrations up to 10 μM. A specific configuration at the four stereogenic centers of the mannitol core (*RRRR*), carrying two hydroxyl groups, of which at least one is



Scheme 1. Reagents and conditions: (a) benzyl trichloroacetimidate or allyl trichloroacetimidate, Tf-OH or $BF_3{\cdot}Et_2O$, dioxane; (b) ValNHMe, CH_2Cl_2.



Scheme 2. Reagents and conditions: (a) (1*S*,2*R*)-1-amino-2-indanol, CH_2Cl_2.

Table 1. Experimental and calculated energetics of inhibition

Compd	Config.	Plm II K_i (μM)		ΔG_{bind} (kcal/mol)			Ligand-surrounding interactions (kcal/mol)[b]			
		Expt	LIE	Score[a]	Expt	LIE	$\langle V_{l-s}^{el} \rangle_p$	$\langle V_{l-s}^{vdW} \rangle_p$	$\langle V_{l-s}^{el} \rangle_w$	$\langle V_{l-s}^{vdW} \rangle_w$
3a	*SRRRRS*	3.8	0.52	−12.8±0.7 (−12.7)	−7.4	−8.6±0.4	−75.6±0.1	−93.9±0.2	−71.8±0.8	−53.2±0.1
3b	*RRRRRR*	>10	160	−10.8±1.3 (−12.0)	nd	−5.2±0.4	−62.5±0.6	−94.8±1.0	−69.8±0.0	−52.7±0.0
3e	*SSSSSS*	>10	$7.8 \cdot 10^3$	−11.6±0.8 (−9.7)	nd	−2.9±0.4	−54.7±0.2	−96.2±1.5	−69.8±0.0	−52.7±0.0
3f	*RSSSSR*	>10	$1.5 \cdot 10^4$	−10.3±0.8 (−11.0)	nd	−2.5±0.6	−55.4±0.5	−96.9±0.7	−71.8±0.8	−53.2±0.1
3c	*SRRRRS*	4.1	11	−6.8±0.7 (−6.5)	−7.4	−6.8±0.4	−73.2±0.3	−79.6±0.0	−71.9±0.8	−44.3±0.0
3d	*RRRRRR*	>10	330	−4.5±1.2 (−7.4)	nd	−4.8±0.5	−71.1±0.9	−80.8±0.3	−77.0±0.1	−43.6±0.3
3g	*SSSSSS*	>10	$1.3 \cdot 10^8$	−4.5±0.9 (−4.6)	nd	2.9±0.7	−47.0±1.5	−82.2±0.4	−77.0±0.1	−43.6±0.3
3h	*RSSSSR*	>10	$9.4 \cdot 10^5$	−4.8±0.9 (−4.8)	nd	0.0±0.7	−51.1±1.0	−82.4±0.3	−71.9±0.8	−44.3±0.0
4	*-RRRR-*	0.096	0.24	−8.5±0.9 (−9.4)	−9.6	−9.1±0.4	−67.2±0.1	−85.9±0.1	−64.5±0.9	−40.6±0.1

[a]The calculated binding affinity using an empirical scoring function,[34] where the average value of 100 energy minimized snapshots and their standard deviation is given. In parentheses is the score of the initial minimized structures of the complexes after MD equilibration.
[b]The calculated average polar $\langle V_{l-s}^{el} \rangle$ and nonpolar $\langle V_{l-s}^{vdW} \rangle$ interactions between the ligand and its surrounding. The subscripts p and w describe the ligand in complex with the protein and free in water, respectively.

postulated to be important for coordination to the aspartic acid residues Asp34 and Asp214, is essential.[30] Furthermore, despite the *RRRR* configuration of the central part of the molecules, the compounds **3b** and **3d** with D-valine rather than L-valine in P2/P2′ were devoid of any activity. A comparison between benzyloxy and allyloxy side chains in the P1/P1′ positions, revealed no significant difference in activity, $K_i = 3.8$ and 4.1 μM respectively (c.f. **3a** and **3c**).

A replacement of valine methylamide by (1*S*,2*R*)-1-amino-2-indanol with the *RRRR* configuration of the scaffold improved the inhibitory activity 40 times (c.f. **4** and **3c**). Unfortunately, as deduced from preliminary experiments the most potent inhibitor in the series, compound **4**, exhibited no activity at 5 μM against *P. falciparum* (clone 3D7) blood-stage parasites.

Molecular dynamics simulations and binding affinity calculations

Since no X-ray structures of the complexes of Plm II with **3a–h** are available, a theoretical prediction of the binding affinities and binding conformations of the complexes was performed using computer simulations. Automated docking of the flexible ligands into the rigid, partially collapsed binding site of the X-ray structure (1SME) of Plm II did not yield useful results. The reason for this is the steric hindrance imposed by the rigid receptor model. Instead, we employed manual docking and subsequent MD relaxation of the protein–ligand complex and the average MD structures of the complexes predict that the protein undergoes very little structural changes upon binding of the ligands, except for rotation of certain side chains in order to accommodate the bulky P1′ groups. Thus, the position of the Met75-Val82 loop was not disrupted in the simulations. The binding free energies of **3a–h** were estimated using the linear interaction energy (LIE) method in combination with molecular dynamics simulations.[22] In Table 1, the calculated average ligand-surrounding interaction energies are shown together with the calculated and experimental inhibition constants. The calculations predicted that the *SRRRRS* isomers **3a** and **3c** have the strongest overall binding affinities of −8.6±0.4 and −6.8±0.4 kcal/mol, respectively. It is noteworthy that the simulations predict the absolute binding free energies quite accurately, without any reparametrization of the LIE equation as used in ref 31. From Table 1 and the LIE equation (eq 1) it is evident that the nonpolar interactions (obtained from the difference in interaction energy between the solvated protein–ligand complex and the free ligand in water) are favorable for the binding of all eight ligands with a contribution to the binding free energy of 6.4–7.9 kcal/mol. In absolute terms, all four stereoisomers of both compounds have similar nonpolar interactions with the protein, the least favorable one being, in fact, with the *SRRRRS* isomers. The electrostatic interactions display at least a 2 kcal/mol repulsive contribution to the binding affinity for all compounds except **3a** and **3c**. Thus, as apparent from Table 1, the hydrophobic interactions seem to give the largest contributions to the absolute affinity, while electrostatic interactions (mainly H-bonding) are important for the specificities. In addition, it is notable that the more hydrophobic compounds **3a**, **b**, **e** and **f** contribute with only ∼1 kcal/mol extra to the nonpolar term of the binding free energy than the smaller allyl compounds **3c**, **d**, **g** and **h**.

It is necessary that the conformational space is sufficiently well sampled in the MD simulations, since the convergence of the calculated binding free energies clearly depends on the extent of conformational sampling. This can be problematic even for moderately large ligands if they carry net charges, so that high absolute values of the interaction energies are involved.[24] Although the present inhibitors are uncharged, to deal with this problem, MD equilibration runs of up to 1.35 ns were performed for the ligands in water before data was collected and the averages were calculated on at least 1.2 ns trajectories after equilibration in order to ensure that the conformational space had indeed been adequately sampled. For the simulations of the ligands in complex with protein, data collected over 300–600 ps cover a significant amount of the conformational space around each equilibrated starting conformation, which was assessed by comparing the average energies for the first and second half of each trajectory.

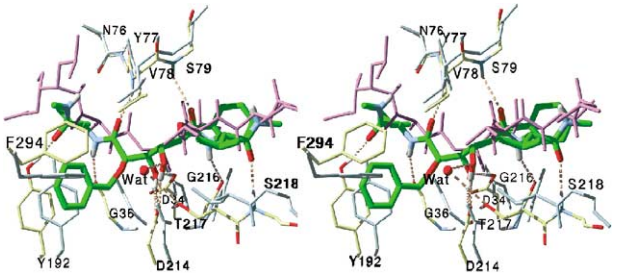


Figure 1. Stereo view of the predicted (MD average) binding conformation of inhibitor **3a** (green) in complex with Plm II (yellow), superimposed on the X-ray structure of pepstatin A (purple) complexed with Plm II (light blue).[3] 'Wat' denotes the water molecule that was found to interact with the P hydroxyl group as well as the side chain of Thr217 in the MD simulations.

To demonstrate the predicted binding conformation of the most potent methylamide inhibitor **3a**, in Figure 1 an average MD structure from 50 ps simulation is shown superimposed on the X-ray structure of Plm II in complex with pepstatin A. For clarity reasons, all water molecules not in direct contact with the ligand and many residues that have nonpolar interactions with the ligand are omitted from the picture. Figure 2 shows a scheme of the interactions in the active site of the same complex together with a similar schematic picture of compound **4**. The compounds **3a** and **3c** display the same binding conformation and the protein structure is, in both cases, very similar to the X-ray structure. In order to make space for the benzyloxy P1′ group, which was docked into the partially collapsed S1′ pocket of the X-ray structure, the side chain of Phe294 is rotated towards Leu292, which in turn rotates towards Ile290. The Phe294 rotation is the only major side chain displacement observed when accommodating the bulky P1′ group. The P1′ side chain is also in close contact with side chain atoms of Val78, Tyr192, Ile212, Thr298 and Ile300, and the P1′ phenyl group of **3a** also has stacking interactions with the aromatic ring of Tyr192. The P2′ valine group is positioned very similarly to the Ala group on pepstatin A, pointing into S2′. The P1 hydrophobic group is positioned in the S1 pocket and turns the phenyl (allyl in **3c**) group towards S3 in the continuous S1–S3 cleft, while the P2 valine group points into S2, albeit slightly displaced in the S3–S4 direction compared to the P2 group of pepstatin A. The protein–ligand backbone H-bond pattern obtained from the simulations is shown in Figure 1 and is similar to that of the initial X-ray structure for the P′ part of the ligand, where the **3a** and pepstatin A backbones are virtually superimposed on each other. A slight rotation of the P′ methylamide group and the loss of H-bond interaction with the Asn76 carbonyl oxygen in the simulated structure is the largest difference between the simulated and the experimental structure. Although the amide group has rotated, the H-bond between the hydroxyl group of Tyr192 and the P2′ carbonyl oxygen is maintained as in the X-ray structure of the Plm II–pepstatin A complex. Furthermore, in the simulated structure the P2′ amide nitrogen interacts with the carbonyl of Gly36 and the nitrogen of the loop residue Val78 donates its H-bond to the P1′ carbonyl. The P1′ hydroxyl is oriented in the same fashion as in the crystal structure: it is located between the carboxylates of Asp34 and Asp214, and donates its H-bond to the negatively charged Asp214. Since the direction of the amide bonds is reversed for the P part of the ligand, the H-bonding interaction pattern at these sites differs from that of the structure of the pepstatin A complex. Here, the P1 hydroxyl donates its H-bond to Asp214 and this hydroxyl also accepts an

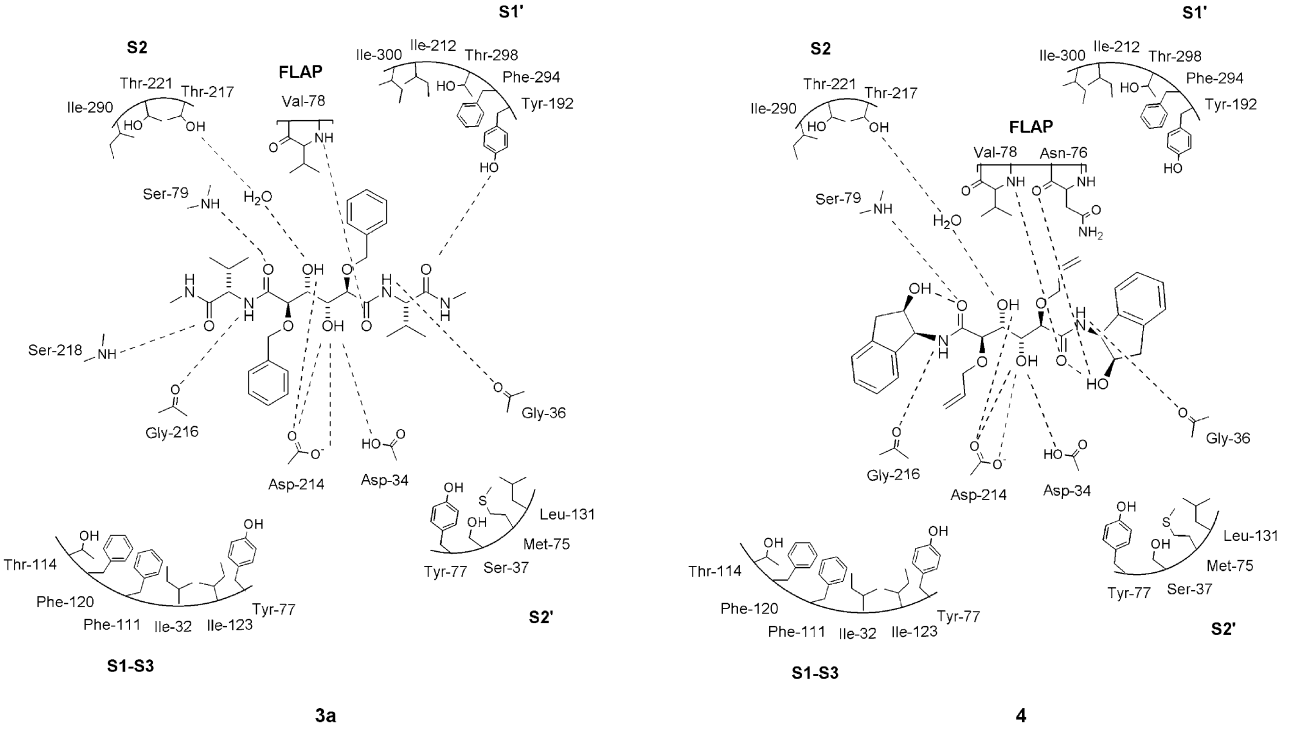


Figure 2. Schematic picture showing the predicted (MD average) interactions of inhibitor **3a** and **4** with the active site.

H-bond from a water molecule. This water molecule donates an H-bond to the sidechain of Thr217 and accepts one from a second water molecule in the binding site (not shown). The MD simulations further suggest that the P1 carbonyl oxygen accepts an H-bond from Ser79 N, and the P2 nitrogen and carbonyl oxygen make H-bonds with Gly216 O and Ser218 N, respectively.

Since the present ligands exhibit only moderate inhibition of Plm II, we aimed at modifying them to create more potent inhibitors. 2-Indanol has been used previously as an effective P2 substituent in HIV-1 protease inhibitors and this substituent was therefore also considered here.[20,32] Thus, we decided to predict the binding constant for **4** where the terminal methylamide had been removed and the P2/P2′ valines had been substituted with (1*S*,2*R*)-1-amino-2-indanol. The result, shown in Table 1, is a 2.3 kcal/mol more negative binding free energy (or a 47 times stronger inhibition constant) than for **3c**. An experimental measurement of the potency of **4** gave a corresponding increase in the binding affinity of 2.2 kcal/mol (or 43 times greater inhibition) relative to **3c**, which essentially confirms the computational result. The simulated structure of the complex with **4** (Fig. 3) reveals a binding mode similar to that of **3a** and **3c**, with the central parts (both hydroxyls, P1/P1′ allyloxy chains, P1/P1′ carbonyls) in essentially the same position as in the valine residue compounds. In addition, the water molecule that donates H-bonds to the P1 and Thr217 hydroxyl oxygen atoms in the simulations of the complexes with **3a** and **3c** was also found in the complex with **4**. The resulting hydrogen bonding patterns of the P2/P2′ groups differ, however, from the case with **3a** and **3c**. Figure 3 shows the H-bonds that are present over at least ~50% of the MD simulation superimposed on the X-ray structure of Plm II in complex with pepstatin A. The P2′ indan ring system in **4** is situated along the position of the **3c** backbone, with its five-membered ring at the valine side-chain site and its aromatic ring along the methylamide group of **3c**. The indan hydroxyl alternates between donating an intramolecular H-bond to the P1′ carbonyl oxygen and one to the Asn76 carbonyl oxygen. Occasionally, a water molecule and the Ser37 side chain also donate H-bonds to this hydroxyl group. The position of the P2′ indan ring system remains the same throughout the whole simulation, as the corresponding part of **3c**. On the other hand, the P2 group is more flexible, which is in contrast to the simulation with **3c**, where the P2 group and methylamide remain essentially in the same place during the whole simulation. In the simulation with **4** the P2 indan ring hydroxyl group interacts with solvent and also forms an intramolecular H-bond, similar to the P2′ group. The indan rings exhibit nonpolar interactions with the side chains of Thr114, Ile290, Ser79 and the P1 group, while one face of the indan moiety is exposed to the solvent.

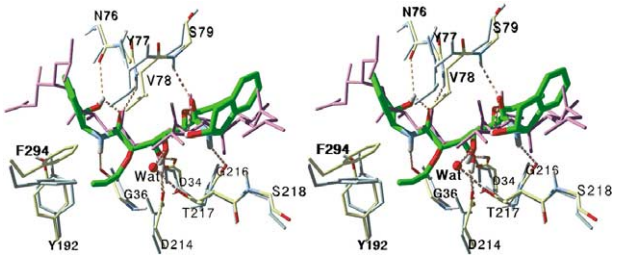


Figure 3. Stereo view of the predicted (MD average) binding conformation of inhibitor **4** (green) in complex with Plm II (yellow), superimposed on the X-ray structure of pepstatin A (purple) complexed with Plm II (light blue).[3] 'Wat' denotes the water molecule that was found to interact with the P hydroxyl as well as the side chain of Thr217 in the MD simulations.

As a comparison to the LIE results the binding affinities were also estimated using an empirical scoring function.[33] This function takes into account H-bonding and lipophilic interactions, as well as the loss of rotational degrees of freedom upon ligand binding (see the Methods section). While scoring functions in general are intended for rapid binding affinity estimates of a single conformation, we decided also to examine the effects of averaging over several snapshots from the MD trajectories since this has been shown previously to improve the results.[34] Hence, the initial equilibrated complexes of Plm II with **3a–h** and **4** were first minimized and scored. Then, 100 snapshots from the MD trajectory of each complex were minimized and scored and the results were averaged. In Table 1, the average values with standard deviations are shown together with the single snapshot scores. Within the isomer series of the allyloxy compounds, a correct affinity ranking was obtained when averaging over snapshots was performed, while scoring only the initial minimized structures did not yield the correct result. The compound **3c** was thus predicted to be the most active inhibitor with a score within 0.6 kcal/mol from the experimental value, using the averaging approach. In addition, the indan compound **4** is reasonably well predicted by the scoring function. The affinities within the benzyloxy series were also correctly ranked. However, all of them have predicted affinities to Plm II that are too great when compared to the experimental values and the average affinity of **3a** is overpredicted by 5.4 kcal/mol. The overestimated high affinity for the benzyloxy series can be attributed to the lipophilic term of the scoring function, which clearly overestimates the hydrophobic binding contribution (predicted to about −10 kcal/mol). This may indicate a more general problem with scoring functions that have a built-in size-dependency of the hydrophobic term through surface area or similar size measures.

Discussion

The ranking of the predicted binding affinities, in addition to the absolute values, agree well with the experimental finding that only the *SRRRRS* isomers **3a** and **3c** display inhibitory effects in the Plm II assay. In addition, among these isomers the same stereo configuration (*SRRRRS*) is the only one that produces active inhibitors in a standardized HIV-1 protease assay.[35,36] An X-ray structure of **3c** with Val substituted for Ile in complex with HIV-1 protease shows that both hydroxyl groups point towards one of the catalytic aspartates.[20]

Also, in the X-ray structure 1SME, the configurations of the pepstatin A P1′ hydroxyl and P2′ Ala correspond to the *SRRRRS* stereoisomers. The LIE model for this class of ligands uses $\alpha = 0.181$, $\beta = 0.33$ and $\gamma = 0$ and was parameterized in ref 23. It provides an excellent estimation of the absolute binding free energies of the methylamide compounds and also captures the effect of the P2/P2′ modified compound **4**. Overall, this LIE model has indeed proven successful in several other cases with the existing parametrization.[24,26,37,38] The simulated structure of Plm II in complex with the ligands **3a–h** and **4** remains essentially as in the X-ray structure, although movement of some side chains was observed in the S1′ pocket upon accommodation of the P1′ group. In addition, a water molecule not present in the X-ray structure of the Plm II/pepstatin A complex was found to bridge between the ligand hydroxyl group and a threonine residue in the simulated complexes of **3a**, **3c**, and **4** as well as in those with **3d** and **3e**. It is noteworthy that the water molecule is not introduced into the binding site in the initial solvation of the complex, but enters during MD equilibration and then remains at its location. In the Plm II/pepstatin A crystal structure, the corresponding position is occupied by the hydrophobic P2 side chain. It may be of interest to note that the intramolecular H-bonds found in compound **4** are neither observed in HIV-1 protease complexes with similar ligands nor with the existing anti-HIV drug Indinavir.[20,39]

In a paper by Haque et al.[8] a crystal structure is described of Plm II in complex with a hydroxyethylamine-based inhibitor that contains a bulky P1′ group. In this structure, which has very recently been made available under the PDB code 1LF3,[40] the authors observe that the distance between the side chain of Val78 and the Phe294 phenyl is 3 Å longer than in the X-ray structure with pepstatin A. It is the backbones of the Met75-Val82 and Leu292-Pro197 loops that move when incorporating the ligand with the bulky P1′ group. In the present study, however, the bulky P1′ groups did not induce any substantial conformational changes of the loops, but the S1′ pocket was expanded to accommodate the P1′ group mainly by movement of the Phe294 side chain. Other very recently published structures are a Plm II/pepstatin A complex, 1M43 solved at 2.4 Å,[40] and two complexes of Met205Ser mutants of Plm II with inhibitors (1LEE and 1LF2).[14] Also a structure of uncomplexed Plm II has been solved (1LF4) at 1.90 Å resolution has become available.[40]

Since the inhibitors **3a–h** cause only moderate inhibition of Plm II, we investigated the effect on the binding constant of a scaffold modification where the P2/P2′ valines of inhibitor **3a** were substituted with (1*S*,2*R*)-1-amino-2-indanol and the terminal methylamide groups were removed. This change introduced large, hydrophobic groups and reduced the peptide-like properties as well as the number of torsional degrees of freedom. The computational result predicts that such a modification yields a 47 times more potent inhibitor than the reference compound **3c**, an effect that was also observed experimentally.

Apart from the LIE method, there are several other techniques available for calculation of protein–ligand binding affinities. The free energy perturbation (FEP) approach[41] has often been used as a tool to calculate differential binding affinities between chemically similar ligands. The method is usually reliable for calculating the effects on binding affinity resulting from small perturbations, for example the addition of a functional group or change of atom type. However, in the present case FEP would not be a realistic choice since the size of the perturbations involved, that is the addition of a phenyl or indan ring, would cause severe convergence problems. Another simplified method for calculating protein–ligand binding affinities is the MM-PBSA approach,[42] which has been successful in predicting absolute binding free energies in several cases.[43] An important difference between the LIE method and MM-PBSA is that the latter employs a continuum solvent description in the final binding energy calculation. While this may improve the convergence, the energetic details of solvation are lost. For instance, the water molecule that was found to bridge between the ligand and Plm II would not give any contribution. Furthermore, the usual implementation of MM-PBSA assumes that the conformations explored by the ligand are the same in the bound and free states and only the complex is simulated. This assumption does not appear warranted in the present case since our ligands are considerably more conformationally restricted in the complexes than in solution. Empirical scoring functions, such as that of Böhm[44] and Eldridge et al.,[33] may be used for estimation of binding free energies on docked complexes. Here, we employed Eldridge's function for scoring snaphots of the different complexes. The results correctly rank the affinities within the isomer series, provided that conformational averaging is carried out, and the affinity of **4** is also reasonably well predicted. The scoring function, however, largely overpredicts the affinities of all benzyloxy ligands, especially those having no experimentally detectable affinity. The large lipophilic term that dominates the scores of the benzyloxy compounds is the main source of the systematic over prediction. Regardless of which method is used for binding energy calculations it seems that conformational averaging tends to improve the result.[34] This is, of course, particularly true for force-field based scoring approaches, such as LIE, since the energy fluctuations due to thermal motion can be substantial. While such methods are more time-consuming than empirical scoring functions their predictive power may still justify the extra effort, as the present calculations attest.

Conclusion

In summary, C_2-symmetric compounds encompassing a mannitol-based scaffold, previously used successfully in HIV-1 protease inhibitors, have been theoretically and experimentally examined as core structures in Plm II inhibitors. Four different stereoisomers with either benzyloxy or allyloxy P1/P1′ side chains were studied. Computational ranking of the binding affinities of the eight compounds was performed and the isomers were

synthesized and evaluated in a Plm II enzyme assay. Since the ligand binding site in the X-ray structure of the protein in complex with pepstatin A is partially collapsed, manual docking of the ligands with subsequent relaxation of the complex was employed rather than automated docking into a rigid protein. Binding affinities were estimated by the linear interaction energy (LIE) method. Within both series of isomers the theoretical binding energies were in agreement with the enzymatic measurements, illustrating the usefulness of the LIE method for the prediction of ligand affinities prior to synthesis. Furthermore, the structural models of the enzyme–inhibitor complexes obtained from the MD simulations provide a framework for the interpretation of further structure–activity relationships. Hence, the affinity of a similar ligand with a different P2/P2′ substituent was estimated using the same procedure, and the predicted improvement in binding constant agrees well with experimental results. In the future, we aim to further improve this diol type of Plm II inhibitors.

Experimental

Chemistry

General information. ^{1}H and ^{13}C NMR spectra were recorded on a Jeol JNM-EX 270 spectrometer at 270.2 and 67.8 MHz, respectively, or on a JEOL JNM-EX400 spectrometer at 399.8 and 100.5 MHz, respectively. Chemical shifts are reported as δ values (ppm) indirectly referenced to TMS by the solvent residual signal. Optical rotations were obtained on a Perkin-Elmer 241 polarimeter. Specific rotations (α_D) are reported in deg/dm, and the concentration (*c*) is given in g/100 mL in the specified solvent. Elemental analysis was performed by Mikro Kemi AB, Uppsala, Sweden. Flash column chromatography was performed on Merck silica gel 60, 0.04–0.063 mm. Thin-layer chromatography was performed using aluminium sheets precoated with silica gel 60 F_{254} (0.2 mm; E. Merck) and visualized with UV light, permanganate or ninhydrin. RP-FPLC was performed on a PepRPC 15 μm HR 16/10 column (Pharmacia Biotech AB) at a flow rate of 3.5 mL/min. SP-HPLC was performed on a Dynamax 60-A 21.4×250 mm, 8 μm silica gel column (Rainin Instrument co. inc) at a flow rate of 20 mL/min.

2,5-Di-*O*-benzyl-D-mannaro-1,4:3,6-di-γ-lactone (2c). D-1,4:3,6-Mannarodilactone **1b** (303 mg, 1.7 mmol) was dissolved in dry dioxane (40 mL) under a nitrogen atmosphere. To the stirred solution was added benzyl trichloroacetimidate (1.4 g, 5.3 mmol) and a catalytic amount of trifluoromethanesulfonic acid (250 μL). The solution was stirred at room temperature for 5 h. After filtration through a pad of K_2CO_3 (2 cm) on silica (2 cm) the filtrate was concentrated and dried under vacuum overnight. The crude product was washed with diethyl ether (2×50 mL) to give pure product **2c** (420 mg, 68%) as white crystals. $[\alpha]_D^{23}$ +129.7° (*c* 1.04, CH_2Cl_2); ^{1}H NMR (DMSO-d_6) δ 7.42–7.31 (m, 10H), 5.27 (AA′ part of AA′XX′, 2H), 4.88 (XX′ part of AA′XX′, 2H), 4.80 (d, *J* = 11.75 Hz, 2H), 4.73 (d, *J* = 11.75 Hz, 2H); ^{13}C NMR (DMSO-d_6) δ 171.8, 136.9, 128.4, 128.1, 128.0, 74.8, 74.4, 72.1. Anal. calcd for $C_{20}H_{18}O_6$: C, 67.79; H, 5.12. Found: C, 67.4; H, 5.4.

2,5-Di-*O*-allyl-L-mannaro-1,4:3,6-di-γ-lactone (2b). L-1,4:3,6-Mannarodilactone **1a** (1.1 g, 6.0 mmol) was dissolved in dry dioxane (50 mL) under a nitrogen atmosphere. The mixture was heated until the dilactone was completely dissolved, and the solution was then allowed to attain room temperature. To the stirred solution was added allyl trichloroacetimidate (3.7 g, 18 mmol) and a catalytic amount of $BF_3{\cdot}Et_2O$ (200 μL). The solution was stirred at room temperature for 3 h. After filtration through a silica pad the filtrate was concentrated and dried under vacuum overnight. The crude product was washed with diethyl ether (3×50 mL) to give pure product **2b** (1.2 g, 76%) as a white solid: $[\alpha]_D^{22}$ −213.15° (*c* 1.034,CH_3CN); ^{1}H NMR (DMSO-d_6) δ 5.93 (dddd, *J* = 2.83, 5.63, 10.39, 17.25 Hz, 2H), 5.36 (ddd, *J* = 2.01, 3.98, 17.25 Hz, 2H), 5.28 (AA′ part of AA′XX′, 2H), 5.25 (ddd, *J* = 1.74, 3.48, 10.39 Hz, 2H), 4.81 (XX′ part of AA′XX′, 2H), 4.31–3.15 (m, 4H); ^{13}C NMR (DMSO-d_6) δ 171.8, 133.9, 118.1, 74.8, 74.4, 71.2. Anal. calcd for $C_{12}H_{14}O_6{\cdot}1/4H_2O$: C, 55.70; H, 5.65. Found: C, 55.7; H, 5.5.

2,5-Di-*O*-allyl-D-mannaro-1,4:3,6-di-γ-lactone (2d). D-1,4:3,6-Mannarodilactone **1b** (101 mg, 0.58 mmol) was dissolved in dry dioxane (6 mL) under a nitrogen atmosphere. The mixture was heated until the dilactone was completely dissolved, and the solution was then allowed to attain room temperature. To the stirred solution was added allyl trichloroacetimidate (444 mg, 2.3 mmol) and a catalytic amount of trifluoromethanesulfonic acid (20 μL). The solution was stirred at room temperature for 7 h. After filtration through a pad of Na_2CO_3 (2 cm) on silica (1 cm) the filtrate was concentrated and dried under vacuum overnight. The crude product was purified on RP-FPLC (15 min gradient of 30–45% CH_3CN in 0.1% aqueous TFA) to give pure product 2d (76 mg, 51%) as a white solid: $[\alpha]_D^{22}$ +213.8° (*c* 1.034, CH_3CN); ^{1}H NMR (DMSO-d_6) δ 5.93 (dddd, *J* = 2.85, 5.69, 10.31, 17.24, 2H), 5.36 (ddd, *J* = 1.98, 3.92, 17.24 Hz, 2H), 5.28 (AA′ part of AA′XX′, 2H), 5.25 (ddd, *J* = 1.73, 3.51, 10.31 Hz, 2H), 5.81 (XX′ part of AA′XX′, 2H), 4.31–4.15 (m, 4H); ^{13}C NMR (DMSO-d_6) δ 171.8, 133.9, 118.1, 74.8, 74.4, 71.2. Anal. calcd for $C_{12}H_{14}O_6$: C, 56.69; H, 5.55. Found: C, 56.3; H, 5.7.

General procedure for preparation of the compounds 3b and 3e–f

Method I. Bislactone **2a** or **2c** was dissolved in dry CH_2Cl_2 (2 mL) under a nitrogen atmosphere, and the solution was cooled to 0 °C. Four equivalents of D- or L-valine methylamide in dry CH_2Cl_2 (2 mL) was added and the reaction mixture was allowed to attain room temperature. Stirring was continued for 9–21 h at room temperature and for another 11–15 h at reflux. Purification by flash chromatography or SP-HPLC gave the products **3b** and **3e–f**.

(2*R*,3*R*,4*R*,5*R*)-*N*1,*N*6-Bis[(1*R*)-2-methyl-1-(methylcarbamoyl)propyl]-2,5-di(benzyloxy)-3,4-dihydroxyhexane-1,6-diamide (3b). The title compound was prepared according to Method I, using 57 mg (0.16 mmol) of bislactone **2a** (stirring 9 h at room temperature and 15 h at reflux). Purification by flash chromatography (CH_2Cl_2/CH_3OH, 20:1–10:1) gave the title compound **3b** (52 mg, 53%) as a white solid: $[\alpha]_D^{22}$ +48.1° (*c* 0.53, $CHCl_3$/CH_3OH, 10:1); ^1H NMR ($CDCl_3$/CD_3OD, 4:2) δ 7.36–7.22 (m, 10H), 4.62 (d, *J* = 11.13 Hz, 2H), 4.49 (d, *J* = 11.13 Hz, 2H), 4.23 (d, *J* = 6.06 Hz, 2H), 4.13 (app s, 4H), 2.71 (s, 6H), 2.14 (dh, *J* = 6.06, 6.89 Hz, 2H), 0.92 (d, *J* = 6.89 Hz, 6H), 0.86 (d, *J* = 6.89 Hz, 6H); ^{13}C NMR ($CDCl_3$/CD_3OD, 4:2) δ 172.9, 172.8, 137.4, 129.1, 128.0, 128.7, 80.0, 73.4, 71.8, 58.9, 31.1, 26.3, 19.6, 17.9. Anal. calcd for $C_{32}H_{46}N_4O_8 \cdot 1/2H_2O$: C, 61.62; H, 7.60; N, 8.98. Found: C, 61.3; H, 7.4; N, 8.7.

(2*S*,3*S*,4*S*,5*S*)-*N*1,*N*6-Bis[(1*S*)-2-methyl-1-(methylcarbamoyl)propyl]-2,5-di(benzyloxy)-3,4-dihydroxyhexane-1,6-diamide (3e). The title compound was prepared according to Method I, using 51 mg (0.14 mmol) of bislactone **2c** (stirring 10 h at room temperature and 13 h at reflux). Purification by flash chromatography (EtOAc/CH_3OH, 50:1–10:1) provided product **3e** (49 mg, 55%) as a white solid: $[\alpha]_D^{22}$ −47.4° (*c* 0.53, $CHCl_3$/CH_3OH, 10:1); ^1H NMR ($CDCl_3$/CD_3OD, 4:2) δ 7.36–7.23 (m, 10H), 4.62 (d, *J* = 11.22 Hz, 2H), 4.49 (d, *J* = 11.22 Hz, 2H), 4.23 (d, *J* = 6.02 Hz, 2H), 4.01 (app s, 4H), 2.71 (s, 6H), 2.15 (dh, *J* = 6.02, 6.89 Hz, 2H), 0.92 (d, *J* = 6.89 Hz, 6H), 0.87 (d, *J* = 6.89 Hz, 6H); ^{13}C NMR ($CDCl_3$/CD_3OD, 4:2) δ 172.9, 172.7, 137.3, 129.1, 128.9, 128.7, 80.0, 73.4, 71.7, 58.9, 31.1, 26.3, 19.6, 17.9. Anal. calcd for $C_{32}H_{46}N_4O_8 \cdot 11/2H_2O$: C, 59.89; H, 7.70; N, 8.73. Found: C, 59.6; H, 7.3; N, 8.4.

(2*S*,3*S*,4*S*,5*S*)-*N*1,*N*6-Bis[(1*R*)-2-methyl-1-(methylcarbamoyl)propyl]-2,5-di(benzyloxy)-3,4-dihydroxyhexane-1,6-diamide (3f). The title compound was prepared according to Method I, using 51 mg (0.15 mmol) of bislactone **2c** (stirring 21 h at room temperature and 11 h at reflux). Purification by SP-HPLC ($CHCl_3$ to $CHCl_3$/CH_3OH, 50:1) gave product **3f** (42 mg, 47%) as a white solid: $[\alpha]_D^{22}$ +3.9° (*c* 0.58, $CHCl_3$/CH_3OH, 10:1); ^1H NMR ($CDCl_3$/CD_3OD, 4:2) δ 7.37–7.24 (m, 10H), 4.65 (d, *J* = 11.55 Hz, 2H), 4.58 (d, *J* = 11.55 Hz, 2H), 4.21 (d, *J* = 5.57 Hz, 2H), 4.14 (AA′ part of AA′XX′, 2H), 4.05 (XX′ part of AA′XX′, 2H), 2.69 (s, 6H), 2.24 (dh, *J* = 5.57, 6.85 Hz, 2H), 0.90 (d, *J* = 6.85 Hz, 6H), 0.82 (d, *J* = 6.85 Hz, 6H); ^{13}C NMR ($CDCl_3$/CD_3OD, 4:2) δ 172.6, 172.4, 137.4, 129.1, 128.8, 128.6, 80.7, 73.4, 72.3, 58.7, 30.5, 26.3, 19.8, 17.7. Anal. calcd for $C_{32}H_{46}N_4O_8 \cdot 1/2H_2O$: C, 61.62; H, 7.60; N, 8.98. Found: C, 61.3; H, 7.4; N, 8.7.

General procedure for preparation of the compounds 3c–d, 3g–h

Method II. Bislactone **2b** or **2d** was dissolved in dry CH_2Cl_2 under a nitrogen atmosphere. Four equivalents of D- or L-valine methylamide in dry CH_2Cl_2 was added and the reaction mixture was stirred at reflux for 21–50 h. Concentration under reduced pressure and different workup procedures gave the pure products **3c–d**, **3g–h**.

(2*R*,3*R*,4*R*,5*R*)-*N*1,*N*6-Bis[(1*S*)-2-methyl-1-(methylcarbamoyl)propyl]-2,5-di(allyloxy)-3,4-dihydroxyhexane-1,6-diamide (3c). The title compound was prepared according to Method II, using 179 mg (0.71 mmol) of bislactone **2b** and totally 9 mL dry CH_2Cl_2 (reflux 21 h). The crude product was suspended in H_2O (5 mL), filtered, dissolved in $CHCl_3$/CH_3OH, 9:1 (20 mL) and filtered through a pad of silica (1.5 cm). The pad was further washed with $CHCl_3$/MeOH, 9:1 (30 mL). Concentration of the filtrate under reduced pressure gave product **3c** (243 mg, 67%) as a white solid: $[\alpha]_D^{22}$ +8.4° (*c* 0.44, $CHCl_3$/CH_3OH, 10:1); ^1H NMR ($CDCl_3$/CD_3OD, 4:2) δ 5.91 (dddd, *J* = 2.80, 5.61, 10.35, 17.24 Hz, 2H), 5.31 (ddd, *J* = 1.94, 3.67, 17.24 Hz, 2H), 5.21 (ddd, *J* = 1.32, 2.80, 10.35 Hz, 2H), 4.22 (d, *J* = 5.40 Hz, 2H), 4.18–4.01 (m, 6H), 3.95 (XX′ part of AA′XX′, 2H), 2.71 (s, 6H), 2.27 (dh, *J* = 5.40, 6.89 Hz, 2H), 0.94 (d, *J* = 6.89 Hz, 6H), 0.87 (d, *J* = 6.89 Hz, 6H); ^{13}C NMR ($CDCl_3$/CD_3OD, 4:2) δ 172.7, 172.5, 134.0, 118.6, 80.5, 72.4, 72.3, 58.7, 30.4, 26.3,19.8, 17.6. Anal. calcd for $C_{24}H_{42}N_4O_8$: C, 56.01; H, 8.23; N, 10.89. Found: C, 56.0; H, 8.3; N, 10.7.

(2*R*,3*R*,4*R*,5*R*)-*N*1,*N*6-Bis[(1*R*)-2-methyl-1-(methylcarbamoyl)propyl]-2,5-di(allyloxy)-3,4-dihydroxyhexane-1,6-diamide (3d). The title compound was prepared according to Method II, using 39 mg (0.71 mmol) of bislactone **2b** and totally 4 mL dry CH_2Cl_2 (reflux 30 h). Purification by SP-HPLC ($CHCl_3$) gave product **3d** (68 mg, 86%) as a white solid: $[\alpha]_D^{21}$ +73.1° (*c* 0.50, $CHCl_3$); ^1H NMR ($CDCL_3$/CD_3OD, 4:2) δ 5.89 (dddd, *J* = 2.89, 5.86, 10.35, 17.24 Hz, 2H), 5.28 (ddd, *J* = 1.57, 3.13, 17.24 Hz, 2H), 5.19 (ddd, *J* = 1.32, 2.80, 10.35 Hz, 2H), 4.24 (d, *J* = 6.02 Hz, 2H), 4.14–3.97 (m, 4H), 3.95 (app s, 4H), 2.72 (s, 6H), 2.16 (dh, *J* = 6.02, 6.85 Hz, 2H), 0.93 (d, *J* = 6.85 Hz, 6H), 0.88 (d, *J* = 6.85 Hz, 6H); ^{13}C NMR ($CDCL_3$/CD_3OD, 4:2) δ 172.9, 172.7, 134.0, 118.9, 79.8, 72.4, 71.6, 58.8, 31.0, 26.2, 19.6, 17.8. Anal. calcd for $C_{24}H_{42}N_4O_8$: C, 56.01; H, 8.23; N, 10.89. Found: C, 55.7; H, 8.1; N, 10.8.

(2*S*,3*S*,4*S*,5*S*)-*N*1,*N*6-Bis[(1*S*)-2-methyl-1-(methylcarbamoyl)propyl]-2,5-di(allyloxy)-3,4-dihydroxyhexane-1,6-diamide (3g). The title compound was prepared according to Method II, using 17 mg (0.067 mmol) of bislactone 2d and totally 4 mL dry CH_2Cl_2 (reflux 50 h). Purification by SP-HPLC ($CHCl_3$ to $CHCl_3$/CH_3OH, 50:1) gave product **3g** (26 mg, 75%) as a white solid: $[\alpha]_D^{21}$ −73.1° (*c* 0.50, $CHCl_3$); ^1H NMR ($CDCl_3$/CD_3OD, 4:2) δ 5.89 (dddd, *J* = 2.89, 5.90, 10.35, 17.24 Hz, 2H), 5.28 (ddd, *J* = 1.53, 3.09, 17.24 Hz, 2H), 5.19 (ddd, *J* = 1.32, 2.80, 10.35 Hz, 2H), 4.24 (d, *J* = 6.02 Hz, 2H), 4.14–3.98 (m, 4H), 3.96 (app s, 4H), 2.72 (s, 6H), 2.16 (dh, *J* = 6.06, 6.85 Hz, 2H), 0.93 (d, *J* = 6.85 Hz, 6H), 0.89 (d, *J* = 6.85 Hz, 6H); ^{13}C NMR ($CDCl_3$/CD_3OD, 4:2) δ 173.0, 172.8, 134.1, 118.9, 79.8, 72.4, 71.7, 58.8, 31.1, 26.3, 19.6, 17.8. Anal. calcd for $C_{24}H_{42}N_4O_8 \cdot 1/4H_2O$: C, 55.53; H, 8.25; N, 10.79. Found: C, 55.3; H, 8.1; N, 10.5.

(2*S*,3*S*,4*S*,5*S*)-*N*1,*N*6-Bis[(1*R*)-2-methyl-1-(methylcarbamoyl)propyl]-2,5-di(allyloxy)-3,4-dihydroxyhexane-1,6-diamide (3h). The title compound was prepared according to Method II, using 47 mg (0.19 mmol) of bislactone

2d and totally 4 mL dry CH_2Cl_2 (reflux 24 h). Purification by SP-HPLC ($CHCl_3$ to $CHCl_3/CH_3OH$, 50:1) gave product **3h** (45 mg, 47%) as a white solid: $[\alpha]_D^{22}$ −8.58° (*c* 0.44, $CHCl_3/CH_3OH$, 10:1); ^1H NMR ($CDCl_3/CD_3OD$, 7:3) δ 5.90 (dddd, *J* = 2.85, 5.61, 10.31, 17.24 Hz, 2H), 5.31 (ddd, *J* = 1.94, 3.67, 17.24 Hz, 2H), 5.21 (ddd, *J* = 1.32, 2.85, 10.31 Hz, 2H), 4.22 (d, *J* = 5.40 Hz, 2H), 4.18–4.01 (m, 6H), 3.95 (XX′ part of AA′XX′, 2H), 2.70 (s, 6H), 2.28 (dh, *J* = 5.40, 6.89 Hz, 2H), 0.94 (d, *J* = 6.89 Hz, 6H), 0.87 (d, *J* = 6.89 Hz, 6H); ^{13}C NMR ($CDCl_3/CD_3OD$, 4:2) δ 172.5, 172.3, 133.8, 118.6, 80.4, 72.4, 72.2, 58.6, 30.2, 26.2, 19.8, 17.5. Anal. calcd for $C_{24}H_{42}N_4O_8 \cdot 1/4H_2O$: C, 55.53; H, 8.25; N, 10.79. Found: C, 55.2; H, 8.1; N, 10.6.

(2*R*,3*R*,4*R*,5*R*)-*N*1,*N*6-Bis[(1*S*,2*R*)-2-hydroxy-1-indanyl]-2,5-di(allyloxy)-3,4-dihydroxyhexane-1,6-diamide (4). A mixture of 2,5-di-*O*-allyl-L-mannaro-1,4:3,6-di-γ-lactone (1.2 g, 4.5 mmol) and (1*S*,2*R*)-1-amino-2-indanol (2.7 g, 18.2 mmol) in dry CH_2Cl_2 (115 mL) was heated to reflux for 16 h under a nitrogen atmosphere. The solvent was removed under reduced pressure and purification by flash chromatography ($CHCl_3/CH_3OH$, 100:3) gave product **4** (1.6 g, 62%) as a white solid: $[\alpha]_D^{30}$ −22.6° (*c* 1.00, $CHCl_3$) ^1H NMR ($CDCl_3$) δ 7.36 (d, *J* = 8.87 Hz, 2H), 7.27–7.18 (m, 8H), 5.91 (ddt, *J* = 5.77, 10.35, 17.28 Hz, 2H), 5.34 (ddd, *J* = 1.52, 3.05, 17.28 Hz, 2H), 5.32–5.24 (m, 2H), 5.23 (ddd, *J* = 1.32, 2.76, 10.35 Hz, 2H), 4.62–4.55 (m, 2H), 4.41 (app s, 4H), 4.21–4.16 (m, 8H), 3.06 (dd, *J* = 5.40, 16.66 Hz, 2H), 2.86 (dd, *J* = 1.94, 16.66 Hz, 2H); ^{13}C NMR ($CDCl_3$) δ 171.6, 140.8, 139.8, 133.2, 128.2, 126.9, 125.3, 123.8, 118.6, 80.9, 72.4, 72.2, 71.5, 57.7, 39.2. Anal. calcd for $C_{30}H_{36}N_2O_8 \cdot 1/4H_2O$: C, 64.68; H, 6.60; N, 5.03. Found: C, 64.6; H, 6.5; N, 5.1.

Plasmepsin assay and K_i determination

Pro-Plm II was a generous gift from Helena Danielson (Department of Biochemistry, Uppsala University, Uppsala, Sweden). The activity of Plm II was measured essentially as described previously,[8] using a total reaction volume of 100 μL. The concentration of pro-Plm II was 3 nM. The pro-sequence of Plm II was cleaved off by preincubation in assay reaction buffer (100 mM sodium acetate buffer (pH 4.5), 10% glycerol and 0.01% Tween 20) at room temperature for 40 min. The reaction was initiated by the addition of 3 μM substrate (DABCYL-Glu-Arg-Nle-Phe-Leu-Ser-Phe-Pro-EDANS, AnaSpec Inc, San Jose, CA, USA) and hydrolysis was recorded as the increase in fluorescence intensity over a 10-min time period, during which the rate increased linearly with time.

Stock solutions of inhibitors in DMSO were serially diluted in DMSO and added directly before addition of substrate, giving a final DMSO concentration of 1%.

IC_{50} values were obtained by assuming competitive inhibition and fitting a Langmuir isotherm [$v_i/v_o = 1/(1 + [I]/IC_{50})$] to the dose response data (Grafit), where v_i and v_o are the initial velocities for the inhibited and uninhibited reaction respectively and [I] is the inhibitor concentration.[45] The K_i was subsequently calculated by using $K_i = IC_{50}/(1 + [S]/K_m)$[46] and a K_m value determined according to Michaelis–Menten.

MD simulations and free energy calculations

The linear interaction energy method (LIE), described in detail elsewhere,[21,22] can be used to predict absolute binding free energies of receptor–ligand complexes. A general version of the LIE formula can be written as

$$\Delta G_{bind} = \alpha(\langle V_{l-s}^{vdW}\rangle_p - \langle V_{l-s}^{vdW}\rangle_w) + \beta(\langle V_{l-s}^{el}\rangle_p - \langle V_{l-s}^{el}\rangle_w) + \gamma \quad (1)$$

where nonbonded ligand-surrounding (l–s) polar (electrostatic) and nonpolar (van der Waals) components of the interaction energies are denoted by the superscripts el and vdW, respectively, while the subscripts p and w describe the ligand in complex with the solvated protein and free in water, respectively. Thus, in each binding calculation an MD simulation of the complex is performed as well as one of the free ligand in water. The averages are calculated as time averages from molecular dynamics trajectories. The previously determined values of α, β and γ from ref 23 are used also here. In this model the β parameter takes into account the deviation from linear response and depends on the chemical nature of the ligand. For all ligands considered in this work β = 0.33. The values for α and γ are 0.181 and 0, respectively.[31]

Automated docking with Autodock3[47] of the ligands with a rigid protein did not yield any useful results, which may be due to steric reasons if the binding pocket in the X-ray structure is partially collapsed. The X-ray structure 1SME[3] of plasmepsin II in complex with pepstatin A was instead used as a starting point for all simulations of the protein–ligand complexes and five internal water molecules were kept (804, 820, 826, 851 and 892 as numbered in the PDB entry). In order to get reliable results from binding free energy calculations it is necessary to use a reasonably well-docked starting conformation of the protein–ligand complex. Since the hydrogen bonding donors and acceptors of the present inhibitor backbone are identical to those of the corresponding pepstatin A P1′–P3′ moieties, it is likely that the binding conformation of the backbone of **3a–h** is similar to that of pepstatin A at these positions. In addition, the P2, P1, P1′ and P2′ side chains are likely to reside in S2, S1, S1′ and S2′, respectively. Therefore the natural docking approach was to superimpose the entire ligand backbone onto that of pepstatin A in the X-ray structure and fit the side chains into the corresponding subsites. Here, nonpolar interaction energies rose significantly due to the repulsive $1/r^{12}$ nonbonded interaction potential term but decreased in the following MD simulations, where the complex was allowed to relax. Since superposition of the ligand backbone, the P2′ side chain and the P′ hydroxyl on the corresponding groups of pepstatin A yield the *SRRRRS* isomer, all other isomers were generated by subjecting the *SRRRRS* isomer to the appropriate potential function and thus changing its stereochemistry. During this process protein and ligand were kept under restraints in order to prevent

sudden structural changes due to high energies. All MD simulations were done using the program Q[48] and the force field parameters of GROMOS87.[49] AM1 calculations were performed to obtain partial charges for the P1/P1′ side chains of the inhibitors while the backbone and hydroxyl groups were assigned standard force field charges. Asp34, which is involved in a hydrogen bond with Asp214 in the X-ray structure, was protonated at the oxygen interacting with Asp214. (The groups surrounding the aspartic acids yield virtually equivalent environments for the two residues and Asp34 was chosen as the one to be protonated.) The Asp303 side chain was also protonated because of its hydrogen bond with the backbone oxygen of Ser215, while Asp214 was assigned a negative charge. A simulation sphere for the molecular dynamics was defined with 18 Å radius and a 1.5 Å restrained (50 kcal/mol Å^2) outer shell and the center of simulation was chosen at the P1′ hydroxyl oxygen of the ligand. The system within the simulation sphere interacted with the system across the boundary only by bonds and angles. The simulation sphere was solvated in an SPC water grid with radius 18 Å, using a minimum solute–solvent distance of 2.4 Å. Long-range electrostatics of the system, except interactions involving the ligand, were treated with the LRF method[50] and the water surface was subjected to SCAAS restraints[51] in order to mimic bulk water at the sphere boundary. MD simulations of the free ligands in water were set up using the same simulation sphere parameters and a central atom of the ligand was restrained to the center of the sphere by 100 kcal/mol Å^2 to keep the ligand from drifting towards the sphere boundary. All MD simulations used a 1 fs timestep and energies were recorded every 10 fs. Each solvated protein–ligand system was gradually heated from 1 to 300 K. During the heating process restraints were imposed on the protein–ligand amide H-bonds and on those between the hydroxyl groups and the catalytic aspartates. Subsequently, the system was equilibrated after which energies were recorded every 50 fs during 300–600 ps. In addition, simulations of the ligands in aqueous solution were performed under the same conditions as the protein simulation. Here, a protocol was used where the water molecules were first equilibrated around the restrained ligand while changing stereochemistry due to the applied potential function to ensure an extended starting conformation, and the whole system was then relaxed and equilibrated. Data collection was performed for 1.2–2.35 ns until stable ligand-surrounding interactions energy averages were obtained. Since the *SRRRRS* and *RSSSSR* compounds are mirror images of each other, only one of the stereoisomers needed to be simulated in water. This is also true for the *RRRRRR* and *SSSSSS* compounds, so in total two simulations were performed for the allyl and two for the benzyl ligands.

The scoring function of Eldridge et al.[33] was applied to each complex after minimization of an initial MD-equilibrated structure, as well as for 100 energy-minimized trajectory snapshots. The minimization of the complexes was performed by running MD at 1 K temperature for 4000 steps using the same setup as for the LIE calculations described above. The estimated affinity using this scoring function has the form

$$\Delta G_{\text{bind}}^{\text{score}} = \Delta G_{\text{H-bond}} N_{\text{H-bond}} + \Delta G_{\text{metal}} N_{\text{metal}} + \Delta G_{\text{lipo}} N_{\text{lipo}} + \Delta G_{\text{rot}} N_{\text{rot}} + \Delta G_0$$

where $N_{\text{H-bond}}$ denotes the number of H-bonds, based on distance and angle criteria for the atoms involved, N_{metal} is the number of interactions between the ligand and metal ions in the receptor, N_{lipo}, is the number of interactions between lipophilic groups in the ligand and receptor and N_{rot} is the number of rotatable bonds in the ligand that are considered frozen due to interactions with the receptor. The constants $\Delta G_{\text{H-bond}}$, ΔG_{lipo}, ΔG_{metal}, ΔG_{rot}, and ΔG_0 have been assigned empirically.[34]

Acknowledgements

We thank the Swedish Foundation for Strategic Research (SSF), the National Graduate School of Scientific Computing (NGSSC) and the Swedish Research Council (VR) for financial support.

References and Notes

1. Bremen, J. *Am. J. Trop. Med. Hyg.* **2001**, *64*, 1.
2. Werbovetz, K. A. *Curr. Med. Chem.* **2000**, *7*, 835.
3. Silva, A. M.; Lee, A. Y.; Gulnik, S. V.; Maier, P.; Collins, J.; Bhat, T. N.; Collins, P. J.; Cachau, R. E.; Luker, K. E.; Gluzman, I. Y.; Francis, S. E.; Oksman, A.; Goldberg, D. E.; Erickson, J. W. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 10034.
4. Moon, R. P.; Tyas, L.; Certa, U.; Rupp, K.; Bur, D.; Jacquet, C.; Matile, H.; Loetscher, H.; Grueninger-Leitch, F.; Kay, J.; Dunn, B. M.; Berry, C.; Ridley, R. G. *Eur. J. Biochem.* **1997**, *244*, 552.
5. Carroll, C. D.; Patel, H.; Johnson, T. O.; Guo, T.; Orlowski, M.; He, Z.-M.; Cavallaro, C. L.; Guo, J.; Oksman, A.; Gluzman, I. Y.; Connelly, J.; Chelsky, D.; Goldberg, D. E.; Dolle, R. E. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 2315.
6. Carroll, C. D.; Johnson, T. O.; Tao, S.; Lauri, G.; Orlowski, M.; Gluzman, I. Y.; Goldberg, D. E.; Dolle, R. E. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 3203.
7. Dolle, R. E.; Guo, J.; O'Brien, L.; Jin, Y.; Piznik, M.; Bowman, K. J.; Li, W.; Egan, W. J.; Cavallaro, C. L.; Roughton, A. L.; Zhao, W.; Reader, J. C.; Orlowski, M.; Jacob-Samuel, B.; DiIanni Carroll, C. *J. Comb. Chem.* **2000**, *2*, 716.
8. Haque, T. S.; Skillman, A. G.; Lee, C. E.; Habashita, H.; Gluzman, I. Y.; Ewing, T. J. A.; Goldberg, D. E.; Kuntz, I. D.; Ellman, J. A. *J. Med. Chem.* **1999**, *42*, 1428.
9. Jiang, S.; Prigge, S. T.; Wei, L.; Gao, Y.-E.; Hudson, T. H.; Gerena, L.; Dame, J. B.; Kyle, D. E. *Antimicrob. Agents Chemother.* **2001**, *45*, 2577.
10. Nezami, A.; Luque, I.; Kimura, T.; Kiso, Y.; Freire, E. *Biochemistry* **2002**, *41*, 2273.
11. Nöteberg, D.; Hamelink, E.; Hulten, J.; Wahlgren, M.; Vrang, L.; Samuelsson, B.; Hallberg, A. *J. Med. Chem.* **2003**, *46*, 734.
12. Dahlgren, A.; Kvarnstrom, I.; Vrang, L.; Hamelink, E.; Hallberg, A.; Rosenquist, A.; Samuelsson, B. *Bioorg. Med. Chem.* **2003**, *11*, 827.
13. Oscarsson, K.; Oscarson, S.; Vrang, L.; Hamelink, E.;

Hallberg, A.; Samuelsson, B. *Bioorg. Med. Chem.* **2003**, *11*, 1235.
14. Asojo, O. A.; Afonina, E.; Gulnik, S. V.; Yu, B.; Erickson, J. W.; Randad, R.; Medjahed, D.; Silva, A. M. *Acta Cryst.* **2002**, *D58*, 2001.
15. Alterman, M.; Andersson, H. O.; Garg, N.; Ahlsen, G.; Lovgren, S.; Classon, B.; Danielson, U. H.; Kvarnstrom, I.; Vrang, L.; Unge, T.; Samuelsson, B.; Hallberg, A. *J. Med. Chem.* **1999**, *42*, 3835.
16. Halperin, I.; Ma, B.; Wolfson, H.; Nussinov, R. *Proteins* **2002**, *47*, 409.
17. Schaffer, L.; Verkhivker, G. M. *Proteins* **1998**, *33*, 295.
18. Wang, J.; Kollman, P. A.; Kuntz, I. D. *Proteins* **1999**, *36*, 1.
19. Zabell, A. P.; Post, C. B. *Proteins* **2002**, *46*, 295.
20. Alterman, M.; Bjoersne, M.; Muehlman, A.; Classon, B.; Kvarnstroem, I.; Danielson, H.; Markgren, P.-O.; Nillroth, U.; Unge, T.; Hallberg, A.; Samuelsson, B. *J. Med. Chem.* **1998**, *41*, 3782.
21. Åqvist, J.; Medina, C.; Samuelsson, J. E. *Protein Eng.* **1994**, *7*, 385.
22. Åqvist, J.; Luzhkov, V. B.; Brandsdal, B. O. *Acc. Chem. Res.* **2002**, *35*, 358.
23. Hansson, T.; Åqvist, J. *Protein Eng.* **1995**, *8*, 1137.
24. Marelius, J.; Graffner-Nordberg, M.; Hansson, T.; Hallberg, A.; Åqvist, J. *J. Comput. Aided Mol. Des.* **1998**, *12*, 119.
25. Graffner-Nordberg, M.; Marelius, J.; Ohlsson, S.; Persson, A.; Swedberg, G.; Andersson, P.; Andersson, S. E.; Åqvist, J.; Hallberg, A. *J. Med. Chem.* **2000**, *43*, 3852.
26. Graffner-Nordberg, M.; Kolmodin, K.; Åqvist, J.; Queener, S. F.; Hallberg, A. *J. Med. Chem.* **2001**, *44*, 2391.
27. Westling, J.; Cipullo, P.; Hung, S. H.; Saft, H.; Dame, J. B.; Dunn, B. M. *Protein Sci.* **1999**, *8*, 2001.
28. Linstead, R. P.; Owen, L. N.; Webb, R. F. *J. Chem. Soc* **1953**, 1225.
29. Patil, V. J. *Tetrahedron Lett.* **1996**, *37*, 1481.
30. Bursavich Matthew, G.; Rich Daniel, H. *J. Med. Chem.* *45*, 541.
31. Hansson, T.; Marelius, J.; Åqvist, J. *J. Comput. Aided Mol. Des.* **1998**, *12*, 27.
32. Vacca, J. P.; Dorsey, B. D.; Schleif, W. A.; Levin, R. B.; McDaniel, S. L.; Darke, P. L.; Zugay, J.; Quintero, J. C.; Blahy, O. M.; Roth, E.; Sardana, V. V.; Schlabach, A. J.; Graham, P. I.; Condra, J. H.; Gotlib, L.; Holloway, M. K.; Lin, J.; Chen, I.-W.; Vastag, K.; Ostovic, D.; Anderson, P. S.; Emini, E. A.; Huff, J. R. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 4096.
33. Eldridge, M. D.; Murray, C. W.; Auton, T. R.; Paolini, G. V.; Mee, R. P. *J. Comput. Aided Mol. Des* **1997**, *11*, 425.
34. Marelius, J.; Ljungberg, K. B.; Åqvist, J. *Eur. J. Pharm. Sci.* **2001**, *14*, 87.
35. The inhibitory effect of the isomers **3a–h** was determined with purified HIV-1 protease in a standardized assay as described in ref 37. Only the *SRRRRS* isomers **3a** and **3c** showed any inhibitory effect, $K_i = 0.0002$ and 1.4 μM, respectively.
36. Nillroth, U.; Besidsky, Y.; Classon, B.; Chattopadhyaya, J.; Ugi, I.; Danielson, U. H. *Drug Des. Discov.* **1995**, *13*, 43.
37. Brandsdal, B. O.; Aqvist, J.; Smalas, A. O. *Protein Sci.* **2001**, *10*, 1584.
38. Luzhkov, V. B.; Åqvist, J. *FEBS Lett.* **2001**, *495*, 191.
39. King, N. M.; Melnick, L.; Prabu-Jeyabalan, M.; Nalivaika, E. A.; Yang, S.-S.; Gao, Y.; Nie, X.; Zepp, C.; Heefner, D. L.; Schiffer, C. A. *Protein Sci.* **2002**, *11*, 418.
40. Asojo, O. A.; Gulnik, S. V.; Afonina, E.; Yu, B.; Ellman, J. A.; Haque, T. S.; Silva, A. M. *J. Mol. Biol.* **2003**, *327*, 173.
41. Kollman, P. *Chem. Rev.* **1993**, *93*, 2395.
42. Srinivasan, J.; Miller, J.; Kollman, P. A.; Case, D. A. *J. Biomol. Struct. Dyn.* **1998**, *16*, 671.
43. Wang, J.; Morin, P.; Wang, W.; Kollman, P. A. *J. Am. Chem. Soc.* **2001**, *123*, 5221.
44. Böhm, H. J. *J. Comput. Aided Mol. Des* **1998**, *12*, 309.
45. Copeland, R. A. (Ed.). *Enzymes: A Practical Introduction to Structure, Mechanism, and Data Analysis*; VCH, New York, N.Y, 1996.
46. Cheng, Y.; Prusoff, W. H. *Biochem. Pharmacol.* **1973**, *22*, 3099.
47. Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. *J. Comput. Chem.* **1998**, *19*, 1639.
48. Marelius, J.; Kolmodin, K.; Feierberg, I.; Åqvist, J. *J. Mol. Graph. Model.* **1998**, *16*, 213.
49. van Gunsteren, W. F.; Brendsen, H. J. C. *Groningen Molecular Simulation (GROMOS) Library Manual*; Biomos: Nijenborgh, Groningen, The Netherlands, 1987.
50. Lee, F. S.; Warshel, A. *J. Chem. Phys.* **1992**, *97*, 3100.
51. King, G.; Warshel, A. *J. Chem. Phys.* **1989**, *91*, 3647.