Azapeptides Structurally Based upon Inhibitory Sites of Cystatins as Potent and Selective Inhibitors of Cysteine Proteases

Ewa Wieczerzak,[†] Piotr Drabik,[†,‡] Leszek Łankiewicz,[†] Stanisław Ołdziej,[†] Zbigniew Grzonka,*[,†] Magnus Abrahamson,[§] Anders Grubb,[§] and Dieter Brömme[||]

Faculty of Chemistry, University of Gdańsk, Gdańsk, Poland, Department of Public Health, University School of Physical Education, Gdańsk, Poland, Department of Clinical Chemistry, University of Lund, University Hospital, Lund, Sweden, and Department of Human Genetics, Mount Sinai School of Medicine, New York

Received February 13, 2002

A series of azapeptides as potential inhibitors of cysteine proteases were synthesized. Their structures, based on the binding center of cystatins, contain an azaglycine residue (Agly) in place of the evolutionarily conserved glycine residue in the N-terminal part of the enzyme binding region of cystatins. Incorporation of Agly should lead to deactivation of the acyl−enzyme complex formed against nucleophilic attack by water molecules in the final step of peptide bond hydrolysis. The majority of synthesized azapeptides shows high inhibitory potency toward the investigated cysteine proteases, papain, cathepsin B, and cathepsin K. One of them, Z-Arg-Leu-Val-Agly-Ile-Val-OMe (compound **17**), which contains in its sequence the amino acid residues from the N-terminal binding segment as well as the hydrophobic residues from the first binding loop of human cystatin C, proved to be a highly potent and selective inhibitor of cathepsin B. It inhibits cathepsin B with a K_i value of 0.088 nM. To investigate the influence of the structure of compound **17** for its inhibitory properties, we determined its conformation by means of NMR studies and theoretical calculations. The Z-Arg-Leu-Val-Agly fragment, covalently linked to Cys29 of cathepsin B, was also developed and modeled, in the catalytic pocket of the enzyme, through a molecular dynamics approach, to analyze ligand−protein interactions in detail. Analysis of the simulation trajectories generated using the AMBER force field provided us with atomic-level understanding of the conformational variability of this inhibitor, which is discussed in the context of other experimental and theoretical data.

Introduction

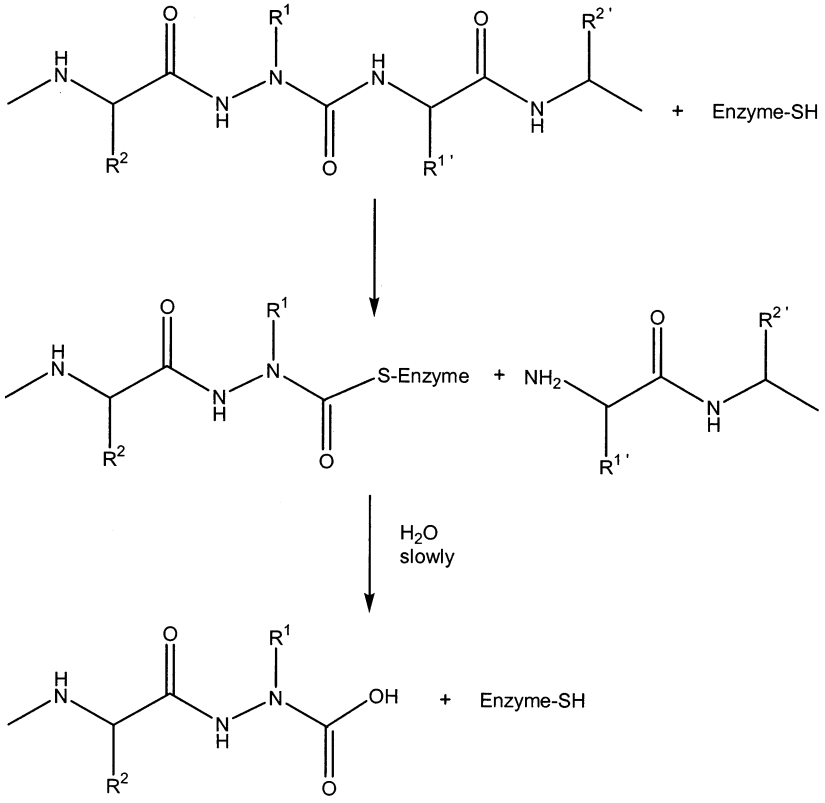
Human papain-like cysteine proteases are mainly localized in lysosomes and play key roles in the intracellular degradation of proteins and peptides (cathepsins B, H, and L).[1,2] They also participate in proteolytic processing of prohormones[3] and proenzymes[4] and appear to be involved in the penetration of normal tissues by macrophages,[5,6] as well as by several types of malignant cells.[7−10] Such proteases are also pivotal in the degradation and remodeling of bone (e.g., cathepsin K)[11−14] and may be instrumental in controlling the MHC class II trafficking in dendritic cells (cathepsin S).[15] Papain-like cysteine proteases are present not only in animals and plants but also in bacteria, fungi, and protozoas,[16] and the proteases of, for example, *Entamoeba histolytica*, *Trypanosoma congolense*, *Leischmania mexicana*, *Trichomonas vaginalis*, and *Plasmodium falciparum* appear to be involved in the replication, migration, and food digestion of these organisms.[16−22]

The activity of such papain-like cysteine proteases (family C1 proteases)[23] is inter alia regulated by their interactions with specific proteinaceous inhibitors, the cystatins,[24−26] and disturbances of the normal protease−protease inhibitor balance may be of pathophysiological significance in several pathological conditions.[26,27] To be able to influence the protease−protease inhibitor balance, a number of low molecular mass peptidyl derivatives structurally based upon the inhibitory sites of cystatins have recently been synthesized and suggested as possible agents for treatment of various diseases.[28−32] However, the peptidyl derivatives produced so far are easily inactivated by proteolytic degradation of the peptidyl portion and contain highly reactive and potentially toxic thiol reactive groups. Several or all of these problems might be overcome by use of peptidomimetic substances[33] with a structural similarity to the previously synthesized low molecular mass cystatin-based peptidyl derivatives. Azapeptides, peptide analogues in which the α-CH group of one or more amino acid residues in the peptide chain is replaced by a nitrogen atom, have proved to be a class of useful peptidomimetics.[34] Azapeptides fulfill all requirements for inhibition of proteases, and several azapeptide type inhibitors of cysteine proteases have been described.[35−37] Incorporation of the azaglycine residue in the P_1 position of a papain substrate allows the formation of an acyl−enzyme complex and on the other hand deactivates the complex toward nucleophilic attack of a water molecule in the last step of peptide bond hydrolysis (Scheme 1). This is due to decreased electrophilicity of the P_1 carbonyl carbon and the change of geometry of the α-atom from tetrahedral (amino acid) to nearly trigonal (azaamino acid). Therefore, the α-CH/N exchange leads to changes in local conformation of the modified peptide.

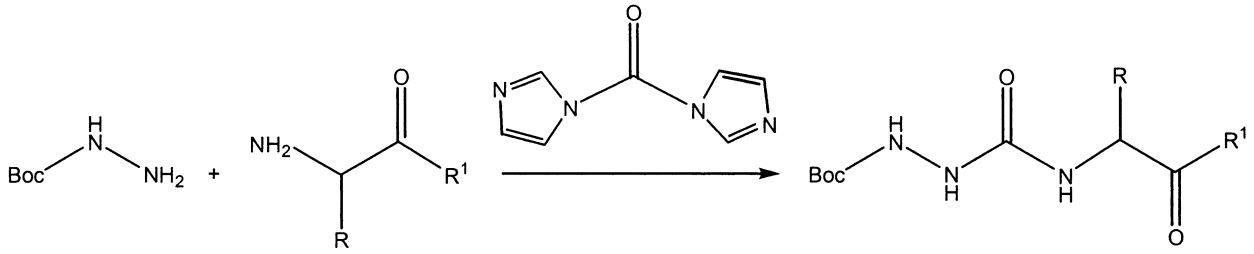
* To whom correspondence should be addressed. Tel: 48-58-345-0369. Fax: 48-58-344-9680. E-mail: grzonka@chemik.chem.univ.gda.pl.
[†] University of Gdańsk.
[‡] University School of Physical Education.
[§] University of Lund.
[||] Mount Sinai School of Medicine.

10.1021/jm020850k CCC: $22.00
Published on Web 08/08/2002

Scheme 1. Postulated Mechanism of Cysteine Proteases Inhibition by Azapeptides



Scheme 2. General Method for Azapeptide Bond Formation



The published azapeptide type inhibitors that deliver the azaamino acid to the S_1 subsite of cysteine proteases are either esters or amides of Ac-Phe-azaamino acids [azaglycine (Agly), azaalanine (Aala), or azaphenylalanine (Aphe)].[35−37] The apparent second-order rate constants for inactivation of papain by azapeptide esters and amides depend greatly on the electronegativity and hydrophobicity properties of the leaving group.[37] Azaglycine derivatives showed the highest affinity for papain and cathepsin B. The reason for this fact is steric in origin, meaning that P_1 substituents larger than hydrogen greatly hinder productive binding to these enzymes.

The objective of the present study was to synthesize azapeptides structurally based upon the inhibitory sites of cystatins and to investigate the efficiency and specificity of such azapeptides as inhibitors of cysteine proteases. The work shows that some of the peptides designed in this way are efficient protease inhibitors and also display specificity for certain cysteine proteases.

Synthesis Chemistry

Azapeptides were synthesized in solution by conventional methods except for the azapeptide bond formation (Scheme 2). The azaglycine residue was incorporated in the growing peptide chain by the coupling of Boc-$NHNH_2$ and an appropriate amino acid/peptide ester/amide with carbonyldiimidazole (CDI) as an azaglycine carbonyl donor. This approach was adopted from the previously described method of preparation of amino acid aldehyde semicarbazone resins.[38] We found that the application of CDI in the azaglycine residue incorporation gave higher yield and purity of the products when compared with the most frequently used di-4-nitrophenyl- or bis-(2,4-dinitrophenyl)carbonates methods.[39,40]

Results and Discussion

In our work, we synthesized some potential aza inhibitors of cysteine proteases that contain the azaglycine residue (Agly) in place of the residue expected to be located in the P_1 position when a substrate or a protein inhibitor interacts with the enzyme. The sequence of the first group of synthesized azapeptides is modeled after the structure of papain substrates based on the Ac-L-Phe-Gly-X motif, where X is an alcohol, amine, or peptide leaving group. Baggio[36] found in his work that esters of N-acetyl-L-phenylalanyl-azaglycyl-L-valine (Ac-Phe-Agly-Val-OMe and Ac-Phe-Agly-Val-

Table 1. Inhibitory Potencies vs Papain of Azapeptides Whose Structure Is Based on the Ac-Phe-**Agly**-Val-X Motif[a]

compd	structure	K_i (nM)
1	Ac-Phe-**Agly**-Val-OH	no inhibition
2	Ac-Phe-**Agly**-Val-OMe	61.0 ± 29.0
3	Ac-Phe-**Agly**-Val-OEt	20.0 ± 10.0
4	Ac-Phe-**Agly**-Val-OBu	59.0 ± 5.9
5	Ac-Phe-**Agly**-Val-OBzl	21.0 ± 2.1
6	Ac-Phe-**Agly**-Val-NHMe	32.0 ± 9.0
7	Ac-Phe-**Agly**-Val-NHBu	28.0 ± 6.2
8	Ac-Phe-**Agly**-Val-NHBzl	30.0 ± 5.5

[a] The K_i values were calculated as detailed in the Materials and Methods section, from results of three continuous rate assays, and presented as mean value ± SD.

OBzl) are good inhibitors of papain. Both compounds were resynthesized by us (Table 1, compounds **2** and **5**) and used as references in our study. First, we decided to synthesize the same type of azapeptides in order to check whether the C-terminal group of Ac-Phe-Agly-Val-X (where X is −OR or −NHR) influenced their potency as papain inhibitors. Because azapeptides are assumed to inactivate cysteine proteases reversibly,[35] we determined the equilibrium constants for dissociation of azainhibitor−enzyme complexes, K_i.[41] They are presented in Table 1. With the exception of compound **1**, which contains the free carboxyl group, the remaining esters or amides of Ac-Phe-Agly-Val are characterized by very similar K_i values for papain. Thus, it seems obvious that the substituent on the C terminus of azapeptide has little effect on its inhibitory potency.

Our studies on low molecular weight inhibitors, the structures of which were based on the N-terminal binding sequence of cystatins and which contained thiol reactive groups (diazomethyl ketones, oxiranes, compounds with activated olefinic double bonds, etc.), have shown that their potency toward cysteine proteases depends strongly upon the structure of their peptidyl fragment.[29−32] The highest inhibitory potencies were found for compounds based on the N-terminal binding structure of human cystatin C. For example, Z-Arg-Leu-Val-Gly diazomethyl ketone[30] is the most potent inhibitor for papain and cathepsin B among peptidyl-diazomethyl ketones, with second-order rate constants (k_{+2}) of the same order of magnitude as those determined for natural inhibitor E-64 [(2*S*,3*S*)-*trans*-epoxysuccinyl-L-leucyl-agmatine], which is used as a standard in the inhibitory bioassays of cysteine proteases.[42] This prompted us to design azapeptides resembling the N-terminal binding center of cystatins (occupying S_n subsites of cysteine proteases) as cysteine protease inhibitors. In addition, we decided to introduce also some hydrophobic amino acid residues from the first (Val, Ile, and/or Ala) and second (Trp) binding loops of cystatins in S_n' positions (Table 2). The evolutionarily conserved P_1 glycine residue in the N-terminal region of human cystatin C (Gly11), which is known to have a crucial role in high-affinity binding and efficient inhibition of cysteine proteases by allowing the N-terminal segment of cystatin C to adopt a conformation suitable for interaction with the substrate binding pockets of the enzymes,[43] has been replaced by its aza analogue.

Compound **9**, Boc-Agly-Val-OBzl, which contains the azaglycine residue in the P_1 position and a hydrophobic valine residue in the P_1' position, is devoid of inhibitory activity toward papain. This result confirms the fact that amino acid residues in positions P_2 and P_3 contribute considerably to the interactions between inhibitors and cysteine proteases. Boc-Leu-Val-Agly-OEt (**10**), which contains a part of the N-terminal binding sequence of human cystatin C (-Leu-Val-Gly-), is found to be an efficient reversible inhibitor of papain and cathepsin K, with K_i values in the nanomolar range. Prolongation of the -Leu-Val-Agly- fragment with a valine residue in the P_1' position (**11**) improves the inhibitory potencies for papain and cathepsin K. This compound also inhibits cathepsin B. Removal of the carbobenzoxy group from this azapeptide gives compound **12**, which is devoid of inhibitory activity. Lack of inhibitory potencies of compounds **13** and **14**, in which the valine residue in position P_2 was replaced by phenylalanine, is rather unexpected because Phe in P_2 is highly favored by papain. It seems to us that the reason for this fact is steric in origin. Incorporation of the azaamino acid residue in position P_1 can change the spatial orientation of the larger side chains of amino acid residues in its neighborhood, leading to the decrease of the affinity of these azapeptides to papain.

The most potent inhibitors of papain, cathepsin B, and cathepsin K proved to be Z-Arg-Leu-Val-Agly-Val-Ala-NH_2, Z-Arg-Leu-Val-Agly-Ala-Gly-NH_2, and Z-Arg-Leu-Val-Agly-Ile-Val-OMe (compounds **15**−**17**, respectively), whose structures include the N-terminal sequence, as well as the amino acids from the first binding loop of

Table 2. Inhibitory Potencies vs Papain, Human Cathepsin B, and Recombinant Human Cathepsin K of Azapeptides with Structures Based on the Binding Region of Cystatin C[a]

		K_i (nM)		
compd	structure	papain	cathepsin B	cathepsin K
hCC	P_5 P_4 P_3 P_2 P_1 P_1' P_2' P_3' Pro Arg Leu Val Gly Gly Pro Met			
9	Boc-**Agly**-Val-OBzl	no inhibition	nd[b]	nd
10	Boc-Leu-Val-**Agly**-OEt	24.0 ± 5.8	no inhibition	61.0 ± 8.0
11	Z-Leu-Val-**Agly**-Val-OBzl	6.7 ± 2.7	40.0 ± 6.1	34.0 ± 9.5
12	H-Leu-Val-**Agly**-Val-OBzl	no inhibition	nd	nd
13	Z-Leu-Phe-**Agly**-Val-OBzl	no inhibition	nd	nd
14	Z-Arg-Leu-Phe-**Agly**-Val-OBzl	no inhibition	nd	nd
15	Z-Arg-Leu-Val-**Agly**-Val-Ala-NH_2	0.56 ± 0.12	0.12 ± 0.02	2.13 ± 1.02
16	Z-Arg-Leu-Val-**Agly**-Ala-Gly-NH_2	0.78 ± 0.40	7.29 ± 3.39	6.76 ± 1.86
17	Z-Arg-Leu-Val-**Agly**-Ile-Val-OMe	0.77 ± 0.28	0.088 ± 0.025	21.0 ± 1.3
18	Z-Arg-Leu-Val-**Agly**-Ile-Val-Trp-OMe	17.0 ± 0.53	13.0 ± 8.5	72.0[c]
19	Z-Arg-Leu-Val-**Agly**-Trp-Val-Ala-NH_2	100.0[c]	43.0[c]	100.0[c]

[a] The K_i values were calculated as detailed in the Materials and Methods section, from results of 3−5 continuous rate assays if not otherwise stated, and presented as mean value ± SD. [b] nd, not determined. [c] Based on single assay.

Table 3. Inhibitory Potencies vs Papain, Human Cathepsin B, and Human Recombinant Cathepsin K of Azapeptides Structurally Based on the Binding Region of Cystatin C, with Various Amino Acid Residues in the P_2 Position[a]

compd	structure	K_i (nM) papain	cathepsin B	cathepsin K
hCC	P_5 P_4 P_3 P_2 P_1 P_1' P_2' P_3' Pro Arg Leu Val Gly Gly Pro Met			
20	Z-Ala-Leu-Ile-**Agly**-Ile-Val-OMe	0.95 ± 0.20	0.30 ± 0.15	2.12 ± 1.79
21	Z-Ala-Leu-Phe-**Agly**-Ile-Val-OMe	0.12 ± 0.05	0.13 ± 0.08	0.18 ± 0.07
22	Z-Ala-Leu-Tyr(Me)-**Agly**-Ile-Val-OMe	0.28 ± 0.07	1.41 ± 0.57	8.42 ± 0.74
23	Z-Ala-Leu-His-**Agly**- Ile-Val-OMe	0.58 ± 0.50	0.03 ± 0.01	0.86 ± 0.21
24	Z-Ala-Leu-Nal-**Agly**- Ile-Val-OMe	0.10 ± 0.01	1.58 ± 1.30	5.75 ± 1.18
25	Z-Ala-Leu-Leu-**Agly**-Ile-Val-NHBzl	1.35 ± 0.20	0.25 ± 0.08	2.45 ± 0.81

[a] The K_i values were calculated as detailed in the Materials and Methods section, from results of 3−5 continuous rate assays, and are presented as mean value ± SD.

cystatin C. Incorporation into compound **17** of a tryptophan residue from the second binding loop of cystatins, either in P_3' (compound **18**) or in P_1' (**19**) positions, decreases their activities as cysteine protease inhibitors. In the latter case, a tryptophan residue is probably too bulky to fit in the S_1' subsite of papain and is not accepted by cathepsin K due to its hydrophobic character. Cathepsin B prefers large hydrophobic residues in position P_1' of a substrate/inhibitor,[44] which is probably responsible for the higher affinity of compound **19** to this enzyme.

The knowledge of substrate binding sites is based on the kinetics and crystal structures of substrate-mimicking inhibitors bound to the active sites of the enzymes. These data demonstrate that substrate P_1 and P_2' residue side chains face the solvent, whereas the P_2 and P_1' side chains contact the enzyme surface. The S_2 and S_1' substrate binding sites are therefore the ones responsible for the diversity and selectivity of the substrate and inhibitor binding.[45,46] The S_2 binding site is probably the best defined substrate binding site. In the case of the majority of cysteine proteases, it has a hydrophobic character and can thus readily adopt the side chain of the valine residue from the P_2 positions of compounds **15**−**17** on the assumption that these azapeptides bind along the active site cleft in an extended conformation. It seems to us that the differences in the selectivity of compounds **15**−**17** toward the investigated cysteine proteases can, to some extent, result from docking of the appropriate inhibitors' amino acid residues in the S_n' subsites of those enzymes before the formation of the enzyme−inhibitor complex according to the zipper mechanism.[47] Cathepsin B displays a preference for large hydrophobic residues in the P_1' position of substrate/inhibitor. The S_1' pocket of papain is slightly smaller and can accept both hydrophilic and smaller hydrophobic amino acid residues.[44] The crystal structure of procathepsin K[48] reveals that in the case of this enzyme, the S_1' binding site is occupied by Thr76 from the proregion and it does not have the hydrophobic character. Compound **17** proved to be approximately 10 times more potent for cathepsin B than for papain and more than 200 times more potent for cathepsin B than for cathepsin K. The reason for this is probably connected with the higher affinity of the hydrophobic isoleucine residue from the P_1' position of this azapeptide to the amino acid residues forming the S_1' binding site of cathepsin B as compared to those building the corresponding subpockets in papain and cathepsin K molecules. Similar conclusions can be drawn from the K_i values obtained for compound **15**. The valine residue in the P_1' position of this aza inhibitor will probably fit well in the subpockets S_1' of papain and cathepsin B; therefore, we can observe smaller differences in selectivity in relation to those two enzymes. In additon, the length of extensions on the P′ side can also improve the inhibitory potency of this azapeptide to cathepsin K. The K_i value for Z-Arg-Leu-Val-Agly-Val-Ala-NH_2 is about 10 times lower than that obtained for Z-Arg-Leu-Val-Agly-Ile-Val-OMe. On the contrary, compound **16** displays much weaker inhibitory activity to cathepsin B, which can be explained by the size of the P_1' alanine residue—too small to interact with the bulky hydrophobic S_1' subsite of the investigated enzyme. To investigate the influence of amino acid residue in azapeptides' P_2 position on their inhibitory properties, we synthesized some potential aza inhibitors with structures based on the sequence of compound **17**, which proved to be the most potent and selective inhibitor of the investigated cysteine proteases (Table 3). To facilitate the synthesis of these compounds, they contain an alanine instead of an arginine residue. Unfortunately, for some synthesized azapeptides, we were not able to measure K_i values because of their insolubility in the solutions used in the enzyme kinetics tests. Among them are Z-Ala-Leu-Leu-Agly-Ile-Val-OMe, Z-Ala-Leu-Gln-Agly-Ile-Val-OMe, and Z-Ala-Leu-Cha-Agly-Ile-Val-OMe. Comparison of the K_i values obtained for the remaining compounds draws attention to the fact that all of them are potent inhibitors of papain, cathepsin B, and only somewhat weaker cathepsin K. The S_2 binding site of papain readily accommodates different hydrophobic amino acid residues, which is reflected in the similar K_i values of complexes of various aza inhibitors with this enzyme. An analogous situation can be observed for cathepsin B, except for compounds **22** and **24**, whose O-methylated tyrosine and 2-naphthylalanine residues probably are too bulky to interact with amino acids forming the S_2 pockets of cathepsins B and K. Compound **23**, Z-Ala-Leu-His-Agly-Ile-Val-OMe, proved to be a very potent inhibitor of cathepsin B. This fact can be explained by the unique feature of the enzyme connected with the presence in its S_2 subsite of a glutamic acid residue in position 245, which is properly located to interact with the positively charged side chain of the histidine residue in position P_2 of a substrate/inhibitor.[49]

The unique conformational properties of the peptide backbone that are provided by the incorporation of an aza amino acid residue can affect biological action of the aza analogues of natural peptides. We have therefore tried to determine the structure of the most potent and

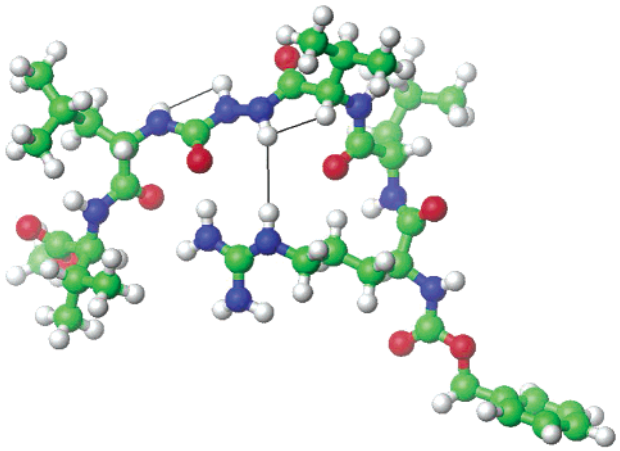


Figure 1. Representative conformation of the Z-Arg-Leu-Val-Agly-Ile-Val-OMe (compound **17**). Black dashed lines mark the hydrogen bonds important from the structural point of view.

cathepsin B selective aza inhibitor, Z-Arg-Leu-Val-Agly-Ile-Val-OMe, by means of NMR studies and theoretical calculations. The representative structure of this compound is presented in Figure 1. The strong nuclear Overhauser effects (NOE) observed between protons from the guanidine moiety, the proton from the azaglycine residue, and the α-proton from the valine residue produce the rigid frame of azapeptide, which is stabilized by the hydrogen bond between the guanidinium group of the arginine and the carbonyl group of the azaglycine residue. The other parts of the molecule are not well-restrained and appeared to be very flexible. The conformation adopted by the investigated azapeptide may be responsible for its selectivity for cathepsin B by allowing some interactions between inhibitor's fragment bound in the catalytic cleft of the enzyme and the appropriate amino acid residues.

To analyze in detail interactions between an aza inhibitor and cathepsin B, the Z-Arg-Leu-Val-Agly fragment, covalently linked to Cys29 in the catalytic pocket of the enzyme, was developed and modeled by molecular dynamics procedures (MD).

The root-mean-square (RMS) deviations along MD trajectories are given in Figure 2. The RMS difference was calculated for structures along the trajectory relative to the crystal structure (Protein Data Bank entry: 1the) for all Cα atoms. In general, the protein−inhibitor complex structure seems to be stabilized.

Time-averaged residue-based deviations as a function of residue number for all MD runs are shown in Figure 3. Changes up to 4.3 Å were observed for some residues (108−125), but the overall Cα mobilities were between 1 and 2.5 Å. The average structure displayed significant changes in some flexible loops of the enzyme (e.g., the occluding loop). It should be stressed that the protein structure reproduced the mobility pattern over the whole molecule, represented experimentally by the atomic temperature factors in the crystal structure of the Z-Arg-Ser(O-Bzl)-cathepsin B complex very well,[50] even if both quantities are proportional to the standard deviation and variance, respectively. This result has validated the use of the AMBER 5.0 force field as suitable for scanning the conformational space of the ligand nested in the catalytic cleft of cathepsin B.

Figure 4 summarizes detailed atomic-level mobilities of the inhibitor's backbone. It can be seen that the overall mobility is quite small, indicating that the inhibitor is stuck in the catalytic pocket. The only exception is the Z residue, which fluctuates freely on the surface of the protein, as can be illustrated by the average mobility of its Cζ atom of 2.8 Å. Moreover, the average distance between this Cζ and Cζ of Tyr75 (whose ring is proposed to interact with the phenyl ring of the Z residue) is 12.3 Å with a considerable standard deviation of 3.7 Å. Both backbone and side chain of the arginine residue (shown as the mobilities of Cα and Cζ) are still almost perfect. To analyze in more detail the latter, an evolution of the distance between Cζ (Arg) and Cδ of the protein's Glu245 (whose side chain, when

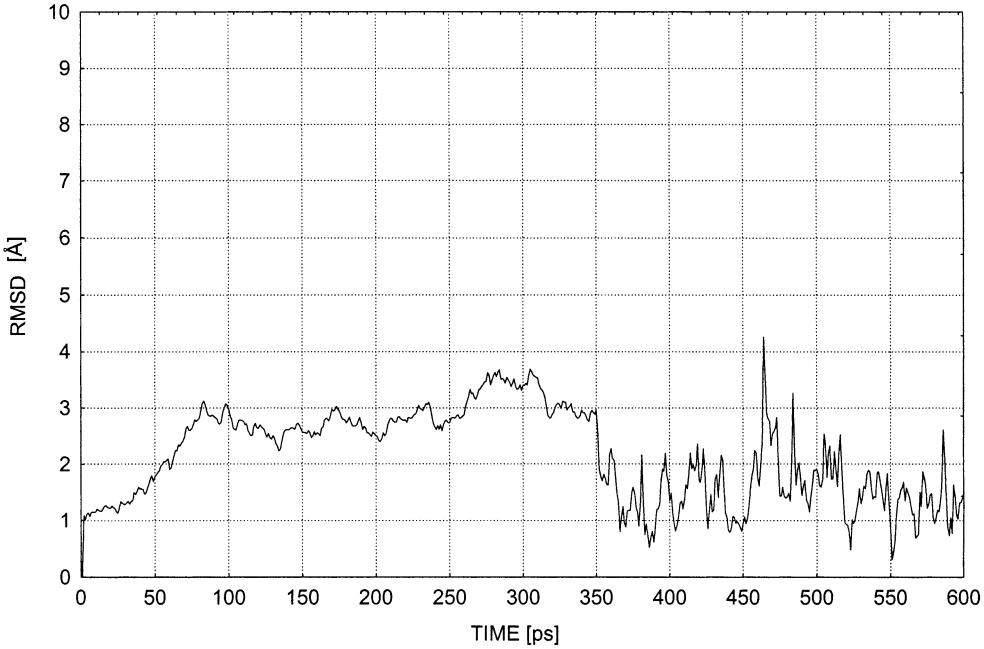


Figure 2. RMS deviations (Å) vs simulation time (ps) of Cα atoms along the MD trajectory. One can see that after ∼80 ps of the MD productive run, the system appears relaxed.

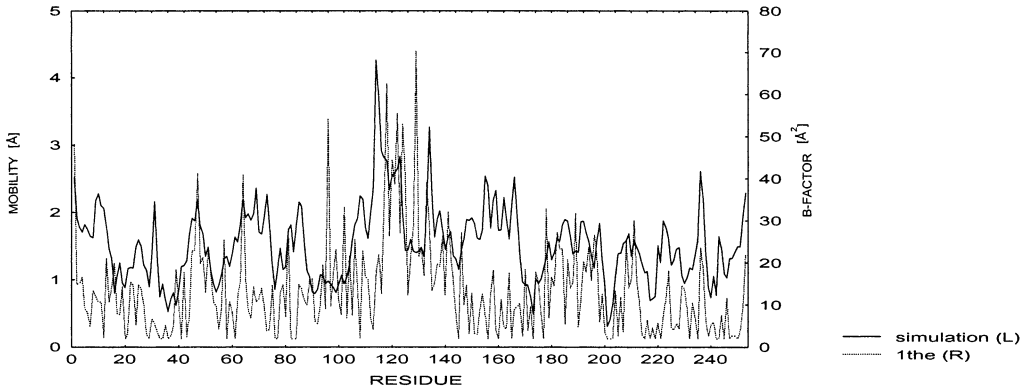


Figure 3. Time-averaged residue-based deviations (mobilities) along the sequence during MD runs. The distribution of the temperature factors along the cathepsin B sequence (as reported in PDB file 1the) is also given. The parallel distribution of both quantities along the sequence is clear. It can be seen that some loops of the cathepsin B (e.g., the occluding loop) undergo high fluctuations during molecular dynamics runs.

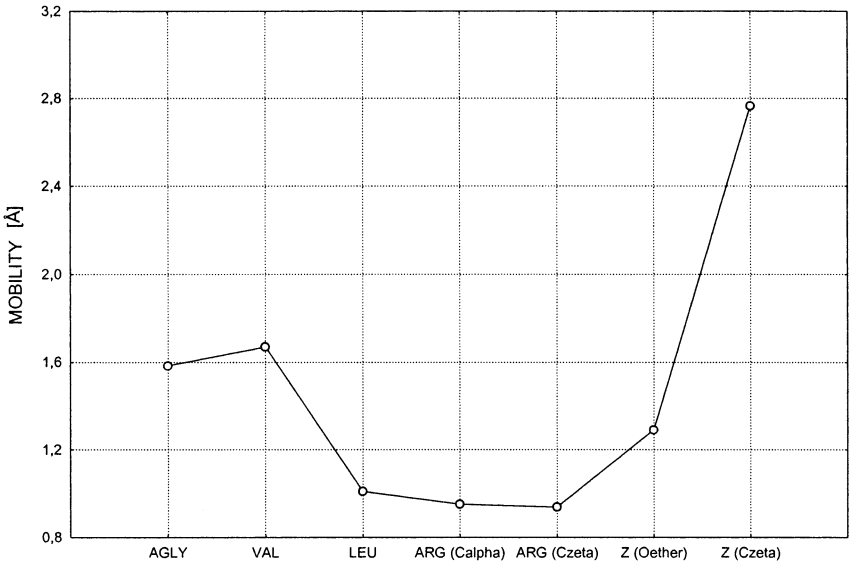


Figure 4. Mobilities of inhibitors' individual residues during the molecular dynamics simulations.

ionized, is able to form a salt bridge with the positively charged arginine side chain[50]) is shown in Figure 5 and the numerical characteristics are given in Table 4. It seems that a stable salt bridge is formed between Arg and Glu245. When put in a different place in inhibitor's sequence, Arg residue is also able to form salt bridges: the guanidinium of P_2 Arg of Z-Arg-Ser(Bzl)-CH_2-rat cathepsin B can reach and form a salt bridge Glu245.[50] Such interaction between Glu245 and P_4 Arg of the Arg-Leu-Val containing highly active substrates[29] and inhibitors[30] is of great importance.

The MD simulations revealed a very surprising structure of the Z-Arg-Leu-Val-Agly cathepsin B complex in comparison with the published experimental data. Crystal structures of various inhibitors with cysteine proteases show that inhibitors bind along the active site cleft in an extended conformation. In the Z-Arg-Leu-Val-Agly cathepsin B complex, the modeled fragment of the inhibitor adopts a bent conformation and there cannot be seen any interactions of the Leu and Val residues of the inhibitor with the amino acids forming the substrate binding sites S_3 and S_2, respectively. That bend conformation of fragment Z-Arg-Leu-Val-Agly may assist the interaction between the P_4 Arg residue and Glu245. The leucine residue is located between the substrate pockets S_1 and S_1', and the valine residue is located between the S_2 and S_1 binding sites of cathepsin B. The structure of the ligand−protein complex is presented in Figure 6. Thus, it can be stated that the results of the simulation show that a salt bridge between the arginine residue of the inhibitor and the Glu245 is likely to be responsible for the inhibitor's low mobility and its overall structure in the catalytic pocket of cathepsin B. This seems to afford a rational cause for the statement that this residue is particularly important in the interaction between the cathepsin B and the investigated inhibitor upon binding.

Conclusions

In this paper, we have disclosed a new series of azapeptide inhibitors of cysteine proteases (papain and cathepsins B and K). They are predicted to reversibly inactivate cysteine proteases. Introduction of an aza-glycine residue in the P_1 position of an enzyme substrate permits the formation of an acyl−enzyme complex, but this complex, for electronic and steric reasons, is deactivated toward nucleophilic attack of a water molecule in the last step of peptide bond hydrolysis. We have found that incorporation into the potential aza inhibitors of the amino acid residues from the sequences of binding regions of the natural cysteine protease inhibitors, the cystatins, greatly affect their inhibitory potency.

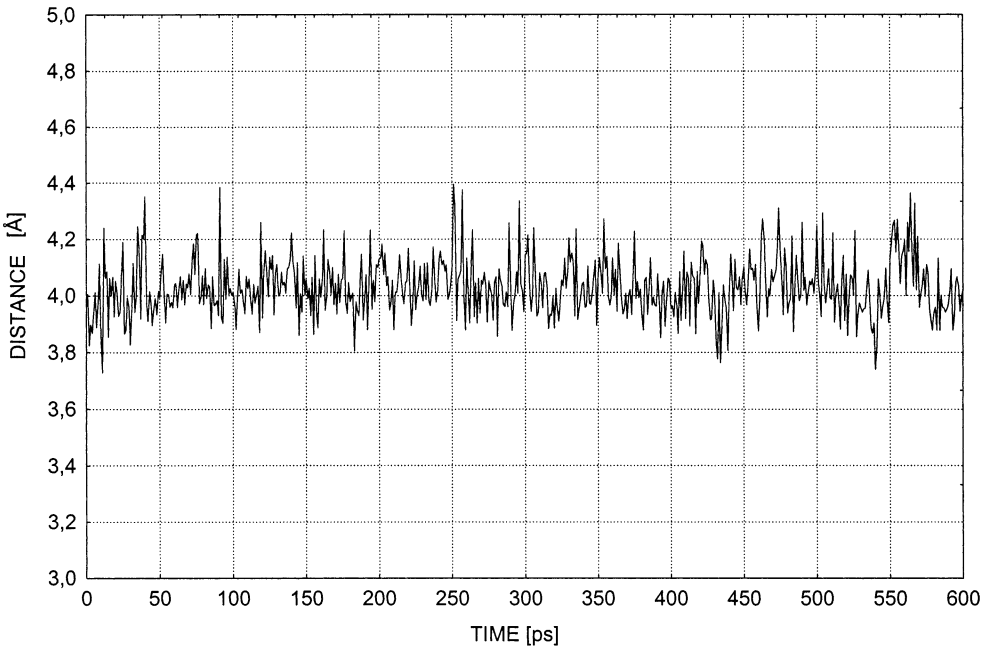


Figure 5. Evolution of the distance between Cζ (Arg in the inhibitor) and Cδ of the protein's Glu245 during the productive MD run.

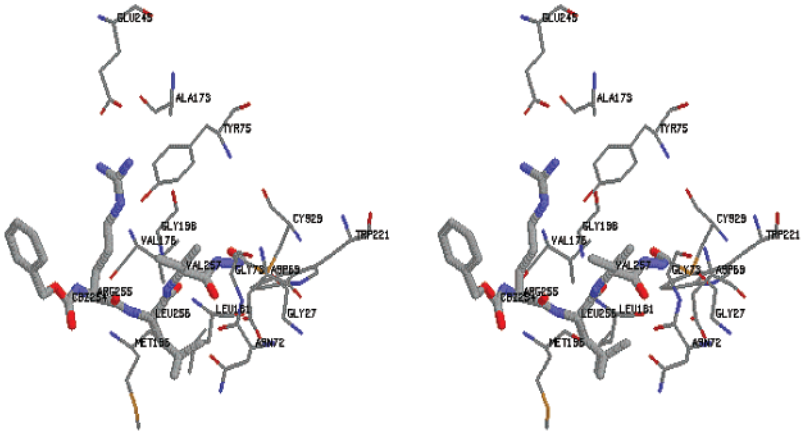


Figure 6. Structure of the Z-Arg-Leu-Val-Agly cathepsin B complex after 200 ps of MD simulations.

Table 4. Descriptive Statistics of the Salt Bridge between Arg (Cζ in the Inhibitor) and the Glu245 (Cδ)

	mean	min	max	SD
distance (Å)	4.029 117	3.728 331	4.394 749	0.102 784

Aza amino acids confer unique conformational properties to the peptide backbone. The most potent and cathepsin B selective aza inhibitor, Z-Arg-Leu-Val-Agly-Ile-Val-OMe (compound **17**), does not adopt an extended conformation, neither in solution nor in complex with cathepsin B, as demonstrated by NMR studies and theoretical calculations. This observation is in contrast to previously published data, which indicate that peptidyl inhibitors bind along the active site cleft of cysteine proteases and display extended conformations. Presented results of the molecular dynamics simulations are in agreement with the experimental data obtained for Z-Arg-Leu-Val-Agly-Ile-Val-OMe and also for some other potent and selective aza inhibitors of cathepsin B, the structure of which is based on the -Arg-Leu-Val-Gly- fragment of cystatin C.

Experimental Section

General. All solvents were of the highest purity. Amino acid derivatives were purchased from Peptides International, Inc. Peptide bonds were formed using DCC (dicyclohexylcarbodiimide) or TBTU [2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium tetrafluoroborate] as the coupling reagents with the addition of HOBt (1-hydroxybenzotriazole). Other procedures were adopted from Bodanszky.[51] Final products were purified by crystallization or reversed-phase high-performance liquid chromatography (HPLC) using a C_8 Kromasil column (25 mm × 250 mm, 7 μm) and a linear gradient of 0−80% acetonitrile in 0.1% aqueous trifluoroacetic acid as a mobile phase. Their purity was assessed by RP-HPLC using a C_8 Kromasil column (4.6 mm × 250 mm, 5 μm) and the linear gradient of the same mobile phase for 60 min with UV detection at 223 nm, by thin-layer chromatography (TLC), by ^{1}H NMR (Varian Mercury 400 MHz) and IR (Bruker spectrometer IFS 66) spectra, and by mass spectrometry (VG Mass Lab). TLC was carried out by using Merck aluminum sheets precoated with silica gel 60 F_{254}.

General Method for Peptide Bond Synthesis Using TBTU/HOBt. *t*-Butoxycarbonyl (Boc)-protected amino acid/peptide (1 mol equiv) was dissolved in dimethylformamide (DMF). The hydrochloride of the ester/amide of the appropriate amino acid/peptide (1 mol equiv) and HOBt (1 mol equiv) were

then added. The mixture was cooled to 0 °C, and triethylamine (3 mol equiv) and TBTU (1.1 mol equiv) were added. The resultant mixture was stirred, and after the TLC showed the completion of the reaction, the solvents were evaporated. The mixture was then diluted with ethyl acetate and washed successively with brine, saturated $NaHCO_3$, brine, 1 M $KHSO_4$, and brine and dried ($MgSO_4$). After the solvent was removed, the product was crystallized or purified by HPLC.

General Method for Peptide Bond Synthesis Using DCC/HOBt. Boc-protected amino acid/peptide (1 mol equiv) was dissolved in DMF. The ester/amide of the appropriate amino acid/peptide (1 mol equiv) and HOBt (1 mol equiv) were then added. The mixture was cooled to 0 °C, and DCC (1.05 mol equiv) was then added over a period of 30 min. The mixture was stirred for 2 h at room temperature, and the precipitated dicyclohexylurea was then filtered off. After the solvents had been evaporated, the mixture was diluted with ethyl acetate and washed successively with brine, saturated $NaHCO_3$, brine, 1 M $KHSO_4$, and brine and dried ($MgSO_4$). After the solvent was removed, the product was crystallized or purified by HPLC.

General Method for Azapeptide Bond Formation. The solution of CDI (1.1 mol equiv) in DMF was treated at −10 °C with a solution of *t*-butoxycarbonyl-hydrazide in DMF. After 15 min, the hydrochloride of the ester of the appropriate amino acid or peptide (1 mol equiv) and triethylamine (1.1 mol equiv) were added. The mixture was stirred overnight, evaporated, and redissolved in ethyl acetate. The solution was washed with brine, saturated $NaHCO_3$, brine, 1 M $KHSO_4$, and brine and dried ($MgSO_4$). After the solvent was removed, the product was crystallized or purified by HPLC and used in the synthesis of the elongated azapeptides.

N-Acetyl-L-phenylalanyl-azaglycyl-L-valine Benzyl Ester (5). RP-HPLC of the crude product gave pure **5** in 51% yield (t_R = 40.1 min). ^{1}H NMR (DMSO-d_6): δ 9.83 (s, 1H), 8.17−8.13 (m, 1H), 7.95 (s, 1H), 7.39−7.31 (m, 5H), 7.30−7.25 (m, 5H), 6.49−6.47 (m, 1H), 5.14 (s, 1H), 4.53−4.47 (m, 1H), 4.15−4.10 (m, 1H), 2.87 (dt, 2H), 2.07−2.01 (m, 1H), 1.75 (s, 3H), 0.88−0.81 (m, 6H). MS (FAB): 454 $(M + H)^+$.

Benzyloxycarbonyl-L-arginyl-L-leucyl-L-valyl-azaglycyl-L-isoleucyl-L-valine Methyl Ester (17). The crude product was purified by means of RP-HPLC (t_R = 41.3 min). Yield: 28%. ^{1}H NMR (DMSO-d_6): δ 9.69 (s, 1H), 8.19 (d, 1H), 7.96 (s, 1H), 7.92 (d, 1H), 7.75 (d, 1H), 7.44−7.39 (m, 2H), 7.37−7.31 (m, 5H), 6.11 (d, 1H), 5.03 (s, 2H), 4.40−4.34 (m, 1H), 4.22−4.19 (m, 1H), 4.14−4.10 (m, 2H), 4.03−4.02 (m, 1H), 3.61 (s, 3H), 3.38−3.36 (m, 2H), 3.09−3.08 (m, 1H), 2.06−2.01 (m, 1H), 1.95−1.90 (m, 1H), 1.73−1.70 (m, 2H), 1.63−1.60 (m, 1H), 1.52−1.39 (m, 2H), 1.29−0.98 (m, 4H), 0.91−0.79 (m, 24H). MS (FAB): 805 $(M + H)^+$.

Benzyloxycarbonyl-L-alanyl-L-leucyl-L-phenylalanyl-azaglycyl-L-isoleucyl-L-valine Methyl Ester (21). The final product was obtained by means of RP-HPLC (t_R = 52.6 min) in a yield of 36%. ^{1}H NMR (DMSO-d_6): δ 9.76 (s, 1H), 8.19 (d, 1H), 7.97 (s, 1H), 7.85 (d, 1H), 7.79 (d, 1H), 7.40−7.38 (m, 1H), 7.34−7.26 (m, 5H), 7.25−7.14 (m, 5H), 6.19 (d, 1H), 5.01 (s, 2H), 4.54−4.50 (m, 1H), 4.28−4.20 (m, 2H), 4.14−4.11 (m, 1H), 4.05−4.02 (m, 1H), 3.61 (s, 3H), 3.04−2.79 (m, 2H), 2.08−1.99 (m, 1H), 1.67−1.63 (m, 1H), 1.54−1.48 (m, 1H), 1.38−1.34 (m, 2H), 1.15 (d, 3H), 1.07−0.98 (m, 2H), 0.91−0.79 (m, 18H). MS (FAB): 768 $(M + H)^+$.

Determination of Enzyme Inhibitory Activity. The general methods for assays to determine the concentrations of active enzyme inhibitor in solutions of peptides by titration of papain and for assay in dilute systems to determine equilibrium constants for interactions between enzymes and peptides have been described.[41] Briefly, benzoyl-DL-arginine 4-nitroanilide (Bz-Arg-pNA, Bachem) was used as the substrate in chromogenic assays at pH 6.5 to determine the active enzyme concentration in solutions of affinity-purified papain[30] by titration with known amounts of the irreversibly binding papain inhibitor, E-64 (Sigma). This papain solution was then used to similarly determine the amount of papain inhibitory sites in solutions of peptides, through a dilution series and measurement of residual papain activity. Stock solutions of peptides were dissolved to 20 mg/mL in dimethyl sulfoxide (DMSO) and further diluted to 1 mg/mL (1−1.5 mM) in water. The stability of dissolved peptides was checked by papain titration immediately after stock solutions were prepared and compared with results after storage of the solutions for 1 week at +4 °C, with results demonstrating that the soluble concentrations remained constant over the time period.

Continuous-rate assays with the fluorogenic substrate, benzyloxycarbonyl-phenylalanyl-arginyl-aminomethylcoumarin (Z-Phe-Arg-NHMec, Bachem), were used to study the rate of inactivation of papain (above), human cathepsin B (Calbiochem), and recombinant human cathepsin K (produced in *Pichia pastoris* as previously[52]) at 37 °C. The buffer used was 50 mM sodium phosphate, containing 1 mM dithiothreitol (DTT) and 1 mM ethylenediaminetetraacetic acid (EDTA), adjusted to pH 6.5 for papain and pH 6.0 for cathepsin B and K. A Perkin-Elmer LS50 fluorimeter set at 360 nm excitation and 460 nm emission wavelengths was used, and data were analyzed with FLUSYS.[53] Observed rates of inactivation (k_{obs}) were plotted against [I] to determine rate constants for association (k_{+1}) and dissociation (k_{-1}). Steady state rates before (V_0) and after (V_i) addition of inhibitor were used to calculate equilibrium constants for dissociation of enzyme−inhibitor complexes ($K_{i(app)}$).[54]

These constants were corrected for the substrate competition using K_m values determined for the substrate batch, under the assay conditions employed (42, 55, and 9.1 μM for papain, cathepsin B, and cathepsin K, respectively) to obtain true K_i values.

NMR Studies. All NMR spectra were recorded on a 400 MHz Varian Mercury spectrometer. The experiments were carried out in deuterated DMSO (DMSO-d_6). The standard procedure for signal identification was used by utilization of two-dimensional proton−proton correlation spectroscopy (COSY) and NOESY spectra. In our study, we used the procedure of structure determination of small flexible peptides developed in our laboratory.[55] All theoretical calculations were carried out using AMBER empirical force field and the AM1 semiempirical method implemented in the MOPAC2000 package.[56]

Computational Methods. All simulations were carried out using the AMBER 5.0 program.[57] The all-atom AMBER force field was used.[58]

Starting Models. Starting coordinates for all heavy atoms of cathepsin B were obtained from the crystal structure from an aqueous solution (Protein Data Bank entry: 1the). The initial model for cathepsin B−inhibitor complex was prepared in Sybyl program[59] on the basis of computer mutations of Z-Arg-Ser(O-Bzl) complex (1the).[50] Crystallographic water molecules, if any, were removed.

Parametrization. Two new residues were absent in the original AMBER force field: (i) residue Z (benzyloxycarbonyl), which terminates the inhibitor, and (ii) azaglycine. Z and Agly were parametrized in the present study according to the recommendations in the AMBER 5.0 manual. Electrostatic potential-derived charges at the 6-31G* level were used. The molecular orbital calculations, using the GAMESS molecular orbital program package,[60] were carried out for several Z and Agly conformations generated with the help of analysis of crystal structures from the Cambridge Crystallographic Data Base[61] and PM3[62] semiempirical calculations performed in MOPAC.[63] The charges were optimized by fitting them to the ab initio molecular electrostatic potentials with subsequent averaging over all conformations, as recommended in the RESP protocols.[64]

Modeling Procedures. The crude cathepsin B−ligand complex was initially minimized in vacuo to eliminate bad steric contacts. The minimization consisted of 20 000 steps (after 1000 steps, the minimizer was switched from the steepest descent to the conjugate gradient mode). Subsequently, the systems were diluted in TIP3P water, minimized, heated, and equilibrated (5 ps of heating, 12 ps for water equilibration, and finally 10 ps for equilibration of the whole system; 27 ps in total) as described in Table 5.

Table 5. Protocol for Heating

stage	duration (ps)	constraints	temp (K)
1	5	Cα	10−300 (linear)[a]
2	12	Cα	300
3	10		300

[a] Initial atom velocities were assigned from a Maxwell−Boltzman distribution at $T = 10$ K, residue-based cutoff = 12 Å, step = 1 fs.

Molecular Dynamics Productive Runs. Simulations of complexes in solution were performed under periodic boundary conditions in a closed, isothermal, isobaric (NTP) ensemble. Throughout the simulation, the solute and solvent were coupled to a constant temperature ($T = 300$ K) heat bath and constant pressure ($P = 1$ ba) bath.[65] All bonds were constrained using the SHAKE algorithm[66] with a relative geometric tolerance for coordinate resetting of 0.0005 Å, allowing a time step of 1 fs. The Leapfrog version[67] of the Verlet algorithm[68] was employed to integrate the equations of motion. A residue-based cutoff distance of 12 Å was used for nonbonded interactions. The TIP3P model was used for water molecules.[69] Approximately 6500 TIP3P water molecules were in the box, i.e., 23 000 atoms in total. Coordinates were saved every 1000 steps. A long-time MD run of 0.6 ns was carried out.

Acknowledgment. This work was supported by the Polish Scientific Research Committee (KBN), Grant 0998/P04/2000/18, and by the Swedish Medical Research Council (Project Nos. 05196 and 09915). The computations were carried out at the Academic Computer Center in Gdańsk TASK, Poland.

Supporting Information Available: Physicochemical properties of the synthesized azapeptides. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Barrett, A. J.; Kirschke, H. Cathepsin B, Cathepsin H, and Cathepsin L. *Methods Enzymol.* **1981**, *80*, 535−561.
(2) Gradehandt, G.; Ruede, E. The endo/lysosomal protease cathepsin B is able to process conalbumin fragmnets for presentation to T cells. *Immunology* **1991**, *74*, 393−398.
(3) Dunn, A. D.; Cruthcfield, H. E.; Dunn, J. T. Thyroglobulin processing by thyroidal proteases. Major sites of cleavage by cathepsins B, D, and L. *J. Biol. Chem.* **1991**, *266*, 20198−20204.
(4) Taugner, R.; Buhrle, C. P.; Nobiling, R.; Kirschke, H. Coexistence of renin and cathepsin B in epithelioid cell secretory granules. *Histochemistry* **1985**, *83*, 103−108.
(5) Imort, M.; Zuhlsdorf, M.; Feige, U.; Hasilik, A.; von Figura, K. Biosynthesis and transport of lysosomal enzymes in human monocytes and macrophages. Effects of ammonium chloride, zymosan and tunicamycin. *Biochem. J.* **1983**, *214*, 671−678.
(6) Reilly, J. J., Jr.; Mason, R. W.; Chen, P.; Joseph, L. J.; Sukhatme, V. P.; Yee, R.; et al. Synthesis and processing of cathepsin L, an elastase, by human alveolar macrophages. *Biochem. J.* **1989**, *257*, 493−498.
(7) Poole, A. R.; Tiltman, K. J.; Recklies, A. D.; Stoker, T. A. Differences in secretion of the proteinase cathepsin B at the edges of human breast carcinomas and fibroadenomas. *Nature* **1978**, *273*, 545−547.
(8) Sloane, B. F.; Rozhin, J.; Johnson, K.; Taylor, H.; Crissman, J. D.; Honn, K. V. Cathepsin B association with plasma membrane in metastatic tumors. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 2483−2487.
(9) Sloane, B. F.; Cathepsin B cystatins: evidence for a role in cancer progression. *Semin. Cancer Biol.* **1990**, *1*, 137−152.
(10) Chauhan, S. S.; Goldstein, L. J.; Gottesman, M. M. Expression of cathepsin L in human tumors. *Cancer Res.* **1991**, *51*, 1478−1481.
(11) Delaisse, J. M.; Eechout, Y.; Vaes, G. In vivo and in vitro evidence for the involvement of cysteine proteinases in bone resorption. *Biochem. Biophys. Res. Commun.* **1984**, *125*, 441−447.
(12) Lerner, U. H.; Grubb, A. Human cystatin C, a cysteine proteinase inhibitor, inhibits bone resorption in vitro stimulated by parathyroid hormone and parathyroid hormone-related peptide of malignancy. *J. Bone Miner. Res.* **1992**, *7*, 433−440.
(13) Drake, F. H.; Dodds, R. A.; James, I. E.; Connor, J. R.; Debouck, C.; Richardson, S.; et al. Cathepsin K, but not cathepsins B, L, or S, is abundantly expressed in human osteoclasts. *J. Biol. Chem.* **1996**, *271*, 12511−12516.
(14) Gelb, B. D.; Shi, G. P.; Chapman, H. A.; Desnick, R. J. Pycnodysostosin, a lysosomal disease caused by cathepsin K deficiency. *Science* **1996**, *273*, 1236−1238.
(15) Pierre, P.; Mellman, I. Developmental regulation of invariant chain proteolysis controls MHC class II trafficking in mouse dendrititc cells. *Cell* **1998**, *93*, 1135−1145.
(16) Rawlings, N. D.; Barrett, A. J. MEROPS: the peptidase database. *Nucleic Acids Res.* **1999**, *27*, 325−331.
(17) Luaces, A. L.; Barrett, A. J. Affinity purification and biochemical characterization of histolysin, the major cysteine proteinase of Entamoeba histolytica. *Biochem. J.* **1988**, *250*, 903−909.
(18) North, M. J. The characcteristics of cysteine proteinases of parasitic protozoa. *Biol. Chem. Hoppe Seyler* **1992**, *373*, 401−406.
(19) Rosenthal, P. J.; McKerrow, J. H.; Aikawa, M.; Nagasawa, H.; Leech, J. H. A malarial cysteine proteinase is necessary for hemoglobin degradation by Plasmodium falciparum. *J. Clin. Invest.* **1988**, *82*, 1560−1566.
(20) Rosenthal, P. J.; Lee, G. K.; Smith, R. E. Inhibition of a Plasmodium vinckei cysteine proteinase sures murine malaria. *J. Clin. Invest.* **1993**, *91*, 1052−1056.
(21) Chagas, J. R.; Authie, E.; Serveau, C.; Lalmanach, G.; Juliano, L.; Gauthier, F. A comparison of the enzymatic properties of the major cysteine proteinases from *Trypanosoma congolense* and *Trypanosoma cruzi*. *Mol. Biochem. Parasitol.* **1997**, *88*, 85−94.
(22) Bourgeau, G.; Lapointe, H.; Peloquin, P.; Mayrand, D. Cloning, expression, and sequencing of a protease gene (trp) from Porphyromonas gingivalis W83 in *Escherichia coli*. *Infect. Immun.* **1992**, *60*, 3186−3192.
(23) Barrett, A. J.; Rawlings, N. D. Types and famielies of endopeptidases. *Biochem. Soc. Trans.* **1991**, *19*, 707−715.
(24) Bode, W.; Engh, R.; Musil, D.; Thiel, U.; Huber, R.; Karshnikov, A.; Brzin, J.; Kos, J.; Turk, V. The 2.0 Å X-ray crystal structure of chicken egg white cystatin and its possible mode of interaction with cysteine proteinases. *EMBO J.* **1988**, *7*, 2593−2599.
(25) Turk, V.; Bode, W. The cystatins: protein inhibitors of cysteine proteinases. *FEBS Lett.* **1991**, *285*, 213−219.
(26) Henskens, M. C.; Veerman, E. C. I.; Amerongen, A. V. N. Cystatins in health and disease. *Biol. Chem. Hoppe Seyler* **1996**, *377*, 71−86.
(27) Grubb, A. O. Cystatin C − Properties and use as diagnostic marker. *Adv. Clin. Chem.* **2000**, *35*, 63−98.
(28) Björck, L.; Åkesson, P.; Bohus, M.; Trojnar, J.; Abrahamson, M.; Olafsson, I.; Grubb, A. Bacterial growth blocked by a synthetic peptide based on the structure of a human proteinase inhibitor. *Nature* **1989**, *337*, 385−386.
(29) Grubb, A.; Abrahamson, M.; Olafsson, I.; Trojnar, J.; Kasprzykowska, R.; Kasprzykowski, F.; Grzonka, Z. Synthesis of cysteine proteinase inhibitors structurally based on the proteinase interacting N-terminal regions of human cystatin C. *Biol. Chem. Hoppe Seyler* **1990**, *371* (Suppl.), 137−144.
(30) Hall, A.; Abrahamson, M.; Grubb, A.; Trojnar, J.; Kania, P.; Kasprzykowska, R.; Kasprzykowski, F. Cystatin C based peptidyl diazomethanes as cysteine proteinase inhibitors: Influence of the peptidyl chain length. *J. Enzyme Inhib.* **1992**, *6*, 113−123.
(31) Czaplewski, C.; Grzonka, Z.; Jaskolski, M.; Kasprzykowski, F.; Kozak, M.; Politowska, E.; Ciarkowski, J. Binding modes of a new epoxysuccinyl-peptide inhibitor of cysteine proteases. Where and how do cysteine proteases express their selectivity? *Biochim. Biophys. Acta* **1999**, *1431*, 290−305.
(32) Grzonka, Z.; Jankowska, E.; Kasprzykowski, F.; Kasprzykowska, R.; Lankiewicz, L.; Wiczk, W.; Wieczerzak, E.; Ciarkowski, J.; Drabik, P.; Janowski, R.; Kozak, M.; Jaskolski, M.; Grubb, A. Structural studies of cysteine proteases and their inhibitors. *Acta Biochim. Pol.* **2001**, *48*, 1−20.
(33) Gante, J. Peptidomimetics − tailored enzyme inhibitors. *Angew. Chem., Int. Ed. Engl.* **1994**, *33*, 1699−1720.
(34) Gante, J. Azapeptides. *Synthesis* **1989**, 405−413.
(35) Magrath, J.; Abeles, R. H. Cysteine protease inhibition by azapeptide esters. *J. Med. Chem.* **1992**, *35*, 4279−4283.
(36) Baggio, R.; Shi, Y.-Q.; Wu, Y.-q.; Abeles, R. H. From poor substrates to good inhibitors: Design of inhibitors for serine and thiol proteases. *Biochemistry* **1996**, *35*, 3351−3353.
(37) Xing, R.; Hanzlik, R. P. Azapeptides as inhibitors and active site titrants for cysteine proteinases. *J. Med. Chem.* **1998**, *41*, 1344−1351.
(38) Murphy, A. M.; Dagnino, R.; Vallar, P. L.; Trippe, A. J.; Sherman, S. L.; Lumpkin, R. H.; Tamura, S. Y.; Webb, T. R. Automated synthesis of peptide C-terminal aldehydes. *J. Am. Chem. Soc.* **1992**, *114*, 3156−3157.
(39) Gupton, F.; Carroll, D. L.; Tuhy, P. M.; Kam, C.-M.; Powers, J. C. Reaction of azapeptides with chymotrypsin-like enzymes. *J. Biol. Chem.* **1984**, *259*, 4279−4281.

(40) Gray, C. J.; Quibell, M.; Baggett, N.; Hammerle, T. Incorporation of azaglutamine residues into peptides synthesised by the ultrahigh load solid (gel)-phase technique. *Int. J. Pept. Protein Res.* **1992**, *40*, 351−362.
(41) Abrahamson, M. Cystatins. *Methods Enzymol*. **1994**, *244*, 685−700.
(42) Matsumoto, K.; Mizoue, K.; Kitamura, K.; Tse, W. C.; Huber, C. P.; Ishida, T. Structural basis of inhibition of cysteine proteases by E−64 and its derivatives. *Biopolymers* **1999**, *51*, 99−107.
(43) Hall, A.; Dalboge, H.; Grubb, A.; Abrahamson, M. Importance of the evolutionarily conserved glycine residue in the N-terminal region of human cystatin C(Gly-11) for cysteine endopeptidase inhibition. *Biochem. J*. **1991**, *291*, 123−129.
(44) Menard, R.; Carmona, E.; Plouffe, C.; Brömme, D.; Konishi, Y.; Lefebvre, J.; Storer, A. C. The specificity of the $S_{1'}$ subsite of cysteine proteases. *FEBS Lett*. **1993**, *328*, 107−110.
(45) Turk, B.; Turk, D.; Turk, V. Lysosomal cysteine proteases: more than scavengers. *Biochim. Biophys. Acta* **2000**, *1477*, 98−111.
(46) Turk, B.; Turk, V.; Turk, D. Structural and functional aspects of papain-like cysteine proteinases and their protein inhibitors. *Biol. Chem*. **1997**, *378*, 141−150.
(47) Burgen, A. S. V.; Roberts, C. C. K.; Feeney, J. Binding of flexible ligands to macromolecules. *Nature* **1975**, *253*, 753−655.
(48) LaLonde, J. M.; Zhao, B.; Janson, C. A.; D'Alessio, K. J.; McQueney, M. S.; Orsini, M. J.; Debouck, C. M.; Smith, W. W. The crystal structure of human procathepsin K. *Biochemistry* **1999**, *38*, 862−869. (b) Sivaraman, J.; Lalumiere, M.; Menard, R.; Cygler, M. Crystal structure of wild-type human procathepsin K. *Protein Sci*. **1999**, *8*, 283−290.
(49) Storer, A. C.; Menard, R. Recent insights into cysteine protease specificity: Lessons for drug design. In *Perspectives in Drug Discovery and Design*; McKerrow, J. H., James, M. N. G., Eds.; ESCOM Science Publishers B. V.: Leiden, 1996; Vol. 6, pp 33−46.
(50) Jia, Z.; Hasnain, S.; Hirama, T.; Lee, X.; Mort, J. S.; To, R.; Huber, C. P. Crystal structures of recombinant rat cathepsin B and a cathepsinB-inhibitor complex. *J. Biol. Chem*. **1995**, *270*, 5527−5533.
(51) Bodanszky, M. *Practice in Peptide Chemistry*; Springer-Verlag: Weinheim, 1984.
(52) Linnevers, C. J.; McGrath, M. E.; Armstrong, R.; Mistry, F. R.; Barnes, M. G.; Klaus, J. L.; Palmer, J. T.; Katz, B. A.; Brömme, D. Expression of human cathepsin K in *Pichia pastoris* and preliminary crystallographic studies of an inhibitor complex. *Protein Sci*. **1997**, *6*, 919−921.
(53) Rawlings, N. D.; Barrett, A. J. FLUSYS: a software package for the collection and analysis of kinetic and scanning data from Perkin-Elmer fluorimeters. *Comput. Appl. Biosci*. **1990**, *6*, 118−119.
(54) Henderson, P. J. F. A linear equation that describes the steady-state kinetics of enzymes and subcellular particles interacting with tightly bound inhibitors. *Biochem. J*. **1972**, *127*, 321−333.
(55) Groth, M.; Malicka, J.; Czaplewski, C.; Oldziej, S.; Lankiewicz, L.; Wiczk, W.; Liwo, A. Maximum entropy approach to the determination of solution conformation of flexible polypeptides by global conformational analysis and NMR spectroscopy − Application to DNS^1-c[D-A$_2$bu^2, Trp^4, Leu^5]enkephalin and DNS^1-c-[D-A$_2$bu^2, Trp^{4,} D-Leu^5]enkephalin. *J. Biomol. NMR* **1999**, *15*, 315−330.
(56) *MOPAC2000*; Fuitsu Ltd., 2000.
(57) Case, D. A.; Pearlman, D. A.; Caldwell, J. W.; Cheatham, T. E., III; Ross, W. S.; Simmerling, C.; Darden, T.; Merz, K. M.; Stanton, R. V.; Cheng, A.; Vincent, J. J.; Crowley, M.; Ferguson, D. M.; Radmer, R.; Seibel, G. L.; Singh, U. C.; Wiener, P.; Kollman, P. A. *Amber 5.0*; University of California: San Francisco, 1997.
(58) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M., Jr.; Fergusson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A second generation force field for the simulation of proteins, nucleic acids and organic molecules. *J. Am. Chem. Soc*. **1995**, *117*, 5179−5197.
(59) *SYBYL 6.1*; Tripos Inc., 1699 S. Manley Rd., St. Louis, MO 63144, 1994.
(60) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S.; Windus, T. L., Dupuis, M.; Montgomery, J. A. general atomic and molecular electronic structure system. *J. Comput. Chem*. **1993**, *14*, 1347−1363.
(61) Allen, F. H.; Bellard, S.; Brice, M. D.; Cartwright, B. A.; Doubleday, A.; Higgs, H.; Hummelink, T.; Hummmelink-Peters, B. G.; Kinnard, O.; Motherwall, W. D. S.; Rogers, J. R.; Watson, D. G. The Cambridge Crystallographic Data Centre: Computer-based search, retrieval, analysis and display of information. *Acta Crystallogr. B* **1979**, *B35*, 2331−2339.
(62) Stewart, J. J. P. Optimization of parameters for semiempirical methods I − method. *J. Comput. Chem*. **1989**, *10*, 209−220. (b) Stewart, J. J. P. Optimization of parameters for semiempirical methods II − applications. *J. Comput. Chem*. **1989**, *10*, 221−264.
(63) Stewart, J. J. P. MOPAC: A semiempirical molecular orbital program. *J. Comput.-Aided Mol. Design* **1990**, *4*, 1−105.
(64) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: The RESP model. *J. Phys. Chem*. **1993**, *97*, 10269−10280.
(65) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem Phys*. **1984**, *81*, 3684−3690.
(66) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. Numerical integration of the Cartesian equation of motion of a system with constraints: molecular dynamics of N-alkanes. *J. Comput. Phys*. **1977**, *23*, 327−341.
(67) Hockney, R. W. The potential calculation and some applications. *Methods Comput. Phys*. **1970**, *9*, 136−211.
(68) Verlet, L. Computer experiments on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules. *Phys. Rev*. **1967**, *159*, 98−103.
(69) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys*. **1983**, *79*, 926−935.

JM020850K