Journal of Peptide Science
J. Pept. Sci. 2007; **13**: 536–543
Published online 6 July 2007 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/psc.883

JOURNAL OF PEPTIDE SCIENCE

An enormously active and selective azapeptide inhibitor of cathepsin B

EWA WIECZERZAK,[a] SYLWIA RODZIEWICZ-MOTOWIDŁO,[a]* ELŻBIETA JANKOWSKA,[a] ARTUR GIEŁDOŃ[a,b] and JERZY CIARKOWSKI[a]

[a] Faculty of Chemistry, University of Gdańsk, Sobieskiego 18, 80-952 Gdańsk, Poland
[b] Johan Wolfgang Goethe University, Institute of Biochemistry, Max-von-Laue-Str. 7, D-60438 Frankfurt a.M., Germany

Received 23 April 2007; *Revised 11 May 2007*; *Accepted 13 May 2007*

Abstract: The peptidomimetic Z-Arg-Leu-Arg-Agly-Ile-Val-OMe (where Agly means α-aza-glycyl, -NHNHCO-) is the strongest ($K_i = 480$ pM) and the most selective inhibitor of cathepsin B to date, being ~2310 times as active to cathepsin B as to cathepsin K. In this paper we introduce the peptide and seek to rationalize its structure-activity relationships using molecular dynamics (MD) and NMR. It is shown that the -Agly-moiety restrains the peptide backbone to a bent shape, contrary to its parent peptide (with Gly in position 4), having its backbone extended and flexible. This fold is maintained in the plug covalently bound to the cathepsin B Cys29, in compliance with similar bends already observed in two other azapeptides attached to the active sites of cathepsin B. The MD simulation of the Z-Arg-Leu-Arg-Agly ~ cathepsin B complex suggests that, contrary to other potent inhibitors of cathepsin B, the current double Arg^1/Arg^3 inhibitor, while maintaining the fold is able to form a unique ion cluster involving both Arg residues on the inhibitor part and two acidic Glu171 and Glu245 on the cathepsin B part, thus enhancing the affinity and subsequently the inhibiting power and selectivity of Z-Arg-Leu-Arg-Agly-Ile-Val-OMe to the observed extreme extent.

Supplementary electronic material for this paper is available in Wiley InterScience at http://www.interscience.wiley.com/jpages/1075-2617/suppmat/

Keywords: cathepsin B and K; inhibitor; peptidomimetic; structure-activity; NMR; molecular modeling; α-aza-glycyl

INTRODUCTION

Cathepsins B (EC 3.4.22.1) and K (EC.3.4.22.38) are cysteine proteases of the papain family [1]. Cathepsin B acts mainly as a component of the intracellular protein degradation system in the lysosomes. Extracellular cathepsin B has been implicated in inflammatory airway disease [2] and bone and joint disorders [3]. Cathepsin K is predominantly expressed in osteoclasts [4,5] where it plays a key role in osteoclast-mediated bone resorption [6,7], which implicates that specific inhibitors of this enzyme could be used for treatment of such diseases as osteoporosis, characterized by an imbalance between the formation and resorption of the bone matrix. Disturbance in the normal balance of enzymatic activity of cysteine proteases leads to pathological conditions. Most of them are connected with the proteolytic function of cathepsins outside lysosomes. Structurally, all cathepsins of the papain family are homologous with the parent enzyme and are characterized by a 'substrate trough', causing an enzyme to take up a two-domain coffee bean-like shape. The bottom of this channel is lined with the catalytic triade, Cys29:His199:Asn219 in cathepsin B. The latter family member is structurally unique as the primed subsites (S1′, S2′ ...) of its substrate trough [8] are, contrary to other family members, filled with a unique sequence insert termed the occluding loop, which accordingly biases the enzyme towards exo-carboxydipeptidase activity [9,10]. The activity of all cathepsins *in vivo* is regulated *inter alia* by their interactions with specific inhibitors, cystatins [11–13]. In recent years, there has been an increasing interest in the design of synthetic cysteine protease inhibitors as agents for treatment of various diseases. Special effort has been put for the search of low molecular mass inhibitors [14]. A number of peptidyl derivatives structurally based on the inhibitory sites of cystatins has been synthesized [15] but these compounds as being prone to proteolytic degradation, are rapidly excreted and poorly bioavailable. The majority of these problems might be overcome by use of peptidomimetics with structure resemblance to these previously synthesized peptidyl derivatives. Among the peptide mimetics there are azapeptides, peptide analogs in which α-CH moiety of one or more amino acid residues in the peptide chain is replaced by an N atom [16]. In our present work we have studied the peptidomimetic Z-Arg-Leu-Arg-Agly-Ile-Val-OMe, where Agly denotes α-aza-glycyl, -NHNHCO-. The parent hexapeptide Z-Arg-Leu-Arg-Gly-Ile-Val-OMe is a modification of a generic substrate sequence -Arg-Leu-Val-Gly-Ile-Val- being a combination of a tetrapeptide

*Correspondence to: Sylwia Rodziewicz-Motowidło, Faculty of Chemistry, University of Gdańsk, Sobieskiego 18, 80-952 Gdańsk, Poland; e-mail: sylwia@chem.univ.gda.pl

-RLVG- and a dipeptide -IV- from the *N*-terminus and Loop 1, respectively, of human cystatin C, an endogenic protein inhibitor of cathepsins [17,18]. In the three-dimensional structure of cystatins these two sequence segments make key structural contribution to the cystatinic 'wedge', fitting the trough to cap the catalytic pocket in cysteine proteases of the papain family.

METHODS

Synthesis

The syntheses were carried out in solution using TBTU or EDC as coupling agents. The incorporation of α-aza-glycyl was accomplished by the coupling of Boc-$NHNH_2$ and an appropriate amino acid/peptide ester with *N*,*N*′-carbonyldiimidazole (CDI) as an α-aza-glycyl carbonyl donor [17]. Crude products were pre-purified by ion exchange chromatography (Sepharose S Fast Flow, linear gradient 0–0.2 M KCl in 50% MeOH, pH 4.5). Final purification was achieved using reverse-phase high performance liquid chromatography (RP-HPLC) with a C8 Kromasil column (25 × 250 mm, 7 μm). Purity of peptides was checked by analytical RP-HPLC (C8 Kromasil column 4.6 × 250 mm, 5 μm) in 30 min linear gradient of 0–80% acetonitrile in 0.1% aqueous trifluoroacetic acid as a mobile phase and with UV detection at 223 nm. The peptides identity was confirmed by mass spectrometry (MALDI TOF, Bruker). Z-Arg-Leu-Arg-Gly-Ile-Val-OMe: $t_R = 17.1$ min, $M(M)^+ = 863.1$ Da, $M_{calc} = 863.2$ Da; Z-Arg-Leu-Arg-Agly-Ile-Val-OMe: $t_R = 19.6$ min, $M(M)^+ = 862.2$ Da, $M_{calc} = 862.2$ Da.

Enzyme Inhibitory Activity

The assays were done as already described [19], using benzoyl-DL-arginine 4-nitroanilide (Bz-Arg-pNA, Bachem) as the substrate in chromogenic assays at pH 6.5 to determine the active enzyme concentration in solutions of papain (Sigma) 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 one 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-aminomethyl-coumarin (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* in the same manner as previously [20]) at 37 °C. The buffer used was 50 mM sodium phosphate, containing 1 mM DTT and 1 mM EDTA, adjusted to pH 6.5 for papain, and pH 6.0 for cathepsin B and K. Perkin-Elmer LS50 fluorimeter set at 360 nm excitation and 460 nm emission wavelengths was used, and data was analyzed with FLUSYS [21]. 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 [$K_{i(app)}$] [22]. 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.

Molecular Modeling

All simulations were carried out using the AMBER 7.0 all-atom force field [23]. Starting coordinates for all heavy atoms of cathepsin B were obtained from the crystal structure of Z-Arg-Ser(Bzl)-cathepsin B complex (Protein Data Bank entry: 1the, [9]). The initial model for cathepsin B-inhibitor complex was prepared in Sybyl program [24]. Crystallographic water molecules, if any, were removed. Two new moieties, absent in the original AMBER force field, i.e. the Z group terminating the inhibitor and the aza-glycyl, have already been parameterized [17,18] as recommended in the AMBER 7.0 manual. The newly built model of cathepsin B-inhibitor was placed in the TIP3P water [25]. To eliminate all bad steric contacts the following optimization was done three times: 500 steepest descent minimization and 1000 molecular dynamic steps with periodic boundary conditions and 30 K temperature. In this procedure all Cα atoms and its *N*-counterpart in Agly were frozen. Subsequently 500 steepest descent and 500 conjugate minimization steps were carried out. During these two steps all atoms were set free to move. Subsequently the system was linearly heated from 10 to 300 K during 10 ps simulation. Finally, the productive 1 ns molecular dynamic (MD) were performed under periodic boundary conditions in a closed, isothermal–isobaric (NTP) ensemble. The solute and solvent were coupled to a constant-temperature ($T = 300$ K) heat bath and constant-pressure (P = 1 ba) bath [26] throughout the simulation. All bonds were constrained using the SHAKE algorithm [27] with a relative geometric tolerance for coordinate resetting of 0.0005 Å, allowing a time step of 1 fs. A residue-based cutoff distance of 12 Å was used for nonbonded interactions. Approximately 6500 TIP3P water molecules were in the box, leading to ~23 000 atoms total. Coordinates were saved every 1000 steps.

NMR Experiment

All NMR measurements were performed in DMSO-d_6 with peptide concentration 14 mM, on Varian Unity 500 Plus spectrometer operating at 500 MHz resonance frequency. All spectra were recorded at 303 K except for the temperature coefficients of the chemical shifts, which were measured for the amide proton resonances throughout the temperature range 295–318 K. The assignment of the proton chemical shifts was accomplished by means of 2D proton spectra: TOCSY (90 ms), NOESY (200 ms, 250 ms, 300 ms), ROESY (200 ms) and DQF-COSY. All data were processed using VNMR [28] and XEASY [29] software. The nuclear overhauser effects (NOE) intensities were picked up from the NOESY spectrum (200 ms). The NOE volumes were integrated and calibrated

DOI: 10.1002/psc

with the XEASY software [29]. Because of the spectral overlap of **4** peptide, only well-separated cross-peaks were integrated and further used in the structure calculations. For 3D structure calculations the NOESY experiments provided for peptide **3** - 79 NOE effects and for peptide **4** - 60 NOE effects. The $^{3}JH_{N\alpha H}$ coupling constants were obtained from DQF-COSY spectrum. Temperature dependences of the chemical shifts of NH protons ($\Delta\sigma/\Delta T$) were calculated from 1D spectra recorded at 295, 303, 308, 313 and 318 K. NMR structure calculations were done using the ANALYZE program dedicated to structure determination of small flexible peptides [30]. All MD simulations were carried out using the AMBER 7.0 program [23]. The all-atom AMBER force field and explicit DMSO molecules as a solvent were used. The set of conformations from the whole MD trajectory was clustered into 50 families of conformations. The theoretical intensities of NOE and the $^{3}J_{HNH\alpha}$ vicinal coupling constants were calculated for the lowest energy conformations from each family. The intensities of NOE effects were computed by solving the system of the Bloch differential equations applying the MORASS program [31]. The vicinal $^{3}J_{HNH\alpha}$ coupling constants of low-energy conformations were calculated from the empirical Karplus equation [32]. On the basis of the sets of measured and theoretically calculated observables the statistical weights of the conformations were fitted to obtain the best agreement between the theoretical and experimental spectra. The best fit to the experimental NOESY spectrum was obtained after superposition of five low-energy conformations which statistical weights were higher than 3%.

RESULTS

Enzyme Inhibitory Activity

In our pursuit for effective and selective human cathepsin inhibitors [17,18] we have found an enormously potent and selective azapeptide inhibitor of cathepsin B, namely Z-Arg-Leu-Arg-Agly-Ile-Val-OMe, introduced in this paper. Its biological properties are compared in Table 1 with two selected inhibitors designed and synthesized earlier. Along with the concept put forth in the introduction on design of papain-like protease inhibitors, those listed in Table 1 contain Agly instead of the evolutionarily conserved P1 glycine-11 from an *N*-terminal tetrapeptide sequence fragment of human cystatin C, see also Refs 17,18. Azapeptides inactivate cysteine proteases reversibly [33] hence the equilibrium constants K_i for dissociation of the inhibitor–enzyme complex make adequate measure of inhibiting potency. A review of crystal structures of substrate-mimicking inhibitors bound to the active sites of cysteine proteases reveal that their P3, P2 and P1′ side chains make contacts with enzymes surface so that the S3, S2 and S1′ subsites of the enzymes make accordingly potential targets for the design of selective inhibitors [34,35].

Molecular Modeling

In our previous papers [17,18] we hypothesized that, contrary to selective cathepsin K inhibitors tending to place their P1, P2 and P3 residues extended in the enzyme-inhibitor complex, the requirements for high selectivity to cathepsin B are different because, (1) the respective chains of residues (e.g. the Z-RLV-Agly- and Z-RLH-Agly- chains of compounds **1** and **2**, respectively) are bent in their complexes with cathepsin B; (2) apart of being hydrophobic in both catepsins, the subsite S2 made of Ala173, Gly198, Glu245 in catepsin B, accepts also aromatic residues [36] (3) an interaction between inhibitor's Arg and Glu245 lining S2 in cathepsin B but not in cathepsin K may enhance the affinity. While the above arguments (1) and (2) have been confirmed, the contact between the inhibitor's Arg and enzyme's Glu245 was found for only **1** [17] but not for **2**. On the contrary, the analysis of preferences of occurrence of the inhibitor's **2** Arg using MD simulation indicated its complete promiscuity [18]. On the other hand, the formally P2 residue His of **2** preferred – while manifesting its aromatic–aromatic empathy – a vicinity of the enzyme's Trp30 and other residues lining the subsite S1, compare Figures 2 and 3 in Ref. 18. Yet, an analogical simulation in the current study of the inhibitor **3**, which comprises the second Arg in the

Table 1 Inhibitory potencies of azapeptides *versus* Papain, human Cathepsin B and recombinant human Cathepsin K

Compound	Structure	K_i (nM)[a]			Selectivity:
		Papain	cathepsin K	cathepsin B	K_i(cathepsin K)/K_i(cathepsin B)
1[b]	Z-Arg-Leu-Val-Agly-Ile-Val-OMe	0.77 ± 0.28	21.0 ± 1.3	0.088 ± 0.025	240
2[c]	Z-Arg-Val-His-Agly- Ile-Val-OMe	1.04 ± 0.06	0.55 ± 0.09	0.011 ± 0.003	500
3	Z-Arg-Val-Arg-Agly- Ile-Val-OMe	8.85 ± 0.36	1.11 ± 0.28	0.00048 ± 00 013	2310
4	Z-Arg-Leu-Arg-Gly-Ile-Val-OMe	n.m.[d]	n.m.	n.m.	n.m.

[a] The K_i values were calculated as detailed in the Methods section from 3 to 5 continuous rate assays, and presented as mean value ± SD [17,19].
[b] Ref. 17.
[c] Ref. 18.
[d] Not measured.

DOI: 10.1002/psc

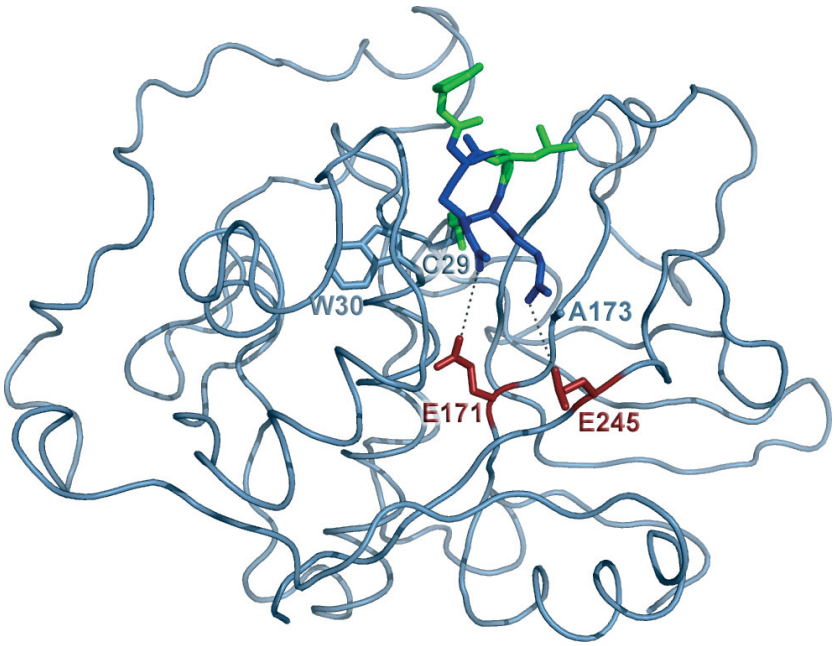

Figure 1 The structure of Z-Arg-Leu-Arg-Agly-cathepsin B complex after 900 ps of MD simulations.

equivalent position to His in **2**, clearly indicates that this other formally P2 Arg, makes indeed an ion pair-type interaction with Glu245 in the subsite S2 of cathepsin B(Figure 1). Moreover, the side chains of both Arg residues of **3** attached to the enzyme (Z-RLR-Agly∼cathepsin B) run extended and in parallel at some 4.5–5.0 Å away of each other. This otherwise surprising observation gets its rationale after having realized that in the MD-optimized complex of **3** with cathepsin B, both inhibitor's Arg residues are involved in a cluster of ions contributed, apart by themselves, by two glutamates, namely Glu245 (see above) and Glu171 from the enzyme,(Figure 1). Interestingly, the most populated structure(s) of the mother azahexapeptide Z-Arg-Leu-Arg-Agly-Ile-Val-OMe found by NMR in DMSO-d_6, see the subsequent section, also show both Arg side chains running parallel extended close to each other, demonstrating that this tendency may be intrinsic to the inhibitor and not just imposed on its chain by the cathepsin B environment.

NMR Experiment

To supplement the observations resulting from MD alone we did NMR studies of **3** and of its parent hexapeptide substrate Z-Arg-Leu-Arg-Gly-Ile-Val-OMe, **4**. Their chemical shift data are given in Tables S1 and S2 (in supplementary materials), respectively while their TOCSY spectra are shown accordingly in Figures 2 and 3. The NMR spectra indicate that more than one set of proton resonances in the spectra is observed for inhibitor **3**. In case of Arg and Leu residues – three, and for Agly, Ile and Val – two sets of proton resonances were identified. The minor conformations

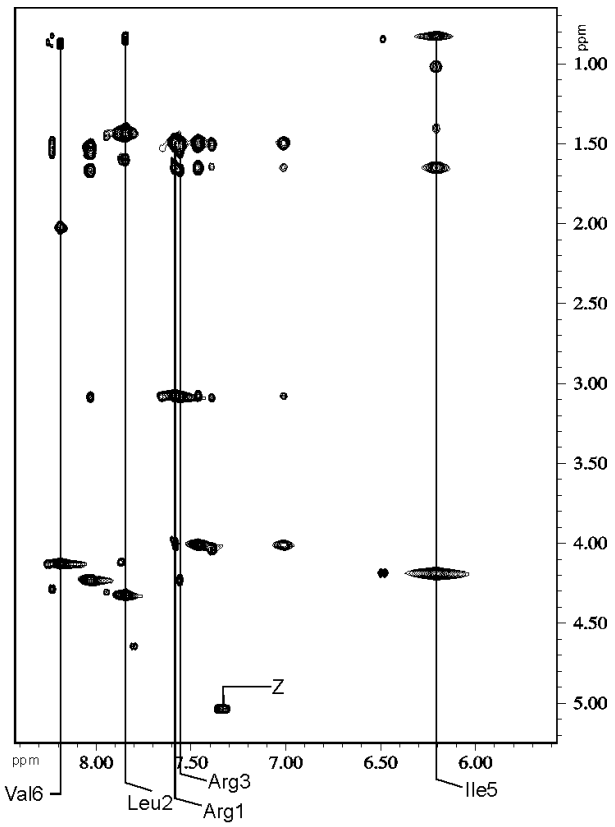

Figure 2 Amino acid spin systems (diagnostic region) in the TOCSY spectrum of **3** in DMSO-d_6 ($T = 303$ K, mixing time 90 ms).

were identified on the basis of the low intensity of the characteristic proton spin systems with similarity to the patterns displayed by the minor conformation. The

DOI: 10.1002/psc

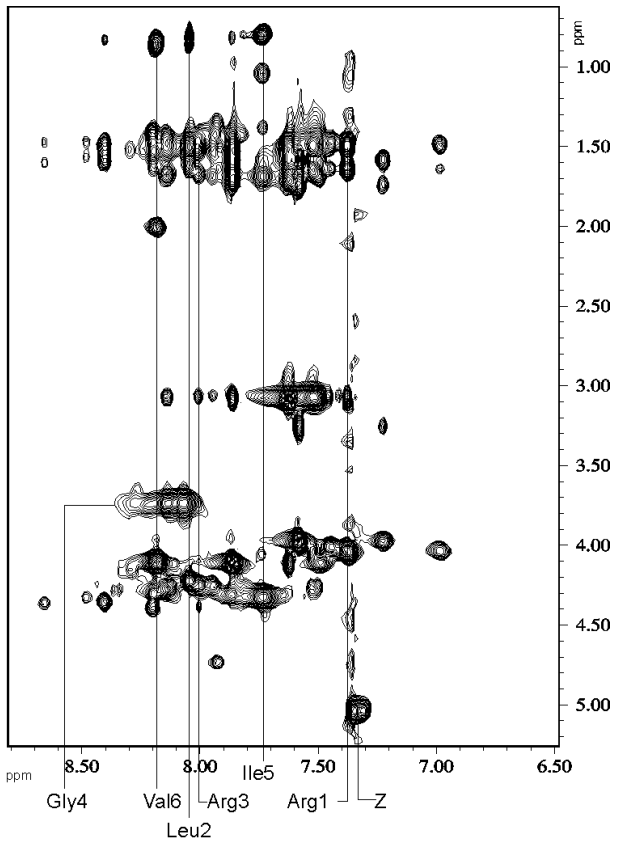

Figure 3 Amino acid spin systems (diagnostic region) in the TOCSY spectrum of **4** in DMSO-d_6 ($T = 303$ K, mixing time 90 ms).

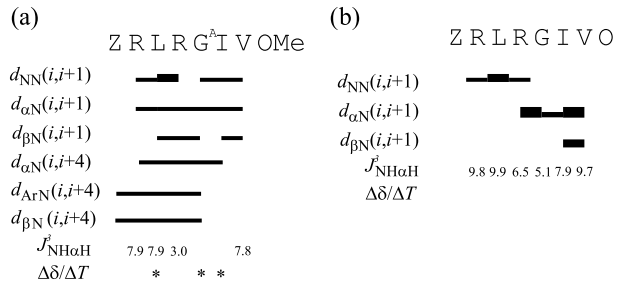

Figure 4 The integral intensities of off-diagonal signals in NOESY spectrum in DMSO-d_6 at 303 K, the vicinal coupling constants ($^3J_{HN\alpha H}$) and the amide protons temperature coefficients ($\Delta\delta/\Delta T$) for major conformations of (a) Z-Arg-Leu-Arg-Agly-Ile-Val-OMe and (b) Z-Arg-Leu-Arg-Gly-Ile-Val-OMe. Small values of the amide protons temperature coefficients are marked with the asterisks; G^A – α-aza-glycyl residue.

percentage contribution of the major and the minor conformations to the peptide structure was measured by comparison between the peak volumes of the identical protons' NOE in the NOESY spectrum (NH–βH signals of the Leu residue). The calculated percentage of minor I and minor II structures was about 4%. All peptide bonds in the major conformation were found to be in *trans* configuration. The strong NOE effects for $H^{\alpha}{}_iHN_{i+1}$ and HN_iHN_{i+1} connectivities through the whole molecule were found (Figure 4(a)). The NOE effects observed for sequential connectivities between $H^{\alpha}{}_iHN_{i+4}$, and also $H^{\beta}{}_iHN_{i+4}$ and $HH_iH^{CH3}{}_{i+4}$ and one medium range $H^{Ar}{}_iH^{\alpha}{}_{i+4}$ contact for Z and Agly protons, indicate the bent structure of the whole peptide. The cross-peaks seen for $H^{\alpha}{}_iHN_{i+1}$ and $H^{\beta}{}_iHN_{i+1}$ point out that the β-turn at the region Arg^1–Val^6 is well formed and stable. The small value of vicinal coupling constant for Arg^4 (3 Hz, Figure 4(a)) indicates a bend structure in this region. Small values of the temperature coefficients of amide protons (Figure 4(a)) suggest the presence of strong hydrogen bonds for NH protons of Leu, Agly and Ile amino acid residues. These protons may be involved in the hydrogen bonds with carbonyl group of other amino acid residues or with the oxygen atoms of DMSO. Superposition of the selected conformations over all Cα atoms, yields a root mean square deviation (RMSD) value of1.66 Å (Figure 5(a)). The *N*- and *C*-terminal fragments of Z-Arg-Leu-Arg-Agly-Ile-Val-OMe molecule in each conformation are close to each other, indicating a bent structure in the middle part of the molecule. All conformations are stabilized by β-turns in the region Arg^1-$Agly^4$ and Leu^2-Ile^5 (Table S3 in supplementary materials). The side chains of the inhibitor's amino acid residues are exposed to the solvent. In the most populated conformation the side chains of Arg^1, Arg^3 and Ile^5 are situated at one side of the molecule whereas the hydrophobic groups of Z and Ile^5 residues are closely oriented to form a hydrophobic cluster(Figure 5(a)). The percentage contribution of the major and the minor conformations to the peptide **4** structure was measured by comparison between the peak volumes of identical protons in the TOCSY spectrum (NH–αH signals of the Gly residue, Figure 3). The calculated percentage of four most abundant structures was: 31 (I), 20 (II), 30, (III) and 19% (IV). All peptide bonds in the major conformation were found to be in *trans* configuration. The inspection of the NOE pattern of **4** (Figure 4(b)) shows lack of diagnostic NOE effects and argues for an unstructured conformation of this peptide. The $^3J_{HNH\alpha}$ vicinal coupling constants (Figure 4(b)), except for Gly^4 amino acid residue, are in the range of 6–9 Hz. These data are consistent with a statistical-coil or extended conformation and corroborate the results from NOE experiments. The values of the temperature coefficients of amide protons suggest no hydrogen bonds for NH protons in peptide **4**. The lack of hydrogen bonds indicates its flexible structure. The superposition of the selected conformations over all Cα atoms, yields a root mean square deviation (RMSD) value of 1.90 Å (Figure 5(b)). The calculated structures of Z-Arg-Leu-Arg-Gly-Ile-Val-OMe molecule are extended and have no regular structure motifs. Conformations are stabilized by occasional β-turns at variable positions (Table S4

DOI: 10.1002/psc

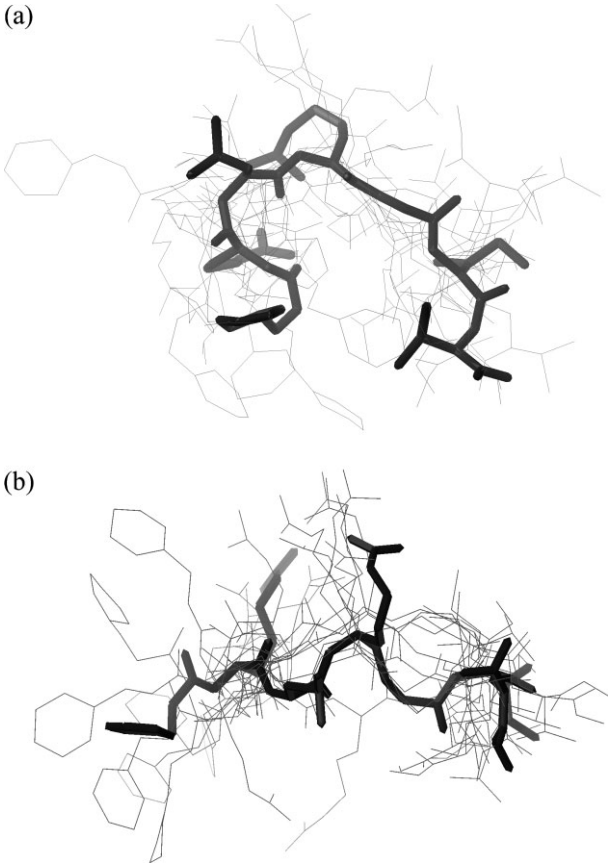

Figure 5 The low-energy conformation sets of (a) Z-Arg-Leu-Arg-Agly-Ile-Val-OMe and (b) Z-Arg-Leu-Arg-Gly-Ile-Val-OMe structures; α-carbon atoms were superposed. The highest-statistical weight conformations are exposed thick.

in supplementary materials). In the most populated conformation the side chains of Arg1 and Arg3 are situated on a common side of the molecule whereas the other residues are oriented in variable directions.

The comparison of the NMR spectra for **3** and **4**, shows that both the inhibitor **3** and its parent substrate **4** are characterized by a major and three or two minor conformations in DMSO. In the case of peptide **3** only one minor conformation for the $Agly^4$–Val^6 sequence was found, whereas in the case of its parent peptide **4**, three minor conformations were observed. This indicates that the replacement of the α-CH group in Gly residue by the N atom results in a rigid conformation and also results with a bent structure of the azapeptide peptides. On the contrary, the structures of the substrate are more disordered and flexible. On Figure 6, the NMR-derived major conformation of **3** superimposed on the conformation with the highest statistical weight of this inhibitor attached in the catalytic center of cathepsin B is shown. The overlap was done over the corresponding Z-RLR-regions using the Cα carbon atoms and appears reasonably good. Thus, it argues not only for a reasonability of the conformation of the inhibitor **3** attached to cathepsin B,

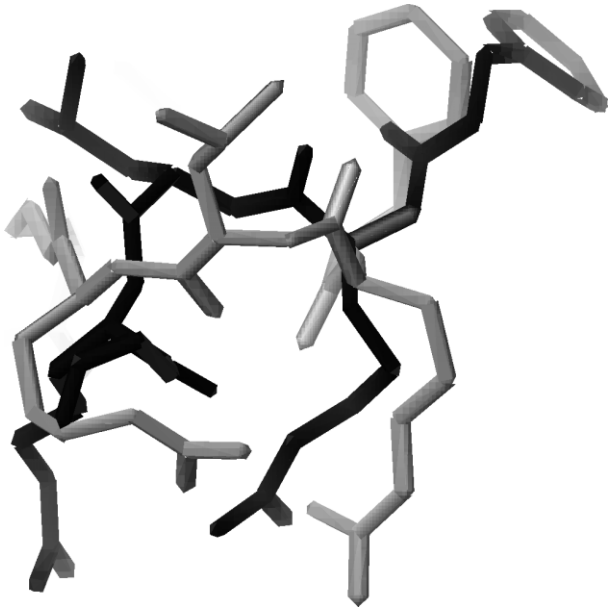

Figure 6 The NMR-derived major conformation with the highest statistical weight of Z-Arg-Leu-Arg-Agly-Ile-Val-OMe (gray) superimposed on the low-energy conformation of this inhibitor attached in the catalytic center of cathepsin B (black). The overlap was done over the heavy atoms within the corresponding Z-RLR-frames. RMS = 4.50 Å.

as shown in Figure 1, but also for an intrinsic tendency of the Z-RLR-region of the azapeptide **3** to form a bent having both Arg side chains extended in parallel not far from each other.

DISCUSSION

Our results indicate an intrinsic structural prevalence of the *N*-terminal part of the azapeptide **3** to stay bent, while exposing its both Arg side chains extended parallel not far from each other. This feature is evident both in a 'free' azahexapeptide dissolved in DMSO and in an *N*-terminal tetrapeptide covalently attached to the active site Cys29 in cathepsin B(Figure 6). The inhibitor **3** shares the tendency of being bent with the azapeptide inhibitors **1** and **2** [17,18]. For instance, a superposition of the highest statistical weight conformers of Z-Arg-Leu-Arg-Agly-Ile-Val-OMe **3** and Z-Arg-Leu-His-Gly-Ile-Val-OMe **2** [18], resolved using NMR in DMSO-d_6 is quite good, as can be realized by viewing Figure 7. Yet the other two analogs (**1** and **2**), although designed from common grounds with **3**, carry only one Arg (in position 1 of the peptide, $Z\text{-}R^1$-, or position P4, according to the Schechter–Berger nomenclature [8]). Although surprising enough in solution (Figures 5(a) and 7), the extended parallel side chains of both Arg^1 and Arg^3 (P4 and P2, respectively, according to [8]) are well warranted in the catalytic pocket of cathepsin B, where as shown above, they are involved in an ion cluster contributed by this

DOI: 10.1002/psc

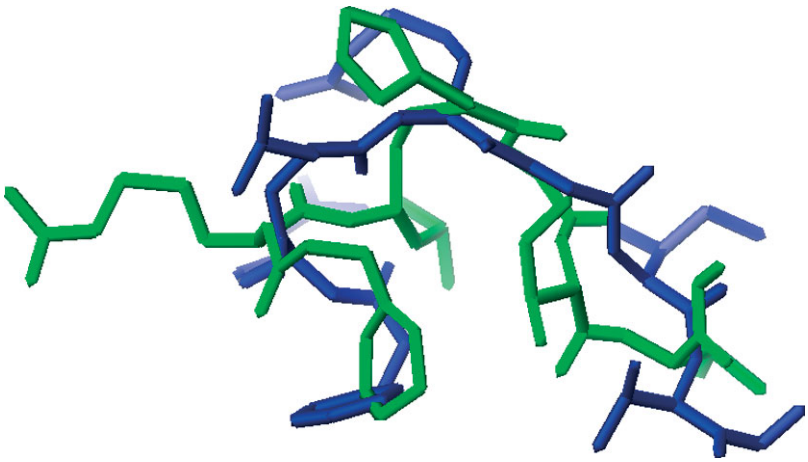

Figure 7 Superposition of the highest statistical weight conformations of Z-Arg-Leu-Arg-Agly-Ile-Val-OMe (blue) and Z-Arg-Leu-His-Gly-Ile-Val-OMe (green) [18] NMR structures. The overlap was done over the backbone Cα atoms RMS = 2.17 Å.

pair of Arg residues from the inhibitor and a pair of Glu171 and Glu245 residues, lining the S2 subsite in the enzyme's substrate trough(Figure 1). Indeed, the former azapeptide inhibitors **1** and **2** while trying to accommodate an ion bridge to the subsite S2 using their either Arg or His residues, respectively (see above and [17,18]), could not accomplish a *tour de force* interaction only feasible for a cluster of a pair of Arg and a pair of Glu in the Z-Arg-Leu-Arg-Agly~cathepsin B complex, subject this work. This would validate an extreme inhibiting power of **3**. On the other hand, the equivalent positions in cathepsin K are occupied by S132 and L209 (in place of E171 and E245, respectively, in cathepsin B [37]), making the S2 subsite in this enzyme inappropriate either for a single or for a double Arg residue. This would provide a rationale for extreme selectivity of the azapeptides **1–3**, and particularly of the latter in favor of cathepsin B against cathepsin K.

Supplementary Material

Supplementary electronic material for this paper is available in Wiley InterScience at: http://www.interscience.wiley.com/jpages/1075-2617/suppmat/

Acknowledgements

This work was supported by University of Gdańsk, Grant BW/8000-5-0280-6 and DS 8372-4-0138-7 and DS 8350-5-0131-7. Enzyme inhibitory activity was determined in the Department of Clinical Chemistry of Lund University Hospital, Sweden. Authors would like to thank Prof. Anders Grubb and Dr Magnus Abrahamson for access to their facilities and expert help and Ewa Piotrowska for the technical support in NMR experiment. The calculations were carried out in the TASK in Gdańsk, Poland.

REFERENCES

1. Barrett AJ, Kirschke H. Cathepsin B, Cathepsin H, Cathepsin L. *Methods Enzymol.* 1981; **80**: 535–561.
2. Burnett D, Abrahamson M, Devalia JL, Sapsford RJ, Davies RJ, Buttle DJ. Synthesis and secretion of procathepsin B and cystatin C by human bronchial epithelial cells in vitro: modulation of cathepsin B activity by neutrophil elastase. *Arch. Biochem. Biophys.* 1995; **317**: 305–310.
3. Buttle DJ. Lysosomal cysteine endopeptidases in the degradation of cartilage and bone. In *Immunopharmacology of Joints and Connective Tissue*, Dingle JT, Davies ME (eds). Academic Press: London, 1994; 225–243.
4. Brömme D, Okamoto K. Human cathepsin O2, a novel cysteine protease highly expressed in osteoclastoma and ovary. Molecular cloning, sequencing and tissue distribution. *Biol. Chem. Hoppe-Seyler* 1995; **376**: 379–384.
5. Drake FH, Dodds RA, James IE, Connor JR, Debouck C, Richardson S, Lee-Rykaczewski E, Coleman L, Rieman D, Barthlow R, Hastings G, Gowen M. Cathepsin K, but not cathepsins B, L, or S, is abundantly expressed in human osteoclasts. *J. Biol. Chem.* 1996; **271**: 12511–12516.
6. Brömme D, Okamoto K, Wang BB, Broc S. Human cathepsin O2, a matrix protein-degrading cysteine protease expressed in osteoclasts. Functional expression of human cathepsin O2 in *Spodoptera frugiperda* and characterization of the enzyme. *J. Biol. Chem.* 1996; **271**: 2126–2132.
7. Garnero P, Borel O, Byrjalsen I, Ferreras M, Drake FH, McQueney MS, Foged NT, Delmas PD, Delaisse JM. The collagenolytic activity of cathepsin K is unique among mammalian proteinases. *J. Biol. Chem.* 1998; **273**: 32347–32352.
8. Schechter I, Berger A. On size of active site in proteases I. Papain. *Biochem. Biophys. Res. Commun.* 1967; **27**: 157–162.
9. Jia Z, Hasnain S, Hirama T, Lee X, Mort JS, To R, Huber CP. Crystal structures of recombinant rat cathepsin B and a cathepsin B-inhibitor complex. *J. Biol. Chem.* 1995; **270**: 5527–5533.
10. Karrer KM, Peiffer SL, DiTomas ME. Two distinct gene subfamilies within the family of cysteine protease genes. *Proc. Natl. Acad. Sci. U.S.A.* 1993; **90**: 3063–3067.
11. 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.
12. Grubb A. Cystatin C – Properties and use as diagnostic marker. *Adv. Clin. Chem.* 2000; **35**: 63–98.
13. Henskens MC, Veerman ECI, Amerongen AVN. Cystatins in health and disease. *Biol. Chem. Hoppe-Seyler* 1996; **377**: 71–86.

14. Leung-Toung R, Li W, Tam TF, Karimian K. Thiol-dependent enzymes and their inhibitors: a review. *Curr. Med. Chem.* 2002; **9**: 979–1002.
15. Björck L, Åkesson P, Bohus M, Trojnar J, Abrahamson M, Olaffson I, Grubb A. Bacterial growth blocked by a synthetic peptide based on the structure of a human proteinase inhibitor. *Nature* 1989; **337**: 385–386.
16. Gante J. Azapeptides. *Synthesis* 1989; 405–413.
17. Wieczerzak E, Drabik P, Lankiewicz L, Oldziej S, Grzonka Z, Abrahamson M, Grubb A, Brömme D. Azapeptides structurally based upon inhibitory sitesof cystatins as potent and selective inhibitors of cysteine proteases. *J. Med. Chem.* 2002; **45**: 4202–4211.
18. Wieczerzak E, Jankowska E, Rodziewicz-Motowidlo S, Gieldon A, Lagiewka J, Grzonka Z, Abrahamson M, Grubb A, Brömme D. Novel azapeptide inhibitors of cathepsins B and K. Structural background to increased specificity for cathepsin B. *J. Pept. Res.* 2006; **66**(Suppl 1): 1–11.
19. Abrahamson M. Cystatins. *Meth. Enzymol.* 1994; **244**: 685–700.
20. Linnevers CJ, McGrath ME, Armstrong R, Mistry FR, Barnes MG, Klaus JL, Palmer JT, Katz BA, Brömme D. Expression of human cathepsin K in *Pichia pastoris* and preliminary crystallographic studies of an inhibitor complex. *Protein Sci.* 1997; **6**: 919–921.
21. Rawlings ND, Barrett AJ. 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.
22. Henderson PJF. 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.
23. Case DA, Pearlman DA, Caldwell JW, Cheatham TE III, Wang J, Ross WS, Simmerling C, Darden T, Merz KM, Stanton RV, Cheng A, Vincent JJ, Crowley M, Tsui V, Gohlke H, Radmer R, Duan J, Pitera J, Massova I, Seibel GL, Singh UC, Weiner PK, Kollman PA. *Amber 7*. University of California: San Francisco, 2002.
24. Tripos Incorporated. *SYBYL 6.6*. St. Louis, MO, 1999.
25. Jorgensen WL, Chandrasekhar J, Madura JD. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 1983; **79**: 926–935.
26. Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* 1984; **81**: 3684–3690.
27. Ryckaert JP, Ciccotti G, Berendsen HJC. Numerical integration of the Cartesian equation of motion of a system with constraints: molecular dynamics of N-alkanes. *J. Comp. Phys.* 1977; **23**: 327–341.
28. *Varian, Nuclear Magnetic Resonance Instruments, VnmrTM Software*, Revision 5.3B 1/1997.
29. Bartles C, Xia T, Billeter M, Günter P, Wütrich K. The program XEASY for the computer-supperted NMR spectral analysis of biological macromolecules. *J. Biomol. NMR* 1995; **1**: 1–10.
30. 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 polypetides by globar conformational analysis and NMR spectroscopy. *J. Biomol. NMR* 1999; **15**: 315–330.
31. Meadows RP, Post CB, Luxon BA, Gorenstein DG. *MORASS 2.1*. Purdue University: West Lafayette, 1994.
32. Karplus M. Contact electron-spin coupling of nuclear magnetic moments. *J. Chem. Phys.* 1959; **30**: 11–15.
33. Magrath J, Abeles RH. Cysteine protease inhibition by azapeptide esters. *J. Med. Chem.* 1992; **35**: 4279–4283.
34. 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.
35. Turk B, Turk D, Turk V. Lysosomal cysteine proteases: more than scavengers. *Biochim. Biophys. Acta* 2000; **1477**: 98–111.
36. Cygler M, Sivaraman J, Grochulski P, Coulombe R, Storer AC, Mort JS. Structure of rat procathepsin B: model for inhibition of cysteine protease activity by the proregion. *Structure* 1996; **4**: 405–416.
37. Czaplewski C, Grzonka Z, Jaskólski 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.

DOI: 10.1002/psc