Acta Biochimica Polonica

Vol. 47 No. 4/2000

1061–1066

QUARTERLY

Communication

Theoretical studies of binding modes of two covalent inhibitors of cysteine proteases★✪

Piotr Drabik[1, 2], Ewa Politowska[1], Cezary Czaplewski[1], Franciszek Kasprzykowski[1], Leszek Łankiewicz[1] and Jerzy Ciarkowski[1]✉

[1]*Faculty of Chemistry, University of Gdańsk, Gdańsk, Poland, and* [2]*Department of Public Health, University School of Physical Education, Gdańsk, Poland*

Received: 13 March, 2000; revised: 24 July, 2000; accepted: 9 November, 2000

Key words: constrained simulated annealing, covalent protease inhibitors, cysteine proteases, papain, molecular dynamics

Physiological and pathological roles of cysteine proteases make them important targets for inhibitor development. Although highly potent inhibitors of this group of enzymes are known, their major drawback is a lack of sufficient specificity. Two cysteine protease covalent inhibitors, viz. (i) Z-RL-deoxo-V-peptide-epoxysuccinyl hybrid, and (ii) Z-RLVG-methyl-, have been developed and modeled in the catalytic pocket of papain, an archetypal thiol protease. A number of configurations have been generated and relaxed for each system using the AMBER force field. The catalytic pockets S_3 and S_4 appear rather elusive in view of the observed inhibitors' flexibility. This suggest rather limited chances for the development of selective structure-based inhibitors of thiol proteases, designed to exploit differences in the structure of catalytic pockets of various members of this family.

The role of cysteine proteases in pathological conditions makes them important targets for the design of new inhibitors to serve as drugs for treatment of a variety of diseases

★Presented at the XV Polish Peptide Symposium, Waplewo, 2–4 September, 1999.

✪Supported by grants No. DS 8245-4-0124-9 (P.D., E.P., C.C., J.C.) and 279/P04/97/13 (F.K., L.Ł.) from the State Committee for Scientific Research (KBN, Poland). The computations were carried out at the Academic Computer Center in Gdańsk TASK, Poland, and the Interdisciplinary Center for Mathematical Modeling (ICM) in Warsaw, Poland.

✉Author for correspondence: Jerzy Ciarkowski, Faculty of Chemistry, University of Gdańsk, J. Sobieskiego 18, 80-952, Gdańsk, Poland, tel. (48 58) 345 0330; fax: (48 58) 341 0357, e-mail: jurek@chemik.chem.univ.gda.pl

Abbreviations: Ac, acetyl; Cbz (also Z), benzyloxycarbonyl; MD, molecular dynamics; PDB, Protein Data Bank.

(Bode & Huber, 1992; Otto & Schirmeister, 1997; Schirmeister, 1999).

Although highly potent inhibitors are known, their major drawback is a lack of sufficient specificity to allow the targeting of a particular thiol protease. The inspection of the reported enzyme–inhibitor structures indicates a wide dispersion of the ligand binding modes (Drenth *et al.*, 1976; Varughese *et al.*, 1989; Stubbs *et al.*, 1990; Yamamoto *et al.*, 1991; 1992; Kim *et al.*, 1992; Schroder *et al.*, 1993). This, in turn, strongly suggests the absence of a general principle governing binding of the inhibitor. In order to analyze in detail the binding of a specific inhibitor, we have modeled two ligands, developed in our laboratory (Hall *et al.*, 1992), in the active site of papain. Thus, the aim of the study was to evaluate the mobility of two new inhibitors of thiol proteases in the catalytic pocket of papain.

Both subjects of the present study: (i) Z-RL-deoxo-V-peptide-epoxysuccinyl hybrid (further referred to as InhA), and (ii) Z-RLVG-methyl- (called InhB), were developed on the assumption that a reactive group responsible for the covalent linkage (i.e., epoxysuccinyl or diazomethyl) should be combined with a peptide fragment, modeled after the cystatin C N-terminal sequence (residues R^8LVG^{11}) which determines structural features responsible for tight non-covalent binding of cystatins to thiol proteases (Hall *et al.*, 1992; Czaplewski *et al.*, 1999).

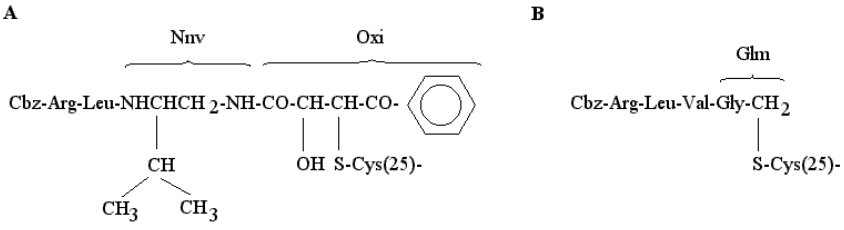

Figure 1. The structure of the cysteine protease inhibitors: (A) InhA and (B) InhB.

COMPUTATIONAL METHODS

All simulations were carried out using the AMBER 5.0 program (Case *et al.*, 1997). The all-atom AMBER force field was used (Cornell *et al.*, 1995).

Starting coordinates for all heavy atoms of papain were obtained from the crystal structures from aqueous solution, Protein Data Bank, PDB (Bernstein *et al.*, 1977), files 1pe6 and 1pad for InhA and InhB, respectively. The topology and definitions of particular residues of both protein–ligand complexes are given in Fig. 1. The design of the starting model and development of specific constraints for the papain–InhA complex (not having the Cbz residue) are described in detail elsewhere (Czaplewski *et al.*, 1999). The initial model for the papain–InhB complex was designed in Sybyl 6.1 program (1994, Tripos, Inc.) on the basis of point-substitutive mutations of Acaapack (Ac-AAPA-CH_2) complex (1pad), and a set of specific constraints for this starting model was set (Table 1) to maintain the papain–inhibitor interactions (Drenth *et al.*, 1976). New residues, absent in the original AMBER force field, have been parameterized according to standard procedures (Bayly *et al.*, 1993).

The details of modeling procedure have been described in our previous article (Czaplewski *et al.*, 1999). Briefly, a sequence of C^{α}–positional and "softly" (as indicated in Table 1) constrained minimization and simulated annealing *in vacuo*, was followed by a series of C^{α}–positional constrained minimizations *in vacuo* and in water, and subsequent unconstrained minimization in water. A consecutive unconstrained thermalization preceded a 230 ps molecular dynamics (MD) runs of the complexes in water, under periodic boundary conditions in a closed, isothermal, isobaric (NTP) ensemble. Throughout the simulation the sol-

ute and solvent were coupled to a constant-temperature (T = 300 K) heat bath and a constant-pressure (P = 1 bar) bath (Berendsen *et al.*, 1984). Counterions (10 Cl^-) to neutralize the system were added using AMBER 5.0 was thermalized and subjected to 230 ps of unconstrained molecular dynamics at 300 K.

Time-averaged residue-based deviations as a function of residue number for all molecular dynamics runs are shown in Fig. 2. Changes

Table 1. The two sets of "soft" constraints for the papain–InhB starting complex.

A specific atom pair was assigned to one or the other constraint category, in accordance with the initial distance linking the constrained atoms. The harmonic force-constant in both sets of constraints equaled 100 kcal/mol.

First atom	Second atom	Initial distance [Å]	Penalty function characteristics
Leu 215 (CG)	Tyr 61 (CG)	5.162	0–3.5/8.5–∞ Å steeply dropping/raising linear;
Leu 215 (CG)	Tyr 67 (CG)	5.190	
Val 216 (CB)	Val 133 (CB)	5.290	3.5–4.5/7.5–8.5 Å dropping/raising harmonic; 4.5–7.5 Å free to vary
Val 216 (CB)	Val 157 (CB)	7.081	
Val 216 (N)	Gly 66 (O)	2.792	
Val 216 (O)	Gly 66 (N)	2.794	0–2.0/3.2–∞ Å steeply dropping/raising linear;
Glm 217 (N)	Asp 158 (O)	2.937	
Glm 217 (O)	Cys 25 (N)	3.130	2.0–2.5/3.0–3.2 Å dropping/raising harmonic; 2.5–3.0 Å free to vary
Glm 217 (O)	Gln 19 (NE)	3.165	

program LEAP. All bonds were constrained using the SHAKE algorithm (Ryckaert *et al.*, 1977) allowing a time step of 1 fs. A double residue-based cutoff distance of 10/14 Å was used for nonbonded interactions. The TIP3P model was used for water molecules (Jorgensen *et al.*, 1983). A typical box size was 78 × 66 × 60 Å. Approximately 8200 TIP3P water molecules were in the box, i.e. the whole system consisted of a total of 28 000 atoms.

RESULTS AND DISCUSSION

The initial structures of the papain–ligand complexes were subjected to the constrained simulated annealing so as to perform the simulation at a very high, physically unrealistic temperature. The additional kinetic energy enhances the ability of the system to explore the energy surface and can prevent the molecule from getting stuck in a localized region of conformational space. Afterwards the system

up to 5 Å were observed for some residues (C-terminus), but the overall C^{α} mobilities figured approximately 1 (for InhA) and 2 (for InhB) Å. The average structures displayed similarly significant changes in some flexible loops of the enzyme. It should be stressed that the protein structures reproduced very well the mobility pattern over the whole molecule, represented experimentally by the atomic temperature factors in the crystal structure of the papain–E-64c complex (Kim *et al.*, 1992). This result validated the use of the AMBER 5.0 force field as a suitable tool for scanning the conformational space of both ligands nested in the catalytic cleft of papain.

The detailed atomic-level analysis of mobilities of the inhibitors' backbones reveals that the dispersion of the relaxed positions of residues increases steeply towards the N-terminus of the inhibitors (Fig. 3). The catalytic pocket S_3, as defined by the pioneering studies of Schechter & Berger (1967), appears rather elusive in view of the evident inhibi-

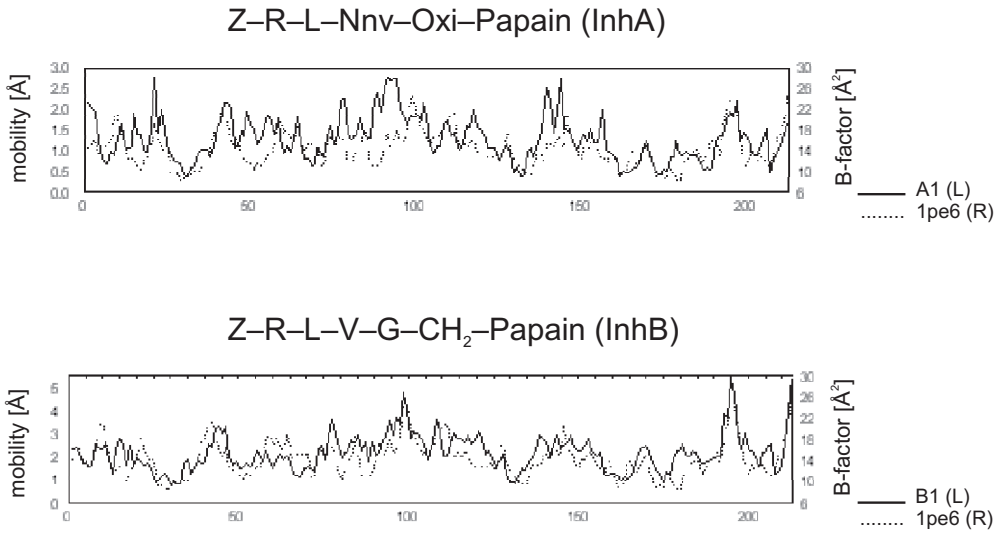

Figure 2. The time-averaged C^{α}-based deviations (mobilities) along the papain sequence accompanying MD runs, solid line.

Symbols A1 and B1 refer to InhA and InhB MD runs, respectively. The distribution of the C^{α} temperature factors (as given in the Protein Data Bank file 1pe6), being an independent experimental measure of the papain-chain mobility, is shown as a dotted line. A correlation between the two plots is evident. It can be seen that some loops of the papain L lobe undergo high fluctuations during the molecular dynamics runs.

tors' flexibility observed in the present study and in the other experimentally solved papain-inhibitor complexes (Drenth *et al.*, 1976; Varu-

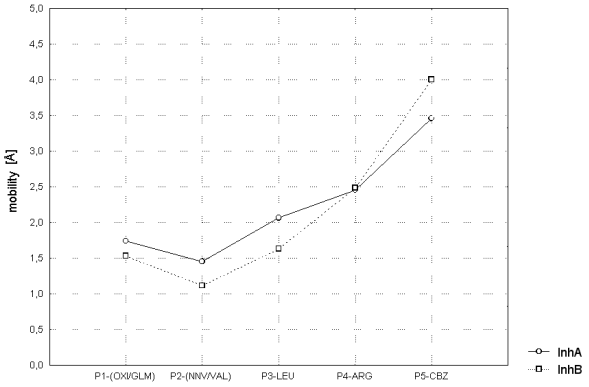

Figure 3. The time-averaged C^{α}-based (except the Cbz and Oxi, residues, for which the benzyloxy oxygen and sulfur, respectively, serve as probes) mobility of inhibitors' individual residues during the molecular dynamics simulations.

ghese *et al.*, 1989; Stubbs *et al.*, 1990; Yamamoto *et al.*, 1991; Kim *et al.*, 1992; Schroder *et al.*, 1993). The location and definition of substrate binding site S_4 is even more questionable. These features are also reflected in the two selected binding modes of papain–InhA and papain–InhB, given in Fig. 4.

The lack of a single site in space that would be clearly defined for binding of inhibitor's main chain may indicate that the surface of the papain–inhibitor interaction spreads over a relatively large area. Keeping in mind the clear diversity in the architecture, accommodated by covalent inhibitors in experimental structures, and the fact that the exact location of the residues and the subsites more distant to the catalytic cysteine cannot be clearly identified (Turk *et al.*, 1998), it should be emphasized that the localization and definition of the substrate binding sites S_3 and S_4 is precarious. As the substrate binding sites cannot be precisely localized, the term "binding site" may be somewhat misleading.

These findings prompt us to suggest that the considerably weak definition of catalytic pockets S_3 and S_4 of cysteine proteases predicts rather limited chances for the development of selective structure-based inhibitors of thiol proteases, designed to exploit tiny differences between the catalytic pockets of various members of this family.

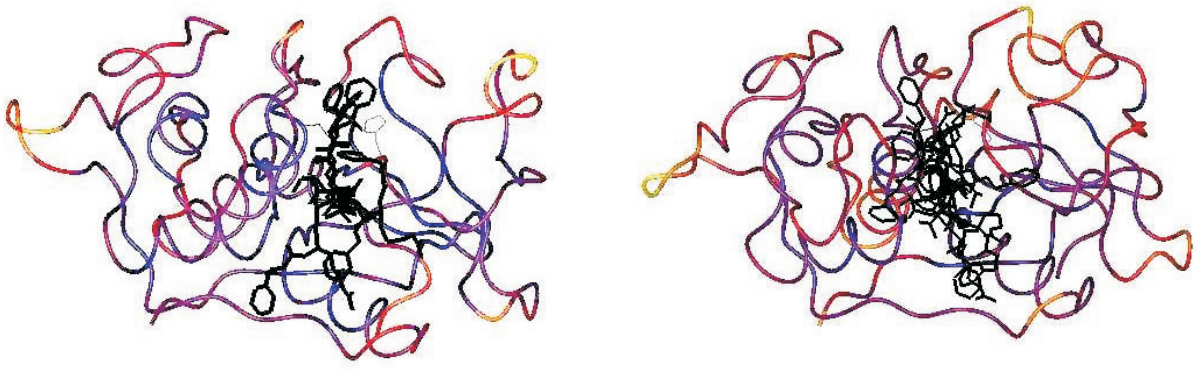

Figure 4. The selected binding modes of relaxed InhA and InhB configurations (black lines), resulting from random sampling of MD productive runs.

Blue and yellow indicate low and high atomic mobility limits, respectively. The catalytic dyad (C25 in blue and H159 in red) is also shown. The obvious scatter of the inhibitors covering a considerable part of space can be easily seen.

REFERENCES

Bayly, C.I., Cieplak, P., Cornell, W.D. & Kollman, P.A. (1993) *J. Phys. Chem.* **97**, 10269–10280.

Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W.F., DiNola, A. & Haak, J.R. (1984) *J. Chem. Phys.* **81**, 3684–3690.

Bernstein, F.C., Koetzle, T.F., Williams, G.J.B., Meyer, E.F., Jr., Brice, M.D., Rodgers, J.R., Kennard, O., Shimanouchi, T. & Tasumi, M. (1977) *J. Mol. Biol.* **112**, 535–542.

Bode, W. & Huber, R. (1992) *Eur. J. Biochem.* **204**, 433–451.

Case, D.A., Pearlman, D.A., Caldwell, J.W., Cheatham III, T.E., 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. (1997) *Amber 5.0*, University of California, San Francisco.

Cornell, W.D., Cieplak, P., Bayly, C.I., Gould, I.R., Merz, K.M., Jr., Ferguson, D.M., Spellmeyer, D.C., Fox, T., Caldwell, J.W. & Kollman, P.A. (1995) *J. Am. Chem. Soc.* **117**, 5179–5197.

Czaplewski, C., Grzonka, Z., Jaskólski, M., Kasprzykowski, F., Kozak, M., Politowska, E. & Ciarkowski, J. (1999) *Biochim. Biophys. Acta* **1431**, 290–305.

Drenth, J., Kalk, K.H. & Swen, H.M. (1976) *Biochemistry* **15**, 3731–3738.

Hall, A., Abrahamson, M., Grubb, A., Trojnar, J., Kania, P., Kasprzykowska, R., Kasprzykowski, F. (1992) *J. Enzyme Inhibition* **6**, 113–123.

Jorgensen, W.L., Chandreskhar, J., Madura, J.D., Impey, R.W. & Klein, M.L. (1983) *J. Chem. Phys.* **79**, 926–935.

Kim, M.-J., Yamamoto, D., Matsumoto, K., Inoue, M., Ishida, T., Mizuno, H., Sumiya, S. & Kitamura, K. (1992) *Biochem. J.* **287**, 797–803.

Otto, H.-H. & Schirmeister, T. (1997) *Chem. Rev.* **97**, 133–171.

Ryckaert, J.P., Ciccotti, G. & Berendsen, H.J.C. (1977) *J. Comput. Phys.* **23**, 327–341.

Schechter, I. & Berger, A. (1967) *Biochem. Biophys. Res. Commun.* **27**, 157–162.

Schirmeister, T. (1999) *Biopolymers* **51**, 87–97.

Schroder, E., Phillips, C., Garman, E., Harlos, K. & Crawford, C. (1993) *FEBS Lett.* **315**, 38–42.

Stubbs, M.T., Laber, B., Bode, W., Huber, R., Jerala, R., Lenarcic, B. & Turk, V. (1990) *EMBO J.* **9**, 1939–1947.

SYBYL 6.1. (1994) Tripos Inc., 1699 S. Manley Rd., St. Louis, MO 63144, U.S.A.

Turk, D., Guncar, G., Podobnik, M. & Turk, B. (1998) *Biol. Chem.* **379**, 137–147.

Varughese, K.I., Ahmed, F.R., Carey, P.R., Hasnain, S., Huber, C.P. & Storer, A.C. (1989) *Biochemistry* **28**, 1330–1332.

Yamamoto, D., Matsumoto, K., Ohishi, H., Ishida, T., Inoue, M., Kitamura, K. & Mizuno, H. (1991) *J. Biol. Chem.* **266**, 14771–14777.

Yamamoto, A., Tomoo, K., Doi, M., Ohishi, H., Inoue, M., Ishida, T., Yamamoto, D., Tsubo, I.S., Okamoto, H. & Okada, Y. (1992) *Biochemistry* **31**, 11305–11309.