Communication

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 268, No. 21, Issue of July 25, pp. 15343–15346, 1993

Structure of a Non-peptide Inhibitor Complexed with HIV-1 Protease

DEVELOPING A CYCLE OF STRUCTURE-BASED DRUG DESIGN*

(Received for publication, May 13, 1993)

Earl Rutenber‡, Eric B. Fauman§, Robert J. Keenan, Susan Fong, Paul S. Furth, Paul R. Ortiz de Montellano, Elaine Meng, Irwin D. Kuntz, Dianne L. DeCamp¶, Rafael Salto‖, Jason R. Rosé, Charles S. Craik, and Robert M. Stroud**

From the Department of Biochemistry and Biophysics and Department of Pharmaceutical Chemistry, University of California, San Francisco, California 94143-0448

A stable, non-peptide inhibitor of the protease from type 1 human immunodeficiency virus has been developed, and the stereochemistry of binding defined through crystallographic three-dimensional structure determination. The initial compound, haloperidol, was discovered through computational screening of the Cambridge Structural Database using a shape complementarity algorithm. The subsequent modification is a non-peptidic lateral lead, which belongs to a family of compounds with well characterized pharmacological properties. This thioketal derivative of haloperidol and a halide counterion are bound within the enzyme active site in a mode distinct from that observed for peptide-based inhibitors. A variant of the protease cocrystallized with this inhibitor shows binding in the manner predicted during the initial computer-based search. The structures provide the context for subsequent synthetic modifications of the inhibitor.

Inhibition of human immunodeficiency virus (HIV-1)[1] protease leads to loss of viable viral particles in infected cells (1). The x-ray crystal structure of HIV-1 protease (2–9) permits structure-based searches for specific inhibitors (10, 11). The dimeric enzyme presents two aspartic acid side chains, one from each monomer, to form a single catalytic site. Two symmetrically disposed β ribbon "flaps" surround the catalytic center and close around peptide-based inhibitors.

The instability of peptidomimetic inhibitors *in vivo*, the inability to deliver them through oral routes, and their poor bioavailability for the virus (12) place a high premium on development of organic compounds as leads that avoid these problems. Over 10 million chemical compounds have been characterized (13), and the crystallographically determined three-dimensional structures of a subset are available in the Cambridge Structural Database (CSD) (14, 15). Other structures can be predicted using software such as CONCORD (16). Effective computational screening of these structures would greatly assist in drug discovery.

A computational search of the CSD was based on shape and to some extent on possible chemical complementarity with the active site of HIV-1 protease (17–19). Approximately 50 of the top scoring 200 compounds were commercially available, and 15 were assayed for inhibition. Of these, five were inhibitory in the range of 10 μM to 1 mM and one, haloperidol (UCSF1) (Fig. 1*A*), was selected for subsequent improvement. UCSF1 is inhibitory with an apparent K_i (inhibition constant) = 100 μM for HIV-1 protease. Replacement of the ketone functional center with a thioketal ring (1,3-dithiolane) resulted in UCSF8 (Fig. 1*B*). UCSF8 is a competitive inhibitor with an apparent K_i of 15 ± 2 μM (Fig. 1*C*) for HIV-1 protease, and the inhibition constant is highly dependent upon the concentration of NaCl, NaBr, NaI, and NaF. The K_i of UCSF8 for HIV-2 protease is 100 μM. Selectivity of UCSF8 for HIV-1 protease over pepsin is over 40-fold, whereas it is approximately 100-fold or greater for renin and cathepsin D based on the ratios of IC_{50} values. The complex of this improved inhibitor with HIV-1 protease was crystallized and the structure solved to 2.2-Å resolution (Table I). The same inhibitor was also cocrystallized with a variant HIV-1 protease designed to be more proteolytically stable and the structure solved to 1.9-Å resolution. Cocrystallization of UCSF8 and HIV-1 protease gave tetragonal bipyramidal crystals in space group $P4_1$, whereas cocrystallization of UCSF8 and HIV-1 protease Q7K gave orthorhombic plates in space group $P2_12_12_1$. Both structures were solved by isomorphous molecular replacement using a partial model of the dimer generated from the chemically synthesized unliganded enzyme structure, 3HVP (6). Together, these structures reveal distinct binding modes of this inhibitor within the active site cavity and define specific strategies for future inhibitor design.

The largest difference between the UCSF8-HIV-1 protease structure and previously reported structures of HIV-1 protease is the unique conformation of the active site flaps (Fig. 2*C*). The conformation of the flaps is intermediate between that in the unliganded enzyme and that of the peptide analog-bound structures, as demonstrated by principal axes analysis (20).

A chloride anion is coordinated between the bound inhibitor and the almost symmetric amide nitrogens of Ile-50 and Ile-150 in the flaps that surround the active site (Fig. 2*C*). The distances between the chloride position and the coordinating heavy atoms are similar to the interatomic distance (~3.2 Å) observed between a chloride and a charged or uncharged nitrogen involved in an electrostatic interaction in crystallographically determined small molecule structures (21, 22).

UCSF8 is clearly defined in continuous density in a complex-minus-protein difference map and lies in relatively extended conformation. The two 2-fold related and partially overlapping binding modes for the inhibitor are each about 50% occupied (Fig. 2*A*). Thus, the inhibitor density was interpreted and refined as two overlapping structures constrained by a non-crystallographic 2-fold axis. After refinement there was little resid-

* This research was supported in part by National Institutes of Health Grant GM 39552. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ American Cancer Society postdoctoral fellow.
§ Howard Hughes predoctoral fellow.
¶ Recipient of a National Institutes of Health Postdoctoral Fellowship.
‖ Fulbright/Ministry of Education and Science postdoctoral fellow.
** To whom correspondence should be addressed. Tel.: 415-476-4224; Fax: 415-476-1902.

[1] The abbreviations used are: HIV-1, human immunodeficiency virus; CSD, Cambridge Structural Database.

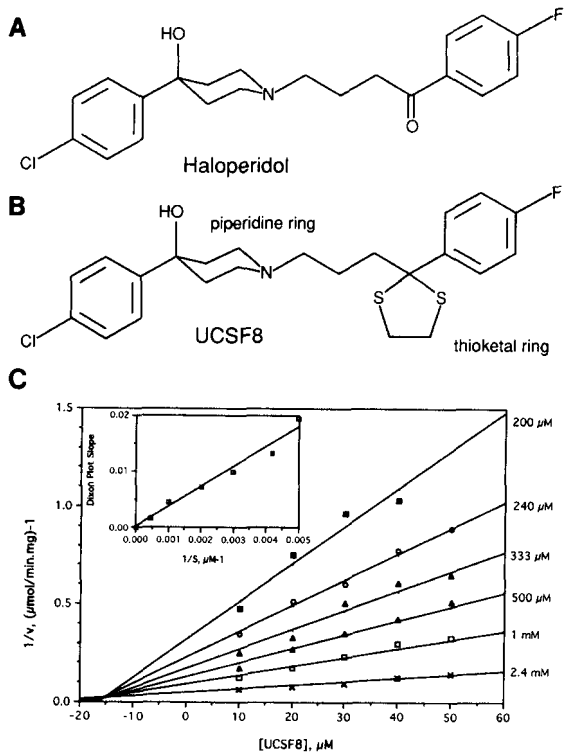

FIG. 1. *A*, the chemical structure of haloperidol (UCSF1) used in the initial assays. *B*, the chemical structure of UCSF8, an inhibitor of HIV-1 protease with a K_i = 15 µM. *C*, Dixon plot of inhibition of HIV-1 protease by UCSF8. A plot (*inset*) of the slopes from the Dixon plot *versus* the reciprocal of substrate concentration indicates that UCSF8 acts as a competitive inhibitor. UCSF8 was obtained by reaction of haloperidol with 1,2-ethanedithiol with a 99% yield. Purification was accomplished by filtration through a silica gel column. The resulting crystalline product had a melting point of 101.0–102.5 °C. Elemental analysis and ^{1}H NMR, infrared, and mass spectra are consistent with the assigned structure. Purified HIV-1 protease (1.93 × 10^{-3} mg/ml) was incubated with UCSF8 in 50 mM sodium acetate buffer, pH 5.5, containing 1 mM dithiothreitol, 1 mM EDTA, 1 M NaCl, and 5% Me_2SO. After 1 min, the substrate (*Peptide I*) was added to give the final substrate concentrations shown (*abscissa*). The assay solutions were incubated for 5 min at 37 °C, and enzyme activity was determined by quantifying the hydrolysis products on high performance liquid chromatography.

ual difference density at the inhibitor sites, indicating that the single binding mode, averaged by statistical occupation, accounts for all the density in the map. The higher thermal *B* factors for atoms of UCSF8 ($\langle B \rangle$ = 68.6 Å^2) *versus* $\langle B \rangle$ = 30.3 Å^2 for all atoms in the protein, indicate that the inhibitor lies in a shallow energy minimum with root mean square amplitude of vibration, u = 0.9Å, but it is not bound in other modes. In sum, the amino acid residues around the inhibitor binding site lie in well ordered density; the inhibitor lies in clear but less resolved density.

The UCSF8 binding site is on the opposite surface of the active site relative to the peptidomimetic inhibitor (JG365) (8) (Fig. 2*A*). The five-membered, two-sulfur-containing thioketal ring extends into the P2 substrate binding pocket of the enzyme (23) and represents the only overlap with the volume of the active site occupied by peptidomimetic inhibitors (Fig. 2*A*).

TABLE I
Crystallographic statistics of structure solution

Hanging drops were formed by mixing 3 µl of solution containing 5.0 mg/ml HIV-1 PR, 1.0 mM DTT, 1.0 mM EDTA, and a saturating concentration of UCSF8 (~400 mM) with 3 µl of buffer A (1.0 M NaCl, 20.0 mM sodium acetate, pH 5.4). The drops were equilibrated against buffer A at 23 °C. Crystals up to 40 µm × 200 µm in dimension generally grew in 3–5 days. Diffraction intensities were recorded on a multiwire detector (Siemens X-100) using Cu Kα x-rays from a rotating anode generator (Rigaku Ru300) with a graphite monochromator. Intensities were integrated using BUDDHA (28), and frame to frame scaling was carried out using XENGEN (29). The structures were solved by isomorphous molecular replacement using a partial model of the dimer generated from the chemically synthesized unliganded enzyme structure, 3HVP (6). Coordinates will be deposited in the Brookhaven Protein Data Bank.

	HIV-1 PR	HIV-1 PR Q7K
Crystallographic data		
Space group	$P4_1$	$P2_12_12_1$
Unit cell *a*, *b*, *c* (Å)	50.2, 50.2, 100.4	51.5, 60.4, 62.6
Resolution (Å)	∞ – 2.2	∞ – 1.9
$N_{observations}$	29,425	26,951
N_{unique}	9,988	12,260
R_{sym} (*I*) (%)[a]	6.3	7.5
Completeness (%)	93	77
Refinement data		
Resolution (Å)	7.0–2.2	7.0–1.9
R_{cryst} of refined structure (%)[b]	17.4	17.9
RMS deviations from ideality in final model		
Bond lengths (Å)	0.011	0.016
Bond angles	2.93°	3.24°
Dihedral angles	27.0°	28.0°
Improper angles[c]	1.34°	1.41°

[a] $R_{sym} = \{\sum_{hkl}\sum_{i=1}^{N}(I_{avg} - I_i)^2/\sum_{hkl}\sum_{i=1}^{N}(I_i)^2\}^{1/2}$
[b] $R_{cryst} = \sum|(|F_o| - |F_c|)|/\sum|F_o|$
[c] Improper angles define chiral centers and planar groups of atoms.

The mutant protein, HIV-1 protease Q7K, has a similar inhibition profile, similar kinetic parameters, and a more than 100-fold greater stability toward proteolytic degradation than the wild-type enzyme (24). The structure of UCSF8 in complex with HIV-1 protease Q7K reveals two molecules of UCSF8 bound in the active site of the enzyme (Fig. 2*B*), each with a chloride ion ligated by the piperidine nitrogen. The orientation of UCSF8 in the primary site is similar to that identified by the DOCK algorithm (Fig. 3). The active site flaps are closed as in peptide-based inhibitor complexes of HIV-1 protease. The crystallographic observation of these various binding modes suggests that they are all of similar energy. These independent sites will be exploited by iterative structure-based drug development.

UCSF8 binds to HIV-1 protease 4.8 Å away and rotated by 79° from the orientation predicted in the computational search. Furthermore, the bound conformation differs from the CSD structure used to make the prediction (17) (Fig. 2*A*). This result can be accounted for in part by the simplifying assumptions used in the original shape complementarity algorithm. 1) Molecular water and counterions were not considered; 2) only steric interactions were considered; 3) a fixed crystallographically determined conformation of a chemically different compound was used; 4) a fixed conformation of the protein, that of the unliganded protease, was assumed.

FIG. 2. *A*, omit map for UCSF8 in complex with native HIV-1 protease shown with the best scoring orientation of bromperidol. The two 2-fold related configurations of UCSF8 (*black* and *gray*) are shown with the difference electron density for UCSF8 at 2.2-Å resolution. The map, shown at a 2.8-σ contour level, was calculated by setting the occupancy of the inhibitor to 0.0 and performing positional refinement on the entire structure. The best scoring DOCK binding mode for bromperidol (*dashed lines*) is also shown for comparison. The catalytic aspartic acid residues (*gray*) are shown for reference. Key residues are labeled. The two 2-fold related positions for the counterion are shown with *dashed lines* indicating electrostatic interactions. *B*, omit map for UCSF8 in complex with HIV-1 protease Q7K as viewed through the flaps. The two binding sites of UCSF8 (*black*) are shown with the difference electron density for UCSF8 at 1.9-Å resolution. The map is contoured at 3.0 σ and was calculated as described above. *C*, schematic overview of interactions in the active site of HIV-1 protease as occupied by UCSF8. A chloride anion coordinates the piperidine nitrogen of UCSF8 to the amide nitrogens of residues Ile-50 and Ile-150. A tightly bound water molecule, Wat-301, remains present between Asp-25 and Asp-125, as observed in the unliganded HIV-1 protease structures.

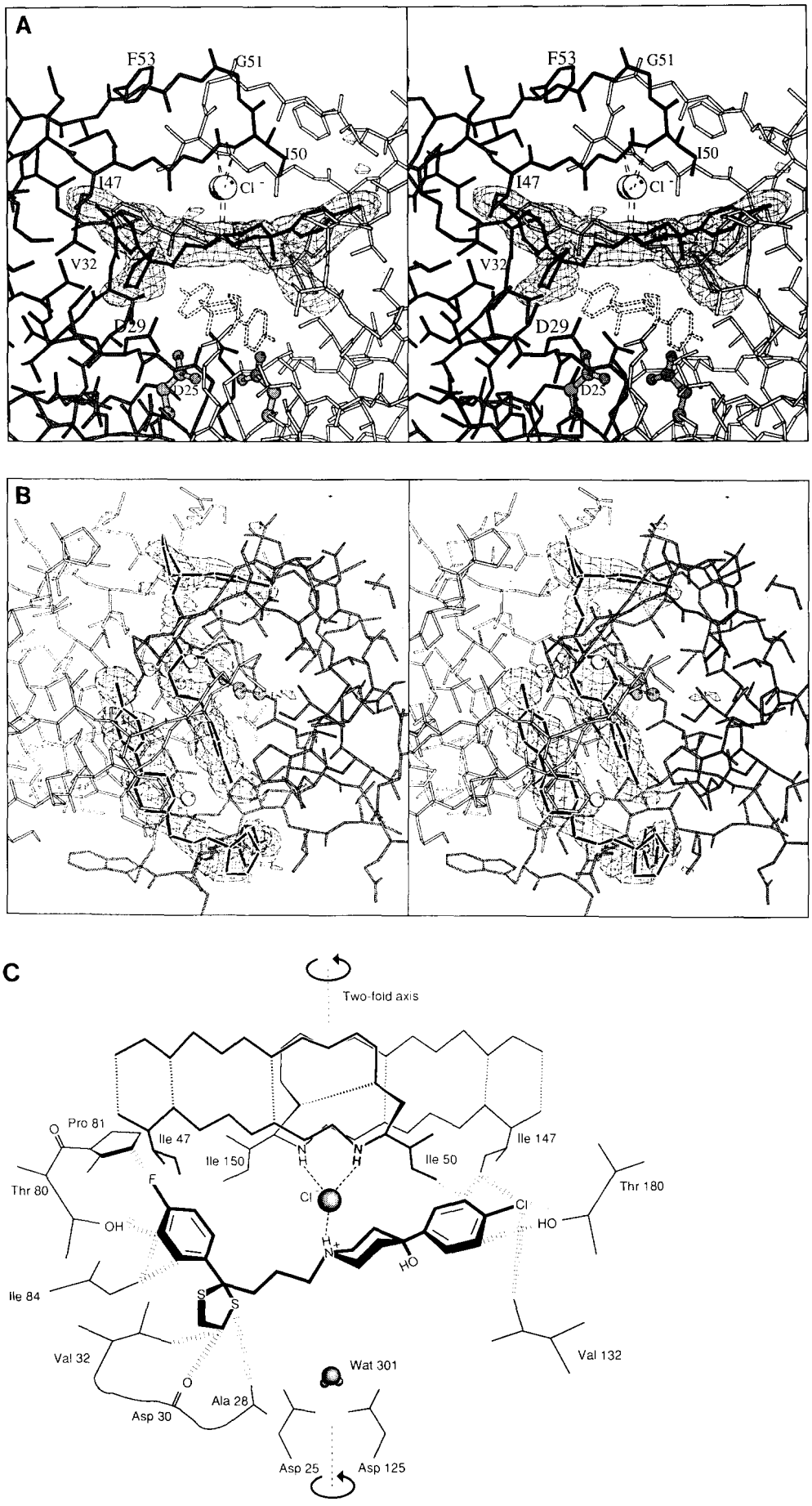

FIG. 2

FIG. 3. **The binding mode of the two UCSF8 molecules (*white, gray*) in complex with HIV-1 protease Q7K as viewed through the flaps.** The primary site is adjacent to the catalytic aspartic acid residues and overlaps the P1 and P1′ binding sites observed for peptide-based inhibitors. The secondary site is further from the active center, extending from the P2 binding site toward solvent. The active site flaps are in the closed position as observed in peptide-based inhibitor complexes of HIV-1 protease. The best scoring DOCK binding mode for bromperidol (*black*) overlaps the experimentally determined primary binding site of UCSF8.

However, using the crystallographically determined conformations of UCSF8 and HIV-1 protease, the correct binding mode is identified as the global minimum without inclusion of the chloride ion or explicit water, suggesting that the first assumption is warranted. Clearly, evaluating alternative orientations and conformations among a diverse set of compounds and implementing energetic considerations is essential for more accurate computational analysis (25). One possible alternative to the use of a fixed protein conformation is to use multiple crystallographically determined structures of the protein as targets to sample more of the conformational space available to the protein, and softer boundaries where the B factors are large (26, 27). Despite these limitations, this structure-based approach aided in developing a unique inhibitor with a K_i = 15 μM and offers an important alternative to the often rate-limiting step in drug development. The structures of the UCSF8-protease complexes offer the opportunity to follow rational routes to iterative development of a more potent, stable, and bioavailable drug.

Acknowledgments—We thank P. Fitzgerald for helpful discussions, P. Ghosh for assistance with synchrotron data, A. Wlodawer for provision of the 7HVP coordinates, and the UCSF Spectroscopy Facility for IR and mass spectral analyses. Figs. 2*A*, 2*B*, and 3 were generated using MolScript (30).

REFERENCES

1. McQuade, T. J., Tomasselli, A. G., Liu, L., Karacostas, V., Moss, B., Sawyer, T. K., Heinrikson, R. L., and Tarpley, W. G. (1990) *Science* **247,** 454–456
2. Navia, M. A., Fitzgerald, P. M., McKeever, B. M., Leu, C. T., Heimbach, J. C., Herber, W. K., Sigal, I. S., Darke, P. L., and Springer, J. P. (1989) *Nature* **337,** 615–620
3. Wlodawer, A., Miller, M., Jaskolski, M., Sathyanarayana, B. K., Baldwin, E., Weber, I. T., Selk, L. M., Clawson, L., Schneider, J., and Kent, S. B. H. (1989) *Science* **245,** 616–621
4. Lapatto, R., Blundell, T., Hemmings, A., Overington, J., Wilderspin, A., Wood, S., Merson, J. R., Whittle, P. J., Danley, D. E., and Geoghegan, K. F. (1989) *Nature* **342,** 299–302
5. Jaskolski, M., Tomasselli, A. G., Sawyer, T. K., Staples, D. G., Heinrikson, R. L., Schneider, J., Kent, S. B., and Wlodawer, A. (1991) *Biochemistry* **30,** 1600–1609
6. Miller, M., Schneider, J., Sathyanarayana, B. K., Toth, M. V., Marshall, G. R., Clawson, L., Selk, L., Kent, S. B. H., and Wlodawer, A. (1989) *Science* **246,** 1149–1152
7. Fitzgerald, P. M. D., McKeever, B. M., VanMiddlesworth, J. F., Spinger, J. P., Heimbach, J. C., Lue, C. T., Herber, W. K., Dixon, R. A., and Darke, P. L. (1990) *J. Biol. Chem.* **265,** 14209–14219
8. Swain, A. L., Miller, M. M., Grenn, J., Rich, D. H., Schneider, J., Kent, S. B., and Wlodawer, A. (1990) *Proc. Natl. Acad. Sci. U. S. A.* **87,** 8805–8809
9. Blundell, T. L., Lapatto, R., Wilderspin, A. F., Hemmings, A. M., Hobart, P. M., Danley, D. E., and Whittle, P. J. (1990) *Trends Biochem. Sci.* **15,** 425–430
10. Goodford, P. (1985) *J. Med. Chem.* **28,** 849–857
11. Perutz, M. (1992) *Protein Structure: New Approaches to Disease and Therapy*, W. H. Freeman & Co., New York
12. Verhoef, J. C., Bodde, H. E., de Boer, A. G., Bouwstra, J. A., Junginger, H. E., Merkus, F. W., and Breimer, D. D. (1990) *Eur. J. Drug Metab. Pharmacokinet.* **15,** 83–93
13. Fisanick, W. (1990) *J. Chem. Inf. Comput. Sci.* **30,** 145–148
14. Allen, F. H., Kennard, O., Motherwell, W. D. S., Town, W. G., and Watson, D. G. (1973) *J. Chem. Doc.* **13,** 119–123
15. Allen, F. H., Bellard, S., Brice, M. D., Cartwright, B. A., Doubleday, A., Higgs, H., Hummelink, T., Hummelink-Peters, B. G., Kennard, O., Motherwell, W. D. S., Rodgers, J. R., and Watson, D. G. (1979) *Acta Crystallogr. Sect. B Struct. Sci.* **35,** 2331–2339
16. Rusinko, A. (1989) *J. Chem. Inf. Comput. Sci.* **29,** 251–264
17. DesJarlais, R. L., Seibel, G. L. Kuntz, I. D., Furth, P. S., Alvarez, J. C., Ortiz de Montellano, P. R., DeCamp, D. L., Babé, L. M., and Craik, C. S. (1990) *Proc. Natl. Acad. Sci. U. S. A.* **87,** 6644–6648
18. Kuntz, I. D., Blaney, J. M., Oatley, S. J., Langridge, R., and Ferrin, T. E. (1982) *J. Mol. Biol.* **161,** 269–288
19. DesJarlais, R., Sheridan, R. P., Seibel, G. L., Dixon, J. S., Kuntz, I. D., and Venkataraghavan, R. (1988) *J. Med. Chem.* **31,** 722–729
20. Browner, M. F., Fauman, E., and Fletterick, R. J. (1992) *Biochemistry* **31,** 11297–11304
21. Furguson, G., Kaitner, B., Connett, B. E., and Rendle, D. F. (1991) *Acta Crystallogr. Sect. B Struct. Sci.* **47,** 479–482
22. Datta, N., Mondal, P., and Pauling, P. (1979) *Acta Crystallogr. Sect. B Struct. Sci.* **35,** 1486–1490
23. Schechter, I., and Berger, A. (1967) *Biochem. Biophys. Res. Commun.* **27,** 157
24. Rosé, J. R., Salto, R., and Craik, C. S. (1993) *J. Biol. Chem.* **268,** 11939–11945
25. Meng, E. C., Shoichet, B. K., and Kuntz, I. D. (1992) *J. Comput. Chem.* **13,** 505–524
26. Perry, K. M., Fauman, E. B., Finer-Moore, J. S., Montfort, W. R., Maley, G. F., Maley, F., and Stroud, R. M. (1990) *Proteins* **8,** 315–333
27. Tilton, R. F., Jr., Singh, U. C., Weiner, S. J., Connolly, M. L., Kuntz, I. D., Kollman, P. A., Max, N., and Case, D. A. (1986) *J. Mol. Biol.* **192,** 443–449
28. Blum, M., Metcalf, P., Harrison, S. C., and Wiley, D. C. (1987) *J. Appl. Crystallogr.* **20,** 235–239
29. Howard, A. J., Neilsen, C., and Xuong, Ng. (1985) *Methods Enzymol.* **114,** 452–465
30. Kraulis, P. J. (1991) *J. Appl. Crystallogr.* **24,** 946–950