Mass-weighted molecular dynamics simulation of the protein-ligand complex of rhizopuspepsin and inhibitor

Boryeu Mao
Upjohn Research Laboratories, Kalamazoo, Michigan 49001 USA

ABSTRACT The mass-weighted molecular dynamics simulation method was developed previously for sampling the multidimensional conformational space of linear and cyclic polypeptides and studying their conformational flexibility. Herein results from molecular dynamics simulations of the protein-ligand complex of the aspartyl protease rhizopuspepsin and a polypeptide inhibitor are reported. The dihedral conformational space sampling for the linear peptide inhibitor in situ was found to be increased in the mass-weighted simulation as in other molecular systems previously studied. More significantly, the physical space of the enzyme binding pocket was also sampled efficiently in the simulations and multiple binding sites were identified for the inhibitor. These results suggest that it may be possible now to study, by computer simulations, the putative initial enzyme-inhibitor complex suggested experimentally from the time-dependent kinetics of enzyme inhibition by slow-binding inhibitors (Morrison, J. F., and C. T. Walsh. 1988. *Adv. Enzymol.* 61:201), and/or conformational substates in protein-ligand complexes suggested in the study of reassociation dynamics of myoglobin and carbon monoxide following photolysis (Austin, R. H., K. W. Beeson, L. Eisenstein, H. Frauenfelder, and I. C. Gunsalus. 1975. *Biochemistry.* 14:5355). Moreover, the intermediate binding steps and the molecular flexibility of the inhibitor shown in the MWMD simulation may have crucial roles in the ligand binding process.

INTRODUCTION

Proteins are intrinsically flexible molecules (Debrunner and Frauenfelder, 1982; Wagner and Wüthrich, 1986). The large number of internal degrees of freedom due to the large number of constituent atoms in proteins also results in complex potential energy surfaces and complex atomic motions (Frauenfelder et al., 1979) and protein-ligand interactions (Austin et al., 1975). On the other hand, dynamics of internal motions of these molecules over a wide range of magnitude and time scale, as well as other problems in protein structure and function, have been studied by a variety of computer simulation methods (McCammon and Harvey, 1987; Brooks et al., 1988; Karplus and Petsko, 1990). For example, results from molecular dynamics simulations showed that atomic motion can be highly complex (Northrup et al., 1981; Mao et al., 1982; Ichiye and Karplus, 1987); collective and large scale structural changes have also been studied (McCammon et al., 1977; Northrup et al., 1982; Mao and McCammon, 1983).

Due to the large number of degrees of freedom in a biomacromolecular system, the sampling of the multidimensional conformational space in molecular dynamics simulations is restricted to the local region near the initial conformation. Usually a crystallographically determined high resolution molecular structure of the protein is used as the starting structure in these simulations. Even for short polypeptide molecules, the conformational search may be limited in certain situations (Mao and Friedman, 1990). The mass-weighted molecular dynamics (MWMD) simulation method has been applied to the investigation of the conformational flexibility in small linear polypeptides (Mao and Friedman, 1990; Mao, 1991) and cyclic and pseudocyclic molecules (Mao et al., 1991). In these simulations, atomic masses in the molecular system were multiplied uniformly by a numerical weighting factor ω ($\omega > 1.0$); rigid internal degrees of freedom of the system (i.e., covalent bonds, bond angles, amide planes, and chiral carbon centers) were maintained by constraints such that these structural features do not become distorted during mass-weighted molecular dynamics simulations. In the otherwise standard molecular dynamics simulation, potential energy surface in the "soft" degrees of freedom (e.g., dihedral rotations and nonbonded interactions, which confer molecular flexibility) is not altered by these constraints and energy barriers are thereby overcome more readily due to large effective momenta of the particles. The mass-weighting scheme has an advantage over the high-temperature scheme in that a large value can be used for ω and atomic velocities can remain in the regime in which the MD integration maintains numerical stability. In comparison with normal-mass molecular dynamics (NMMD) simulations, the full 360° range of

0006-3495/91/10/966/08 $2.00

individual backbone dihedral angles ϕ's and ψ's in a linear polypeptide was sampled thoroughly when the atomic masses were weighted from their respective normal values; moreover, the correlation of the values of ϕ and ψ torsional angles (ϕ-ψ plot) of individual amino acid residues in the MWMD simulation shows that the underlying potential energy surface is consistent with the adiabatic potential energy contours in Ramachandran plots and not affected by the constraints (Mao, 1991), and that the ϕ-ψ conformational space is efficiently sampled. In simulations of pseudocyclic molecules, the extensive sampling of the physical space about the fixed points in MWMD suggests that the method can generate polypeptide conformations for modeling protein surface loops (Mao et al., 1991).

From these earlier investigations in which NMMD and MWMD of simple polypeptide systems were studied comparatively, it appears that atomic mass weighting in molecular dynamics simulations could be a general concept for sampling conformational space of molecular systems that contain flexible degrees of freedom such as protein-ligand interactions in a biomolecular complex. Results from the rhizopuspepsin-inhibitor complex in this work indicate efficient sampling of both the internal (dihedral angle) coordinate space and the physical space of the binding pocket. Moreover, a secondary inhibitor binding site was identified which may be hypothesized to be related to the slow-binding inhibition observed for aspartyl proteases and other enzymes. In molecular complexes for which no structural data are available but for which molecular models can be constructed, the method can be employed to evaluate the stability of binding of a proposed ligand binding geometry, and/or to generate other potential ligand binding sites.

METHODS

For investigating the application of our MWMD simulation method to protein-ligand complexes, the crystal structure of a fungal aspartyl protease, rhizopuspepsin, complexed with a polypeptide inhibitor (Suguna et al., 1987) was chosen as a prototypical system (Fig. 1 *a*). For the purpose of exploring and developing the methodology, atoms of the enzyme were held fixed near their respective crystallographically determined positions, with all of the crystalline water molecules removed (see Fig. 1 for details). For comparison with the MWMD results, a normal-mass molecular dynamics (NMMD) simulation (Brooks et al., 1983) was first generated in which only inhibitor atoms were allowed to move freely. In the MWMD simulation, atoms in the polypeptide inhibitor were mass weighted and internal coordinates were constrained as in previous MWMD studies (Mao and Friedman, 1990; Mao, 1991); details of the simulation protocols are described in Fig. 2.

RESULTS

As shown in Fig. 2, the molecular dynamics integration has yet to reach equilibration after 200 ps of NMMD simulation under the conditions prescribed for this set of simulations; in contrast, the MWMD simulation became "equilibrated" at $\sim$75 ps, and the total energy of system remained essentially constant as expected for microcanonical ensembles generated from molecular dynamics simulations. The sampling of the full 360° range for a representative backbone dihedral angle and that for a representative sidechain dihedral angle during the two simulations are shown in Fig. 3. The sampling of the dihedral angle space can be summarized in Table 1 in which two indicators for the range and the coverage of the conformational sampling (Mao et al., 1991) were computed for unconstrained dihedral angles in the backbone (ϕ's and ψ's) and the sidechain (χ_1's) in the two simulations; the significant improvement in the sampling in MWMD is shown by the higher values of both indicators from the MWMD simulation than those from the NMMD simulation.

In addition to the improved sampling of internal dihedral angle degrees of freedom in MWMD, the sampling of physical space in the binding pocket is also increased in the mean time. The position of the Cα atom in Phe_5 of the inhibitor during the normal-mass MD simulation is shown in Fig. 4; in contrast, the position distribution not only becomes broadened in MWMD, but also shows two preferred locations in the binding pocket for this backbone atom (Fig. 5 *a*). Of these two preferred locations for the inhibitor, the primary site is located near the crystallographic position, and the secondary site (with a lower population) is located between the primary binding site and the opening of the binding pocket. The position distribution of Cγ of Phe_5 (Fig. 5 *b*) shows that the sidechain group also has two preferred locations, indicating that there are two specificity pockets for this residue. At the secondary binding site located near the exit to the solvent space, the kinetically energized inhibitor was in fact able to assume conformations that are required for eventually reaching the opening and escaping the enzyme binding pocket (Fig. 6). In the MWMD simulation in which the opening of the binding pocket was blocked by water molecules, which as a first approximation were kept fixed in their respective spatial positions as protein atoms were, the inhibitor molecule was, as expected, prevented from leaving the binding pocket; moreover, the same set of primary and secondary binding sites were also identified (data not shown).

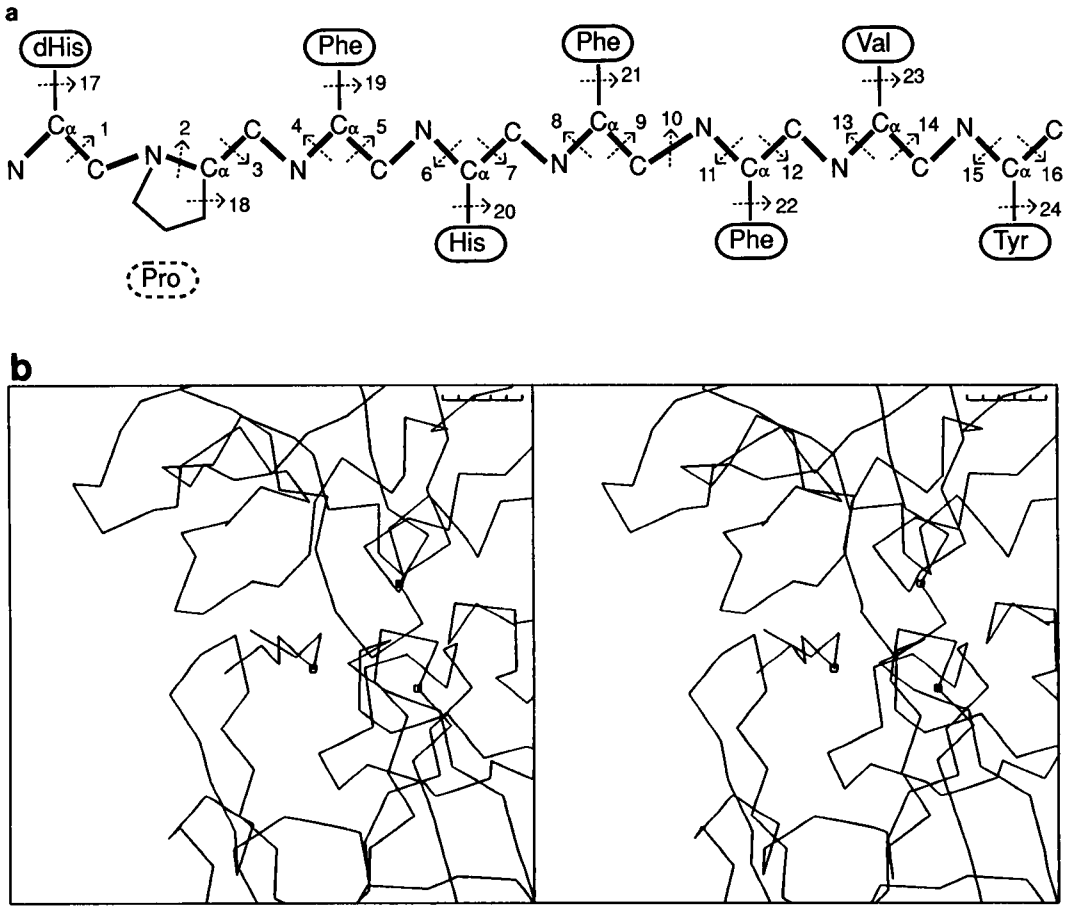

FIGURE 1 (*a*) Schematic drawing of the peptide inhibitor D-His-Pro-Phe-His-PheΨ[CH_2-NH]Phe-Val-Tyr. Each arrow indicates a dihedral angle that is allowed to freely rotate during simulations, except dihedral angle #18 which is monitored for proline ring conformation. All proper dihedral angles of sidechains are allowed to rotate; only χ_1's are explicitly indicated here. (*b*) Schematic representation of ligand binding region of the rhizopuspepsin-inhibitor complex, in stereoscopic images. The enzyme and the inhibitor are represented by the Cα tracing; sidechain groups are present in the simulations but not explicitly drawn here. The Cα atom of Phe_5 of the inhibitor and the Cα atoms of the catalytic residues of rhizopuspepsin, Asp_{35} and Asp_{218}, are marked with small squares. The scale on the upper right corner is a 5-Å bar for reference. In the crystallographically determined structure of the complex, the two terminal residues of the inhibitor did not have well-defined positions. Atomic coordinates for these two residues were first generated by standard geometry of the two residue types (Mao, unpublished data); the positions of their backbone and sidechains were then optimized by rigid rotation of the backbone and side-chain dihedral angles. Crystalline water molecules were then removed, and the complex energy minimized so that the molecular dynamics simulation could be started without extreme initial atomic velocities; the RMS position difference between the energy-minimized structure of the complex and the crystallographic coordinate set is 0.16 Å.

DISCUSSION

The results presented above showed the advantage of the MWMD method over the normal-mass MD simulation in sampling the conformational space of the ligand molecule in the complex. The system becomes equilibrated more rapidly. The sampling of internal and external degrees of freedom of the ligand is significantly increased; the sampling of the physical space of the binding pocket appears to be coupled to the sampling of dihedral conformational space in that the latter allows the inhibitor molecule as a whole to assume conformations that are significantly different from the crystallographic structure and thereby explore other binding geometries within the enzyme binding pocket. During the MWMD simulation, the inhibitor overcomes a potential energy barrier that apparently exists between the primary and the secondary binding sites. Moreover, the sampling of the conformational space for the inhibitor is such that, in the simulation in which the bulk solvent is not explicitly represented and the inhibitor molecule is energized kinetically by the weighting of

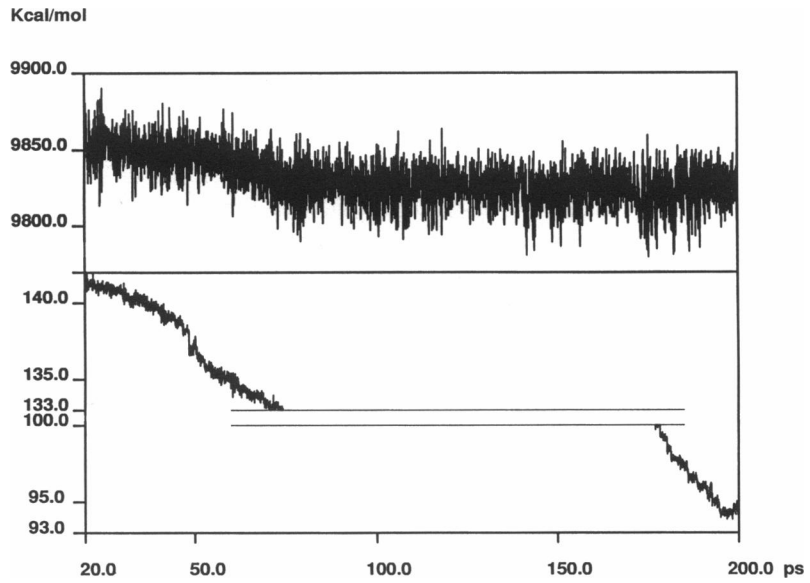

FIGURE 2 Total energy of MWMD (*upper curve*) and normal MD (*lower curve*) simulations. The simulation step is 0.9766 fs. Nonbonded interatomic interactions were cutoff at 8.0 Å and hydrogen-bond interactions were cutoff at 6.0 Å, and distance-dependent dielectric constant was used for electrostatic interactions; lists of hydrogen-bond and nonbonded atom pairs were regenerated at every 20 steps. The system was thermalized during the first 20 ps by heating in 30-K increments at 0.5-ps intervals, followed by reequilibrations at 0.5-ps intervals; the final nominal equivalent system temperature is 600 K for both simulations. The mass weighting factor for the inhibitor atoms is 30.0 in the MWMD simulation; harmonic potentials were set at 10,000.0 Kcal/mol-Å^2 for constraining covalent bonds, 4,000.0 Kcal/mol-rad^2 for valence angles, and 400.0 Kcal/mol-rad^2 for improper dihedral angles and amide planes.

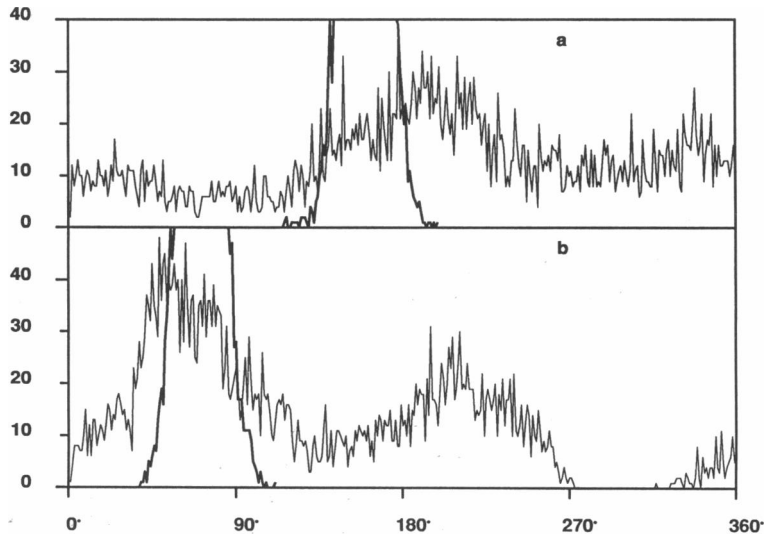

FIGURE 3 Distribution of numerical values in the 360° range for two of the dihedral angles shown in Fig. 1 *a*; (*a*) for dihedral angle #12, (*b*) for #22. For each dihedral angle, the 360° space is divided into 1° bins and number of time steps at which the dihedral angle has the numerical value at each bin is counted for the length of an MD simulation and plotted against the 360 bins. For each dihedral angle, the curve in thick line is obtained from the 200 ps normal mass molecular dynamics simulation, and the curve in thin line is obtained from the first 100 ps mass-weighted MD simulation. See Table 1 for details of the statistics for all dihedral angles shown in Fig. 1 *a*.

TABLE 1 **The range and the coverage of backbone and sidechain dihedral angle space of D-His-Pro-Phe-His-PheΨ[CH_2-NH]Phe-Val-Tyr in normal-mass MD (1–200 ps) and mass-weighted MD (1–100 ps) simulations**

	NMMD		MWMD		ratio	
	(*R*)	(*C*)	(*R*)	(*C*)	(*R*)	(*C*)
		%		%		
1	223	10.04	323	28.80	1.45	2.87
2	92	7.53	158	14.73	1.72	1.95
3	98	9.97	359	34.10	3.66	3.42
4	83	7.28	356	24.00	4.29	3.30
5	73	6.64	302	25.41	4.14	3.82
6	72	6.82	309	20.90	4.29	3.06
7	101	11.27	321	22.73	3.18	2.02
8	79	7.24	305	17.75	3.86	2.45
9	100	6.75	296	21.24	2.96	3.15
10	93	6.78	219	22.73	2.35	3.35
11	116	7.20	312	27.00	2.69	3.75
12	76	6.86	360	34.10	4.74	4.97
13	114	13.22	358	35.99	3.14	2.72
14	103	9.82	352	25.92	3.42	2.64
15	134	10.71	360	46.28	2.69	4.32
16	342	30.14	360	46.28	1.05	1.54
17	181	10.13	359	35.99	1.98	3.56
18	111	7.16	73	6.86	0.66	0.96
19	97	8.36	350	22.73	3.61	2.72
20	170	13.09	347	26.44	2.04	2.02
21	80	6.58	271	20.90	3.39	3.18
22	67	6.23	306	27.00	4.57	4.33
23	169	9.60	358	33.23	2.12	3.46
24	142	8.25	343	22.73	2.42	2.75

For either simulation, the *range (R)* and the *coverage (C)* indicators are computed for each dihedral angle as labeled in Fig. 1 *a*. From the distribution of numerical values in the full 360° range for a given dihedral angle in either NMMD or MWMD, such as those shown in Fig. 3, *R* is the number of 1° bins in the 360° range that are occupied (i.e., the number of points along the *x*-axis that have nonzero *y*-values); *C* is defined as the percentage of the rectangular area (from 0 to 360 in the *x*-axis, and from 0 to the maximum *y*-value for the curve in the *y*-axis) that is under the curve. The *C* indicator has a minimum value of 0.278% (1/360) for an extremely narrow distribution, and a maximum value of 100.0% for an entirely uniform distribution. The ratios in *R* and *C* (MWMD value/NMMD value) are also listed for each dihedral angle.

atomic masses, the inhibitor, while in the secondary binding site, was able to locate the opening of the binding pocket and assume conformations suitable for negotiating the path for leaving the binding pocket.

The absence of solvent bulk and the rigidity of the enzyme binding pocket in the present study obviously require more sophisticated treatments. The description of the primary and the secondary binding sites for inhibitors of aspartyl proteases presented here nevertheless may remain unchanged when these factors are included in the simulation, for the following reasons. Position fluctuations of atoms in a protein are estimated, both experimentally and theoretically, to be generally on the order of 0.7 Å (Northrup et al., 1982). Because the binding sites are ~3–4 Å in size and are separated by a distance of 4–5 Å (Fig. 5), the flexibility of the enzyme binding pocket due to atomic position fluctuations on the order of 0.7 Å will be expected to modulate somewhat, but not alter dramatically, the potential energy surface and the two binding sites revealed in the present MWMD simulation. On the other hand, the inclusion of fixed water molecules in the simulation prevents the inhibitor from leaving the binding site and forces the inhibitor to sample only the interior of the binding pocket; still the same set of primary and secondary binding sites were identified, without any additional sites for the inhibitor. Thus only two such binding sites appear to exist for certain inhibitors of aspartyl proteases such as that in the bimolecular complex of rhizopuspepsin studied here; this would be consistent with the two-state model of slow-binding inhibition (i.e., initial and final enzyme-inhibitor complexes) which was proposed from kinetics studies of slow inhibition of aspartyl proteases (Kati et al., 1987) and other enzymes (Morrison and Walsh, 1988). Moreover, given that these two sites have been identified with fixed protein atoms in the simulations, it may be hypothesized further that the conformational change of the enzyme that might be required for ligand binding (Morrison, 1982; Kati *et al.*, 1987) may have occurred before the formation of the initial enzyme-inhibitor complex; i.e., the conformational change of the enzyme is completed on or before the formation of the complex in which the inhibitor binds in the secondary binding site, and the inhibitor subsequently moves to the primary binding site in the slow step and becomes more tightly bound without any major conformational change of the enzyme (Fig. 5). The secondary binding site, predicted from the crystal structure of the complex and the sampling of the potential energy surface at its binding pocket, can be verified experimentally, for example, in a system in which an extremely slow-binding inhibitor of an aspartyl protease (Bartlett and Kezer, 1984) is cross-linked to enzyme residues via suitable photoaffinity labels (Chowdhry and Westheimer, 1979).

In summary, we described the mass-weighted molecular dynamics simulation of an enzyme-inhibitor complex; results from this and previous studies using the MWMD method show that a variety of biophysical and biochemical problems in the structure and function of biomacromolecular systems can be approached by this molecular dynamics-based simulation method. With a more sophisticated treatment of solvent effects and enzyme binding site flexibility, it may also become possible to study protein conformational substates in ligand binding (Austin et al., 1975; Frauenfelder et al., 1988), as well as the

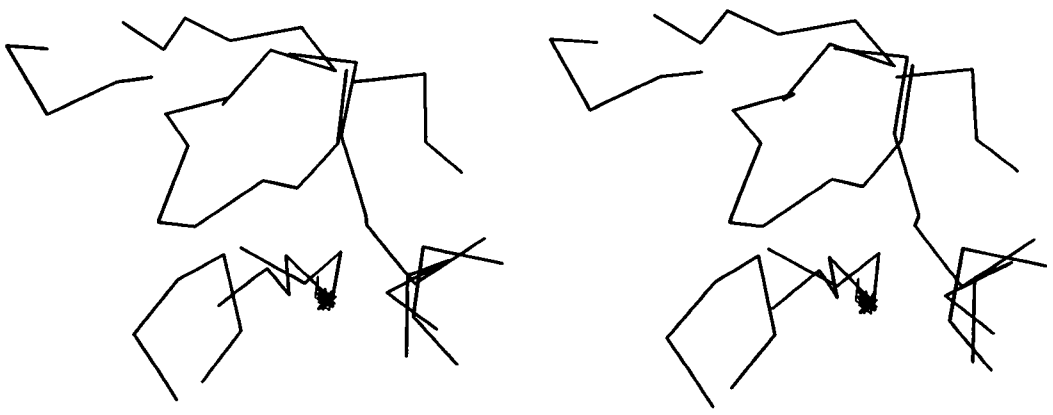

FIGURE 4 Trajectory of atomic position, in stereoscopic images, for the Cα atom of Phe_5 in the inhibitor shown in Fig. 1 *a* during normal mass MD simulation. The enzyme binding pocket and the initial inhibitor structure are shown in thick lines. Similarly tight distributions are observed for other Cα atoms of the inhibitor, except for Val at the amino and Tyr at the carboxyl terminus.

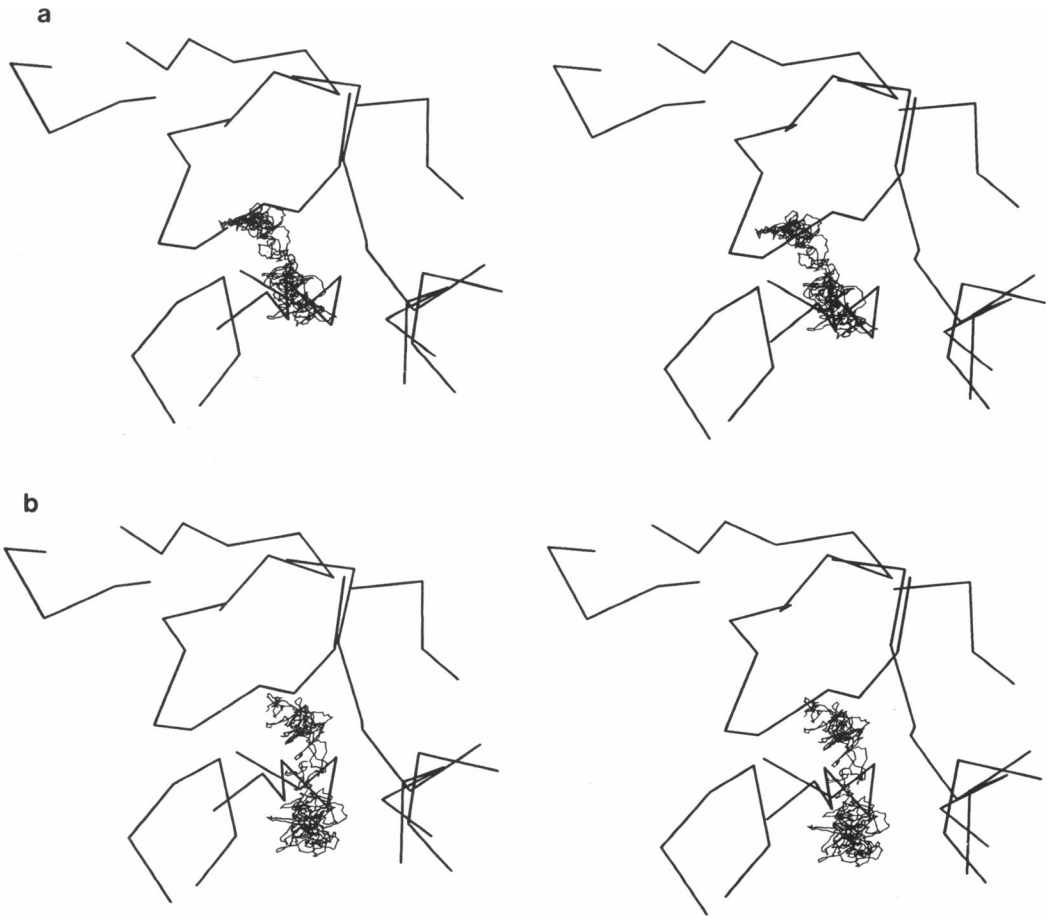

FIGURE 5 (*a*) Trajectory of atomic position for the Cα atom of Phe_5 in the inhibitor during the period from 70 to 120 ps of MWMD simulation. Two preferred locations in the binding pocket are identified for this atom. The Cα atoms of the four central residues, from His to Val, also have two preferred locations, whereas terminal residues have one but generally broad distribution. (*b*) Trajectory of atomic position for the Cγ atom of Phe_5.

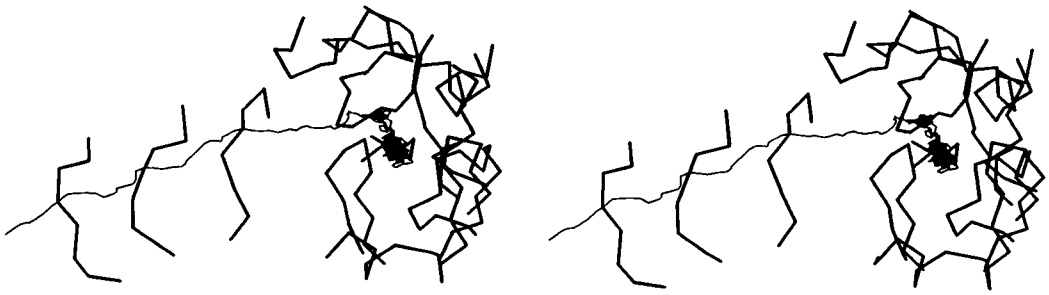

FIGURE 6 Trajectory of atomic position for the Cα atom of Phe_5 in the inhibitor during the period from 20 to 130 ps of MWMD simulation. The three structures of the inhibitor located outside of the binding pocket are extracted from the MWMD simulation at 123.82, 127.02, and 129.82 ps, respectively; the amino terminal of the inhibitor is pointing downward. The last structure is located ~40 Å away from the binding site.

role of such substates in ligand binding kinetics. As a practical note, for many enzyme-ligand complexes for which structural data (e.g., NMR spectroscopy, x-ray crystallography) are not available but for which molecular models have been constructed, it has been difficult to evaluate the accuracy of these models due to the limitation of energy-minimization and molecular dynamics simulation in sampling only the local conformational space. The extensive conformational sampling in MWMD simulations of a protein-ligand complex will help find appropriate binding site(s) for the ligand, independent of its starting conformation and physical location in the binding pocket; alternative binding sites can be identified from the simulation results. Moreover, the intermediate binding steps and the molecular flexibility of the inhibitor shown in the MWMD simulation may have crucial roles in the ligand binding process. These considerations are important factors in successful, structurally motivated inhibitor designs.

The author thanks F. J. Kézdy for directing his attention to the literature on slow-binding inhibition and for general discussions. He also would like to acknowledge helpful discussion with D. J. Duchamp, and thank M. Karplus for making the Harvard macromolecular computation program CHARMM available.

Received for publication 8 February 1991 and in final form 28 May 1991.

REFERENCES

Austin, R. H., K. W. Beeson, L. Eisenstein, H. Frauenfelder, and I. C. Gunsalus. 1975. Dynamics of ligand binding of myoglobin. *Biochemistry.* 14:5355–5373.

Bartlett, P. A., and W. B. Kezer. 1984. Phosphinic acid dipeptide analogues: potent, slow-binding inhibitors of aspartic peptidases. *J. Am. Chem. Soc.* 106:4282–4283.

Brooks, B. R., R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus. 1983. CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. *J. Comp. Chem.* 4:187–217.

Brooks, C. L. III, M. Karplus, and B. M. Pettitt. 1988. Proteins: A Theoretical Perspective of Dynamics, Structure, and Thermodynamics. John Wiley & Sons, New York.

Chowdhry, V., and F. H. Westheimer. 1979. Photoaffinity labeling of biological systems. *Annu. Rev. Biochem.* 48:293–325.

Debrunner, P. G., and H. Frauenfelder. 1982. Dynamics of proteins. *Annu. Rev. Phys. Chem.* 33:283–299.

Frauenfelder, H., F. Parak, and R. D. Young. 1988. Conformational substates in proteins. *Annu. Rev. Biophys. Biophys. Chem.* 17:451–479.

Frauenfelder, H., G. A. Petsko, and D. Tsernoglou. 1979. Temperature-dependent x-ray diffraction as a probe of protein structural dynamics. *Nature (Lond.).* 280:558–563.

Ichiye, T., and M. Karplus. 1987. Anisotropy and anharmonicity of atomic fluctuations in proteins: analysis of a molecular dynamics simulation. *Proteins.* 2:236–259.

Karplus, M., and G. A. Petsko. 1990. Molecular dynamics simulations in biology. *Nature (Lond.).* 347:631–639.

Kati, W. M., D. T. Pals, and S. Thaisrivongs. 1987. Kinetics of the inhibition of human renin by an inhibitor containing a hydroxyethylene dipeptide isostere. *Biochemistry.* 26:7621–7626.

Mao, B. 1991. Mass-weighted molecular dynamics simulation and conformational analysis of polypeptide. *Biophys. J.* 60:611–622.

Mao, B., and A. R. Friedman. 1990. Molecular dynamics simulation by atomic mass weighting. *Biophys. J.* 58:803–805.

Mao, B., and J. A. McCammon. 1983. Theoretical study of hinge bending in L-arabinose-binding protein. *J. Biol. Chem.* 258:12543–12547.

Mao, B., G. M. Maggiora, and K. C. Chou. 1991. Mass-weighted molecular dynamics simulation of cyclic polypeptides. *Biopolymers.* 31:1077–1086.

Mao, B., M. R. Pear, J. A. McCammon, and S. H. Northrup. 1982. Molecular dynamics of ferrocytochrome *c*: anharmonicity of atomic displacements. *Biopolymers.* 21:1979–1989.

McCammon, J. A., and S. C. Harvey. 1987. Dynamics of Proteins and Nucleic Acids. Cambridge University Press, Cambridge.

McCammon, J. A., B. R. Gelin, M. Karplus, and P. G. Wolynes. 1977. The hinge-bending mode in lysozyme. *Nature (Lond.).* 262:325–326.

Morrison, J. F. 1982. The slow-binding and slow, tight-binding inhibition of enzyme-catalyzed reactions. *Trends Biochem. Sci.* 7:102–105.

Morrison, J. F., and C. T. Walsh. 1988. The behavior and significance of slow-binding enzyme inhibitors. *Adv. Enzymol.* 61:201–301.

Northrup, S. H., M. R. Pear, C. Y. Lee, J. A. McCammon, and M. Karplus. 1982. Dynamical theory of activated processes in globular proteins. *Proc. Natl. Acad. Sci. USA.* 79:4035–4039.

Northrup, S. H., M. R. Pear, J. A. McCammon, M. Karplus, and T. Takano. 1982. Internal mobility of ferrocytochrome *c*. *Nature (Lond.).* 287:659–660.

Northrup, S. H., M. R. Pear, J. D. Morgan, and J. A. McCammon. 1981. Molecular dynamics of ferrocytochrome *c*. Magnitude and anisotropy of atomic displacements. *J. Mol. Biol.* 153:1087–1109.

Suguna, K., E. A. Padlan, C. W. Smith, W. D. Carlson, and D. R. Davies. 1987. Binding of a reduced peptide inhibitor to the aspartic proteinase from *Rhizopus chinensis:* implication for a mechanism of action. *Proc. Natl. Acad. Sci. USA.* 84:7009–7013.

Wagner, G., and K. Wüthrich. 1986. Observation of internal motility of proteins by nuclear magnetic resonance in solution. *Methods Enzymol.* 131:307–326.