Dynamic Flexibility of Protein−Inhibitor Complexes: A Study of the HIV-1 Protease/KNI-272 Complex

Xincai Luo,† Ryohei Kato,‡ and Jack R. Collins*

Contribution from the Frederick Biomedical Supercomputer Center, SAIC−Frederick/NCI, 430 Miller Drive, Frederick, Maryland 21702

Received July 8, 1998. Revised Manuscript Received September 29, 1998

Abstract: The dynamics and flexibility of protein−ligand complexes is central to understanding and predicting binding geometries and energetics. We have calculated various measures of the dynamic flexibility of a pseudo-*C*2-symmetric protein, HIV-1 protease, complexed with the asymmetric inhibitor KNI-272 based on molecular dynamics simulations. This system is expected to be an excellent candidate for observing asymmetric dynamics between the two monomers due to the differences in the interactions between the two monomers of the protease and the inhibitor. Experimental methods have thus far been unable to observe the expected asymmetry in this system. Our calculated results are in excellent agreement with the available experimental data for the main-chain order parameters from a parallel ^{15}N NMR study of the same inhibitor−protein complex, as well as the Debye−Waller temperature factors from X-ray crystallography. In our simulations, asymmetry between the monomers is found almost exclusively in the side-chain order parameters of the inhibitor and protease (especially residues 84A and 84B), for which experimental data are not yet available. We analyze the dynamic information obtained from the different methods and discuss protein−ligand interactions responsible for the dynamical behavior of the complex.

Introduction

Ligand and protein flexibility is a major challenge to the computational prediction and design of new and more therapeutically desirable drug candidates. It is well established that ligand binding can induce structural changes in both the ligand and the protein with which it is interacting and that these internal motions can determine a protein's unique function and properties.[1−4] Internal motions occur over a wide range of time scales and are often crucial for allowing a protein to adopt a specific conformation for ligand binding, release, and function.[5,6] Information about protein motion can be obtained from experimentally derived values such as NMR order parameters, time-resolved fluorescence depolarization rates, and crystallographic *B*-factors. Computer simulations can also probe questions of flexibility in protein−ligand interactions and can serve as a useful tool in the prediction, analysis, and interpretation of the aforementioned experimental methods. In this study, we perform molecular dynamics (MD) simulations on a protein−ligand complex and compare them with concurrent NMR measurements[7] and previously published crystal data[8] in an effort to ascertain the best properties and analyses for elucidating important regions of flexibility in the bound complex.

Nuclear magnetic resonance is a particularly valuable probe of the fast internal protein motions that occur in the solution state. Recent advances in NMR spectroscopy provide sensitive techniques for probing local N−H and C−H motions on the picosecond time scale, not visible to long-time-averaged methods such as X-ray crystallography, by calculating order parameters (S^2) derived from the NMR data.[9] Despite the power of these NMR studies, however, they are limited by the constraints imposed by experimental design considerations such as sample labeling, solubility, signal resolution, and approximations in data analysis. Often, much more information can be obtained by coupling the experimental NMR data with X-ray crystallographic structures and *B*-factors as well as detailed computational analyses based on simulation.[10,11]

When protein and ligand complexes can be crystallized, motions on a time scale significantly longer than that of NMR techniques can generally be correlated to the mean-square atomic fluctuations, related to the Debye−Waller temperature factors

* Corresponding author: e-mail collinsj@ncifcrf.gov.

† Smith-Kline Beecham postdoctoral fellow, funded by NIH Grant GM50579.

‡ Permanent address: Pharmaceuticals and Biotechnology Laboratory, Japan Energy Corp., Toda, Saitama 335, Japan.

(1) Lesk A. M.; Chothia, C. *J. Mol. Biol.* **1984**, *174*, 175−191.
(2) Sali, A.; Veerapandian, B.; Cooper, J. B.; Moss, D. S.; Hofmann, T.; Blundell, T. L. *Proteins: Struct., Funct., Genet.* **1992**, *12*, 158−170.
(3) Stillman, T. J.; Baker, B. J.; Britton, K. L.; Rice, D. W. *J. Mol. Biol.* **1993**, *234*, 1131−1139.
(4) Anderson, B. F.; Baker, H. M.; Norris, G. E.; Rumball, S. V.; Baker, E. N. *Nature* **1990**, *344*, 784−787.
(5) Baldwin, J.; Chothia, C. *J. Mol. Biol.* **1979**, *129*, 175−220.
(6) Rasmussen, B. F.; Stock, A. M.; Ringe, D.; Petsko, G. A. *Nature* **1992**, *357*, 423−424.
(7) Freedberg, D. I.; Wang, Y.; Stahl, S. J.; Kaufman, J. D.; Wingfield, P. T.; Kiso, Y.; Torchia, D. A. *J. Am. Chem. Soc.* **1998**, *120*, 7916−7923.
(8) Nicholson, L. K.; Yamazaki, T.; Torchia, D. A.; Grzesiek, S.; Bax, A.; Stahl, S. J.; Kaufman, J. D.; Wingfield, P. T.; Lam, P. Y. S.; Jadhav, P. K.; Hodge, C. N.; Domaille, P. J.; Chang, C. H. *Nat. Struct. Biol.* **1995**, *2*, 274−280.
(9) Frank, M. K.; Clore, G. M.; Gronenborn, A. M. *Protein Sci.* **1995**, *4*, 2605−2615.
(10) McCammon, A. J.; Harvey, S. C. *Dynamics of Proteins and Nucleic Acids*; Cambridge University Press: Cambridge, UK, 1986.
(11) Brooks, C. L.; Karplus, M.; Pettitt, B. M. *Proteins: A Theoretical Perspective of Dynamics, Structure, and Thermodynamics*; Advances in Chemical Physics LXXI; John Wiley & Sons: New York, 1988.

10.1021/ja9824066 CCC: $15.00
Published on Web 11/18/1998

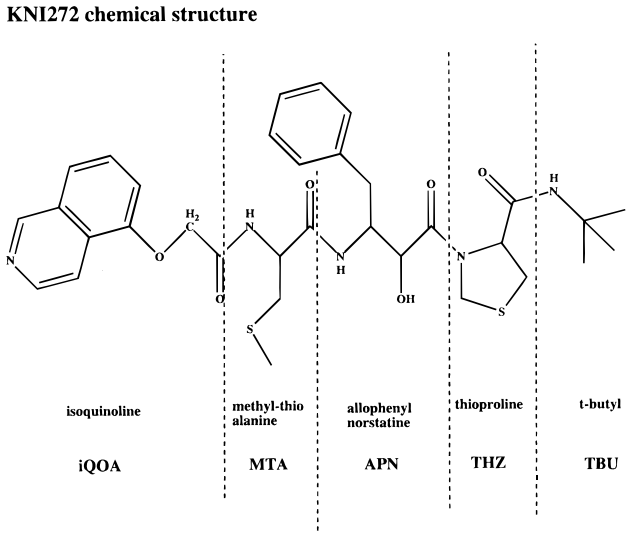

Figure 1. Chemical structure of KNI-272. The name and abbreviation for each of the chemical groups interacting with the subsites of HIV-1pr are shown for reference.

(*B*-factors), determined from fitting the three-dimensional X-ray model to the observed electron density.[12] *B*-factors can be good relative indicators of conformational flexibility and motion within a molecule and useful in certain aspects of structure-based drug design strategies. However, insufficient structure resolution, intermolecular crystal contacts, solvent scattering, and conformational heterogeneity within the unit cell can present difficulties in accurately determining the *B*-factors.

Computationally, MD simulations can provide detailed information complementary to each of the experimental techniques mentioned above.[10,11] Using the appropriate statistical average or correlation function, the experimental observable of interest can be evaluated from the detailed molecular motions generated from a MD simulation.[13] Fast molecular fluctuations, less than 100 ps, are particularly well-characterized by conventional MD simulations, and many studies have shown that the mean-square fluctuations derived from sufficiently long simulations compare well with the crystallographic *B*-factors.[11]

Detailed comparisons between order parameters calculated from MD simulations and the experimentally derived values have been reported by several groups.[14−18] In general, good agreement between experiment and simulation is found for well-defined systems. Recently, Eriksson et al.[15] compared the experimental order parameters with those derived from MD simulations for the glucocorticoid receptor DNA-binding domain and found that corrections were necessary to account for certain residues that converted between distinct conformations during the simulation. They argue that the experimental order parameters are heavily weighted by motions in the 10−50-ps range

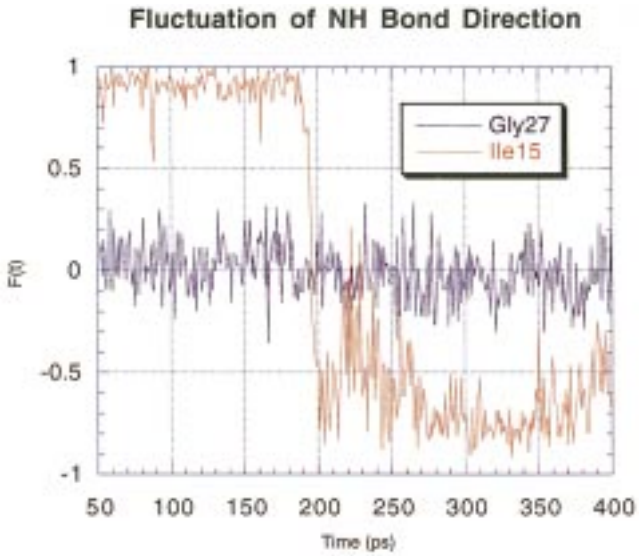

Figure 2. During the dynamics simulation, the NH bond direction may change from one state to another or be quite stable during the course of the trajectory. The transition of the NH bond direction, near 200 ps, for Ile15 in monomer B is shown in red. Gly 27, shown in blue, is characteristic of a highly ordered NH bond with excellent convergence in the reorientational correlation function.

and that large, infrequent transitions are insufficiently sampled in the practical limit of MD simulations. They suggest that these transitions should be omitted from the simulation analysis and only short, stable time frames of the trajectories should be used when calculating the order parameters from MD. Such corrections were shown to lead to better agreement between the experimentally derived order parameters and those calculated from MD simulation. Chandrasekhar et al.[14] have also shown that caution must be used when calculating the reorientational correlation functions, used to calculate the order parameters, from simulations that include single large-amplitude transitions or infrequently sampled events. The authors argue that these events lead to statistically unmeaningful, and unconverged, correlation decays. From their detailed analysis of the correlation functions, they caution that the three cases we investigate in our analysis (discussed in the Methods section) must be considered separately to reliably predict order parameters. When care is taken to account for these cases, the calculated order parameters show quite good agreement with the NMR-derived values. From these previous studies, we are confident that careful simulation followed by a thorough analysis of the resulting correlation functions can lead to meaningful and reliable predictions of NMR order parameters, crystallographic *B*-factors, and other dynamic properties.

To investigate various measures of dynamic flexibility on a ligand-bound protein complex, we have studied the well-characterized system of human immunodeficiency virus type-1 protease (HIV-1pr)[19−21] bound to an asymmetric inhibitor KNI-272 (see Figure 1). This system is chosen for these studies since both NMR and crystallographic data are available for comparison. In addition, this complex is particularly well-suited for our studies since the protease forms a C_2-symmetric homodimer in solution and is expected to interact asymmetrically with KNI-272, an inhibitor that forms an exceptionally

(12) Blundell, T. L.; Johnson, L. N. *Protein Crystallography*; Academic Press Inc.: London, 1976. Creighton, T. E. *Proteins. Structures and Molecular Properties*; W. H. Freeman & Co.: New York, 1984; pp 204−220.

(13) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Oxford University Press: Oxford, UK, 1989.

(14) Chandrasekhar, I.; Clore, G. M.; Szabo, A.; Gronenborn, A. M.; Brooks, B. R. *J. Mol. Biol.* **1992**, *226*, 239−250.

(15) Eriksson, M. A. L.; Berglund, H.; Hard, T.; Nilsson, L. *Proteins: Struct., Funct., Genet.* **1993**, *17*, 375−390.

(16) Smith, L. J.; Mark, A. E.; Dobson, C. M.; van Gunsteren, W. F. *Biochemistry* **1995**, *34*, 10918−10931.

(17) Yamasaki, K.; Saito, M.; Oobatake, M.; Kanaya, S. *Biochemistry* **1995**, *34*, 6587−6601.

(18) Philippopoulos, M.; Mandel, A. M.; Palmer, A. G., III; Lim, C. *Proteins: Struct., Funct., Genet.* **1997**, *28*, 481−493.

(19) Tomasselli, A. G.; Howe, W. J.; Sawyer, T. K.; Wlodawer, A.; Heinrikson, R. L. *Chim. Oggi* **1991**, *9*, 6−27.

(20) Wlodawer, A.; Erickson, J. W. *Annu. Rev. Biochem.* **1993**, *62*, 543−585.

(21) Appelt, K. *Perspect. Drug Discov. Des.* **1993**, *1*, 23−48.

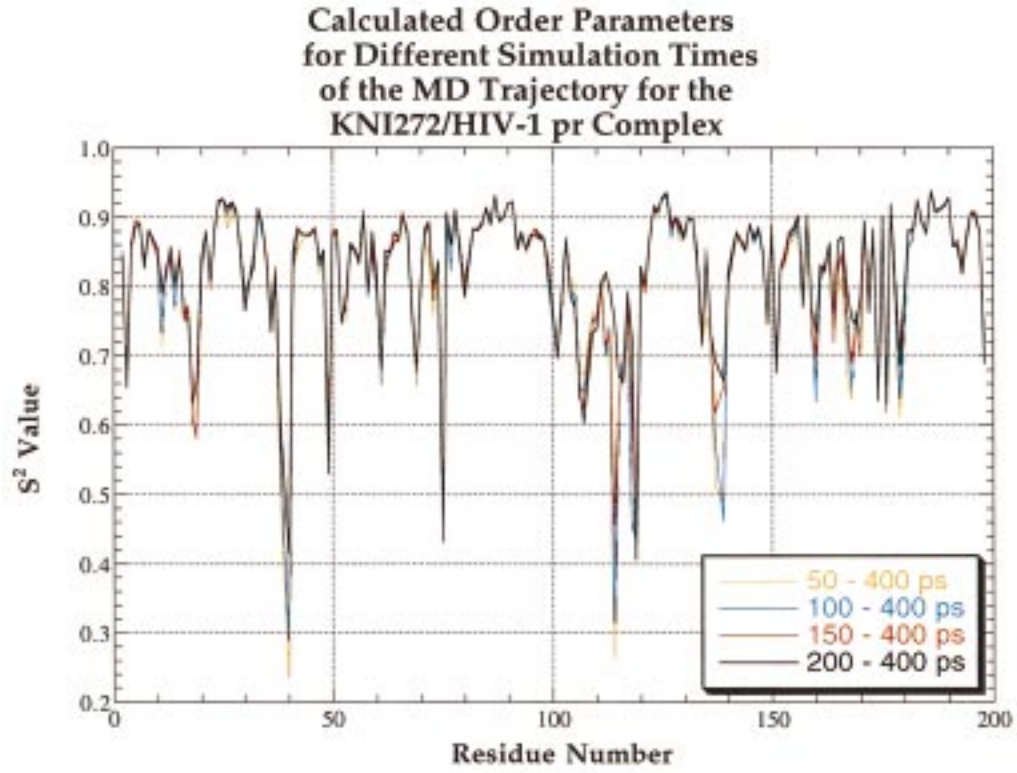

Figure 3. Backbone order parameters calculated from four different lengths of trajectories. Most of the residues show no significant changes in the calculated values of their order parameters. Those showing large deviations corresponded to residues where the reorientational correlation function did not converge. The order parameters for these latter residues were calculated over shorter ranges of the trajectory as described in the text.

tight complex (~4 pM) with the protease.[22] We describe MD simulations and the calculation of main-chain and side-chain order parameters of the protein and inhibitor along with the mean-square fluctuation of all of the atoms in the complex, focusing on the bound complex of KNI-272 with HIV-1pr to address the question of amplitude, time scale, and types of motion of both the inhibitor and the protein.

Methods and Calculation

The order parameters for the HIV-1pr/KNI-272 complex are calculated from MD simulations using the model-free formulation developed by Lipari and Szabo.[23] Starting from the X-ray crystal structure of the HIV-PR complex with KNI-272, reported as 1HPX in the Brookhaven protein data bank,[24] all hydrogens and a periodic box of solvent are added using the EDIT module of AMBER,[25] and the protein, inhibitor, and solvent are modeled using the AMBER all-atom force field.[26] Charges for KNI-272 are calculated with Gaussian 92[27] and the Merz−Kollman/Singh method[28] of fitting pointcharges to the quantum mechanical electrostatic potential calculated at the 6-31G* level. One of the aspartic acids in the catalytic dyad (monomer A, Asp25) is explicitly protonated in the model. Missing force field parameters are estimated from similar chemical species in the AMBER database. Three crystallographically determined water molecules (301, 607, and 608) found in the active site are retained. The model system is solvated in a rectangular box ($77.6 \times 55.5 \times 45.3$ Å^3) of 4586 water molecules, resulting in a total of 16 979 atoms in the system. The enzyme, inhibitor, and water molecule are initially minimized for 500 iterations of steepest descents, followed by conjugate gradients minimization until the root-mean-square deviation (RMSD) becomes less than 0.05 (total 3955 steps). The MD simulations are performed at a constant temperature of 300 K with a 9 Å cutoff, a constant dielectric of 1, a time step of 1 fs, and periodic boundary conditions keeping the system at a pressure of 1 atm. The initial heating is performed in the first 5 ps, followed by 50 ps of equilibration. Coordinate sets of the enzyme, the inhibitor, and crystallographically determined water molecules are saved every 0.1 ps for 350 ps following the equilibration. Before the order parameters are calculated, the translations and rotations of the protein−inhibitor complex are removed. The translation is removed by aligning the center of mass of each frame with the center of mass of the first frame, and rotations are removed after adjusting the translation by performing a mass-weighted least-squares fit following the procedure of Kabsch.[29]

It is known that there are instances where the standard procedure of calculating order parameters does not converge to a constant value.[14,15] In these cases, the NH bond direction jumps between two or more distinct minima and results in significantly lower order parameters if the range of the trajectory includes the two states. As shown by Eriksson et al.,[15] good estimates of the order parameters can be

(22) Baldwin, E. T.; Bhat, T. N.; Gulnik, S.; Liu, B.; Topol, I. A.; Kiso, Y.; Mimto, T.; Mitsuya, H.; Erickson, J. W. *Structure* **1995**, *3*, 581−590.

(23) Lipari, G.; Szabo, A. *J. Am. Chem. Soc.* **1982**, *104*, 4546−4559.

(24) Abola, E. E.; Bernstein, F. C.; Bryant, S. H.; Koetzle, T. F.; Weng, J. Protein Data Bank. In *Crystallographic Databases—Information Content, Software Systems, Scientific Applications*; Allen, F. H., Bergerhoff, G., Sievers, R., Eds.; Data Commission of the International Union of Crystallography: Bonn/Cambridge/Chester, 1987; pp 107−132. 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. *J. Mol. Biol.* **1977**, *112*, 535−542.

(25) Pearlman, D. A.; Case, D. A.; Caldwell, J. W.; Ross, W. S.; Cheatham, T. E., III; Ferguson, D. M.; Seibel, G. L.; Singh, U. C.; Weiner, P. K.; Kollman, P. A. *AMBER 4.1*; University of California, San Francisco, 1995.

(26) Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case. D. A. *J. Comput. Chem.* **1986**, *7*, 230.

(27) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Wong, M. W.; Foresman, J. B.; Robb, M. A.; Head-Gordon, M.; Replogle, E. S.; Gomperts, R.; Andres, J. L.; Raghavachari, K.; Binkley, J. S.; Gonzalez, C.; Martin, R. L.; Fox, D. J.; Defrees, D. J.; Baker, J.; Stewart, J. J. P.; Pople, J. A. *Gaussian 92/DFT*, revision G.2; Gaussian, Inc., Pittsburgh, PA, 1993.

(28) Besler, B. H.; Merz, K. M., Jr.; Kollman, P. A. *J. Comput. Chem.* **1990**, *11*, 431.

(29) Kabsch, W. *Acta Crystallog. Sect. A* **1976**, *32*, 922−923.

calculated only for a shorter range of the trajectory in which no transition occurs. To identify residues and ranges in the trajectory where such transitions may be occurring, we calculate the average of the NH bond directions for the first 50 ps of the trajectory and the average for the whole trajectory as follows:

$$\vec{r}_{50} = \frac{\sum_{50\text{ ps} \le t \le 100\text{ ps}} (\vec{r}_h(t) - \vec{r}_n(t))}{\left| \sum_{50\text{ ps} \le t \le 100\text{ ps}} (\vec{r}_h(t) - \vec{r}_n(t)) \right|} \quad (1)$$

$$\vec{r}_a = \frac{\sum_{50\text{ ps} \le t \le 400\text{ ps}} (\vec{r}_h(t) - \vec{r}_n(t))}{\left| \sum_{50\text{ ps} \le t \le 400\text{ ps}} (\vec{r}_h(t) - \vec{r}_n(t)) \right|} \quad (2)$$

With these two vectors, we generate a reference (normalized) vector such that

$$\vec{V} = \alpha\vec{r}_{50} + \beta\vec{r}_a \quad (3)$$

$$\vec{V}\cdot\vec{r}_a = 0 \quad (4)$$

$$\vec{V}\cdot\vec{r}_{50} > 0 \quad (5)$$

With this reference vector, we calculate a function $F(t)$ as follows:

$$F(t) = V\cdot(\vec{r}_h(t) - \vec{r}_n(t)) \quad (6)$$

The function has the properties that $-1 \le F(t) \le 1$ and the average is zero. When there is no transition in the NH bond direction, $F(t)$ fluctuates around zero (Figure 2). In contrast, when there is a transition in the NH bond direction, $F(t)$ deviates significantly from zero (Figure 2). For any residue where a transition occurs, we select a shorter range of the trajectory to calculate the order parameters.

It has been shown previously[14] that the shapes of the correlation functions used to calculate order parameters generally fall into three categories: (1) rapid initial decay followed by a stable plateau; (2) initial rapid decay, intermediate plateau, and then a slow decay; and (3) continuous decay with no plateau region. Only the first category leads to reliably converged order parameters, whereas the latter two categories indicate contributions from motions other than fast vibration within a relatively narrow well and can lead to difficulties in interpreting the experimental NMR data. Often these motions are in functionally important regions of the protein. To fully elucidate these motions, a synergy between experiment and simulation is often required to experimentally verify the reliability of the MD simulations and to extract information from the simulation that is not easily obtained from experiment. A detailed account of the calculation and analysis of reorientational correlation functions obtained from MD and related to NMR has been published previously.[14]

To test whether the computed results depend on the length of our simulation, we calculate the order parameter values using four different lengths (200, 250, 300, and 350 ps) of the trajectory. These four lengths correspond to the simulation times of 200−400, 150−400, 100−400, and 50−400 ps, respectively. In general, the calculated order parameters are not sensitive to the length of the trajectory used in our calculation, as can be seen in Figure 3. Residues with transitions such as those shown in Figure 2 for Ile15 are calculated using a part of the trajectory in which the reorientational correlation function converged.

Results and Discussion

Based on our MD trajectory, NMR order parameters (S^2) and crystallographic *B*-factors are calculated. Calculation of the S^2 parameters require a two-step analysis due to the rapid time scale probed by the NMR measurements. Figure 2 shows two different cases that result from the MD simulations. In the first case, represented by a transition plot for Gly 27, the orientation of the NH bond vector is relatively constant throughout the simulation. For residues corresponding to these cases, the order parameters are calculated on the basis of the entire 350-ps trajectory. For most of the residues that fall into this case, the agreement between theory and experiment is excellent. The second case, represented by a transition plot for Ile15, results from multiple orientations of the NH bond vector. Residues falling into this class switch between distinct minima during the course of the trajectory. Order parameters calculated from the entire trajectory result in values much lower than the experimentally determined values, as seen in earlier works by other researchers.[14,15] Therefore, following procedures similar to earlier works,[14,15] the MD-derived order parameters are calculated from the converged part of the trajectory (e.g., 200−

Table 1. Calculated Order Parameters for the KNI-272/HIV-1pr Complex for the Main-Chain NH Groups[a]

res	S^2	res	S^2	res	S^2	res	S^2
1A	−1.000	51A	0.879	1B	−1.000	51B	0.886
2A	0.849	52A	0.745	2B	0.696	52B	0.686
3A	0.683	53A	0.750	3B	0.788	53B	0.820
4A	0.861	54A	0.857	4B	0.871	54B	0.827
5A	0.900	55A	0.851	5B	0.898	55B	0.857
6A	0.891	56A	0.845	6B	0.876	56B	0.864
7A	0.844	57A	0.911	7B	0.822	57B	0.904
8A	0.884	58A	0.788	8B	0.868	58B	0.799
9A	−1.000	59A	0.855	9B	−1.000	59B	0.904
10A	0.849	60A	0.795	10B	0.806	60B	0.762
11A	0.719	61A	0.737	11B	0.845	61B	0.819
12A	0.833	62A	0.823	12B	0.813	62B	0.812
13A	0.853	63A	0.854	13B	0.827	63B	0.835
14A	0.765	64A	0.879	14B	0.795	64B	0.863
15A	0.858	65A	0.867	15B	0.749	65B	0.830
16A	0.774	66A	0.903	16B	0.808	66B	0.864
17A	0.762	67A	0.888	17B	0.747	67B	0.851
18A	0.787	68A	0.782	18B	0.800	68B	0.793
19A	0.714	69A	0.659	19B	0.806	69B	0.756
20A	0.816	70A	0.794	20B	0.837	70B	0.767
21A	0.885	71A	0.869	21B	0.809	71B	0.746
22A	0.809	72A	0.843	22B	0.792	72B	0.885
23A	0.895	73A	0.759	23B	0.863	73B	0.850
24A	0.918	74A	0.837	24B	0.919	74B	0.886
25A	0.929	75A	0.811	25B	0.910	75B	0.854
26A	0.887	76A	0.901	26B	0.927	76B	0.897
27A	0.911	77A	0.828	27B	0.932	77B	0.898
28A	0.904	78A	0.912	28B	0.869	78B	0.917
29A	0.863	79A	−1.000	29B	0.899	79B	−1.000
30A	0.797	80A	0.798	30B	0.867	80B	0.820
31A	0.798	81A	−1.000	31B	0.870	81B	−1.000
32A	0.850	82A	0.886	32B	0.900	82B	0.859
33A	0.908	83A	0.885	33B	0.895	83B	0.870
34A	0.882	84A	0.894	34B	0.857	84B	0.926
35A	0.846	85A	0.909	35B	0.711	85B	0.893
36A	0.768	86A	0.888	36B	0.852	86B	0.880
37A	0.799	87A	0.930	37B	0.739	87B	0.938
38A	0.580	88A	0.895	38B	0.700	88B	0.905
39A	−1.000	89A	0.898	39B	−1.000	89B	0.910
40A	0.578	90A	0.919	40B	0.655	90B	0.920
41A	0.823	91A	0.922	41B	0.824	91B	0.927
42A	0.855	92A	0.862	42B	0.858	92B	0.858
43A	0.866	93A	0.876	43B	0.880	93B	0.869
44A	−1.000	94A	0.846	44B	−1.000	94B	0.834
45A	0.874	95A	0.864	45B	0.855	95B	0.878
46A	0.886	96A	0.884	46B	0.887	96B	0.909
47A	0.820	97A	0.877	47B	0.868	97B	0.902
48A	0.842	98A	0.872	48B	0.879	98B	0.874
49A	0.741	99A	0.810	49B	0.856	99B	0.820
50A	0.864			50B	0.757		

[a] Values of −1 indicate a proline. Monomers A and B are listed separately.

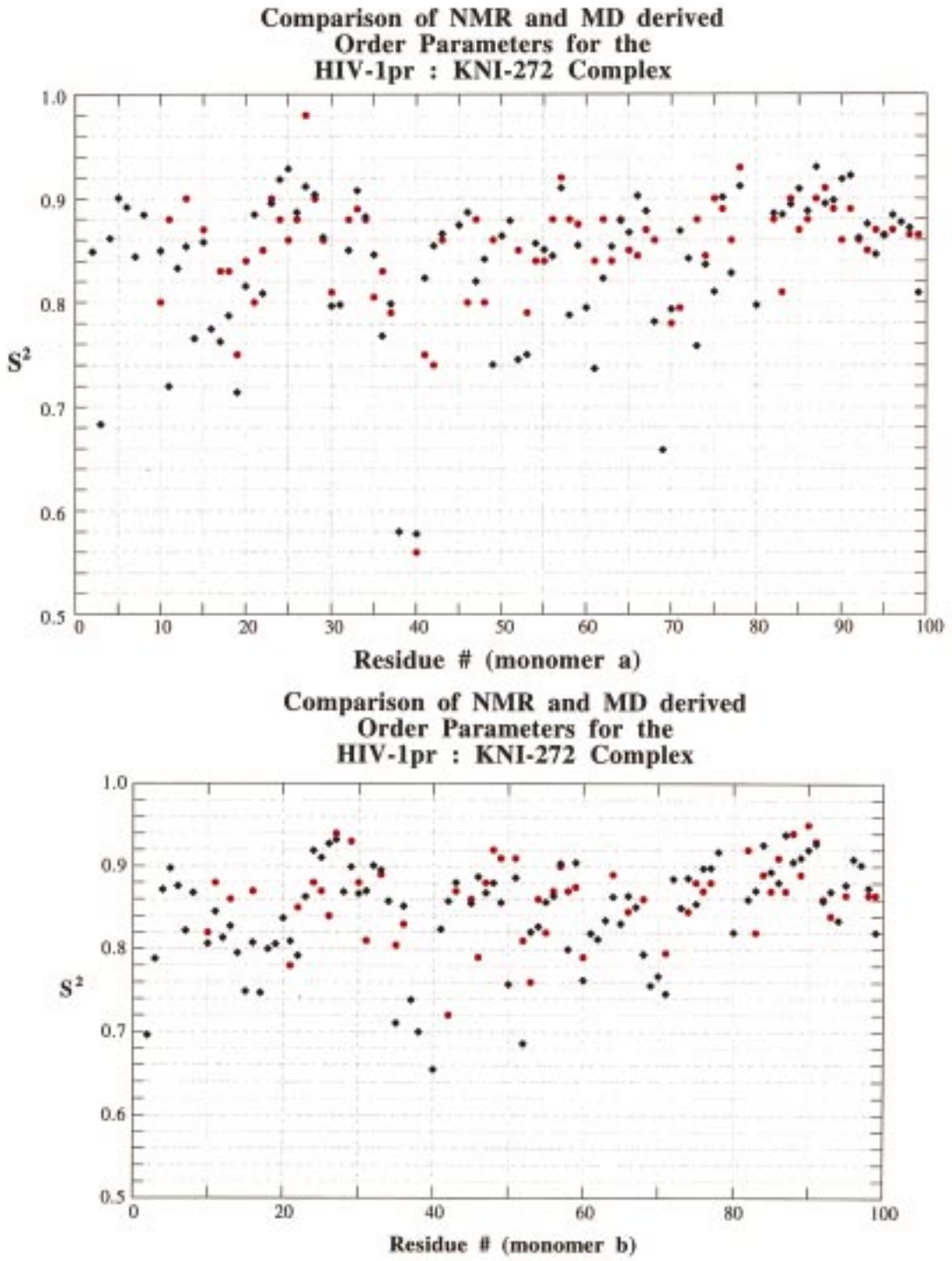

Figure 4. Comparison of the MD (blue diamonds)- versus NMR (red circles)-derived order parameters for the KNI-272/HIV-1pr complex. (A, top) Residues 1−99, corresponding to monomer A, and (B, bottom) monomer B.

400ps for Ile15). The final calculated order parameters (Table 1) are compared with those experimentally derived from NMR in Figure 4. The overall agreement between the calculated and NMR values is excellent (see Figure 5), with an average absolute error of 0.042 and an RMS deviation of 0.055. The magnitude of the errors can be compared to the NMR results where only intermonomer differences greater than 0.09 were considered to be statistically significant.[7] The crystallographic *B*-factors are calculated using standard formulas and procedures after exploring different protocols for removing the protein translation and rotation. The calculated *B* factors display an overall agreement with the experimental values (shown in Figure 6), with trends consistent with the NMR results. Notable differences between the crystallographic and MD-derived *B*-factors occur around residues Gly16 and Gly17 in each of the two monomers. These differences can be explained in terms of crystal contacts that are included in the X-ray data but are absent in the MD simulations. Residues 16 and 17 each have intermolecular contacts within the crystal lattice where the Cα−Cα are less than 6 Å. In fact, the main-chain carbonyl oxygen of Gly16 is only 3 Å from Pro39 in the neighboring molecule. These contacts are expected to reduce motion in the crystal that may

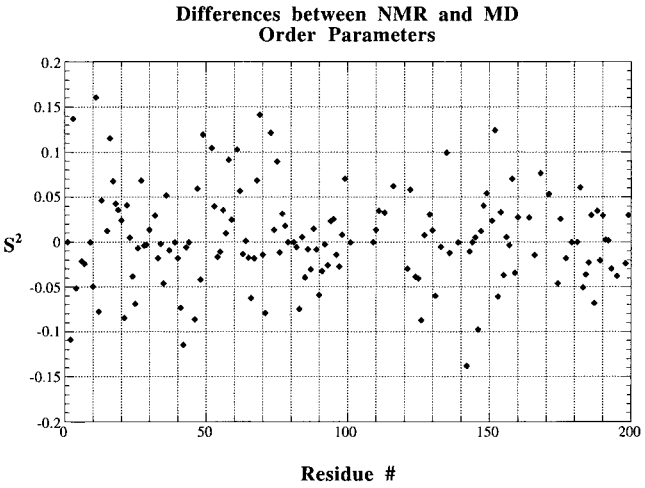

Figure 5. Differences between the NMR and MD order parameters. The RMSD for all of the values is 0.055, and the average absolute error is 0.042. Negative values correspond to the cases where the MD values are larger than those from NMR.

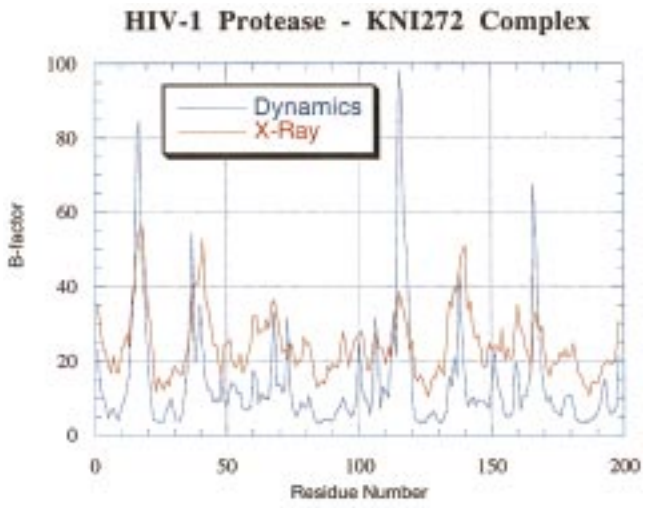

Figure 6. *B*-Factors of the Cα atoms from MD (blue) and the 1HPX crystal structure (red) are compared. Regions of the largest deviation (near residues 16 and 17) are involved in close interdimer contacts in the crystal structure and are expected to be significantly depressed compared to the solution-phase simulation.

be inherent in the free protein. Figure 7 shows a color-coded Cα trace of the calculated flexibility of the protein; Figure 7A is based on *B*-factors, and Figure 7B corresponds to the order parameters. Residues in contact with the inhibitor are generally well ordered as revealed in both the order parameters and the *B*-factors, in agreement with experiment.

Comparison with the NMR results[7] reveals that NMR and MD agree that the greatest dynamic motion in the protein, as measured by the order parameters, is in the flap elbows around residues 37−42. Our MD-derived order parameters also suggest a more flexible area in the flap region of monomer B than the NMR indicates. This cannot be directly compared to experiment, however, since the NMR order parameters for many residues around the tips of the flaps are missing. Our MD-derived order parameters suggest that the flap tip residues Ile50A, Gly51A, and Gly51B are well ordered, while Ile50B is more dynamic. Previous NMR studies of the DMP323−HIVpr complex have speculated that the flaps are involved in a conformational change between β-I and β-II type turns.[8] Despite the greater flexibility in monomer B, we find no direct evidence for a conformational change after inspection of our trajectory for the KNI-272-HIVpr complex. Our calculated order parameters indicate that Gly52 is less ordered in both monomers than other residues around the tips of the flaps. This flexibility at Gly52 may be required for protein function since any mutation at this position renders the protein inactive.[30]

It was originally expected that the dynamics of the active site residues in the two monomers would be different due to the asymmetric nature of KNI-272. The calculated order parameters and *B*-factors fail to reveal this and are remarkably similar between the two monomers. Our calculated results suggest that the flexibility of the two chains of the protein backbone is only slightly perturbed by the asymmetric inhibitor, in agreement with experiment. This result is quite surprising given the nature of the interactions between the inhibitor and the enzyme active site, and the fact that the inhibitor has no P3′ residue[31] to interact with the protein. Figure 8 shows a plot of the differences between monomers A and B for both the order parameters and *B*-factors derived from MD. Based on the calculated results of the backbone atoms, we speculate that the asymmetry of the inhibitor may be more apparent in the side-chain dynamics than in the main chain of the protein.

To investigate the dynamics of the inhibitor and the side chains of the protein, we calculated the order parameters of the protein side chains and the inhibitor atoms. These results are given in Tables 2−4. Our primary focus is on those residues that come into close contact with the bound inhibitor. The simulations reveal a distinct asymmetry in the order parameters of the methyl hydrogens for Val32, Ile50, and Ile84 (Table 2). The calculated order parameters suggest that both Ile84A and Ile84B undergo considerable motion during the simulation. It is not clear, however, whether the motion is due only to methyl hydrogen rotation or if the heavy atoms themselves are moving. The motion of the C−C bond vectors separates out these motions and provide a more detailed picture of the dynamic nature of these side-chain residues. These results are given in Table 3. The order parameters calculated from the C−C bond vectors reveal significant motion of the carbon atoms in Ile84B but not in Ile84A. The *B*-factors from simulation, and to a much lesser extent from experiment, reveal greater motion in Ile84B than Ile84A as well. The reason for these differences probably lies in the observation that residue Ile84B is in contact with the thioproline ring of the inhibitor and water 607, whereas Ile84A is much more ordered due to interactions with the APNS and TBU groups of KNI-272. Figure 9 shows the time course of the Ile84B side-chain dihedral angles and reveals a transition between two different conformational states. Given the magnitude of the motion revealed from the MD simulation in the Ile84B side chain, we expect compensatory motion in the inhibitor as well.

The backbone N−H order parameters of the inhibitor indicate that the inhibitor is quite well ordered and displays little flexibility (see Table 4). However, upon further investigation, the side-chain bond vectors reveal quite low order parameters, indicating considerable flexibility of the inhibitor side chains despite the rather rigid mainchain (Table 4). In contrast, the *B*-factors from X-ray and simulation are not significantly larger than those of nearby atoms. The methylthioalanine side chain

(30) Shao, W.; Everitt, L.; Manchester, M.; Loeb, D. D.; Hutchison, C. A., III; Swanstrom, R. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 2243−2248.

(31) Schechter, I.; Berger, A. *Biochem. Biophys. Res. Commun.* **1967**, *27*, 157−162. Nomenclature from this paper is used to describe the inhibitor−protein interactions. P_n refers to residues on the N-terminal side of the inhibitor starting at the site of the scissile bond in a typical substrate. P_n' refers to residues on the C-terminal side of the inhibitor with respect to the scissile bond in a typical substrate.

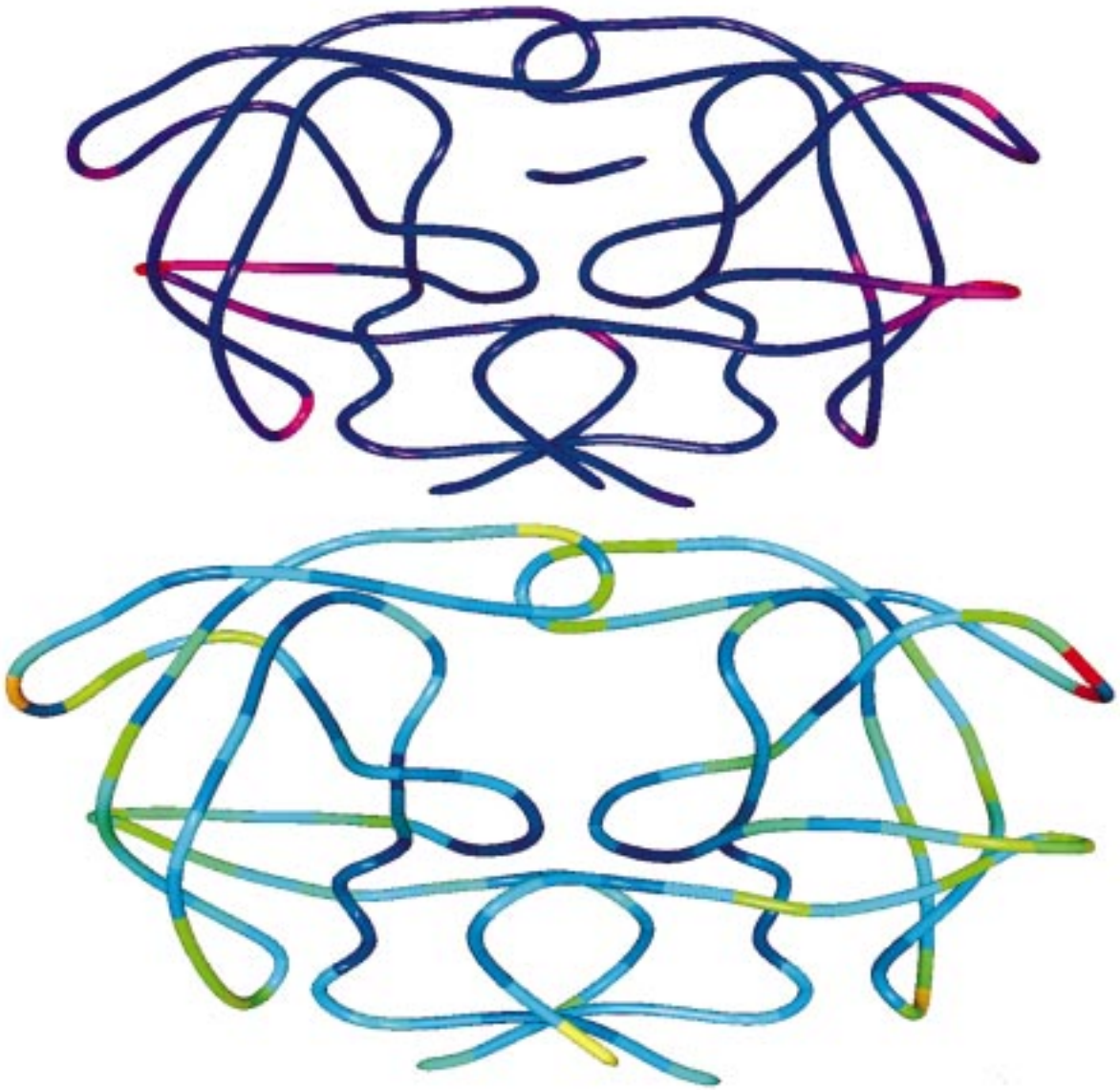

Figure 7. Color coded Cα trace of the *B*-factors (A, top) and the order parameters (B, bottom) calculated from the MD simulation. Regions of the greatest mobility are colored red and those with the least motion are colored blue. It is immediately apparent that the regions away from the binding site of the inhibitor are the most dynamic.

changes conformation several times during the course of the trajectory. This is displayed graphically in Figure 10. The dihedral angle between the Cβ and the Sδ atoms changes dramatically at approximately 100 ps and oscillates up to 100° many times during the remainder of the simulation. It is clear that there are at least two low-energy conformational states for the methylthioalanine that contribute to its interactions with the protease. Proper account of the energetics and geometric analysis needs to include both of these conformations in order to model or compare the inhibitory efficacy of KNI-272, and possibly other molecules with similar structural features, in any drug design efforts. In contrast to the methylthioalanine group, the *tert*-butyl C-terminus of the inhibitor is relatively stable during the course of the trajectory. We originally expected to see considerable motion in the end regions but little motion near the closely packed residues near the catalytic aspartates. Our simulations reveal, however, that regions of the inhibitor important to binding and in direct contact with the protein may be quite flexible. Our simulations also suggest that the associated water molecules may be considered a relatively rigid part of the protein in certain circumstances.

The two tightly bound water molecules (301 and 607) associated with the active site are extremely static and show no significant deviations from their positions or orientations throughout the simulation (see Table 4). This is in contrast to previously reported simulations on the KNI-272/HIV-1pr complex[32] which considered only the flap water in the simulation. Our results suggest that these two water molecules may play a significant role in the binding interactions between KNI-272 and HIV-1pr and should be explicitly included in all studies attempting to account for the energetics and geometry of binding in this and similar complexes.

Conclusions

The dynamic properties of the KNI-272/HIV-1 complex have been simulated, resulting in calculated order parameters and *B*-factors that are in excellent agreement with the available experimental data based on order parameters from a parallel NMR study of the same inhibitor−protein complex as well as

(32) Kato, R.; Takahashi, O.; Kiso, Y.; Moriguchi, I.; Hirono, S. *Chem. Pharm. Bull. (Tokyo)* **1994**, *42*, 176−178.

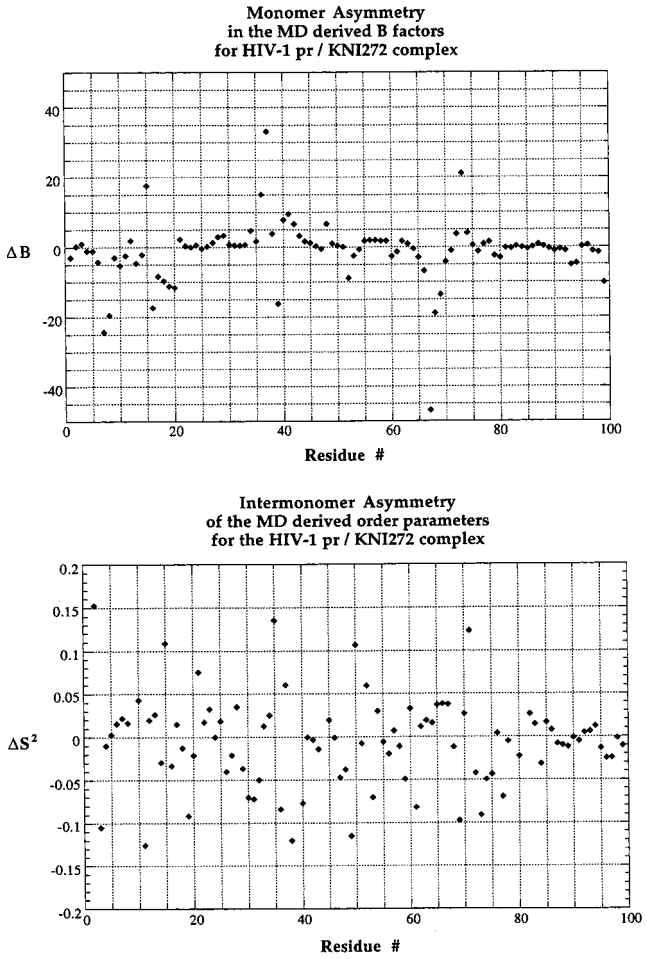

Figure 8. Plots showing the asymmetry between monomers A and B in the HIV-1pr/KNI-272 complex as measured by the calculated *B*-factors (top) and the MD-derived order parameters (bottom). Residues displaying the greatest differences between monomers are in close contact with other molecules within the crystal.

Table 2. Calculated Order Parameters of Selected HIV-1pr Protein Side-Chain Proton-Bound Vectors[a]

bond	S^2	contacts with KNI272 <4.5 Å
V32A Cγ1−H	(0.80,0.86,0.81)	TBU
V32A Cγ2−H	(0.49,0.49,0.46)	TBU
V32B Cγ1−H	(0.51,0.48,0.50)	
V32B Cγ2−H	(0.62,0.68,0.58)	MTA
I84A Cγ1−H	(0.61,0.62)	APN
I84A Cγ2−H	(0.71,0.80,0.73)	TBU
I84A Cδ1−H	(0.15,0.17,0.12)	APN, TBU
I84B Cγ1−H	(0.27,0.32)	THZ, WAT607
I84B Cγ2−H	(0.09,0.07,0.09)	THZ, WAT607
I84B Cδ1−H	(0.42,0.39,0.42)	THZ, MTA
I50A Cγ1−H	(0.55,0.53)	MTA
I50A Cδ1−H	(0.16,0.24,0.12)	MTA
I50A Cγ2−H	(0.88,0.87,0.87)	MTA, WAT301
I50B Cγ1−H	(0.84,0.82)	APN, THZ, TBU, WAT301
I50B Cδ1−H	(0.37,0.34,0.38)	APN, TBU
I50B Cγ2−H	(0.32,0.30,0.30)	TBU

[a] Each proton of the methyl or methylene group is listed (in parentheses). Also listed are the inhibitor residues within 4.5 Å of the carbon atom defining the bond vector. Abbreviations for contacts are defined in Figure 1.

the Debye−Waller temperature factors from X-ray crystallography. Further, several insights not amenable to experiment have emerged from analysis of the MD data. From our analysis, we observe bond vector motion that was not apparent from the *B*-factors derived from crystallographic studies nor from order

Table 3. Calculated Order Parameters for Heavy Atom Bond Vectors of Side Chains in Contact with the Inhibitor KNI-272 Complexed with HIV-1pr

	chain	
	A	*B*
I50 Cβ−Cγ2	0.91	0.93
I50 Cβ−Cγ1	0.92	0.91
I50 Cγ1−Cδ1	0.54	0.80
V32 Cβ−Cγ1	0.89	0.86
V32 Cβ−Cγ2	0.90	0.87
I84 Cβ−Cγ1	0.89	0.35
I84 Cβ−Cγ2	0.89	0.31
I84 Cγ1−Cδ1	0.73	0.81

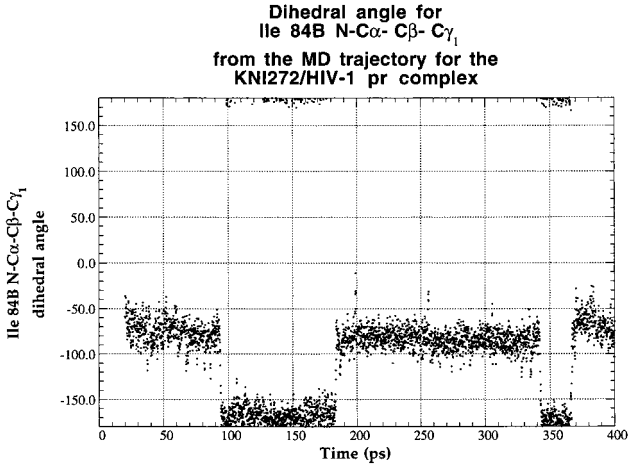

Figure 9. Plot of the χ1 dihedral angle of Ile84B during the course of the MD trajectory. Ile84B shows sharp conformational transitions, in contrast to Ile84A, and may be correlated with transitions seen in the side chain of the inhibitor fragment MTA (see Figures 1 and 10).

Table 4. Calculated Order Parameters for the Inhibitor KNI272 and the Nearby Water Molecules[a]

bond	S^2
Backbone	
MTA (N−H)	0.79
APW (N−H)	0.91
TBU (N−H)	0.80
Side Chains	
MTA Cβ(C−H)	(0.79,0.80)
MTA Cϵ(C−H)	(0.04,0.11,0.08)
TBU C1 (C−H)	(0.41,0.38,0.41)
TBU C2 (C−H)	(0.86,0.59,0.63)
TBU C3 (C−H)	(0.85,0.63,0.62)
MTA (Sδ−Cϵ)	0.38
Water	
301(O−H)	(0.85,0.82)[b]
607(O−H)	(0.76,0.81)

[a] All hydrogen atoms belonging to a methyl or methylene group or water molecule are listed in parentheses. The water molecule forming a bridge between the inhibitor and the flaps of the protein is specially designated due to its importance. [b] Flap water.

parameters based on NH vectors. Specifically, the MD simulations reveal asymmetric motions in the KNI-272 complex that are not readily visible from the NMR or X-ray data. This suggests that detailed motion involving protein−ligand complexes must be derived from several different experimental and computational techniques. We speculate that NMR experiments employing ^{13}C-labeled inhibitor and amino acids may observe some of the motions found in our simulations. We also observed two water molecules (301 and 607) with high order

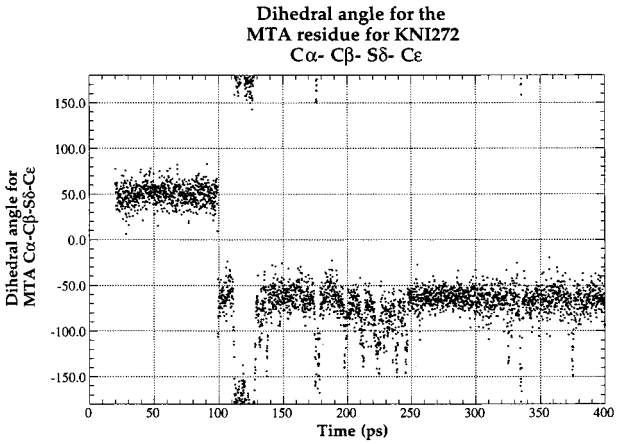

Figure 10. Plot of the $\chi 2$ dihedral angle of the MTA group in KNI-272 during the course of the MD trajectory. The largest transition, near 100 ps, occurs around the same time that Ile84B changes conformation from its initial geometry. During the last 300 ps of the trajectory, the $C\beta-S\delta$ dihedral angle varies by up to 150°, indicating significant conformational flexibility not seen in the *B*-factors or backbone order parameters.

parameters throughout the simulation, in contrast to previous studies. This indicates that water must be explicitly included in simulations of protein−ligand complexes where the flexibility and energetics of the system are of interest. Our study further demonstrates that, in a situation where the experimental study is difficult, such as that with the native enzyme of HIV-1 protease, dynamics simulation could be a good alternative to NMR spectroscopy and crystallography in providing information about the protein−ligand dynamics.

Acknowledgment. We thank Drs. Daron Freedberg and Dennis Torchia (located at the FDA and NIH, respectively) for sharing their NMR results prior to publication and engaging in many fruitful and substantive discussions. The authors also thank Drs. Stan Burt and Brian Luke of the FBSC for reading the manuscript for clarity and content. X.L. was supported by a grant from the NIH to Smith-Kline Beecham (GM50579). This project has been funded in whole or in part with Federal funds from the National Cancer Institute, National Institutes of Health, under Contract No. NO1-CO-56000. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

JA9824066