Method for Including the Dynamic Fluctuations of a Protein in Computer-Aided Drug Design

Heather A. Carlson,* Kevin M. Masukawa, and J. Andrew McCammon

Department of Chemistry and Biochemistry, Department of Pharmacology, University of California, San Diego, 9500 Gilman Drive, La Jolla, California 92093-0365

We have recently presented a new pharmacophore design method that allows for the incorporation of the inherent flexibility of a target active site. The flexibility of the enzymatic system is described by collecting many conformations of the uncomplexed protein; this ensemble of conformational states can come from a molecular dynamics (MD) simulation, multiple crystal structures, or many NMR structures. Binding sites for functional groups that complement the active site are determined through multiple-copy calculations. These calculations are conducted for each protein conformation, providing a large collection of potential binding sites. The Cartesian coordinates from each protein conformation are overlaid through RMS fitting of essential catalytic residues, and the pharmacophore model is described by binding regions that are conserved over many protein conformations. Previously, we developed a "dynamic" pharmacophore model for HIV-1 integrase using 11 conformations of the protein from an MD simulation; the MUSIC procedure was used to calculate binding positions for methanol molecules in each configuration of the active site. Here we present "static" pharmacophore models developed with a single conformation of the protein from two new crystal structures (standard protocol for multiple-copy methods). The static models do not perform as well as the previous dynamic model in fitting known inhibitors of HIV-1 integrase. To test the applicability of the dynamic pharmacophore method and the assumption that any reliable source of protein conformations is applicable, we have now developed a second dynamic pharmacophore model based on the two crystal structures also used for the development of the static models. Though the dynamic model based on the two crystal structures does not fit as many known inhibitors as the previous dynamic model, it is a significant improvement over the static models. Even better performance is expected with the addition of new crystal structures as they become available. However, it is notable that using only two structures leads to great improvement in the models.

Introduction

Unresolved segments of protein crystal structures can be very problematic when using those structures in computer modeling. We were recently faced with such a problem in our studies of HIV-1 integrase.[1,2] Until only recently, all available crystal structures of the catalytic core were incomplete with unresolved flexible loops. Particularly problematic was the fact that the largest section of missing structure was a flexible loop adjacent to the active site. With a portion of the active site environment unresolved, our efforts in computer-aided inhibitor design were greatly hindered. To overcome this, the missing loops were modeled on the basis of the conformation of the homologous loops in the crystal structure of the integrase from Avian Sarcoma Virus.[3] This completed structure was then used to initiate two 1 ns molecular dynamics (MD) simulations of the core domain, with and without a catalytic metal ion present in the active site.[1] Conformations from these simulations were used in the development of the "dynamic" pharmacophore model for HIV-1 integrase.[2]

One of our goals in studying the integrase is to develop receptor-based pharmacophore models to identify compounds that complement the active site. Given the uncertainty in the homology-modeled loop adjacent to the active site, it was important to develop a new method that could reduce the propagation of errors in an individual protein conformation into the resulting pharmacophore model. By using multiple structures of the protein, an "averaged" picture emerges, potentially focusing the design of inhibitors to the most important features of the active site. Current methods to develop receptor-based models usually rely on a single representation of the protein conformation.[4−12] Our previous study demonstrated that a "static" pharmacophore model, based on the crystal structure used to initiate the MD study, exhibited a poorer performance than the dynamic model when fitting known inhibitors of the integrase.[2]

HIV-1 integrase is an interesting test case for developing dynamic methods because it has an active site that is shallow, solvent exposed, and minimally restricted in conformational sampling. Also, the flexibility of the active site loop is required for catalytic activity; this must be incorporated into any reliable pharmacophore model of the system.[13] Studies have demonstrated the need for protein flexibility in ligand docking and multiple-copy simulations in order to achieve proper results;[14−17] however, the large majority of publications continue to report the use of only one static structure for multiple-copy simulations.[4,5,7−12] Because dynamic behavior is very important for regulating function in many protein systems, this method has

* E-mail: hcarlson@chemcca10.ucsd.edu. Phone: (858) 822-1469. Fax: (858) 534-7042.

10214 *J. Phys. Chem. A, Vol. 103, No. 49, 1999*

Carlson et al.

been proposed to accommodate such information in drug design. The goal of developing a dynamic pharmacophore model is to identify compounds that complement the protein while causing minimal disruption of the conformation and flexibility of the active site, potentially reducing the entropic penalties[4,5] incurred by the protein upon binding a ligand. While rigid versions of the ligands can be synthesized to reduce such penalties on the part of the ligand, this is the first method introduced to reduce similar entropic penalties incurred by the receptor. Multiple configurations of a protein receptor can be obtained from an MD simulation, multiple crystal structures, or NMR structures[5,14] to describe the inherent flexibility of the active site. These methods each provide the uncomplexed receptor under the influence of explicit solvent molecules. Binding sites for various functional groups within each receptor configuration can be determined with calculations employing multiple-copy methods.

In our first study,[2] the multi-unit search for interacting conformers (MUSIC) method was used to determine the binding sites for hydrogen-bond donating groups within the active site of multiple protein configurations provided from an MD simulation.[1] The MUSIC procedure is a Monte Carlo (MC) simulation that simultaneously calculates multiple, gas-phase minimizations for hundreds of probe molecules within the active site. The many configurations of the protein are then overlaid to reveal conserved binding sites that are highly occupied over the course of the MD simulation despite the motion of the active site. These conserved binding sites define a pharmacophore model for inhibitors that should in theory bind to the active site and still allow for almost the same flexibility.

In this study, we examine the possibility of using the multiple crystal structures to create a second dynamic pharmacophore model (Dynamic-Crystal). Shortly after the completion of the MD simulations,[1] crystal structures became available for the complete catalytic core of HIV-1 integrase with resolved flexible regions.[18,19] Two static models are presented, each based on only one of the two available crystal structures. The Dynamic-Crystal pharmacophore was developed by combining the data for the two static models.

Computational Details and Results

Preparation of the Protein Structures. As the dynamic model from the MD simulations has been described elsewhere,[2] we will limit this discussion to the development of pharmacophore models based on the crystal structures. For comparison, the dynamic model based on the MD simulations (Dynamic-MD) is provided in Table 1 with the models derived from the crystal structures. To test the models, the Catalyst program[20] was used to fit a set of known inhibitors to the pharmacophore models; see Table 2.

In the original MD simulations, the system was fully solvated with explicit water molecules and a divalent ion was present in the active site.[1] The ion and all solvent molecules were removed to provide a "bare" active site for the MUSIC studies. All ionization states were kept the same in the MUSIC studies except that H66 (adjacent to the active site) was modeled as protonated, which is its charge state in the absence of the divalent ion.[2] Also, the N- and C-termini of the catalytic domain were modeled as neutral methyl amides to avoided any false minima associated with the terminal charges at residues 57 and 210, since they are not present in the full integrase enzyme. The pepz utility[27] available with the BOSS[28] and MCPRO[29] programs was used to add hydrogens to the crystal structures of the catalytic core from Maignan et al.[18] (monomer C of 1BI4 in the protein data bank) and Goldgur et al.[19] (monomer B in 1BIS).

MUSIC. In 1991, Miranker and Karplus introduced multiple-copy methods[6] based on ideas presented with the development of the GRID method.[30] The method was quickly embraced and is now a standard practice in computational drug design.[4,5,7,31−34] MUSIC is the multiple-copy method that is employed in the development of the dynamic and static pharmacophore models; it is a procedure available in the Monte Carlo program BOSS.[28] Methanol molecules have been used to describe binding sites that complement the catalytic residues D64 and D116 in the active site. The probe molecules and the protein were described with the all-atom OPLS force field.[35,36] Large cutoff radii (50 Å) were used so that no nonbonded interactions were neglected. The protein configurations were held rigid because the flexibility was represented through the use of multiple configurations of the protein; however, it should be noted that the side chains can be conformationally sampled within the MUSIC procedure.

The protein structures were not rotated or translated during the MUSIC studies, and the methanol probe molecules were held internally rigid. Sampling moves for the probes were limited to displacements of 0.15 Å or less and rotations of 15° or less. The system was initiated with a 17.0 Å sphere of densely packed methanol molecules centered at the active site (backbone nitrogen of Q62 in the Goldgur et al. structure and $O\epsilon$ of D64 of the Maignan et al. structure), resulting in 245 and 383 copies of methanol, respectively. A half-harmonic potential was applied at the boundary of the 17.0 Å sphere (force constant of 5 kcal/(mol·Å^2)) to keep the probes from possibly dissociating from the protein. The following simulated annealing protocol was used to minimize the hundreds of probe molecules: 10^6 iterations of MC sampling at 300 °C, 10^6 iterations at 200 °C, 10^6 iterations at 100 °C, 10^6 iterations at 0 °C, and finally 2×10^6 iterations at −100 °C.

Pharmacophore Models. MUSIC calculations with methanol molecules were used to identify binding sites for functional groups to complement the active site. The hundreds of probes clustered into many local minima within the receptor; given that hydroxyl groups can be hydrogen-bond donors and acceptors, their orientation relative to the active site residues could dictate either type of interaction site. Because the essential residues of HIV-1 integrase are carboxylates, only hydrogen-bond donor sites were identified (MUSIC calculations with acetone probe molecules were also used to identify hydrogen-bond acceptor sites, but none were located close enough to the essential residues to be incorporated into the pharmacophore models). The centers and radii of the hydrogen-bond donor sites were calculated from the Cartesian coordinates of the oxygen atoms of all the methanol probes in an individual cluster. The center of each hydrogen-bond donor site in the pharmacophore model was equal to the average position of the methanol oxygens in each site. The radii of the hydrogen-bond donor sites were set to double the RMS deviation of the oxygen coordinates. The second and third models in Table 1 are from the crystal structures of Maignan et al.[18] and Goldgur et al.,[19] respectively, Static Model 1 and Static Model 2.

For the Dynamic-Crystal pharmacophore model, the Midas-Plus[37,38] program was used to overlay the MUSIC results for the two protein structures by an RMS fit of the $C\beta$, $C\gamma$, and $O\delta$ atoms of the essential residues, D64 and D116, in both conformations of the protein. In the development of the dynamic pharmacophore model based on the MD simulations, only conserved regions with probes from several protein configurations were used. However, only two binding sites are conserved regions when the two crystal structures are overlaid (HBd1 of

Dynamic Fluctuations of a Protein

J. Phys. Chem. A, Vol. 103, No. 49, 1999 **10215**

TABLE 1: Dynamic Pharmacophore Model Based on the MD Simulations and the Static and Dynamic Pharmacophore Models Based on the Crystal Structures[e]

	Radius	X	Y	Z (Å)
Dynamic Model from MD structures				
HBdonor1	*1.316*	7.249	4.783	2.405
HBdonor2	*1.088*	9.543	4.462	1.849
HBdonor3	*0.388*	11.037	3.382	1.497
HBdonor4	*1.106*	10.877	1.845	-2.410
HBdonor5	*1.088*	8.232	7.622	-1.851
HBdonor6	*1.494*	9.072	8.404	-4.178
Q62	*1.5*	5.842	4.291	-3.330
D64	*1.5*	7.819	2.218	0.311
D116	*1.5*	10.461	5.219	-2.795
Static Model 1 from crystal structure of Maignan et al.[a]				
HBdonor1	*0.464*	6.346	-4.762	10.626
HBdonor2	*0.591*	9.455	-4.219	14.306
HBdonor3	*1.511*	11.052	-2.038	9.695
HBdonor4	*0.804*	11.056	1.222	9.286
O of L63 [b]	*1.5*	7.763	0.359	16.393
D64	*1.5*	7.409	-2.046	12.892
D116	*1.5*	12.001	0.425	12.406
Static Model 2 from crystal structure of Goldgur et al.[c]				
HBdonor1	*0.920*	-7.970	2.051	-3.353
HBdonor2	*0.470*	-9.712	-0.847	-4.840
HBdonor3	*0.740*	-9.794	-2.587	0.591
HBdonor4	*1.008*	-10.858	-6.280	-5.223
Q62	*1.5*	-8.909	-0.797	4.047
D64	*1.5*	-8.019	-1.392	-1.845
D116	*1.5*	-11.446	-5.914	-1.851
Dynamic Model from the crystal structures[d]				
HBdonor1	*1.351*	4.405	1.945	29.259
HBdonor2	*0.470*	1.942	2.140	30.504
HBdonor3	*1.511*	-0.572	0.994	32.099
HBdonor4	*0.884*	-3.487	1.825	30.844
HBdonor5	*0.591*	2.662	-2.515	29.691
HBdonor6	*0.740*	0.603	-2.958	28.327
Q62	*1.5*	2.708	-5.740	26.390
O of L63	*1.5*	-0.254	-2.232	25.267
D64	*1.5*	1.668	0.013	27.808
D116	*1.5*	-2.926	-1.454	30.359
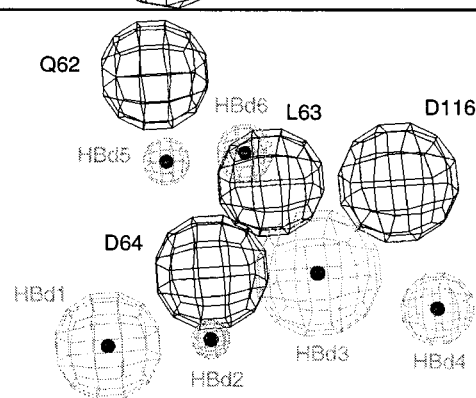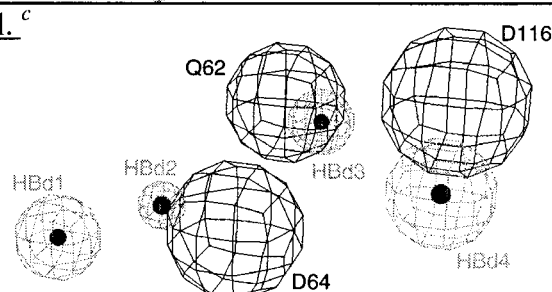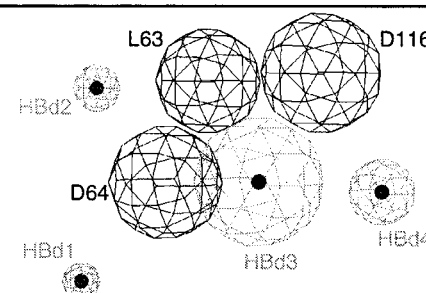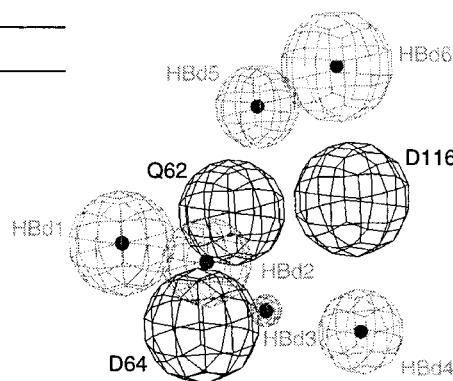
[a] Reference 18. [b] The side chain of Q62 is in an orientation away from D64 and D116 in the structure by Maignan et al., so the carbonyl oxygen of L63 was chosen to represent the bottom of the active site. [c] Reference 19. [d] Both L63 and Q62 are retained from the static models. The change in the coordinates reflects the change in origin and orientation when RMS-fitting the Oδ of D64 and D116, but similar views of the models are presented above for clarity. [e] Atoms with gray spheres are hydrogen-bond donors; the black spheres are excluded volumes based on active site residues.

Static Model 1 and Static Model 2, HBd4 of Static Model 1 and Static Model 2). If the model were to be restricted to only conserved binding regions, as in the dynamic model based on the MD simulations, it would result in a model with only two hydrogen-bond donor sites. A two-site model is not specific enough for drug design, so the two static models were simply combined and the conserved regions treated as larger binding sites with the centers and radii recalculated.

On the basis of the Cartesian coordinates from the overlays, the average position for Cγ of D64 and Cγ of D116 were used as the centers for two excluded volumes with radii of 1.5 Å in the pharmacophore model. Two additional excluded volumes were added at the relative positions of the backbone O of L63 in Static Model 1 and the Cδ of Q62 in Static Model 2 to represent the bottom of the active site in the two crystal structures. The excluded volumes in the static and dynamic

TABLE 2: Performance of the Pharmacophore Models Tested against Compounds from the Literature that Have Been Tested for Inhibitory Activity (Compounds Ordered by Inhibitory Activity)

Compound	IC$_{50}$s 3'/ST	Dynamic-MD	Static 1 Maignan et al.[a]	Static 2 Goldgur et al.[b]	Dynamic-crystal
Chicoric Acid[c]	0.15/0.13	X			
107[c]	0.23/0.11	X			X
4,5-DCQA[c]	0.25/0.46	X			
81[c]	0.4/0.2	X	X		X
67[c]	0.5				
71[c]	0.5	X			
85[c]	0.5/0.2	X	X		X
3,5-DCQA[c]	0.64/0.66	X			X
3,4-DCQA[c]	0.79/0.54	X			X
1,5-DCQA[c]	0.68/1.08	X			
NSC 118695[d]	0.9/0.3	X	X		X
Quercategetin[e]	1.3/0.6	X			
105[c]	0.98/0.81	X	X		X
NSC 64205[d]	1.1/0.5	X	X		X
NSC 158393[f]	1.5/0.8				
NSC 607319[d]	1.4/1.0	X	X	X	X
NSC 309121[d]	1.7/1.0	X	X		X
68[c]	1.7	X			X
Myricetin[e]	2/0.6	X			
Doxorubicin[c]	0.9/2.4	X	X		X
			Very Active Compounds		
Purpurogallin[g]	2.1		*Active Compounds*		
111[c]	2.3/1.1				
NSC 310217[f]	2/1.5				
NSC 64452[d]	1.2/3.6	X			X
Quinalizarin[e]	4/1				
83[c]	3.3/1.9	X			X
NSC 261045[d]	2.3/4.1	X	X		X
90[c]	1.38/4.71	X			
NSC 642710[f]	5.3/5.0	X	X		X
NSC 115290[h]	5	X			X
Ellagic Acid[g]	5.1				
115[c]	6.7/5.2				
Mitoxantrone[c]	3.8/8.0	X	X	X	X
Rosmarinic Acid[c]	9/4	X			X
Tyrophostin A51[e]	10/3				
110[c]	9.1/5.8				
1,4-DCQA[c]	9.5/7.8	X			X
141[c]	11.6/7.9	X			X
Hypericin[g]	10				
TMS[g]	17/5				
NSC 318213[d]	23.9/14.0				
66[c]	21.4/5.4	X			X
NSC 233026[d]	20.6/19.7				
NSC 371056[f]	29.9/16.5	X	X	X	X
NSC 48240[f]	26/20.6	X			X
97[c]	33/33	X	X	X	X
			Active Compounds		
Chlorogenic Acid[c]	87.8/45.8	X	*Ineffective Compounds*		
13[h]	120/96				
NSC 641547[f]	224/134	X			
NSC 674503[d]		X			
NSC 635971[f]		X			X
NSC 642651[d]		X	X		X
112[c]		X	X		X
9[h]		X			X
52[h]		X	X	X	X
NSC 281311[d]		X	X		X
103[c]		X	X	X	X
5[f]					
NDGA[g]		X			X

[a] Reference 18. [b] Reference 19. [c] Preference 21. [d] Reference 22. [e] Reference 23. [f] Reference 24. [g] Reference 25. [h] Reference 26. The numbers listed for some of the compounds are the labels given in the referenced papers.

Dynamic Fluctuations of a Protein

J. Phys. Chem. A, Vol. 103, No. 49, 1999 **10217**

models were used to eliminate compounds likely to have steric conflicts with the protein. Though the binding regions are specifically for hydroxyl groups, the criteria were extended for the pharmacophore model to include any oxygen, nitrogen, or sulfur that could donate a hydrogen bond (bound to one, two, or three hydrogens). This resulted in a new dynamic pharmacophore model with 10 sites, four excluded volumes and six hydrogen-bond donor sites. The two static models both consist of three excluded volumes and four hydrogen-bond donor sites.

Using the Catalyst Program. For the test set of known inhibitors of HIV-1 integrase, 59 compounds from the literature[21−26] were built into a user-defined database through the 3-D viewer interface in the Catalyst program.[20] Structures of all the compounds are given as Supporting Information; readers are referred to that source and the cited sources in Table 2 for additional information about the compounds used in the test set. To test the selectivity of the pharmacophore models, noninhibitory compounds with structures similar to known inhibitors were included in the test set. A total of 20 compounds in the test set had IC_{50}'s under 1 μM for 3′ preprocessing or strand transfer (referred to as very active compounds in Table 2 and the following discussion). An additional 26 compounds had IC_{50}'s between 1 and 35 μM for both catalyzed processes (active compounds). Three ineffective inhibitors (IC_{50}'s of 46− 224 μM) and 10 noninhibitors were also included. All compounds chosen to test the pharmacophore models contained at least four hydrogen-bond donors. The inhibitors that are most likely to bind to the side chains of D64 and D116 must contain many hydrogen-bond donors to complement the carboxylate groups.

Conformations of the inhibitors were created with the fast conformational generator employing the default limits and the built-in force field used in Catalyst.[20] They were fit to the pharmacophore models using the "best" routine. The best routine allows for an additional cycle of fitting to refine the conformations of the compound to better fit the pharmacophore model. Results for fitting the inhibitors to the pharmacophore models are given in Table 2.

Discussion

We have previously reported the excellent performance of the Dynamic-MD pharmacophore model to fit the test set;[2] the excellent fit rate for the very active compounds is particularly encouraging (see Table 2). Somewhat discouraging is the number of ineffective compounds that fit the Dynamic-MD model (11 of 13), but it is emphasized that the ineffective compounds chosen for the study closely resembled the very active compounds. They were specifically chosen to be a difficult countertest of the models. Each model in Table 2 that fits significant numbers of very active compounds also fits many ineffective compounds. Most of the compounds in the test set were identified by searching the nonproprietary half of a small molecule database maintained by the National Cancer Institute (NCI). If the Dynamic-MD, Dynamic-Crystal, or Static 1 models were used to search the NCI database, they would have also identified many of these same successful compounds.

Poor performance is seen for Static Model 2 in fitting the set of published inhibitors. Only six compounds fit the model, most of which have IC_{50}'s well over 10 μM. The structure solved by Goldgur et al.[19] was used to develop Static Model 2. It is very different from the MD structures and the crystal structure by Maignan et al.[2] In the structure by Goldgur et al., the flexible loop adjacent to the active site is oriented away from the catalytic residues. Though the flexible loops are very different
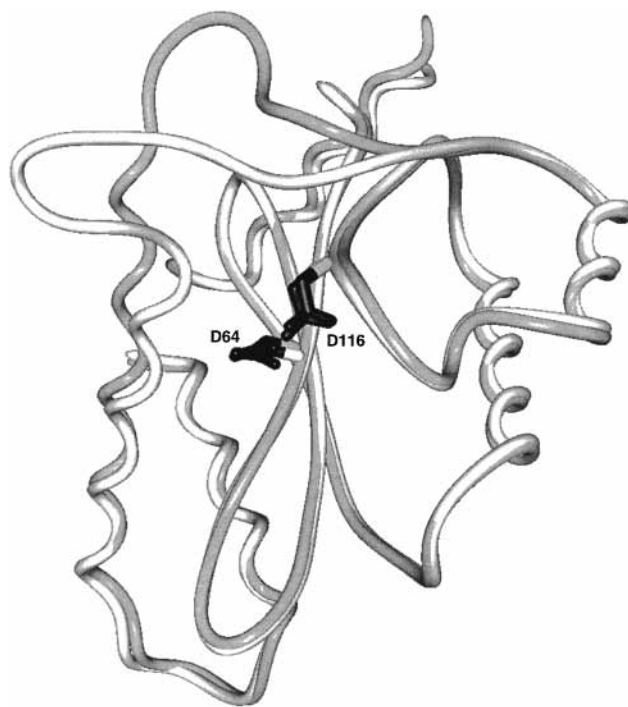

Figure 1. Crystal structure solved by Maignan et al.[18] (shown in white) and the Goldgur et al.[19] structure (shown in gray). The structures show excellent agreement in the secondary structure except in the flexible loop over the active site (a second smaller flexible loop is also in slight disagreement, but it is far from the active site and not shown for clarity). The side chains of D64 and D116 are in black. The orientation of these two essential aspartic acids is in excellent agreement between the two structures.

(see Figure 1), the relative positions of D64 and D116 are almost the same in the two structures; it is the orientations of other nearby residues that disagree and lead to different pharmacophore models. While the structure solved by Goldgur et al. does not appear to be appropriate for this particular set of inhibitory compounds, it might well be appropriate for other classes of inhibitors. Many inhibitors of HIV-1 integrase contain mostly hydrogen-bond acceptor functionalities and would be expected to dock in sites other than the active site or perhaps force the active site into a different conformation. Furthermore, the flexibility in this loop is essential for catalytic activity.[13] It is quite possible that the conformation is valid for ligand binding even though not reflected in fitting the test set, and its incorporation in to a dynamic model could lead to improvement.

Static Model 1 has the better performance of the static models, identifying 18 compounds. Eight of which are very active compounds. The structure solved by Maignan et al.,[18] that was used to develop Static Model 1, is very similar to the MD structures.[2] It is reasonable that this static model has the more similar performance to the dynamic model. Table 3 presents a comparison of the performance of the dynamic models and Static Model 1. Static Model 1 is comparable in percent yield and enrichment to the Dynamic-MD model. The excellent enrichment values for both models fitting the very active compounds are particularly notable. However, the Dynamic-MD model is clearly superior in the number of active compounds that it identifies (% actives in Table 3).

Though we are quite pleased with the success of the Dynamic-MD model, it is important to evaluate the general method using a different protocol. It would be best if the method of overlaying many structures would lead to a better performance of the pharmacophore models whether using MD simulations or

TABLE 3: Comparison of the Performance of the Dynamic Models versus Static Model 1 in Fitting the Compounds in the Test Set Given in Table 2[a]

% yield = (active compounds fit to model/all compounds fit to model
% actives = (active compounds fit/all active compounds in test set)
enrichment = (active fits/total fits)/(all active compounds in test set/all in test set)[b]

	dynamic-MD	static 1	dynamic-crystal
Active Compounds = 20 Very Active Compounds			
% yield	42	44	36
% actives	90	40	60
enrichment	1.23	1.31	1.07
Active Compounds = All 46 Active Compounds			
% yield	74	72	76
% actives	70	28	54
enrichment	0.95	0.93	0.97

[a] The performance is evaluated for the fit of the very active compounds and the fit of all active compounds. [b] An enrichment of 1.0 indicates that the model is fitting compounds with the same ratio of active compounds as exist in the original test set, while values over 1 reveal that the model is identifying a higher percentage of active compounds (model is preferentially identifying active compounds).

experimentally determined structures. For this reason, we created and evaluated the Dynamic-Crystal model. Even though Static Model 2 only fit six compounds (all of which were also fit with Static Model 1), combining the information leads to many more inhibitors fitting the Dynamic-Crystal model. Twelve of the 20 very active compounds are identified, a 50% improvement over Static Model 1. Thirteen of the active compounds in Table 2 fit the Dynamic-Crystal model, which is a very large improvement over Static Model 1 (260%). The Dynamic-Crystal model appears less selective for the very active compounds (lower % yield and enrichment values in the first half of Table 3). However, the lower % yield only reflects the improvement in fitting active compounds in Table 2, and the enrichment value is approximately 1, indicating that the model is not preferentially binding the active and inactive compounds over the very active ones. Furthermore, the % actives values in Table 3 for both very active and all active compounds is significantly improved over Static Model 1.

Conclusion

The method for developing dynamic pharmacophores has been shown to be successful when using experimentally determined protein structures or structures from MD simulations. The Dynamic-Crystal model is significantly improved over the static models. What is surprising is the excellent performance of the Dynamic-Crystal model given that only two protein structures were used to represent the flexibility of the system. This points to the versatility of the method and its applicability even when few structures are available for a given system.

The Dynamic-MD, static 1, or Dynamic-Crystal pharmacophore models could all be used in database searching to identify new inhibitors. Each would have different advantages. Dynamic-MD appears capable of identifying many very active compounds; 90% of the very active compounds fit the model, and the enrichment value is 1.23. But the high number of compounds fitting this model may imply that a very large number of compounds would be identified in database searching, making experimental testing of the compounds time-consuming. Static Model 1 is also selective for very active compounds with lower fit rates that the Dynamic-MD model, but the number of compounds identified could be too small, leading to an

insufficient number of compounds being identified. The Dynamic-Crystal model is most likely preferable over Static Model 1, as it identifies many more active and very active compounds. The size and nature of the database would dictate which dynamic pharmacophore model would be best. We are currently involved in searching the Available Chemicals Database for inhibitory compounds, and our collaborators have begun experimental testing of those compounds. We eagerly await further experimental studies employing these models.

Supporting Information Available: Chemical structures, nomenclature, and references are given for the compounds in the test set of known inhibitors of HIV-1 integrase. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Lins, R. D.; Briggs, J. M.; Straatsma, T. P.; Carlson, H. A.; Greenwald, J.; Choe, S.; McCammon, J. A. *Biophys. J.* **1999**, *76*, 2999.

(2) Carlson, H. A.; Masukawa, K. M.; Jorgensen, W. L.; Lins, R. D.; Briggs, J. M.; McCammon, J. A. *J. Med. Chem.*, in press.

(3) Bujacz, G.; Jaskólski, M.; Alexandratos, J.; Wlodawer, A.; Merkel, G.; Katz, R. A.; Skalka, A. M. *J. Mol. Biol.* **1995**, *253*, 333.

(4) Böhm, H.-J.; Klebe, G. *Angew. Chem., Int. Ed. Engl.* **1996**, *35*, 2588.

(5) Walters, W. P.; Stahl, M. T.; Murcko, M. A. *DDT* **1998**, *3*, 160.

(6) Miranker, A.; Karplus, M. *PROTEINS* **1991**, *11*, 29.

(7) Zheng, Q.; Kyle, D. J. *DDT* **1997**, *6*, 229.

(8) Miranker, A.; Karplus, M. *PROTEINS* **1995**, *23*, 472.

(9) Joseph-McCarthy, D.; Fedorov, A. A.; Almo, S. C. *Protein Eng.* **1996**, *9*, 773.

(10) Joseph-McCarthy, D.; Hogle, J. M,; Karplus, M. *PROTEINS* **1997**, *29*, 32.

(11) Leclerc, F.; Karplus, M. *Theor. Chem. Acc.* **1999**, *101*, 131.

(12) Castro, A.; Richards, W. G.; Lyne, P. D. *Med. Chem. Res.* **1999**, *9*, 98.

(13) Greenwald, J.; Le, V.; Butler, S. L.; Bushman, F. D.; Choe, S. *Biochemistry* **1999**, *38*, 8892.

(14) Knegtel, R. M. A.; Kuntz, I. D.; Oshiro, C. M. *J. Mol. Biol.* **1997**, *266*, 424.

(15) Sandak, B.; Wolfson, H. J.; Nussinov, R. *PROTEINS* **1998**, *32*, 159.

(16) Zheng, Q.; Kyle, D. J. *PROTEINS* **1994**, *19*, 324.

(17) Totrov, M.; Abagyan, R *PROTEINS* **1997**, *Suppl 1*, 215.

(18) Maignan, S.; Guilloteau, J.-P.; Zhou-Liu, Q.; Clément-Mella, C.; Mikol, V. *J. Mol. Biol.* **1998**, *282*, 359.

(19) Goldgur, Y.; Dyda, F.; Hickman, A. B.; Jenkins, T. M.; Craigie, R.; Davies, D. R. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 9150.

(20) *Catalyst 3.10;* Molecular Simulations Inc.: San Diego, CA, 1996.

(21) Neamati, N.; Sunder, S.; Pommier, Y. *DDT* **1997**, *2*, 487.

(22) Neamati, N.; Hong, H.; Sunder, S.; Milne, G. W. A.; Pommier, Y. *Mol. Pharmacol.* **1997**, *52*, 1041.

(23) Farnet, C. M.; Wang, B.; Lipford, J. R.; Bushman, F. D. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 9742.

Dynamic Fluctuations of a Protein

J. Phys. Chem. A, Vol. 103, No. 49, 1999 **10219**

(24) Hong, H.; Neamati, N.; Wang, S.; Nicklaus, M. C.; Mazumder, A.; Zhao, H.; Burke, T. R., Jr.; Pommier, Y.; Milne, G. W. A. *J. Med. Chem.* **1997**, *40*, 930.

(25) Farnet, C. M.; Wang, B.; Hansen, M.; Lipford, J. R.; Zalkow, L.; Robinson, W. E. Jr.; Siegel, J.; Bushman, F. D. *Antimicrob. Agents Chemother.* **1998**, *42*, 2245.

(26) Nicklaus, M. C.; Neamati, N.; Hong, H.; Mazumder, A.; Sunder, S.; Chen, J.; Milne, G. W. A.; Pommier, Y. *J. Med. Chem.* **1997**, *40*, 920.

(27) Tirado-Rives, J.; Jorgensen, W. L. *pepz.*; Yale University: New Haven, CT, 1998.

(28) Jorgensen, W. L. *BOSS,* version 3.8; Yale University: New Haven, CT, 1996.

(29) Jorgensen, W. L. *MCPRO,* version 1.5; Yale University: New Haven, CT, 1997.

(30) Goodford, P. J. *J. Med. Chem.* **1985**, *28*, 849.

(31) Kuntz, I. D. *Science* **1992**, *257*, 1078.

(32) Colman, P. M. *Curr. Opin. Struct. Biol.* **1994**, *4*, 868.

(33) Blundell, T. L. *Nature* **1996**, *Suppl. 384*, 23.

(34) Marrone, T. J.; Briggs, J. M.; McCammon, J. A. *Annu. Rev. Pharmacol. Toxicol.* **1997**, *37*, 71.

(35) Kaminski, G.; Duffy, E. M.; Matsui, T.; Jorgensen, W. L. *J. Phys. Chem.* **1994**, *98*, 13077.

(36) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225.

(37) *MidasPlus*; Computer Graphics Lab: University of California, San Francisco.

(38) Ferrin, T. E.; Huang, C. C.; Jarvis, L. E.; Langridge, R. *J. Mol. Graphics.* **1988**, *6*, 13.