Protein Science (1998), 7:573–579. Cambridge University Press. Printed in the USA.

Solvation studies of DMP323 and A76928 bound to HIV protease: Analysis of water sites using grand canonical Monte Carlo simulations

TAMI J. MARRONE,[1,4] HALUK RESAT,[2] C. NICHOLAS HODGE,[3] CHONG-HWAN CHANG,[3] AND J. ANDREW McCAMMON[1]

[1]Department of Chemistry and Biochemistry and Department of Pharmacology, University of California at San Diego, La Jolla, California 92093-0365
[2]Koc University, Istinye, Istanbul 8086, Turkey
[3]The DuPont Merck Pharmaceutical Company, DuPont Merck Experimental Station, Wilmington, Delaware 19880-0500

(RECEIVED September 4, 1997; ACCEPTED November 13, 1997)

Abstract

We examine the water solvation of the complex of the inhibitors DMP323 and A76928 bound to HIV-1 protease through grand canonical Monte Carlo simulations, and demonstrate the ability of this method to reproduce crystal waters and effectively predict water positions not seen in the DMP323 or A76928 structures. The simulation method is useful for identifying structurally important waters that may not be resolved in the crystal structures. It can also be used to identify water positions around a putative drug candidate docked into a binding pocket. Knowledge of these water positions may be useful in designing drugs to utilize them as bridging groups or displace them in the binding pocket. In addition, the method should be useful in finding water sites in homology models of enzymes for which crystal structures are unavailable.

Keywords: grand canonical Monte Carlo; HIV protease; simulation; water

Water molecules can play an important role in the binding affinity or specificity of an inhibitor. A prominent example is the structural water, WAT 301, seen in the crystal structure of nearly every HIV protease–inhibitor complex (Wlodawer & Erickson, 1993). WAT 301 bridges the backbone amide protons of Ile 50 and Ile 150 to inhibitor (and presumably substrate) carbonyls through hydrogen bonding. Mammalian protease–inhibitor complexes cannot accommodate it, so incorporation of a bidentate hydrogen-bond acceptor into an inhibitor to displace this water would achieve greater specificity for the viral protease over the mammalian proteases (Lam et al., 1994). A novel class of potent cyclic urea HIV protease inhibitors were rationally designed to displace WAT 301 to achieve greater specificity (Lam et al., 1994). This example clearly demonstrates the importance of the knowledge of water positions in a binding site for use in a structure-based drug design strategy.

Because of the importance of water positions, several structural techniques such as NMR spectroscopy and X-ray crystallography have been used to map out the hydration water molecules in protein and inhibitor complexes. Here we investigate the water positions within HIV-1 protease complexes with two inhibitors, DMP323 (Ala et al., 1997) and A76928 (Hosur et al, 1994), through Grand Canonical Monte Carlo (GCMC) simulations. This simulation method is complementary to the experimental determinations of water positions, and can be used to help interpret the experimental data. It can also be used to predict water positions that would be difficult to obtain experimentally, for example, when resolution or partial occupancy limit the assignment of water positions. Finally, there are many cases where it is not possible to obtain an X-ray or an NMR structure of a ligand–protein complex, but an estimate of the stability of the complex is desirable: examples include novel ligands proposed in improving binding or other properties of drug leads in the active site of enzymes, and models of therapeutically important enzymes based on homology to related enzymes with known structures.

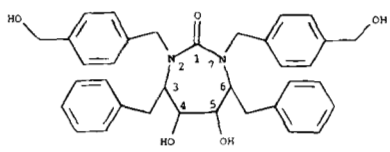
GCMC (Adams, 1975) has been used successfully to determine the water positions in small polydisaccharides (Resat & Mezei, 1994), in DNA crystal structures (Resat & Mezei, 1996), and in the association of benzamidine with trypsin (Resat et al., 1997). This simulation technique is performed in the grand canonical ensemble, i.e., the number of water molecules is allowed to vary while the chemical potential is kept constant (Adams, 1975; Cagin &

Reprint requests to: Professor J. Andrew McCammon, Department of Chemistry and Biochemistry, 9500 Gilman Drive, 0365, University of California San Diego, La Jolla, California 92093-0365; e-mail: jmccammo@chem.ucsd.edu.

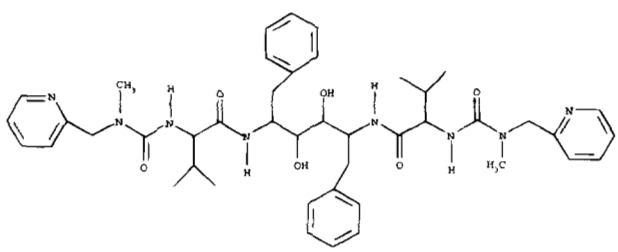
[4]Present address: Agouron Pharmaceuticals, 3301 North Torrey Pines Court, La Jolla, California 92037.

Pettitt, 1991; a closely related approach is the Gibbs ensemble simulations, see Panagiotopoulos, 1992). During the GCMC simulation, waters are translated, rotated, inserted, or deleted within the simulation cell provided the moves satisfy the Markov chain rules of Monte Carlo simulations. The protein and inhibitor remained fixed. The chemical potential is coupled to a constant chemical potential reservoir and a simulation parameter is adjusted to give a bulk water density in the simulation cell that is equal to that in the chemical potential reservoir. Because water molecules are inserted and deleted, water molecules in deep pockets within the protein are more easily sampled using GCMC than traditional Monte Carlo or molecular dynamics techniques (Resat & Mezei, 1994). The entire trajectory is analyzed, and water sites are predicted using a graph theory analysis (Berge, 1962). The occupancy of these sites and the fluctuations of water positions around these sites are also calculated. These quantities reflect the probability of finding a water at that site at any given time and the mobility of the simulation waters assigned to that site. Further details of these simulations can be found elsewhere (Resat & Mezei, 1996).

We investigate the water positions within complexes of HIV-1 protease with two inhibitors DMP323 and A76928, through GCMC simulations. Essentially, we use this technique to predict water sites in both HIV complexes. Schematic diagrams for these inhibitors are shown in Figure 1. DMP323 is a cyclic urea inhibitor designed to displace WAT 301. A76928 does not displace WAT 301, which is seen in the crystal structure of the complex. These complexes were chosen for study because their high-resolution X-ray structures (Hosur et al., 1994; Ala et al., 1997) are available (pdb entries: 1hvk and 1qbs), and they represent two different classes of inhibitors. Also, there is extensive NMR data for the DMP323 complex (Wang et al., 1996a). The simulation results for these two systems are used to address (a) the utility of the GCMC technique to reproduce and predict experimentally observed water molecules in a protein and inhibitor complex; and (b) to explore how the inhibitor affects the distribution of water molecules within the binding site and throughout the HIV protease.



DMP 323



A76928

Fig. 1. Schematic diagram of the inhibitors examined in this study, DMP323 and A76928.

Results

Table 1 contains the identity of each crystal water of the DMP323-HIV protease complex, the occupancy of the computationally predicted site assigned to the crystal water, the distance between the crystal water and the predicted site, and the root-mean-square (RMS) fluctuations of the simulated waters around the predicted site. Here we define the water sites as the predicted oxygen positions. The hydrogens of the simulated water molecules occupy many locations around the oxygen site during the simulations, and the average positions do not always represent a reasonable water conformation. Hence, all predictions and comparisons presented will refer to the water oxygen site only. Comparing the distance between the crystal waters and the predicted sites, we reproduce the crystal waters quite well. In cases where the crystal water and the predicted site were greater than 1.5 Å apart, a visual inspection was made of these crystal waters, and it was found that all of these waters reside at the protein surface. Surface waters are less well-defined crystallographically (Lounnas et al., 1994). Also, our computer simulation more closely represents a protein in dilute solution than a crystalline environment, so one would expect some differences between predicted surface waters and those determined crystallographically, due to the influence of neighboring protein on the surface water structure in the crystal.

The catalytically competent form of HIV protease exists on average as a dimer with C2 symmetry; hence, certain waters seen in one monomer are also seen in a symmetry-related position in the other monomer. Curiously, a water site that is symmetry related to WAT 401 in the DMP323 structure was not seen in the crystal structure. The symmetrically equivalent position is occupied by the Arg87 sidechain. With a simple rotation, the position could accommodate a water. Analysis of the DMP323 simulation reveals a water site near Arg87. Although this site was not identified in the crystal structure, it is clearly predicted by the simulation. The position of this water can be seen in Figure 2. Recent NMR studies (Wang et al., 1996a), which examined the water structure of the DMP323–HIV protease complex, show the symmetry-related pair of WAT 477 and WAT 415 to exhibit weak signals and have short residence times, although they are buried waters and are seen in the crystal structure. The GCMC simulation results reproduce these waters, but predict them to be occupied only 60% of the time. These results are in agreement with the experimental NMR observations.

The DMP323 GCMC simulation reproduces the 122 crystal waters seen in the DMP323 crystal structure. Several of these waters positions are also seen in the A76928 results, although only one structural water (WAT 301) is given in the deposited A76928 crystal structure. The simulation of A76928 reproduces WAT 301 extremely well. The distance between the predicted site and the water oxygen is 0.12 Å. The occupancy of the simulated site was 1.00, and fluctuations around the site were 0.27 Å RMS.

Given that we have demonstrated the ability of this simulation method to reproduce experimentally determined water positions, we now examine the water positions within 5 Å of each protease inhibitor. Figures 3A and B show all water molecules within 5 Å of any atom of the inhibitors DMP323 and A76928, respectively. Ile50, Ile 150, Asp 25, and Asp 125 are highlighted in thick lines, while other protein residues are shown in blue. The inhibitors are shown in thick lines as well. Crystal waters are shown in magenta, while the predicted water sites are shown in red in Figure 3A, the DMP323 complex. The simulation results

Table 1. *Comparison of the predicted simulation sites with the DMP323-HIV protease complex crystal structure*

Crystal water	Occupancy	Distance (Å)	RMS (Å)	Crystal water	Occupancy	Distance (Å)	RMS (Å)
401	1.000	0.984	0.363	462	1.000	0.353	1.010
402	1.000	0.315	0.238	463	1.000	1.869	0.693
403	1.000	0.425	0.248	464	1.000	0.739	0.365
404	1.000	1.077	1.081	465	1.000	1.706	0.779
405	1.000	0.401	0.347	466	1.000	0.434	0.712
406	1.000	0.339	0.315	467	1.000	0.714	0.701
407	0.957	0.927	0.951	468	1.000	0.414	0.382
408	1.000	0.890	0.682	469	0.507	1.391	0.455
409	0.952	2.437	0.992	470	1.000	1.032	0.469
410	1.000	1.647	1.004	471	0.875	1.819	1.015
411	0.793	0.356	0.294	472	1.000	1.522	0.695
412	1.000	0.148	0.426	473	1.000	2.095	1.312
413	1.000	0.627	0.793	474	1.000	1.944	1.111
414	0.738	1.179	0.564	475	1.000	1.803	1.379
415	0.627	0.407	0.850	476	1.000	0.931	0.650
416	1.000	0.917	0.582	477	0.570	0.518	0.303
417	1.000	1.344	0.978	478	1.000	0.259	0.406
418	1.000	1.773	0.860	479	1.000	0.632	0.643
419	0.655	0.685	0.789	480	1.000	1.179	0.381
420	1.000	2.079	0.670	481	1.000	1.982	0.693
421	0.490	1.379	0.641	482	1.000	1.330	0.908
422	0.842	0.237	0.815	483	1.000	1.834	0.602
423	1.000	2.405	0.928	484	1.000	0.950	0.516
424	1.000	2.196	1.226	485	1.000	0.380	0.475
425	0.992	1.560	1.410	486	1.000	0.813	0.298
426	1.000	0.242	0.220	487	1.000	0.641	0.686
427	1.000	0.771	0.733	488	1.000	1.862	0.587
428	0.488	1.162	0.652	489	1.000	0.534	0.287
429	1.000	2.026	0.329	490	0.657	1.279	1.251
430	1.000	1.458	0.594	491	0.680	2.239	1.229
431	1.000	0.747	1.045	492	1.000	1.727	1.059
432	0.550	2.286	0.538	493	1.000	1.065	0.560
433	0.965	0.577	0.884	494	0.902	2.799	1.739
434	1.000	1.279	1.141	495	1.000	1.364	0.717
435	0.327	2.117	0.365	496	1.000	0.291	0.635
436	0.913	0.234	0.476	497	1.000	1.006	0.317
437	1.000	1.056	0.784	498	1.000	0.753	0.530
438	1.000	1.716	0.621	499	0.995	0.752	0.795
439	1.000	1.118	0.421	500	0.452	1.150	0.419
440	0.895	0.846	0.733	501	1.000	0.341	0.659
441	1.000	0.978	0.466	502	1.000	1.119	1.455
442	1.000	1.289	0.481	503	0.973	1.728	0.842
443	1.000	1.589	0.677	504	1.000	0.987	0.416
444	1.000	1.604	0.377	505	0.350	2.050	0.318
445	0.997	1.259	0.991	506	1.000	1.087	0.754
446	1.000	0.881	1.454	507	1.000	0.903	1.119
447	1.000	2.793	0.611	508	1.000	1.825	0.381
448	1.000	0.518	0.651	509	1.000	1.590	0.583
449	0.988	2.369	1.086	510	1.000	1.207	1.120
450	1.000	1.252	0.415	511	1.000	0.389	1.268
451	1.000	1.209	0.616	512	1.000	1.101	0.734
452	1.000	1.097	0.474	513	1.000	1.972	0.738
453	1.000	1.556	0.795	514	1.000	1.540	0.495
454	1.000	1.166	1.012	515	0.982	2.199	1.972
455	1.000	0.603	0.478	516	1.000	1.481	0.434
456	0.915	0.756	0.325	517	0.428	1.173	0.245
457	1.000	0.206	0.718	518	1.000	1.625	0.875
458	1.000	1.168	1.100	519	0.978	0.872	1.109
459	1.000	0.999	0.393	520	1.000	1.563	0.771
460	1.000	1.293	0.638	521	1.000	0.257	0.603
461	1.000	1.435	0.882	522	1.000	1.223	0.441

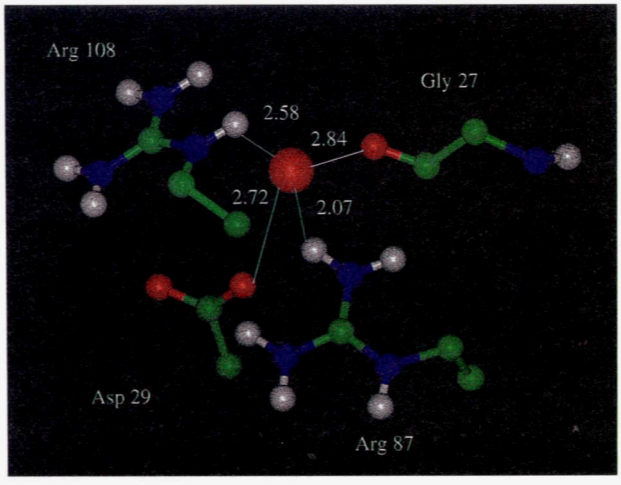
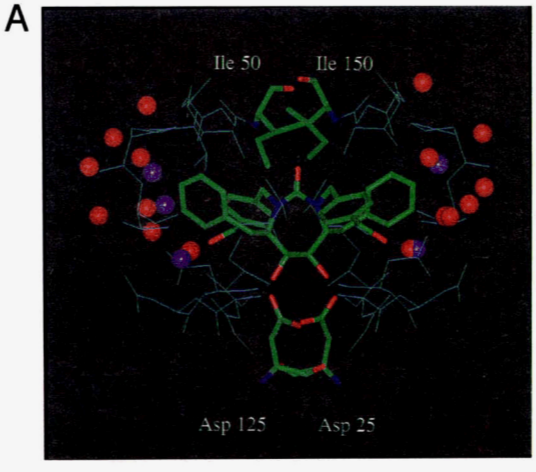
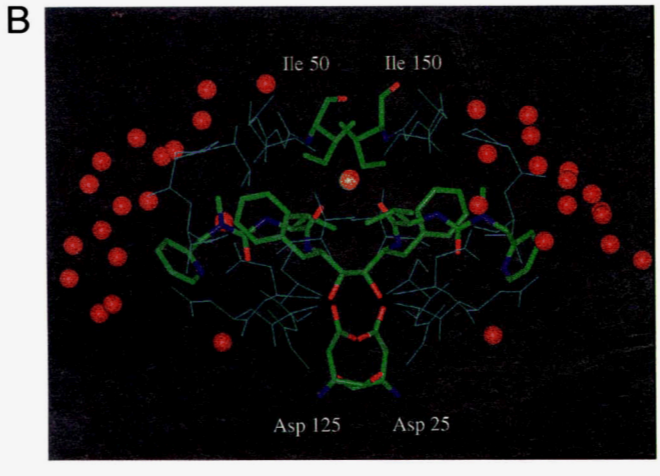


Fig. 2. Water site symmetry related to WAT 401. This site was not observed in the DMP323 crystal structure, but was predicted by the GCMC simulation of the DMP323 HIV protease complex.

reproduce the crystal waters well. The simulation also predicts several water molecules within 5 Å that are not seen in the crystal structure. These waters reside at the surface of the protein and would be difficult to verify experimentally. Figure 3B shows the A76928 structure in the same orientation as the DMP323 complex. The crystal water, WAT 301 is shown in yellow and the predicted sites are shown in red. Comparing the simulation results for DMP323 with A76928, there are different solvation patterns within 5 Å of the inhibitors. Because it is a much larger molecule than DMP323, A76928 would displace some of the crystal waters observed in Figure 3A.

The simulation also reproduced waters 422 and 456 in the DMP323 complex (waters nearest to the bottom of the active site in Fig. 3A). These waters make extensive contact with the complex: WAT 422 (occupancy 0.842) interacts with the inhibitor benzylic hydroxyl group (3.62 Å), both carboxylate oxygens of Asp 129 (3.10, 3.54 Å), the amide nitrogen of Asp 129 (3.07 Å), the carbonyl oxygen of Gly 127 (3.33 Å), and a terminal nitrogen of Arg 8 (3.5 Å). The extremely tight fit of this water into its pocket is shown in Figure 3C. This site is not predicted to be occupied in the simulation of A76928, which is reasonable because the amide nitrogen between P1 and P2 donates a hydrogen bond to Gly 127 (Fig. 3B), making the pocket smaller and removing two sources of electrostatic stabilization. In fact, none of the crystal structures of linear inhibitors with a secondary amide linkage between P1 and P2 show a water in the vicinity of WAT 422; but WAT 422 does appear in 1hvp, wherein the P1′–P2′ linkage is a tertiary sulfonamide without a hydrogen to donate. Thus, the sim-





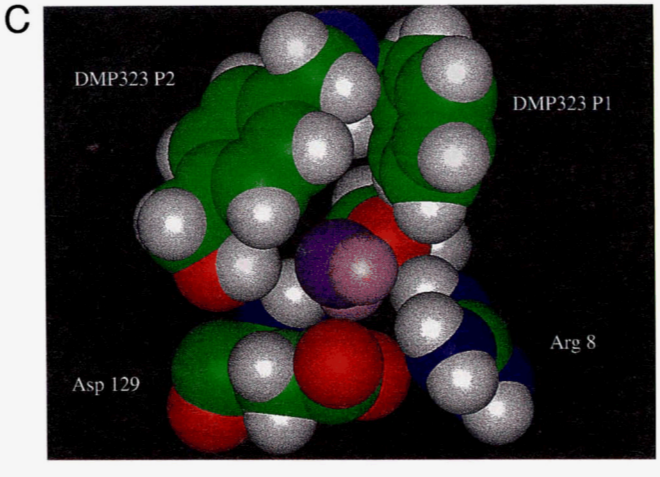


Fig. 3. A: Residues and waters within 5 Å of the DMP323 inhibitor bound to HIV protease. Predicted water sites are shown in red, while crystal waters for the DMP323 structure are shown in magenta. Asp 25, Asp 125, Ile 50, Ile 150, and DMP323 are shown as thick lines, while other protein residues are shown as pale blue line for clarity. **B:** Residues and waters within 5 Å of the A76928 inhibitor bound to HIV protease. Predicted water sites are shown in red, while WAT 301 is shown in yellow. Asp 25, Asp 125, Ile 50, Ile 150, and A76928 are shown as thick lines, while other protein residues are shown as pale blue line for clarity. **C:** CPK model of DMP323–HIV protease protease showing all atoms within 4 Å of the tightly encapsulated water 422 (magenta). Hydrogens were added to empty valences (pH 6) to better represent steric contacts.

Table 2. *Brookhaven Protein Data Bank entries superimposed and analyzed with respect to the DMP323 structure*

PDB entry	Backbone RMS deviation (Å)	Resolution (Å)	Inhibitor	Reference
1HBV	0.406	2.3	SB203238	Hoog et al., 1995
1HIV	0.768	2.0	U75875	Thanki et al., 1992
1HOS	0.461	2.3	SB204144	Abdel-Meguid et al., 1993
1HPS	0.494	2.3	SB206343	Thompson et al., 1994
1HPV	0.331	1.9	VX-478	Kim et al., 1995
1HTF	0.471	2.2	GR126045	Jhoti et al., 1994
1HTG	0.646	2.0	GR137615	Jhoti et al., 1994
4HVP	0.697	2.3	MVT-101	Miller et al, 1989
4PHV	0.649	2.1	L-700,417	Bone et al., 1991
1HVK	0.719	1.8	A76928 (s,s)	Hosur et al., 1994
1QBS	—	1.9	DMP323	Ala et al., 1997

ulation is successful in selectively placing a water into a highly constrained cavity in the DMP323 complex.

To test the ability of the GCMC simulation method to predict water positions that were not known a priori, we examined the crystal waters of several HIV protease and inhibitor complexes deposited in the Brookhaven Protein Data Bank. The crystal structures of several complexes, which are described in Table 2, were superimposed onto the DMP323 crystal structure using the backbone coordinates only. The RMS deviations for this superimposition are found in Table 2. The DMP323 simulation was analyzed within the context of the crystal waters of these protease complexes. Although the simulation was performed on the DMP323 complex and not on these other complexes, we analyzed their crystal waters as if they were the experimentally determined DMP323 crystal waters. This analysis identifies crystal waters seen in other protease complexes, but not seen in the DMP323 structure. Given that the structure of the protease complexes are similar, we can garner confidence that our method is correctly predicting water sites, because these sites are observed in other crystal structures.

Because there is difficulty in reproducing and predicting surface waters reliably, we only analyzed sites within 4 Å of the protein and inhibitor complex that had a buried van der Waals surface area greater than 25%. Although this cutoff is somewhat arbitrary, it was chosen because oxygen sites predicted in the DMP323 simulation that had greater than 25% buried surface areas were less than 1.5 Å away from the assigned crystal water. Oxygens with buried van der Waals surface areas less than 25% ranged from 0.4 to 2.8 Å away from the crystal waters.

Figure 4 shows a ribbon diagram of the GR126045 HIV protease complex. Shown in red are water sites that were predicted by the DMP323 simulation, yet not seen in the DMP323 crystal structure. The crystal waters are shown in yellow. The fact that these waters were observed in other crystal structures in areas of the protein distant from the inhibitor suggests that the GCMC simulation technique can be used for reliable water site prediction. Note that we applied the criteria that a simulated water must have a buried van der Waals surface area greater than 25% for our analysis. Some of the symmetry-related waters one would expect to see in this figure are not shown, because their buried surface areas did not meet the specified criteria.

The water distributions in and around the binding pocket are shown to vary with the inhibitor, as seen in Figure 3. The next issue to be addressed is how much do the water distributions vary in other locations within HIV protease that are distal to the binding pocket. The 122 crystal waters of DMP323 were examined and it was determined if they were observed in the various crystal structures taken from the Brookhaven Protein Data Bank. A water is considered to be seen in the other crystal structure if there is a water less than 1.5 Å from the DMP323 crystal water. The deposited crystal structure of the complex with A76928, 1HVK, reported only one crystal water, WAT 301, so the simulated water sites were used for the comparison. These results can be found in Table 1 of the Supplementary Material. Several waters are observed experimentally in all or many of the structures: 401, 402, 403, 406, 409, 415, 436, and 477. WAT 301 is also present in structures for most inhibitors. Figure 5 shows the position of these water molecules. Not surprisingly, conserved water molecules are seen in regions of the protein distal to the inhibitor binding pocket and the flaps of the protease. However, WAT 301 is conserved in inhibitors that can accommodate it.

Discussion

We have demonstrated that water sites determined from the grand canonical Monte Carlo simulations of two HIV protease–inhibitor complexes reproduce the water sites observed in the crystal structures of these complexes quite well. Our simulation results show that this method is predictive by determining sites observed in other HIV protease complexes with similar structures to the protein–inhibitor complexes studied here. The results also indicate a water symmetric to WAT 401 that was not seen in the DMP323 crystal structure.

Comparing the water distributions within 5 Å of the DMP323 and the A76928 show that these inhibitors have a profound effect on the local water distribution. However, comparing the predicted waters with crystal waters from several HIV structures, the inhibitors have very little effect on the water distribution in regions of the protein distal to the binding pocket. This suggests that future

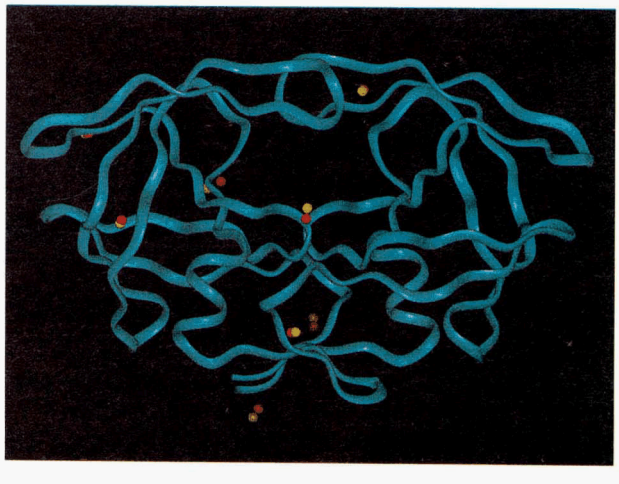


Fig. 4. A ribbon diagram of the GR126045 HIV protease structure. Water molecules predicted in the DMP323 simulation that are observed in the GR126045 crystal structure, but not in the DMP323 crystal structure, are shown in red. Crystal waters are shown in yellow.

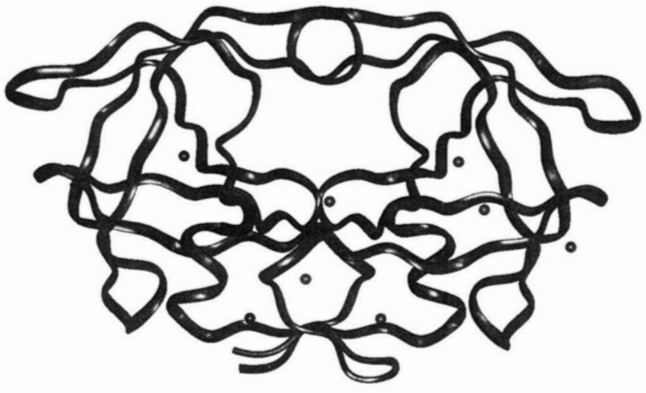


Fig. 5. A ribbon diagram of the DMP323 HIV protease structure. Water molecules that are conserved throughout the HIV protease complexes given in Table 2 are shown in black.

simulation studies may only need to include a sphere centered on the binding pocket to identify important waters that coexist near the inhibitor. Although these simulations took ~3 weeks of computer time on an SGI R4400 with a 200 MHz clock, reducing the system size should drastically reduce the computational cost and analysis time. This reduced model could be used to determine the water structure around putative drug candidates docked into the active site of an enzyme as part of a structure-based drug design strategy. The method may also be useful for post-processing homology models to assign likely hydration sites prior to docking or dynamics simulations in the absence of explicit solvent.

Materials and methods

Force field

The protein atoms were represented by the united atom OPLS (Jorgensen & Tirado-Rives, 1988) force field, in both simulations. The side chains of aromatic residues were represented explicitly using all-atom OPLS parameters. The TIP3P (Jorgensen et al., 1983) water model was used. Charges of the central ring of DMP323 were fit to the potential of the hydrogen-capped central ring using the Merz-Kollman method (Bessler et al., 1990) The electrostatic potential was generated from an ab initio calculation utilizing the 6-31G* basis set within Gaussian 94 (Frisch et al., 1994) Lennard-Jones parameters for the central ring were assigned from the OPLS force field (Jorgensen & Tirado-Rives, 1988) using values for analogous atoms. The aromatic groups of DMP323 that are peripheral to the central ring were assigned charges and Lennard-Jones values for all-atom benzene parameters from the OPLS force field. The parameters of the remaining peripheral atoms were assigned from analogous atoms in the OPLS force field. The Lennard-Jones parameters and the charges for A76928 were taken from analogous groups in the OPLS force field.

Simulation protocol

Protons were added to Asp 25 and Asp 125 in the crystal structure of the DMP323 complex and oriented to obtain the best hydrogen-bonding interaction with the hydroxyl groups of the inhibitors. The catalytic aspartates were also modeled as protonated for the A76928 complex. pH-dependent changes in carbon chemical shifts suggest that Asp 25 and Asp 125 are protonated when DMP323 is bound to HIV protease (Yamazaki et al., 1994). A similar experiment showed shifts that are consistent with a monohydroxy inhibitor, KNI-272, binding to HIV protease with singly deprotonated catalytic aspartates (Wang et al., 1996b). Because there is some uncertainty as to how these data relate to the actual ionization state under inhibitor assay conditions (C.N. Hodge, unpubl. obs.) the decision was made to use the same protonation state for both complexes in this study. Polar hydrogens were added to both structures based on geometrical rules and were relaxed in the presence of a fixed protein and inhibitor.

The crystal structure of each HIV protease complex was stripped of its crystal waters and placed in a box with dimensions $72 \times 58 \times 46$ Å. Periodic boundary conditions were employed. The protease and the inhibitor remained fixed in space while only the water molecules were moved during the Monte Carlo (MC) run. The solute–solvent and solvent–solvent group-based cutoffs were set to 10 Å. The chemical potential of the waters was adjusted to give an appropriate density (1.0 g/mL). There were 5,138 and 5,144 water molecule in the simulations cell on average for the DMP323 and A76928 protease–inhibitor complex simulations, respectively. The system was equilibrated for 42 million MC steps and sampled for 15 million steps. Each MC move was followed by an insertion attempt and a deletion attempt during both the equilibration and sampling phases. The coordinates of the atoms in the system were stored every 50,000 steps during sampling, resulting in 300 stored coordinate sets.

Electronic supplementary material

A table containing the DMP323 crystal water positions that are also seen in other HIV protease crystal structures is given in the electronic supplementary material. This table can be found in the file named hiv_solv.doc.

Acknowledgments

We are grateful to Professor Mihaly Mezei of Mount Sinai School of medicine for the use of his grand canonical Monte Carlo code. We also acknowledge NIH and NSF for financial support and the NSF MetaCenter Program for computer time. T.J.M. was supported in part by an NIH postdoctoral fellowship and wishes to thank P. Carty for helpful discussions.

References

Abdel-Meguid SS, Zhao BG, Murthy KHM, Winborne E, Choi J-K, DesJarlais RL, Minnich MD, Culp JS, DeBouck C, Tomaszek TA Jr, Meek TD, Dreyer GB. 1993. Inhibition of human immunodeficiency virus-1 protease by a c2 symmetric phosphinate-synthesis and crystallographic analysis. *Biochemistry 32*:7972–7980.

Adams DJ. 1975. Grand canonical ensemble Monte Carlo for a Lennard-Jones fluid. *Mol Phys 29*:307–311.

Ala PJ, Huston EE, Klabe RM, McCabe DD, Duke JL, Rizzo CJ, Korant BD, DeLoskey RJ, Lam PYS, Hodge CN, Chang C-H. 1997. Molecular basis of HIV-1 protease drug resistance: Structural analysis of mutant proteases complexed with cyclic urea inhibitors. *Biochemistry 36*:1573–1580.

Berge, C. 1962. *The theory of graphs and its applications*. New York: John Wiley and Sons.

Bessler BH, Merz KM Jr, Kollman PA. 1990. Atomic charges derived from semi-empirical methods. *J Comput Chem 11*:431–439.

Bone R, Vacca JP, Anderson PS, Holloway MK. 1991. X-ray crystal structure of the HIV protease complex with L-700,417, an inhibitor with pseudo C2 symmetry. *J Am Chem Soc 113*:9382–9384.

Cagin T, Pettitt BM. 1991. Grand molecular dynamics: A method for open systems. *Mol Phys 6*:5–26.

Frisch MJ, Trucks HB, Schlegel HB, Gill PMW, Johnson, BG, Robb MA, Cheeseman JR, Keith TA, Petersson, Montgomery JA, Raghavachari K, Al-laham MA, Zakrzewski VG, Ortiz JV, Foresman JB, Cioslowski J, Stefanov BB, Nanayakkara A, Challacombe M, Peng CY, Ayala PY, Chen W, Wong MW, Andres JL, Perplogle ES, Gomperts R, Martin RL, Fox DJ, Binkley JS, DeFrees DJ, Baker J, Stewart JP, Head-Gordon M, Gonzalez C, Pople JA. 1995. *Gaussian 94 Revision E.1.* Pittsburgh, Pennsylvania: Gaussian Inc.

Hoog SS, Zhao BG, Winborne E, Fisher S, Green DW, DesJarlais RL, Newlander KA, Callahan JF. 1995. A check on rational drug design-crystal structure of a complex of human immunodeficiency virus-1 protease with a novel gamma-turn mimetic inhibitor. *J Med Chem 38*:3246–3252.

Hosur MV, Bhat TN, Kempf DJ, Baldwin ET, Liu B, Gulnik S, Wideburg NE, Norbeck DW, Appelt K, Erickson JW. 1994. Influence of stereochemistry on activity and binding modes for C(2) symmetry-based diol inhibitors of HIV-1 protease. *J Am Chem Soc 116*:847–855.

Jhoti H, Singh OMP, Weir MP, Cooke R, Murray-Rust P, Wonacott A. 1994. X-ray crystallographic studies of a series of penicillin-derived asymmetric inhibitors of HIV-1 protease. *Biochemistry 33*:8417–8427.

Jorgensen WL, Tirado-Rives J. 1988. The OPLS potential functions for proteins. Energy minimizations for crystals of cyclic peptides and crambin. *J Am Chem Soc 110*:1657–1666.

Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. 1983. Comparison of simple potential functions for simulating liquid water. *J Chem Phys 79*:926–933.

Kim EE, Baker CT, Dwyer MD, Murcko MA, Rao BG, Tung RD, Navia MA. 1995. Crystal structure of HIV-1 protease in complex with VX-478, a potent and orally bioavailable inhibitor of the enzyme. *J Am Chem Soc 117*:1181–1182.

Lam PYS, Jadhav PK, Eyermann CJ, Hodge CN, Ru Y, Becheler LT, Meek JL, Otto MJ, Rayner MM, Wong YN, Chang C-H, Weber PC, Jackson DA, Sharpe TR, Erickson-Viitanen S. 1994. Rational design of potent, bioavailable, nonpeptide cyclic ureas as HIV protease inhibitors. *Science 263*:380–384.

Lounnas V, Pettitt MB, Phillips GN Jr. 1994. A global model of the solvent–protein interface. *Biophys J 66*:601–604.

Miller M, Schneider J, Sathyanarayana BK, Toth MU, Marshall GR, Clawson L, Selk L, Kent SBH, Wlodawer A. 1989. Structure of a complex of synthetic HIV-1 protease with a substrate-based inhibitor at 2.3 Å resolution. *Science 246*:1149–1152.

Panagiotopoulos AZ. 1992. Direct determination of fluid phase equilibria by simulation in the Gibbs ensemble: A review. *Mol Simul 9*:1–23.

Resat H, Mezei M. 1994. Grand canonical Monte Carlo simulation of water positions in crystal hydrates. *J Am Chem Soc 116*:7451–7452.

Resat H, Mezei M. 1996. Grand canonical ensemble Monte Carlo simulation of the dCpG/proflavine crystal hydrate. *Biophys J 71*:1179–1190.

Resat H, Marrone TJ, McCammon JA. 1997. Enzyme-Inhibitor association thermodynamics: Explicit and continuum solvent studies. *Biophys J 72*:522–532.

Thanki N, Rao JKM, Foundling SI, Howe WJ, Tomasselli AG, Heinrikson RC, Thaisrivongs S, Wlodawer A. 1992. Crystal structure of a complex of HIV-1 protease with a dihydroxyethylene-containing inhibitor—Comparisons with molecular modeling. *Protein Sci 8*:1061–1072.

Thompson SK, Murthy KHM, Zhao BG, Winborne E, Green DW, Fisher S, DesJarlais RL, Tomaszek TA Jr, Meek TD. 1994. Rational design, synthesis, and crystallographic analysis of a hydroxyethylene-based HIV-1 protease inhibitor containing a heterocyclic p-1′, P-2′ amide bond isostere. *J Med Chem 37*:3100–3107.

Wang Y-X, Freedberg DI, Grzesiak S, Torchia DA, Wingfield PT, Kaufman JD, Stahl SJ, Chang C-H, Hodge CN. 1996a. Mapping hydration water molecules in the HIV-1 protease/DMP323 complex in solution by NMR spectroscopy. *Biochemistry 35*:12694–12704.

Wang Y-X, Freedberg DI, Yamazaki T, Wingfield PT, Stahl SJ, Kaufman JD, Kiso Y, Torchia DA. 1996b. Solution NMR evidence that HIV-1 protease catalytic aspartyl groups have different ionization states in the complex formed with the asymmetric drug KNI-272. *Biochemistry 35*:9945–9950.

Wlodawer A, Erickson JW. 1993. Structure-based inhibitors of HIV-1 protease. *Annu Rev Biochem 62*:543–585.

Yamazaki T, Nicholson LK, Torchia DA, Wingfield P, Stahl SJ, Kaufman JD, Eyermann CJ, Hodge CN, Lam PYS, Ru Y, Jadhav PK, Chang C-H, Weber PC. 1994. NMR and X-ray evidence that the HIV protease catalytic aspartyl groups are protonated in the complex formed by the protease and a non-peptide cyclic urea-based inhibitor. *J Am Chem Soc 116*:101791–101792.